METALLICITY GRADIENTS IN THE MILKY WAY DISK AS OBSERVED BY THE SEGUE SURVEY

JUDY Y. CHENG[1], CONSTANCE M. ROCKOSI[2,9], HEATHER L. MORRISON[3], RALPH A. SCHÖNRICH[4], YOUNG SUN LEE[5],
TIMOTHY C. BEERS[5,6], DMITRY BIZYAEV[7], KAIKE PAN[7], AND DONALD P. SCHNEIDER[8]
[1] Department of Astronomy and Astrophysics, University of California Santa Cruz, Santa Cruz, CA 95064, USA; jyc@ucolick.org
[2] UCO/Lick Observatory, Department of Astronomy and Astrophysics, University of California, Santa Cruz, CA 95064, USA
[3] Department of Astronomy, Case Western Reserve University, Cleveland, OH 44106, USA
[4] Max-Planck-Institute für Astrophysik, Karl-Schwarzschild-Str. 1, D-85741 Garching, Germany
[5] Department of Physics and Astronomy and JINA (Joint Institute for Nuclear Astrophysics), Michigan State University, East Lansing, MI 48824, USA
[6] National Optical Astronomy Observatory, Tucson, AZ 85719, USA
[7] Apache Point Observatory, Sunspot, NM 88349, USA
[8] Department of Astronomy & Astrophysics, Pennsylvania State University, University Park, PA 16802, USA
Received 2011 October 6; accepted 2011 December 24; published 2012 February 1

ABSTRACT

The observed radial and vertical metallicity distribution of old stars in the Milky Way disk provides a powerful constraint on the chemical enrichment and dynamical history of the disk system. We present the radial metallicity gradient, $\Delta[\text{Fe/H}]/\Delta R$, as a function of height above the plane, $|Z|$, using 7010 main-sequence turnoff stars observed by the Sloan Extension for Galactic Understanding and Exploration survey. The sample consists of mostly old thin and thick disk stars, with a minimal contribution from the stellar halo, in the region $6\,\text{kpc} < R < 16\,\text{kpc}$, $0.15\,\text{kpc} < |Z| < 1.5\,\text{kpc}$. The data reveal that the radial metallicity gradient becomes flat at heights $|Z| > 1$ kpc. The median metallicity at large $|Z|$ is consistent with the metallicities seen in outer disk open clusters, which exhibit a flat radial gradient at $[\text{Fe/H}] \sim -0.5$. We note that the outer disk clusters are also located at large $|Z|$; because the flat gradient extends to small R for our sample, there is some ambiguity in whether the observed trends for clusters are due to a change in R or $|Z|$. We therefore stress the importance of considering both the radial and vertical directions when measuring spatial abundance trends in the disk. The flattening of the gradient at high $|Z|$ also has implications on thick disk formation scenarios, which predict different metallicity patterns in the thick disk. A flat gradient, such as we observe, is predicted by a turbulent disk at high redshift, but may also be consistent with radial migration, as long as mixing is strong. We test our analysis methods using a mock catalog based on the model of Schönrich & Binney, and we estimate our distance errors to be $\sim$25%. We also show that we can properly correct for selection biases by assigning weights to our targets.

Key words: Galaxy: abundances – Galaxy: disk – Galaxy: evolution – Galaxy: formation

Online-only material: color figures

1. INTRODUCTION

1.1. Metallicity Gradients

The spatial variation in the metallicity distribution of old stars in the Milky Way disk system is linked to the formation and evolution of the Galaxy. The metallicity of stars at a particular place in the disk depends on the gas accretion rate, star formation history, and subsequent evolution at that location. For example, in the simplest picture of "inside-out" disk formation, low angular momentum gas falls to the center of the halo first, forming stars earlier and becoming chemically enriched much faster than the outer disk (e.g., Larson 1976; Matteucci & Francois 1989; Chiappini et al. 1997; Prantzos & Boissier 2000). In this scenario, heavy element abundances in the interstellar medium generally decrease as a function of Galactocentric radius, i.e., the metallicity gradient is negative. The presence of radial flows (e.g., Lacey & Fall 1985; Goetz & Koeppen 1992; Portinari & Chiosi 2000; Spitoni & Matteucci 2011) and the nature of the early infalling gas (e.g., Cresci et al. (2010)), however, have significant impacts on the chemical evolution of the disk. The exact nature of these processes is not yet fully understood and can lead to a gradient in the early disk that is weaker or even reversed compared to the simplest picture. Thus, observations of both the slope and temporal evolution of the radial metallicity gradient of the disk provide strong

observational constraints for chemical evolution models (e.g., Chiappini et al. 2001; Cescutti et al. 2007; Magrini et al. 2009).

The radial metallicity gradient of the Milky Way disk has been measured using a number of different tracers, including Cepheids and open clusters, yielding a value between ~ -0.01 and -0.09 dex kpc^{-1} (e.g., Caputo et al. 2001; Friel et al. 2002; Chen et al. 2003; Luck et al. 2006; Lemasle et al. 2008; Sestito et al. 2008; Pedicelli et al. 2009; Pancino et al. 2010; Carrera & Pancino 2011; Luck & Lambert 2011). These tracers represent the composition of the interstellar gas at the time that they were formed, thus the wide variety of tracers studied probe the metallicity gradient at different times. This simple picture, however, can be complicated by processes that change the orbits of stars, such as dynamical heating from perturbations like spiral structure, molecular clouds, or minor mergers, which can make gradients shallower or wash them out completely.

In addition to the steepness and time evolution of the gradient, nonlinear features in the metallicity distribution provide further observational constraints. Some authors, for example, have noted the presence of a discontinuity in the radial metallicity gradient at a Galactocentric radius of $R \sim 10\text{–}12$ kpc, beyond which the metallicity gradient becomes shallower or flat (i.e., slope close to or equal to zero) for both open clusters (e.g., Twarog et al. 1997; Yong et al. 2005; Carraro et al. 2007b; Sestito et al. 2008) and Cepheids (e.g., Andrievsky et al. 2002b; Yong et al. 2006; Pedicelli et al. 2009). Whether this transition is characterized by a sharp break or by a smoother transition

---
[9] Packard Fellow.

is still unclear (see, for example, Pancino et al. 2010; Jacobson et al. 2011a).

Several possible explanations for the reported discontinuity have been put forward. Andrievsky et al. (2004) argue that the presence of the Galactic bar and spiral arms may influence the star formation rate and flow of gas throughout the disk, affecting the amount of chemical enrichment that occurs at different radii. Yong et al. (2005) favor star formation in the outer disk, triggered by small accretion events. This scenario was supported by the enhanced abundances of α- and r-process elements of open clusters in the outer disk (e.g., Carraro et al. 2004; Yong et al. 2005). More recent work, however, indicates that these initial measurements may have been too high, and the outer disk clusters have abundances consistent with those in the inner disk (e.g., Carraro et al. 2007b; Bragaglia et al. 2008; Friel et al. 2010).

In this paper, we present the radial metallicity gradient, $\Delta[\text{Fe/H}]/\Delta R$, of the Milky Way disk, as a function of height above the plane $|Z|$, using a sample of 7010 field stars from the Sloan Extension for Galactic Understanding and Exploration (SEGUE; Yanny et al. 2009), part of the Sloan Digital Sky Survey (SDSS; York et al. 2000). The sample covers Galactic coordinates $6\,\text{kpc} < R < 16\,\text{kpc}$, $0.15\,\text{kpc} < |Z| < 1.5\,\text{kpc}$, where R is the cylindrical Galactocentric radius and $|Z|$ is absolute distance from the plane.

Our field stars are older than the open clusters and Cepheids used in previous gradient measurements in the literature and serve to extend the observations of the metallicity distribution of the disk to older tracers, which can provide constraints on the strength of the gradient at early times and on how much radial mixing occurred. Previous work with field stars has indicated that the disk has no radial metallicity gradient far from the midplane (Allende Prieto et al. 2006; Jurić et al. 2008; Katz et al. 2011). Our sample will be useful as a comparison for both studies of open clusters and Cepheids close to the midplane and studies of field stars at large heights above the plane.

Additionally, we examine whether a discontinuity in the radial metallicity gradient exists in the old disk stars. While the distances derived for individual field stars are less accurate than for other tracers, our sample is sufficiently large to divide into bins of $|Z|$, which allows for an examination of the metallicity distribution in the disk as a function of both R and $|Z|$. The distinction between R and $|Z|$ is important, as many of the outer disk tracers in the literature are also located far from the midplane, and the question of whether the reported trends are a function of R, $|Z|$, or both needs to be assessed.

1.2. Thick Disk Formation

This work is further motivated by the idea that the metallicity gradient of the old disk may be used as an observational constraint to distinguish between possible formation mechanisms for the thick disk. Traditionally, the Galactic disk can be thought of as the sum of two components: a thin disk of young metal-rich stars and a thick disk of older, more metal-poor stars (e.g., Gilmore et al. 1995; Chiba & Beers 2000; Bensby et al. 2004; Ivezić et al. 2008). The existence of the Milky Way's thick disk was first noted by Yoshii (1982) and Gilmore & Reid (1983), and thick disks with similar kinematics, structure, and stellar populations have since been observed to be a common feature in nearby spiral galaxies (Dalcanton & Bernstein 2002; Yoachim & Dalcanton 2005, 2006, 2008).

The ubiquity of thick disks in external galaxies, as well as the similarity of their properties, suggests that whatever process is responsible for their existence is important in the formation and evolution of disk galaxies. Furthermore, because the thick disk is old, the properties of its stars can be used as a "fossil record" of the disk's early formation. How stars end up in a thick disk, far from the plane of the Galactic disk, remains an open question.

Several mechanisms for thick disk formation have been proposed, four of which are discussed below. Within the context of the hierarchical structure formation predicted by ΛCDM cosmology, a thick disk may arise through (1) the puffing up or vertical heating of a pre-existing thin disk during a minor merger (e.g., Villalobos & Helmi 2008; Read et al. 2008; Kazantzidis et al. 2008, 2009; Purcell et al. 2009; Bird et al. 2012), (2) the direct accretion of stars formed in satellites that merged with the Galaxy (Abadi et al. 2003), or (3) star formation in an early turbulent disk phase during a period of high gas accretion (e.g., Brook et al. 2004, 2005; Bournaud et al. 2009). Even in the absence of cosmological accretion, a thick disk may also arise through (4) radial migration of stellar orbits (e.g., Schönrich & Binney 2009a, 2009b; Loebman et al. 2011).

Each of these scenarios is motivated by both theory and observations. Halo merger histories in cosmological N-body simulations suggest that the types of mergers required by scenarios 1 and 2 are common; Stewart et al. (2008) estimate that 70% of Milky Way sized halos have experienced a 1:10 merger within the last 10 Gyr. Streams in the halo of the Milky Way (e.g., Newberg et al. 2002; Belokurov et al. 2007) and other galaxies (e.g., Ibata et al. 2001; Martínez-Delgado et al. 2010) provide observational evidence of such accretion events. Other recent work, however, has emphasized the importance of smooth gas accretion in the growth of disk galaxies (Brooks et al. 2009; Dekel et al. 2009), which can lead to a turbulent disk with high velocity dispersion as in scenario 3. Observational support for this picture includes the clump-cluster galaxies (Elmegreen & Elmegreen 2005) and the thick chain and spiral galaxies (Elmegreen & Elmegreen 2006) seen at high redshift. Lastly, Sellwood & Binney (2002) and Roškar et al. (2008b, 2008a) showed that resonant interactions between stars and transient spiral waves can change the radii of stellar orbits while keeping them on circular orbits, leading to the kind of radial mixing necessary for scenario 4. Observations of nearby stars indicate that radial migration is an important process that may shape the correlations between the kinematics, metallicities, and ages of stars in the solar neighborhood (Haywood 2008).

While there is evidence that the mechanisms described above—minor mergers, early gas accretion, and radial migration—are at play in galaxy formation, the question of which mechanism, if any, is the *dominant* force behind thick disk formation remains unanswered. Recent work by Sales et al. (2009) is an example of how the kinematics predicted by the four different scenarios—in particular, the distribution of stars' orbital eccentricities—can be used to test the various scenarios. This approach has been taken observationally by Dierickx et al. (2010) and Wilson et al. (2011), using stars from the SDSS and the RAdial Velocity Experiment (RAVE; Steinmetz et al. 2006), respectively. Both studies disfavor the minor merger scenarios, as they do not observe enough stars with high orbital eccentricities. Numerical simulations by Di Matteo et al. (2011), however, suggest that the observed eccentricity distribution can be obtained in the minor merger scenario given different orbital parameters and satellite properties than those used in the Sales et al. (2009) analysis. The conflicting interpretations show the need for other observational constraints. In this paper, we examine the radial

metallicity gradient, $\Delta[\mathrm{Fe/H}]/\Delta R$, as a function of height above the Galactic plane, $|Z|$, in order to further distinguish between different thick disk formation scenarios.

For example, if thick disk stars originated in a thin disk and were subsequently heated by a minor merger, as in scenario 1, then the observed radial metallicity gradient depends on the amount of mixing in the radial direction. If most of the heating occurs in the vertical direction, the thick disk will have the same metallicity gradient as the initial thin disk. Simulations, however, show that there can be substantial heating in the radial direction (e.g., Hayashi & Chiba 2006; Kazantzidis et al. 2009; Bird et al. 2012), which suggests that the gradient may be more shallow or flat than expected. If thick disk stars originated outside of the Galaxy and were deposited in the thick disk via accretion events, as in scenario 2, then the metallicity distribution may exhibit clumpiness. The simulations of Abadi et al. (2003) showed that a single disrupted satellite roughly ends up in a torus of stars; several disrupted satellites would make up the thick disk by contributing stars of different metallicities at different radii.

If thick disk stars originated in a turbulent gas disk at high redshift, as in scenario 3, the short timescale for star formation makes the thick disk chemically homogeneous, with no metallicity gradient (Brook et al. 2005; Bournaud et al. 2009). If thick disk stars originated in a thin disk and were pushed to larger radii through radial migration, as in scenario 4, then the original gradient would become washed out into a shallow or nonexistent radial metallicity gradient (Roškar et al. 2008a; Sánchez-Blázquez et al. 2009). The presence of a gradient, then, would rule out chemical homogeneity (scenario 3), while the strength of the gradient would constrain the amount of disk heating by minor mergers (scenario 1) and radial migration (scenario 4). Examining old disk stars in a large volume, beyond the solar neighborhood, will allow one to distinguish between the various scenarios.

In contrast to many previous studies, we do not assign our stars to a thin disk or thick disk component. For samples of nearby stars in the solar neighborhood, this division is often done by assuming that the thick disk has a larger velocity dispersion and a slower mean rotation (i.e., thick disk stars are kinematically hot), as in the cases of Bensby et al. (2003) and Venn et al. (2004). Another method of separating thin and thick disk stars is by their chemistry, as Lee et al. (2011b) do using [Fe/H] and [α/Fe]; they favor this type of division because a star's composition is less likely to change than its spatial location or kinematics. Schönrich & Binney (2009b) have used mock observations to show that separating thin and thick disk stars using chemistry versus kinematics yields samples with different properties.[10] In addition, whether or not the thin disk and thick disk are truly distinct components is still an open question, with some studies arguing that the two components arise from a smooth correlation between chemical and kinematic properties (see discussions by Haywood 2008 and Ivezić et al. 2008).

Because our sample is not restricted to the solar neighborhood, we can compare the stellar populations of the thin disk and thick disk based on stars' locations instead of their kinematics or chemistry, which allows us to avoid assuming a specific

model for the Milky Way disk. In this paper, we do not assign individual stars to the thin disk or thick disk. Instead, we will use the term *thick disk* to refer to stars that are currently found at large distances from the plane on their orbits; the term *thin disk* refers to stars that are found close to the Galactic midplane. Based on double-exponential fits to the vertical scale heights of the stellar density distribution in the disk (e.g., Jurić et al. 2008; de Jong et al. 2010), we expect the thick disk to be the dominant population above $|Z| \sim 1.0$ kpc.

Previous analyses of samples outside of the solar neighborhood have found no radial metallicity gradient at vertical heights $|Z| > 1.0$ kpc (Allende Prieto et al. 2006; Jurić et al. 2008; Katz et al. 2011). Our sample is complementary because our lines of sight are located at relatively low Galactic latitude and we can directly compare the radial metallicity gradients of the thin and thick disks (up to $|Z| = 1.5$ kpc) using the same sample. In addition, we can explore whether the reported discontinuity in the outer disk is a purely *radial* trend or if a *vertical* trend is contributing to the observed flattening of the gradient at large R.

The paper is organized as follows: the sample selection and data are described in Section 2. We then describe our methods of determining distances and correcting for the selection function in Sections 3 and 4, respectively. Our gradient measurements are presented in Section 5. Error analysis is presented in Section 6. We discuss the results in Section 7 and conclude with a summary in Section 8. For readers who are only interested in the results, we recommend skipping Sections 4 and 6. Further description of our weighting scheme, introduced in Section 4, is provided in the Appendix. Throughout our analysis we adopt the Galactocentric radius of the Sun, $R_{\mathrm{GC},\odot} = 8.0$ kpc.

2. DATA

2.1. Sample Selection

We measure the metallicity gradient of old main-sequence turnoff (MSTO) stars in low Galactic latitude fields from the SEGUE survey (Yanny et al. 2009; Aihara et al. 2011; Eisenstein et al. 2011; C. M. Rockosi et al. 2012, in preparation). These stars allow us to reach the largest distances probed by main-sequence stars within the fixed magnitude limits of the survey. The data were obtained using the same telescope (Gunn et al. 2006), camera (Gunn et al. 1998), and filter system (Fukugita et al. 1996) as the SDSS. The old MSTO is selected to be in the blue part of the color–magnitude diagram (CMD), as described in detail below, and can be identified using the SDSS Data Release 7 (DR7; Abazajian et al. 2009) version of the Catalog Archive Server[11] as targets with `sspParams.zbclass = STAR`, `SpecObjAll.primTarget = 2048`, and `PlateX. programName = seglow%`. We also require that there are no repeat observations so that each star is only counted once. An equivalent query in Data Release 8[12] (DR8; Aihara et al. 2011) is `SpecObjAll.class = STAR`, `SpecObjAll.primTarget = 2048`, `PlateX.programName = seglow%`, and `PlateX. isPrimary = 1`, where the last requirement removes repeat observations.

The `programName` qualifier selects our targets from a subset of 22 SEGUE plug-plates[13] that comprise the "low-latitude" pointings, which are restricted to Galactic latitudes $8° < |b| < 16°$ (see Section 3.15 of Yanny et al. 2009). These lines of sight

---

[10] For example, a star assigned to the thin disk using chemical criteria may be assigned to the thick disk using kinematic criteria if it belongs to the inner disk and is in the tail of the rotational velocity distribution. According to the Schönrich & Binney (2009b) model, this explains the tail of thick disk stars with high [Fe/H] and why most of the thick disk stars in the Bensby et al. (2003) and Venn et al. (2004) samples are located at radii within the solar circle.

[11] http://casjobs.sdss.org/CasJobs/
[12] http://skyservice.pha.jhu.edu/casjobs/
[13] Hereafter, we simply refer to the SDSS/SEGUE "plug-plates" as "plates."

Table 1
Properties for 11 Lines of Sight

Plates		R.A. (deg)	Decl. (deg)	l (deg)	b (deg)	$E(B-V)^a$	$N_{spectra}$
2712	2727	105.6	12.4	203.0	8.0	0.09	744
2536	2544	286.7	39.1	70.0	14.0	0.15	661
2534	2542	277.6	21.3	50.0	14.0	0.17	672
2554	2564	303.0	60.0	94.0	14.0	0.19	734
2678	2696	98.1	26.7	187.0	8.0	0.24	766
2556	2566	330.2	45.1	94.0	−8.0	0.31	728
2668	2672	79.5	16.6	187.0	−12.0	0.33	830
2681	2699	71.5	22.0	178.0	−15.0	0.41	758
2537	2545	334.2	69.4	110.0	10.5	0.49	708
2538	2546	323.1	73.6	110.0	16.0	0.65	716
2555	2565	312.4	56.6	94.0	8.0	0.82	511

Note. [a] Median value for spectra in line of sight using values from SFD98.

are high enough to avoid the young star-forming disk, as well as the regions with the most crowding and highest reddening, but also sufficiently low that they have a long sightline through the disk. The lines of sight fall into roughly two groups in Galactic longitude: seven at $50° < l < 110°$ and another four toward the anticenter, $170° < l < 210°$.

Each plate covers 7 deg^2 on the sky, with targets in the magnitude range $16 < g < 20$, where the magnitudes have not been corrected for extinction. In this paper, we will refer to any reddening- or extinction-corrected magnitudes and colors with subscripts g_{SFD} and g_0, for corrections derived from Schlegel et al. (1998, hereafter SFD98) and isochrone fitting (see Section 3), respectively. Table 1 lists the properties of the 11 lines of sight (two plates per pointing) included in our sample, ordered by the median extinction $E(B-V)$, obtained from SFD98. For the total sample, $E(B-V)$ varies between 0.05 and 1.07 mag. On average, there were 600–700 spectra obtained per line of sight.

For most of the SEGUE survey, which was at high Galactic latitude, targets are identified as MSTO stars based on their $(u-g)_{SFD}$ and $(g-r)_{SFD}$ colors (see Yanny et al. 2009). In ugr color space, it is possible to separate the MSTO stars from metal-poor halo stars because of the large ultraviolet excess of metal-poor stars. In the low-latitude pointings, however, two issues arise. First, the u-band magnitudes and their uncertainties are unreliable due to the large extinction in these regions. Second, it is impossible to use a single constant selection in $(g-r)_{SFD}$ that will yield the same stellar population along every line of sight because the reddening in these fields is, on average, much higher and more variable than in the high-latitude fields. For these low-latitude fields, we use a targeting procedure that is more robust to reddening, which will reliably choose the stars at the blue edge of the CMD. Starting with the photometric objects identified as stars in the imaging, this selection procedure is as follows.

1. We remove all stars with $g > 20$ and $i < 14.2$ (using magnitudes *uncorrected for extinction*) to ensure that the targets are sufficiently bright for high-quality spectroscopy in the expected exposure time.
2. The 7 deg^2 area of each plate is large enough that the extinction can be highly variable across the plate, and there are always many more targets than fibers available. We remove the regions of highest extinction from consideration to maximize the number of useful spectra. For each

half of the plate, we calculate the 75th percentile of the $E(B-V)$ distribution using the total line-of-sight extinction from SFD98. This procedure is done for each half independently to ensure that the targets are approximately evenly distributed over the plates, since the reach of the fibers is only about half of the plate diameter.

3. We remove all objects with $E(B-V)$ larger than the higher of the two 75th percentile values. Taking the higher value ensures that there are enough usable targets on each half of the plate to fill all the fibers given their limited reach across the plate. This should not bias the sample, as we do not expect that the objects behind more extinction are intrinsically any different from those that are unobscured. This is especially true for distant objects that are located far behind the dust. *The magnitudes used throughout the rest of this procedure are corrected using the SFD98 extinction values.* The SFD98 extinction was applied so that stars in the same approximate luminosity range were targeted along each line of sight, despite the large variation in extinction among the different lines of sight.

4. We examine the $(g-r)_{SFD}$ distribution in bins of g_{SFD}, each 1 mag wide. For each distribution, we find the peak, which is the $(g-r)_{SFD}$ color of the MSTO in each g-magnitude bin. In addition, we determine the half-maximum on the blue side; this is, $(g-r)_{half-max}$.

5. The red cut for each bin is defined as $(g-r)_{cut} \equiv (g-r)_{half-max} + 0.25$. We fit a line to $(g-r)_{cut}$ as a function of g_{bin}, where g_{bin} is the mean g_{SFD} of all the stars in each bin.

6. All stars on the blue side of the line are defined as candidate spectroscopic targets with equal probability of being selected. Targets are randomly chosen from the resulting candidate list.

Though the $(g-r)_{SFD}$ color of the population may change from field to field because of varying amounts of extinction, the identification of the MSTO stars as the bluer population holds for all lines of sight. As a result, this method is more robust to reddening than the standard color cuts for normal SEGUE plates. Halo contamination is expected to be low in these plates, compared with the higher-latitude pointings (see further discussion in Section 6.2). The color cut, however, will bias our sample against metal-rich stars, which have redder colors. The severity of the bias depends on how much the MSTO color changes with metallicity, which in turn depends on the ages of stars at each metallicity (see further discussion in Sections 6.1 and 6.3.2).

Figure 1 shows the results of the procedure outlined above for two lines of sight with low and high extinction (median $E(B-V) = 0.17$ and 0.41, respectively). The density of objects identified as stars in the photometry are plotted in gray scale and contours, while MSTO stars in our spectroscopic sample are plotted as blue circles. All photometric objects bluer than the red limit of the spectroscopic sample were considered as candidates for spectroscopy, but only a randomly selected subset of those were actually observed. In the 22 low-latitude plates, spectra of 7828 MSTO stars were taken. We keep the targets with good photometry and spectra with signal-to-noise ratio (S/N) > 10 pixel^{-1}, where each pixel corresponds to ~1 Å. We remove any stars which have large discrepancies (>0.1 mag) in $g-r$ color between different photometric reductions. We also remove rare blue stars with $(g-r)_{SFD} < −0.25$. The resulting sample contains 7655 spectra with a mean S/N ~ 30 pixel^{-1}.
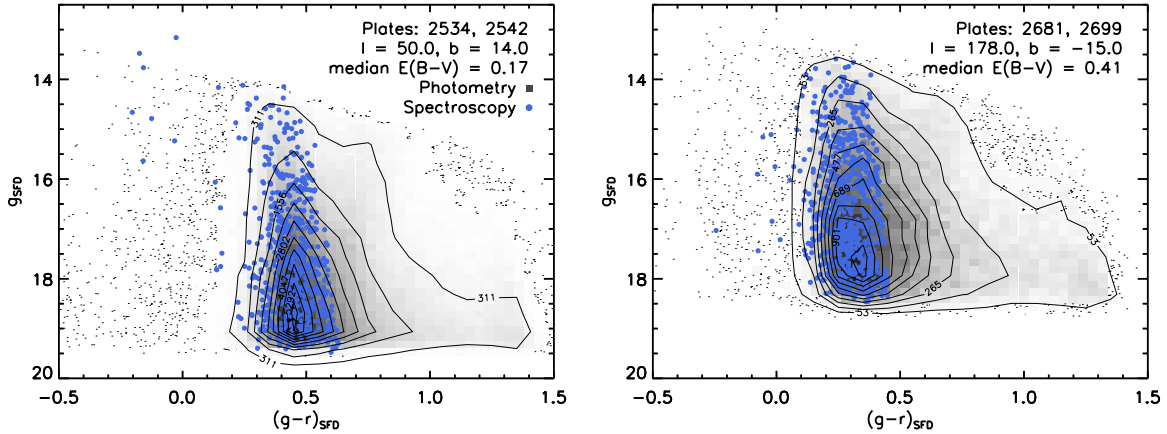
Figure 1. Main-sequence turnoff (MSTO) selection. Spectroscopic targets (blue circles) are a randomly selected subset of the MSTO stars identified using the photometry of all objects in the field (gray scale and contours). The contour labels indicate the number of stars per 0.375 by 0.1 mag g_{SFD}–$(g − r)_{SFD}$ bin. The results are shown for two lines of sight with low and high extinction (left and right, respectively), as measured by SFD98.

Table 2
Cluster Metallicities Measured by the SSPP

Cluster[a]	[Fe/H]$_{\text{Lit}}$	[Fe/H]$_{\text{SSPP}}$	[Fe/H]$_{\text{Lit}}$ − [Fe/H]$_{\text{SSPP}}$
M3	−1.50	−1.43	−0.07
M71	−0.82	−0.74	−0.08
NGC 2158	−0.25	−0.29	+0.04
NGC 2420	−0.20[b]	−0.31	+0.11
M35	−0.16	−0.24	+0.08
M67	+0.05[c]	+0.00	+0.05
NGC 6791	+0.30	+0.29	+0.01

Notes.
[a] Cluster data are also presented in Smolinski et al. (2011), with the exceptions of NGC 2420 and M67. A full comparison of the SSPP for all data in the SEGUE cluster samples will be presented in C. M. Rockosi et al. (2012, in preparation).
[b] Jacobson et al. (2011a).
[c] Randich et al. (2006).

2.2. SEGUE Stellar Parameter Pipeline: Accuracy in Regions of High Extinction

The SEGUE Stellar Parameter Pipeline (SSPP; Lee et al. 2008a) estimates the effective temperature T_{eff}, surface gravity log g, and metallicity [Fe/H] for each spectroscopic target in the survey. We use stellar parameters from the version of the SSPP used for DR8, which includes improved [Fe/H] estimates at both high and low metallicities (Smolinski et al. 2011). The SSPP has been extensively tested using globular and open clusters, where true cluster members are identified using their metallicities and radial velocities (Lee et al. 2008b; Smolinski et al. 2011).

We verify that these results hold for cluster members in our temperature and surface gravity range ($5000 < T_{\text{eff}} < 7000$ K, log $g > 3.3$). Table 2 shows the comparison between the literature values (Column 2) and the SSPP for cluster members in our temperature and surface gravity range (Columns 3 and 4). The offsets in [Fe/H] between the literature values and the SSPP are small (within 0.1 dex), and we see no trends with T_{eff}, [Fe/H], and S/N. These tests show that we can reliably measure trends in [Fe/H] throughout our entire sample volume, and that the absolute values of the metallicities presented in the paper are accurate to 0.1 dex or better.

Each parameter is estimated using multiple methods: 11 for T_{eff}, 10 for log g, and 12 for [Fe/H]. For T_{eff} in particular, these methods include spectral fitting and χ^2-minimization using grids of synthetic spectra (ki13, k24, NGS1), measuring line indices (WBG, HA24, HD24), neural networks using training sets of both real and synthetic spectra (ANNRR, ANNSR), and $g − r$ color predictions (T_K, T_G, T_I). See Lee et al. (2008a) for complete details on each of these individual methods. Some methods use only the spectra as input, some use only the photometry, and some use both. The individual estimates of each parameter are averaged to obtain a final adopted value. Each individual estimate is valid for some range of $(g − r)_{SFD}$ color and S/N, which determines whether it will be included in the final average. A more complicated decision tree is used for [Fe/H] and is described in detail in the Appendix of Smolinski et al. (2011). We have 7605 stars that have good stellar parameters; their temperatures fall in the range $5000 K < T_{\text{eff}} < 7000$ K, making them F and G dwarfs.

Before using the SSPP parameters, we test whether the parameter estimates are affected by high Galactic extinction. The SSPP was designed to analyze the normal SDSS and SEGUE data at high Galactic latitude and uses photometry that has been corrected for extinction using the reddening maps of SFD98. These extinction values reflect the total line-of-sight extinction, which means that the colors of less distant stars will be overcorrected; they will be too blue. This effect is likely to result in a systematic error in the parameters estimated by those methods in the SSPP that use the photometry.

If this overcorrection affects our sample, we expect estimates that use the photometry to be systematically different when the extinction is high. The left panels of Figure 2 show the discrepancy between the individual T_{eff} estimates and the adopted value as a function of $E(B − V)$. Whether or not the estimate includes the photometry is indicated in the bottom left of each right-hand panel. We have only plotted estimates that were accepted by the SSPP (i.e., the target falls in the $(g − r)_{SFD}$ or S/N range in which the particular method is reliable). As expected, the photometry-dependent estimates of T_{eff} (k24, WBG, T_K, T_G, T_I) are systematically higher (i.e., the color is bluer), compared to the adopted value, for the highest values of extinction.

To remove this effect, we calculate new averaged values of T_{eff} using only the estimates from methods that exclude the photometry (ki13, ANNSR, ANNRR, NGS1, HA24, HD24). The right panels of Figure 2 show the discrepancy between the individual T_{eff} estimates and the newly calculated spectra-only value as a function of $E(B − V)$. In contrast to the left
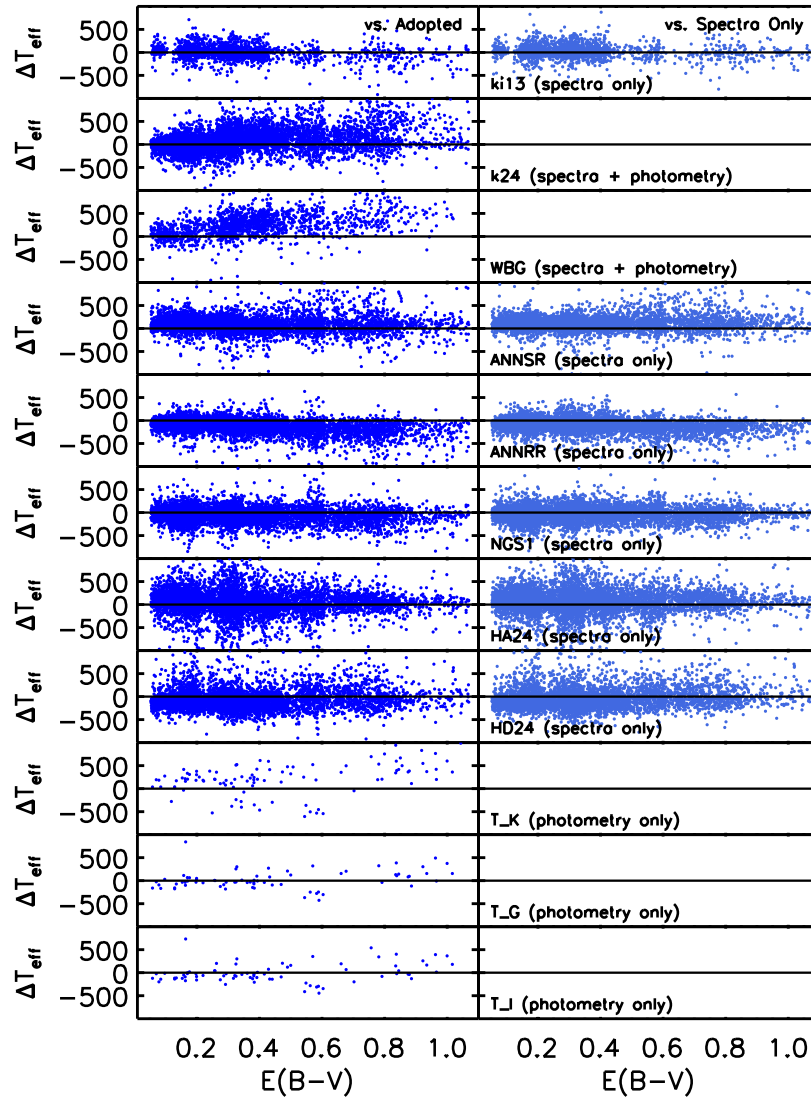
Figure 2. Left panels: differences between individual SSPP temperature estimates and the adopted values as a function of SFD98 extinction $E(B - V)$. For estimates that include photometry (k24, WBG, T_K, T_G, T_I), the deviation from the adopted value increases with extinction, indicating that the temperatures are overestimated due to overcorrection when using the SFD98 extinction values. Right panels: differences between individual SSPP temperature estimates and the spectra-only values as a function of extinction. The trend is no longer evident in the spectra-only estimates, indicating that the spectra-only temperature is more reliable for highly extincted objects.

(A color version of this figure is available in the online journal.)

panels, there is no clear trend in $\Delta T_{\rm eff}$ with extinction for the individual spectra-only estimates, which suggests that our new spectra-only average of the temperature is more reliable. At $E(B - V) > 0.8$, the median discrepancy between the new spectra-only estimate and the original adopted value is about 100–200 K, which is comparable to the expected errors of the SSPP $T_{\rm eff}$. We show in Section 6.3.1 that this could amount to a systematic error in the distance of $\sim 20\%$–25%. For the remainder of the paper, $T_{\rm eff}$ will refer to the spectra-only value.

No trend with extinction is observed in the photometry-dependent estimates of [Fe/H] (k24, WBG, CaIIK2, CaIIK3, ACF, CaII), so we keep the adopted values. This provides a more robust result, as the adopted value is an average of a larger number of estimators.

3. DISTANCES

To calculate distances to each target, we use the spectra-only $T_{\rm eff}$ and SSPP [Fe/H], plus the theoretical isochrones of An

et al. (2009),[14] which have been shown to be good matches to *ugriz* cluster fiducials. We do not make use of the SSPP log *g* estimate, as it is relatively inaccurate near the turnoff, where the expected range of gravities is small compared to the errors. Twelve sets of An et al. (2009) isochrones at metallicities in the range $-3.0 < {\rm [Fe/H]} < +0.4$ are available, with each set having a range of ages up to 15.8 Gyr. The distance uncertainties are discussed in Section 6.3.1.

We assign all stars in the sample to the isochrone with the closest metallicity. We then find the mean temperature for the stars in each metallicity bin by fitting a Gaussian to the distribution of effective temperatures. We identify the age of the isochrone with the turnoff temperature closest to the measured mean temperature; we refer to this as the *turnoff age of the mean temperature*, or TAMT, for each metallicity. Targets hotter than the mean temperature cannot be placed on the TAMT isochrone, so younger isochrones must be used; these targets are assigned

---

[14] http://www.astronomy.ohio-state.edu/iso/sdss.html

Table 3
Mean Isochrone Ages as a Function of [Fe/H]

[Fe/H]	[α/Fe]	log ⟨T_{eff}⟩	⟨T_{eff}⟩ (K)	log ⟨age⟩	⟨age⟩ (Gyr)	N_{obj}
[Fe/H] < −2.50	+0.4	3.768	5864	10.2	15.8	9
−2.50 < [Fe/H] < −0.75	+0.3	3.768	5864	10.20	15.8	880
−0.75 < [Fe/H] < −0.40	+0.2	3.775	5958	10.05	11.2	2695
−0.40 < [Fe/H] < −0.25	0.0	3.772	5912	10.05	11.2	1527
−0.25 < [Fe/H] < −0.15	0.0	3.772	5912	10.00	10.0	857
−0.15 < [Fe/H] < −0.05	0.0	3.772	5912	9.95	8.9	732
−0.05 < [Fe/H] < +0.05	0.0	3.763	5796	10.00	10.0	480
+0.05 < [Fe/H] < +0.15	0.0	3.763	5796	9.95	8.9	229
+0.15 < [Fe/H] < +0.30	0.0	3.763	5796	9.85	7.1	142
+0.30 < [Fe/H]	0.0	3.763	5796	9.70	5.0	54

to the oldest possible age (i.e., the oldest isochrone where the target is cooler than the turnoff).

Figure 3 shows a schematic picture of how ages are assigned to targets depending on their effective temperatures. The top panel shows six representative isochrones at solar metallicity, while the bottom three panels show how stars are assigned to these isochrones based on their place in the T_{eff} distribution. The procedure described above is shown in the panel labeled "TAMT" while the panels labeled "TO" and "ZAMS" show two other age assumptions that we use to estimate our distance errors (Section 6.3.1). Table 3 lists the TAMT determined by finding the mean temperature for each metallicity bin for the 7605 stars in our sample. To test the effect of using different stellar evolutionary models, we also calculate the TAMT using the isochrones of the Dartmouth[15] group (Dotter et al. 2008) and find that the results are within 0.05 dex of those listed in Table 3.

Once we have assigned a given target to an isochrone with a particular age and metallicity, we use the isochrone to obtain the predicted $g − r$ color for the target's spectra-only T_{eff} by linearly interpolating on the isochrone in temperature–color space. A comparison of the predicted and observed $g − r$ gives an estimate of the *isochrone extinction* in $g − r$, which is also used to determine the extinction in the g band. The isochrone extinction is an improvement over the SFD98 values because it does not assume that the target lies behind all of the dust in the line of sight. We step through this procedure for the two isochrones with the nearest values of [Fe/H] and then linearly interpolate to find the predicted apparent and absolute magnitudes. The apparent magnitude, now corrected using the isochrone extinction, along with the predicted absolute magnitude in the g band yield the distance.

In addition to being used in the distance calculation, the isochrone extinction provides information about the dust distribution along different lines of sight in the field. The left panel of Figure 4 shows the isochrone extinction as a function of the derived distance for one line of sight, with the color indicating the SFD98 value. There is good general agreement between the two extinction estimates, especially on a relative scale. Quantitatively, the agreement is poorest for small distances where the targets are in front of some of the dust, but asymptotes to better agreement at large distances. This is consistent with the idea that the SFD98 values are overestimates because they include all of the dust in the given line of sight. For the line of sight shown, most of the dust is located within 2 kpc of the Sun, beyond
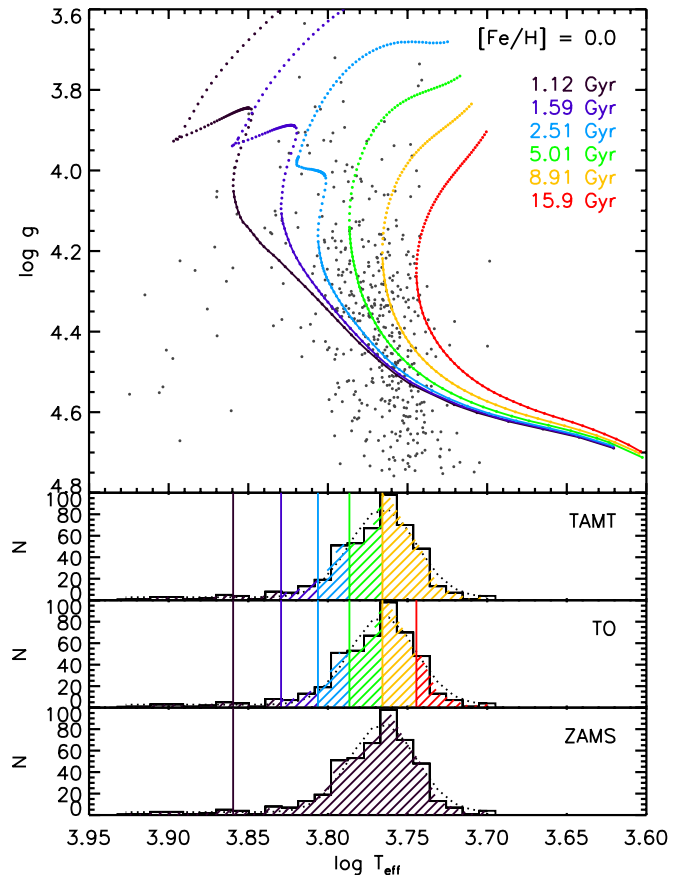

Figure 3. Determination of isochrone ages for solar metallicity targets (−0.05 < [Fe/H] < 0.05, gray dots). Isochrones are shown in the top panel, with the temperature distribution in the bottom three panels. In practice, the set of isochrones used at each metallicity includes all available ages, but for clarity we show only six—the oldest, the youngest, and those located at the mean and 1σ, 2σ, and 3σ values of the temperature distribution. The colors in the lower panels indicate which isochrones in the top panel are used to calculate distances for those targets. We show how ages are assigned to our targets making three possible age assumptions: turnoff age of the mean temperature (TAMT), turnoff (TO), and zero-age main sequence (ZAMS). TAMT: for each metallicity, most targets are assigned to a single age (the TAMT) at which the turnoff temperature is closest to the measured mean of the temperature distribution of the SEGUE targets. Targets hotter than the mean are assigned the oldest possible age (i.e., the oldest isochrone where the target is cooler than the turnoff). TO: all targets are assigned to the oldest possible age, even those cooler than the mean of the temperature distribution. ZAMS: all targets are assigned to the youngest possible age. The TAMT assumption is used in our analysis, while the TO and ZAMS assumptions are used to test the accuracy of our distance estimates (see Section 6.3.1).
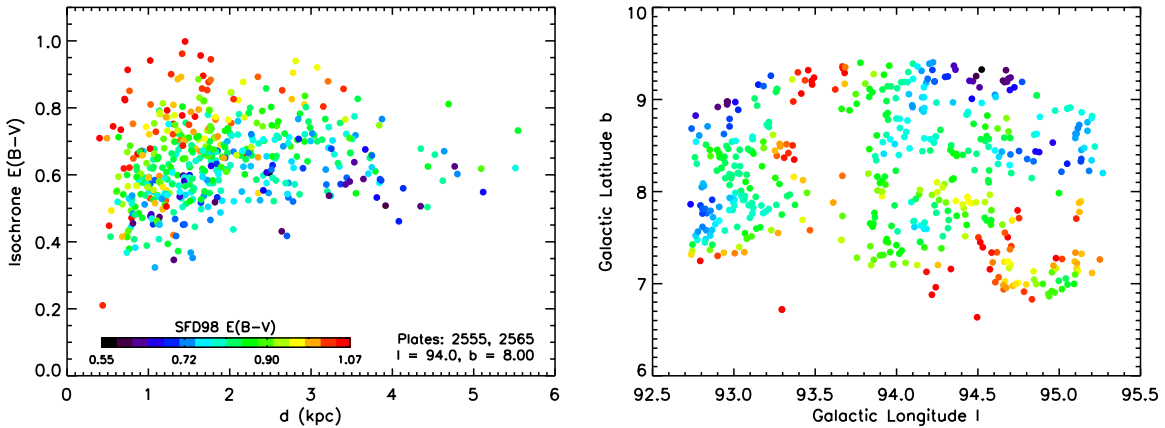
15 http://stellar.dartmouth.edu/~models/grid.html

Figure 4. Estimates of the extinction derived from isochrone fitting provide a picture of the dust distribution along a given line of sight. The left panel shows the isochrone extinction as a function of derived distance, d, from the Sun. At small d the isochrone extinction is smaller than the SFD98 value (indicated by color symbols), but at large d it asymptotes to the SFD98 value. For this line of sight, which has the highest median $E(B - V)$ of our 11 fields, most of the dust is found within 2 kpc of the Sun. The scatter in isochrone extinction for a given value of SFD98 extinction may reflect the patchiness of the dust on the plane of the sky. This effect is seen in the right panel, which shows multiple separate patches that have the same SFD98 extinction values, but may have different distributions of dust along the line of sight. The color coding in the right panel is the same as in the left.

which the isochrone extinction is approximately constant as a function of distance.

The scatter in the left panel of Figure 4 may be explained by the patchiness of the extinction on the plane of the sky. The right panel of Figure 4 shows that the distribution of $E(B - V)$ varies on small scales. The scatter in the left panel is expected if each region of high extinction has a different dust distribution along the line of sight. The agreement between the isochrone extinction and the SFD98 values provides a sanity check which indicates that our isochrone extinction estimates are reliable.

Sixty-two stars in our sample (0.8%) end up with negative values of isochrone extinction. Three kinds of stars fall into this category: (1) 17 stars are outliers with very blue $g - r$ colors. These stars do not contribute to the final measurement; after we apply our weighting scheme they receive a CMD weight of zero (see Section 4). (2) 7 stars exhibit large changes in $g - r$ color between different photometric reductions and are likely to be catastrophic errors, possibly due to blending in the relatively crowded, low-latitude fields. (3) The remaining 38 stars tend to be faint, and likely have negative reddening because the errors in their temperatures cause their predicted $g - r$ colors to be redder than their observed colors. We only see these stars in fields where the reddening is low; in fields where reddening is higher, the observed $g - r$ is much redder than the predicted $g - r$, so the temperature errors are not large enough to cause stars to have negative reddening. We expect just as many objects to have temperature errors that cause their predicted $g - r$ colors to be bluer, and we leave all of these objects in our sample, which should give us a more symmetric distribution of errors in our distances.

With distances for all of our targets, we obtain the spatial distribution of our sample in Galactic coordinates R and $|Z|$, as shown in Figure 5. The SEGUE targets are shown as blue dots. Our radial coverage extends mostly to Galactocentric radii at the solar circle and beyond, though at low $|Z|$ it is confined just to the solar circle. About one-third of our sample is located below the Galactic plane; the $b < 0°$ lines of sight cover the entire radial range outside of the solar circle. Forty-eight open clusters and 190 Cepheids analyzed in the literature using high-resolution spectra are also shown (open symbols and crosses, respectively). These data have been used to study the radial metallicity gradient of the disk and will serve as comparison

samples. The properties of the open clusters are listed in Table 4, and the data are described in more detail in Section 7.2.

We use our sample of MSTO stars to measure the metallicity gradient of the old disk in four different $|Z|$ slices, as indicated by the black lines in Figure 5. Taking thin and thick disk scale heights to be ∼300 and 900 pc, respectively, the lower two slices (0.15 kpc < $|Z|$ < 0.25 kpc and 0.25 kpc < $|Z|$ < 0.5 kpc) are dominated by the thin disk. The third slice (0.5 kpc < $|Z|$ < 1.0 kpc) is made up of a mix of the thin and thick disks, and the fourth slice (1.0 kpc < $|Z|$ < 1.5 kpc) is dominated by the thick disk. Of our sample, 7180 stars fall into these four slices.

Dividing our sample in this way allows a comparison of the radial metallicity gradient of the thin disk to that of the thick disk, as well as to distinguish between *radial* and *vertical* trends. We note that *all* of the distant clusters that have been used to study the behavior of the *radial* metallicity gradient in the outer disk ($R > 15$ kpc) are located at least 1.5 kpc from the Galactic midplane; these are shown in gray in Figure 5.

4. CORRECTING FOR SELECTION BIASES: WEIGHTS

To use field stars to determine the metallicity distribution in the disk, we must understand how the spectroscopic sample is drawn from the underlying population. As there are many more stars than it was possible for SEGUE to obtain spectra, we must assess whether the spectroscopic sample is truly representative of all of the stars in the disk. It likely is not, because our stars are selected to be the bluer stars in the CMD, making our selection biased against metal-rich and older stars, which have redder colors. The severity of the metallicity bias will depend on the ages of metal-rich stars; the older and more metal rich they are, the more likely they are to fall out of our sample. To correct for this bias, we employ a weighting scheme to step backward in our sample selection and reconstruct the properties of the underlying parent population. We describe the scheme briefly below; further details are in the Appendix.

There are three major ways in which the spectroscopic sample is different from the parent population along each line of sight: (1) the photometric objects in regions with the highest extinction were not considered for spectroscopy. (2) Not all candidates for spectroscopy are observed. (3) We observe only MSTO stars using a color cut that is biased against redder metal-rich stars.

Table 4
Open Clusters with High-resolution Observations

Cluster	l	b	d (kpc)	R (kpc)	Z (kpc)	[Fe/H]	Age (Gyr)	Reference
Be17	176.0	−3.6	2.7	10.7	−0.17	−0.15	10.1	Friel et al. (2005)
Be20	203.5	−17.3	8.6	15.9	−2.55	−0.49	4.1	Yong et al. (2005)
	203.5	−17.4	8.3	15.6	−2.49	−0.30	6.0	Sestito et al. (2008)
Be22	199.9	−8.1	6.0	13.7	−0.84	−0.32	3.3	Villanova et al. (2005)
Be25	226.6	−9.7	11.4	17.7	−1.92	−0.20	5.0	Carraro et al. (2007b)
Be29	197.9	8.0	14.8	22.4	+2.05	−0.54	4.3	Yong et al. (2005)
	198.0	8.0	13.2	20.8	+1.83	−0.31	4.0	Sestito et al. (2008)
	198.0	8.1	13.2	20.8	+1.85	−0.44	4.5	Carraro et al. (2004)
Be31	206.2	5.1	5.3	12.9	+0.47	−0.53	5.3	Yong et al. (2005)
Be32	208.0	4.4	3.1	10.9	+0.24	−0.29	5.5	Sestito et al. (2006)
	208.0	4.4	3.4	11.1	+0.26	−0.30	5.9	Friel et al. (2010)
Be39	223.5	10.1	4.3	11.4	+0.75	−0.21	7.0	Friel et al. (2010)
Be73	215.3	−9.4	9.8	16.8	−1.60	−0.22	1.5	Carraro et al. (2007b)
Be75	234.3	−11.1	9.1	15.1	−1.76	−0.22	4.0	Carraro et al. (2007b)
Blanco1[a]	15.6	−79.3	0.3	8.0	−0.26	+0.04	0.1	Ford et al. (2005)
	15.6	−79.3	0.3	8.0	−0.26	+0.20	0.1	Edvardsson et al. (1995)
Cr261	301.7	−5.5	2.8	7.0	−0.26	−0.22	6.0	Friel et al. (2003)
	301.7	−5.5	2.8	7.0	−0.26	+0.13	6.0	Sestito et al. (2008)
	301.7	−5.5	2.8	7.0	−0.27	−0.03	6.0	Carretta et al. (2005)
	301.7	−5.5	2.8	7.0	−0.26	−0.01	6.0	De Silva et al. (2007)
M67	215.6	31.7	0.9	8.6	+0.45	+0.00	4.3	Santos et al. (2009)
	215.7	31.9	0.8	8.6	+0.42	+0.02	4.3	Yong et al. (2005)
	215.6	31.7	0.9	8.6	+0.45	+0.03	4.3	Friel et al. (2010)
	215.6	31.7	0.8	8.6	+0.44	−0.01	4.3	Jacobson et al. (2011b)
	215.6	31.7	0.9	8.6	+0.45	+0.02	4.3	Friel & Boesgaard (1992)
	215.6	31.7	0.9	8.6	+0.45	−0.03	4.3	Tautvaišiene et al. (2000)
	215.6	31.7	0.9	8.6	+0.45	+0.03	4.3	Randich et al. (2006)
	215.6	31.7	0.9	8.6	+0.45	+0.03	4.3	Pace et al. (2008)
Mel66	260.5	−14.2	4.4	9.6	−1.07	−0.38	4.0	Gratton & Contarini (1994)
	260.5	−14.2	4.4	9.6	−1.07	−0.33	4.0	Sestito et al. (2008)
Mel71[a]	228.9	4.5	3.2	10.3	+0.25	−0.30	0.2	Brown et al. (1996)
NGC 1193	146.8	−12.2	5.8	13.1	−1.23	−0.22	4.2	Friel et al. (2010)
NGC 1245	146.6	−8.9	3.0	10.6	−0.46	−0.04	1.1	Jacobson et al. (2011b)
NGC 1817	186.1	−13.1	1.5	9.5	−0.34	−0.16	1.1	Jacobson et al. (2011b)
	186.1	−13.1	1.5	9.5	−0.34	−0.07	1.1	Jacobson et al. (2009)
NGC 188	122.8	22.5	1.7	8.9	+0.65	−0.03	6.3	Jacobson et al. (2011b)
	122.8	22.5	1.7	8.9	+0.65	+0.12	6.3	Friel et al. (2010)
	122.8	22.5	1.7	8.9	+0.65	+0.01	6.3	Randich et al. (2003)
NGC 1883	163.1	6.2	3.9	11.8	+0.42	−0.20	0.7	Villanova et al. (2007)
	163.1	6.2	3.9	11.8	+0.42	−0.01	0.7	Jacobson et al. (2009)
NGC 1901[a]	279.0	−33.6	0.4	8.0	−0.23	−0.08	0.8	Carraro et al. (2007a)
NGC 2112[a]	205.9	−12.6	0.9	8.8	−0.19	−0.09	2.0	Brown et al. (1996)
	205.9	−12.6	0.9	8.8	−0.19	+0.16	2.0	Carraro et al. (2008)
NGC 2141	198.1	−5.8	3.9	11.8	−0.39	−0.14	2.5	Yong et al. (2005)
	198.0	−5.8	3.9	11.8	−0.39	+0.00	2.4	Jacobson et al. (2009)
NGC 2158	186.6	1.8	4.0	12.0	+0.13	−0.28	1.9	Jacobson et al. (2011b)
	186.6	1.8	4.0	12.0	+0.13	−0.03	1.9	Jacobson et al. (2009)
NGC 2194	197.3	−2.3	1.9	9.8	−0.08	−0.08	0.9	Jacobson et al. (2011b)
NGC 2204	226.0	−16.2	4.1	11.1	−1.14	−0.23	2.0	Jacobson et al. (2011a)
NGC 2243	239.5	−18.0	3.6	10.2	−1.11	−0.42	4.7	Jacobson et al. (2011a)
	239.5	−18.0	3.6	10.2	−1.11	−0.48	4.7	Gratton & Contarini (1994)
NGC 2324	213.4	3.3	4.2	11.7	+0.24	−0.17	0.6	Bragaglia et al. (2008)
NGC 2355	203.4	11.8	1.9	9.7	+0.39	−0.08	0.8	Jacobson et al. (2011b)
NGC 2420	198.1	19.6	2.5	10.3	+0.84	−0.57	2.2	Jacobson et al. (2011b)
	198.1	19.6	2.5	10.3	+0.84	−0.20	2.2	Jacobson et al. (2011b)
NGC 2425	231.5	3.3	3.3	10.4	+0.19	−0.15	2.5	Jacobson et al. (2011b)
NGC 2477	253.6	−5.8	1.2	8.4	−0.12	+0.07	1.0	Bragaglia et al. (2008)
NGC 2506	230.6	9.9	3.3	10.4	+0.57	−0.24	1.7	Mikolaitis et al. (2011)
	230.6	9.9	3.3	10.4	+0.57	−0.20	1.7	Carretta et al. (2004)
NGC 2539[a]	233.7	11.1	1.4	8.9	+0.26	+0.13	0.4	Santos et al. (2009)
NGC 2660	265.9	−3.0	2.8	8.6	−0.14	+0.04	1.0	Sestito et al. (2006)
NGC 3680[a]	286.8	16.9	0.9	7.8	+0.27	−0.04	1.2	Pace et al. (2008)
	286.8	16.9	0.9	7.8	+0.27	−0.03	1.2	Santos et al. (2009)
	286.8	16.9	0.9	7.8	+0.27	+0.04	1.2	Smiljanic et al. (2009)
NGC 3960	294.4	6.2	2.1	7.4	+0.22	+0.02	0.9	Sestito et al. (2006)

Table 4
(Continued)

Cluster	l	b	d (kpc)	R (kpc)	Z (kpc)	[Fe/H]	Age (Gyr)	Reference
NGC 6253	335.5	−6.2	1.6	6.6	−0.17	+0.46	3.0	Carretta et al. (2007)
	335.5	−6.3	1.6	6.6	−0.17	+0.36	3.0	Sestito et al. (2007)
NGC 6791[a]	70.0	10.9	5.9	8.1	+1.11	+0.37	4.4	Peterson & Green (1998)
	70.0	10.9	5.9	8.1	+1.11	+0.30	4.4	Boesgaard et al. (2009)
	70.0	10.9	5.9	8.1	+1.11	+0.47	4.4	Carretta et al. (2007)
	70.0	10.9	5.9	8.1	+1.11	+0.38	4.4	Carraro et al. (2006)
	70.0	10.9	5.9	8.1	+1.11	+0.35	4.4	Origlia et al. (2006)
NGC 6819	74.0	8.5	8.2	9.7	+1.21	+0.09	2.7	Bragaglia et al. (2001)
NGC 6939[a]	95.9	12.3	1.2	8.2	+0.25	+0.00	2.2	Jacobson et al. (2007)
NGC 7142	105.0	9.0	3.2	9.3	+0.50	+0.13	4.0	Jacobson et al. (2008)
	105.0	9.0	3.2	9.3	+0.50	+0.08	4.0	Jacobson et al. (2007)
NGC 752[a]	137.1	−23.2	0.5	8.3	−0.18	+0.01	1.1	Sestito et al. (2004)
	137.1	−23.2	0.5	8.3	−0.18	−0.09	1.1	Hobbs & Thorburn (1992)
NGC 7789	115.5	−5.4	2.2	9.2	−0.21	−0.04	1.8	Jacobson et al. (2011b)
	115.5	−5.4	2.2	9.2	−0.21	+0.02	1.8	Jacobson et al. (2011b)
Ru4	222.0	−5.3	4.7	11.9	−0.43	−0.09	0.8	Carraro et al. (2007b)
Ru7	225.4	−4.6	6.0	12.9	−0.48	−0.26	0.8	Carraro et al. (2007b)
Sau1	214.7	7.4	13.2	20.2	+1.70	−0.38	5.0	Carraro et al. (2004)
To2	232.0	−6.9	8.3	14.6	−0.99	−0.50	2.2	Brown et al. (1996)
	232.0	−6.9	8.3	14.6	−0.99	−0.28	2.2	Frinchaboy et al. (2008)
	232.0	−6.9	8.3	14.6	−0.99	−0.31	2.2	Villanova et al. (2010)

Note. [a] Spatial and age information from WEBDA.
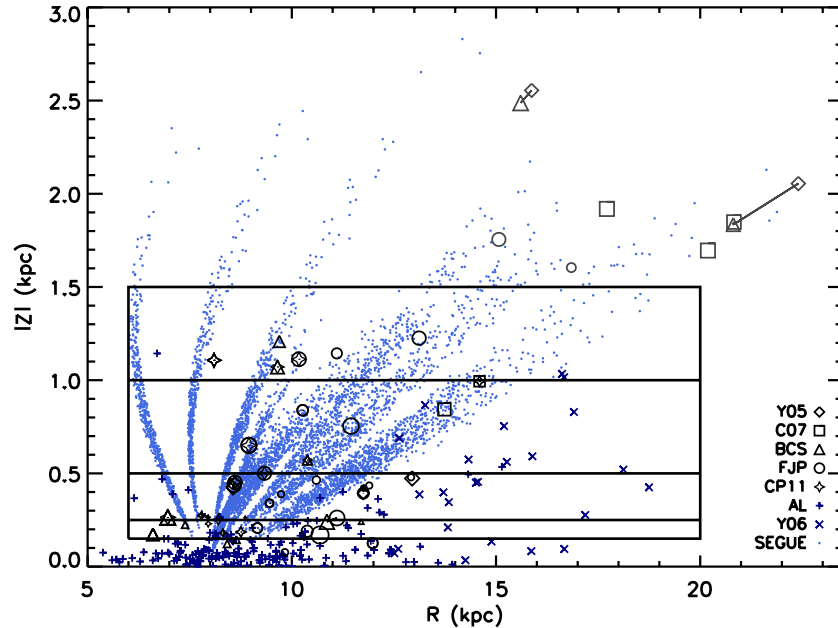
Figure 5. Spatial distribution of our sample (blue dots) in Galactic coordinates R and $|Z|$. We divide our sample into four $|Z|$ slices, indicated by the black boxes. Open cluster and Cepheid data are plotted with open symbols and pluses/crosses, respectively (see Section 7.2 for details). Objects from the literature located at $|Z| > 1.5$ kpc are plotted in gray. We note that all of the outer disk clusters from the literature are located at large $|Z|$. Multiple literature values for a single cluster are connected by lines.

(A color version of this figure is available in the online journal.)

Each star in our sample is given three weights corresponding to the three differences listed above: (1) the *area weight*, W_A, which depends on the coverage of targets on the plane of the sky in each line of sight; (2) the *CMD weight*, W_{CMD}, which depends on the target's location in the CMD and corrects for the random selection of a subsample of all candidates for spectroscopy that pass the MSTO selection; and (3) the *LF weight*, W_{LF}, which depends on the target's T_{eff}, [Fe/H], and location in the CMD, and corrects for the metallicity bias of the MSTO selection. The

total weight, W, is the product of the three weights W_A, W_{CMD}, and W_{LF}. After removing targets with $W = 0$, we are left with a sample of 7010 stars. Details about how each of these weights is calculated can be found in the Appendix.

Figure 6 shows the distribution of the total weight (black) and each individual weight, as a function of R, in four slices of $|Z|$. The panels on the right show the distribution of weights. W_A (green) is the smallest contribution and does not vary significantly. W_{LF} (red) is relatively constant as a function of
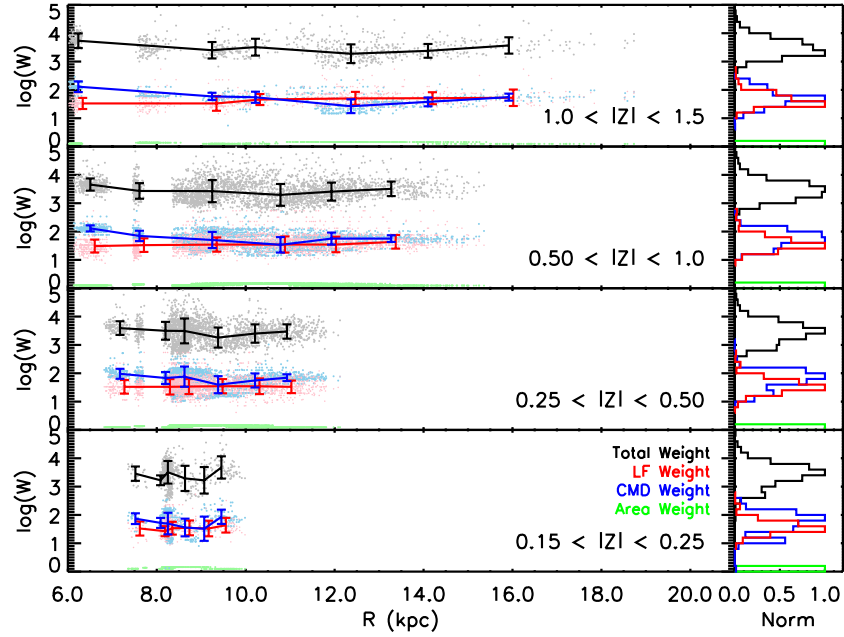
Figure 6. Variation in weights as a function of R and $|Z|$. The variation in the total weight W (black) is mostly dependent on the CMD weight W_{CMD} (blue). The LF weight W_{LF} (red) is fairly constant at all locations, with the width of the distribution being ~ 0.5–1.0 dex. The area weight W_A (green) is the smallest contribution and does not vary greatly between different lines of sight. While W_{CMD} shows the most dramatic variation, it is also less uncertain than W_{LF} because it only requires counting objects in CMD bins.
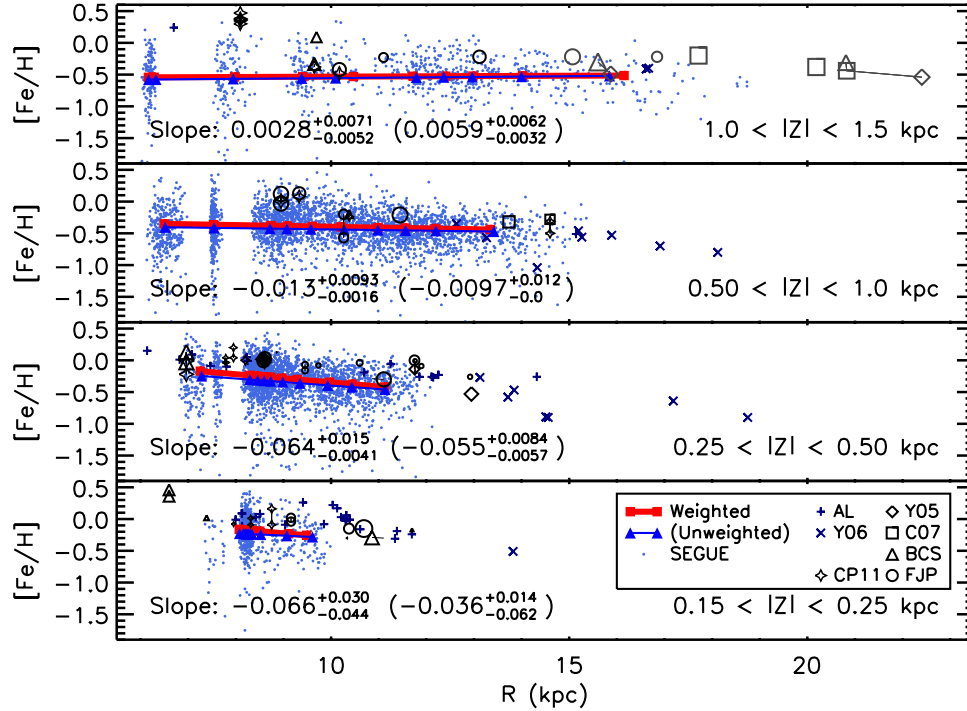
Figure 7. Metallicity [Fe/H] vs. Galactocentric radius R in four $|Z|$ slices. Light blue points indicate the SEGUE data. The weighted median metallicity and the derived linear fit are shown as red squares, with the numerical values in the bottom left of each panel. The blue triangles and values in parentheses show the results we would have obtained if no corrections for known selection effects had been applied. The spacing of the symbols indicates the radial distribution of the targets. Open symbols and pluses/crosses are open clusters and Cepheids from the literature (see Section 7.2 for details). The sizes of the open cluster symbols indicate their ages (smaller symbols for younger clusters). At low $|Z|$ (< 0.5 kpc), our derived gradient is consistent with published values. At high $|Z|$ (> 0.5 kpc), the constant [Fe/H] is consistent with the cluster metallicities reported by Yong et al. (2005) in the outer disk.

R and $|Z|$, with the distribution having a width of ~ 0.5–1.0 dex. Since W_{LF} is relatively constant, a systematic error in W_{LF}, which could arise from using the wrong luminosity function, will not cause a spurious change in the ratio of metal-poor to metal-rich stars.

W_{CMD} (blue) shows the most variation because it normalizes for the fact that there are more stars in the inner disk than the outer disk. Although the variations in W_{CMD} are large, this weight is less uncertain than W_{LF}, as it only requires counting objects in each bin. Figure 6 shows that
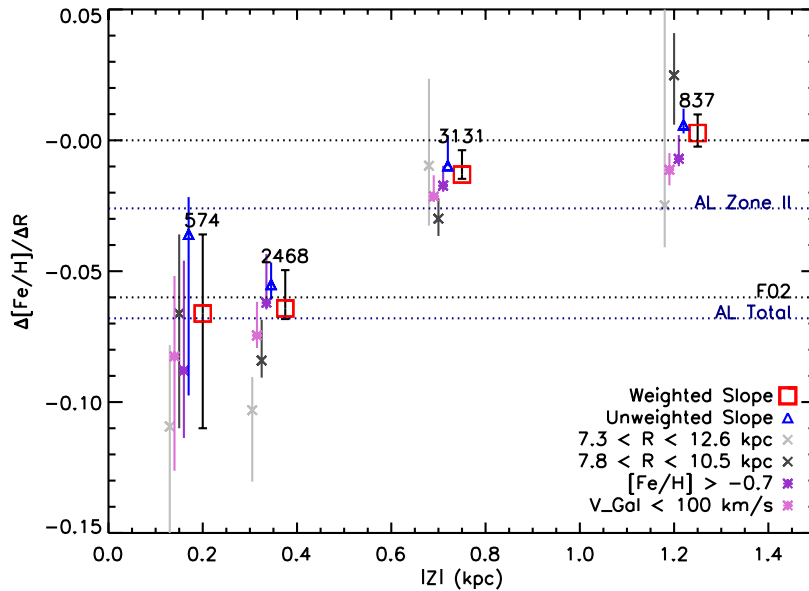
Figure 8. Measured gradient, $\Delta[\mathrm{Fe/H}]/\Delta R$, vs. vertical height, $|Z|$. The horizontal black dotted lines indicate a flat gradient and the gradient measurement published by Friel et al. (2002) using open clusters. The horizontal blue dotted lines indicate the gradient measurements published by Luck et al. (2006) using Cepheids for the ranges $6.6\,\mathrm{kpc} < R < 10.6\,\mathrm{kpc}$ (zone II) and $4.0\,\mathrm{kpc} < R < 14.6\,\mathrm{kpc}$ (their total sample). The red squares and blue triangles show the measured gradients (weighted and unweighted, respectively) as a function of height above the Galactic plane presented in this work. The slope becomes flat at high $|Z|$ using both the weighted and unweighted results. The number of objects in each $|Z|$ slice is indicated above the symbols. The gray symbols show the results obtained for restricted ranges in R. The purple symbols show the results obtained when probable halo contaminants are removed (see Section 6.2). In each case, the trend of the gradient becoming more shallow with distance from the plane is still observed, which indicates that our result does not arise due to the increase in the radial range or contamination from the halo.

(A color version of this figure is available in the online journal.)

the change in the total weight W mostly follows the change in W_{CMD}.

5. RESULTS: RADIAL METALLICITY GRADIENTS

The total weights are applied to each target, allowing us to use our sample of MSTO stars to estimate the properties of the underlying parent population. We divide our sample into four slices of $|Z|$: (1) $0.15\,\mathrm{kpc} < |Z| < 0.25\,\mathrm{kpc}$, (2) $0.25\,\mathrm{kpc} < |Z| < 0.5\,\mathrm{kpc}$, (3) $0.5\,\mathrm{kpc} < |Z| < 1.0\,\mathrm{kpc}$, and (4) $1.0\,\mathrm{kpc} < |Z| < 1.5\,\mathrm{kpc}$. Within each $|Z|$ slice, we fit a linear gradient to the data, weighting each target by the total weight determined using the scheme described in Section 4.

Figure 7 shows the radial metallicity gradients for all four $|Z|$ slices for both the unweighted and weighted cases (blue triangles and red squares, respectively). The weighted slopes are indicated in the bottom left corner of each panel, with the unweighted values in parentheses. The quoted errors are derived from the Monte Carlo simulations described in Section 6.3.2 and include only the random errors from uncertainties in the stellar parameters. See Section 6.3.2 for a discussion of the systematic errors. The large symbols and navy blue pluses/crosses show the positions and metallicities for open clusters and Cepheids from the literature (see Section 7.2).

In the low $|Z|$ slices ($|Z| < 0.5\,\mathrm{kpc}$), we obtain values that are consistent with the -0.06 dex kpc^{-1} determined by Friel et al. (2002) for open clusters and by Luck et al. (2006) for Cepheids. Our major result is that in the high $|Z|$ slices ($|Z| > 0.5$ kpc), the slope is flat for the entire radial range $6\,\mathrm{kpc} < R < 16\,\mathrm{kpc}$. The constant median [Fe/H] in the highest slice ($|Z| > 1.0\,\mathrm{kpc}$) is consistent with the metallicities reported by Yong et al. (2005) and others at large R.

Our results are summarized in Figure 8, which shows the measured slopes as a function of vertical height $|Z|$. The horizontal dotted lines show the values in the literature published

by Friel et al. (2002) and Luck et al. (2006) for open clusters and Cepheids, respectively, which are consistent with the values we obtained for the low $|Z|$ slices. Again, the red squares show the values obtained after we apply our weights, and the blue triangles show the values obtained using the unweighted data without any corrections for the selection function. The observed trend—flattening slope with increasing $|Z|$—is seen in both the unweighted and weighted cases.

To test the robustness of this result, we re-measured the gradient for various subsamples of our data. The purple symbols in Figure 8 show the results if we exclude likely halo stars (see Section 6.2). The gray symbols show the values obtained for restricted ranges in R ($7.3\,\mathrm{kpc} < R < 12.6$ kpc and $7.8\,\mathrm{kpc} < R < 10.5$ kpc). These measurements were done to test that the flattening of the radial gradient is truly a trend in $|Z|$, and not just a result of the wider range in R that is observed in the high $|Z|$ bins. The two restricted ranges in R correspond to the extent of observations in the two lower $|Z|$ bins. Figure 8 shows that gradient becomes flat with or without weighting and independent of the range in R observed.

6. ERRORS

6.1. Metallicity Bias

As described in Section 2.1, our sample of MSTO stars is chosen by making a cut in $(g - r)_{\mathrm{SFD}}$, which likely biases it against redder, more metal-rich stars. Theoretical isochrones give us an idea of the significance of this bias. Figure 9 shows the temperature and $(g - r)_{\mathrm{SFD}}$ distributions of metal-rich stars in our sample ([Fe/H] $\geqslant +0.2$, gray circles) against the theoretical isochrones of An et al. (2009, color). We estimate that the maximum observable ages of stars at [Fe/H] $= +0.4$ and $+0.2$ are ~ 8 and 11 Gyr, respectively. At each metallicity, this corresponds to the turnoff age of the
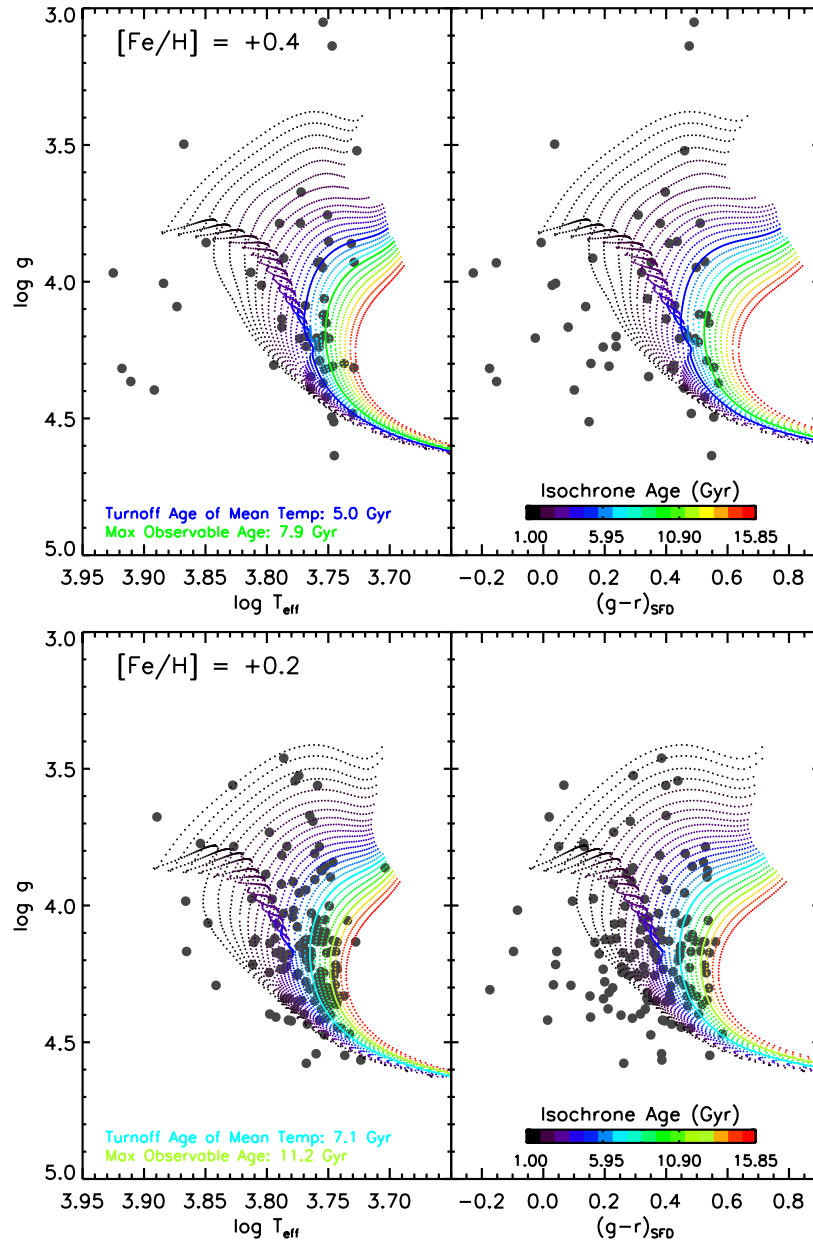
Figure 9. Maximum observable ages for metal-rich stars. We compare the SEGUE data (gray circles) with theoretical isochrones (color; An et al. 2009) to estimate the oldest main-sequence turnoff stars that would fall in our sample. The TAMT (see Table 3) and maximum observable age are shown as solid lines. For [Fe/H] = +0.4 (left panel), the coolest/reddest stars may be as old as ∼8 Gyr; for [Fe/H] = +0.2 (right panel), the coolest/reddest stars that fall in our sample may be as old as ∼11 Gyr. Based on the measured ages of nearby stars (Bensby et al. 2005; Casagrande et al. 2011), we do not expect there to be a significant population of old, metal-rich stars which would be excluded from our sample.

coolest/reddest stars that pass the selection criteria for our sample. The TAMT and maximum observable age are shown as solid lines. Based on the measured ages of nearby stars (Bensby et al. 2005; Casagrande et al. 2011), very few stars with metallicities [Fe/H] > +0.2 are older than 8 Gyr. We therefore expect that metal-rich stars are well represented in our sample. For stars with metallicities [Fe/H] ⩽ 0.0, there is no significant bias against stars of any age. We do the same test using isochrones from both Dotter et al. (2008) and the Padova[16] group (Girardi et al. 2004), and we find that the temperature of the turnoff for a given age does not vary significantly between the different stellar evolution models, so the above conclusions about the metallicity bias are robust.

The arguments presented above suggest that the metallicity bias resulting from our color selection does not eliminate a significant fraction of the disk population from our sample. However, the color selection also causes metal-rich stars to be underrepresented in our data. We use mock catalogs from the models of Schönrich & Binney (2009a) to quantify how well we correct for this effect and recover the true metallicity distribution in our sample. We show below (Section 6.3.2) that we are able to measure the true gradients within the errors and reproduce the trends as a function |Z| when we apply the weighting scheme described in Section 4 and the Appendix.

6.2. Halo Contamination

Since we are interested in the metallicity distribution of the Galactic disk, we must assess whether our sample may

---

[16] http://stev.oapd.inaf.it/cmd

be contaminated by halo objects. Our sample does not reach $R < 5$ kpc, where the bulge is expected to be a significant population. One way to quantify the amount of contamination is by examining different multi-component models for the Galaxy, in particular, those of Jurić et al. (2008) and de Jong et al. (2010). For the lines of sight in our sample, these two models predict total halo contaminations of ∼2% and ∼0.8%, respectively. In the most distant $|Z|$ bin, the predicted contaminations increase to ∼11% and ∼5.6%, respectively. The difference in the predictions can be almost entirely attributed to the discrepant local densities for the halo that were derived (0.51% by Jurić et al. 2008 and 0.17% by de Jong et al. 2010).

Both models, however, predict that the anticenter lines of sight, aimed toward the outer disk, will be more contaminated by halo stars. This is especially true in the Jurić et al. (2008) model, which predicts that the two lines of sight that reach the largest values of R will have 17%–18% contamination (compared to 9%–13% in the other lines of sight). If there is indeed more halo contamination in the outer regions, this should push our measured gradient to be steeper, since there will be more metal-poor halo stars at large R compared to the inner disk. However, we see the opposite effect that the gradient is flatter in the highest $|Z|$ bins.

Looking at the data, we do not see evidence for significant halo contamination. Halo stars, which have large velocities and are more metal poor than disk stars, can generally be identified using their kinematics or chemistry. To assess the effect that such stars may have on our results, we recalculate the gradient after applying two different cuts to remove potential halo stars from the data: (1) a metallicity cut that removes all stars with [Fe/H] < -0.7, and (2) a kinematic cut that removes all stars with $V_{Gal} < 100$ km s^{-1} to remove stars with the largest velocity offset relative to the projection of the local standard of rest, where $V_{Gal} = V_R + 220 \cdot \cos b \cdot \sin l$ and V_R is the line-of-sight velocity measured from the SEGUE spectra. We only remove stars with $V_{Gal} < 100$ km s^{-1} along lines of sight with $50° < l < 130°$. We do not include the lines of sight directed toward the Galactic anticenter because the local standard of rest is tangent to those directions and the projection does not give a meaningful velocity. The two cuts remove 963 and 423 stars, respectively. The gradients we measure using each of these cuts are not significantly different from our main results, as shown by the purple symbols in Figure 8.

6.3. Mock Catalog Analysis

To quantify the errors in our analysis, we utilize mock observations generated from the models of Schönrich & Binney (2009a). By using a model where we know the true stellar parameters, distances, and metallicity distributions of the targets, we can test whether we are able to reliably reproduce the ground truth after applying our observational selection to the mock catalogs and performing the same analysis procedures. As the purpose is merely to assess how accurately we can measure quantities such as distances and metallicity gradients, this way of testing our methods is not dependent on having a *correct* model for the Galaxy. We do, however, need a model that can reproduce some basic properties of the observed stellar populations. Schönrich & Binney (2009a) have shown that their model provides a good match to the properties of stars in the solar neighborhood as observed by the Geneva–Copenhagen Survey (Nordström et al. 2004).

We use a mock catalog, provided by R. Schönrich, of 6,701,170 stars with *ugriz* colors, stellar parameters, and dis-
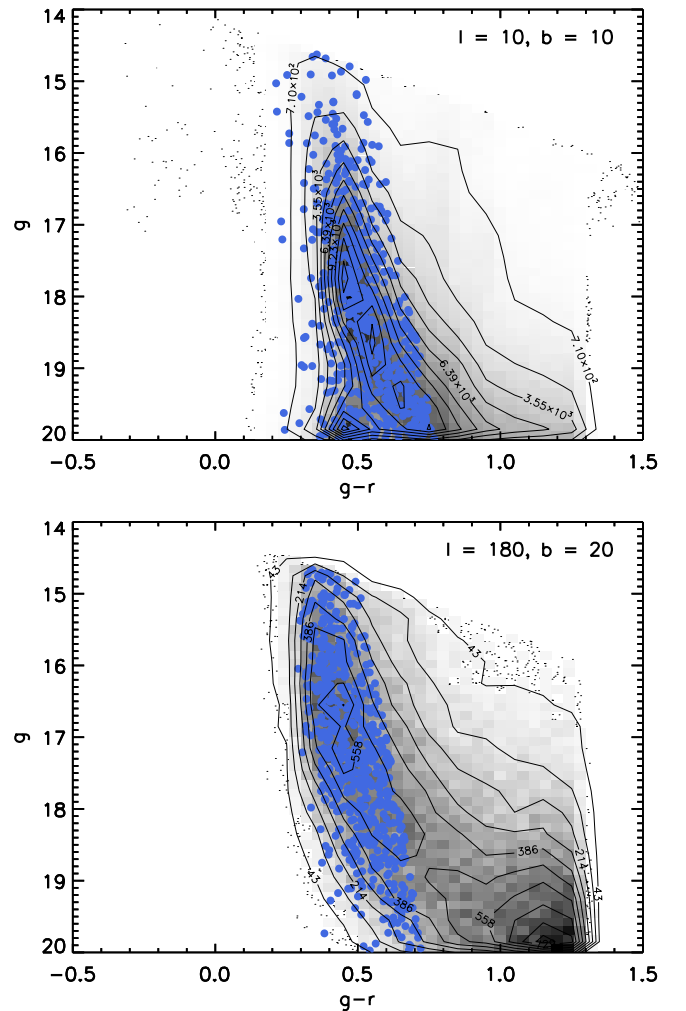

Figure 10. Mock catalog MSTO selection. This figure is analogous to Figure 1, but using simulated observations from Schönrich & Binney (2009a). The full sample of stars (gray scale and contours) is put through the same sample selection as the real data to yield an analogous MSTO sample (blue circles). The similarity between the CMDs of the real and mock data shows that we are using a reasonable model of the Galaxy to test our errors.

tances along 10 lines of sight. Targets in the mock catalogs are assigned stellar parameters using the BaSTI[17] isochrones. We repeat the same sample selection described in Section 2.1 on the mock catalog to replicate the effects of using MSTO stars as tracers. Figure 10 is analogous to Figure 1 and shows the results of the random subsampling for two lines of sight. Though the model was not tuned to look like the SEGUE data, the model CMDs are good matches to the observations.

In our analysis, we use all 111,640 objects that fulfill the MSTO selection to estimate errors on the distance (Section 6.3.1). We draw many random subsamples of 6500 MSTO stars to estimate the systematic and random errors on our gradient measurement (Section 6.3.2), to simulate the effect of having a limited number of spectroscopic targets along each line of sight. We repeat the analysis using different random subsamples to assess how much the results change based on which particular targets are chosen for spectroscopy.

6.3.1. Errors in Distance Estimates

Systematic errors. The main source of systematic error in the distance arises from assumptions we must make when
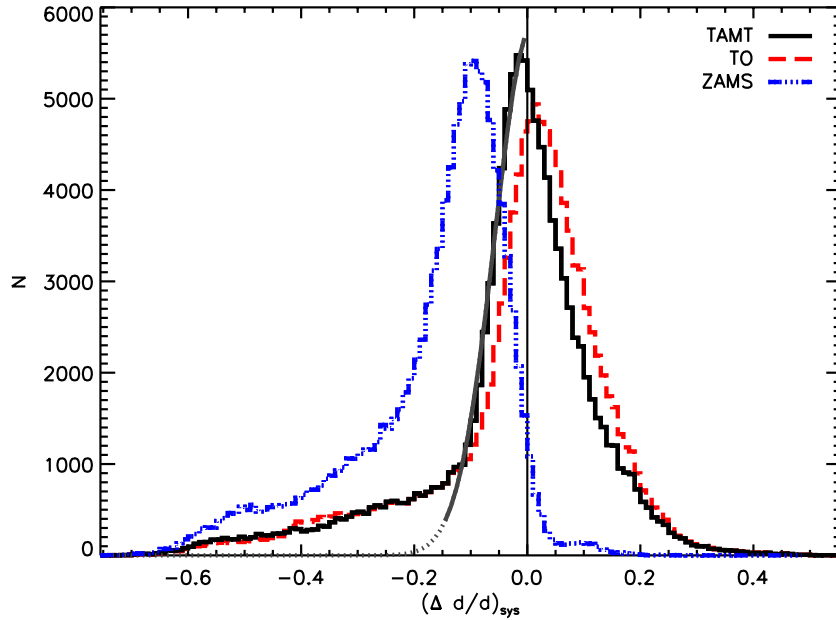

Figure 11. Fractional error distributions of solar metallicity stars ($-0.05 < $ [Fe/H] $ < 0.05$) for three different age assumptions, as described in the text (see also Figure 3). The TAMT age assumption gives the best agreement between the calculated and true model distances. Based on the width of the distribution, we estimate the systematic error on our distance estimates to be $\sim$10%. The vertical line indicates where the fractional error is zero. The gray line is a Gaussian fit to the TAMT distribution where $-0.15 < (\Delta d/d)_{\rm sys} < 0.0$, which we use to estimate the fraction of subgiants in our sample ($\sim$15%).

(A color version of this figure is available in the online journal.)

choosing isochrones to estimate the luminosities of our stars. One source of error is the α-enhancement; An et al. (2009) adopt an α-enhancement scheme where each value of [Fe/H] has an assumed value of [α/Fe] (see Table 2). We do not expect the discrepancy between the target and isochrone [α/Fe] to have a large effect on the distance estimate. The [α/Fe] of stars in our sample (Lee et al. 2011a) are generally within 0.2 dex of the values assigned to the An et al. (2009) isochrones. Using the [Fe/H]–[α/Fe] grid of Dotter et al. (2008), we estimate that a 0.2 dex change in [α/Fe], at worst, leads to a $\sim$10% change in the distance for a star on the zero-age main sequence.

For our sample, a larger source of uncertainty is the ages we assign to our targets; the slope of the main sequence becomes very steep at the turnoff, making the predicted distance a strong function of the chosen age. We compare the distance estimates obtained using three different age assumptions, which are shown schematically for stars at solar metallicity ($-0.05 < $ [Fe/H] $ < 0.05$) in Figure 3.

1. *Turnoff age of the mean temperature (TAMT).* We use one isochrone for most of the stars in the metallicity bin—this is determined by finding the mean of the temperature distribution. Targets hotter than the MSTO are assigned to the oldest possible isochrone. This was the scheme used throughout our analysis (described in Section 3) and illustrated in the second panel of Figure 3.

2. *Turnoff (TO).* We use the oldest possible isochrone consistent with their measured temperature for all targets—this assumes that every star is located at the turnoff. In comparison to the TAMT assumption, this approach changes the distances for stars cooler than the mean temperature of the sample, as shown in the third panel of Figure 3.

3. *Zero-age main sequence (ZAMS).* We use the youngest possible isochrone for all targets—this assumes that most targets are located on the zero-age main sequence, as shown in the fourth panel of Figure 3.

To assess the effect of the age assumption alone, we compare our calculated distance, d, to the true distance from the model, $d_{\rm model}$, for 111,640 mock targets at solar metallicity ($-0.05 < $ [Fe/H] $ < 0.05$), where theoretical isochrones from different groups show the best agreement. Any disagreement between the isochrones used to generate the mock catalog and the isochrones used to calculate the distances will introduce an additional systematic error, which we discuss at the end of this section. The distributions of the systematic fractional error, $(\Delta d/d)_{\rm sys}$, for all three age assumptions are shown in Figure 11, where the systematic fractional error is given by

$$\left(\frac{\Delta d}{d}\right)_{\rm sys} = \frac{d - d_{\rm model}}{d_{\rm model}}. \tag{1}$$

Based on the offsets of the distributions, the TO (red dashed line) and ZAMS (blue dash-dotted line) assumptions overestimate and underestimate the distances, respectively, though the ZAMS assumption does much worse. The TAMT (black solid line) assumption gives the best agreement between the calculated and model distances, with an error of $\sim$10%, based on the width of the $(\Delta d/d)_{\rm sys}$ distribution. Mock targets at other metallicities exhibit the same behavior, with the TAMT assumption giving the best distances with a mean error of $\sim$10%. This error includes the discrepancy between the isochrone [α/Fe] and that of individual stars, as the model stars have a range of [α/Fe] that do not exactly match the sequence of the An et al. (2009) isochrones.

Using the $(\Delta d/d)_{\rm sys}$ distribution, we can also estimate the number of subgiants in our sample by investigating the long tail of negative fractional errors. The gray line in Figure 11 shows a Gaussian fit to the TAMT histogram where $-0.15 < (\Delta d/d)_{\rm sys} < 0.0$—this range was chosen by eye to reflect the range where subgiants were not contributing to the distribution. By examining the discrepancy between the Gaussian fit and the long tail of negative $(\Delta d/d)_{\rm sys}$ values, we estimate the
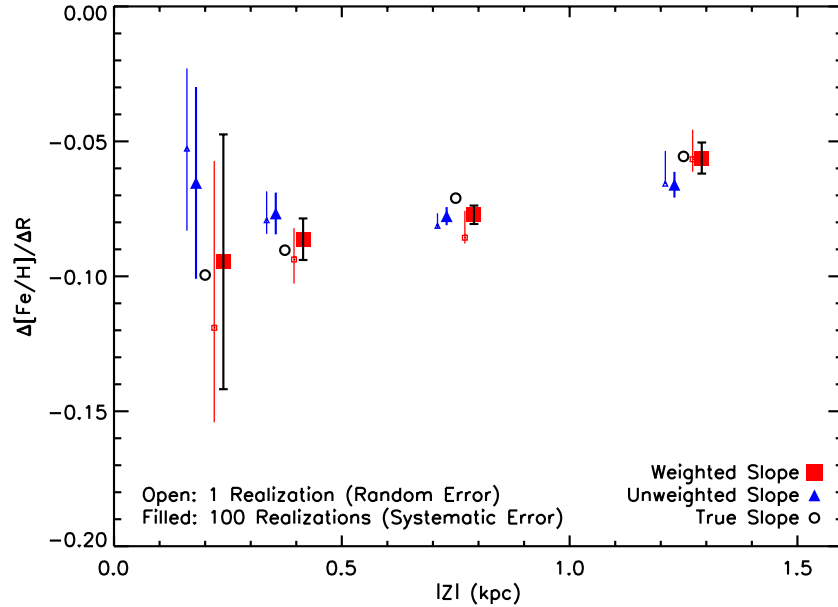
Figure 12. Measured gradient, $\Delta[Fe/H]/\Delta R$, vs. vertical height, $|Z|$, using mock samples. This figure is analogous to Figure 8, except the samples used are drawn from mock catalogs of the galaxy model of Schönrich & Binney (2009a). The mean and standard deviation for 100 such samples of MSTO stars are shown as filled symbols. The open symbols show the measured gradients for one particular random sample, with the error bars showing the random errors derived from Monte Carlo realizations using perturbed stellar parameters. The black circles show the true gradient, which is measured using all of the targets in the catalog. The true gradient generally falls within the errors presented for the weighted gradient (red squares) measured using the MSTO samples, which indicates that our method of measuring the gradient gives a reliable result. More importantly, we are able to reproduce the flattening trend seen in the true gradient. This is not true for the unweighted gradient measurements (blue triangles). Note that the vertical scale is different from that of Figure 8.

(A color version of this figure is available in the online journal.)

contamination by subgiants to be ~15%; we obtain the same result if we take subgiants to be all stars with $(\Delta d/d)_{\rm sys} \leqslant -0.15$. We derive similar values for other metallicities as well.

We can also assess how an error in the value of the TAMT affects the distances that we obtain in our calculation. If we view the TO assumption as the TAMT assumption with an incorrect TAMT—one that is too old by 0.2 dex—we can estimate how much the distances change as a result. Comparing the TAMT and TO results in Figure 11, the peak $(\Delta d/d)_{\rm sys}$ values of the two assumptions are within 5%, which suggests that any error in the TAMT is smaller than the error from the age assumption scheme used.

The above analysis can be applied to stars at all metallicities. At each metallicity, however, the peaks of the distributions are shifted up to ~10% because of differences between the stellar evolutionary models used (i.e., BaSTI and An et al. 2009 isochrones). This has the effect of increasing the width of the $(\Delta d/d)_{\rm sys}$ distribution to ~15% for the total sample, compared to ~10% for the solar metallicity sample. The peak of the distribution is at ~ − 8%. The results are similar when using the Dotter et al. (2008) isochrones: at a given metallicity, the peak of the distribution is shifted up to ~10%. For the total sample, the width is ~15% and the peak of the distribution is at ~ − 3%.

Random errors. The random error in the distances is dominated by uncertainties in the stellar parameters derived by the SSPP; errors in the magnitudes are trivial in comparison. We examine how the parameter errors propagate through our analysis by generating 500 Monte Carlo realizations of the mock catalog, with slightly perturbed values of the stellar parameters, and repeating the same distance measurements each time. For each realization we assign a Gaussian distribution of perturbations with widths of 200 K and 0.3 dex for $T_{\rm eff}$ and [Fe/H], respectively. These values are motivated by the errors estimated by Lee et al. (2008a). The resulting random fractional error,

$(\Delta d/d)_{\rm rand}$, is given by

$$\left(\frac{\Delta d}{d}\right)_{\rm rand} = \frac{d_{\rm perturbed} - d}{d}, \qquad (2)$$

where $d_{\rm perturbed}$ is the distance calculated with the new perturbed values of the stellar parameters. From the distribution of $(\Delta d/d)_{\rm rand}$, we estimate that the uncertainty in the SSPP parameters translates to a ~15%–20% error in the distance. These values correspond to the 68% confidence levels derived from the width of the distribution.

We can now calculate the total error in the distance, which combines the systematic and random errors from the age assumption and the stellar parameters, respectively. Because we include stars at all metallicities, we are also accounting for the uncertainties from discrepancies between the BaSTI and An et al. (2009) isochrones. The result is the total fractional error $(\Delta d/d)_{\rm total}$, given by

$$\left(\frac{\Delta d}{d}\right)_{\rm total} = \frac{d_{\rm perturbed} - d_{\rm model}}{d_{\rm model}}. \qquad (3)$$

We estimate that, including both the systematic and random errors, we have a total distance error of ~20%–25%, with a larger contribution from the uncertainty in the stellar parameters determined by the SSPP. This total error is comparable to the size of the errors in the distances estimated by Gilmore et al. (1995) for their sample of MSTO stars.

6.3.2. Errors in the Metallicity Gradient Measurement

Systematic errors. Using the same mock catalog, we also assess how well we measure the radial metallicity gradients in the disk. Figure 12 is the same as Figure 8, but shows the gradient results for the mock catalogs in four $|Z|$ slices. The true

gradients, measured using the entire mock catalog, are shown as black circles. We draw 100 random samples (6500 MSTO stars each) from the mock catalog using the same selection criteria as the real data (Section 2.1) and follow the same analysis procedures, including accounting for the weights (Section 4 and the Appendix).

The filled symbols in Figure 12 show the mean gradient of the 100 realizations; the error bars indicate the standard deviation of the distribution. The smaller open symbols show the gradient measurements for one particular realization, with the error bar indicating the random error due to the uncertainty in the stellar parameters (see below).

The true gradients (open black circles) are generally within the errors of the weighted measurements (filled red squares), indicating that our measurements are reliable. More importantly, the weighted measurements show the same flattening trend with $|Z|$ as the true gradient. This is not seen in the unweighted gradient measurements (filled blue triangles). While the metallicity gradient in the model is steep compared to recent observations (e.g., Luck & Lambert 2011), our mock catalog analysis demonstrates that the weighting procedure is necessary to reproduce the flattening trend.

Random errors. The error bars on the open symbols represent the random errors, which are derived by running our full analysis on one sample of 6500 MSTO stars from the mock catalog, where we generate 500 Monte Carlo realizations of the mock data with perturbed stellar parameters. The error bars indicate the 68% confidence levels derived from the width of the distribution of slopes we obtain in the 500 realizations. This is the same procedure followed to determine the random distance errors in Section 6.3.1. The random errors are comparable to the systematic errors on the gradient, although in some cases they are larger. The large error bars for the lowest $|Z|$ bin is due to a combination of the smaller number of stars and the smaller range in R.

7. DISCUSSION

7.1. Comparison with Previous Studies

The median metallicity we find at $|Z| > 1$ kpc ([Fe/H] ~ -0.5) is consistent with the value published by Gilmore et al. (1995) for their stars located 1.5 kpc above the plane; the median $|Z|$ for our sample above 1 kpc is 1.24 kpc. Their sample also consisted of an in situ sample of F/G dwarfs, allowing for a direct comparison. This value is also in agreement with the mean of Soubiran et al. (2003, −0.48) and the mode of Katz et al. (2011, −0.52) at similar $|Z|$ heights. The G-dwarf sample of Lee et al. (2011b), also drawn from SEGUE, has median values of [Fe/H] within 0.1 dex of that of our sample as a function of $|Z|$.

Based on recent observations along two lines of sight, Katz et al. (2011) find that the similar metallicity distributions of their two fields at 2 kpc $< |Z| < 4$ kpc are suggestive of a weak or non-existent radial gradient, although they cannot unambiguously rule it out. Our results are also consistent with the lack of a radial metallicity gradient found by both Allende Prieto et al. (2006), using spectra of 12,483 F/G stars from the SDSS, and Jurić et al. (2008), using photometric metallicities for millions of stars in the SDSS. However, both studies examine a much higher $|Z|$ sample with a limited R range at the vertical heights that our sample covers. The present study is complementary to these two studies in that we observe lower Galactic latitudes, and we are able to make a direct comparison

of the thin disk and thick disk using the same homogeneous sample. Our results are based on a sample that has been carefully corrected for selection effects, and we use an improved version of the SSPP that is more accurate for metal-rich stars, which was not available for these previous analyses.

In their analysis of thick disk stars in the solar neighborhood, Bensby et al. (2003, 2005) find that abundance trends using O, Mg, and Fe do not vary as a function of R or $|Z_{max}|$ in the region 5 kpc $< R < 7$ kpc, 0 kpc $< |Z_{max}| < 1.1$ kpc, where $|Z_{max}|$ is the maximum vertical height reached by the calculated stellar orbit. These detailed abundances, together with our finding of a flat metallicity gradient at large $|Z|$, are suggestive of a chemically homogeneous thick disk, although more observational data are needed to confirm this.

7.2. Comparison with Open Clusters and Cepheids

In Figures 5 and 7 we show open cluster and Cepheid data from the literature. These two classes of objects have been studied extensively with high-resolution spectroscopy and also span a large range in age, giving them the power to probe the chemistry of the interstellar medium at different times. Cepheids typically have lifetimes on the order of ∼100 Myr and trace the present metallicity distribution. Open clusters can have ages anywhere from less than a Gyr to greater than 10 Gyr and have been used to examine the temporal evolution of the disk's metallicity gradient.

We compare our old disk stars to a sample of open clusters with abundance determinations from high-resolution spectra. We note in particular 39 open clusters from the work of four groups that have studied clusters at $R > 10$ kpc: 5 from Yong et al. (2005, Y05), 9 from Carraro et al. (Carraro et al. 2004, 2007b; Villanova et al. 2005; Frinchaboy et al. 2008: C07), 12 from Bragaglia, Carretta, Sestito, et al. (Bragaglia et al. 2001; Carretta et al. 2004, 2005, 2007; Sestito et al. 2006, 2007, 2008: BCS), and 19 from Friel, Jacobson, and Pilachowski (Friel et al. 2005, 2010; Jacobson et al. 2008, 2009, 2011a, 2011b: FJP). Additional clusters from the compilation of Carrera & Pancino (2011, CP11, see their Table 12) are also included; we only consider those spatially coincident with our sample, i.e., those that are at least 150 pc from the midplane. Where we did not already have spatial and age information, parameters were obtained using the WEBDA[18] open cluster database. Twenty-three clusters have measured abundances from more than one study, which provides an indication of the size of the uncertainties and systematics between groups. For clusters with more than one measurement, a line connects the symbols representing the different determinations.

Properties of the clusters are listed in Table 4. The open clusters have ages ranging from 0.6 to 10.1 Gyr; the median age is 2.0 Gyr, with only two clusters as old as 7 Gyr, which makes this cluster sample younger than the typical age of the stars in our sample (see Table 3). The ages of the clusters are indicated by the size of the symbols in Figures 5 and 7, with older clusters having larger symbols. It should be noted, however, that the cluster ages are derived using a variety of methods and are not on the same scale as those of the field stars considered in this study.

The Cepheid data are taken from Andrievsky et al. (2002a, 2002b, 2002c, 2004) and Luck et al. (2003, 2006) (AL), and Yong et al. (2006, Y06). Most of the Cepheids in these samples are too close to the midplane ($|Z| < 0.15$ kpc) to

---
[18] http://www.univie.ac.at/webda/

be directly compared to our sample. Those in the higher bins (0.5 kpc < |Z| < 1.0 kpc), which are mostly from the Yong et al. (2006) sample, tend to be at larger radii than our MSTO stars. For consistency we have re-calculated R, |Z| values using published l, b, and distances with $R_{GC,\odot} = 8.0$ kpc, where necessary.

Both Cepheids and open clusters appear to be slightly more metal rich than the median metallicities of the old disk stars. Systematic differences in [Fe/H] between different groups can be up to 0.2 dex, as shown by the clusters with multiple measurements (see Table 4 as well as the discussion in the Appendix of Friel et al. 2010). Given the good agreement between the SSPP and literature values, however, these systematic differences may not explain the discrepancy. As we do not see any obvious differences in metallicities between old and young clusters (large versus small open symbols), the cause of a possible discrepancy is unknown at this time. Understanding the differences between the different tracers will be crucial to understanding how the disk formed and its subsequent evolution.

At low |Z|, the slopes of the open clusters and field stars do not appear to be dramatically different, as expected given the good agreement between our measurement of the radial gradient in the low |Z| bins and the values of Friel et al. (2002) and Luck et al. (2006, dotted lines in Figure 8). However, at high |Z|, there are three clusters with metallicities ~0.5 dex higher than the median [Fe/H] for field stars at small R (< 10 kpc). While a steeper slope in the inner regions of the disk (-0.13 dex kpc^{-1} at R < 8 kpc) has been reported for both open clusters and Cepheids (e.g., Pedicelli et al. 2009; Magrini et al. 2010), we do not see a steeper metallicity gradient at R < 8 kpc in our sample of old disk stars.

The question of how the radial metallicity gradient may change over time is still debated in the literature, with conflicting reports of flattening (e.g., Friel et al. 2002; Chen et al. 2003; Maciel et al. 2005), steepening (e.g., Stanghellini & Haywood 2010), and constant slopes (e.g., Magrini et al. 2009) as a function of time. If the gradient is steeper in the inner disk, as has been reported, the flatness of the gradient for old disk stars, at face value, implies that the radial metallicity gradient in the inner disk has grown steeper with time. Radial migration complicates this issue greatly, however, as the movement of stars from their initial orbits may wash out a previously existing gradient (e.g., Roškar et al. 2008a). Based on the observations presented here, we are not able to draw any clear conclusions about the time evolution of the metallicity gradient in the inner disk.

The disagreement in abundances between young and old tracers is not present in the outer disk, where the median metallicity of the old disk stars ([Fe/H] ~ -0.5) is consistent with the metallicities reported by Yong et al. (2005) for outer disk open clusters. Furthermore, the median metallicity of our sample at |Z| > 1.0 kpc is constant at *all* values of R, which suggests that the flattening of the gradient is due to a trend in |Z|. Given the change in the radial range spanned in each |Z| bin, a superposition of the negative gradient at small |Z| with a flat gradient at large |Z| could result in an apparent discontinuity in the radial gradient.

This result suggests that when examining trends in the *radial* direction, it is important to take the *vertical* information into account. As shown in Figure 5, all of the outer disk clusters from the literature (R > 15 kpc) are located far from the Galactic plane; this has been noted in previous studies (e.g., Jacobson et al. 2011a; Carrera & Pancino 2011). In the present

work, we find that the mean metallicity of these clusters is consistent with field stars located at *all R* at similar vertical heights (|Z| > 1.0 kpc), although our field star sample extends only to $R \sim 15$ kpc, whereas the clusters reach radial distances of 20 kpc or more. Based on our observations of old disk stars, whether the reported discontinuity is truly a feature of the *radial* metallicity gradient is unclear.

Our sample does not reach to sufficiently large radii at low latitudes; with the limited range in R at low |Z|, we cannot confirm that the trend is purely vertical. If the radial gradients we measure at |Z| < 1.0 kpc extend to large R (>15 kpc) *for all* |Z|, then the discontinuity in the radial metallicity gradient of the disk actually reflects the change in slope at different |Z|. Friel et al. (2010) have stressed the need for more high-resolution data of open clusters to build up a homogeneous, statistical sample; our work shows that it is also necessary to fully sample a range of both R and |Z| to understand the metallicity trends in the disk. The large R, low |Z| region of the Galaxy will be probed by future surveys such as APOGEE (Eisenstein et al. 2011).

7.3. Implications for Thick Disk Formation

The lack of a radial metallicity gradient far from the Galactic plane provides an observational constraint that must be matched by any viable scenario for thick disk formation, such as the four described in Section 1. A flat gradient, as we measure, may come most naturally out of a turbulent disk at high redshift (scenario 3). If the thick disk formed rapidly at early times (e.g., Brook et al. 2004, 2005), then the thick disk would be chemically homogeneous, and the radial metallicity gradient far from the plane of the disk would be flat.

The observations can only be explained by current models of radial migration in isolated disks (e.g., Roškar et al. 2008a; Schönrich & Binney 2009a) if radial mixing is strong (scenario 4). These models show that dynamical interactions with spiral arms can move stars from their initial orbits and make the gradient shallow, but the mechanism must be efficient if it is the dominant mechanism for forming the thick disk. The degree of radial mixing may be increased through dynamical interactions with the Galactic bar (Friedli et al. 1994; Martin & Roy 1994; Minchev & Famaey 2010).

Another possible way to increase the amount of radial mixing in a disk is to place the disk within a cosmologically motivated accretion history (scenario 1). The simulations of Bird et al. (2012) show that stars can move far from their initial orbits in a disk that is bombarded by multiple minor mergers (e.g., Kazantzidis et al. 2008). Furthermore, the bombarded disk experiences more radial mixing at high |Z| than an identical disk in isolation. At the midplane, however, both scenarios show about the same amount of mixing. The increased radial mixing at high |Z| could be responsible for the lack of a radial gradient.

With our data alone, we cannot rule out the direct accretion of stars that originally formed in satellites that have been disrupted (scenario 2). To assess whether the thick disk exhibits the predicted clumpiness, we would need to investigate the azimuthal variation within our sample or at more detailed abundances with follow-up observations. But observations of the orbital properties of stars in SDSS and RAVE (Dierickx et al. 2010; Wilson et al. 2011) show that there are not enough high-eccentricity stars to match what is seen in the simulations of Abadi et al. (2003). Furthermore, recent simulations (Read et al. 2008; Villalobos & Helmi 2008) predict lower numbers of stars being directly accreted during minor mergers than the 2003 simulations.

Other recent work using kinematic constraints—correlations between rotational velocity V_ϕ, [Fe/H], R, and $|Z|$—have also favored a cosmological formation scenario for the thick disk (e.g., Spagna et al. 2010; Lee et al. 2011b), although they do not rule out scenario 4. Both of these papers stress the importance of further theoretical work. Only recently have simulations begun to include prescriptions for the formation of stars and their impact on the expected chemistry of the disk populations (e.g., Brook et al. 2005; Stinson et al. 2010; Loebman et al. 2011; and references therein). The measurement of the radial metallicity gradient at different heights above the plane is an additional observational constraint that can be used to test new, improved models for disk evolution and the next generation of simulations.

8. SUMMARY

Using a sample of 7010 MSTO stars from the SEGUE survey, we measure the radial metallicity gradient, Δ[Fe/H]/ΔR, as a function of height above the plane $|Z|$, in the Milky Way disk. Near the midplane, where our sample is dominated by the thin disk, we see a negative radial metallicity gradient that is consistent with previously published values (Friel et al. 2002; Luck et al. 2006; Luck & Lambert 2011). At large vertical heights, where our sample is dominated by the thick disk, the radial metallicity gradient becomes flat, consistent with previous work (Allende Prieto et al. 2006) using a sample located at larger $|Z|$. Our sample, located at low Galactic latitude, covers a larger range in R at small $|Z|$ and allows us to make a direct comparison between the thin disk and thick disk using the same sample.

At $|Z| > 1.0$ kpc, the median metallicity of old disk stars is consistent with the open cluster metallicities reported by Yong et al. (2005) and others at large R. In addition, our sample of disk stars shows that the flat gradient at large vertical height $|Z|$ extends to small R. Because the outer disk clusters are all located at large $|Z|$, the reported discontinuity in the radial gradient is consistent with a transition found using tracers at small $|Z|$ to large $|Z|$. We stress that abundances need to be examined as a function of both R and $|Z|$ in order to truly understand the observed trends.

In contrast to the outer disk, open clusters and Cepheids in the inner disk at high $|Z|$ have median metallicities ~0.5 dex higher than old disk stars at the same R and $|Z|$; thus, far from the Galactic plane, the younger tracers do not exhibit the same flat metallicity gradient that is seen in the old disk stars. Whether this is indicative of a metallicity gradient that becomes steeper with time is unclear, as radial migration may play a role in erasing a pre-existing gradient in the old disk stars, and what that initial gradient in the early disk may be is still uncertain.

A flat radial metallicity gradient in the thick disk is consistent with the predictions of a gas-rich, turbulent disk at high redshift (Brook et al. 2005; Bournaud et al. 2009). It may also be consistent with the scenarios of radial migration (Roškar et al. 2008a; Schönrich & Binney 2009b; Minchev & Famaey 2010) or minor mergers (Kazantzidis et al. 2008; Bird et al. 2012), provided that mixing in the radial direction is strong. We also cannot exclude the direct accretion of stars from satellites in minor mergers (Abadi et al. 2003). While we are not able to conclusively rule out any of these scenarios, the change in the radial gradient as a function of height above the plane is an important observational constraint for future theoretical work.

APPENDIX

CALCULATING WEIGHTS

As discussed in Section 4, we assign weights to each of the MSTO stars to reconstruct the properties of the underlying parent population. There are three major ways in which the spectroscopic sample is different from the full population along each line of sight: (1) the photometric objects in regions with the highest extinction were not considered for spectroscopy. (2) Not all candidates for spectroscopy are observed. (3) We
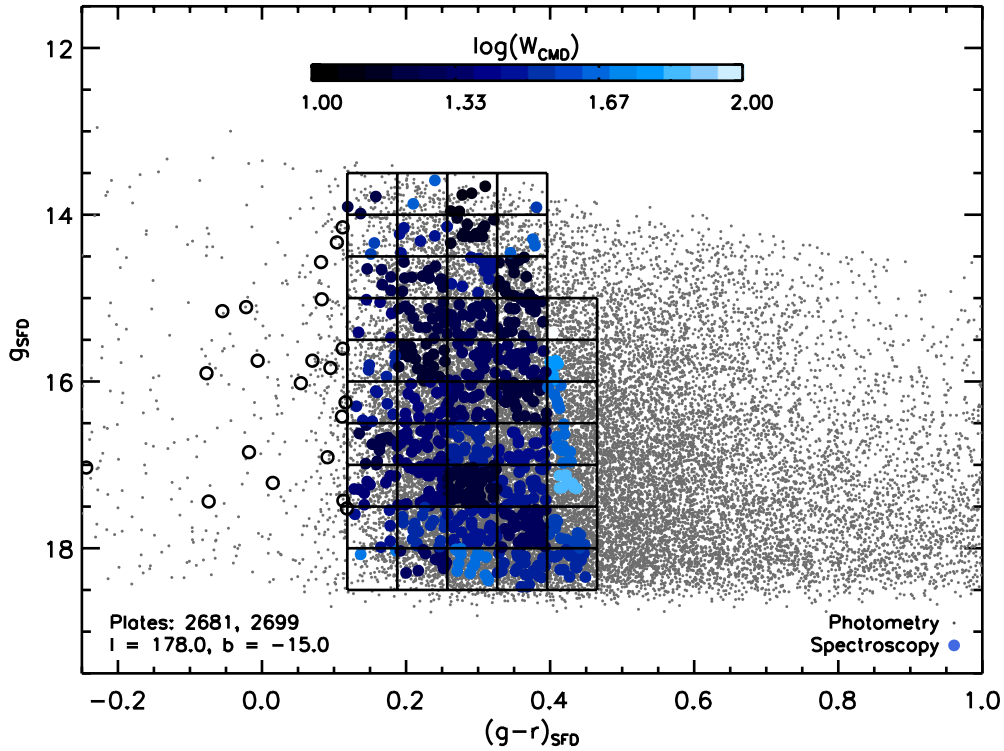
Figure 13. Weighting for CMD sampling along one line of sight. For each bin (black lines) the CMD weight W_{CMD} is calculated by taking the ratio of the number of photometric (gray) to spectroscopic (blue) objects. Outliers bluer or redder than 2σ from the mean value of $(g - r)_{SFD}$ are given a weight of zero (open circles). The values of W_{CMD} vary by less than a factor of 10, with the highest values in the reddest and faintest bins. This correction accounts for uneven sampling of the CMD, and provides the normalization for the variation in the number of objects between different lines of sight.

observe only MSTO stars using a color cut which is biased against redder metal-rich stars.

Each star in our sample is given three weights which correspond with the three differences listed above: (1) the *area weight*, W_A, which depends on the coverage on the plane of the sky in each line of sight; (2) the *CMD weight*, W_{CMD}, which depends on the target's location in the CMD; (3) the *LF weight*, W_{LF}, which depends on the target's T_{eff}, [Fe/H], and location in the CMD. The total weight, W, is the product of the three weights W_A, W_{CMD}, and W_{LF}.

A.1. Area Weight

The *area weight*, W_A, corrects for the area not covered by the spectroscopic survey because the 25% most extincted photometric objects were not considered for spectroscopy. These objects were removed using the cut described in steps 2 and 3 of Section 2.1. The right-hand panel of Figure 4 shows an example of how the missing area is distributed in the field. Since the missing area along each line of sight is slightly different, this weight is needed to ensure that every line of sight effectively probes the same volume of the Galaxy and has equal influence on the final gradient measurement. Since the dust is mainly in the foreground and our sample is primarily distant stars, we assume that the volume of the Galaxy behind the high-extinction patches is the same as the rest of the volume probed by stars along the same line of sight.

We use the extinction map of SFD98 to calculate the area with extinction lower than the cut. W_A is the ratio of the total area (7 deg²) to the low-extinction area. We note that the angular resolution of the SFD98 maps is 6.1 arcmin, so that the most extincted regions are always in irregular, contiguous patches on the sky.

A.2. Color–Magnitude Weight

The *CMD weight*, W_{CMD}, normalizes between the different lines of sight; while each field has roughly the same number of spectra, the total number of photometric objects varies a great deal due to the structure of the Galaxy. This weight also accounts for any uneven sampling due to the stochastic nature of the random selection of spectroscopic targets, especially at the faint and red limits, where targets are less likely to end up in our sample because they have low-quality spectra. We divide the CMD into bins of g_{SFD} and $(g-r)_{SFD}$, as shown in Figure 13 (black lines). We use the SFD98 colors and magnitudes because this is the CMD in which the $g - r$ cut was applied, meaning that the randomly selected spectra are an unbiased sample of the underlying CMD; this procedure does not depend on whether the SFD98 colors and magnitudes are correct. Since the sampling is a smooth function of color and magnitude and changes slowly, the bins are sufficiently small that we can assume the sampling is constant within each bin.

The magnitude bins are 0.5 mag wide and span the entire range of the sample. Because the CMD bins use corrections from SFD98, which can be affected by reddening, the color bins are different for each line of sight. To determine the color bins, we calculate the mean and standard deviation of the $(g - r)_{SFD}$ colors of stars in each line of sight. Spectra with colors more than 2σ from the mean have $W_{CMD} = 0$; this removes 217 stars from the sample (open circles). Most of these are very blue objects (hotter stars that are likely not on the main sequence), although some red objects are removed as well. The remaining sample in the line of sight is divided into five equal-sized color bins. In each CMD bin, W_{CMD} is the ratio of the number of photometric objects (small gray dots) to the number of spectroscopic objects (blue circles) and is shown by the color coding in Figure 13. The
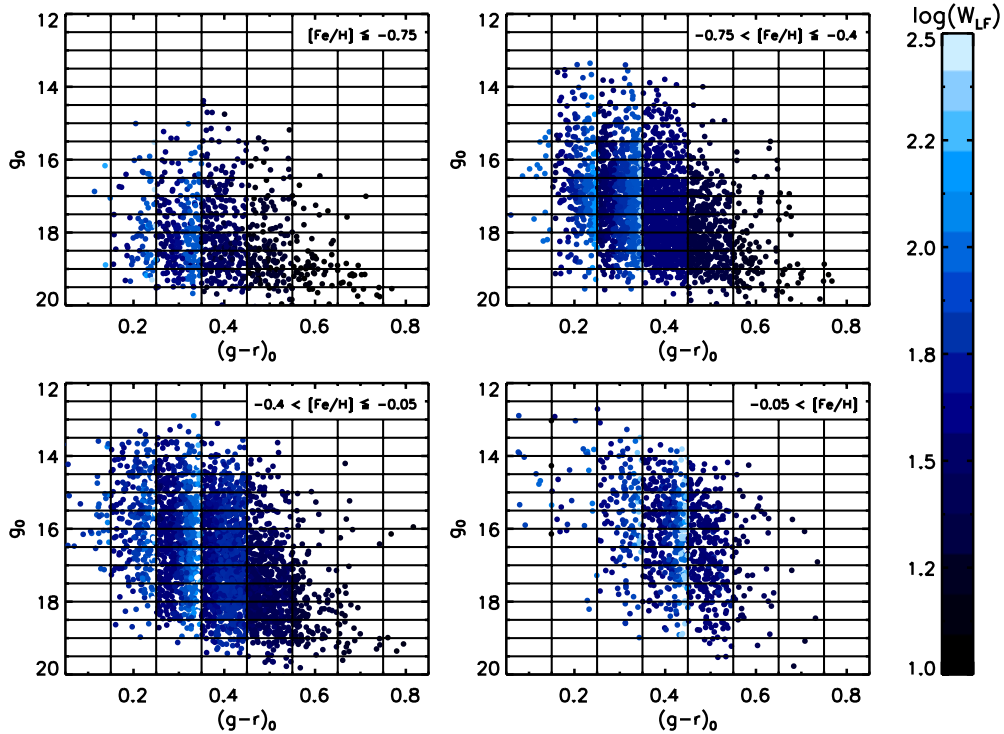
Figure 14. Weighting for LF coverage for the total sample. For each isochrone extinction-corrected CMD bin (black lines), the LF weight W_{LF} corrects for the fraction of the luminosity function that is observed. The luminosity function used varies with both metallicity (shown in four separate panels) and age, which depends on the temperature. This correction accounts for the bias against redder, metal-rich stars that results from the $(g - r)_{SFD}$ color selection.

(A color version of this figure is available in the online journal.)

highest values of W_{CMD} are found in the reddest and faintest bins, while the variation in the middle of CMD is relatively small. The difference between the smallest and largest values is less than a factor of 10.

A.3. Luminosity Function Weight

The *LF weight*, W_{LF}, allows us to use the MSTO sample as tracers for the total underlying population. The CMD bins used to calculate W_{LF} use the isochrone extinction-corrected g_0 and $(g - r)_0$ (see Section 3) and are different than those used to calculate W_{CMD}. For W_{CMD}, we had to account for the random sampling from the total photometric sample using the SFD98 corrections, but for W_{LF}, we use the isochrone extinction corrections because they provide the best estimates of the absolute magnitude and $g - r$ colors. The color bins are 0.1 mag wide in the range $0.05 < (g - r)_0 < 0.85$, while the magnitude bins are 0.5 mag wide in the range $12 < g_0 < 20$; these bins apply for all 11 lines of sight.

In each CMD bin, we find the fraction of the luminosity function that is observed in the given $(g - r)_0$ range. The weight is simply the reciprocal of the fraction. We use luminosity functions assuming a Chabrier (2001) lognormal initial mass function, generated by the Padova group (Girardi et al. 2004), where we have modified the faint end of the theoretical luminosity functions to more closely reflect the shapes of the luminosity functions for disk field stars reported by Reid et al. (2002) and Bochanski et al. (2010). The modifications are made at magnitudes fainter than the peak of the luminosity function. The luminosity function used depends on the age and metallicity of the target.

The derived values for W_{LF} are shown in Figure 14 for the entire sample; each panel shows a different range in metallicity.

For a given $(g - r)_0$ color, W_{LF} is roughly constant as a function of magnitude. Within a given color bin, the change in W_{LF} corresponds to different MSTOs for isochrones of different ages, which are assigned based on the temperatures of the targets as described in Section 3 and Figure 3. In a given CMD bin, stars with older ages have larger LF weights because a smaller fraction of the luminosity function is observable within the bin. This accounts for the vertical striping pattern seen in the four panels. We assign $W_{LF} = 0$ to the three spectra that fall outside the bounds of the CMD bins.

Using the luminosity functions to correct for the fraction of unobserved stars represented by the stars in each CMD bin corrects for the bias against selecting metal-rich stars for the spectroscopic sample. The identification of MSTO stars relies on a cut in $(g - r)_{SFD}$, which preferentially removes metal-rich stars that fall on isochrones with redder turnoff colors. Thus, for metal-rich stars, more of the population is removed by the color cut, and a smaller fraction of the luminosity function is observed. Consequently, metal-rich targets have larger values of W_{LF} (lighter blue circles) at the same $(g - r)_0$ in the four panels in Figure 14. In order for W_{LF} to make the proper corrections, we need to have a sufficient number of metal-rich stars in our sample to which we can apply the weights. Based on the arguments presented in Section 6.1, our sample should satisfy this requirement.

REFERENCES

Abadi, M. G., Navarro, J. F., Steinmetz, M., & Eke, V. R. 2003, ApJ, 597, 21
Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, ApJS, 182, 543
Aihara, H., Allende Prieto, C., An, D., et al. 2011, ApJS, 193, 29
Allende Prieto, C., Beers, T. C., Wilhelm, R., et al. 2006, ApJ, 636, 804
An, D., Pinsonneault, M. H., Masseron, T., et al. 2009, ApJ, 700, 523

Andrievsky, S. M., Bersier, D., Kovtyukh, V. V., et al. 2002a, A&A, 384, 140
Andrievsky, S. M., Kovtyukh, V. V., Luck, R. E., et al. 2002b, A&A, 392, 491
Andrievsky, S. M., Kovtyukh, V. V., Luck, R. E., et al. 2002c, A&A, 381, 32
Andrievsky, S. M., Luck, R. E., Martin, P., & Lépine, J. R. D. 2004, A&A, 413, 159
Belokurov, V., Evans, N. W., Irwin, M. J., et al. 2007, ApJ, 658, 337
Bensby, T., Feltzing, S., & Lundström, I. 2003, A&A, 410, 527
Bensby, T., Feltzing, S., & Lundström, I. 2004, A&A, 421, 969
Bensby, T., Feltzing, S., Lundström, I., & Ilyin, I. 2005, A&A, 433, 185
Bird, J. C., Kazantzidis, S., & Weinberg, D. H. 2012, MNRAS, in press (arXiv:1104.0933)
Bochanski, J. J., Hawley, S. L., Covey, K. R., et al. 2010, AJ, 139, 2679
Boesgaard, A. M., Jensen, E. E. C., & Deliyannis, C. P. 2009, AJ, 137, 4949
Bournaud, F., Elmegreen, B. G., & Martig, M. 2009, ApJ, 707, L1
Bragaglia, A., Carretta, E., Gratton, R. G., et al. 2001, AJ, 121, 327
Bragaglia, A., Sestito, P., Villanova, S., et al. 2008, A&A, 480, 79
Brook, C. B., Gibson, B. K., Martel, H., & Kawata, D. 2005, ApJ, 630, 298
Brook, C. B., Kawata, D., Gibson, B. K., & Freeman, K. C. 2004, ApJ, 612, 894
Brooks, A. M., Governato, F., Quinn, T., Brook, C. B., & Wadsley, J. 2009, ApJ, 694, 396
Brown, J. A., Wallerstein, G., Geisler, D., & Oke, J. B. 1996, AJ, 112, 1551
Caputo, F., Marconi, M., Musella, I., & Pont, F. 2001, A&A, 372, 544
Carraro, G., Bresolin, F., Villanova, S., et al. 2004, AJ, 128, 1676
Carraro, G., de La Fuente Marcos, R., Villanova, S., et al. 2007a, A&A, 466, 931
Carraro, G., Geisler, D., Villanova, S., Frinchaboy, P. M., & Majewski, S. R. 2007b, A&A, 476, 217
Carraro, G., Villanova, S., Demarque, P., Moni Bidin, C., & McSwain, M. V. 2008, MNRAS, 386, 1625
Carraro, G., Villanova, S., Demarque, P., et al. 2006, ApJ, 643, 1151
Carrera, R., & Pancino, E. 2011, A&A, 535, A30
Carretta, E., Bragaglia, A., & Gratton, R. G. 2007, A&A, 473, 129
Carretta, E., Bragaglia, A., Gratton, R. G., & Tosi, M. 2004, A&A, 422, 951
Carretta, E., Bragaglia, A., Gratton, R. G., & Tosi, M. 2005, A&A, 441, 131
Casagrande, L., Schönrich, R., Asplund, M., et al. 2011, A&A, 530, A138
Cescutti, G., Matteucci, F., François, P., & Chiappini, C. 2007, A&A, 462, 943
Chabrier, G. 2001, ApJ, 554, 1274
Chen, L., Hou, J. L., & Wang, J. J. 2003, AJ, 125, 1397
Chiappini, C., Matteucci, F., & Gratton, R. 1997, ApJ, 477, 765
Chiappini, C., Matteucci, F., & Romano, D. 2001, ApJ, 554, 1044
Chiba, M., & Beers, T. C. 2000, AJ, 119, 2843
Cresci, G., Mannucci, F., Maiolino, R., et al. 2010, Nature, 467, 811
Dalcanton, J. J., & Bernstein, R. A. 2002, AJ, 124, 1328
de Jong, J. T. A., Yanny, B., Rix, H., et al. 2010, ApJ, 714, 663
Dekel, A., Sari, R., & Ceverino, D. 2009, ApJ, 703, 785
De Silva, G. M., Freeman, K. C., Asplund, M., et al. 2007, AJ, 133, 1161
Dierickx, M., Klement, R., Rix, H.-W., & Liu, C. 2010, ApJ, 725, L186
Di Matteo, P., Lehnert, M. D., Qu, Y., & van Driel, W. 2011, A&A, 525, L3
Dotter, A., Chaboyer, B., Jevremović, D., et al. 2008, ApJS, 178, 89
Edvardsson, B., Pettersson, B., Kharrazi, M., & Westerlund, B. 1995, A&A, 293, 75
Eisenstein, D. J., Weinberg, D. H., Agol, E., et al. 2011, AJ, 142, 72
Elmegreen, B. G., & Elmegreen, D. M. 2005, ApJ, 627, 632
Elmegreen, B. G., & Elmegreen, D. M. 2006, ApJ, 650, 644
Ford, A., Jeffries, R. D., & Smalley, B. 2005, MNRAS, 364, 272
Friedli, D., Benz, W., & Kennicutt, R. 1994, ApJ, 430, L105
Friel, E. D., & Boesgaard, A. M. 1992, ApJ, 387, 170
Friel, E. D., Jacobson, H. R., Barrett, E., et al. 2003, AJ, 126, 2372
Friel, E. D., Jacobson, H. R., & Pilachowski, C. A. 2005, AJ, 129, 2725
Friel, E. D., Jacobson, H. R., & Pilachowski, C. A. 2010, AJ, 139, 1942
Friel, E. D., Janes, K. A., Tavarez, M., et al. 2002, AJ, 124, 2693
Frinchaboy, P. M., Marino, A. F., Villanova, S., et al. 2008, MNRAS, 391, 39
Fukugita, M., Ichikawa, T., Gunn, J. E., et al. 1996, AJ, 111, 1748
Gilmore, G., & Reid, N. 1983, MNRAS, 202, 1025
Gilmore, G., Wyse, R. F. G., & Jones, J. B. 1995, AJ, 109, 1095
Girardi, L., Grebel, E. K., Odenkirchen, M., & Chiosi, C. 2004, A&A, 422, 205
Goetz, M., & Koeppen, J. 1992, A&A, 262, 455
Gratton, R. G., & Contarini, G. 1994, A&A, 283, 911
Gunn, J. E., Carr, M., Rockosi, C., et al. 1998, AJ, 116, 3040
Gunn, J. E., Siegmund, W. A., Mannery, E. J., et al. 2006, AJ, 131, 2332
Hayashi, H., & Chiba, M. 2006, PASJ, 58, 835
Haywood, M. 2008, MNRAS, 388, 1175
Hobbs, L. M., & Thorburn, J. A. 1992, AJ, 104, 669
Ibata, R., Irwin, M., Lewis, G., Ferguson, A. M. N., & Tanvir, N. 2001, Nature, 412, 49
Ivezić, Ž., Sesar, B., Jurić, M., et al. 2008, ApJ, 684, 287
Jacobson, H. R., Friel, E. D., & Pilachowski, C. A. 2007, AJ, 134, 1216

Jacobson, H. R., Friel, E. D., & Pilachowski, C. A. 2008, AJ, 135, 2341
Jacobson, H. R., Friel, E. D., & Pilachowski, C. A. 2009, AJ, 137, 4753
Jacobson, H. R., Friel, E. D., & Pilachowski, C. A. 2011a, AJ, 141, 58
Jacobson, H. R., Pilachowski, C. A., & Friel, E. D. 2011b, AJ, 142, 59
Jurić, M., Ivezić, Ž., Brooks, A., et al. 2008, ApJ, 673, 864
Katz, D., Soubiran, C., Cayrel, R., et al. 2011, A&A, 525, A90
Kazantzidis, S., Bullock, J. S., Zentner, A. R., Kravtsov, A. V., & Moustakas, L. A. 2008, ApJ, 688, 254
Kazantzidis, S., Zentner, A. R., Kravtsov, A. V., Bullock, J. S., & Debattista, V. P. 2009, ApJ, 700, 1896
Lacey, C. G., & Fall, S. M. 1985, ApJ, 290, 154
Larson, R. B. 1976, MNRAS, 176, 31
Lee, Y. S., Beers, T. C., Allende Prieto, C., et al. 2011a, AJ, 141, 90
Lee, Y. S., Beers, T. C., An, D., et al. 2011b, ApJ, 738, 187
Lee, Y. S., Beers, T. C., Sivarani, T., et al. 2008a, AJ, 136, 2022
Lee, Y. S., Beers, T. C., Sivarani, T., et al. 2008b, AJ, 136, 2050
Lemasle, B., François, P., Piersimoni, A., et al. 2008, A&A, 490, 613
Loebman, S. R., Roškar, R., Debattista, V. P., et al. 2011, ApJ, 737, 8
Luck, R. E., Gieren, W. P., Andrievsky, S. M., et al. 2003, A&A, 401, 939
Luck, R. E., Kovtyukh, V. V., & Andrievsky, S. M. 2006, AJ, 132, 902
Luck, R. E., & Lambert, D. L. 2011, AJ, 142, 136
Maciel, W. J., Lago, L. G., & Costa, R. D. D. 2005, A&A, 433, 127
Magrini, L., Randich, S., Zoccali, M., et al. 2010, A&A, 523, A11
Magrini, L., Sestito, P., Randich, S., & Galli, D. 2009, A&A, 494, 95
Martin, P., & Roy, J.-R. 1994, ApJ, 424, 599
Martínez-Delgado, D., Gabany, R. J., Crawford, K., et al. 2010, AJ, 140, 962
Matteucci, F., & Francois, P. 1989, MNRAS, 239, 885
Mikolaitis, Š., Tautvaišienė, G., Gratton, R., Bragaglia, A., & Carretta, E. 2011, MNRAS, 416, 1092
Minchev, I., & Famaey, B. 2010, ApJ, 722, 112
Newberg, H. J., Yanny, B., Rockosi, C., et al. 2002, ApJ, 569, 245
Nordström, B., Mayor, M., Andersen, J., et al. 2004, A&A, 418, 989
Origlia, L., Valenti, E., Rich, R. M., & Ferraro, F. R. 2006, ApJ, 646, 499
Pace, G., Pasquini, L., & François, P. 2008, A&A, 489, 403
Pancino, E., Carrera, R., Rossetti, E., & Gallart, C. 2010, A&A, 511, A56
Pedicelli, S., Bono, G., Lemasle, B., et al. 2009, A&A, 504, 81
Peterson, R. C., & Green, E. M. 1998, ApJ, 502, L39
Portinari, L., & Chiosi, C. 2000, A&A, 355, 929
Prantzos, N., & Boissier, S. 2000, MNRAS, 313, 338
Purcell, C. W., Kazantzidis, S., & Bullock, J. S. 2009, ApJ, 694, L98
Randich, S., Sestito, P., & Pallavicini, R. 2003, A&A, 399, 133
Randich, S., Sestito, P., Primas, F., Pallavicini, R., & Pasquini, L. 2006, A&A, 450, 557
Read, J. I., Lake, G., Agertz, O., & Debattista, V. P. 2008, MNRAS, 389, 1041
Reid, I. N., Gizis, J. E., & Hawley, S. L. 2002, AJ, 124, 2721
Roškar, R., Debattista, V. P., Quinn, T. R., Stinson, G. S., & Wadsley, J. 2008a, ApJ, 684, L79
Roškar, R., Debattista, V. P., Stinson, G. S., et al. 2008b, ApJ, 675, L65
Sales, L. V., Helmi, A., Abadi, M. G., et al. 2009, MNRAS, 400, L61
Sánchez-Blázquez, P., Courty, S., Gibson, B. K., & Brook, C. B. 2009, MNRAS, 398, 591
Santos, N. C., Lovis, C., Pace, G., Melendez, J., & Naef, D. 2009, A&A, 493, 309
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
Schönrich, R., & Binney, J. 2009a, MNRAS, 396, 203
Schönrich, R., & Binney, J. 2009b, MNRAS, 399, 1145
Sellwood, J. A., & Binney, J. J. 2002, MNRAS, 336, 785
Sestito, P., Bragaglia, A., Randich, S., et al. 2006, A&A, 458, 121
Sestito, P., Bragaglia, A., Randich, S., et al. 2008, A&A, 488, 943
Sestito, P., Randich, S., & Bragaglia, A. 2007, A&A, 465, 185
Sestito, P., Randich, S., & Pallavicini, R. 2004, A&A, 426, 809
Smiljanic, R., Gauderon, R., North, P., et al. 2009, A&A, 502, 267
Smolinski, J. P., Lee, Y. S., Beers, T. C., et al. 2011, AJ, 141, 89
Soubiran, C., Bienaymé, O., & Siebert, A. 2003, A&A, 398, 141
Spagna, A., Lattanzi, M. G., Re Fiorentin, P., & Smart, R. L. 2010, A&A, 510, L4
Spitoni, E., & Matteucci, F. 2011, A&A, 531, A72
Stanghellini, L., & Haywood, M. 2010, ApJ, 714, 1096
Steinmetz, M., Zwitter, T., Siebert, A., et al. 2006, AJ, 132, 1645
Stewart, K. R., Bullock, J. S., Wechsler, R. H., Maller, A. H., & Zentner, A. R. 2008, ApJ, 683, 597
Stinson, G. S., Bailin, J., Couchman, H., et al. 2010, MNRAS, 408, 812
Tautvaišiene, G., Edvardsson, B., Tuominen, I., & Ilyin, I. 2000, A&A, 360, 499
Twarog, B. A., Ashman, K. M., & Anthony-Twarog, B. J. 1997, AJ, 114, 2556
Venn, K. A., Irwin, M., Shetrone, M. D., et al. 2004, AJ, 128, 1177
Villalobos, Á., & Helmi, A. 2008, MNRAS, 391, 1806

Villanova, S., Baume, G., & Carraro, G. 2007, MNRAS, 379, 1089
Villanova, S., Carraro, G., Bresolin, F., & Patat, F. 2005, AJ, 130, 652
Villanova, S., Geisler, D., & Piotto, G. 2010, ApJ, 722, L18
Wilson, M. L., Helmi, A., Morrison, H. L., et al. 2011, MNRAS, 413, 2235
Yanny, B., Rockosi, C., Newberg, H. J., et al. 2009, AJ, 137, 4377
Yoachim, P., & Dalcanton, J. J. 2005, ApJ, 624, 701

Yoachim, P., & Dalcanton, J. J. 2006, AJ, 131, 226
Yoachim, P., & Dalcanton, J. J. 2008, ApJ, 683, 707
Yong, D., Carney, B. W., & Teixera de Almeida, M. L. 2005, AJ, 130, 597
Yong, D., Carney, B. W., Teixera de Almeida, M. L., & Pohl, B. L. 2006, AJ, 131, 2256
York, D. G., Adelman, J., Anderson, J. E., Jr., et al. 2000, AJ, 120, 1579
Yoshii, Y. 1982, PASJ, 34, 365