



# Identifying Sagittarius Stream Stars by Their APOGEE Chemical Abundance Signatures

Sten Hasselquist<sup>1,2,20</sup>, Jeffrey L. Carlin<sup>3</sup>, Jon A. Holtzman<sup>1</sup>, Matthew Shetrone<sup>4</sup>, Christian R. Hayes<sup>5</sup>, Katia Cunha<sup>6</sup>, Verne Smith<sup>7</sup>, Rachael L. Beaton<sup>8</sup>, Jennifer Sobek<sup>9</sup>, Carlos Allende Prieto<sup>10,11</sup>, Steven R. Majewski<sup>5</sup>, Borja Anguiano<sup>5</sup>, Dmitry Bizyaev<sup>12,13</sup>, D. A. García-Hernández<sup>10,11</sup>, Richard R. Lane<sup>14,15</sup>, Kaike Pan<sup>12</sup>, David L. Nidever<sup>7,16</sup>, José. G. Fernández-Trincado<sup>17,18,19</sup>, John C. Wilson<sup>5</sup>, and Olga Zamora<sup>10,11</sup>

<sup>1</sup>New Mexico State University, Las Cruces, NM 88003, USA; [sten@nmsu.edu](mailto:sten@nmsu.edu)

<sup>2</sup>Department of Physics & Astronomy, University of Utah, Salt Lake City, UT, 84112, USA; [stenhasselquist@astro.utah.edu](mailto:stenhasselquist@astro.utah.edu)

<sup>3</sup>LSST, Tucson, AZ 85719, USA; [jcarlin@lsst.org](mailto:jcarlin@lsst.org)

<sup>4</sup>University of Texas at Austin, McDonald Observatory, Fort Davis, TX 79734, USA; [shetrone@astro.as.utexas.edu](mailto:shetrone@astro.as.utexas.edu)

<sup>5</sup>Dept. of Astronomy, University of Virginia, Charlottesville, VA 22904-4325, USA; [crh7gs@virginia.edu](mailto:crh7gs@virginia.edu), [srm4n@virginia.edu](mailto:srm4n@virginia.edu), [ba7t@virginia.edu](mailto:ba7t@virginia.edu), [jcw6z@virginia.edu](mailto:jcw6z@virginia.edu)

<sup>6</sup>Observatório Nacional/MCTI, Rua Gen. José Cristino, 77, 20921-400, Rio de Janeiro, Brazil; [cunha@email.noao.edu](mailto:cunha@email.noao.edu)

<sup>7</sup>National Optical Astronomy Observatories, Tucson, AZ 85719, USA; [vsmith@email.noao.edu](mailto:vsmith@email.noao.edu), [dnidever@email.noao.edu](mailto:dnidever@email.noao.edu)

<sup>8</sup>Department of Astrophysical Sciences, Princeton University, 4 Ivy Lane, Princeton, NJ 08544, USA; [rachael.l.beaton@gmail.com](mailto:rachael.l.beaton@gmail.com)

<sup>9</sup>Department of Astronomy, Box 351580, University of Washington, Seattle, WA 98195, USA; [jsobek@uw.edu](mailto:jsobek@uw.edu)

<sup>10</sup>Instituto de Astrofísica de Canarias, E-38205 La Laguna, Tenerife, Spain; [callende@iac.es](mailto:callende@iac.es), [agarcia@iac.es](mailto:agarcia@iac.es), [ozamora@iac.es](mailto:ozamora@iac.es)

<sup>11</sup>Universidad de La Laguna (ULL), Departamento de Astrofísica, E-38206 La Laguna, Tenerife, Spain

<sup>12</sup>Apache Point Observatory and New Mexico State University, P.O. Box 59, Sunspot, NM 88349-0059, USA

<sup>13</sup>Sternberg Astronomical Institute, Moscow State University, Moscow, Russia

<sup>14</sup>Instituto de Astrofísica, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, 782-0436 Macul, Santiago, Chile; [rlane@astro.puc.cl](mailto:rlane@astro.puc.cl)

<sup>15</sup>Millennium Institute of Astrophysics, Av. Vicuña Mackenna 4860, 782-0436 Macul, Santiago, Chile

<sup>16</sup>Department of Physics, Montana State University, P.O. Box 173840, Bozeman, MT 59717-3840, USA

<sup>17</sup>Departamento de Astronomía, Casilla 160-C, Universidad de Concepción, Concepción, Chile; [jfernandezt@astro-udec.cl](mailto:jfernandezt@astro-udec.cl)

<sup>18</sup>Instituto de Astronomía y Ciencias Planetarias, Universidad de Atacama, Copayapu 485, Copiapó, Chile

<sup>19</sup>Institut Utinam, CNRS UMR 6213, Université Bourgogne-Franche-Comté, OSU THETA Franche-Comté, Observatoire de Besançon, BP 1615, F-25010 Besançon Cedex, France

Received 2018 September 25; revised 2018 December 19; accepted 2019 January 3; published 2019 February 11

## Abstract

The SDSS-IV Apache Point Observatory Galactic Evolution Experiment (APOGEE) survey provides precise chemical abundances of 18 chemical elements for  $\sim 176,000$  red giant stars distributed over much of the Milky Way Galaxy (MW), and includes observations of the core of the Sagittarius dwarf spheroidal galaxy (Sgr). The APOGEE chemical abundance patterns of Sgr have revealed that it is chemically distinct from the MW in most chemical elements. We employ a  $k$ -means clustering algorithm to six-dimensional chemical space defined by  $[(C+N)/Fe]$ ,  $[O/Fe]$ ,  $[Mg/Fe]$ ,  $[Al/Fe]$ ,  $[Mn/Fe]$ , and  $[Ni/Fe]$  to identify 62 MW stars in the APOGEE sample that have Sgr-like chemical abundances. Of the 62 stars, 35 have *Gaia* kinematics and positions consistent with those predicted by  $N$ -body simulations of the Sgr stream, and are likely stars that have been stripped from Sgr during the last two pericenter passages ( $< 2$  Gyr ago). Another 20 of the 62 stars exhibit chemical abundances indistinguishable from the Sgr stream stars, but are on highly eccentric orbits with median  $r_{apo} \sim 25$  kpc. These stars are likely the “accreted” halo population thought to be the result of a separate merger with the MW 8–11 Gyr ago. We also find one hypervelocity star candidate. We conclude that Sgr was enriched to  $[Fe/H] \sim -0.2$  before its most recent pericenter passage. If the “accreted halo” population is from one major accretion event, then this progenitor galaxy was enriched to at least  $[Fe/H] \sim -0.6$ , and had a similar star formation history to Sgr before merging.

**Key words:** galaxies: dwarf – galaxies: individual (Sagittarius dSph) – galaxies: stellar content

**Supporting material:** machine-readable tables

## 1. Introduction

The stellar halo of the Milky Way (MW) contains only a small fraction of the total stellar mass of our Galaxy, but serves as an important laboratory to study the accretion history of the MW.  $N$ -body simulations suggest that galaxies like our own are formed, at least in part, through hierarchical buildup from smaller stellar systems (e.g., Klypin et al. 1999). Therefore, accreted systems should be able to be observed in various states of dissolution as they are destroyed upon entering the MW potential well. In fact, deep photometric surveys have revealed that the distribution of stars that make up the MW stellar halo

actually contains spatially extended and relatively massive stellar substructure, typically in the form of stellar streams (see, e.g., Belokurov et al. 2006; Schlafman et al. 2011, 2012, or a review of known systems in Grillmair & Carlin 2016). Some of these streams can be linked to obvious progenitors, such as the Sagittarius dwarf spheroidal galaxy (Sgr, Ibata et al. 1994; Majewski et al. 2003), while other streams appear to be “orphans,” with the progenitor system unknown (e.g., Belokurov et al. 2007).

While strong evidence for substructure in the MW stellar halo has been observed, it is not clear as to how much of the halo was formed in situ versus being accreted, or even what fraction of the halo consists of disk stars that have been kicked out into highly eccentric orbits (see, e.g., Sheffield et al. 2012).

<sup>20</sup> NSF Astronomy and Astrophysics Postdoctoral Fellow.

Simulations have revealed that the stellar halo was likely formed through some combination of in situ star formation and accretion, with the inner halo ( $R \lesssim 20$  kpc) having a larger fraction of stars formed in situ than the outer halo (e.g., Bell et al. 2008; Zolotov et al. 2009; Font et al. 2011). Such features have been found in observations of the MW halo (e.g., Carollo et al. 2007; Beers et al. 2012). Other observations have revealed that there is a break in the density profile of the halo that is coincident with a change in slope of the metallicity gradient, which shifts from nearly flat to negative at  $R \gtrsim 20$  kpc (e.g., Allende Prieto et al. 2014; Fernández-Alvar et al. 2015; Xue et al. 2015), lending support to these simulations. Simulations of accretion scenarios find that the substructure should remain spatially and kinematically coherent throughout much of the lifetime of the MW (e.g., Johnston et al. 1996), so it is likely that more substructure will be identified as photometric and spectroscopic surveys of the entire sky continue (e.g., Pan-STARRS, *Gaia*, and LSST). In fact, recent works utilizing *Gaia* data suggest that the inner Galactic halo is dominated by debris from a massive, LMC-sized merger some 8–11 Gyr ago (e.g., Belokurov et al. 2018; Deason et al. 2018; Helmi et al. 2018; Koppelman et al. 2018) that may have also deposited several globular clusters into the MW potential (Myeong et al. 2018).

It is also possible to find halo substructure in chemical abundance space rather than position–velocity space. In principle, stars should contain a unique fingerprint in their detailed chemical abundance patterns that link them to the same molecular cloud of their birth, or at least, a molecular cloud that was polluted by metals in similar ways. If true, then this suggests that measuring the detailed chemical abundances of stars can reveal which stars were born from the same molecular cloud. This technique, known as “chemical tagging” (Freeman & Bland-Hawthorn 2002), requires large spectroscopic surveys that can observe hundreds of thousands, or even millions of stars, and that are capable of providing precise chemical abundances. Fortunately, surveys like these have recently been completed or are ongoing (e.g., RAVE, SDSS, LAMOST, GALAH, etc.), and this technique of chemical tagging can be explored (e.g., Mitschang et al. 2014).

Here, we focus on the Apache Point Observatory Galactic Evolution Experiment (APOGEE, Majewski et al. 2017). APOGEE has observed 176,000 red giant stars across the MW from the northern hemisphere, and as part of APOGEE-2, has recently begun observations in the Southern Hemisphere. By 2020, APOGEE will have detailed chemical abundances for  $\sim 500,000$  stars sampling the entire MW, opening up a vast Galactic volume to chemical tagging experiments. Because APOGEE observes in the near-infrared ( $H$ -band,  $1.5\text{--}1.7\ \mu\text{m}$ ) and targets red giant stars that are brighter in the NIR than in the optical, it is able to see stars as far as  $\sim 10$  kpc in the dusty plane of the MW and over 60 kpc out into the halo. This makes it an ideal instrument to study the accretion of the MW halo, especially if accreted populations do dominate at  $R > 20$  kpc.

Chemical tagging techniques have already been explored using the APOGEE data set. Hogg et al. (2016) applied a  $k$ -means clustering algorithm to show that abundances provided by the data-driven algorithm “The Cannon” (Ness et al. 2015) are precise enough to recover groups of stars that can be traced to Sgr and some globular clusters. However, Ness et al. (2017) found that chemical tagging, in its purest form of linking stars to the same birth molecular cloud, is not feasible with the

APOGEE data. They found that with the APOGEE abundances from The Cannon, where chemical elements are determined down to levels of 0.04 dex or less, chemically similar pairs of stars are more likely to be “doppelgangers” (stars that just happen to exhibit the same abundance patterns) rather than siblings actually born from the same molecular cloud. Despite this realization, so called “weak” chemical tagging (chemically tagging stars born in the same type of stellar system rather than the same unique stellar system) has been used in the APOGEE data set to discover field stars that exhibit abundance patterns similar to second-generation globular clusters (e.g., Fernández-Trincado et al. 2016, 2017; Schiavon et al. 2017), providing constraints on what fraction of the MW bulge was formed from accreted globular clusters. Hayes et al. (2018) also found that APOGEE stars with  $[\text{Fe}/\text{H}] < -0.9$  can be split into a high-magnesium population and a low-magnesium population, the latter of which is likely an accreted population based on its chemical abundance patterns, which are similar to those of dwarf galaxies.

In this work, we explore whether or not APOGEE has the ability to chemically tag stars that have formed in dwarf galaxies by performing a chemical tagging analysis targeted at the Sgr system. Observations as well as models of the Sgr system suggest that Sgr tidal debris covers a large fraction of the MW (e.g., Ibata et al. 1995; Majewski et al. 2003; Belokurov et al. 2006; Law & Majewski 2010; Tepper-García & Bland-Hawthorn 2018). The Law & Majewski (2010) model predicts that many Sgr stars should be coincident on the sky with fields observed by APOGEE. It has been shown that the chemical abundance patterns of Sgr are quite unique. The metal-rich Sgr stars ( $[\text{Fe}/\text{H}] > -0.8$ ) exhibit deficiencies in all chemical elements (expressed as  $[\text{X}/\text{Fe}]$ ) relative to the MW (see, e.g., McWilliam et al. 2013; Hasselquist et al. 2017). Recent work confirms these abundance patterns extend to the Sgr tidal streams as well (Carlin et al. 2018).

The paper is organized as follows. In Section 2 we explain the sample of APOGEE data in which we look for Sgr stream members. The process for identifying Sgr stream candidates in chemical abundance space is explained in Section 3. Results are presented in Section 4, where we present the discovery of 68 potential Sgr stream members (62 of which have *Gaia* proper motion measurements). Implications for chemical tagging methods as well as constraints on models for the Sgr system are described in Section 5.

## 2. Observations and Data Reduction

### 2.1. APOGEE

The APOGEE survey was part of Sloan Digital Sky Survey III (Eisenstein et al. 2011), and observed 146,000 stars in the MW galaxy (Majewski et al. 2017) from 2011 to 2014. APOGEE-2 began observations in 2014 as part of the Sloan Digital Sky Survey IV (Blanton et al. 2017), and the latest data release, DR14, contains APOGEE observations of an additional  $\sim 100,000$  stars in the Northern Hemisphere (Abolfathi et al. 2018). The APOGEE instrument is a high-resolution ( $R \sim 22,500$ ) near-infrared ( $1.51\text{--}1.70\ \mu\text{m}$ ) spectrograph described in detail in Wilson et al. (2019, in preparation). For the main survey the instrument was connected to the Sloan 2.5 m telescope (Gunn et al. 2006). Targeting strategies for APOGEE and APOGEE-2 are described in Zasowski et al. (2013, 2017), respectively.

The APOGEE data are reduced through methods described by Nidever et al. (2015), and stellar parameters/chemical abundances are extracted using the APOGEE Stellar Parameters and Chemical Abundances Pipeline (ASPCAP, García Pérez et al. 2016). ASPCAP interpolates in a grid of synthetic spectra (Zamora et al. 2015) generated from the APOGEE *H*-band line list (Shetrone et al. 2015) to find the best fit (through  $\chi^2$  minimization) to the observed spectrum by varying  $T_{\text{eff}}$ , surface gravity, microturbulence, metallicity, carbon abundance, nitrogen abundance, and  $\alpha$ -element abundance. In this analysis we use results from the 14th data release of SDSS (DR14, Abolfathi et al. 2017; Holtzman et al. 2018; Jönsson et al. 2018).

To identify Sgr stream stars in the DR14 sample, we use the Sgr core sample from Hasselquist et al. (2017) that was selected using the methods described in Majewski et al. (2013) as a representation of the chemical abundance patterns of Sgr. This sample contains 158 stars with spectra having signal-to-noise ratios (S/Ns)  $> 70$  per half resolution element, which the APOGEE DR14 documentation defines as the lower threshold for which the APOGEE detailed chemical abundances are reliably measured. Because the bulk of the Sgr members are in the low-temperature regime ( $T_{\text{eff}} < 4200$  K) where some elemental abundance determinations may exhibit systematic differences from higher temperature stars, we limit our MW sample in which we search for Sgr stream stars to the same stellar parameters of APOGEE Sgr stars. This sample is defined as follows:

1.  $3550 < T_{\text{eff}} < 4200$ .
2.  $S/N > 70$  per half resolution element.
3. No “ASPCAPBAD” flag set.<sup>21</sup>

Because the APOGEE Sgr core sample does not contain stars with  $[\text{Fe}/\text{H}] < -1.2$ , we also only consider MW stars with  $[\text{Fe}/\text{H}] > -1.2$ . Distances to the APOGEE stars are derived by the methods described in Hayden et al. (2014), and are accurate to  $\sim 20\%$ . These distances can be found as part of a value-added catalog (VAC) from DR14.<sup>22</sup> While the *Gaia* distances from Bailer-Jones et al. (2018) are more precise at close distances, at  $d > 5$  kpc, the distance uncertainties become larger than  $\sim 20\%$ , and the distances become prior-dominated. More than half of our APOGEE sample on which we perform the clustering analysis consists of stars with distances  $> 5$  kpc, so we adopt the spectrophotometric distances from the DR14 VAC to ensure our distance source is homogeneous across the entire distance range we are able to probe with the APOGEE stars.

## 2.2. Gaia

We supplement the APOGEE radial velocities and distance estimates with proper motions from *Gaia* Data Release 2 (DR2; Gaia Collaboration et al. 2018). Using the nearest neighbors matching tool within TOPCAT (Taylor 2005), we identify the best *Gaia* match (nearest neighbor within  $1''$ ) for each APOGEE star. Of the 68 chemical candidates we identify in Section 3, 62 have *Gaia* proper motion matches. All but two of the sources matched within  $0''.4$  (the other two are at  $0''.73$  and  $0''.86$ ). We

<sup>21</sup> This flag, described in detail in Holtzman et al. (2015), indicates whether a star falls near the edges of the synthetic spectra grids, or has a low S/N spectrum.

<sup>22</sup> [http://www.sdss.org/dr14/data\\_access/value-added-catalogs/](http://www.sdss.org/dr14/data_access/value-added-catalogs/)

then use the full 6D phase-space data (i.e., 3D positions and velocities) to integrate orbits in a model Galactic potential, and further check the membership of our APOGEE-selected sample in the Sgr stream (described in further detail in Section 4).

To further investigate whether we have recovered true Sgr stream stars, we use the proper motions from *Gaia* and the distances and radial velocities from APOGEE to calculate the orbits of the Sgr stream candidates. Orbits were integrated in the `MWPotential2014` potential that is implemented in the `galpy` (Bovy 2015) suite of Python routines.<sup>23</sup> This potential is a composite of a spheroidal bulge, a Miyamoto–Nagai disk, and a Navarro–Frenk–White halo (see the default parameters in Bovy 2015). We note that while many recent works suggest that the halo mass of `MWPotential2014` may be too low to match MW observations (e.g., Fritz et al. 2018; Gaia Collaboration et al. 2018; Watkins et al. 2018), we are mainly interested in the relative orbital properties of stars, so the absolute scale of the potential is unimportant. Throughout this work, we assume a circular velocity of  $220 \text{ km s}^{-1}$  at a Solar radius of 8.0 kpc from the Galactic center, and the Solar motion as measured by Schönrich et al. (2010). We follow the `galpy` usage of a left-handed Galactic Cartesian coordinate (velocity) system with  $X(U)$  positive toward the Galactic center,  $Y(V)$  along Galactic rotation, and  $Z(W)$  in the direction of the north Galactic pole. We initialize the orbit of each star at its current position and velocity, then integrate its orbit both forward and backward for 1.5 Gyr. From this, we estimate the orbital parameters, including the eccentricity, the pericenter and apocenter distances ( $r_{\text{peri}}$  and  $r_{\text{apo}}$ ; relative to the Galactic center), and orbital energy and angular momentum.

## 3. Selecting Sgr Stream Candidates in the APOGEE Data

### 3.1. Kinematic Control Sample

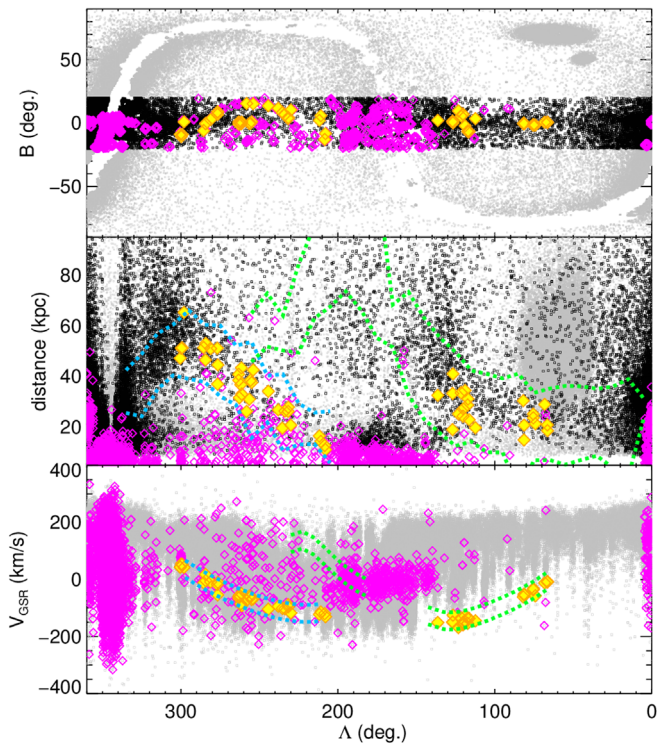
The goal of this work is to find Sgr members through chemistry alone, i.e., without any knowledge of kinematics or positions. However, to verify that the clustering algorithm used in this work is properly returning Sgr stream members by recovering the most obvious Sgr stream members, we select a “kinematic” control sample of likely Sgr stream stars from the APOGEE data. Sgr stream stars in the direction out of the MW plane can be easily distinguished from MW stars based on velocities and color.

The method for selecting the kinematic control sample is as follows:

1. We begin by applying cuts on distance, color, extinction, and proper motion to the entire APOGEE DR14 catalog (using distances from the VAC’s “NMSU\_dist” estimator). We transform the coordinates into the Sagittarius stream-aligned system devised by Majewski et al. (2003),<sup>24</sup> and only select stars with  $|B| < 20^\circ$  (must be within  $20^\circ$  of the Sgr debris plane). To avoid nearby disk stars, we require  $d > 5$  kpc. We further limit the sample to red objects with  $(J - K)_0 > 0.8$ , with low total extinction  $E(B - V) < 0.75$ , and minimal proper motion (to remove nearby high proper motion stars)

<sup>23</sup> Available at <https://github.com/jobovy/galpy>.

<sup>24</sup> In the coordinate system aligned with the Sagittarius stream, the coordinate  $\Lambda$  is the angle along the stream, increasing along the trailing Sgr debris stream, and  $B$  is defined as the angular distance from the Sagittarius debris plane defined by 2MASS M-giants, analogous to Galactic latitude ( $B = 0^\circ$  is the Sagittarius debris plane).



**Figure 1.** Illustration of our process to select Sgr stream candidates based on their kinematics. For comparison, the upper two panels show (as gray points) photometrically selected M-giant candidates based on 2MASS+WISE color-color criteria (as in Carlin et al. 2018; see also Koposov et al. 2014; Li et al. 2016); distances are estimated based on the photometric parallax relation derived for Sgr stream M-giants by Li et al. (2016). In the bottom panel, the gray points are all APOGEE DR14 velocities. The small black points are a subset of the photometric sample selected based on position to be consistent with Sgr stream membership; the Sgr stream is clear in the distance panel (middle) as two swaths of black points arcing from  $320^\circ \gtrsim \Lambda_\odot \gtrsim 200^\circ$  (leading arm) and  $200^\circ \gtrsim \Lambda_\odot \gtrsim 30^\circ$  (trailing tail), mostly at distances beyond  $\sim 20$  kpc. Applying the criteria detailed in Section 3.1 to the APOGEE DR14 catalog yields the magenta sample in all three panels. Finally, we sub-select from the APOGEE sample using the known velocity (Belokurov et al. 2014) and distance (Hermitschek et al. 2017) trends of the Sgr stream. The dashed lines represent the selections applied; each of these encompasses the  $3\sigma$  range about the measured mean values. Our sample of kinematically selected Sgr candidates contains 58 stars, which are shown as yellow/orange diamonds in all three panels above.

$|\mu_{\alpha \cos \delta}, \mu_\delta| < 15 \text{ mas yr}^{-1}$ . The resulting sample is plotted as magenta points in all three panels of Figure 1.

- From this set of APOGEE candidates, we use the known properties of the Sgr stream to select candidates (also retaining only stars whose APOGEE spectra have  $S/N > 30$ ). The selections are based on velocity trends derived from SDSS members (Belokurov et al. 2014, dashed lines in the lower panel of Figure 1), and RR Lyrae distances from PanSTARRS PS1 (Hermitschek et al. 2017, dashed lines in center panel). For each of these sub-selections, we include a range of  $\pm 3\sigma$  about the mean values from each study, using their reported  $1\sigma$  errors. This yields a final sample of 58 kinematically selected Sgr stream candidates, which are shown as yellow/orange diamonds in Figure 1. Of these 58 stars, 21 have sufficient S/N and stellar parameters set by our criteria above, which we adopt as the “kinematic” sample.

This is not a complete sample of Sgr candidates from the APOGEE DR14 database, as the selection criteria are limited to regions with previous measurements. However, it is a reliable set of candidates to use as a test of our chemical tagging selection. We have confirmed that these selection criteria recover the 42 Sgr stream stars observed at high spectral resolution in the optical by Carlin et al. (2018).

### 3.2. k-means Clustering

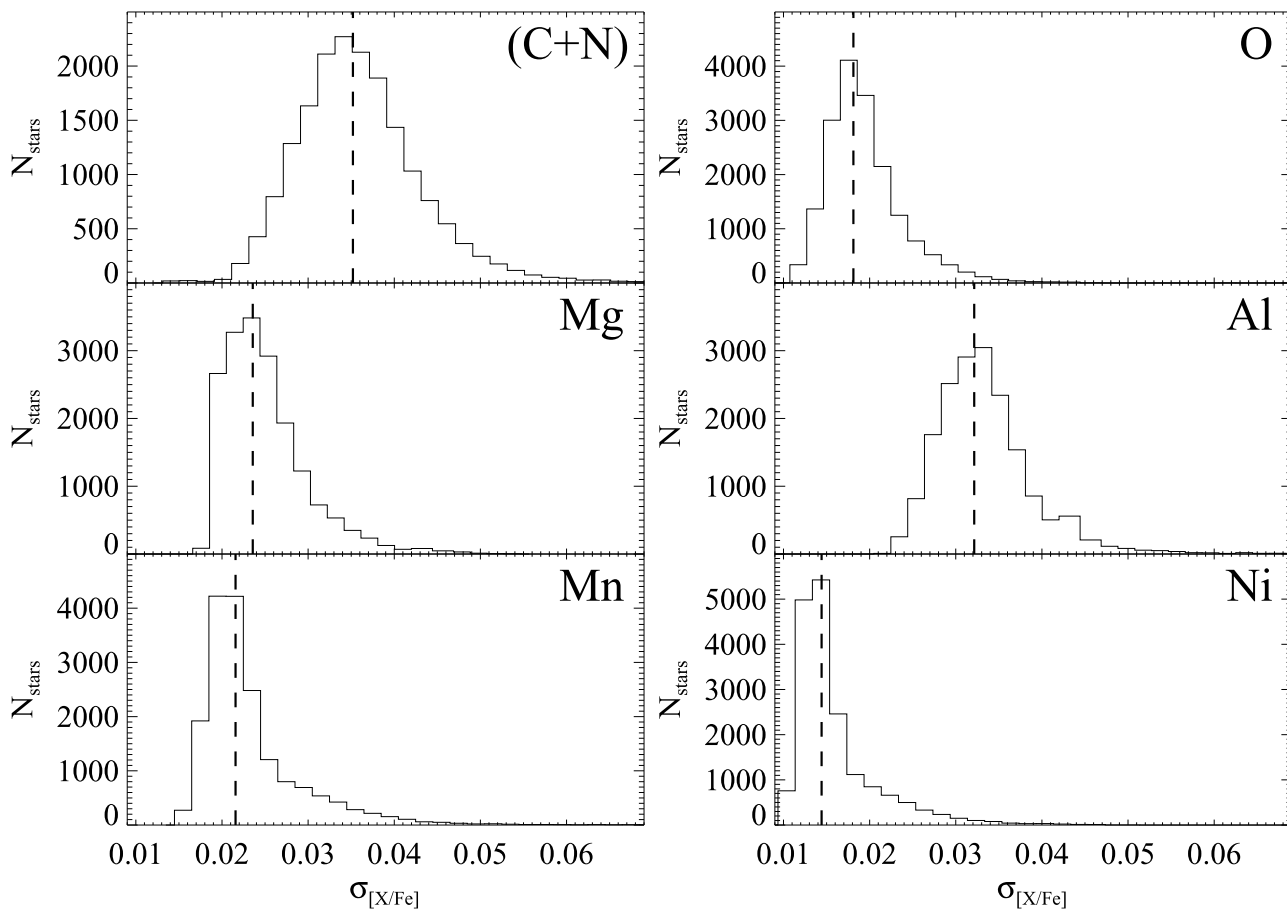
To conduct our search for Sgr stream members in chemical abundance space, we use the  $k$ -means “CLUST\_WTS” function<sup>25</sup> in IDL. This function randomly chooses the starting points of  $k$  clusters and then moves stars in and out of clusters to minimize variance within each cluster and maximize variance between clusters. We set this function to run with 100 different random starting guesses, which we find consistently returns the same clusters for  $k < 50$ . We choose the APOGEE elemental abundances in which Sgr exhibits clear deficiencies, and have low reported uncertainties from the APOGEE pipeline ( $\sim 0.05$  dex or less). The  $N$ -dimensional space in which we cluster is then defined by  $[O/Fe]$ ,  $[(C+N)/Fe]$ ,  $[Mg/Fe]$ ,  $[Al/Fe]$ ,  $[Mn/Fe]$ , and  $[Ni/Fe]$ . The uncertainties for each element are shown in Figure 2. The median uncertainties of each element are used as weights in the  $k$ -means clustering analysis such that elements with lower uncertainties, such as Ni and O, are given larger weight than elements with higher uncertainties, such as Al and C+N.

### 3.3. Selecting $k$ Clusters

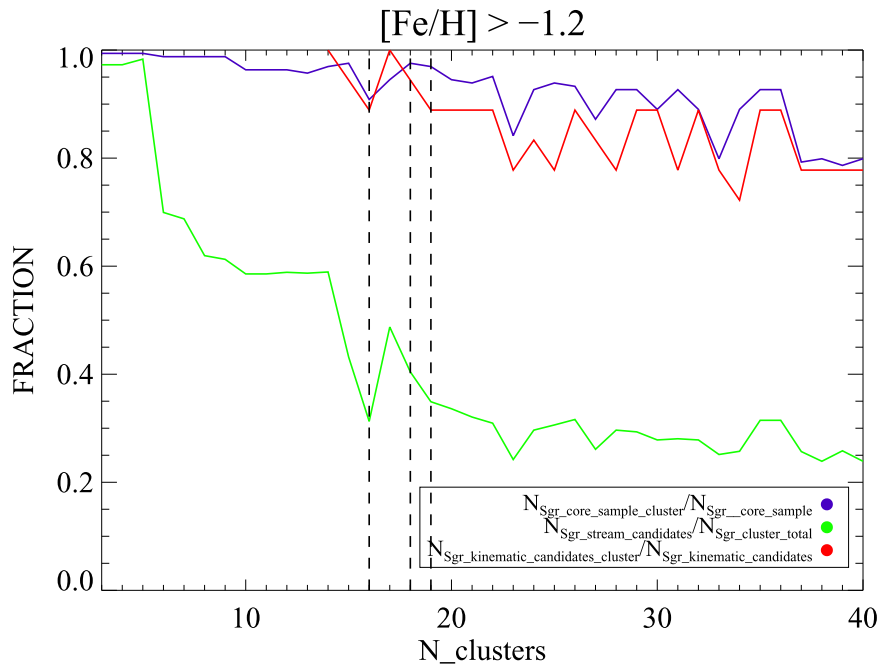
With a kinematic sample of expected stream members and the core sample defining the abundance patterns of Sgr, we determine the number of clusters,  $k$ , in which to separate the data. To select an appropriate value for  $k$ , we run the clustering algorithm for multiple values of  $k$  from 3 to 50. For each value of  $k$ , we track the Sgr core sample from Hasselquist et al. (2017) and the kinematic control sample defined above. The cluster that contains the bulk of the Sgr core sample is referred to as the “Sgr cluster,” and should contain the majority of the kinematic control sample if we are to identify potential stream members. We choose the ideal value of  $k$  to be the number of clusters that results in a marginal splitting (10%) of the Sgr core sample from the Sgr cluster. In this way, we are “tuning” the  $k$ -means algorithm.

The motivation for our choice of  $k$  is shown graphically in Figure 3, where we plot the fraction of stars retained in the Sgr cluster as a function of  $k$  for each sample. The blue line, which represents the fraction of Sgr core members retained in the Sgr cluster, exhibits a decrease at  $k = 16$ , which corresponds with a similar decrease in the fraction of kinematic sample retained (red line). We adopt  $k = 16$  as our best number of clusters in which to separate the data, but note that 18 and 19 are just as viable choices based on our criteria set here. However, these choices result in very little difference in the amount of “chemical” stream candidates; the fraction of the Sgr cluster that is made up of stream candidates (stars that are not Sgr core members) is denoted by the green line in Figure 3.

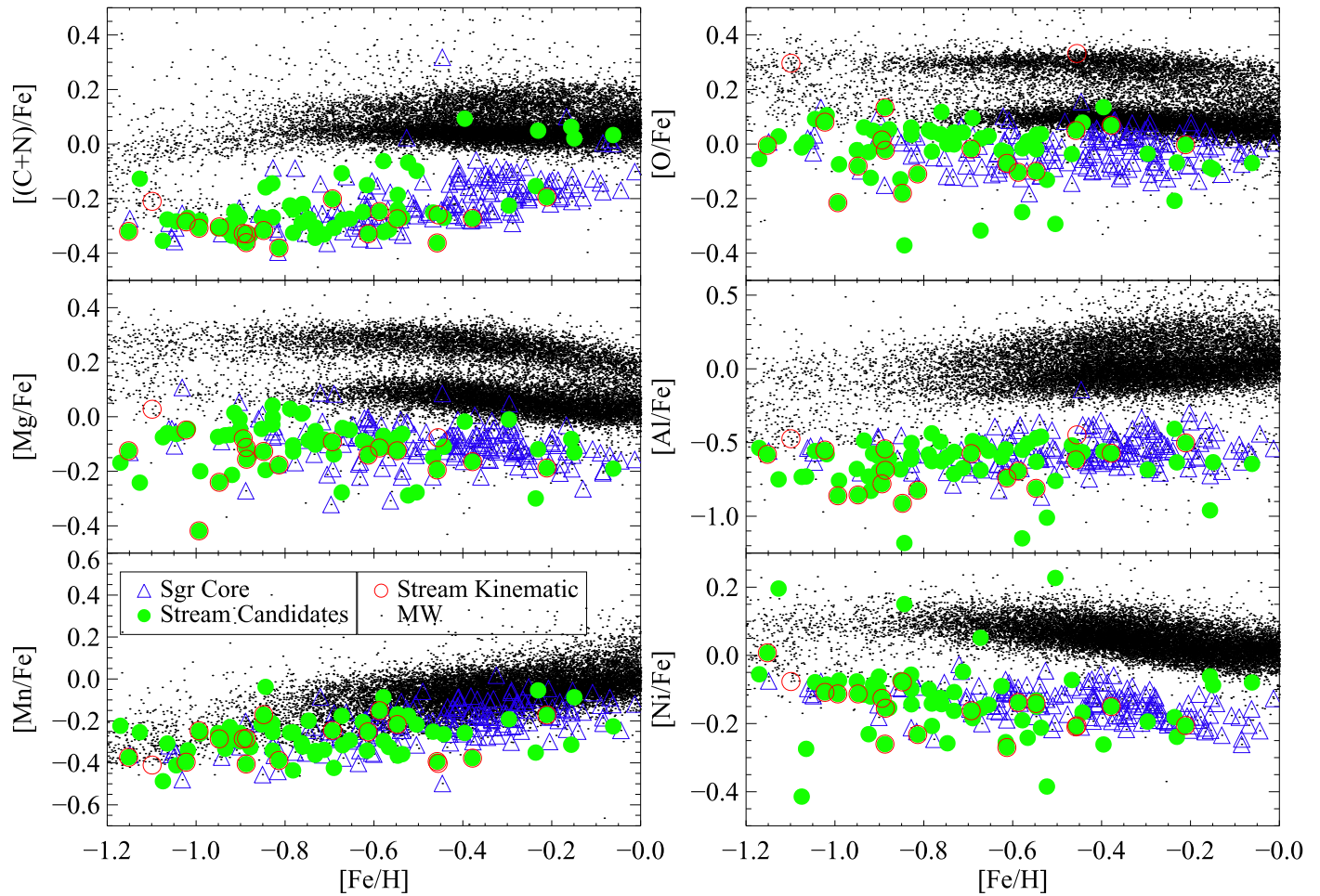
<sup>25</sup> [http://www.harrisgeospatial.com/docs/CLUST\\_WTS.html](http://www.harrisgeospatial.com/docs/CLUST_WTS.html)



**Figure 2.** Distributions of the APOGEE chemical abundance uncertainties for the sample on which we perform our clustering analysis. The dashed lines mark the median uncertainty, which we adopt as weights in the  $k$ -means clustering analysis.



**Figure 3.** Ratios of stellar samples plotted as a function of  $k$  clusters. The lines are colored according to the legend in the lower right panel. The dashed lines mark the best-choice values for  $k$ , defined as where 10% of the Sgr core sample separates from the “Sgr cluster.” In this work, we present results for  $k = 16$ .



**Figure 4.** Two-dimensional chemical abundance projections for the APOGEE sample on which the clustering algorithm was performed. Points are colored according to the legend in the lower left panel.

#### 4. Results

Performing our clustering analysis with  $k = 16$  results in a cluster containing 90% of the Sgr core sample along with an additional 68 stars that do not reside in the Sgr core. Of the 21 kinematic candidates selected in Section 3.1, 19 are in this sample of 68, demonstrating that our method reliably selects Sgr stream stars. These 68 stars are henceforth considered to be “Sgr stream candidates.” The chemical abundance patterns for the Sgr core and Sgr stream candidates are shown in Figure 4, along with the chemical abundance patterns of the entire MW sample on which the clustering algorithm was performed.

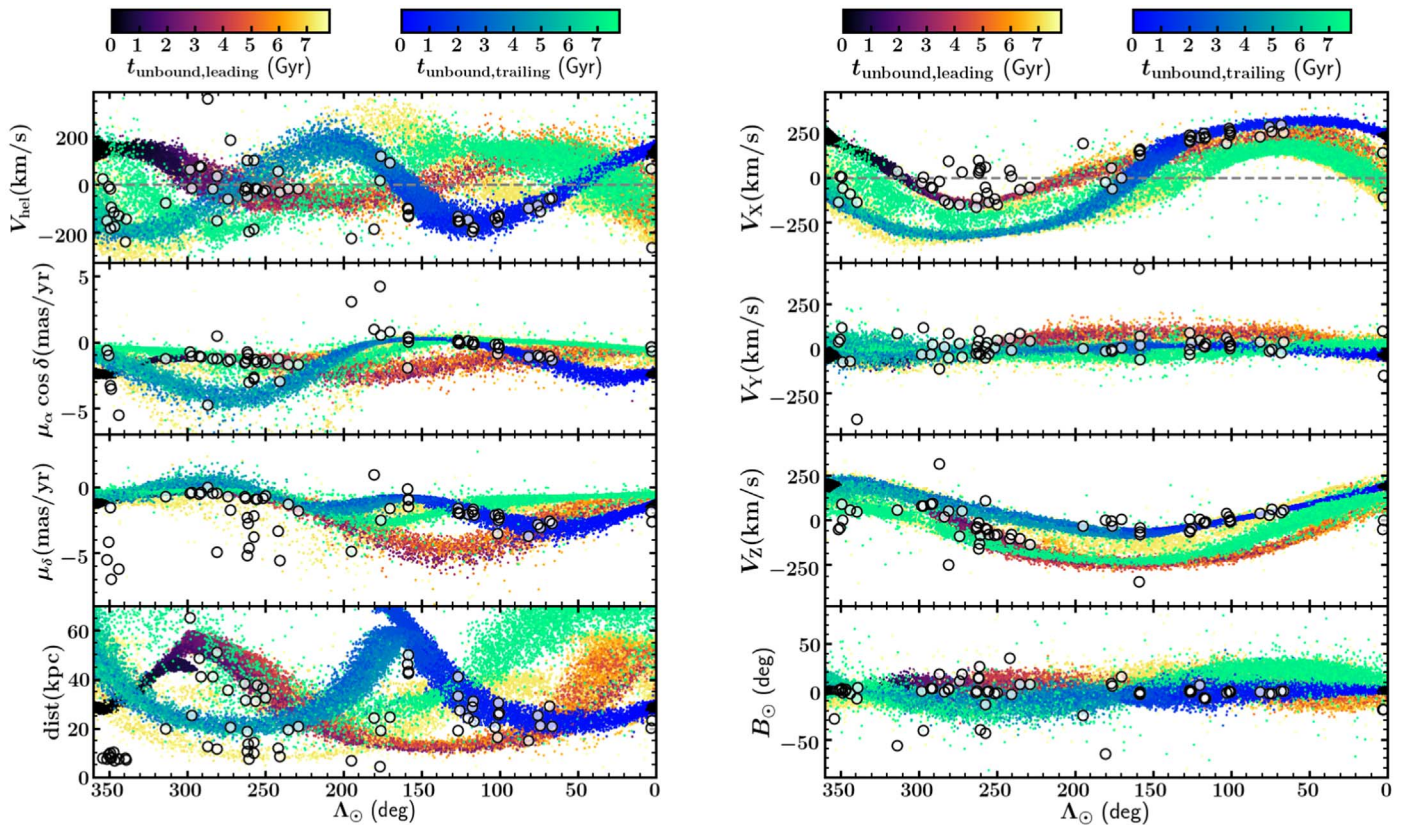
The Sgr stream candidates fall in the same regions of these 2D abundance projections as the Sgr core sample, with the exception of a handful of cases as we now describe. There are 4 stars with  $[\text{Fe}/\text{H}] > -0.25$  that exhibit  $[(\text{C}+\text{N})/\text{Fe}]$  abundances consistent with the MW trend, unlike the majority of Sgr core stars at that metallicity. However, there are three Sgr core stars that also occupy this region of  $[(\text{C}+\text{N})/\text{Fe}]$ – $[\text{Fe}/\text{H}]$  abundance space. There are also four stars with enhanced  $[\text{Ni}/\text{Fe}]$  that do not follow the bulk Sgr core trends.

The Sgr stream candidates are generally more metal-poor than the Sgr core, which is consistent with previous abundance studies of the Sgr stream (see, e.g., Chou et al. 2007; Carlin et al. 2018). In fact, the stream has stars more metal-poor than our cutoff of  $[\text{Fe}/\text{H}] > -1.2$  (see, e.g., Chou et al. 2007), so

we are likely missing metal-poor Sgr stream members that reside in the APOGEE data (discussed further in Section 5).

Using the *Gaia* proper motions, we compare the positions and velocities of the Sgr stream candidates to those predicted for Sgr tidal debris from the Law & Majewski (2010) model of the Sgr system. This  $N$ -body model fits the positions and kinematics of the most recent Sgr tidal debris by introducing a non-axisymmetric component to the MW gravitational potential. The comparisons are shown in Figure 5. Sgr stream candidates are plotted as white circles superposed on the model debris from Law & Majewski (2010), which are colored by when the debris was stripped from Sgr. From these comparisons, we find that with the exception of a group of stars clearly residing in the bulge ( $\Lambda_{\odot} \sim 350^{\circ}$  and  $d = 8$  kpc), most of the stream candidates have positions and velocities consistent with those predicted by the model. One obvious exception is the group of stars with  $240 < \Lambda_{\odot} < 290$  with  $U$  (i.e.,  $V_x$ ) velocities that are higher than those predicted by the model.

In the top row of Figure 6 we plot the orbital eccentricity as a function of orbit apocenter ( $r_{\text{apo}}$ ) as calculated in the `MWPotential2014` gravitational potential for the chemically selected Sgr stream candidates. In the left column, the points are colored by the Sgr  $B$  coordinate, such that the color saturation corresponds to the distance from the Sgr debris plane; stars with white or light-colored shading are near the Sgr



**Figure 5.** Positions and velocities of the Sgr stream candidates (white circles) compared to the predicted positions and velocities of the Law & Majewski (2010) model of the Sgr system (colored dots). Left: comparisons of the APOGEE radial velocities, *Gaia* proper motions, and APOGEE distances plotted as a function of  $\Lambda_{\odot}$ . Right: comparisons of calculated UVW velocities and  $|B|$  as a function of  $\Lambda_{\odot}$ . The Law & Majewski (2010) points are colored according to whether they belong to the leading/trailing debris, as well as when they were stripped from Sgr, as indicated by the color bar above.

orbital plane, while those that are darker red or blue are far from the Sgr plane. In the right column the points are colored by the Sgr  $\Lambda_{\odot}$  coordinate, such that the blue and red points indicate whether the stars have  $\Lambda_{\odot}$  consistent with Sgr trailing arm or leading arm debris, respectively.

Analyzing the stars in this way, we find that the Sgr stream candidates split into three orbital groups along with one hypervelocity star (HVS). The catalog of APOGEE IDs and membership assignments is shown in Table 1.

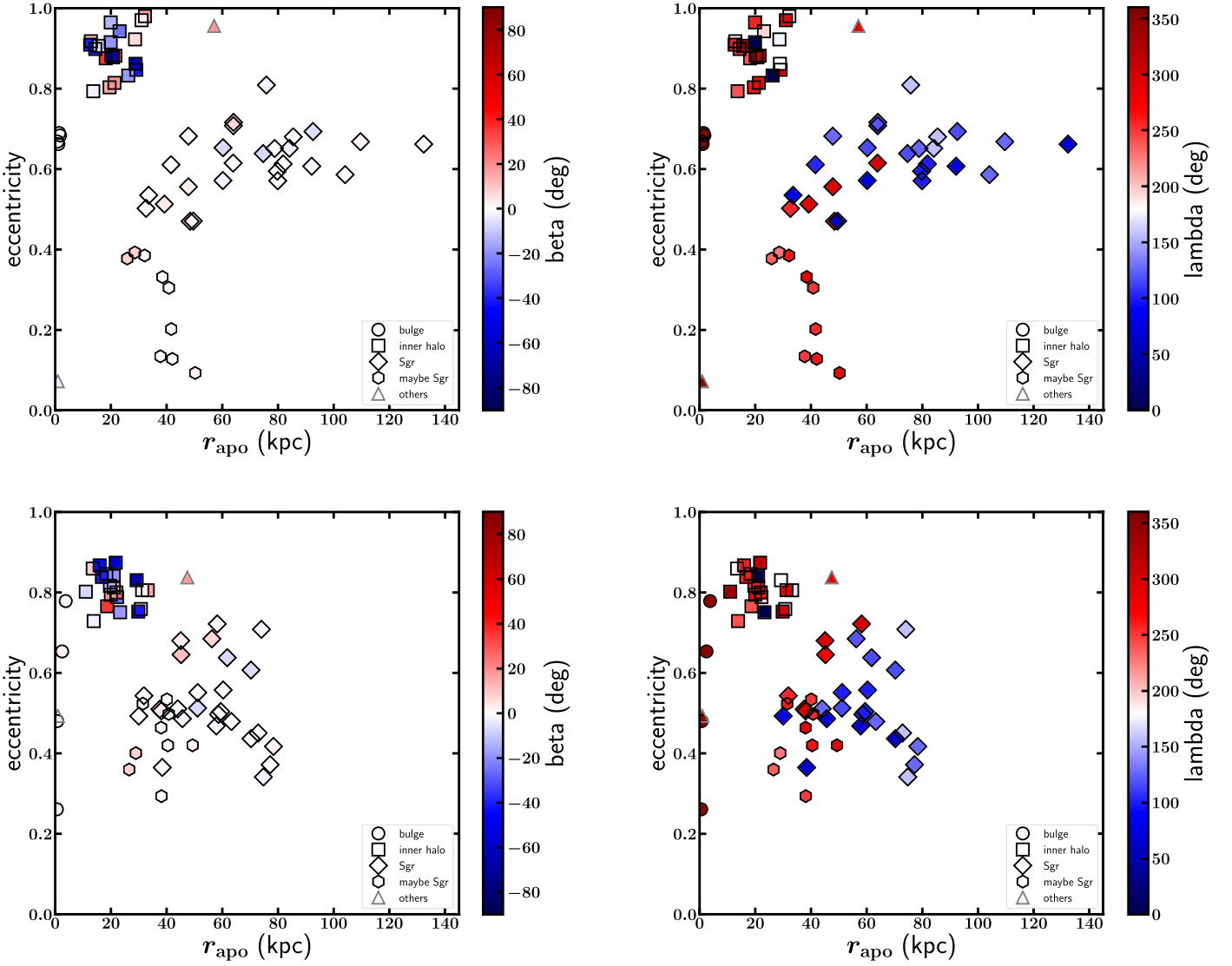
1. Sgr stream (35 stars)—this includes 26 stars (diamonds in Figure 6) with eccentricities between 0.4–0.85 and apocenters between 25–140 kpc that correspond to typical orbital properties of Sgr debris. The nine additional points shown as hexagons in Figure 6 are on more circular orbits than the main Sgr sample, but are all near  $B = 0^{\circ}$  and have apocenters consistent with Sgr debris. We call these “maybe Sgr,” but believe they are likely to be related to the stream.
2. Accreted halo (20 stars)—a distinct group of stars (shown as squares in Figure 6) that have high eccentricities ( $>0.75$ ) and apocenters of  $\sim 15$ –30 kpc stands out in Figure 6. Note that stars in this category appear on both sides of the Sgr plane, and are mostly located at  $\Lambda > 200^{\circ}$  in the North Galactic cap.
3. Bulge (4 stars)—four stars (circles in Figure 6) are on eccentric orbits, but with very small apocenters, as is typical of stars in the Galactic bulge.
4. HVS (1 star)—this star does not appear on this plot because it has no measured apocenter (at the last time step

of our orbit integrations, it is  $\sim 760$  kpc from the Galactic center)—its observed velocities suggest it is unbound in `MWPotential2014`.

5. Others (2 stars)—Two stars do not fit into any of the groups we identified above; we classify these simply as “others” and do not consider them further.

As previously noted, the choice of gravitational potential has little to no effect on our results. We classified stars based on their *relative* orbital parameters; groupings in parameter space such as eccentricity versus  $r_{\text{apo}}$  are likely to be similar even when orbits are integrated in different potentials. Nonetheless, because we compare to the Law & Majewski (2010) model in Figure 5, we integrate orbits in the potential that was used in that work. To do so, we use the `gala`<sup>26</sup> software (Price-Whelan et al. 2017; Price-Whelan 2017), which has explicitly implemented the Law & Majewski (2010) potential. Orbits integrated in the Law & Majewski (2010) gravitational potential yield the results in the lower panels of Figure 6. The different shapes in these panels are based on the classifications from the `MWPotential2014` results in the upper panels. While the positions of individual stars shift between the upper and lower panels in Figure 6, the classifications are retained. In fact, it is encouraging to see that the “maybe Sgr” sample overlaps the “Sgr stream” sample more when integrated in the LM10 potential. Finally, we note that while the uncertainties on most stars’ orbits are rather large, the separation between populations in Figure 6 is

<sup>26</sup> Available at <http://gala.adrian.pw>.



**Figure 6.** Calculated orbital eccentricity plotted as a function of orbit apocenter, colored by  $|B|$  (left) and  $\Lambda_{\odot}$  (right). The point symbols correspond to which orbital group the stream candidates belong to, as described in the text, and indicated in the legend. The top row shows orbital results using the MWPotential2014 gravitational potential, and the bottom row shows orbital results using the gravitational potential used in Law & Majewski (2010).

**Table 1**  
Sgr Stream Cluster

APOGEE ID	Membership <sup>a</sup>	$r_{\text{apo}}$ (kpc)	$\sigma_{r_{\text{apo}}}^+$	$\sigma_{r_{\text{apo}}}^-$	$e$	$\sigma_e^+$	$\sigma_e^-$
2M00151324-1430281	a	60.20	28.70	15.87	0.57	0.08	0.05
2M00194682-1345014	a	92.10	165.03	49.75	0.61	0.20	0.17
2M00445288-1244488	a	33.57	6.87	4.83	0.53	0.02	0.00
2M01512785-0352240	a	79.75	112.74	32.76	0.59	0.15	0.03
2M01532039-0327406	a	81.86	120.18	34.69	0.61	0.15	0.03
2M01535096-0320382	a	79.90	127.77	33.77	0.57	0.16	0.01

**Notes.**

<sup>a</sup> a = Sgr stream, b = accreted halo, c = bulge, d = HVS, e = others.

(This table is available in its entirety in machine-readable form.)

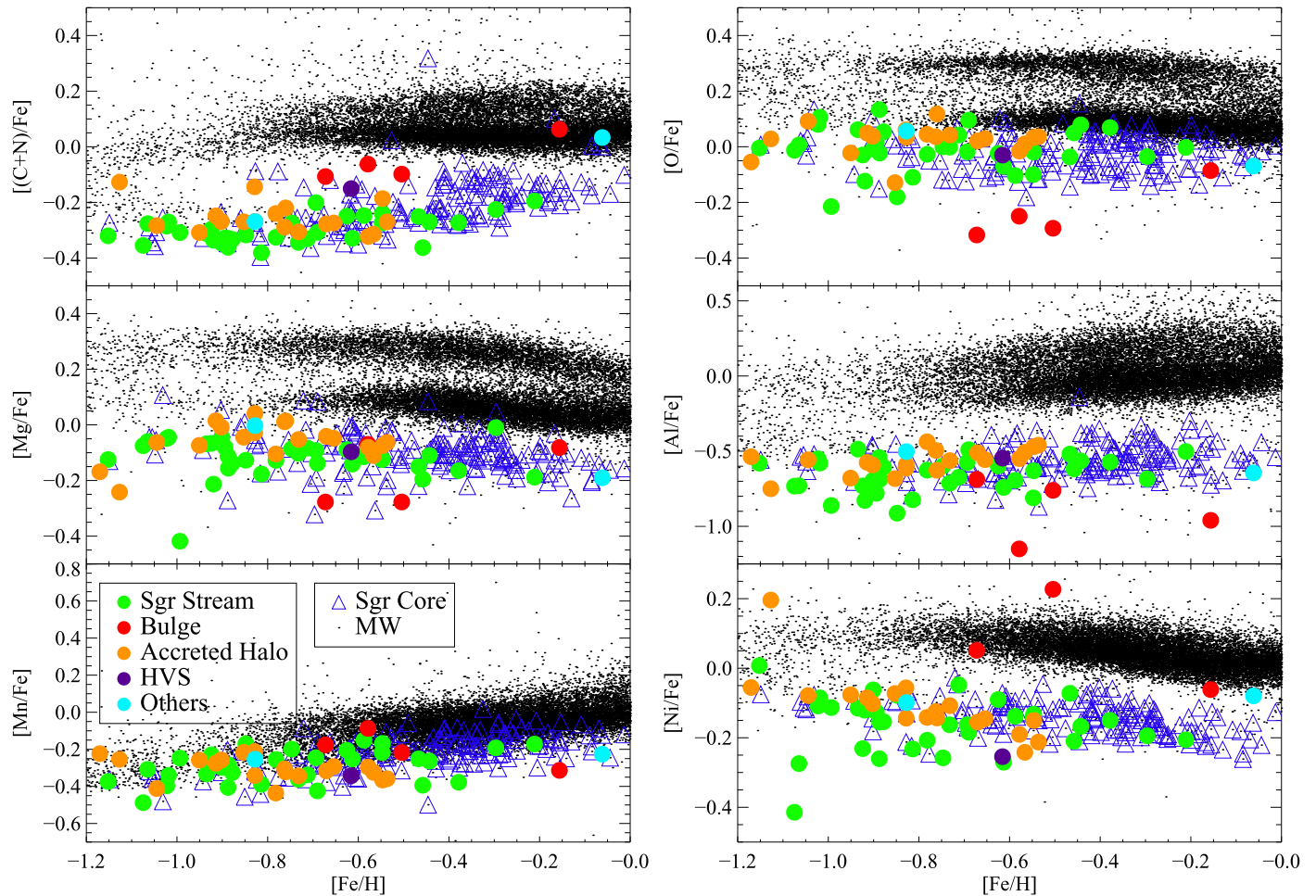
large enough that errors would have little effect on our classifications.

In Figure 7 we plot again the 2D abundance patterns, this time coloring the points based on the orbit divisions described above. We find that the apparent chemical outliers mentioned above appear to reside in the bulge sample. In the following

sections, we describe each sample in more detail and comment on whether or not they can be associated with the Sgr system.

#### 4.1. Sgr Stream

We find that 35 stars (including 19 stars from the kinematic control sample) of the 62 chemical candidates with *Gaia* proper



**Figure 7.** Same as in Figure 4 except the Sgr stream candidates are now colored according to the orbital group defined in the text (see the legend in the lower left panel).

motions follow the expected positions and kinematics of the Law & Majewski (2010) model of the system. Of these, 21 stars belong to the trailing arm and 14 stars belong to the leading arm, including all  $\sim 9$  of the low-eccentricity “maybe Sgr” stars. This is indicated by the coloring in the top right panel of Figure 6. According to this model, these stars were all stripped in the most recent pericenter passage. In Figure 6 we find that the trailing debris exhibit a range of apocenters from 60 to 120 kpc, and the leading arm debris have apocenters of 40–60 kpc. These are more in line with the apocenters found by Belokurov et al. (2014) and Fardal et al. (2018;  $\sim 48$  kpc and  $\sim 100$  kpc, respectively).

We find Sgr stream stars as metal-rich as  $[\text{Fe}/\text{H}] = -0.2$ , with the majority of stream stars exhibiting  $[\text{Fe}/\text{H}] < -0.6$ . The mean metallicity of our Sgr stream sample is more metal-poor than the Sgr core sample, suggesting that there was a radial metallicity gradient across the Sgr progenitor when it began merging with the MW such that the more metal-poor stars were stripped first.

#### 4.2. Accreted Halo

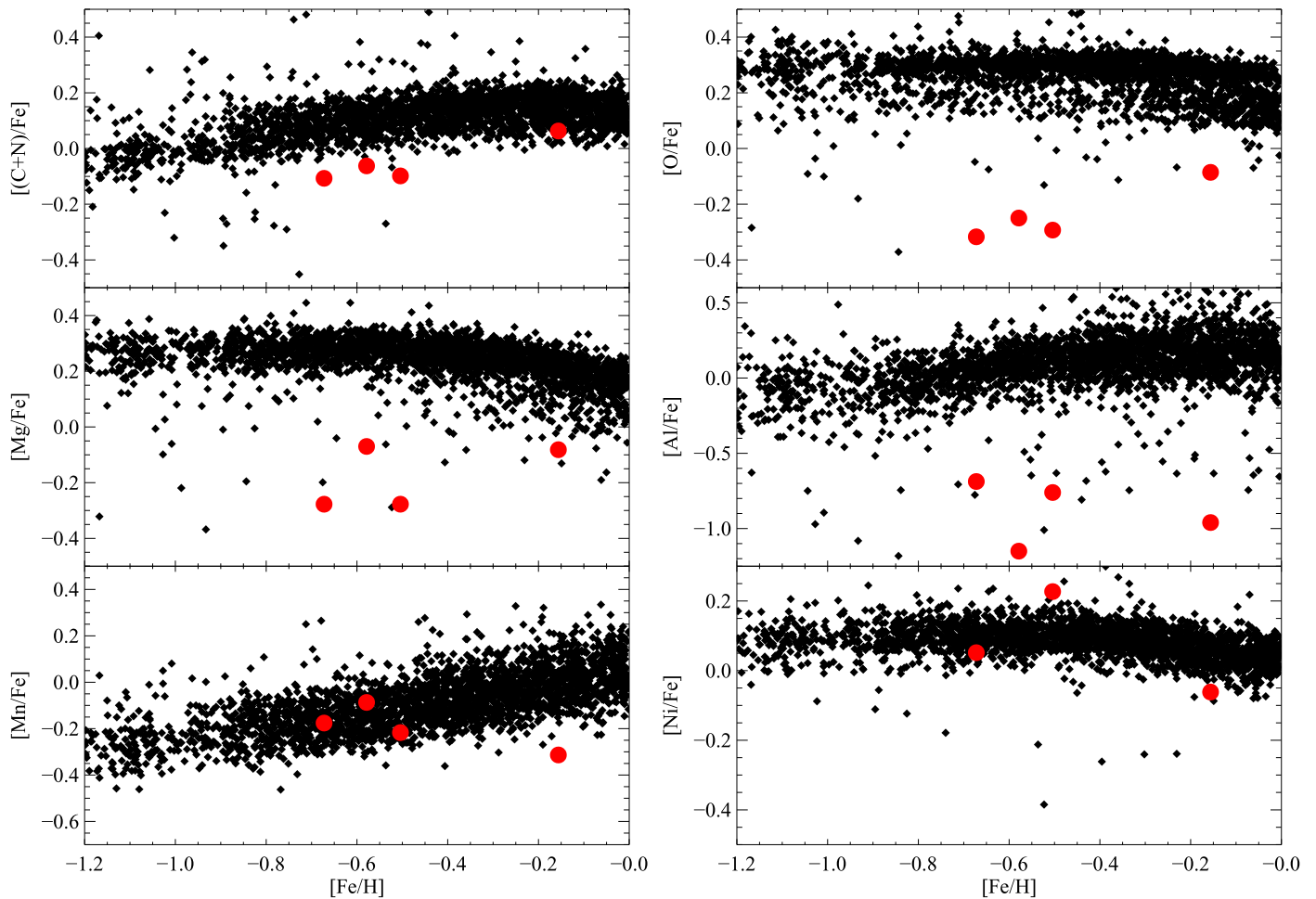
In Figure 6, we define the group of stars that have highly eccentric orbits and median  $r_{\text{apo}} = 20.2$  kpc as the “accreted halo” stars. These stars are similar to the accreted halo stars described in the literature (e.g., Schuster et al. 2012; Belokurov et al. 2018; Deason et al. 2018), where multiple studies

concluded that the pileup of metal-rich ( $[\text{Fe}/\text{H}] > -1.5$ ) halo stars at apocenters of  $\sim 20$ –25 kpc originates from a major merger some 8–11 Gyr ago. They argued that a massive satellite interacting with the disk can deposit stars on highly radial orbits (see, e.g., Belokurov et al. 2018). Further support for this interpretation can be found elsewhere in the literature (see, e.g., Helmi et al. 2018; Kruijssen et al. 2018; Mackereth et al. 2019, and Chiba & Beers 2000; Brook et al. 2003 for earlier reports), although we explore other potential origins of these stars in Section 5.2.

The accreted halo stars are unlikely to be Sgr stream stars, as they generally have  $|B|$  values that put them well outside the predicted plane of Sgr debris. However, given the nature of our analysis, these stars have the same chemistry as the Sgr stream stars. This suggests that the dwarf galaxy that merged with the MW some 8–11 Gyr ago may have a very similar star formation history (SFH) as Sgr, at least until an enrichment of  $[\text{Fe}/\text{H}] \sim -0.6$ , the highest metallicity among “accreted halo” stars.

#### 4.3. Bulge

There are four Sgr chemical candidates that actually reside in the Galactic bulge and are on bulge-like orbits. Thus, they are unlikely to be actual stream members. We find that by increasing  $k$  clusters, we cannot remove these stars from the Sgr chemical candidate cluster without removing likely stream



**Figure 8.** Two-dimensional chemical abundance plots of the bulge stars that were tagged as Sgr chemical candidates (red), compared to a bulge subsample of our MW comparison sample selected as MW stars with  $R < 3$  kpc (black).

members, suggesting a potential limitation to using the  $k$ -means technique to conduct chemical tagging analysis.

These stars are likely grouped with Sgr stream stars due to their exceptionally deficient  $[\text{O}/\text{Fe}]$  abundance ratios, as well as deficient  $[\text{Mg}/\text{Fe}]$  and  $[\text{Al}/\text{Fe}]$ , as seen in Figure 7. They may not be Sgr stream stars, but they also are not typical bulge stars. In Figure 8 we compare the bulge stars (red) to a strictly “bulge” MW sample, which is a subsample of our previous MW comparison sample (MW stars with the same parameters as the Sgr core sample), but with  $R_{\text{Gal}} < 3$  kpc. While there are some bulge stars with similarly deficient abundance ratios, the bulge Sgr stream candidates are deficient by  $\sim 0.5$  dex in  $[\text{O}/\text{Fe}]$ , by  $\sim 0.4$  dex in  $[\text{Mg}/\text{Fe}]$ , and by  $\sim 0.7$  dex in  $[\text{Al}/\text{Fe}]$  compared to the bulk of the bulge stars. They appear to have the same  $[\text{Mn}/\text{Fe}]$ , and are slightly deficient in  $[(\text{C}+\text{N})/\text{Fe}]$ .

The bulge Sgr stream chemical candidates actually have  $[\text{O}/\text{Fe}] \sim 0.2$  dex lower than the actual Sgr stream candidates. They are also enhanced in Ni and Mn relative to the Sgr stream candidates, with Ni and Mn abundances that more closely track the other stars in the bulge. Two of these stars actually have enhanced  $[\text{Ni}/\text{Fe}]$  relative to the bulge stars. This suggests that these may be some class of stars that formed from material containing excess SNe Ia yields as compared to the majority of the MW bulge stars.

While determining the origin of these stars is beyond the scope of this work, the presence of these stars in the Galactic

bulge suggests that either there was an epoch of star formation in the bulge that formed stars from gas that was unusually enriched in SNe Ia, or the MW underwent a merger of a dwarf galaxy that was as metal-rich as Sgr, but formed stars from gas that had a higher ratio of Type Ia/Type II ejecta than Sgr.

#### 4.4. HVS

One star in the Sgr chemical sample is at a Galactocentric distance of  $\sim 54$  kpc, near the position of the most recent apocenter of the LM10 Sgr orbit, which happened  $\sim 450$  Myr ago. Contrary to expectations for tidal debris at apocenter, this star has a total Galactocentric space velocity of  $\sim 577_{-129}^{+209}$   $\text{km s}^{-1}$ . In fact, an orbit integration suggests that this star’s closest approach will occur in  $\sim 20$  Myr, at  $r_{\text{peri}} \sim 53$  kpc, and that it is completely unbound from the MW. If this HVS’s orbit does not trace back to the Galactic center, then what might its origin be? The backward integration of its orbit carries it from its current location near the Galactic anticenter in the third quadrant of the North Galactic Hemisphere, meaning that its orbit passes nowhere near the LMC or M31. Thus, if this star is an interloper from beyond the MW, its origin will require further study.

Given that this HVS candidate is an early M-giant, it is unlikely to originate as a “runaway” in a binary ejection scenario (e.g., Blaauw 1961), because this formation mechanism requires young, massive stars. We have already ruled out

scenarios involving ejection from a black hole at either the center of the MW, M31, or the LMC. Thus, the most likely origin of this distant HVS is a dynamical interaction (e.g., between the components of a triple system, or between a binary and a passing massive star; e.g., Leonard & Duncan 1990; Gvaramadze et al. 2009). However, these mechanisms often require a very massive star to perturb the system; given the old, metal-poor nature of the Sgr debris (and the surrounding halo stars) in the region of sky where the HVS candidate is located, this mechanism also seems unlikely to have created this HVS candidate.

There is some support in the literature for HVSs originating from mergers. Abadi et al. (2009) found that disrupting dwarf galaxies can contribute HVSs; however, the mechanism requires that the HVSs were stripped from their host dwarf on a recent pericentric passage (i.e., near the Galactic center). Given that the star we have found never goes within  $\sim 50$  kpc of the Galactic center, this scenario is unlikely. Similarly, Piffl et al. (2011) suggested that massive dwarf galaxy mergers ( $M > 10^9 M_\odot$ ) can result in a population of HVS that is essentially the high-velocity tail of the debris. However, at the apocenter of the Sgr orbit, where we find the HVS candidate, the mean Galactocentric velocity should be roughly zero. A high-velocity tail relative to the mean velocity is unlikely to produce a star with velocity  $> 600 \text{ km s}^{-1}$ . Again, the simulations discussed by Piffl et al. (2011) require the stars to be debris stripped on the most recent pericentric passage, which seems to be ruled out because our candidate HVS’s orbit does not trace back along the Sgr orbit.

It is possible that this candidate HVS simply has a systematic error in its proper motion, which contributes most of the star’s 3D space velocity (though its  $v_{\text{gsr}} = -111 \text{ km s}^{-1}$  alone is much larger than expected at the Sgr orbital apocenter). This star’s *Gaia* proper motions are  $(\mu_{\alpha\cos\delta}, \mu_\delta) = (-1.91, -0.10) \pm (0.62, 0.44) \text{ mas yr}^{-1}$ , and its distance error is  $\sim 15\%$  (the radial velocity error from APOGEE is negligible). We integrated orbits for the min/max space velocities based on the uncertainties in all measured parameters, and in neither case was the orbit bound to the MW. Nonetheless, if the proper motions are systematically offset, rather than a large random error, we must await future improvements in *Gaia* proper motions to address this question. Future investigation into the chemistry of HVSs across the Galaxy may shed light on the likely origin of these stars.

## 5. Discussion

### 5.1. Implications for Chemical Tagging

Using this *k*-means clustering technique, we find that 35 of the 62 Sgr stream candidates with reliable *Gaia* proper motions are likely Sgr stream members. An additional 20 stars are accreted halo stars with indistinguishable abundance patterns from the Sgr tidal debris. Through our “weak” chemical tagging technique, we are able to recover 55 out of 62 accreted dwarf galaxy stars in the MW halo, identified by clustering on chemical space defined by six chemical elements. Because we do not have a good idea of what the chemical abundance patterns of the Sgr core stars with  $[\text{Fe}/\text{H}] < -1.2$  in the APOGEE sample look like as compared to the MW, we did not extend down our analysis. However, the Sgr stream contains stars at least as metal-poor as  $[\text{Fe}/\text{H}] = -1.6$  (Chou et al. 2007; Carlin et al. 2018), and the core contains stars as

metal-poor as  $[\text{Fe}/\text{H}] \sim -2.0$  (Mucciarelli et al. 2017; Hansen et al. 2018). Future observations of the Sgr core using APOGEE-2 that target more metal-poor K giants will help demonstrate whether this technique can separate Sgr core stars from the MW stars across a much wider metallicity range to fully discover Sgr stream stars.

Thus far, we have only analyzed 1 cluster out of 16. While a comprehensive analysis of each individual cluster is beyond the scope of this paper, we did search the other clusters for the more metal-poor “accreted halo” population, identified in the APOGEE sample by Hayes et al. (2018) and Mackereth et al. (2019). We analyzed the  $[\text{Mg}/\text{Fe}]$  versus  $[\text{Fe}/\text{H}]$  abundance space for each individual cluster and found that one cluster was apparently dominated by the “Low-Mg” population (“LMg”) identified and described by Hayes et al. (2018) (but also described in APOGEE by Hawkins et al. 2015 and Fernández-Alvar et al. 2017) as halo stars that exhibit relatively low  $[\text{Mg}/\text{Fe}]$  abundances and halo-like kinematics. This is shown in the left panel of Figure 9.

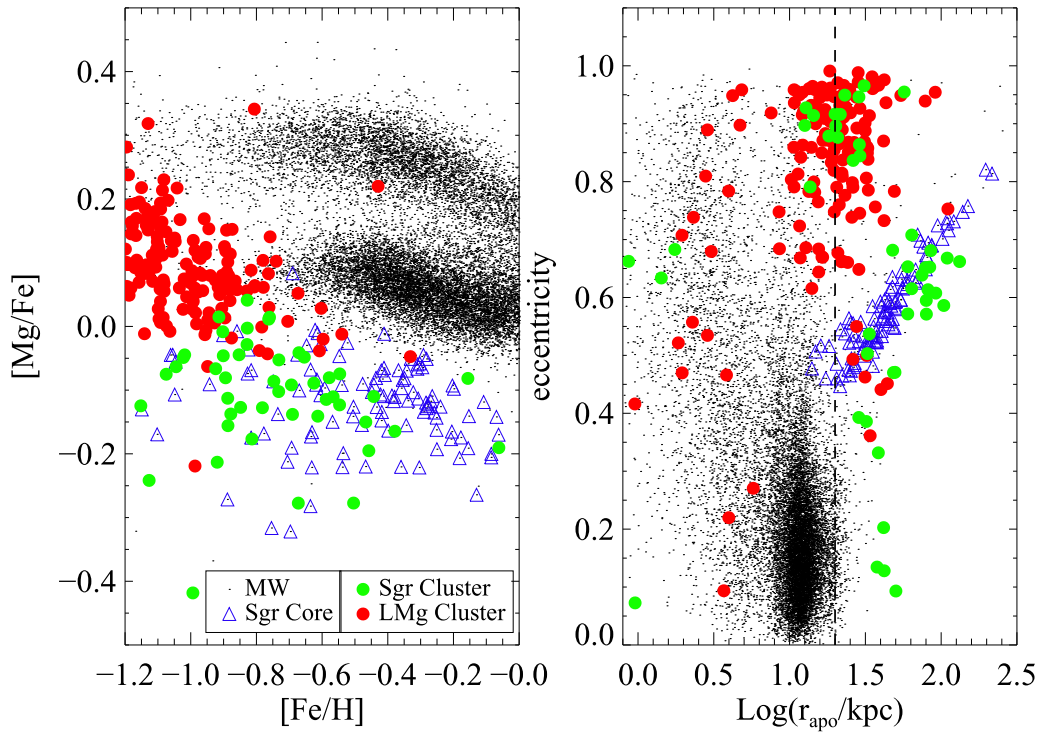
Mackereth et al. (2019) was able to recover this population by conducting a *k*-means clustering analysis on five-dimensional space defined by four chemical elements and orbital eccentricity. In the right panel of Figure 9, we show that the LMg population that naturally occupies one of our clusters does indeed exhibit high orbital eccentricity and median  $r_{\text{apo}} = 20.8$  kpc, which are both characteristics of the “accreted halo” population thought to dominate the inner halo (Belokurov et al. 2018; Deason et al. 2018; Helmi et al. 2018). We show in Figure 10 that these stars are not in a particular region of the sky, nor are they obviously spatially clustered, as described by Deason et al. (2018). The recovery of this population suggests that the chemical tagging technique applied in this work is useful for identifying stars formed in relatively massive dwarf galaxies. The APOGEE IDs for the stars that fall into our LMg cluster are presented in Table 2.

### 5.2. Chemical Evolution of Sgr and the Assembly History of the MW

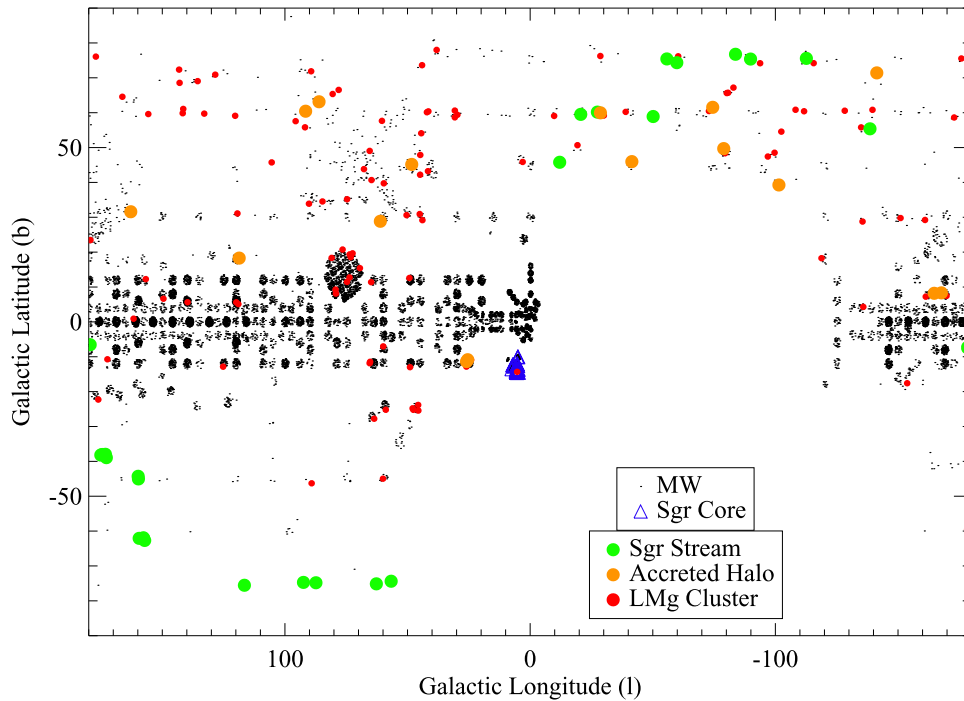
Because we find no stream stars with  $[\text{Fe}/\text{H}] > -0.2$ , which are found in abundance in the APOGEE Sgr core sample, we offer a constraint on how chemically evolved Sgr was during its most recent pericenter passage. The SFH of the Sgr core from Siegel et al. (2007) suggests that these stars formed  $\sim 2$  Gyr ago, which might imply that we may expect to see some of these stars stripped as we are looking at debris that Law & Majewski (2010) declared was stripped as early as 1 Gyr ago. However, if the starburst that formed this population with  $[\text{Fe}/\text{H}] > -0.2$  only occurred in the very central regions some 2 Gyr ago, then we might not see any in the stream structure.

Alternatively, Tepper-García & Bland-Hawthorn (2018) argue that Sgr came into the MW gas-rich and was fully stripped of gas during its last disk passage approximately 1 Gyr ago. In this scenario, the most metal-rich stars were formed during a last burst of SF coincident with this passage. The lack of stars with  $[\text{Fe}/\text{H}] > -0.2$  in our stream sample is also consistent with this scenario. This would require a slight tweak to the Siegel et al. (2007) SFH of Sgr, such that stars with  $[\text{Fe}/\text{H}] > -0.2$  formed 1 Gyr ago rather than 2–3 Gyr ago.

The “accreted halo” stars, which exhibit the same abundance patterns as Sgr, but are on different orbits, are likely the metal-rich end of the accreted halo population described in



**Figure 9.** Left:  $[\text{Mg}/\text{Fe}]$  vs.  $[\text{Fe}/\text{H}]$  abundance plane for the MW stars (black dots), Sgr core sample (purple open triangles), Sgr stream cluster (green filled-circles), and LMg cluster (red filled-circles). Right: Orbital eccentricity vs.  $\text{Log}(r_{\text{apo}})$  for the same samples. The dashed line signifies  $r_{\text{apo}} = 20$  kpc.



**Figure 10.** Galactic map of the identified stellar populations, along with the MW sample on which we performed the cluster algorithm. The points are colored according to the legend.

Hayes et al. (2018), Deason et al. (2018), Helmi et al. (2018), and Mackereth et al. (2019). We have shown that one of our clusters contains the more metal-poor debris. We find that these “accreted halo” stars, identified through chemistry alone, do indeed exhibit apocenters of  $\sim 20$  kpc along with orbital eccentricity  $> 0.8$ . The chemical abundance patterns of the “accreted halo” stars appear to be indistinguishable from the

Sgr core and stream stars at  $-1.0 < [\text{Fe}/\text{H}] < -0.6$ . There are potential differences in the stars with  $[\text{Fe}/\text{H}] < -1.0$ , but we only have a small handful of Sgr stars to compare to.

The indistinguishable chemical abundance patterns suggest that Sgr and the progenitor of the “accreted halo” debris exhibited very similar star formation histories. The most metal-rich of the “accreted halo” stars in our sample have

**Table 2**  
LMg Cluster

APOGEE ID	$r_{\text{apo}}$ (kpc)	Ecc
2M00095633+6739200	11.59	0.95
2M00110806+8047243	27.65	0.55
2M00175311+6816598	10.71	0.79
2M01051470+4958078	10.66	0.86
2M02354988-0910320	16.98	0.91

(This table is available in its entirety in machine-readable form.)

$[\text{Fe}/\text{H}] = -0.6$ . If we assume that the star formation histories must have been the same, we can use the known SFH of Sgr to time when the progenitor of the “accreted halo” debris merged with the MW. Based on the SFH of Sgr, as presented by Siegel et al. (2007), Sgr did not form stars as metal-rich as  $[\text{Fe}/\text{H}] \sim -0.6$  until 8–9 Gyr ago. This is in good agreement with simulations of the “accreted halo” debris, which find that the progenitor was likely an LMC-sized galaxy that merged with the MW some 8–11 Gyr ago (e.g., Belokurov et al. 2018; Mackereth et al. 2019).

However, from our data, we cannot determine whether this accreted halo population is from one major accretion event or a handful of smaller accretion events. While we do find stars as metal-rich as  $[\text{Fe}/\text{H}] \sim -0.6$ , the mean  $[\text{Fe}/\text{H}]$  of the accreted halo population is much more metal-poor than this. If we assume that the accreted halo stars are indeed the metal-rich extension of the LMg population, then we find a mean metallicity of this population of  $\langle [\text{Fe}/\text{H}] \rangle \sim -1.0$ . This is an upper limit, as the true mean metallicity is likely lower than this because the LMg population extends to metallicities more metal-poor than our  $[\text{Fe}/\text{H}] = -1.2$  cutoff (Hayes et al. 2018). Based on the stellar mass–stellar metallicity relationship described by Kirby et al. (2013), a mean metallicity of  $\lesssim -1.0$  suggests the progenitor galaxy had a stellar mass of  $\sim 10^7 M_{\odot}$ , similar to that of Fornax. While good arguments have been made as to why this progenitor galaxy need be massive (e.g., Belokurov et al. 2018; Mackereth et al. 2019), such a galaxy would fall considerably off the stellar mass–stellar metallicity relationship, being too metal-poor for its mass.

## 6. Conclusions

We have demonstrated a “weak chemical tagging” method that efficiently identifies Sgr stream stars from chemical abundances alone. Specifically, we performed a  $k$ -means clustering analysis on six APOGEE chemical abundance ratios in which Sgr core stars are distinct from MW stellar populations (Hasselquist et al. 2017). As a test, our technique recovers 19 of 21 kinematically selected Sgr stream members. Of the 62 stars in our chemically selected APOGEE sample that have reliable *Gaia* proper motions, 35 (56%) are confirmed to be Sgr stream stars, and 20 more belong to an “accreted halo” population that apparently spawned from a progenitor with a very similar SFH to Sgr in its first 2–5 Gyr of existence. Therefore, our detection reliability of “Sgr-like” galaxy stars is  $\sim 89\%$ .

To explore the nature of the chemically selected candidates, we integrated orbits using *Gaia* proper motions and APOGEE radial velocities and distance estimates. A total of 35 candidates have positions and 3D velocities consistent with the Law & Majewski (2010) model of Sgr disruption. The majority of these correspond to Sgr tidal debris from the last two pericenter

passages of Sgr’s orbit. An additional 20 stars are on highly eccentric orbits (median  $e = 0.9$ ) with median  $r_{\text{apo}} = 20.2$  kpc. We refer to this sample as the “accreted halo” stars, because their kinematics appear similar to those of the stars recently claimed to originate in an ancient major merger (e.g., Schuster et al. 2012; Belokurov et al. 2018; Deason et al. 2018). These stars are mostly not located near the Sgr debris plane, but because they were selected to have chemistry similar to Sgr, it is possible that they are related to the oldest debris from the Sgr disruption. It may also be that they were contributed by a different merger event whose progenitor was similar to Sgr.

Also included in the Sgr chemical cluster are four apparent bulge stars. These stars were included in the Sgr cluster due to their exceptionally low  $[\text{O}/\text{Fe}]$  abundance, but exhibited slightly enhanced  $[\text{Ni}/\text{Fe}]$  abundance, uncharacteristic of Sgr stars. However, they are outliers in chemical space, as compared to the bulk of the bulge stars in APOGEE, and suggest that additional chemical information is required to separate chemically similar (but not chemically identical) groups. We additionally find one star whose 3D position in the Galaxy is consistent with being part of the Sgr debris at the most recent apocenter, but whose total Galactocentric velocity is  $\sim 600 \text{ km s}^{-1}$ . The orbit of this star does not approach within 50 kpc of the Galactic center, making it inconsistent with most scenarios for producing HVSs. Because it has abundances consistent with Sgr, it is possible that this star was somehow ejected from the Sgr stream, but further refinement of the proper motion may be necessary to understand its origin.

Future APOGEE-2 observations of the Sgr core that extend to lower metallicities will allow us to test this “tuned”  $k$ -means clustering algorithm to see if we can recover more metal-poor Sgr stars that we know to be in the streams. This will also allow for a more in-depth analysis of the similarities between the “accreted halo” population of stars thought to be from a massive progenitor from 8+ Gyr ago and the Sgr stars.









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 website is [www.sdss.org](http://www.sdss.org).

SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU)/University of Tokyo, 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 Observatory of China, New Mexico State University, New York University, University of Notre Dame, Observatório Nacional/MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of

Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

S.H. is supported by an NSF Astronomy and Astrophysics Postdoctoral Fellowship under award AST-1801940. A.G.H. and O.Z. acknowledge support provided by the Spanish Ministry of Economy and Competitiveness (MINECO) under grant AYA-2017-88254-P. We thank Adrian Price-Whelan for assistance with running and modifying the `gala` software.

### ORCID iDs

Jeffrey L. Carlin  <https://orcid.org/0000-0002-3936-9628>  
 Jon A. Holtzman  <https://orcid.org/0000-0002-9771-9622>  
 Matthew Shetrone  <https://orcid.org/0000-0003-0509-2656>  
 Christian R. Hayes  <https://orcid.org/0000-0003-2969-2445>  
 Rachael L. Beaton  <https://orcid.org/0000-0002-1691-8217>  
 Steven R. Majewski  <https://orcid.org/0000-0003-2025-3147>  
 Kaike Pan  <https://orcid.org/0000-0002-2835-2556>  
 David L. Nidever  <https://orcid.org/0000-0002-1793-3689>

### References

- Abadi, M. G., Navarro, J. F., & Steinmetz, M. 2009, *ApJL*, 691, L63  
 Abolfathi, B., Aguado, D. S., Aguilar, G., et al. 2017, arXiv:1707.09322  
 Abolfathi, B., Aguado, D. S., Aguilar, G., et al. 2018, *ApJS*, 235, 42  
 Allende Prieto, C., Fernández-Alvar, E., Schlesinger, K. J., et al. 2014, *A&A*, 568, A7  
 Bailer-Jones, C. A. L., Rybizki, J., Foesneau, M., Mantelet, G., & Andrae, R. 2018, *AJ*, 156, 58  
 Beers, T. C., Carollo, D., Ivezić, Ž., et al. 2012, *ApJ*, 746, 34  
 Bell, E. F., Zucker, D. B., Belokurov, V., et al. 2008, *ApJ*, 680, 295  
 Belokurov, V., Erkal, D., Evans, N. W., Koposov, S. E., & Deason, A. J. 2018, *MNRAS*, 478, 611  
 Belokurov, V., Evans, N. W., Irwin, M. J., et al. 2007, *ApJ*, 658, 337  
 Belokurov, V., Koposov, S. E., Evans, N. W., et al. 2014, *MNRAS*, 437, 116  
 Belokurov, V., Zucker, D. B., Evans, N. W., et al. 2006, *ApJL*, 642, L137  
 Blaauw, A. 1961, *BAN*, 15, 265  
 Blanton, M. R., Bershady, M. A., Abolfathi, B., et al. 2017, *AJ*, 154, 28  
 Bovy, J. 2015, *ApJS*, 216, 29  
 Brook, C. B., Kawata, D., Gibson, B. K., & Flynn, C. 2003, *ApJL*, 585, L125  
 Carlin, J. L., Sheffield, A. A., Cunha, K., & Smith, V. V. 2018, *ApJL*, 859, L10  
 Carollo, D., Beers, T. C., Lee, Y. S., et al. 2007, *Natur*, 450, 1020  
 Chiba, M., & Beers, T. C. 2000, *AJ*, 119, 2843  
 Chou, M.-Y., Majewski, S. R., Cunha, K., et al. 2007, *ApJ*, 670, 346  
 Deason, A. J., Belokurov, V., Koposov, S. E., & Lancaster, L. 2018, *ApJL*, 862, L1  
 Eisenstein, D. J., Weinberg, D. H., Agol, E., et al. 2011, *AJ*, 142, 72  
 Fardal, M. A., van der Marel, R. P., Law, D. R., et al. 2018, arXiv:1804.04995  
 Fernández-Alvar, E., Allende Prieto, C., Schlesinger, K. J., et al. 2015, *A&A*, 577, A81  
 Fernández-Alvar, E., Carigi, L., Allende Prieto, C., et al. 2017, *MNRAS*, 465, 1586  
 Fernández-Trincado, J. G., Robin, A. C., Moreno, E., et al. 2016, *ApJ*, 833, 132  
 Fernández-Trincado, J. G., Zamora, O., García-Hernández, D. A., et al. 2017, *ApJL*, 846, L2  
 Font, A. S., McCarthy, I. G., Crain, R. A., et al. 2011, *MNRAS*, 416, 2802  
 Freeman, K., & Bland-Hawthorn, J. 2002, *ARA&A*, 40, 487  
 Fritz, T. K., Battaglia, G., Pawłowski, M. S., et al. 2018, arXiv:1805.00908  
 Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, *A&A*, 616, A1  
 Gaia Collaboration, Helmi, A., van Leeuwen, F., et al. 2018, arXiv:1804.09381  
 García Pérez, A. E., Allende Prieto, C., Holtzman, J. A., et al. 2016, *AJ*, 151, 144  
 Grillmair, C. C. J., & Carlin, J. L. J. 2016, Stellar Streams and Clouds in the Galactic Halo, arXiv:1603.08936  
 Gunn, J. E., Siegmund, W. A., Mannery, E. J., et al. 2006, *AJ*, 131, 2332  
 Gvaramadze, V. V., Gualandris, A., & Portegies Zwart, S. 2009, *MNRAS*, 396, 570  
 Hansen, C. J., El-Souri, M., Monaco, L., et al. 2018, *ApJ*, 855, 83  
 Hasselquist, S., Shetrone, M., Smith, V., et al. 2017, *ApJ*, 845, 162  
 Hawkins, K., Jofré, P., Masseron, T., & Gilmore, G. 2015, *MNRAS*, 453, 758  
 Hayden, M. R., Holtzman, J. A., Bovy, J., et al. 2014, *AJ*, 147, 116  
 Hayes, C. R., Majewski, S. R., Shetrone, M., et al. 2018, *ApJ*, 852, 49  
 Helmi, A., Babusiaux, C., Koppelman, H. H., et al. 2018, *Natur*, 563, 85  
 Hernitschek, N., Sesar, B., Rix, H.-W., et al. 2017, *ApJ*, 850, 96  
 Hogg, D. W., Casey, A. R., Ness, M., et al. 2016, *ApJ*, 833, 262  
 Holtzman, J. A., Hasselquist, S., Shetrone, M., et al. 2018, *AJ*, 156, 125  
 Holtzman, J. A., Shetrone, M., Johnson, J. A., et al. 2018, *AJ*, 150, 148  
 Ibata, R. A., Gilmore, G., & Irwin, M. J. 1994, *Natur*, 370, 194  
 Ibata, R. A., Gilmore, G., & Irwin, M. J. 1995, *MNRAS*, 277, 781  
 Johnston, K. V., Hernquist, L., & Bolte, M. 1996, *ApJ*, 465, 278  
 Jönsson, H., Allende Prieto, C., Holtzman, J. A., et al. 2018, *AJ*, 156, 126  
 Kirby, E. N., Cohen, J. G., Guhathakurta, P., et al. 2013, *ApJ*, 779, 102  
 Klypin, A., Kravtsov, A. V., Valenzuela, O., & Prada, F. 1999, *ApJ*, 522, 82  
 Koposov, S. E., Irwin, M., Belokurov, V., et al. 2014, *MNRAS*, 442, L85  
 Koppelman, H., Helmi, A., & Veljanoski, J. 2018, *ApJL*, 860, L11  
 Kruijssen, J. M. D., Pfeffer, J. L., Reina-Campos, M., Crain, R. A., & Bastian, N. 2018, *MNRAS*, arXiv:1806.05680  
 Law, D. R., & Majewski, S. R. 2010, *ApJ*, 714, 229  
 Leonard, P. J. T., & Duncan, M. J. 1990, *AJ*, 99, 608  
 Li, J., Smith, M. C., Zhong, J., et al. 2016, *ApJ*, 823, 59  
 Mackereth, J. T., Schiavon, R. P., Pfeffer, J., et al. 2019, *MNRAS*, 482, 3426  
 Majewski, S. R., Hasselquist, S., Łokas, E. L., et al. 2013, *ApJL*, 777, L13  
 Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., et al. 2017, *AJ*, 154, 94  
 Majewski, S. R., Skrutskie, M. F., Weinberg, M. D., & Ostheimer, J. C. 2003, *ApJ*, 599, 1082  
 McWilliam, A., Wallerstein, G., & Mottini, M. 2013, *ApJ*, 778, 149  
 Mitschang, A. W., De Silva, G., Zucker, D. B., et al. 2014, *MNRAS*, 438, 2753  
 Mucciarelli, A., Bellazzini, M., Ibata, R., et al. 2017, arXiv:1705.03251  
 Myeong, G. C., Evans, N. W., Belokurov, V., Sanders, J. L., & Koposov, S. E. 2018, *ApJL*, 863, L28  
 Ness, M., Hogg, D. W., Rix, H.-W., Ho, A. Y. Q., & Zasowski, G. 2015, *ApJ*, 808, 16  
 Ness, M., Rix, H., Hogg, D. W., et al. 2017, arXiv:1701.07829  
 Nidever, D. L., Holtzman, J. A., Allende Prieto, C., et al. 2015, *AJ*, 150, 173  
 Piffl, T., Williams, M., & Steinmetz, M. 2011, *A&A*, 535, A70  
 Price-Whelan, A., Sipocz, B., Major, S., & Oh, S. 2017, *adm/gala*: v0.2.1, Zenodo, doi:10.5281/zenodo.833339  
 Price-Whelan, A. M. 2017, *Gala*: A Python package for galactic dynamics, *Journal of Open Source Software*, 2, 388  
 Schiavon, R. P., Zamora, O., Carrera, R., et al. 2017, *MNRAS*, 465, 501  
 Schlauffman, K. C., Rockosi, C. M., Lee, Y. S., et al. 2012, *ApJ*, 749, 77  
 Schlauffman, K. C., Rockosi, C. M., Lee, Y. S., Beers, T. C., & Allende Prieto, C. 2011, *ApJ*, 734, 49  
 Schönrich, R., Binney, J., & Dehnen, W. 2010, *MNRAS*, 403, 1829  
 Schuster, W. J., Moreno, E., Nissen, P. E., & Pichardo, B. 2012, *A&A*, 538, A21  
 Sheffield, A. A., Majewski, S. R., Johnston, K. V., et al. 2012, *ApJ*, 761, 161  
 Shetrone, M., Bizyaev, D., Lawler, J. E., et al. 2015, *ApJS*, 221, 24  
 Siegel, M. H., Dotter, A., Majewski, S. R., et al. 2007, *ApJL*, 667, L57  
 Taylor, M. B. 2005, in ASP Conf. Ser. 347, *Astronomical Data Analysis Software and Systems XIV*, ed. P. Shopbell, M. Britton, & R. Ebert (San Francisco, CA: ASP), 29  
 Tepper-García, T., & Bland-Hawthorn, J. 2018, *MNRAS*, 478, 5263  
 Watkins, L. L., van der Marel, R. P., Sohn, S. T., & Evans, N. W. 2018, arXiv:1804.11348  
 Wilson, J. C., Hearty, F. R., Skrutskie, M. F., et al. 2019, *PASP*, in press  
 Xue, X.-X., Rix, H.-W., Ma, Z., et al. 2015, *ApJ*, 809, 144  
 Zamora, O., García-Hernández, D. A., Allende Prieto, C., et al. 2015, *AJ*, 149, 181  
 Zasowski, G., Cohen, R. E., Chojnowski, S. D., et al. 2017, *AJ*, 154, 198  
 Zasowski, G., Johnson, J. A., Frinchaboy, P. M., et al. 2013, *AJ*, 146, 81  
 Zolotov, A., Willman, B., Brooks, A. M., et al. 2009, *ApJ*, 702, 1058