The Astrophysical Journal Supplement Series, 259:60 (31pp), 2022 April https://doi.org/10.3847/1538-4365/ac5323

SEGUE-2: Old Milky Way Stars Near and Far

Constance M. Rockosi[1], Young Sun Lee[2], Heather L. Morrison[3], Brian Yanny[4], Jennifer A. Johnson[5,6], Sara Lucatello[7], Jennifer Sobeck[8], Timothy C. Beers[9], Carlos Allende Prieto[10,11], Deokkeun An[12], Dmitry Bizyaev[13,14], Michael R. Blanton[15], Luca Casagrande[16], Daniel J. Eisenstein[17], Andrew Gould[5], James E. Gunn[18], Paul Harding[3], Inese I. Ivans[19], H. R. Jacobson[20], William Janesh[3], Gillian R. Knapp[18], Juna A. Kollmeier[21], Sébastien Lépine[22], Martín López-Corredoira[10,11], Zhibo Ma[23], Heidi J. Newberg[24], Kaike Pan[13], Jakub Prchlik[3], Conor Sayers[8], Katharine J. Schlesinger[16], Jennifer Simmerer[19], and David H. Weinberg[5,6]

[1] University of California Observatories and Department of Astronomy and Astrophysics, University of California Santa Cruz, Santa Cruz, CA 95064, USA
[2] Department of Astronomy & Space Science, Chungnam National University, Daejeon 34134, Republic of Korea
[3] Department of Astronomy, Case Western Reserve University, Cleveland, OH 44106, USA
[4] Fermi National Accelerator Laboratory, Batavia, IL 60510, USA
[5] Department of Astronomy, The Ohio State University, Columbus, OH 43210, USA
[6] Center for Cosmology and AstroParticle Physics, The Ohio State University, Columbus, OH 43210, USA
[7] Osservatorio Astronomico di Padova—INAF, Vicolo dell'Osservatorio 5, I-35122, Padova, Italy
[8] Department of Astronomy, University of Washington, Seattle, WA 98195, USA
[9] Department of Physics and JINA Center for the Evolution of the Elements, University of Notre Dame, South Bend, IN 46556, USA
[10] Instituto de Astrofísica de Canarias, E-38205 La Laguna, Tenerife, Spain
[11] Departamento de Astrofísica, Universidad de La Laguna (ULL), E-38206 La Laguna, Tenerife, Spain
[12] Department of Science Education, Ewha Womans University, 52 Ewhayeodae-gil, Seodaemun-gu, Seoul 03760, Republic of Korea
[13] Apache Point Observatory and New Mexico State University, P.O. Box 59, Sunspot, NM 88349-0059, USA
[14] Sternberg Astronomical Institute, Moscow State University, Moscow, Russia
[15] Center for Cosmology and Particle Physics, Department of Physics, New York University, 726 Broadway, Room 1005, New York, NY 10003, USA
[16] Research School of Astronomy & Astrophysics, Australian National University, Cotter Road, Weston Creek, ACT 2611, Australia
[17] Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA
[18] Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA
[19] Department of Physics and Astronomy, University of Utah, 115 S. 1400 E., Salt Lake City, UT 84112, USA
[20] Massachusetts Institute of Technology, Kavli Institute for Astrophysics and Space Research, Cambridge, MA 02139, USA
[21] Observatories of the Carnegie Institution for Science, 813 Santa Barbara Street, Pasadena, CA 91101, USA
[22] Department of Physics & Astronomy, Georgia State University, Atlanta, GA, USA
[23] TBanic Data, 39 Jianwai Street, Bldg 15-1106, Beijing, People's Republic of China
[24] Department of Physics, Applied Physics and Astronomy, Rensselaer Polytechnic Institute, Troy, NY 12180, USA
Received 2021 July 20; revised 2021 December 10; accepted 2021 December 22; published 2022 April 7

Abstract

The Sloan Extension for Galactic Understanding and Exploration 2 (SEGUE-2) obtained 128,288 low-resolution spectra ($R \sim 1800$) of 118,958 unique stars in the first year of the Sloan Digital Sky Survey III (2008–2009). SEGUE-2 targeted prioritized distant halo tracers (blue horizontal-branch stars, K giants, and M giants) and metal-poor or kinematically hot populations. The main goal of SEGUE-2 was to target stars in the distant halo and measure their kinematics and chemical abundances to learn about the formation and evolution of the Milky Way. We present the SEGUE-2 field placement and target selection strategies. We discuss the success rate of the targeting based on the SEGUE-2 spectra and other spectroscopic and astrometric surveys. We describe the final SEGUE-2/SDSS-III improvements to the stellar parameter determinations based on the SEGUE Stellar Parameter Pipeline. We report a $(g-i)$ color−effective temperature relation calibrated to the IRFM. We evaluate the accuracy and uncertainties associated with these stellar parameters by comparing with fundamental parameters, a sample of high-resolution spectra of SEGUE stars analyzed homogeneously, stars in well-studied clusters, and stars observed in common by the APOGEE survey. The final SEGUE spectra, calibration data, and derived parameters described here were released in SDSS-III Data Release 9 and continue to be included in all subsequent SDSS Data Releases. Because of its faint limiting magnitude and emphasis on the distant halo, the public SEGUE-2 data remain an important resource for the spectroscopy of stars in the Milky Way.

Unified Astronomy Thesaurus concepts: Milky Way Galaxy (1054); Radial velocity (1332); Sky surveys (1464); Spectroscopy (1558); Chemical abundances (224)

Supporting material: machine-readable tables

1. Introduction

Spectroscopic observations of stars in old stellar populations are of fundamental importance for understanding the formation and evolution of the Milky Way and the cosmological history of galaxy formation. Recent years have seen several massive ($>10^5$ stars) spectroscopic surveys that have produced age, composition, and kinematic data charting the Galaxy's formation history back to early epochs: the buildup of mass and the mass density profile, the accretion, dynamical, and chemical history of the Galaxy, and the formation of the stellar populations.

This paper describes the Sloan Extension for Galactic Understanding and Exploration 2 (SEGUE-2), the spectroscopic survey of Galactic stars carried out by the Sloan Digital Sky Survey (SDSS) during the first year of SDSS-III. SEGUE-2 shifted emphasis from that of SEGUE-1 (Newberg et al. 2002; Yanny et al. 2009) to target stars in the outer Galactic halo.

Section 2 discusses other recent Galactic surveys and their major focus to put the motivating ideas behind SEGUE-2 in context. Section 3 begins the discussion of SEGUE-2 by summarizing the observing plan and choice of spectroscopic fields. The selection of the target stars, the criteria used, and the size and global properties of the final sample are summarized in Section 4. In addition to the new observations, SEGUE-2 updated the SEGUE Stellar Parameter Pipeline (SSPP) to improve the estimates of the stellar parameters [Fe/H],[25] $T_{\rm eff}$, and $\log g$[26] for all stellar spectra obtained with the original SDSS spectrographs.[27] The updates to the SSPP, the new temperature scale, and the calibration of this scale are described in Section 5. The improved SSPP was run for all SDSS/SEGUE stellar spectra (SDSS, SEGUE-1, and SEGUE-2) and the results released and made available as part of SDSS Data Release 9 (DR9, Ahn et al. 2012). The accuracy and precision of stellar parameters from the new SSPP are described in Section 6. Section 7 describes sources of incompleteness and bias in the spectroscopic samples and how they can be corrected.

The DR9 SSPP parameters have remained available, unchanged, in subsequent SDSS-III and SDSS-IV data releases. The radial velocity (RV) measurement pipeline for SEGUE-2 is unchanged from SEGUE-1, and the RV uncertainties as tabulated in Yanny et al. (2009) also describe the SEGUE-2 RV errors; see also Section 6 of this paper. The reduced stellar spectra, as well as all the raw data and calibration information, are also available as part of the SDSS data releases for use in independent analysis (e.g., Ludwig et al. 2008; Caffau et al. 2011). These data remain a resource for stellar spectroscopy as new imaging surveys and the next generation of spectroscopic surveys become available.

2. SEGUE-2 in Context

2.1. The Galactic Halo

Stars in the solar neighborhood have long been known to belong to more than one population. The classic study by Eggen et al. (1962) established two populations: low-metallicity, high-velocity dispersion stars identified with the Galactic halo and solar-metallicity, low-velocity dispersion stars associated with the disk, leading to a model in which the initial gravitational collapse produced low-metallicity old stars and the gas-rich rotating disk with continuing star formation. In this picture, the formation process created a centrally concentrated density distribution, high-velocity dispersion, and generally low metallicity for the halo stars. It is now thought that halo formation was more complex, with the accretion of smaller (hence metal-poor) galaxies playing an important part (Searle & Zinn 1978; Ibata et al. 1994). The expected split between halo and disk in metallicity and kinematics led studies of the halo relying on solar neighborhood stars to select them to have low metallicity (e.g., Bond 1980; Beers et al. 1985, 1992; Norris et al. 1999; Chiba & Beers 2000) or hot kinematics (e.g., Ryan & Norris 1991; Carney et al. 1996).

The advent of sensitive detectors allowing much fainter stars to be observed, of accurate photometry for reliable color selection of stellar samples, of deep wide-field photometric surveys for the identification of very large, homogeneous samples over large spatial scales, and of multiobject spectral surveys have carried these investigations far out into the Galactic halo and demonstrated the presence of substantial substructure. Photometry in the SDSS Early Data Release (Stoughton et al. 2002) proved sufficiently accurate to separate halo main-sequence turnoff stars from those in the disk and thick disk by their bluer colors, making it possible to map the spatial substructure in the halo on large scales (Yanny et al. 2003). This motivated SEGUE (Yanny et al. 2009), carried out as part of SDSS-II, and has been extended by several other large surveys of stellar properties and kinematics. Dynamical time arguments and numerical simulations show that the coherence of accreted stars will be longer-lived kinematically than spatially and that accretion time, orbit, and mass ratios all affect where and when the accreted stars are deposited in the halo and how long the structures remain coherent (Helmi 2008; Johnston et al. 2008; Zolotov et al. 2009; Tissera et al. 2014; Karademir et al. 2019). Thus, the amount of substructure in the stellar halo varies with distance from the Galactic center (GC) and is sensitive to the accretion history. These simulations also predict that the metallicity distribution function (MDF) provides a crucial diagnostic of the accretion history of the Galactic halo. Johnston et al. (2008) describe the relationship between mass, accretion time, and chemical enrichment of stellar halo progenitors and find that mergers contribute more metal-rich stars to the halo. The Aquarius (Cooper et al. 2010) and Auriga (Monachesi et al. 2019) simulations show that a spiral galaxy's halo is a relic of its accretion history, with fewer accretion events leading to steeper metallicity gradients. In the FIRE simulations, El-Badry et al. (2018) find that the fraction of stars with formation redshifts greater than 5 is substantially greater in the outer halo and that the best way to find them is to search at high Galactic latitudes at distances greater than 10 kpc from the Sun. All of these motivate a spectroscopic survey of stars in the distant halo of the Milky Way.

2.2. Spectroscopic Surveys of the Galactic Halo

In this section, we list the main parameters of other large Milky Way halo surveys and summarize the complementary nature of SEGUE-2. Phase 1 of the Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST) observed stars with spectral resolving power[28] $R \sim 1500$ with 4000 fibers per field (Zhao et al. 2012) and included high-latitude ($|b| > 30°$) fields targeting the Galactic halo to $r < 16.8$ and fainter stars in the northern and southern Galactic caps. Data Release 5, the end of LAMOST-1, contains spectra of 8.2 million stars (Luo et al. 2015). LAMOST-2 has added medium-

[25] We use the standard notation: $[X/Q] = \log_{10}N(X)_* - \log_{10}N(X)_\odot - \log_{10}N(Q)_* + \log_{10}N(Q)_\odot$, where N is the number density of atoms.

[26] The surface gravity of a star is reported as $\log_{10}g$, where $g = GM_*/R_*^2$ in cgs units.

[27] Hereafter, stellar spectra acquired with the original SDSS spectrographs are referred to as SDSS/SEGUE spectra when the details of their specific selection in SDSS, SEGUE-1, and SEGUE-2 are not important. In fall 2009, after the end of the SEGUE-2 observations, the spectrographs were upgraded for the BOSS survey (Smee et al. 2013).

[28] Resolving power R is defined as $\lambda/\delta\,\lambda$.

resolution ($R \sim 7500$) spectroscopy, and its Data Release 6 contains more than 9 million low-resolution and 1.3 million medium-resolution spectra (Luo et al. 2018).

The Radial Velocity Experiment (RAVE) used the 6DF spectrograph to obtain $R \sim 7500$ spectra of 130 stars per field with $9 < I < 12$. Designed as a precursor to Gaia spectroscopy, RAVE observed the same wavelength region as Gaia, around the 8550 Å Ca II triplet (Steinmetz et al. 2006). The complete sample, released in DR5, gives spectra, stellar parameters, RVs, and abundances for about 450,000 stars. The Gaia-ESO (Gilmore et al. 2012) and GALactic Archaeology with Hermes (GALAH; De Silva et al. 2015; Martell et al. 2017) survey strategies began a new era of surveys designed around the expectation of exquisitely accurate parallax and proper-motion data for nearby stars from Gaia. Their samples of halo stars are within a few kiloparsecs of the Sun. The Gaia-ESO stars were observed with the GIRAFFE and UVES spectrographs on VLT-Kueyen in two optical bands with $R \sim 18{,}000$ and 28,000 and, for the brighter stars, at $R \sim 45{,}000$. The GALAH survey uses no color selection, while the Gaia-ESO survey targets turnoff stars with 80% of the fibers and brighter red stars in the red clump with the remaining 20% (Stonkutė et al. 2016). The H3 Survey (Conroy et al. 2019a) also uses Gaia parallaxes to eliminate nearby stars in order to focus on the distant halo. The H3 observations use the medium-resolution ($R \sim 23{,}000$) Hectochelle spectrograph to observe in a single echelle order spanning 5150–5300 Å. The main H3 sample is selected in the magnitude range $15 < r < 18$ from the Pan-STARRS PS1 photometry with a Gaia parallax cut. There is also an H3 "high value" sample of luminous halo tracers. The H3 survey plans to measure stellar parameters and distances for ∼200,000 stars.

SDSS-III (Eisenstein et al. 2011) and SDSS-IV (Blanton et al. 2017) included stellar observations obtained as calibration targets (Fernández-Alvar et al. 2016) and in ancillary programs (Dawson et al. 2013) using the upgraded BOSS spectrographs (Smee et al. 2013). SDSS-III featured the large-scale APOGEE-1 and APOGEE-2 surveys (Zasowski et al. 2013, 2017; Majewski et al. 2017). The APOGEE spectrographs (Wilson et al. 2019) observe 300 spectra per field at $R \sim 22{,}500$ in the H band and target intrinsically red stars, selected in the JHK_s bands, in the Galactic disk and bulge. The APOGEE observations at high latitudes target globular clusters and halo field stars.

These surveys generally sample nearby stars. Queiroz et al. (2018) find that for $|b| > 20°$, 99% of the GALAH DR1 stars are closer than 3.3 kpc and found the distance for the 99% fractions of other surveys as follows: 6.5 kpc (RAVE DR5), 9.7 kpc (Gaia-ESO DR3), and 15.4 kpc (APOGEE-SDSS DR14). SEGUE-2 is one of the four surveys of the third incarnation of SDSS, SDSS-III (Eisenstein et al. 2011), and is a continuation and extension of SEGUE-1 (Yanny et al. 2009). SEGUE-2 is focused on stars in the outer halo. Observing in dark time and at moderate resolution, to limiting magnitudes as faint as $g = 20$ for some target classes, SEGUE-2 was able to target distant halo stars out to more than 100 kpc and is complementary to the surveys described above.

2.3. The New Galactic Halo

The major results of these surveys, including SEGUE-1 and subsequently SEGUE-2, include discoveries of stellar streams and other substructure resulting from the tidal destruction of dwarf satellite galaxies; warps, flares, and waves in the Galactic disk; the demonstration that the Milky Way has accreted populations of different metallicity; the discovery of extremely metal-poor, almost always carbon-enhanced, stars; measurements of the Milky Way's mass distribution over a large radius range; the discovery of stars with speeds faster than the escape speed; and the likely remnants of the original disk, now found in the halo. Together, these results lead to a picture of the Galactic halo as formed by accretion events throughout the Galaxy's history, look back toward the "first stars" and early nucleosynthesis, provide a picture of star formation in the extreme environment of the GC, measure the effect of dark-matter substructure on the formation of the smallest galaxies, and give insight into galaxy formation and evolution at high redshifts. This picture of the Galactic halo is the motivation and context for SEGUE-2 spectroscopy in the distant halo.

Substructure: Among the first SDSS results, Newberg et al. (2002) found density enhancements associated with tidal tails from the Sagittarius dwarf galaxy (Mateo et al. 1998; Dohm-Palmer et al. 2001; Majewski et al. 2004), while Yanny et al. (2003) discovered the Monoceros ring, which may be the remains of another accreted satellite or part of a warped and flared outer disk (e.g., Helmi et al. 2003; Ibata et al. 2003; López-Corredoira & Molgó 2014; Xu et al. 2015). These results and others show that the Sagittarius dwarf galaxy is currently being accreted by the Galaxy, producing tidal tails that have been found to stretch to at least 50 kpc from the dwarf galaxy itself (Newberg et al. 2003; Starkenburg et al. 2009; Ruhland et al. 2011; Belokurov et al. 2014; Janesh et al. 2016). The ubiquity of tidal accretion in the halo is demonstrated by the SDSS "Field of Streams" image (Belokurov et al. 2006) and by the analysis of SDSS photometry of the halo by Bell et al. (2008) and de Jong et al. (2010). The analysis of halo substructure by Naidu et al. (2020) using Gaia and the H3 survey finds that the entire distant halo was built from accreted satellites. A recent discovery combining Gaia proper motions with SDSS/SEGUE spectroscopy is the Gaia-Enceladus "sausage" (GSE; Belokurov et al. 2018a; Helmi et al. 2018) containing stars from an early accretion event that dominates the nearby halo and is responsible for the chemical dichotomy of nearby halo stars (Nissen & Schuster 2010; Jackson-Jones et al. 2014; Hayes et al. 2018). This accretion event was also likely responsible for the formation of the "Splash" (Bonaca et al. 2017; Belokurov et al. 2020), stars with metallicities more typical of the thick disk that have, however, very low angular momentum, making them bona fide halo stars now. The heating of the Galaxy's original disk during the early accretion of GSE would result in their current occupation of the halo.

Stellar population gradients: There is observational evidence, though not always consistent, that the halo's composition and kinematics—net rotation and anisotropy—change as a function of Galactocentric distance (e.g., Sirko et al. 2004a; Carollo et al. 2007, 2010, 2011; Beers et al. 2012; An et al. 2013, 2015; Deason et al. 2013; Belokurov et al. 2018b; Cunningham et al. 2019; Lee et al. 2019; An & Beers 2020, 2021). To distances of 15–20 kpc, the stellar halo has been well characterized by in situ measurements (e.g., Carollo et al. 2007; Deason et al. 2013; Belokurov et al. 2018b; Yoon et al. 2018; Fernández-Alvar et al. 2019). An important goal of SEGUE-2 is a larger sample of halo stars at greater distances, especially stars for which we can combine proper-motion measurements with RVs and metallicities from spectroscopy (e.g., Hattori et al. 2017). The H3 Survey

spectroscopy reaches the distant halo, and using H3, Conroy et al. (2019b) find a mean [Fe/H] = −1.2 for the Milky Way halo with no metallicity gradient out to 100 kpc.

Mass of the Milky Way: Distant halo tracers, including globular clusters and dwarf galaxies, are key to establishing the total mass of the Galaxy (e.g., Gaia Collaboration et al. 2018), while its radial density profile can be traced by measuring the velocity dispersion or escape velocity at different distances (Battaglia et al. 2005; Xue et al. 2008; Deason et al. 2012; Taylor et al. 2016; Williams et al. 2017; Callingham et al. 2019). These measurements are essential for studies of the Galaxy in the contexts of galaxy evolution and cosmology (Abadi et al. 2010; Boylan-Kolchin et al. 2010; Wang et al. 2012) and to the analysis of the 6D motions of high-velocity stars in the tails of the velocity distribution and the local escape speed (e.g., Smith et al. 2007; Piffl et al. 2014; Monari et al. 2018).

High-velocity stars: There is a second origin for high-velocity stars, including stars moving at greater than the escape speed. Hills (1988) predicted that three-body interactions involving the supermassive black hole at the GC could inject stars into the Galaxy at speeds exceeding the escape speed. The first such star discovered (Brown et al. 2005) is a B-type star with an RV of 853 km s^{-1}. Additional early-type hypervelocity stars (HVS) have been found since then (Brown et al. 2014; Zheng et al. 2014). Kollmeier & Gould (2007) pointed out that if the stellar population in the GC region is similar to that of the local disk, there should be a detectable population of lower-mass HVS. Koposov et al. (2020) recently found an HVS of type A whose 3D orbit points to an origin in the GC. However, RV searches for lower-mass HVS have so far found no further candidates. This suggests that the properties of binary stars, or of the initial mass function, are different in the GC environment. Proper motions from Gaia have provided a new way to search for HVS candidates (e.g., de la Fuente Marcos & de la Fuente Marcos 2019; Du et al. 2019) and, when combined with RV and even chemical abundance data, can reveal whether a star is escaping and whether it originated in the GC (Boubert et al. 2018; Hawkins & Wyse 2018; Marchetti et al. 2019). Beyond Galactocentric distances of ∼20 kpc, however, RVs remain the best way to identify candidate HVS (Kenyon et al. 2018).

Metal-poor stars: The most metal-poor stars in the halo offer the possibility of investigating the past history of nucleosynthesis. While metal-poor stars are not always the oldest stars (e.g., Tumlinson 2010; Starkenburg et al. 2017; El-Badry et al. 2018), they are chemically old and polluted by the nucleosynthesis products of fewer previous generations of stars. In particular, stars with extremely low metallicity ([Fe/H] < −3) are likely to be polluted only by the first generation of stars, offering the best information so far available on the masses, structure, and deaths of Population III stars (e.g., Beers & Christlieb 2005). However, these stars are very rare and cannot be distinguished from the far more numerous stars with [Fe/H] ∼ −2 using broadband photometry alone (e.g., Beers et al. 1985; Ivezić et al. 2008; An et al. 2013; Schlaufman & Casey 2014; Casagrande et al. 2019), requiring spectroscopic observations of large samples.[29] Further, the most metal-poor stars, with [Fe/H] < −3, require high-dispersion spectroscopy to detect the extremely weak lines and to separate interstellar and stellar Ca II absorption, particularly for warm stars. Observations of the smallish number of known very-metal-poor stars show that almost all are carbon rich (e.g., Beers & Christlieb 2005; Lee et al. 2013; Placco et al. 2014; Yoon et al. 2018), can have large overabundances of the *r*-process heavy elements (Sneden et al. 1996; Barklem et al. 2005; Lai et al. 2008; Li et al. 2015a; Hansen et al. 2018), have unexpected abundance trends for elements such as Cr and Co (McWilliam et al. 1995), and show large dispersions in the relative abundances of light elements such as O, Mg, and Ca (e.g., Cohen et al. 2013; Aoki et al. 2014; Li et al. 2015b).

The targeting and sky coverage philosophy of SEGUE-2 capitalizes on the capabilities of the Sloan Foundation Telescope and survey instrumentation: a wide field, broad wavelength coverage, efficient moderate-resolution spectrographs, the multiplex advantage of the optical fiber system, and high total throughput to cover a large angular area and reach the distant halo. The SEGUE-2 strategy of maximizing the volume sampled in the distant halo is aimed at understanding the formation of the Galaxy's stellar halo by in situ star formation and by accretion of dwarf galaxies. Details of the survey design are discussed in the next section.

3. Survey Design and Data

The SEGUE-2 survey used the 2.5 m Sloan Foundation Telescope (Gunn et al. 2006) and the two original Sloan Digital Sky Survey fiber spectrographs (Smee et al. 2013), as did SEGUE-1 and the SDSS. Each spectrograph has a red- and blue-optimized channel and acquires data for 320 of the 640 fibers in the 3° diameter spectroscopic field. The fibers are plugged into aluminum plates mounted at the telescope focal plane. A set of 640 spectra acquired simultaneously is hereafter referred to as a "plugplate" or "plate." The fibers are 180 μm in diameter, 3″ on the sky, and the spectrograph resolution is 2.8 Å at 5000 Å, or $R \sim 1800$. SEGUE-2, like SEGUE-1, is a magnitude-limited survey. Each set of 640 spectra is integrated to a signal-to-noise ratio (S/N) per angstrom of > 10, averaged across the *g*-filter bandpass, for targets $g < 19.5$. The S/N target was set by the lower limit for reliable stellar parameter determinations from the SSPP. This results in a consistent relation between S/N and magnitude over the survey for stars of similar spectral type.

SEGUE-2 observed for the first year of SDSS-III, from 2008 August through 2009 July. The survey obtained 128,288 spectra of 118,184 unique stars on 211 plates, 204 of which were unique and 7 repeated. Two of those 204 plates were specially designed for the open cluster Berkeley 29 to add to the sample of cluster stars from earlier SDSS data (see Section 5.4.2).

In SEGUE-2, the 640 fibers available for each plate were allocated to 10 different target categories, each with their own *ugriz* color, magnitude, and proper-motion limits. These are described in more detail in Section 4. The focus on obtaining spectra in the distant halo of the Galaxy was the primary consideration in deciding how to select the spectroscopic targets, as well as in planning the layout of the survey on the sky. A star near the old-population main-sequence turnoff has absolute magnitude $M_g \sim 4.0$ (An et al. 2008), so a turnoff star at apparent magnitude $g = 19.9$ is at a distance of 15 kpc. This is approximately the magnitude at which the sky and source flux are equal in the 3″ SDSS fibers. The bright limit of the

[29] Recent photometric survey efforts involving one or more narrowband filters have pushed the limiting metallicity achievable down to [Fe/H] ∼ −3.0 to −3.5 (e.g., Huang et al. 2021).

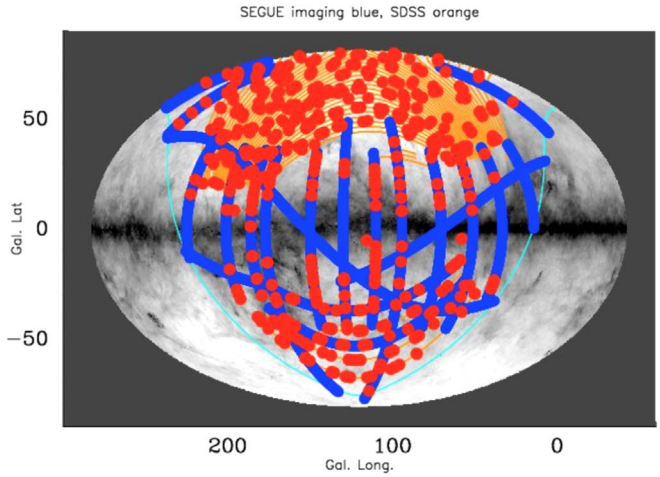

Figure 1. Layout of the SEGUE-1 and SEGUE-2 spectroscopy in Galactic coordinates in an Aitoff projection. The red circles are at the locations of the 7 deg^2 spectroscopic fields that comprise the SEGUE-1 and SEGUE-2 surveys. These are not to scale but show the SEGUE-2 field layout strategy of small overlapping groups of SEGUE-1 and SEGUE-2 fields to allow repeat observations of SEGUE-2 and SEGUE-1 targets. Each field has 640 spectroscopic targets. The orange stripes trace the original (contiguous) SDSS imaging survey area in the north Galactic sky. The blue stripes are the individual 2.5 wide stripes of SDSS imaging data in the south Galactic cap, through the Galactic plane, and along known stellar streams that were taken in the SDSS and SDSS-II surveys. All the SEGUE-1 and SEGUE-2 spectroscopic targets were selected from the SDSS imaging data. The grayscale background is the Schlegel et al. (1998) dust extinction map.

SDSS imaging data is about 14th magnitude in the g, r, and i filters and sets the bright limit for the spectroscopic targets. The faint spectroscopic limit at $g = 19.5$ was chosen as a compromise between survey depth, S/N, and exposure time per plate. Given these magnitude limits, obtaining spectra for stars more distant than 15 kpc requires targeting stars more intrinsically luminous than the main sequence. For an old stellar population such as the halo, the choices are blue horizontal-branch (BHB) stars and red giants; SEGUE-2 targets both. Because these are short phases late in stellar evolution, the number of these stars in any population is small and the projected density on the sky low. The BHB stars and red giants that can be selected most efficiently using *ugriz* imaging (see Section 4) have projected densities much less than 100 per square degree at $g < 19.5$ for Galactic latitudes that avoid the disk. Therefore, maximizing the survey area in order to target as many of these rare tracers as possible was the highest priority when planning the layout of the plates on the sky. The distribution of the SDSS imaging area on the sky, from which the targets were selected, the time limit of one year, and the need to observe near the meridian (because there is no atmospheric dispersion correction over the 3° diameter field) were also important practical constraints. The final layout of the SEGUE-2 survey is shown in Figure 1. In contrast to the bright and faint plates of SEGUE-1, each SEGUE-2 field is covered by only one plate of 640 spectra. Where possible the plates were placed such that there was a small overlap with other SEGUE-2 plates or existing SEGUE-1 data to obtain repeat observations of about 10 stars per plate as a check on the RV zero points and other measurement accuracy.

The average projected stellar density in the magnitude range $14 < g < 19.5$ is much greater than the available fiber density. The desire to maximize the area and the constraint of limited observing time over only one year all but eliminated return visits to any field. Therefore, SEGUE-2 acquired spectra for only a subsample of the stars in any field. The projected stellar density varies by factors of several over the SDSS imaging footprint, making the subsampling fraction different for every SEGUE field. This mismatch between the variable stellar density and the fixed number of fibers per field requires that investigations to measure the distribution function of a Galactic stellar population, or that otherwise require knowing the true density of a population or stellar tracer, must correct for this sampling factor as described in Section 7. It may also be necessary to account for the color, magnitude, and proper-motion selection of the tracer, also described in Section 7. The need to account for the survey sampling is a direct consequence of the priority put on a large survey area to maximize the number of distant halo tracers in limited observing time.

The imaging and astrometric data used to select targets, as well as the spectrographs and the fiber system, were the same for both SEGUE-1 and SEGUE-2, and both surveys were processed with the same software pipelines for Data Releases 8 and 9. Except for the specific selection used for the different kinds of spectroscopic targets, the data from the SEGUE-1 and SEGUE-2 surveys are otherwise very similar. Combining data from the two surveys in the analysis is no different than combining the different kinds of spectroscopic target types from one of the two surveys.

4. SEGUE-2 Spectroscopic Samples: Target Selection and Verification

The SEGUE-2 spectroscopic targets are selected from pre-DR8 SDSS photometry (hereafter SEGUE-2 photometry) because the SEGUE-2 targeting had to be complete before the final DR8 photometry was available. This target photometry is included in the public SDSS data for the specific areas of sky with SEGUE-2 spectroscopic observations and can be found in the SDSS database in the sppTargets table. The sppTargets table identifies all stars in the areas targeted by SEGUE-2 that passed the target selection criteria (as determined using the SEGUE-2 photometry) as well as the stars actually observed. In addition, the SEGUE-1 and SEGUE-2 target selection information is available for the final DR8 photometry over the entire SDSS imaging area. This full-footprint targeting information for DR8 is in the segueTargetAll table in the SDSS database. Using the target selection information and target photometry is discussed further in Section 7.

Some SEGUE-2 target categories used proper-motion information, either to select stars likely to have large space velocities or to remove local dwarf stars that contaminate the color and magnitude selection of more distant objects. As Gaia data were not yet available, the proper motions used to select SEGUE-2 targets were from the USNO-B proper-motion catalog (Monet et al. 2003), recalibrated as described in Munn et al. (2004), and available with DR7 and all subsequent data releases. All target categories that used proper-motion information selected only those objects with reliable proper-motion estimates, as in Kilic et al. (2006) (see footnote to Table 1). The errors of the recalibrated USNO-B catalog are ∼4 mas per coordinate (Munn et al. 2004). At a distance of 2 kpc, this is a transverse velocity error of ∼38 km s^{-1} per coordinate. The SEGUE-2 targets were selected using the DR7 version of the USNO-B proper motions, so are not affected by the systematic astrometric errors in DR8, which were resolved for

Table 1
SEGUE-2 Selection by Target Type

Target Class	Bit Name	Binary Digit	Selection Criteria[a]	Total[b]
Main-sequence Turnoff	SEGUE2_MSTO	0	$18 < g < 19.5$ $0.1 < (g - r) < 0.48$ psfmagErr(g, r, i) < 0.05	39,493
Red K Giants	SEGUE2_RKG	1	$15.5 < g < 18.5$ $r > 15$ $0.8 < (g - r) < 1.3$ $(u - g) > 0.84(g - r) + 1.758$ $(u - g) < 2.4(g - r) + 0.73$ psfmagErr(g, r, i) < 0.05 PMtotal < 11 mas yr^{-1} good PM[c]	1391
l-color K Giants	SEGUE2_LKG	2	$15.5 < g < 18.5$ $r > 15$ $0.7 < (u - g) < 3$ $0.5 < (g - r) < 0.8$ $0.1 < (r - i) < 0.6$ l-color > 0.09[d] psfmagErr(g, r, i) < 0.05 PMtotal < 11 mas yr^{-1} good PM[c]	23,891
Proper-motion K Giants	SEGUE2_PMKG	3	$15.5 < g < 18.5$ $r > 15$ $0.8 < (u - g) < 1.2$ $(u - g) > 2.375(g - r) - 0.45$ $(u - g) < 0.84(g - r) + 1.758$ psfmagErr(g, r, i) < 0.05 PMtotal < 7 mas yr^{-1} good PM[c]	17,978
Low-metallicity Candidate	SEGUE2_LM	4	$15.5 < g < 18$ $0.5 < (u - g) < 2.0$ $r > 15$ $0.3 < (g - r) < 0.8$ l-color > 0.115[d] psfmagErr(g, r, i) < 0.05 psfmagErr(u) < 0.2	16,507
Hypervelocity Star	SEGUE2_HVS	5	$17 < g < 20$ $V_{tot} > 400$ km s^{-1} PMtotal > 8 mas yr^{-1} good PM[c] EITHER: $\mu_{\perp} < 6$ mas yr^{-1} OR $V_{\perp} > 400$ km s^{-1} EITHER: $0.35 < (g - r)_0 < 0.40$ AND $0.375 < (u - g)_0 - 2.5(g - r)_0 < 0.525$ OR $0.40 < (g - r)_0 < 0.60$ AND $0.225 < (u - g)_0 - 2.5(g - r)_0 < 0.425$	581
Extreme M Subdwarf	SEGUE2_XDM	6	$g > 15.5$ $(g - r) > 1.4$ $(r - i) > 0.4$ $15 < r < 20$ $(r - i) < 3.0(g - r) - 3.5$ PMtotal > 10 mas yr^{-1} $H_r > 10.0 + 2.5(g - r)$ [e] psfmagErr(r, i) < 0.05 good PM[c]	10,691
M Giant	SEGUE_MII	7	$15.5 < g < 19.25$ $i > 14.5$ $(u - g) > 1.8 + 0.9(g - r)$	638

Table 1
(Continued)

Target Class	Bit Name	Binary Digit	Selection Criteria[a]	Total[b]
			$(g-r) > 1.3$ psfmagErr$(g, r) < 0.05$ PMtotal < 11 mas yr^{-1} good PM[c]	
High-velocity Halo Star	SEGUE2_HHV	8	$17 < g < 20$ $0.1 < (g-r) < 0.48$ $V_{\rm tan} > 300$ km s^{-1} $V_{\rm tan}/\sigma(V_{\rm tan}) > 3.0$ good PM[c]	4297
Blue Horizontal Branch	SEGUE2_BHB	13	$15.5 < g < 20.3$ $-0.5 < (g-r) < 0.1$ $0.8 < (u-g) < 1.5$	10,660

Notes.
[a] Recall that all target selection is made with dereddened magnitudes, as described in Section 4.
[b] Total number of spectra of stars with this targeting bit set, including duplicate observations. Stars may have more than one targeting bit set.
[c] Acceptance criteria for data in SDSS/USNO-B catalog: (1) Match = 1; (2) dist22 $> 7\farcs0$; (3) pmSigmaRa < 525 mas yr^{-1}; (4) pmSigmaDec < 525 mas yr^{-1}; (5) nFit = 6.
[d] l-color: $-0.436u + 1.129g - 0.119r - 0.574i + 0.1984$; defined for $0.5 < (g-r) < 0.8$.
[e] Reduced proper motion in the noted band: $H_{\rm band} = {\rm mag}_{\rm band} + 5 + 5\log_{10}$(PM (arcseconds yr^{-1})).

DR9 and later. The SEGUE-1 and SEGUE-2 target selection information in the sppTargets and segueTargetAll tables also uses the DR7 version of the recalibrated USNO-B proper motions. Comparisons of the quality and accuracy of the successive versions of the recalibrated USNO-B proper-motion catalog released with DR7, DR8, and DR9 are given in the DR9 data release paper (Ahn et al. 2012), and an overall assessment of the uncertainties is provided in Dong et al. (2011).

All of the SDSS imaging data were rereduced, and the global photometric calibration was updated in DR8 (Aihara et al. 2011; Padmanabhan et al. 2008). The DR8 version of the imaging data was the default imaging catalog for the SDSS imaging data until the DR13 release. For DR13, the imaging data were recalibrated to the Pan-STARRS 1 survey (Finkbeiner et al. 2016; Albareti et al. 2017). These changes in the photometry in turn change the values of parameters used to select stars as targets for SEGUE-2 spectroscopy. This complicates matching the most recent imaging data to the spectra, as stars that made the target section cuts using the SEGUE-2 photometry sometimes do not make the cuts with later versions of the imaging data and vice versa. This is an especially large effect in the categories that look for outliers and rare objects in the photometry and proper-motion data, for example, the M-giant category (Section 4.2.2 and Appendix A). We discuss below in Section 7 how the targeting information was included in the data releases, as well as strategies for combining the information in the imaging and spectroscopy data to create the best samples for particular types of analysis.

The SEGUE-2 photometry in the sppTargets table includes only objects classified as stars in the imaging data, and only objects classed as stars were considered as potential spectroscopic targets. All the SEGUE-2 selection criteria use magnitudes measured by fitting to the point-spread function of the data (Stoughton et al. 2002). The SDSS photometric quality flags are used to reject any object that is saturated in the imaging data. Because too much light through the fibers can cause unacceptable contamination between spectra of stars on adjacent fibers and scattered light, a bright limit of $g_{\rm fib3} > 15.5$ and $r_{\rm fib3} > 15.0$ is imposed on all targets, where fib3 is the $3''$ synthesized fiber magnitudes (Stoughton et al. 2002) and is not corrected for extinction. The $3''$ fiber magnitudes are 0.3 mag fainter on average than the point-spread function magnitudes. Except where explicitly specified, all magnitudes used to select spectroscopic targets in SEGUE-2 are corrected for extinction using the dust maps of Schlegel et al. (1998) so that the sample depth is approximately independent of the foreground extinction. Thus, we omit the usual subscript that indicates that a magnitude has been corrected for extinction when discussing targeting.

The target selection categories for the SEGUE-2 survey are summarized in Table 1. The target classes are shown in color–apparent-magnitude space in Figure 2 and in color–color space in Figure 3. Each class is described in more detail below.

4.1. Blue Horizontal-branch Candidates

BHB stars are powerful tracers of the halo because they are usually produced by an old, metal-poor population and have a narrow range of absolute visual magnitudes, increasing the precision of distance determinations (e.g., Xue et al. 2008; Santucci et al. 2015a). The SEGUE-2 BHB category selects stars bluer than the main-sequence turnoff but not as blue as white dwarfs: $-0.5 < g-r < 0.1$, $0.8 < u-g < 1.5$. This selection contains true BHB stars with low surface gravity as well as higher surface gravity A-colored objects, which we refer to here as "blue stragglers" (BS) or "A main-sequence" objects. We note that for surveys with no u band available, BHB and BS stars may be effectively separated using sufficiently deep z-band observations in combination with g- and either r- or i-band observations (Vickers et al. 2012; Whitten et al. 2019).

The limiting magnitude for this category was $g = 20.3$. Objects that pass the color criteria with $g < 19.5$ have higher priority than objects with $19.5 < g < 20$, which in turn have

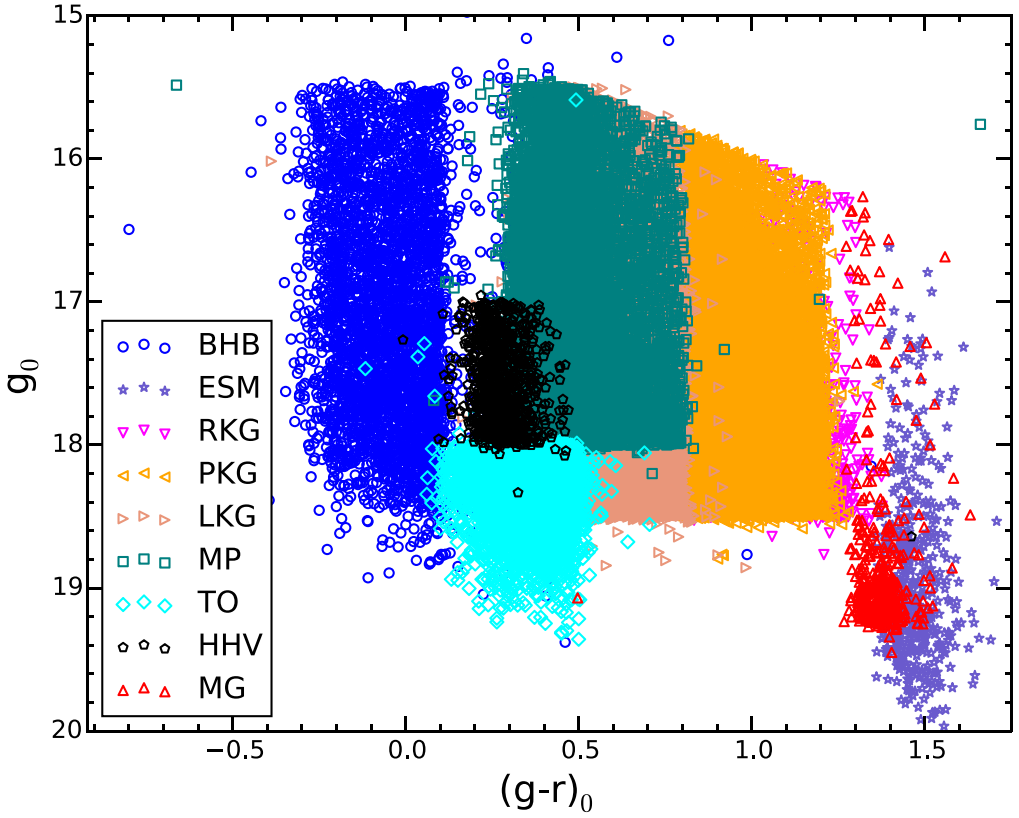

Figure 2. Color–magnitude diagram for the SEGUE-2 sample with major target classes indicated. BHB stands for blue horizontal-branch stars, ESM extremely cool subdwarfs, TO main-sequence turnoff stars, RKG red K giants, LKG *l*-color K giants, PKG proper-motion K giants, MP low-metallicity targets, HHV high-velocity halo stars, and MG M giants.

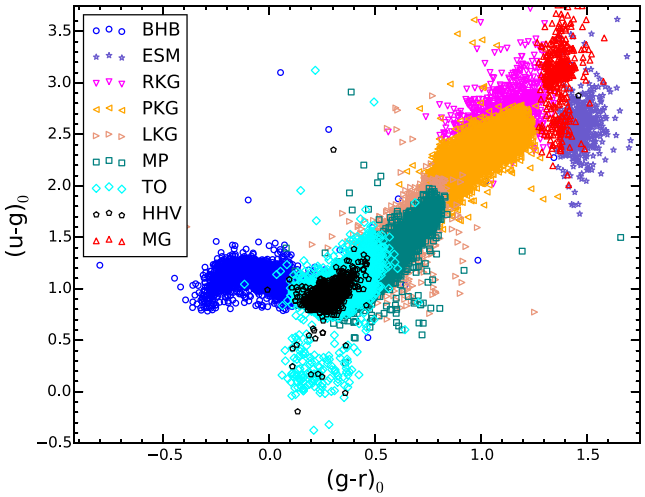

Figure 3. Color–color diagram for the SEGUE-2 sample with major target classes indicated, as described in Figure 2.

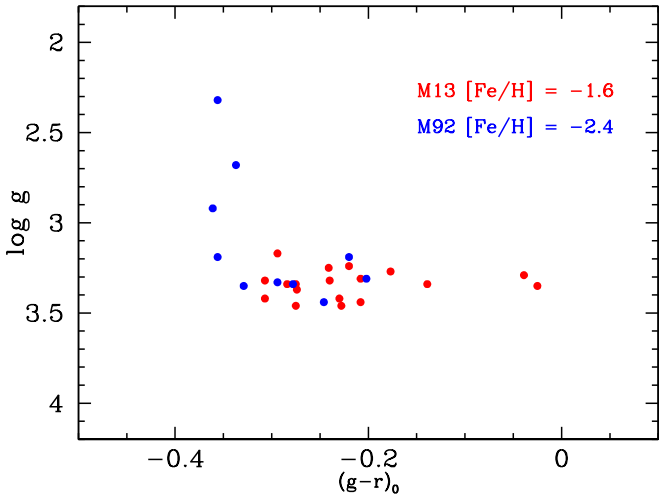

Figure 4. SSPP log g values for spectroscopically verified BHBs that are confirmed members in clusters M13 and M92. It can be seen that most BHB stars have SSPP log g values between 3.1 and 3.5, but that the hottest BHB stars in M92 have lower values, ranging up to 2.3. All of these stars have spectra with S/N > 35.

higher priority than objects with $20 < g < 20.3$. The priorities are used to ensure that all the brighter stars in the footprint are observed because their spectra yield the most accurate velocity and parameter measurements.

The SSPP uses the method of Wilhelm et al. (1999) to combine *ugr* colors and the Balmer line profiles ($D_{0.2}$) to estimate $T_{\rm eff}$, log g, and [Fe/H] for these hot BHB candidates. To isolate a sample of BHB stars from BS or A main-sequence stars, cuts in stellar parameters are required. To investigate the reliability of the parameters for BHB stars, SEGUE-1 targeted them in two clusters: M13 and M92. We identified genuine members of the clusters using a combination of RV, proper motion, and position on the color–magnitude diagram (CMD) (full details are given in Morrison et al. 2016). We show in Figure 4 that these confirmed BHB members of the globular clusters have SSPP log g ranging from 2.3 to 3.5, with only the bluest stars in the very-metal-poor cluster M92 having log g less than 3.0.

Yanny et al. (2009), Harrigan et al. (2010), Deason et al. (2010, 2011), and Fermani & Schönrich (2013) all used cuts in log g, $T_{\rm eff}$, and ugr to identify purer BHB samples to study the halo's kinematics and density distribution. Deason et al. (2011) used a log g range from 2 to 4, while Fermani & Schönrich (2013) suggested a narrower range of log $g = 2$–3.3, the latter being more conservative than the results in Figure 4 suggest.

In addition, some investigators have used reanalysis of the SEGUE spectra to reduce contamination by gaining more information. Sirko et al. (2004b) fit the Balmer line profiles, following Clewley et al. (2002). Adopting this technique can reduce the contamination rate to significantly less than 10%, allowing Xue et al. (2008) to measure the mass of the Milky Way out to 60 kpc, and Ruhland et al. (2011) and Kafle et al. (2012) to study halo streams. Santucci et al. (2015b) added another criterion to the Clewley et al. technique, basing their additional method on the depth and width of Balmer lines, first used by Pier (1983). Santucci et al. (2015a) then use the BHB sample to map the halo's age distribution.

A total of 9744 unique objects in this category were targeted for SEGUE-2. Matched to the DR8 photometry, which is not identical to the targeting photometry, 9660 candidates in this category have DR8 photometry available. Of these 9660 objects with matched DR8 photometry, 7327 have a valid (non –9999) measurement for the parameter `LOGGADOP`.

4.2. Luminous Red-giant Stars

Red giants are the most-luminous tracers available in old populations and so are a particularly important target category for the study of the distant halo. Depending on the metallicity, the tip of the giant branch of an old population falls either in the K-giant region (for metal-poor stars) or the M-giant region (for metal-richer stars). SEGUE-2 used different target selection criteria for K and M giants, described separately below.

4.2.1. K Giants

The most-luminous metal-poor K giants have $M_g \sim -2$, which means that even a star as bright as $g = 18.5$ can be seen to 120 kpc, making these stars the tracers of choice for in situ studies of the outer halo. Unfortunately, at these magnitudes, nearby K dwarfs vastly outnumber distant K giants. SEGUE-2 used three different criteria to separate K giants from K dwarfs.

The three K-giant categories are shown in color–color space in Figure 5. The giant locus in the $u-g$, $g-r$ diagram is affected both by luminosity and metallicity. Giants are found below the locus of foreground dwarfs for the bluer part of the K-giant range, cross the locus near $g-r = 0.8$, and are found above the locus for redder colors.

The rarest targets (at the tip of the giant branch) belong to the red K-giant category, with fewer than 20 per 7 deg^2 plate, except on the lines of sight with the highest stellar density. All stars that passed the red K-giant target criteria were observed (unless blocked by fiber collisions). Within the l-color K-giant region (the l-color is defined in Lenz et al. 1998; see footnote to Table 1), stars with $g-r$ greater than 0.6 were given priority over the bluer stars because the redder giants are more luminous. Unfortunately, there is no way to use ugr to favor the selection of high-luminosity, low-metallicity K giants in the proper-motion K-giant region, but as the name implies, it is still possible to achieve reasonable efficiency for targeting giants in this region with a proper-motion cut to reject foreground

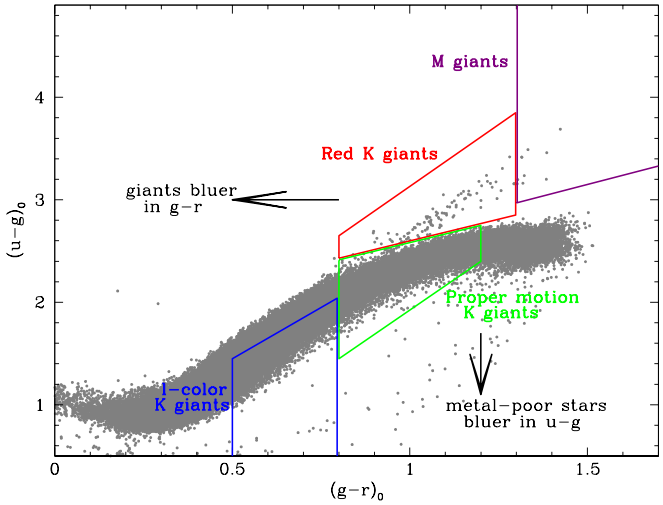

Figure 5. $u-g$ vs. $g-r$ diagram illustrating the K- and M-giant target types for SEGUE-2. Gray points are stars from the Stripe 82 catalog of Ivezić et al. (2007), which have particularly accurate photometry. The red K giants occupy the region where K giants are found above the locus (of predominantly foreground dwarfs), the proper-motion K giants sample this locus at the red end of the K-giant color range, and the l-color K giants occupy the region where metal-poor stars are found below the locus. The M giants form the continuation to the red of the red K-giant sequence.

dwarfs. The limit used was total proper motion <11 mas yr^{-1}, which is a 2σ limit for the SDSS proper motions (see the discussion of the recalibrated USNO-B proper motions earlier in this section). Stars in the proper-motion K-giant region were assigned to fibers until the total number of K-giant targets was 150 per 7 deg^2 plate. If other SEGUE-2 target categories did not have enough stars to meet their fiber quotas, which happened at very high latitudes, the extra fibers were assigned to this category.

Each of the three target types imposes different metallicity biases. The l-color sample is only complete below [Fe/H] $= -1.5$. The red K giants include the most metal-rich stars in the SEGUE-2 K-giant sample because the metal-rich red-giant branch stands out more from the stellar locus and the luminosity separation is largest for metal-rich giants. The targeting of K giants using the $u-g$ versus $g-r$ diagram is discussed in more detail in Yanny et al. (2009). While the combination of the red K-giant and the proper-motion K-giant classes can be used to produce a relatively unbiased sample of the K-giant metallicity distribution, it should be remembered that the actual numbers of stars observed in each category are quite different, and so it is necessary to correct for sampling biases following either the weighting method of Schlesinger et al. (2012) or a Bayesian formalism such as that used by Xue et al. (2015).

The probability of actually observing a K giant is a strong function of magnitude and color. The number of foreground dwarfs increases toward redder colors and fainter magnitudes and in fields that are at lower Galactic latitudes and Galactic longitudes far from the anticenter. In addition, the halo is very centrally concentrated (e.g., Vivas & Zinn 2006; Xue et al. 2015), so the redder, fainter stars that belong to the outer halo are very rare. Therefore, selection efficiency ranges from a few percent for the reddest stars near the faint magnitude limit of our selection at $g < 18.5$ to greater than 50% for the brightest stars in the red K-giant region ($g < 17$), which stand out best