



Elemental Abundances in M31: Gradients in the Giant Stellar Stream*

Iyanna Escala^{1,2,7} , Karoline M. Gilbert^{3,4} , Jennifer Wojno⁴ , Evan N. Kirby⁵ , and Puragra Guhathakurta⁶ ¹The Observatories of the Carnegie Institution for Science, 813 Santa Barbara Street, Pasadena, CA 91101, USA; iescala@carnegiescience.edu²Department of Astrophysical Sciences, Princeton University, 4 Ivy Lane, Princeton, NJ 08544, USA³Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA⁴Department of Physics & Astronomy, Bloomberg Center for Physics and Astronomy, Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218, USA⁵Department of Astronomy, California Institute of Technology, 1200 E California Boulevard, Pasadena, CA 91125, USA⁶UCO/Lick Observatory, Department of Astronomy & Astrophysics, University of California Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA

Received 2021 March 15; revised 2021 May 4; accepted 2021 May 5; published 2021 July 6

Abstract

We analyze existing measurements of $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$ for individual red giant branch (RGB) stars in the Giant Stellar Stream (GSS) of M31 to determine whether spatial abundance gradients are present. These measurements were obtained from low- ($R \sim 3000$) and moderate- ($R \sim 6000$) resolution Keck/DEIMOS spectroscopy using spectral synthesis techniques as part of the Elemental Abundances in M31 survey. From a sample of 62 RGB stars spanning the GSS at 17, 22, and 33 projected kpc, we measure a $[\text{Fe}/\text{H}]$ gradient of $-0.018 \pm 0.003 \text{ dex kpc}^{-1}$ and negligible $[\alpha/\text{Fe}]$ gradient with M31-centric radius. We investigate GSS abundance patterns in the outer halo using additional $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$ measurements for six RGB stars located along the stream at 45 and 58 projected kpc. These abundances provide tentative evidence that the trends in $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$ beyond 40 kpc in the GSS are consistent with those within 33 kpc. We also compare the GSS abundances to 65 RGB stars located along the possibly related Southeast (SE) shelf substructure at 12 and 18 projected kpc. The abundances of the GSS and SE shelf are consistent, supporting a common origin hypothesis, although this interpretation may be complicated by the presence of $[\text{Fe}/\text{H}]$ gradients in the GSS. We discuss the abundance patterns in the context of photometric studies from the literature and explore implications for the properties of the GSS progenitor, suggesting that the high $\langle[\alpha/\text{Fe}]\rangle$ of the GSS ($+0.40 \pm 0.05 \text{ dex}$) favors a major merger scenario for its formation.

Unified Astronomy Thesaurus concepts: [Stellar abundances \(1577\)](#); [Galaxy abundances \(574\)](#); [Stellar streams \(2166\)](#); [Galaxy stellar halos \(598\)](#); [Andromeda Galaxy \(39\)](#); [Galaxy formation \(595\)](#)

1. Introduction

Stellar streams originate from the ongoing tidal disruption of accreted galaxies and globular clusters, providing an instantaneous view of the hierarchical formation of the host galaxy (e.g., Freeman & Bland-Hawthorn 2002; Bullock & Johnston 2005; Helmi 2020). In the Milky Way (MW), the discovery of the Sagittarius stream (Ibata et al. 2001a) provided an early indication of the importance of mergers in Galactic formation history. The contemporaneous discovery of M31's Giant Stellar Stream (GSS; Ibata et al. 2001b) further indicated that stellar streams are a common feature of galaxies beyond the MW and that mergers have also played a significant role in M31's evolution.

The GSS is a conspicuous tidal structure in M31's southeastern quadrant that spans at least 6° (~ 80 projected kpc) on the sky and $\gtrsim 100$ kpc in line-of-sight distance over its extent (McConnachie et al. 2003; Conn et al. 2016). The stream appears to be characterized by a metal-rich, high surface brightness core ($\Sigma_V \sim 30 \text{ mag arcsec}^{-2}$; Ibata et al. 2001b) and

an asymmetric envelope that has both lower metallicity and surface brightness (Ibata et al. 2007). In comparison to the phase-mixed component of M31's stellar halo, photometric and spectroscopic studies of the GSS's resolved stellar populations have revealed that it is more metal-rich and kinematically colder and possesses more dominant intermediate-age stellar populations (e.g., Guhathakurta et al. 2006; Brown et al. 2006; Kalirai et al. 2006; Ibata et al. 2007; Gilbert et al. 2009; Tanaka et al. 2010; Ibata et al. 2014). Based on these properties, the GSS was inferred to originate from the recent ($\lesssim 1 \text{ Gyr}$) disruption of a distinct satellite progenitor on a highly radial orbit with a lower stellar mass limit of $\sim 10^8 M_\odot$ (Ibata et al. 2004; Fardal et al. 2006; Font et al. 2006b).

However, the nature of the GSS accretion event is likely more complex than initially surmised. Spectroscopic surveys of M31's stellar halo have uncovered a number of faint kinematical features that are tidal debris possibly related to the GSS. Kalirai et al. (2006) first detected a second kinematically cold component (KCC) in a field probing the GSS at 20 projected kpc that was not a prediction of concurrent dynamical models (Ibata et al. 2004; Fardal et al. 2006; Font et al. 2006b), despite the similarity of its photometric metallicity to the primary GSS substructure. Gilbert et al. (2009) later traced the KCC inward to 17 projected kpc, showing that the feature was consistently separated from the GSS by $\sim 100 \text{ km s}^{-1}$ in line-of-sight velocity over its spanned radial range, thus providing compelling evidence in favor of a direct physical connection between the GSS and KCC.

Following the discovery of the KCC, Gilbert et al. (2007) kinematically detected a faint substructure component located $\sim 11\text{--}18$ projected kpc along M31's southeastern minor axis. Unlike in the case of the KCC, this feature matched predictions

* The data presented herein were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California and the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation.

⁷ Carnegie-Princeton Fellow.



from models of the GSS accretion event; specifically, for the Southeast (SE) shelf generated by the fourth pericentric passage of the GSS progenitor (Fardal et al. 2006, 2007). The similarity of the photometric metallicity and age distributions of stellar populations in the SE shelf and GSS (Brown et al. 2003, 2006; Gilbert et al. 2007) further bolstered the hypothesis that the SE shelf and GSS were tidal debris from the same event. The prediction of the SE shelf illustrates that minor merger models for the formation of the GSS ($M_* \sim (1-5) \times 10^9 M_\odot$; Fardal et al. 2006, 2007, 2008, 2013; Mori & Rich 2008; Kirihiara et al. 2014; Sadoun et al. 2014; Kirihiara et al. 2017; Miki et al. 2016) can successfully reproduce the broad morphological and kinematical features of the stream while accounting for diffuse shell-like features such as the Northeast (Ferguson et al. 2002, 2005) and West (W; Fardal et al. 2007) shelves as part of the forward continuation of the stream. In further support of this hypothesis, Fardal et al. (2012) showed that the kinematics of the W shelf were strikingly similar to predictions for the feature and that the shelf’s metallicity was consistent with that of the GSS.

Nevertheless, minor merger models for the GSS’s formation are unable to simultaneously provide a concise explanation for the origin of the KCC. Gilbert et al. (2019) speculated that an asymmetric extension of the W shelf toward M31’s south-eastern quadrant could potentially account for the KCC within this framework, although multiple superposed loops of the GSS also provide a feasible explanation for the KCC in a major merger scenario ($M_* \sim 10^{10} M_\odot$; Bhattacharya 2019, 2021; Hammer et al. 2010, 2018; D’Souza & Bell 2018). The formation of the GSS via a major merger had not been explored earlier in order to preserve the integrity of M31’s disk (e.g., Mori & Rich 2008), although simulations have since demonstrated that gas-rich mergers can enable disk survival (e.g., Hopkins et al. 2009). Without disk intactness as a constraining factor, the GSS and its associated shells can be reproduced by merger ratios varying from 300:1 to 2:1 (Hammer et al. 2018), casting uncertainty on whether a major (<10:1) or minor ($\geq 10:1$) merger is responsible for the stream.

Chemical abundance measurements ([Fe/H] and $[\alpha/\text{Fe}]$) of individual red giant branch (RGB) stars in the GSS have the potential to elucidate the properties of the progenitor by breaking the degeneracy between formation models. Simulations of MW-mass galaxies have shown that the mass and accretion time distributions of external progenitors can imprint strong chemical signatures in a galaxy’s accreted stellar populations in terms of Fe and α elements (O, Ne, Mg, Si, S, Ar, Ca, and Ti), respectively (e.g., Robertson et al. 2005; Font et al. 2006a; Johnston et al. 2008). Using an extrapolation of the stellar mass–metallicity relation for Local Group dwarf galaxies (Kirby et al. 2013), Gilbert et al. (2019) estimated a stellar mass for the progenitor of $(1-5) \times 10^9 M_\odot$ based on the first spectral synthesis–based [Fe/H] measurements from a field located at 17 projected kpc in the GSS. Gilbert et al. also found that the GSS has a high average α enhancement (~ 0.4 dex), indicating that its progenitor formed stars with high efficiency. Escala et al. (2020a) later confirmed that the chemical abundance patterns found by Gilbert et al. (2019) extended to a GSS field at 22 projected kpc. Although the stellar mass predicted by iron abundance in the GSS is consistent with minor merger models, this cannot be interpreted as direct evidence in favor of such a scenario if the progenitor had a metallicity gradient.

Indeed, massive satellite galaxies in the Local Group such the Large and Small Magellanic Clouds (LMC and SMC), M33, and Sagittarius (Sgr) are known to possess negative radial metallicity gradients in their RGB populations. For the LMC, SMC, and M33, radial metallicity gradients of $-(0.06-0.08)$ dex kpc^{-1} have been detected out to several disk scale lengths (LMC: Choudhury et al. 2016; SMC: Dobbie et al. 2014; Parisi et al. 2016; Choudhury et al. 2018; M33: Kim et al. 2002; Tiede et al. 2004; Barker et al. 2007) that are similar to the gradient in the MW’s disk (-0.06 dex kpc^{-1} ; e.g., Cheng et al. 2012; Hayden et al. 2014). In addition, metallicity differences of 0.4–0.6 have been observed between the Sgr core and Sgr streams (Chou et al. 2007; Monaco et al. 2007; Keller et al. 2010; Hayes et al. 2020), which translate to an intrinsic gradient of about -0.2 dex kpc^{-1} in the Sgr progenitor based on dynamical modeling (Law & Majewski 2010).⁸ Although only weak internal gradients are measured along the Sgr streams ($(-1.2-(-1.4)) \times 10^{-3}$ dex deg^{-1} , but consistent with zero within the uncertainties; Hayes et al. 2020), combining these measurements with modeling provides evidence for a gradient with dynamical age (i.e., initial orbital radius; -0.12 ± 0.03 dex Gyr^{-1}) in the Sgr progenitor. In comparison, smaller Local Group dwarf galaxies ($M_* \lesssim 10^{8.5} M_\odot$) have diverse gradients, ranging from flat to as steep as -0.4 dex per half-light radius (e.g., Kirby et al. 2011, 2017; Leaman et al. 2013; Vargas et al. 2014b; Ho et al. 2015; Kacharov et al. 2017), with no clear relationship between the magnitude of a gradient and luminosity, host distance, or morphology (Ho et al. 2015, cf. Leaman et al. 2013), although there may be a trend with median stellar age (Mercado et al. 2021).

In accordance with expectations based on the GSS progenitor’s inferred mass, an early photometric survey of stellar populations along the line of sight to the GSS (Ferguson et al. 2002) noted the presence of color variations over the stream, which were attributed to metallicity variations. Subsequently, Ibata et al. (2007) inspected such variations between the high surface brightness core of the GSS and its extended envelope, noting that the latter had lower average photometric metallicity. Gilbert et al. (2009) provided additional support for this dichotomy by measuring photometric metallicities of spectroscopically confirmed RGB stars in the core and envelope of the GSS. More recently, photometric studies have embarked on increasingly detailed explorations of GSS metallicity variations as a function of two-dimensional position on the sky (Conn et al. 2016; Cohen et al. 2018).

Thus, detailed observations of abundance gradients in the GSS are necessary to map on-sky variations to the initial abundance properties of the progenitor. In this work, we present a comprehensive analysis of spatial [Fe/H] and $[\alpha/\text{Fe}]$ gradients in the GSS and likely associated substructures using spectral synthesis–based abundance measurements from the Elemental Abundances in M31 survey (Escala et al. 2019, 2020a, 2020b; Gilbert et al. 2019, 2020; Kirby et al. 2020; Wojno et al. 2020) with the aim of providing further constraints for GSS formation models. In Section 2, we provide an overview of the spectroscopic data and chemical abundance measurements, which we use to investigate the GSS’s

⁸ Possible explanations for this steep gradient include a massive, disk Sgr progenitor similar to the LMC or, more likely, the lack of self-consistent modeling of continuing star formation in the Sgr core as the progenitor tidally disrupts over several gigayears (Hayes et al. 2020).

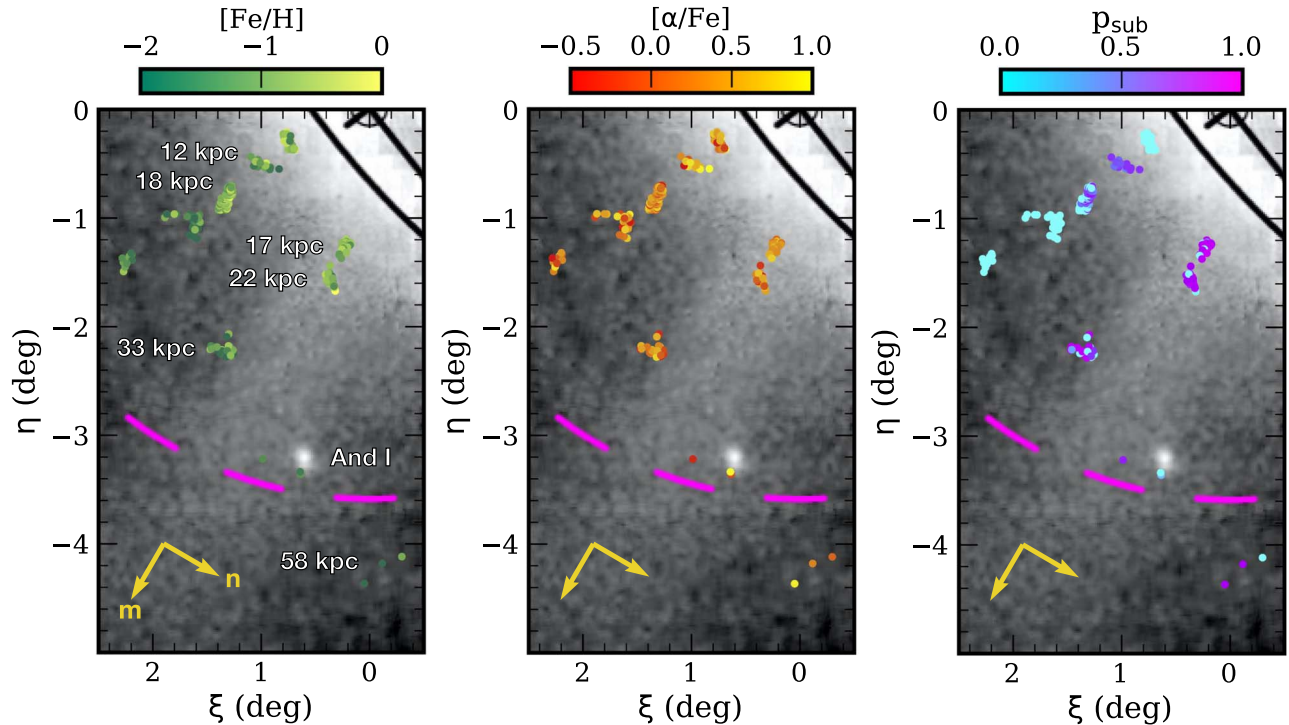


Figure 1. The spatial distribution of 203 RGB stars in M31’s SE quadrant, given in M31-centric coordinates, with measurements of $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$ from the Elemental Abundances in M31 survey (Gilbert et al. 2019, 2020; Escala et al. 2020a, 2020b; J. Wojno et al. 2021, in preparation). The spectroscopic fields utilized in this work (Table 1) are labeled according to their projected M31-centric distance, excepting And I at 45 projected kpc (left panel). The RGB stars are overlaid on the PAndAS star count map (McConnachie et al. 2018). In each panel, stars are color coded by (left) $[\text{Fe}/\text{H}]$, (middle) $[\alpha/\text{Fe}]$, and (right) probability of belonging to kinematically cold substructure. The thick, solid black lines represent the edge of M31’s classical disk ($i = 77^\circ$, $r = 17$ kpc) and the orientation of its minor axis. The dashed magenta lines delineate 50 projected kpc. The gold vectors represent GSS-aligned coordinate axes (Fardal et al. 2006).

abundance properties between 17 and 58 kpc in Section 3. We conclude by discussing our results in the context of both the observational and theoretical literature in Section 4 before summarizing in Section 5.

2. Data

We utilized existing measurements of $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$ for individual RGB stars in M31’s stellar halo obtained from low- ($R \sim 3000$) and moderate- ($R \sim 6000$) resolution Keck/DEIMOS spectroscopy as part of the Elemental Abundances in M31 survey (Gilbert et al. 2019, 2020; Escala et al. 2019, 2020a, 2020b). In total, 200 RGB stars in our sample have published measurements of $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$ in the southeastern quadrant of M31’s stellar halo. We also include unpublished measurements (J. Wojno et al., 2021, in preparation) for three M31 RGB stars in a spectroscopic field overlapping with the GSS envelope at 58 projected kpc. Figure 1 illustrates the spatial distribution of these stars compared to the star count map from the Pan-Andromeda Archaeological Survey (PAndAS; McConnachie et al. 2018), while providing a sense of the variation in $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$ over the probed region.

In contrast to previous work by Escala et al. (2020b) using a nearly identical sample, we focused our analysis on M31 RGB stars with a high probability of belonging to kinematically identifiable substructure based on their heliocentric radial velocities (p_{sub} ; right panel of Figure 1; Section 2.2). The majority of these stars are located in spectroscopic fields along the GSS at 17, 22, and 33 projected kpc from the center of M31, with a few stars located in the outer halo at 45 and 58

projected kpc. Additional stars with nonzero substructure probability, which are likely associated with the SE shelf substructure (Gilbert et al. 2007; Fardal et al. 2007; Escala et al. 2020a), are located in fields at 12 and 18 kpc along M31’s minor axis. We provide a brief summary of the spectroscopic observations and abundance measurements below and refer the reader to Gilbert et al. (2019, 2020) and Escala et al. (2019, 2020a, 2020b) for further details.

2.1. Spectroscopy

All spectroscopic fields, except a13 and And I,⁹ were observed for a minimum of 5 hr with the 600 line mm^{-1} or 1200 line mm^{-1} grating for the case of low- and moderate-resolution spectroscopy, respectively. These configurations result in spectra with an FWHM spectral resolution of 2.8 Å ($R \sim 3000$) and 1.2 Å ($R \sim 6000$). Additionally, each deep (5+ hr) field was designed from previous shallow (~ 1 hr) DEIMOS observations from the Spectroscopic and Photometric Landscape of Andromeda’s Stellar Halo (SPLASH) survey (Guhathakurta et al. 2006; Kalirai et al. 2006; Gilbert et al. 2007, 2009) in order to maximize the yield of spectroscopically confirmed M31 RGB stars. Data for fields a13 and And I were obtained as part of the SPLASH survey, where a handful of stars have spectra with fortuitously high signal-to-noise ratios (S/Ns) such that measuring abundances is feasible. The spectra were reduced using a modified version of the spec2d pipeline

⁹ The field And I is based on a mixture of both shallow and deep spectroscopic data from the SPLASH survey (Gilbert et al. 2009) and the Elemental Abundances in M31 survey (Kirby et al. 2020).

Table 1
Properties of Substructure in Spectroscopic Fields

Field	r_{proj} (kpc)	Comp.	μ (km s ⁻¹)	σ (km s ⁻¹)	f	$\langle[\text{Fe}/\text{H}]\rangle$	$\langle[\alpha/\text{Fe}]\rangle$	N_{M31}	$N_{[\alpha/\text{Fe}]}$	Ref.
f207	17	GSS	-529.4	24.5	0.33	$-0.87^{+0.09}_{-0.10}$	$+0.44^{+0.04}_{-0.05}$	108	21	1, 2
		KCC	-427.3	21.0	0.32	-0.79 ± 0.07	$+0.54 \pm 0.06$
S	22	GSS	-489.0	26.1	0.49	$-1.02^{+0.15}_{-0.14}$	$+0.38^{+0.17}_{-0.19}$	87	20	3
		KCC	-371.6	17.6	0.22	-0.71 ± 0.11	$+0.35^{+0.08}_{-0.09}$
a3	33	GSS	-444.6	15.7	0.56	$-1.11^{+0.12}_{-0.13}$	$+0.34^{+0.08}_{-0.09}$	75	21	1, 4
And I	45	GSS	-383.3	32.4	0.56	$-1.49^{+0.10}_{-0.02}$	$-0.19^{+0.21}_{-0.05}$	38	3	1, 5
a13	58	GSS	-301.6	29.2	0.62	$-2.50^{+0.40}_{-0.54}$	$+0.71^{+0.19}_{-0.34}$	31	3	1, 6, 7
H	12	SE Shelf	-295.4	65.8	0.56	$-1.30^{+0.13}_{-0.12}$	$+0.53^{+0.08}_{-0.10}$	104	16	3
f123	18	SE Shelf	-279.9	11.0	0.32	-0.71 ± 0.07	$+0.41^{+0.04}_{-0.05}$	74	49	1, 4

Note.—The columns of the table refer to spectroscopic field name; projected radius from M31’s galactic center; substructure component; the median heliocentric velocity (μ), dispersion (σ), and fractional contribution (f) of the velocity model for a given component (Section 2.2); average $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$ for a given component (Section 2.3), weighted by the inverse variance of the measurement uncertainty and substructure probability; number of spectroscopically confirmed RGB members (excluding dwarf galaxy members for And I; Section 2.2); and number of RGB stars with successful $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$ measurements. References: (1) Gilbert et al. (2018), (2) Gilbert et al. (2019), (3) Escala et al. (2020a), (4) Escala et al. (2020b), (5) Gilbert et al. (2020), (6) J. Wojno et al., 2021, in preparation, (7) this work.

(Cooper et al. 2012; Newman et al. 2013 for the original pipeline; Simon & Geha 2007; Kirby et al. 2020 for modifications specific to stellar point sources). Table 1 provides a summary of the properties for each spectroscopic field containing kinematically identifiable substructure.

2.2. Radial Velocity Measurements and Membership Determination

Radial velocities were measured via cross-correlation with templates using the procedures described in Simon & Geha (2007) and Kirby et al. (2015). The statistical uncertainty is calculated from Monte Carlo trials in which a given observed spectrum is perturbed according to its standard error, whereas the systematic uncertainty is calculated via repeat velocity measurements of the same stars. We adopted a systematic velocity term of 1.49 km s⁻¹ for 1200 line mm⁻¹ grating spectra (Kirby et al. 2015) and 5.6 km s⁻¹ for 600 line mm⁻¹ grating spectra (Collins et al. 2011).

Using a combination of heliocentric radial velocity, color-magnitude diagram (CMD) position, Na I $\lambda\lambda 8190$ equivalent widths, and photometric and calcium triplet-based metallicity estimates, we assigned a probability of belonging to M31 to each star with a successful velocity measurement. We utilized the Bayesian inference method of Escala et al. (2020b) to determine membership for all stars except those in the 17, 45, and 58 kpc fields, for which we used the maximum likelihood-based technique of Gilbert et al. (2006). For the 45 kpc field, we additionally required that M31 RGB candidates had $[\text{Fe}/\text{H}]_{\text{phot}} > -0.95$ in order to separate M31 halo stars from stars belonging to the And I dwarf spheroidal galaxy. Gilbert et al. (2009, 2020) demonstrated that this additional photometric metallicity criterion clearly demarcates the two populations, regardless of velocity or apparent proximity to the dwarf galaxy.

As discussed by Escala et al. (2020b), the Bayesian inference and maximum likelihood-based methods of M31 membership determination produce generally consistent results, where Escala et al.’s classification of stars as M31 members is slightly more conservative. In general, we consider stars to be M31 RGB stars if they are more likely to belong to M31 than to the MW foreground. The membership criterion for the 45 and

58 kpc fields is more stringent, requiring that stars are at least three times more likely to belong to M31 than to the MW, owing to the increased likelihood of contamination by MW foreground dwarfs in M31’s sparse outer halo.

Figure 2 shows the heliocentric radial velocity distribution of RGB stars in each spectroscopic field containing substructure (Table 1), where stars with velocities consistent with that of kinematically cold components have a higher probability of belonging to substructure (p_{sub} ; Gilbert et al. 2019; Escala et al. 2020a, 2020b). Figure 2 also shows Gaussian mixture models of the velocity distribution for each field, where each model contains both halo and substructure components. We adopted the 50th percentile values of the marginalized posterior probability distributions from Escala et al. (2020a) to model the substructure components in the 12 and 22 kpc fields, whereas all other component models (including halo components) are from Gilbert et al. (2018). The substructure probability for a star with a given radial velocity is thus the odds ratio of the Bayes factor under the assumption of the substructure versus halo models.

2.3. Chemical Abundance Measurements

Chemical abundance ($[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$) and stellar parameter measurements (T_{eff}) were obtained from spectral synthesis of low- and medium-resolution stellar spectroscopy for individual RGB stars in each field. In summary, each observed spectrum is compared to a grid of synthetic spectra using Levenberg-Marquardt minimization to identify the best-fit stellar parameters and abundances. Throughout this procedure, the spectroscopic effective temperature (T_{eff}) is loosely constrained by photometry, whereas the surface gravity ($\log g$) is fixed to its photometric value, assuming a distance modulus of $(m - M) = 24.63 \pm 0.2$ (Clementini et al. 2011) for M31. Measurements of $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$ obtained for identical stars from low- and medium-resolution spectra (Section 2.1) are generally consistent within the uncertainties (Escala et al. 2020a). Systematic uncertainties on the abundance measurements are added in quadrature to the random component of the uncertainty from the fitting procedure. We adopted systematic error terms of 0.130 (0.101) and 0.107 (0.084) for $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$ measurements, respectively, obtained from 600 (1200) line mm⁻¹ spectra (Gilbert et al. 2019;

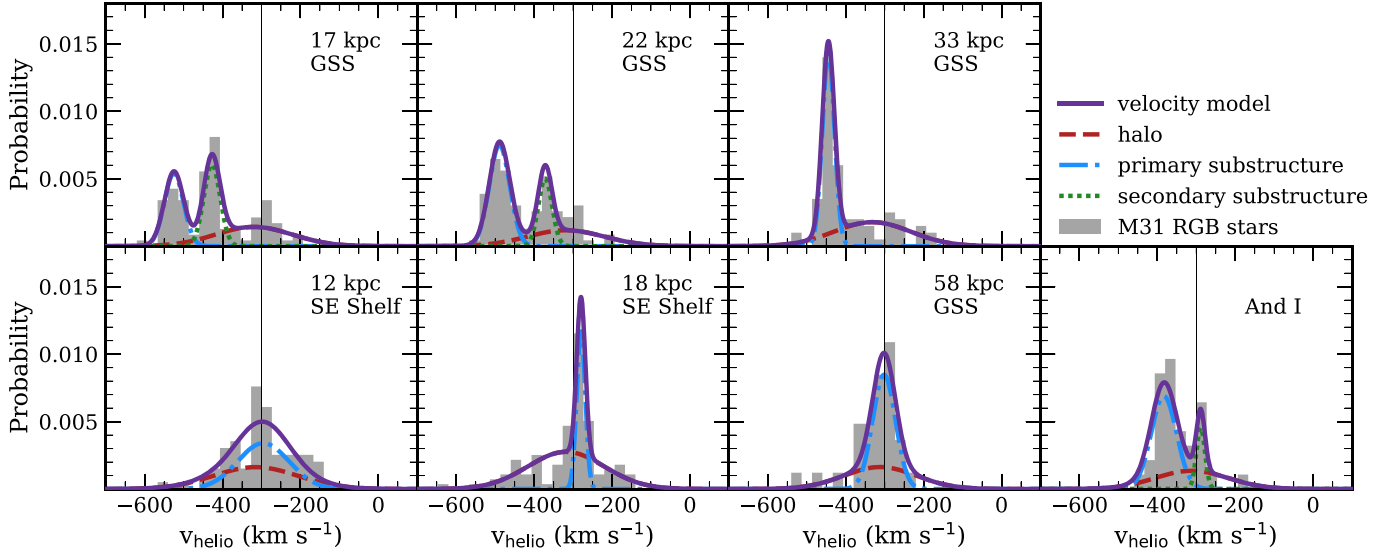


Figure 2. Heliocentric radial velocity distributions (Section 2.2) for M31 RGB stars (gray histograms) in spectroscopic fields with detected kinematic substructure (Table 1; Gilbert et al. 2019; Escala et al. 2020a, 2020b). We show the adopted velocity model for each field (purple solid lines; Gilbert et al. 2018; Escala et al. 2020a), including kinematically hot halo components (dashed red lines) and cold components (dashed-dotted blue lines and dotted green lines) corresponding to primary and secondary substructures. The substructure components present in these fields are the GSS, KCC, and SE shelf. The And I field at 45 projected kpc contains a dwarf galaxy but also overlaps with the GSS (Table 1). RGB stars that are likely And I members are excluded from the field’s velocity distribution.

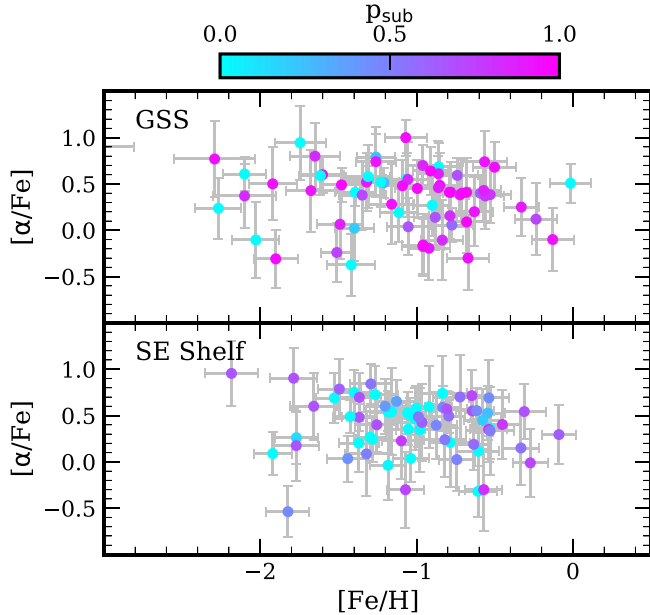


Figure 3. $[\text{Fe}/\text{H}]$ vs. $[\alpha/\text{Fe}]$ for M31 RGB stars in spectroscopic fields spanning the GSS (top panel) and SE shelf (bottom panel; Gilbert et al. 2019, 2020; Escala et al. 2020a, 2020b; J. Wojno et al. 2021, in preparation). Each star is color coded by its kinematically based probability of belonging to substructure (Section 2.2), i.e., stars with $p_{\text{sub}} > 0.5$ ($p_{\text{sub}} < 0.5$) are likely associated with GSS-related tidal debris (the smooth halo).

Escala et al. 2020a). We refer the reader to Escala et al. (2019, 2020a) and Kirby et al. (2008, 2009) for detailed descriptions of the low- and medium-resolution spectral synthesis techniques.

Figure 3 shows $[\alpha/\text{Fe}]$ versus $[\text{Fe}/\text{H}]$ for RGB stars in spectroscopic fields targeting the GSS and SE shelf in M31’s stellar halo (Gilbert et al. 2019, 2020; Escala et al. 2020a, 2020b; J. Wojno et al., 2021, in preparation), where each star is color coded by its probability of belonging to the given substructure component(s) present in each field. These

final samples consist of M31 RGB stars with reliable stellar parameter and abundance measurements that do not show clear evidence of strong TiO absorption in their spectra. Such TiO stars are omitted from the final sample because we did not model absorption from the molecule when generating our grid of synthetic spectra. Furthermore, the size of a potential validation sample of TiO stars that could be used to evaluate the accuracy of these abundance measurements is currently limited. In order to select the final sample of unpublished measurements in the 58 kpc field, we employed our standard criteria ($\delta[\text{Fe}/\text{H}] < 0.4$, $\delta[\alpha/\text{Fe}] < 0.4$ and well-constrained χ^2 contours in each fitted parameter). The only exception is that we used a color cut ($(V - I)_0 < 2$) to exclude possible TiO stars from our final sample in this field, where we have shown that the majority of TiO stars have colors redder than this threshold (e.g., Escala et al. 2020a). Table 1 summarizes the chemical abundance properties of the GSS and SE shelf as probed at the locations of our spectroscopic fields.

3. Chemical Abundance Gradients in the GSS

We measured spatial abundance gradients in the GSS from a sample of 62 M31 RGB stars with $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$ measurements located in fields spanning the feature at 17, 22, and 33 projected kpc (Figure 1, Table 1). As described in Section 3.1, we also considered the impact of a small sample of abundance measurements spanning the GSS in the outer halo at 45 and 58 projected kpc on the spatial gradients. We modeled the gradients by fitting a line to the data, allowing for uncertainties on both the dependent (y) and independent (x) axes (Section 3.5.2). As opposed to describing the line by a slope (k) and intercept (b), we utilized the angle ($\phi = \tan^{-1} k$) and the orthogonal distance of the line from the origin ($b_{\perp} = b \cos \phi$) as model parameters. We used a Markov Chain Monte Carlo ensemble sampler (Foreman-Mackey et al. 2013) to draw from the posterior probability distribution defined by

Table 2
Spatial Abundance Gradients in the GSS (17–33 kpc)

	Slopes			Intercepts		
	r_{proj} (kpc)	m (kpc)	$ n $ (kpc)	r_{proj} (kpc)	m (kpc)	$ n $ (kpc)
	GSS+KCC			GSS+KCC		
[Fe/H]	-0.018 ± 0.003	-0.015 ± 0.003	-0.065 ± 0.012	-0.52 ± 0.08	-0.59 ± 0.06	-0.76 ± 0.04
[Fe/H] _{bias}	-0.026 ± 0.003	-0.021 ± 0.003	-0.081 ± 0.012	-0.25 ± 0.08	-0.36 ± 0.06	-0.61 ± 0.04
$[\alpha/\text{Fe}]$	-0.007 ± 0.007	-0.006 ± 0.006	-0.016 ± 0.024	$+0.57 \pm 0.17$	$+0.54 \pm 0.14$	$+0.45 \pm 0.08$
	GSS			GSS		
[Fe/H]	-0.016 ± 0.004	-0.013 ± 0.003	-0.053 ± 0.013	-0.61 ± 0.09	-0.68 ± 0.07	-0.83 ± 0.04
[Fe/H] _{bias}	-0.022 ± 0.003	-0.018 ± 0.003	-0.071 ± 0.012	-0.36 ± 0.08	-0.46 ± 0.07	-0.67 ± 0.04
$[\alpha/\text{Fe}]$	-0.006 ± 0.007	-0.005 ± 0.006	-0.016 ± 0.026	$+0.52 \pm 0.14$	$+0.51 \pm 0.16$	$+0.44 \pm 0.10$

Note.—Each row and column pair indicates the distance coordinate of the corresponding gradient slope and intercept for a given elemental abundance. The distance coordinates are given as projected radius (r_{proj}), and distance along (m) and across (n) the GSS (Fardal et al. 2006, 2013). We measured gradients both including (top rows) and excluding (bottom rows) the KCC as part of the GSS. We also include values for [Fe/H] gradients measured by incorporating maximal bias estimates owing to selection effects (Section 3.5.1). Note that the $|n|$ -intercepts depend on the adopted zero-point of the GSS-aligned coordinate system.

the log likelihood under this model (Hogg et al. 2010),

$$\ln \mathcal{L} = -\frac{1}{2} \sum_{i=1}^N \left(\frac{\Delta_i^2}{\Sigma_i^2} - \ln(\Sigma_i^2) \right), \quad (1)$$

$$\Delta_i = y_i \cos \phi - x_i \sin \phi - b_{\perp}, \quad (2)$$

$$\Sigma_i^2 = \frac{1}{p_{i,\text{sub}}} (\delta y_i^2 \cos^2 \phi + \delta x_i^2 \sin^2 \phi), \quad (3)$$

where the index i corresponds to a given RGB star with position x_i , abundance ratio y_i , and associated uncertainties (δx_i , δy_i). We employed 10^2 walkers and 10^3 steps for a total of 5×10^4 samples of each parameter when using the latter 50% of each chain. We assumed flat priors on the model parameters (ϕ , b_{\perp}) and incorporated the substructure probability ($p_{i,\text{sub}}$) as an additional weighting term. Thus, the fitting procedure penalizes [Fe/H] and $[\alpha/\text{Fe}]$ measurements for RGB stars that are highly probable members of the kinematically hot stellar halo. Following the conclusion of the fitting procedure, we transformed the marginalized posterior probability distributions back to the more traditional (k , b) parameterization. We adopted the 50th percentiles and 68% confidence intervals of these distributions as the final values and uncertainties for each model parameter.

For our fiducial case, we fit for abundance gradients with respect to projected M31-centric radius (r_{proj}) and defined $p_{i,\text{sub}}$ as the probability that a RGB star belongs to *any* substructure component, inclusive of the KCC. The physical motivation for this approach is the chemical similarity between the GSS and KCC, where current evidence suggests that their [Fe/H] and $[\alpha/\text{Fe}]$ distributions do not differ substantially between 17 and 22 kpc (Gilbert et al. 2009, 2019; Escala et al. 2020a). Our analysis provides further support for this conclusion, where we found that weighting gradient measurements solely toward RGB stars with a high probability of belonging to the GSS produces fully consistent results for the slopes. The gradient intercepts are marginally consistent (within $\sim(1-2)\sigma$) for [Fe/H], where including the KCC results in more metal-rich values for the normalization, and are statistically consistent for $[\alpha/\text{Fe}]$. The abundance gradient slopes and intercepts and their uncertainties are presented in Table 2, where the top panels of

Figures 4 and 5 show the relationship between [Fe/H] and $[\alpha/\text{Fe}]$ and r_{proj} when including and excluding the KCC, respectively, as a contributor to the substructure probability. We measured a relatively steep, negative [Fe/H] gradient as a function of projected radius in the GSS, whereas we did not find evidence of a statistically significant (i.e., inconsistent with zero by at least 3σ) radial $[\alpha/\text{Fe}]$ gradient.

In order to distinguish between abundance gradients present along the high surface brightness core of the GSS and across the GSS envelope, we then transformed the M31-centric coordinates (ξ , η) for each RGB star into a GSS-aligned coordinate system (m , n) defined by Fardal et al. (2006, 2013). This system is described by the unit vectors $\hat{m} = (0.504, -0.864)$ (along the GSS core) and $\hat{n} = (-0.864, -0.504)$ (across the GSS envelope) in the (ξ , η) plane (Figure 1), where we have adopted $(\alpha_{J2000}, \delta_{J2000}) = (00^{\text{h}}42^{\text{m}}44^{\text{s}}, +41^{\text{d}}16^{\text{m}}09^{\text{s}}.0)$ for the position of M31’s center. Fardal et al. (2006) calculated these unit vectors from the slope traced out by the (ξ , η) sky positions of Canada–France–Hawaii Telescope (CFHT) imaging fields targeting the GSS core (McConnachie et al. 2003). We then shifted the center of the GSS-aligned coordinate system from $(m, n) = (0, 0)$ to $(m, n) = (0, 0.34)^{\circ}$ to correspond to the location of the transverse peak of GSS RGB star counts, which Fardal et al. (2013) determined from background-subtracted imaging of M31’s SE quadrant (Irwin et al. 2005). We converted the m and n coordinates from degrees to kiloparsecs using a line-of-sight distance to M31 of 785 kpc (McConnachie et al. 2005). In the subsequent analysis, we present transverse gradients in terms $|n|$ (the absolute coordinate) as opposed to n to clearly reflect trends between the GSS core and envelope, although we plot data points with respect to n to preserve the spatial orientation of the GSS on the sky.

The middle (lower) panels of Figures 4 and 5 show the resulting abundance gradients computed along (across) the GSS, and Table 2 summarizes the relevant parameters. As before, we did not detect statistically significant $[\alpha/\text{Fe}]$ gradients in either dimension of the GSS-aligned coordinate system. We detected negative [Fe/H] gradients both across and along the GSS. This former trend reflects a steep decline in the metallicity between the core and envelopes of the stream as previously observed in photometric metallicities (Ibata et al. 2007; Gilbert et al. 2009). If confirmed, the latter trend would represent the first detection of a significant spectroscopic

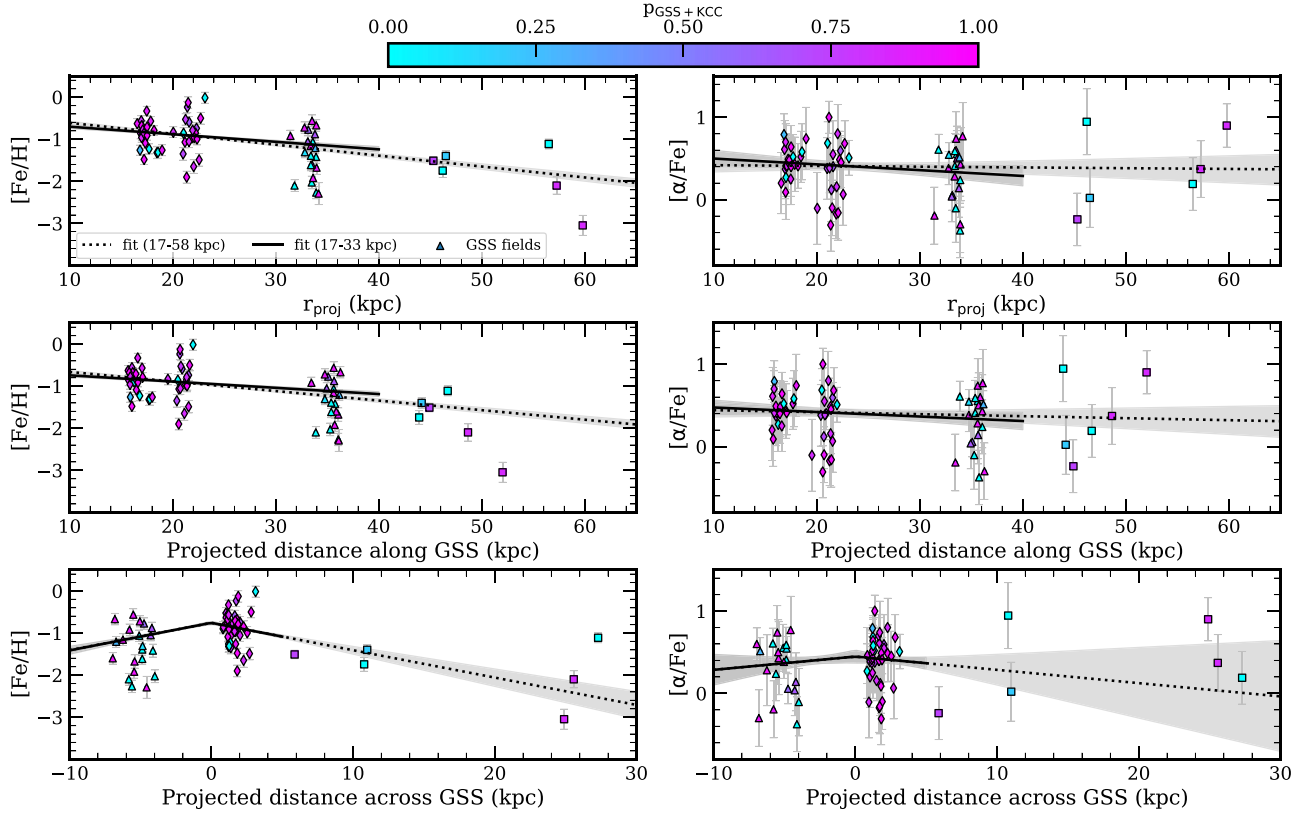


Figure 4. Spatial gradients of $[\text{Fe}/\text{H}]$ (left column) and $[\alpha/\text{Fe}]$ (right column) in the GSS, where the GSS and KCC are treated as a single component. Data points correspond to abundance measurements for M31 RGB stars in spectroscopic fields spanning the GSS (Figure 1, Table 1), where each point is color coded according to its probability of belonging to any given substructure component present in a field. Marker shape (triangle, diamond, square) denotes position across the GSS (eastern edge, core, and western envelope). Solid (dotted) lines and gray envelopes represent gradients measured considering only the inner-halo GSS fields (17–33 kpc) and including the outer-halo GSS fields (17–58 kpc). (Top row) Gradients measured as a function of projected distance from the center of M31. (Middle row) Gradients measured along an axis aligned with the high surface brightness core of the GSS, using the coordinate transformations defined by Fardal et al. (2006, 2013). (Bottom row) Gradients measured perpendicular to the GSS core. The gradients are consistent between including and excluding the outer-halo GSS stars.

$[\text{Fe}/\text{H}]$ gradient along the stream, where this gradient is consistent with the radial $[\text{Fe}/\text{H}]$ gradient within 1σ . Despite this similarity, it is unclear whether the apparent radial gradient is driven primarily by the gradient aligned with *or* transverse to the GSS. In Section 3.3, we show that current data are consistent with the radial $[\text{Fe}/\text{H}]$ gradient originating solely from an intrinsic $[\text{Fe}/\text{H}]$ gradient in only one of these spatial dimensions, where larger samples are required to distinguish between these trends.

3.1. The GSS in the Outer Halo

In order to explore chemical abundance trends in the GSS over a larger projected area, we expanded our analysis of gradients to include six M31 RGB stars with measurements of $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$ (Gilbert et al. 2020; J. Wojno et al., 2021, in preparation; Table 1) present in spectroscopic fields beyond 40 projected kpc that are known to probe the GSS. Figures 4 and 5 show the gradients measured between 17 and 58 projected kpc, which include the outer-halo GSS stars, compared to our fiducial gradients measured between 17 and 33 projected kpc. Including the outer-halo GSS stars results in $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$ gradients with respect to r_{proj} and the GSS-aligned coordinates that are marginally consistent (within 1.6σ) with the parameters in Table 2. Although the sign of the transverse $[\alpha/\text{Fe}]$ gradient changes from negative to positive upon inclusion of the outer GSS stars, each case is consistent with a flat $[\alpha/\text{Fe}]$ gradient within the 1σ uncertainties. Thus, the incorporation of GSS stars beyond

40 kpc suggests that the declining trends of $[\text{Fe}/\text{H}]$ with respect to projected distance across the GSS *and* projected distance along the GSS continue out to the farthest positions probed by our data.

Additionally, we note that there is no statistically significant difference in the gradient slopes between including and excluding the KCC when measuring the gradients between 17 and 58 kpc, where this feature is not present in the line-of-sight velocity distributions (Figure 2) of the 45 and 58 kpc fields. The gradient intercepts maintain marginal consistency (within $(1-2)\sigma$) regardless of inclusion of the KCC, where the most notable change occurs in the normalization of the transverse $[\text{Fe}/\text{H}]$ gradient. In summary, larger samples spanning the GSS in the outer halo are necessary to confirm the identified trends from spectral synthesis-based abundance measurements. We explore whether such trends with $[\text{Fe}/\text{H}]$ persist in $[\text{Fe}/\text{H}]_{\text{phot}}$ measurements of probable GSS stars within our set of spectroscopic fields (Table 1) in Section 3.2.

3.2. Photometric Metallicity Gradients in the GSS

We further investigated spatial trends in the metallicity distribution of the GSS by repeating the above analysis using measurements of $[\text{Fe}/\text{H}]_{\text{phot}}$ for all 270 (339) spectroscopically identified RGB stars excluding (including) the outer-halo GSS fields. We measured $[\text{Fe}/\text{H}]_{\text{phot}}$ by interpolating the color and magnitude of each star on a grid of 9 Gyr PARSEC isochrones (Marigo et al. 2017) with $[\alpha/\text{Fe}] = 0$ as described by Escala et al. (2020a). Figure 6 shows the spatial $[\text{Fe}/\text{H}]_{\text{phot}}$ gradients fit

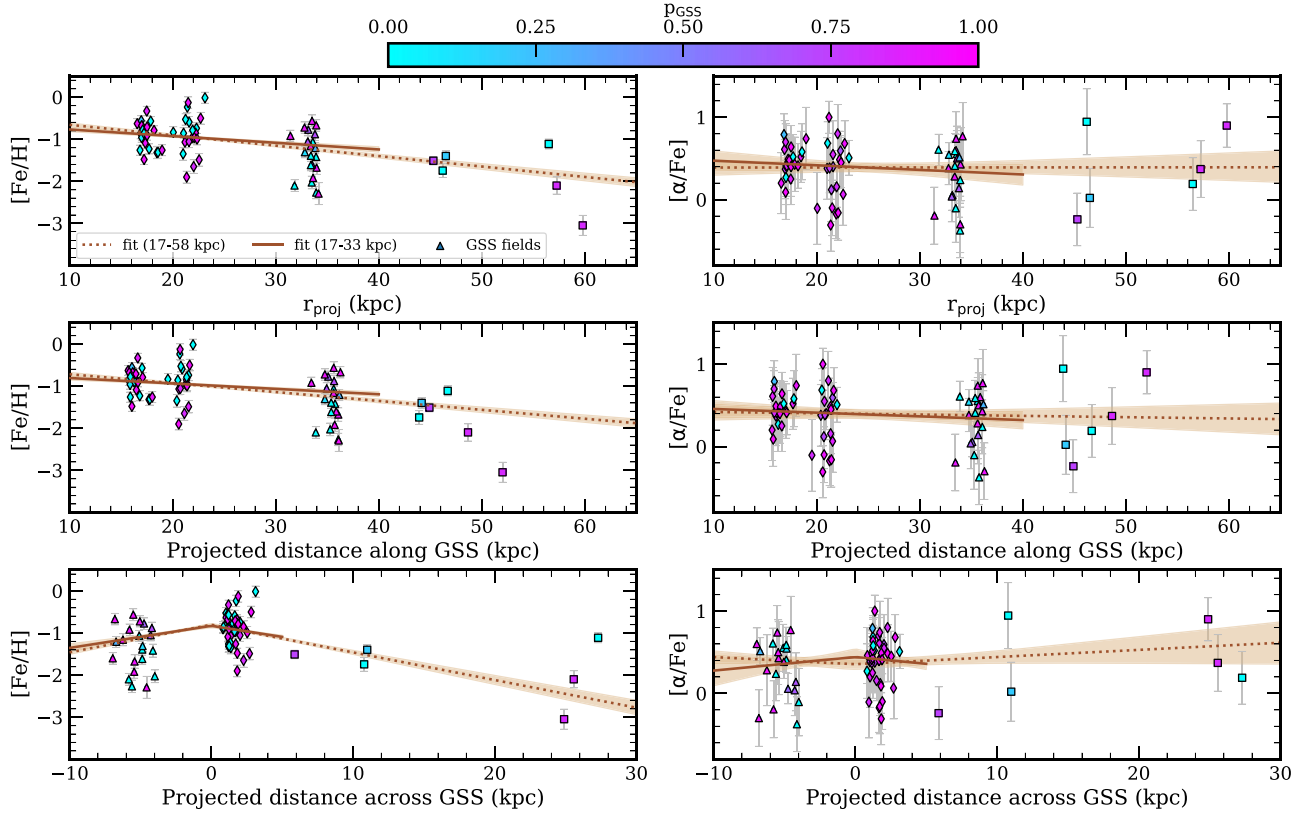


Figure 5. Same as Figure 4, except the GSS and KCC are treated as separate components. All gradients measured between 17 and 33 projected kpc, which exclude the KCC, are statistically consistent with the case of including the KCC (Figure 4). The same is true for gradients measured between 17 and 58 projected kpc.

between 17–33 kpc and 17–58 kpc compared against $[\text{Fe}/\text{H}]_{\text{phot}}$ measurements for the final abundance sample (Section 2.3) and for all spectroscopically identified RGB member stars (Section 2.2) across fields spanning the GSS (Table 1). The abundance sample excludes TiO stars, which have high $[\text{Fe}/\text{H}]_{\text{phot}}$ and includes only stars that have reliable $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$ measurements (Section 2.3). The latter sample includes all RGB members regardless of whether they show spectral TiO signatures or have successful $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$ measurements.

We chose to examine the $[\text{Fe}/\text{H}]_{\text{phot}}$ distribution of all RGB stars along the line of sight to the GSS, as opposed to solely RGB stars in the abundance sample, given that the former sample is larger and less subject to bias against stars with high $[\text{Fe}/\text{H}]_{\text{phot}}$. The spatial $[\text{Fe}/\text{H}]_{\text{phot}}$ trends are qualitatively the same between all member stars and the abundance sample. The most notable difference is that the gradient intercepts are more metal-poor for the abundance sample, which omits TiO stars. We treated the GSS and KCC as the same substructure component, where disregarding the KCC as a contributor to the substructure probability does not result in a significant difference ($\gtrsim 1\sigma$) in the inferred gradients with projected radius and along the GSS core when measured out to 33 or 58 kpc.

As shown in Figure 6, considering only the inner GSS sample yields marginally positive $[\text{Fe}/\text{H}]_{\text{phot}}$ gradients with respect to projected radius and projected distance along and across the GSS. However, when including RGB stars in the outer GSS, the $[\text{Fe}/\text{H}]_{\text{phot}}$ gradients become marginally negative. This suggests that the photometric metallicity along the GSS increases out to ~ 30 –40 projected kpc before decreasing at larger distances. Furthermore, the photometry predicts a minor asymmetry in the

metallicity distribution across the GSS core, where the eastern edge of the GSS appears to have higher $[\text{Fe}/\text{H}]_{\text{phot}}$ than both the core at $n \sim 1$ kpc and the extended western envelope.

The trends gathered from photometry are at odds with those derived from our spectral synthesis–based metallicity measurements (e.g., Figure 4), which show consistently negative $[\text{Fe}/\text{H}]$ gradients along and across the GSS. Given that the $[\text{Fe}/\text{H}]_{\text{phot}}$ gradients measured from the abundance sample show the same qualitative behavior as the sample of RGB members and the photometric and spectroscopic $[\text{Fe}/\text{H}]$ measurements are positively correlated, it is unlikely that biases incurred by selection effects (Section 3.5.1) such as the omission of TiO stars can explain this discrepancy. The probable culprit is the disproportionately metal-rich disparity (0.60 ± 0.11 for the abundance sample) between $[\text{Fe}/\text{H}]_{\text{phot}}$ and $[\text{Fe}/\text{H}]$ measurements in the 33 kpc GSS field compared with other fields. A possible explanation is that the necessary assumptions of constant stellar age and α enhancement to determine CMD-based metallicity estimates are particularly inappropriate for the stellar populations probed in this spectroscopic field. In order to minimize this disparity, we would need to assume both older and α -enhanced isochrones when measuring $[\text{Fe}/\text{H}]_{\text{phot}}$ for this field. Adopting $t = 14$ Gyr and $[\alpha/\text{Fe}] = +0.3$,¹⁰ as opposed to $t = 9$ Gyr and $[\alpha/\text{Fe}] = 0$,

¹⁰ The PARSEC isochrones do not have an α -enhanced option. We chose this isochrone set despite this because of the need for stellar evolutionary models that include molecular TiO in M31’s stellar halo. Thus, we estimated the effect of assuming α -enhanced isochrones on $[\text{Fe}/\text{H}]_{\text{phot}}$ from Gilbert et al. (2014). Using Vandenberg et al. (2006) models, they found that assuming $[\alpha/\text{Fe}] = +0.3$ decreases $[\text{Fe}/\text{H}]_{\text{phot}}$ by ~ 0.2 for M31 RGB stars. Given that this estimate is a relative quantity, it should not depend significantly on the adopted isochrone set.

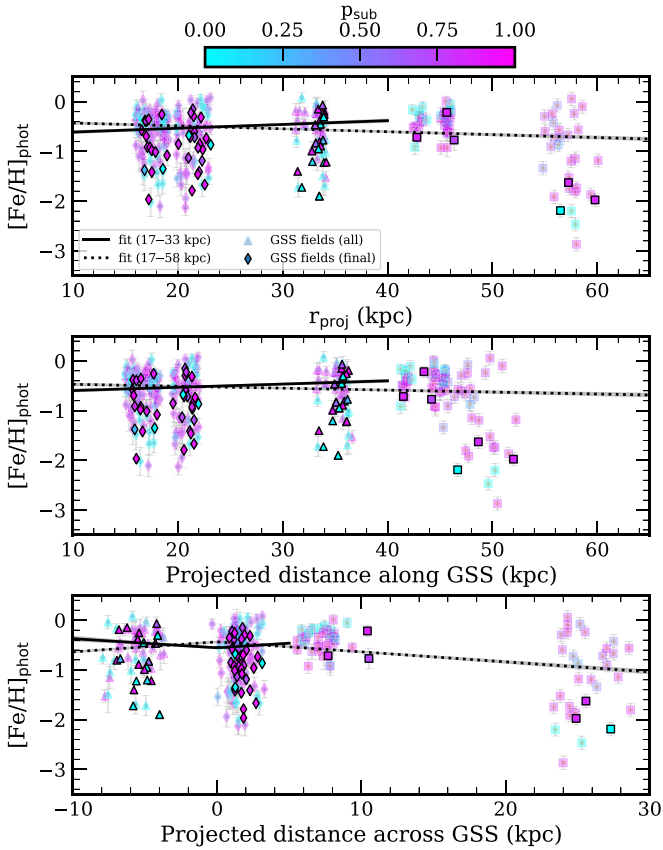


Figure 6. Photometric metallicity ($[\text{Fe}/\text{H}]_{\text{phot}}$) gradients in GSS (Section 3.2) measured from all 270 (339) RGB stars spectroscopically identified as M31 members (Section 2.2) across the GSS fields (Table 1), assuming 9 Gyr PARSEC isochrones (Marigo et al. 2017) with $[\alpha/\text{Fe}] = 0$. We show the fitted gradients including (dotted line) and excluding (solid line) the outer-halo GSS fields. RGB stars excluded from the final abundance sample (opaque points; Section 2.3) are shown as transparent points, where the marker shapes and color coding are the same as Figure 4. The inferred metallicity trends differ between photometric and spectroscopic metallicity measurements even when controlling for selection effects, suggesting that the difference is intrinsic to the measurement methodologies.

reduces the difference between $[\text{Fe}/\text{H}]_{\text{phot}}$ and $[\text{Fe}/\text{H}]$ by ~ 0.3 to ~ 0.2 – 0.4 dex. We emphasize that using different values for the stellar age and α enhancement will not necessarily resolve this metallicity discrepancy, given that the assumption of mono-age and mono- $[\alpha/\text{Fe}]$ stellar populations is unrealistic for GSS stars with a range of stellar ages (Brown et al. 2006; Tanaka et al. 2010) and $[\alpha/\text{Fe}]$ (Gilbert et al. 2019; Escala et al. 2020a).

We refer the interested reader to Escala et al. (2020b) for a detailed discussion of the systematics between spectral synthesis- and CMD-based metallicity measurements in the context of M31’s stellar halo. We compare the spatial metallicity trends implied by both the spectroscopic and photometric metallicity measurements to those from the literature in Section 4.1.

3.3. Apparent Transverse versus Aligned Metallicity Gradients

In Section 3, we measured $[\text{Fe}/\text{H}]$ gradients out to 33 and 58 kpc as a function of projected M31-centric radius (r_{proj}), projected GSS-aligned distance (m), and projected absolute distance orthogonal to the GSS ($|n|$). The statistical consistency of the radial and m gradients (Table 2) prompts the question of whether the observed radial gradients are primarily driven by

the gradients along or across the GSS. The former (latter) case would indicate that there is little to no intrinsic transverse (aligned) $[\text{Fe}/\text{H}]$ gradient but rather that the observed transverse (aligned) gradient is an apparent consequence of an intrinsic GSS-aligned (GSS-transverse) gradient combined with the particular spatial sampling of the spectroscopic fields (Figure 1). Thus, we utilized 5×10^4 pairs of transformed parameters sampled from the posterior probability distribution of our gradient model (Section 3) to infer the expected behavior of an apparent transverse $[\text{Fe}/\text{H}]$ gradient in the GSS by assuming that only an intrinsic aligned gradient is present.

For each slope-intercept pair drawn from the aligned gradient model, we calculated the predicted $[\text{Fe}/\text{H}]$ value at the observed GSS-aligned coordinate of each RGB star and then assigned this value to the corresponding transverse coordinate. Figure 7 shows the 68% confidence intervals for the $[\text{Fe}/\text{H}]$ values predicted from the aligned coordinates alone (green envelopes) compared to the “true” $[\text{Fe}/\text{H}]$ values inferred from the transverse coordinates (gray envelopes). Within both 33 and 58 projected kpc, the GSS-aligned position appears to predict $[\text{Fe}/\text{H}]$ at a given transverse position, although it is less immediately clear in the 58 kpc case. This indicates that the observed radial gradient in our data is most likely driven by the gradient *either* along *or* across the GSS in the inner halo, and tentatively the outer halo, with the caveat that larger samples in the outer halo are required to distinguish between these two trends.

3.4. Relationship to the SE Shelf

Motivated by the multiple lines of evidence for an association between the SE shelf and GSS (Gilbert et al. 2007; Escala et al. 2020a), we compared their chemical abundances, defining the sample for each feature from all stars in fields where it is present (Table 1). We computed $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$ for the GSS and SE shelf via 10^4 bootstrap resamplings of their abundance distributions, weighting by substructure probability and the inverse variance of the measurement uncertainty. Table 3 presents the 50th percentiles of the resulting distributions (and the associated uncertainties from the 16th and 84th percentiles), where we included halo stars in the 45 and 58 kpc fields in our calculations (Section 3.1). In summary, the mean chemical properties of the GSS and SE shelf agree within 1σ , regardless of whether we include or exclude the KCC. There is tentative evidence that the SE shelf is both more metal-rich and α enhanced than the GSS and furthermore that the KCC is more metal-rich than the GSS alone, but larger sample sizes are required to confirm these possibilities.

Figure 8 shows the $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$ distribution functions for the GSS, both including and excluding the KCC, and the SE shelf. We constructed the histograms from all $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$ measurements in spectroscopic fields known to contain a given feature (Table 1), where we utilized substructure probability (Section 2.2) and the inverse variance of the measurement uncertainty as weights. In order to evaluate whether the $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$ distributions are statistically consistent between the GSS and SE shelf, we generated a distribution of p -values using the k -sample Anderson-Darling test. First, we selected all stars that are likely associated with a given feature ($p_{\text{sub}} > 0.5$). We then perturbed the $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$ measurements of these stars by 10^4 random draws from their Gaussian uncertainties and computed the test statistic between the GSS (including the KCC) and SE shelf for each

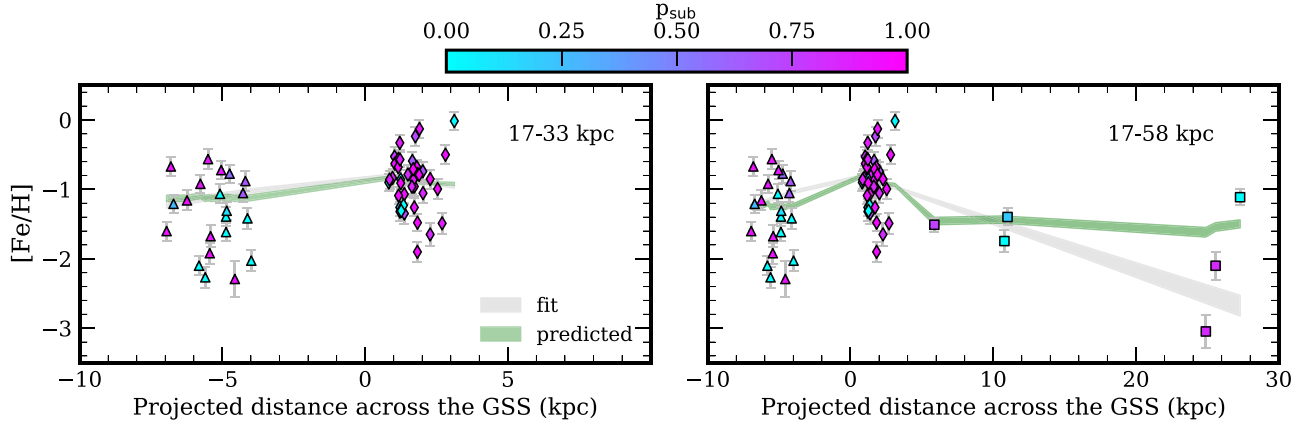


Figure 7. $[\text{Fe}/\text{H}]$ gradient measured out to 33 kpc (left panel) and 58 kpc (right panel) as a function of projected distance across the GSS. The data points and color coding are the same as in Figure 4. The shaded gray envelopes are the 68% confidence intervals of the fitted $[\text{Fe}/\text{H}]$ gradients in the transverse direction (Table 2). The shaded green envelopes represent the apparent transverse gradient assuming that only an aligned gradient is present in the data (Section 3.3). Within both 33 and 58 kpc, a star’s projected distance along the GSS appears to be predictive of $[\text{Fe}/\text{H}]$ regardless of its projected distance across the GSS, indicating that an intrinsic $[\text{Fe}/\text{H}]$ gradient likely exists in only one of these coordinates, although larger samples are needed to distinguish between these trends.

Table 3
Chemical Properties of the GSS and SE Shelf (17–58 kpc)

Comp.	$\langle [\text{Fe}/\text{H}] \rangle$	$\sigma([\text{Fe}/\text{H}])$	$\langle [\alpha/\text{Fe}] \rangle$	$\sigma([\alpha/\text{Fe}])$
GSS+KCC	-0.96 ± 0.06	0.44 ± 0.05	0.40 ± 0.05	0.31 ± 0.03
GSS	-1.03 ± 0.07	0.47 ± 0.05	0.39 ± 0.06	0.34 ± 0.04
SE Shelf	$-0.89^{+0.07}_{-0.08}$	$0.47^{+0.04}_{-0.05}$	$0.45^{+0.04}_{-0.05}$	0.29 ± 0.05

iteration. We found that we could not reject the null hypothesis that $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$ for the GSS and SE shelf are drawn from the same distribution at or below a 10% significance level within the 1σ confidence intervals. This is the case even when adopting a more stringent threshold for substructure membership ($p_{\text{sub}} > 0.75$) or when excluding the KCC as a contributor to the GSS substructure probability.

Based on current measurements, it is therefore feasible for the SE shelf to originate from the same progenitor as the GSS. This is consistent with the finding by Gilbert et al. (2007) that the $[\text{Fe}/\text{H}]_{\text{phot}}$ distributions of M31 RGB stars kinematically associated with the GSS and SE shelf agree when correcting for contamination by the dynamically hot stellar halo, thereby bolstering support for the chemical similarity of the GSS and SE shelf. However, we acknowledge that this apparent similarity between the GSS and SE shelf may be complicated by the presence of spatial $[\text{Fe}/\text{H}]$ gradients in the GSS, which originate in its progenitor (Section 4.2). Such large-scale $[\text{Fe}/\text{H}]$ gradients in the GSS may therefore prohibit the existence of a clear $[\text{Fe}/\text{H}]$ signature for debris related to the merger event, making it more difficult to definitively associate substructure such as the SE shelf with the GSS.

3.5. Sources of Systematic Uncertainty

3.5.1. Sample Selection

The selection criteria for our final sample (Section 2.3) introduce two primary sources of potential bias into our abundance measurements owing to (1) the exclusion of red, presumably metal-rich stars with strong TiO absorption in their atmospheres (b_{TiO}) and (2) S/N limitations, which preferentially affect our ability to measure abundances for metal-poor stars ($b_{\text{S/N}}$). As in Escala et al. (2020a, 2020b), we assessed the impact of these

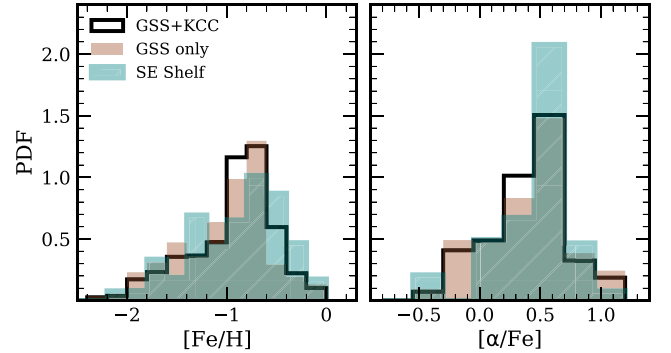


Figure 8. $[\text{Fe}/\text{H}]$ (left) and $[\alpha/\text{Fe}]$ (right) distribution functions in the GSS, including (black outlined histogram) and excluding (shaded brown histogram) the KCC and the SE shelf (shaded blue histogram). We adopted 0.20 (0.25) dex bins for $[\text{Fe}/\text{H}]$ ($[\alpha/\text{Fe}]$), which are comparable to the typical measurement uncertainty, and weighted the histograms by substructure probability and the inverse variance of the measurement uncertainty. The $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$ distributions of the GSS and SE shelf appear to be consistent (Section 3.4).

sources of bias on our measured $[\text{Fe}/\text{H}]$ gradients by shifting each $[\text{Fe}/\text{H}]$ measurement in a given spectroscopic field by its corresponding maximal estimate for $b_{\text{TiO}} + b_{\text{S/N}}$. We determined b_{TiO} from the $\langle [\text{Fe}/\text{H}]_{\text{phot}} \rangle$ difference between all RGB stars and those in the final sample. We computed $b_{\text{S/N}}$ from the $\langle [\text{Fe}/\text{H}] \rangle$ difference between RGB stars with successful $[\text{Fe}/\text{H}]$ measurements (regardless of $[\alpha/\text{Fe}]$) and the final sample.

We include the $[\text{Fe}/\text{H}]$ gradients measured between 17 and 33 kpc with bias estimates ($[\text{Fe}/\text{H}]_{\text{bias}}$) in Table 2. Regardless of whether RGB stars beyond 40 kpc are included, the $[\text{Fe}/\text{H}]$ and $[\text{Fe}/\text{H}]_{\text{bias}}$ gradient slopes are consistent within 2σ . For both excluding or including outer-halo GSS stars, the $[\text{Fe}/\text{H}]_{\text{bias}}$ gradient slopes agree within 1.7σ regardless of whether we treat the GSS and KCC as a single component. However, the normalization of the $[\text{Fe}/\text{H}]_{\text{bias}}$ gradients increases by $\sim(2\text{--}2.5)\sigma$ compared with the $[\text{Fe}/\text{H}]$ gradients as a result of statistically taking into account the red, photometrically metal-rich TiO stars omitted from the final sample. Thus, we can conclude that our findings of negative metallicity gradient slopes with respect to projected radius and along and across the GSS are relatively robust against the exclusion of TiO stars as the dominant source of bias.

The $[\alpha/\text{Fe}]$ gradients are unaffected by S/N limitations as a source of bias. However, they could be affected by the omission of relatively metal-rich TiO stars, assuming a correlation between $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$ in the GSS, such that metal-rich stars tend to be less α enhanced. Based on current data, it is unclear whether this trend is uniformly present among GSS stars in all spectroscopic fields (Figure 3). Escala et al. (2020a) did not find evidence of a statistically significant decline in $[\alpha/\text{Fe}]$ with respect to $[\text{Fe}/\text{H}]$ for either the GSS or the KCC in the 22 kpc field, whereas the characteristic “knee” feature in the $[\alpha/\text{Fe}]$ versus $[\text{Fe}/\text{H}]$ plane may be more apparent (at $[\text{Fe}/\text{H}] \sim -0.9$) in the 17 kpc field (Gilbert et al. 2019). Escala et al. (2020b) additionally found a visible decline in $[\alpha/\text{Fe}]$ with $[\text{Fe}/\text{H}]$ in the 33 kpc field. It is therefore challenging to predict the net impact of excluding TiO stars on the $[\alpha/\text{Fe}]$ gradients for the GSS.

3.5.2. Definition of GSS-Aligned Axes

To determine whether the abundance gradients are robust to different definitions of GSS-aligned coordinate axes, we remeasured the gradients while introducing positional uncertainty terms to the fitting procedure. We did this for all combinations of cases including and excluding the KCC, as well as including and excluding the outer-halo GSS fields. By employing Gaussian fits to imaging data from McConnachie et al. (2003); Font et al. (2006b) found that 80% (1.28σ) of the GSS’s luminosity was contained within $\pm 0.25^\circ$ of the core. Thus, we propagated an error of $\delta\theta = 0.2^\circ$ through our coordinate transformations ($m = \cos(\theta)\xi + \sin(\theta)\eta$, where $\theta \sim -149.8^\circ$ east of north for our adopted coordinate system defined by Fardal et al.),¹¹ which translates to median errors of $\delta m = 0.02$ kpc and $\delta n = 0.07$ kpc. Incorporating these position-dependent errors results in gradient slopes and intercepts that are unchanged within the quoted uncertainties (Table 2).

3.5.3. Distance Variations along the GSS

Early studies of resolved stellar populations in the GSS revealed the three-dimensional structure of the stream, where the line-of-sight distance to the stream increases with increasing projected distance along the stream from the center of M31 (McConnachie et al. 2003). Given that we have assumed a constant distance modulus for all spectroscopic fields (Section 2.3), we assessed the impact of line-of-sight distance variations along the GSS on our measured abundance gradients. We adopted updated distances derived from the CMD position of the tip of the RGB along the GSS (Conn et al. 2016), as probed by the PAndAS survey. Similarly to McConnachie et al., Conn et al. found that the line-of-sight distance to the GSS increases as a function of angular separation from M31, with a distance gradient of 20 kpc per degree over an angular extent of 6° .

Based on these values, the 22 kpc (33 kpc) GSS field (Figure 1; Table 1) is located approximately 23 (38) kpc behind M31, corresponding to an increase of 0.06 (0.10) magnitudes compared to our assumed distance modulus.¹² This translates to a weighted average difference in the photometric effective temperature and surface gravity of $\Delta T_{\text{eff,phot}} = 1.33 \pm 0.79$ (0.89 ± 0.99) K and

$\Delta \log g = -0.02 \pm 0.04$ (-0.04 ± 0.04) dex, respectively, for M31 RGB stars present in the field, using 9 Gyr PARSEC isochrones (Marigo et al. 2017) with $[\alpha/\text{Fe}] = 0$. Assuming values of $T_{\text{eff,phot}}$ and $\log g$ corresponding to the increased heliocentric distance to the GSS yields $\Delta T_{\text{eff}} = 6.36 \pm 2.51$ (0.98 ± 0.40) K, $\Delta[\text{Fe}/\text{H}] = 0.0 \pm 0.12$ (-0.01 ± 0.03) dex and $\Delta[\alpha/\text{Fe}] = -0.01 \pm 0.27$ (0.01 ± 0.07) dex (considering only statistical errors), where the abundance variations are well within our systematic uncertainties.

Thus, the ~ 20 kpc difference in line-of-sight distance between our innermost and outermost GSS fields within 40 projected kpc has a negligible impact on our derived stellar parameters and, consequently, on our measured abundance gradients within this radial range. Gilbert et al. (2009) found comparable results regarding the impact of GSS distance variations on differences in the photometric metallicity between the core and envelope of the stream. Furthermore, Vargas et al. (2014a) performed a supporting analysis, in which they varied the assumed line-of-sight distance to M31 halo stars (by ~ 150 kpc in either direction), and found that it does not alter spectral synthesis-based abundance measurements within their uncertainties.

4. Discussion

4.1. Comparison to Previous Studies

Early spectroscopic studies of individual RGB stars in the GSS at 22 and 33 projected kpc revealed a photometric metallicity difference of 0.09 dex (Guhathakurta et al. 2006; Kalirai et al. 2006), supporting the possibility of metallicity variations in the stream as seen from photometry alone (Ferguson et al. 2002; Ibata et al. 2007; hereafter I07). Using a large sample of photometric metallicities of spectroscopically confirmed GSS stars, Gilbert et al. (2009; hereafter G09) corroborated I07’s core versus envelope metallicity dichotomy (top left panel of Figure 9) by finding that GSS stars located at 17, 22, and 33 projected kpc near the core were more metal-rich by ~ 0.10 (0.53 ± 0.13) dex than GSS stars located at 45 (58) projected kpc (without line-of-sight distance corrections). Gilbert et al. concluded that their defined GSS core has an identical metallicity distribution to the 45 kpc field and is significantly more metal-rich than the envelope as represented by the 58 kpc field.

The G09 fields are nearly identical to those utilized in this work and target the same stellar populations. Indeed, with regard to $[\text{Fe}/\text{H}]_{\text{phot}}$ measurements, we find a difference of -0.10 ± 0.05 (0.56 ± 0.15) dex,¹³ such that the G09 core fields are nearly as metal-rich as the 45 kpc field (more metal-rich than the 58 kpc field). From our spectral synthesis-based metallicity measurements, we find that the G09 core fields are more metal-rich than the 45 (58) kpc fields by 0.63 ± 0.10 (1.62 ± 0.48) dex, which at face value suggests a steeper decline in metallicity between the GSS core and envelope. The difference between the trends predicted by the photometric and

¹¹ The choice of the zero-point for the GSS-aligned axes does not affect the determination of the gradient slopes, although it alters the values of the intercepts.

¹² We omitted the 17 kpc field from this analysis because it is separated from the 22 kpc field by ~ 3 kpc in line-of-sight distance (Conn et al. 2016), corresponding to a difference in distance moduli of only ~ 0.01 mag. We also excluded the GSS fields in the outer halo (Table 1) because our fiducial gradients consider only the inner-halo GSS fields.

¹³ The details of sample selection are the most likely explanation for the slight discrepancy between this work and Gilbert et al. (2009) for the $[\text{Fe}/\text{H}]_{\text{phot}}$ difference between the G09 core fields and the 45 kpc field. We incorporated additional RGB stars published by Kirby et al. (2020). Differences in the assumed isochrone age or model set should not significantly alter relative measures of $[\text{Fe}/\text{H}]_{\text{phot}}$ computed within a given data set. Although G09 considered only RGB stars within $\pm 2\sigma_r$ of the GSS, in contrast to our usage of the KCC-inclusive substructure probability, upweighting likely GSS stars in our analysis would exacerbate the discrepancy because the KCC is more metal-rich than the GSS (Table 3).

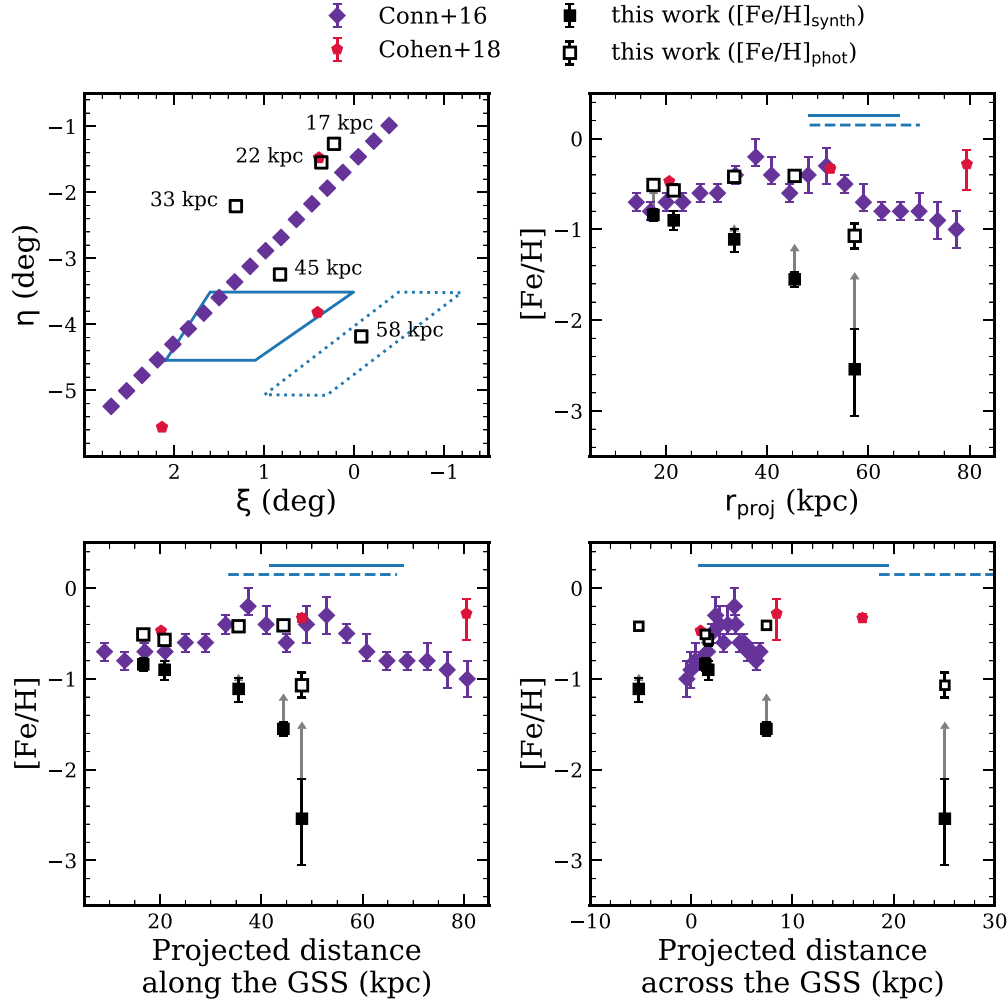


Figure 9. Comparison between this work and previous studies of spatial metallicity variations in the GSS (Section 4.1). (Top left) Approximate locations of photometric (purple diamonds, red pentagons; Conn et al. 2016; Cohen et al. 2018) and spectroscopic (black open squares; this work, Table 1) fields probing the GSS in M31-centric coordinates. The blue polygons denote the “core” (solid lines) and “envelope” (dashed lines) regions of the GSS from Ibata et al. (2007). (Top right, bottom left, bottom right) $[\text{Fe}/\text{H}]$ as a function of projected M31-centric radius, projected distance along the GSS, and projected distance across the GSS. The distance range spanned by the core (envelope) is shown in each panel as a solid (dashed) blue line. $[\text{Fe}/\text{H}]$ refers to photometric metallicity for Conn et al. (2016) and Cohen et al. (2018), whereas we show both spectral synthesis–based ($[\text{Fe}/\text{H}]_{\text{synth}}$; black filled squares; Section 2.3) and photometric ($[\text{Fe}/\text{H}]_{\text{phot}}$; black open squares; Section 3.2) metallicities for the spectroscopic fields. Gray arrows represent maximal bias estimates for $[\text{Fe}/\text{H}]_{\text{synth}}$ in each field (Section 3.5.1).

spectroscopic metallicities cannot be accounted for by field-to-field variations in estimates of the $[\text{Fe}/\text{H}]$ bias resulting primarily from the omission of red TiO stars (Section 3.5.1; Figure 9), which modifies the $[\text{Fe}/\text{H}]$ difference between the core and 45 (58) kpc fields to 0.26 ± 0.10 (0.57 ± 0.48) dex. However, when comparing results from various studies on spatial metallicity variations in the GSS, it is important to acknowledge varying definitions of the stream’s core. For example, the G09 fields that define the GSS core are not spatially colocated with the core from I07 (top panels of Figure 9), where the region spanned by the former (latter) covers ~ 17 –33 (48–66) projected kpc. Thus, the $[\text{Fe}/\text{H}]_{\text{phot}}$ difference examined by I07 primarily reflects orthogonal metallicity variations beyond 40 kpc in the GSS (bottom right panel of Figure 9), whereas the spatial distribution of the G09 fields presents a more complex picture.

Figure 9 provides a view of metallicity variations in the GSS on equivalent spatial footing, as a function of projected radius, GSS-aligned distance, and GSS-transverse distance, while also placing our spectral synthesis–based $[\text{Fe}/\text{H}]$ measurements in the context of the literature (Ibata et al. 2007; Conn et al. 2016;

Cohen et al. 2018). We substituted G09’s $[\text{Fe}/\text{H}]_{\text{phot}}$ measurements with those from this work (Section 3.2) for a similar set of spectroscopic fields (Table 1) for the sake of homogeneity. We transformed the M31-centric coordinates of the imaging fields from Conn et al. (2016; hereafter C16) and Cohen et al. (2018; hereafter C18) and the area spanned by I07’s core and envelope regions into the GSS-aligned coordinate system of Fardal et al. (2006, 2013) for direct comparison with our results.

First, we summarize the methodology and main results of the relevant photometric studies. C16 derived azimuthally averaged RGB metallicities spanning 70 projected kpc along the GSS by modeling PAndAS CMDs as a combination of weighted isochrones and an MW foreground contamination model (Martin et al. 2013). C18 obtained CMD-based metallicities for individual RGB candidates in pencil-beam Hubble Space Telescope/Advanced Camera for Surveys fields from Project AMIGA (Lehner et al. 2020) and Brown et al. (2006) targeting the GSS at 21, 52, and 80 projected kpc. Neither C16 nor C18 correct for contamination of the GSS by M31’s kinematically hot stellar halo, although they show that

the influence of M31’s halo on their results within 50 kpc should not be significant. Both studies found evidence for an increase in $[\text{Fe}/\text{H}]_{\text{phot}}$ with projected distance along the GSS out to $\sim 45\text{--}50$ kpc, after which the behavior of $[\text{Fe}/\text{H}]_{\text{phot}}$ with GSS-aligned distance becomes less certain owing to heavy MW contamination.¹⁴ Thus, it is currently unclear whether CMD-based metallicities predict a plateau or a decline in the GSS-aligned gradient beyond ~ 50 kpc. As for GSS-transverse distance, the range spanned by the C16 and C18 data is limited to that of the I07 core region, where the net $[\text{Fe}/\text{H}]_{\text{phot}}$ trend seems to be at most marginally positive.

Our $[\text{Fe}/\text{H}]_{\text{phot}}$ measurements broadly agree with C16 and C18 between ~ 0 and 10 kpc across the GSS and within ~ 45 kpc along the GSS (Figure 9). However, our results diverge beyond this latter point, where we find an $\sim 0.60\text{--}0.90$ dex lower average metallicity at ~ 50 kpc. Potential reasons for this difference could be (1) unaccounted-for contamination in the PAndAS/Hubble Space Telescope data by red MW dwarf stars with high inferred $[\text{Fe}/\text{H}]_{\text{phot}}$ or (2) issues regarding sample selection and the associated Poisson noise in the sparse outer regions of the GSS. Although neither C16 nor C18 provide constraints in the I07 envelope region, the combination of these measurements with those from this work (and equivalently G09) appear to suggest that the “edge” of the photometrically metal-rich core occurs between ~ 20 and 25 kpc across the GSS (see also C18). However, we have shown that it is unclear whether the core-envelope dichotomy visible from photometric metallicities clearly extends to spectroscopic metallicities based on currently available data (Section 3.3), where we cannot distinguish between an intrinsic gradient along or across the GSS.

Figure 9 also demonstrates that the $[\text{Fe}/\text{H}]$ measurements show an apparent decline with projected distance along the GSS that is inconsistent with the qualitative trends predicted by $[\text{Fe}/\text{H}]_{\text{phot}}$ measurements in this work, C16, and C18. If the radial $[\text{Fe}/\text{H}]$ gradient of the stream is intrinsic to the GSS-transverse distance (and not the GSS-aligned distance; Section 3.3), some of this inconsistency could result from our pencil-beam spectroscopic fields at 33, 45, and 58 projected kpc probing metallicity variations between the core and the envelope rather than those between the inner and outer GSS. However, this cannot entirely explain the discrepancy between trends deduced from CMD-based and spectral synthesis-based metallicities, given that it persists for the fields at 17 and 22 projected kpc near the GSS core. Thus, at least some of this discrepancy is likely fundamental to the measurement methodologies (Section 3.2), where this interpretation is supported by the general similarity between $[\text{Fe}/\text{H}]_{\text{phot}}$ gradients from various studies. As we have previously discussed (Section 3.1, 3.3), additional spectroscopy in the outer GSS is required to provide improved constraints on the stream’s spatial abundance patterns.

4.2. Implications for the GSS Progenitor

Both major and minor merger models for the formation of the GSS broadly reproduce the observed morphological and kinematical features of the stream and its associated shells (Fardal et al. 2006, 2007, 2008, 2013; Mori & Rich 2008; Kirihaara et al. 2014; Sadoun et al. 2014; Kirihaara et al. 2017;

Miki et al. 2016 for minor mergers; Hammer et al. 2010, 2018; Bhattacharya et al. 2019, 2021; D’Souza & Bell 2018 for major mergers). Among minor merger models, rotating, disk progenitors better match the observed asymmetric structure of the GSS than spheroidal counterparts (Fardal et al. 2008, 2013; Kirihaara et al. 2017), although neither class of progenitor models can currently account for the existence of the KCC (Gilbert et al. 2019) or the disturbed nature of M31’s disk (e.g., Quirk & Patel 2020; Dorman et al. 2015; Bernard et al. 2015; Williams et al. 2015). To first order, major merger models explored thus far can simultaneously explain M31’s disk and halo properties, although this does not necessarily disqualify a minor merger from being responsible for the GSS’s formation.

Thus, it is currently unknown whether the GSS progenitor had a stellar mass of $(1\text{--}5) \times 10^9 M_{\odot}$, or $\sim 10^{10} M_{\odot}$, as, respectively, predicted by minor and major merger models (see above references). Current observational constraints on the stellar mass of the GSS progenitor from chemical abundance measurements place it between that of the LMC and M32 ($(1\text{--}5) \times 10^9 M_{\odot}$; Gilbert et al. 2019) when correcting for potential sources of observational bias (Section 3.5.1), which is consistent with predictions of minor merger models for the formation of the GSS. However, Gilbert et al. caution that this cannot be interpreted as direct evidence in favor of a minor merger scenario without knowledge of where the GSS stars originate from in the progenitor, if the progenitor possessed a metallicity gradient. Along with prior studies (Section 4.1), this work has shown that this situation is indeed the case given the observed presence of spatial metallicity gradients in the GSS.

Simulations of minor and major merger scenarios for the formation of the GSS that track stellar metallicity ubiquitously predict the existence of strong gradients in the progenitor in order to approximately match observations (Fardal et al. 2008; Mori & Rich 2008; Miki et al. 2016; Kirihaara et al. 2017; D’Souza & Bell 2018; Hammer et al. 2018). Nonetheless, they differ in the details regarding the exact magnitude of the gradient (but less so in its direction; cf. Miki et al. 2016) and the original location in the progenitor of GSS core stars. For example, some simulations posit that the GSS core is constituted by stars originating near the metal-rich center of the progenitor (Fardal et al. 2008; Miki et al. 2016; Kirihaara et al. 2017), whereas others postulate that the stream debris comes from more metal-poor regions corresponding to a larger radial range within or on the outskirts of the progenitor (Mori & Rich 2008; D’Souza & Bell 2018; Hammer et al. 2018). Thus, an understanding of how the distribution of GSS-related tidal debris on the sky maps to galactocentric radius in the progenitor is crucial for reconstructing the progenitor’s metallicity gradient—and subsequently its average metallicity and inferred stellar mass—from available observational data.

Although the lack of a consensus on the original location of GSS stars in the progenitor limits our ability to directly constrain its metallicity gradient, comparisons between current model predictions and data are informative for identifying potential areas of disagreement. Figure 10 shows CMD-based ($[\text{Fe}/\text{H}]_{\text{phot}}$) and spectral synthesis-based ($[\text{Fe}/\text{H}]$) metallicity measurements for the GSS in our spectroscopic fields (Table 1) as a function of projected radius and azimuthal angle (defined such that 0° is east and the GSS core is located at $\sim 65^\circ$) alongside trends from the models of Kirihaara et al. (2017). Their model assumes a minor merger with a GSS progenitor described by a rotating thick disk with a stellar mass of $7.3 \times 10^8 M_{\odot}$, a central value of $\langle [\text{Fe}/\text{H}] \rangle = -0.5$, and a gradient of $\Delta[\text{Fe}/\text{H}] = -0.5$ in units of disk

¹⁴ C16 note that the MW contamination fraction in their outermost imaging subfields exceeds 80% and may not therefore be representative. C18 similarly comment that the results for their 80 kpc field are highly sensitive to assumptions regarding their adopted MW foreground contamination model.

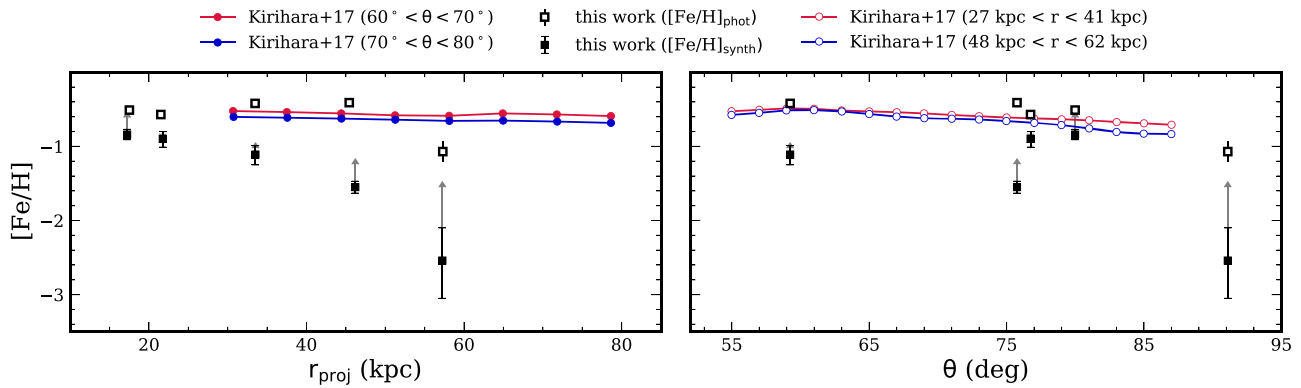


Figure 10. Comparison between observed (this work) and predicted (Kirihiara et al. 2017) radial (left) and azimuthal (right) metallicity variations in the GSS (Section 4.2). The Kirihiara et al. model assumes a minor merger with a GSS progenitor described by a thick disk of stellar mass $M_d = 7.3 \times 10^8 M_\odot$, central $\langle [\text{Fe}/\text{H}] \rangle = -0.5$, and gradient of $\Delta[\text{Fe}/\text{H}] = -0.5$ in units of disk scale length. The azimuthal angle θ is defined such that $\theta = 0^\circ$ is east and the GSS core is located at $\theta \sim 65^\circ$. The left panel shows predicted $[\text{Fe}/\text{H}]$ trends for both the core (red points) and envelope (blue points) of the GSS, whereas the right panel shows $[\text{Fe}/\text{H}]$ trends for inner (open red points) and middle (open blue points) radial regions of the GSS. We include both spectral synthesis-based ($[\text{Fe}/\text{H}]_{\text{synth}}$) and CMD-based ($[\text{Fe}/\text{H}]_{\text{phot}}$) metallicities for our spectroscopic fields. Gray arrows represent maximal bias estimates for $[\text{Fe}/\text{H}]_{\text{synth}}$ in each field (Section 3.5.1).

scale length (where $R_d = 1.1$ kpc). Kirihiara et al. investigated the metallicity patterns in their simulated GSS analog, which resulted from the initial gradient in the progenitor, predicting that the strongest metallicity variations were azimuthal and located at large projected radii (48–62 kpc). Furthermore, stronger gradients in their model translated to more pronounced metallicity differences between the GSS core and envelope, and metallicity differences along the stream were most prominent in its innermost regions.

Although the above scenario could be qualitatively consistent with our measurements, Figure 10 clearly illustrates that this model is not able to provide a quantitative match. Considering only our fields within 33 kpc, the predicted trends are generally too metal-rich for the $[\text{Fe}/\text{H}]$ measurements, even when taking into account $[\text{Fe}/\text{H}]$ bias terms (Section 3.5.1), although it is more similar to the equivalent $[\text{Fe}/\text{H}]_{\text{phot}}$ measurements. Additionally, the observed azimuthal behavior of $[\text{Fe}/\text{H}]$ is more complicated than can be accounted for by the model. Although the former discrepancy could be minimized by assuming a more metal-poor center for the progenitor, a similar effect could presumably be achieved if the GSS core originates from further out in the progenitor’s disk than is the case in this model. Additionally considering fields out to 58 kpc highlights the fact that the observed radial metallicity gradient may be much steeper than that predicted by this model, which could indicate a need for a stronger initial gradient in the progenitor.

Given that few GSS formation models that track stellar metallicity take the additional step of quantifying the predicted abundance ratios (Fardal et al. 2008; Miki et al. 2016; Kirihiara et al. 2017), it is unclear whether they can generally reproduce sufficiently strong gradients in comparison to our spectroscopic and photometric metallicity measurements. From a statistical sample of major merger scenarios for M31’s formation, D’Souza & Bell (2018) found that tidal debris from GSS progenitor analogs exhibited metallicity variations as large as 1 dex but did not further quantify such results.¹⁵ Furthermore, although this class of simulations demonstrate core-envelope dichotomies (Fardal et al. 2008; Mori & Rich 2008; Kirihiara et al. 2017; D’Souza &

Bell 2018), they do not generally predict observed gradients *along* the stream, as may exist in our data. This is excepting the models of Miki et al. (2016), which produced negative radial gradients of approximately $-0.01 \text{ dex kpc}^{-1}$ (compared to $-0.018 \pm 0.003 \text{ dex kpc}^{-1}$; Table 2). In the case of Fardal et al. (2008), the initial gradient in the progenitor is calibrated to the results of Ibata et al. (2007), as opposed to being set by the relationship between its stellar mass and metallicity. In general, current GSS formation models appear to be capable of generating the morphological structure of the stream and its associated shells despite assuming a wide range of mass and metallicity properties for the progenitor (e.g., Hammer et al. 2018), therefore limiting the predictive power of any given modeled metallicity gradient for the GSS.

Additional studies that perform detailed modeling of the GSS metallicity distribution and careful comparisons to observations are therefore needed. In particular, models that also track α elements will be instructive. The lack of significant spatial $[\alpha/\text{Fe}]$ gradients in the GSS (Section 3) suggests that its progenitor may have been uniformly α enhanced or that its $[\alpha/\text{Fe}]$ variations are below the detectable threshold set by our typical measurement uncertainty (i.e., $\lesssim 0.3$). The presence of spatial $[\alpha/\text{Fe}]$ variations in Local Group dwarf galaxies, such as MW dwarf spheroidal satellite galaxies and the Magellanic Clouds (e.g., Kirby et al. 2011; Nidever et al. 2020), has generally not been quantified, thus largely precluding comparisons of observational expectations for $[\alpha/\text{Fe}]$ gradients to the GSS. The exceptions include chemical abundance studies of M31 satellite dwarf galaxies (Vargas et al. 2014b), where no strong evidence for significant $[\alpha/\text{Fe}]$ gradients was found, and Sgr (Hayes et al. 2020). Hayes et al. measured an $[\alpha/\text{Fe}]$ gradient of $(0.1 - 1.2) \times 10^{-3} \text{ dex deg}^{-1}$ (or 0.06–0.15 dex in absolute difference) between the Sgr core and Sgr streams—in addition to weaker internal $[\alpha/\text{Fe}]$ gradients in the streams—that they interpreted as reflecting Sgr’s $[\text{Fe}/\text{H}]$ gradient combined with the characteristic anticorrelation between $[\alpha/\text{Fe}]$ and $[\text{Fe}/\text{H}]$ in dwarf galaxies (e.g., Shetrone et al. 2001; Venn et al. 2004; Kirby et al. 2011). Given that the GSS possesses a significant $[\text{Fe}/\text{H}]$ gradient (Table 2) and shows evidence for a decline in $[\alpha/\text{Fe}]$ with $[\text{Fe}/\text{H}]$ in some spectroscopic fields (Section 3.5.1), the GSS could therefore feasibly exhibit $[\alpha/\text{Fe}]$ variations of $\lesssim 0.3$ dex.

¹⁵ A limitation of current (cosmological, hydrodynamical) simulations that explore major merger scenarios for the GSS’s formation (D’Souza & Bell 2018; Hammer et al. 2010, 2018) is that they are computationally expensive, such that the resolution is necessarily lower than that of N -body minor merger models. Thus, kinematical and chemical structure cannot be resolved in these simulations at the level of detail dictated by the observations.

Regardless of whether an $[\alpha/\text{Fe}]$ gradient exists in the GSS, its high average α enhancement ($+0.40 \pm 0.05$; Table 3) can provide constraints on the nature of the GSS progenitor and thus formation scenarios for the stream. From the first $[\alpha/\text{Fe}]$ measurements of individual RGB stars in the GSS, Gilbert et al. (2019) concluded that the GSS progenitor must have formed stars efficiently enough to enrich to high metallicity ($[\text{Fe}/\text{H}] \sim -0.9$) before experiencing a precipitous decline in its star formation rate such that the yields of Type Ia supernovae dominated over those of core-collapse supernovae. Indeed, the GSS progenitor must have had more efficient star formation than that of the present-day massive dwarf galaxies of the Local Group (Hasselquist et al. 2017; Mucciarelli et al. 2017 for Sgr; Pompéia et al. 2008; Lapenna et al. 2012; Van der Swaelmen et al. 2013; Nidever et al. 2020 for the Magellanic Clouds) or even the dominant progenitor of the MW’s stellar halo (Gaia-Enceladus-Sausage; Haywood et al. 2018; Helmi et al. 2018; Naidu et al. 2020), which have $\langle[\alpha/\text{Fe}]\rangle \lesssim +0.2$ dex. The observed average α enhancement of the GSS—and by extension, its progenitor—is therefore unusual compared to expectations of its stellar mass from dynamical modeling in a minor merger scenario ($M_\star \sim (1-5) \times 10^9 M_\odot$).

Even a scenario in which a massive progenitor dwarf galaxy ($M_\star \sim 10^9 M_\odot$) is accreted sufficiently early to truncate its star formation history on the high $[\alpha/\text{Fe}]$ plateau (e.g., Johnston et al. 2008; Lee et al. 2015) is unlikely to explain the observed $\langle[\alpha/\text{Fe}]\rangle$ of the GSS. Minor merger models for the GSS place its first pericentric passage and accompanying formation of the stream at $\lesssim 1$ Gyr ago (Fardal et al. 2006, 2007, 2008, 2013; Mori & Rich 2008; Kirihaara et al. 2014; Miki et al. 2016), where the cosmologically motivated models of Sadoun et al. (2014) time the initial accretion of the progenitor at ~ 3 Gyr ago. In addition, the most recent star formation in the GSS occurred ~ 4 Gyr ago, where the GSS has a typical stellar age of ~ 8 Gyr (Brown et al. 2006). Thus, a lower-mass progenitor would have produced lower $\langle[\alpha/\text{Fe}]\rangle$ than is observed in the GSS: the progenitor’s star formation would have quenched via interaction with M31’s ionized circumgalactic medium (Lehner et al. 2020) only within the last few gigayears, providing sufficient time for Type Ia supernovae to deplete $[\alpha/\text{Fe}]$ with respect to $[\text{Fe}/\text{H}]$. The extended star formation history of the GSS similarly constrains the scenario of a high-mass progenitor ($M_\star \sim 10^{10} M_\odot$), although the interaction between M31 and the progenitor can begin as early as $\gtrsim 5-10$ Gyr ago in this case (D’Souza & Bell 2018; Hammer et al. 2018).

The most significant difference between a major versus minor merger scenario for the average α enhancement of the GSS is therefore not the accretion time of the event but rather the ability of the progenitor to sustain efficient star formation—and high $\langle[\alpha/\text{Fe}]\rangle$ —over many gigayears such that the progenitor could simultaneously enrich to high $[\text{Fe}/\text{H}]$ (up to at least -0.96 dex; Table 3). Simulations have shown that massive, star-forming MW-like galaxies ($M_\star \sim 10^{(9.7-10.7)} M_\odot$) can produce $[\alpha/\text{Fe}] \sim +0.4$ dex at $[\text{Fe}/\text{H}] \sim -1$ dex (Mackereth et al. 2018; Naiman et al. 2018; Gebek & Matthee 2021), in broad agreement with the abundance ratios observed in the GSS. On the observational front, Gallazzi et al. (2021) recently presented the first measurements of $[\alpha/\text{Fe}]$ in star-forming, massive galaxies ($M_\star \sim 10^{(9.5-11.5)} M_\odot$) beyond the Local Group using $>110,000$ $z=0$ galaxies in SDSS DR7 (Abazajian et al. 2009), confirming that the positive correlation between $[\alpha/\text{Fe}]$ and stellar mass observed for quiescent massive galaxies (e.g., Thomas et al. 2005; Gallazzi

et al. 2006; Conroy et al. 2014; Segers et al. 2016) extends to these systems. However, Gallazzi et al. found that star-forming massive galaxies tend to have lower star-formation-history-integrated $[\alpha/\text{Fe}]$ at a given stellar mass, with a mean value of $[\alpha/\text{Fe}] \sim +0.15$ dex at $M_\star \sim 10^{10.5} M_\odot$ and 1σ upper limits of $\sim +0.3$ dex. Assuming that the average α enhancement of the GSS is representative of the progenitor,¹⁶ the GSS progenitor would be within $\lesssim 1.5\sigma$ of this relation in a major merger scenario (with the caveat that the progenitor halted star formation at $z \sim 0.4$, although it was star forming at the time of accretion). We therefore conclude that a massive GSS progenitor ($M_\star \sim 10^{10} M_\odot$) provides a more natural framework for explaining the high α enhancement and metallicity gradient of the GSS.

5. Summary

The GSS (Ibata et al. 2001b) is the most prominent tidal structure in M31, covering a significant portion of its southeastern quadrant and likely polluting much of its stellar halo (e.g., Brown et al. 2006; Richardson et al. 2008; Gilbert et al. 2009). Until recently, studies of the GSS’s chemical composition were limited to photometric and calcium triplet-based metallicity estimates, where Gilbert et al. (2019) presented the first $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$ abundances in the stream. From an existing sample of 62 RGB stars with measurements of $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$ from the Elemental Abundances in M31 survey (Escala et al. 2019, 2020a, 2020b; Gilbert et al. 2019, 2020; Kirby et al. 2020; Wojno et al. 2020), we have investigated the two-dimensional chemical abundance distribution of the GSS from a set of spectroscopic fields (Table 1) spanning 17–33 projected kpc (Section 3). We have expanded this data set to include $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$ measurements for six additional RGB stars in the western envelope of the GSS (Gilbert et al. 2020; J. Wojno et al., 2021, in preparation) in order to extend our analysis beyond 40 kpc (Section 3.1). We have measured a pronounced negative $[\text{Fe}/\text{H}]$ gradient (-0.018 ± 0.003 dex kpc $^{-1}$; Table 2) and a negligible $[\alpha/\text{Fe}]$ gradient as a function of projected radius in the GSS. Although limited by sample size, the outer GSS data support a continuation of the inner GSS abundance trends. We have also shown that the measured $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$ gradients are largely insensitive to whether the GSS and the KCC (Kalirai et al. 2006; Gilbert et al. 2009, 2019) are treated as a single feature, suggesting that they indeed share a common origin.

The spectroscopic metallicity measurements show evidence for an apparent negative gradient between the inner and outer GSS along an axis defined by the high surface brightness core of the GSS, although it is unclear whether this trend is a manifestation of intrinsic metallicity variations between the core and the envelope of the GSS combined with the spatial sampling of the spectroscopic fields (Section 3.3). Recent photometric metallicity measurements of the GSS show evidence for a *positive* gradient over a similar radial range (Conn et al. 2016). By measuring the photometric metallicity for 339 RGB stars in our spectroscopic fields spanning the GSS (Section 3.2), we have confirmed that $[\text{Fe}/\text{H}]_{\text{phot}}$ trends in our data are similar to those in the literature (Section 4.1) and thus conclude that differences between metallicity patterns predicted

¹⁶ Major merger models for the GSS’s formation predict that the GSS has significant contributions from the more metal-poor outskirts of the progenitor (Hammer et al. 2018; D’Souza & Bell 2018). Assuming that $[\alpha/\text{Fe}]$ declines at high $[\text{Fe}/\text{H}]$, it is therefore possible that the GSS stars are biased toward higher $[\alpha/\text{Fe}]$ relative to the progenitor as a whole.

by spectroscopic and photometric measurements are likely intrinsic to the measurement methodologies.

Although we do not detect a significant $[\alpha/\text{Fe}]$ gradient in the GSS, the high average α enhancement of the feature ($\langle[\alpha/\text{Fe}]\rangle = +0.40 \pm 0.05$; Table 3) argues in favor of an origin in a major merger ($M_* \sim 10^{10} M_\odot$), as opposed to a minor merger ($M_* \sim 10^9 M_\odot$), when combined with constraints regarding its star formation history (Brown et al. 2006) and relatively high mean metallicity ($\langle[\text{Fe}/\text{H}]\rangle = -0.96 \pm 0.06$; Table 3). A massive, disk, star-forming galaxy could enrich to high $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$ (e.g., Gallazzi et al. 2021) by maintaining a high efficiency of star formation for many gigayears (Section 4.2).

In addition, we have demonstrated that the $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$ distributions of the GSS are statistically consistent with those of the SE shelf (Section 3.4; Table 3), a tidal feature predicted by GSS formation models (Fardal et al. 2006, 2007) and subsequently discovered from spectroscopy (Gilbert et al. 2007), thereby providing support for a common origin scenario. However, metallicity gradients originating in the progenitor are a common feature of GSS formation models (Fardal et al. 2008; Mori & Rich 2008; Miki et al. 2016; Kirihaara et al. 2017; Hammer et al. 2018; D’Souza & Bell 2018), such that it is unclear how an initial gradient translates to an observed gradient among the tidal debris (Section 4.2), thus limiting the ability to make chemical connections between features. Future advances in understanding the abundance patterns of the GSS will be instigated by larger samples of $[\text{Fe}/\text{H}]$ and $[\alpha/\text{Fe}]$ measurements in the outer GSS paired with increasingly sophisticated models of its formation.

We thank the anonymous referee, whose careful reading of this paper improved its clarity. We thank Emily C. Cunningham for helpful comments on the manuscript and Stephen Gwyn for reducing the photometry for slitmask f123_1. I.E. was generously supported by a Carnegie-Princeton Fellowship through the Carnegie Observatories. This material is based upon work supported by the NSF under grants Nos. AST-1614081 (E.N.K.), AST-1614569 (K.M.G., J.W.), and AST-1412648 (P.G.). E.N.K. gratefully acknowledges support from a Cottrell Scholar award administered by the Research Corporation for Science Advancement, as well as funding from generous donors to the California Institute of Technology.

We are grateful to the many people who have worked to make the Keck Telescope and its instruments a reality and to operate and maintain the Keck Observatory. The authors wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Maunakea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain.

Facility: Keck (DEIMOS).

Software: Astropy (Astropy Collaboration et al. 2013, 2018), emcee (Foreman-Mackey et al. 2013).

ORCID iDs

Ivanna Escala  <https://orcid.org/0000-0002-9933-9551>
 Karoline M. Gilbert  <https://orcid.org/0000-0003-0394-8377>
 Jennifer Wojno  <https://orcid.org/0000-0002-3233-3032>
 Evan N. Kirby  <https://orcid.org/0000-0001-6196-5162>
 Puragra Guhathakurta  <https://orcid.org/0000-0001-8867-4234>

References

- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, **558**, A33
- Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, *AJ*, **156**, 123
- Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, *ApJS*, **182**, 543
- Barker, M. K., Sarajedini, A., Geisler, D., et al. 2007, *AJ*, **133**, 1125
- Bernard, E. J., Ferguson, A. M. N., Chapman, S. C., et al. 2015, *MNRAS*, **453**, L113
- Bhattacharya, S., Arnaboldi, M., Caldwell, N., et al. 2019, *A&A*, **631**, A56
- Bhattacharya, S., Arnaboldi, M., Gerhard, O., et al. 2021, *A&A*, **647**, A130
- Brown, T. M., Ferguson, H. C., Smith, E., et al. 2003, *ApJL*, **592**, L17
- Brown, T. M., Smith, E., Guhathakurta, P., et al. 2006, *ApJL*, **636**, L89
- Bullock, J. S., & Johnston, K. V. 2005, *ApJ*, **635**, 931
- Cheng, J. Y., Rockosi, C. M., Morrison, H. L., et al. 2012, *ApJ*, **746**, 149
- Chou, M.-Y., Majewski, S. R., Cunha, K., et al. 2007, *ApJ*, **670**, 346
- Choudhury, S., Subramaniam, A., & Cole, A. A. 2016, *MNRAS*, **455**, 1855
- Choudhury, S., Subramaniam, A., Cole, A. A., et al. 2018, *MNRAS*, **475**, 4279
- Clementini, G., Contreras Ramos, R., Federici, L., et al. 2011, *ApJ*, **743**, 19
- Cohen, R. E., Kalirai, J. S., Gilbert, K. M., et al. 2018, *AJ*, **156**, 230
- Collins, M. L. M., Chapman, S. C., Ibata, R. A., et al. 2011, *MNRAS*, **413**, 1548
- Conn, A. R., McMonigal, B., Bate, N. F., et al. 2016, *MNRAS*, **458**, 3282
- Conroy, C., Graves, G. J., & van Dokkum, P. G. 2014, *ApJ*, **780**, 33
- Cooper, M. C., Griffith, R. L., Newman, J. A., et al. 2012, *MNRAS*, **419**, 3018
- D’Souza, R., & Bell, E. F. 2018, *NatAs*, **2**, 737
- Dobbie, P. D., Cole, A. A., Subramaniam, A., et al. 2014, *MNRAS*, **442**, 1680
- Dorman, C. E., Guhathakurta, P., Seth, A. C., et al. 2015, *ApJ*, **803**, 24
- Escala, I., Gilbert, K. M., Kirby, E. N., et al. 2020a, *ApJ*, **889**, 177
- Escala, I., Kirby, E. N., Gilbert, K. M., et al. 2019, *ApJ*, **878**, 42
- Escala, I., Kirby, E. N., Gilbert, K. M., et al. 2020b, *ApJ*, **902**, 51
- Fardal, M. A., Babul, A., Geethan, J. J., et al. 2006, *MNRAS*, **366**, 1012
- Fardal, M. A., Babul, A., Guhathakurta, P., et al. 2008, *ApJL*, **682**, L33
- Fardal, M. A., Guhathakurta, P., Babul, A., et al. 2007, *MNRAS*, **380**, 15
- Fardal, M. A., Guhathakurta, P., Gilbert, K. M., et al. 2012, *MNRAS*, **423**, 3134
- Fardal, M. A., Weinberg, M. D., Babul, A., et al. 2013, *MNRAS*, **434**, 2779
- Ferguson, A. M. N., Irwin, M. J., Ibata, R. A., et al. 2002, *AJ*, **124**, 1452
- Ferguson, A. M. N., Johnson, R. A., Faria, D. C., et al. 2005, *ApJL*, **622**, L109
- Font, A. S., Johnston, K. V., Bullock, J. S., et al. 2006, *ApJ*, **638**, 585
- Font, A. S., Johnston, K. V., Guhathakurta, P., et al. 2006, *AJ*, **131**, 1436
- Foreman-Mackey, D., Hogg, D. W., Lang, D., et al. 2013, *PASP*, **125**, 306
- Freeman, K., & Bland-Hawthorn, J. 2002, *ARA&A*, **40**, 487
- Gallazzi, A., Charlot, S., Brinchmann, J., et al. 2006, *MNRAS*, **370**, 1106
- Gallazzi, A. R., Pasquali, A., Zibetti, S., et al. 2021, *MNRAS*, **502**, 4457
- Gebek, A., & Matthee, J. 2021, arXiv:2102.04561
- Gilbert, K. M., Fardal, M., Kalirai, J. S., et al. 2007, *ApJ*, **668**, 245
- Gilbert, K. M., Guhathakurta, P., Kalirai, J. S., et al. 2006, *ApJ*, **652**, 1188
- Gilbert, K. M., Guhathakurta, P., Kollipara, P., et al. 2009, *ApJ*, **705**, 1275
- Gilbert, K. M., Kalirai, J. S., Guhathakurta, P., et al. 2014, *ApJ*, **796**, 76
- Gilbert, K. M., Kirby, E. N., Escala, I., et al. 2019, *ApJ*, **883**, 128
- Gilbert, K. M., Tollerud, E., Beaton, R. L., et al. 2018, *ApJ*, **852**, 128
- Gilbert, K. M., Wojno, J., Kirby, E. N., et al. 2020, *AJ*, **160**, 41
- Guhathakurta, P., Rich, R. M., Reitzel, D. B., et al. 2006, *AJ*, **131**, 2497
- Hammer, F., Yang, Y. B., Wang, J. L., et al. 2010, *ApJ*, **725**, 542
- Hammer, F., Yang, Y. B., Wang, J. L., et al. 2018, *MNRAS*, **475**, 2754
- Hasselquist, S., Shetrone, M., Smith, V., et al. 2017, *ApJ*, **845**, 162
- Hayden, M. R., Holtzman, J. A., Bovy, J., et al. 2014, *AJ*, **147**, 116
- Hayes, C. R., Majewski, S. R., Hasselquist, S., et al. 2020, *ApJ*, **889**, 63
- Haywood, M., Di Matteo, P., Lehnert, M. D., et al. 2018, *ApJ*, **863**, 113
- Helmi, A. 2020, *ARA&A*, **58**, 205
- Helmi, A., Babusiaux, C., Koppelman, H. H., et al. 2018, *Natur*, **563**, 85
- Ho, N., Geha, M., Tollerud, E. J., et al. 2015, *ApJ*, **798**, 77
- Hogg, D. W., Bovy, J., & Lang, D. 2010, arXiv:1008.4686
- Hopkins, P. F., Cox, T. J., Younger, J. D., et al. 2009, *ApJ*, **691**, 1168
- Ibata, R., Chapman, S., Ferguson, A. M. N., et al. 2004, *MNRAS*, **351**, 117
- Ibata, R., Irwin, M., Lewis, G., et al. 2001b, *Natur*, **412**, 49
- Ibata, R., Irwin, M., Lewis, G. F., et al. 2001a, *ApJL*, **547**, L133
- Ibata, R., Martin, N. F., Irwin, M., et al. 2007, *ApJ*, **671**, 1591
- Ibata, R. A., Lewis, G. F., McConnachie, A. W., et al. 2014, *ApJ*, **780**, 128
- Irwin, M. J., Ferguson, A. M. N., Ibata, R. A., et al. 2005, *ApJL*, **628**, L105
- Johnston, K. V., Bullock, J. S., Sharma, S., et al. 2008, *ApJ*, **689**, 936
- Kacharov, N., Battaglia, G., Rejkuba, M., et al. 2017, *MNRAS*, **466**, 2006
- Kalirai, J. S., Guhathakurta, P., Gilbert, K. M., et al. 2006, *ApJ*, **641**, 268

- Keller, S. C., Yong, D., & Da Costa, G. S. 2010, [ApJ](#), **720**, 940
- Kim, M., Kim, E., Lee, M. G., et al. 2002, [AJ](#), **123**, 244
- Kirby, E. N., Cohen, J. G., Guhathakurta, P., et al. 2013, [ApJ](#), **779**, 102
- Kirby, E. N., Gilbert, K. M., Escala, I., et al. 2020, [AJ](#), **159**, 46
- Kirby, E. N., Guhathakurta, P., Bolte, M., et al. 2009, [ApJ](#), **705**, 328
- Kirby, E. N., Guhathakurta, P., & Sneden, C. 2008, [ApJ](#), **682**, 1217
- Kirby, E. N., Lanfranchi, G. A., Simon, J. D., et al. 2011, [ApJ](#), **727**, 78
- Kirby, E. N., Rizzi, L., Held, E. V., et al. 2017, [ApJ](#), **834**, 9
- Kirby, E. N., Simon, J. D., & Cohen, J. G. 2015, [ApJ](#), **810**, 56
- Kirihara, T., Miki, Y., & Mori, M. 2014, [PASJ](#), **66**, L10
- Kirihara, T., Miki, Y., Mori, M., et al. 2017, [MNRAS](#), **464**, 3509
- Lapenna, E., Mucciarelli, A., Origlia, L., et al. 2012, [ApJ](#), **761**, 33
- Law, D. R., & Majewski, S. R. 2010, [ApJ](#), **718**, 1128
- Leaman, R., Venn, K. A., Brooks, A. M., et al. 2013, [ApJ](#), **767**, 131
- Lee, D. M., Johnston, K. V., Sen, B., et al. 2015, [ApJ](#), **802**, 48
- Lehner, N., Berek, S. C., Howk, J. C., et al. 2020, [ApJ](#), **900**, 9
- Mackereth, J. T., Crain, R. A., Schiavon, R. P., et al. 2018, [MNRAS](#), **477**, 5072
- Marigo, P., Girardi, L., Bressan, A., et al. 2017, [ApJ](#), **835**, 77
- Martin, N. F., Ibata, R. A., McConnachie, A. W., et al. 2013, [ApJ](#), **776**, 80
- McConnachie, A. W., Ibata, R., Martin, N., et al. 2018, [ApJ](#), **868**, 55
- McConnachie, A. W., Irwin, M. J., Ferguson, A. M. N., et al. 2005, [MNRAS](#), **356**, 979
- McConnachie, A. W., Irwin, M. J., Ibata, R. A., et al. 2003, [MNRAS](#), **343**, 1335
- Mercado, F. J., Bullock, J. S., Boylan-Kolchin, M., et al. 2021, [MNRAS](#), **501**, 5121
- Miki, Y., Mori, M., & Rich, R. M. 2016, [ApJ](#), **827**, 82
- Monaco, L., Bellazzini, M., Bonifacio, P., et al. 2007, [A&A](#), **464**, 201
- Mori, M., & Rich, R. M. 2008, [ApJL](#), **674**, L77
- Mucciarelli, A., Bellazzini, M., Ibata, R., et al. 2017, [A&A](#), **605**, A46
- Naidu, R. P., Conroy, C., Bonaca, A., et al. 2020, [ApJ](#), **901**, 48
- Naiman, J. P., Pillepich, A., Springel, V., et al. 2018, [MNRAS](#), **477**, 1206
- Newman, J. A., Cooper, M. C., Davis, M., et al. 2013, [ApJS](#), **208**, 5
- Nidever, D. L., Hasselquist, S., Hayes, C. R., et al. 2020, [ApJ](#), **895**, 88
- Parisi, M. C., Geisler, D., Carraro, G., et al. 2016, [AJ](#), **152**, 58
- Pompéia, L., Hill, V., Spite, M., et al. 2008, [A&A](#), **480**, 379
- Quirk, A. C. N., & Patel, E. 2020, [MNRAS](#), **497**, 2870
- Richardson, J. C., Ferguson, A. M. N., Johnson, R. A., et al. 2008, [AJ](#), **135**, 1998
- Robertson, B., Bullock, J. S., Font, A. S., et al. 2005, [ApJ](#), **632**, 872
- Sadoun, R., Mohayaee, R., & Colin, J. 2014, [MNRAS](#), **442**, 160
- Segers, M. C., Schaye, J., Bower, R. G., et al. 2016, [MNRAS](#), **461**, L102
- Shetrone, M. D., Côté, P., & Sargent, W. L. W. 2001, [ApJ](#), **548**, 592
- Simon, J. D., & Geha, M. 2007, [ApJ](#), **670**, 313
- Tanaka, M., Chiba, M., Komiyama, Y., et al. 2010, [ApJ](#), **708**, 1168
- Thomas, D., Maraston, C., Bender, R., et al. 2005, [ApJ](#), **621**, 673
- Tiede, G. P., Sarajedini, A., & Barker, M. K. 2004, [AJ](#), **128**, 224
- Van der Swaelmen, M., Hill, V., Primas, F., et al. 2013, [A&A](#), **560**, A44
- VandenBerg, D. A., Bergbusch, P. A., & Dowler, P. D. 2006, [ApJS](#), **162**, 375
- Vargas, L. C., Geha, M. C., & Tollerud, E. J. 2014, [ApJ](#), **790**, 73
- Vargas, L. C., Gilbert, K. M., Geha, M., et al. 2014, [ApJL](#), **797**, L2
- Venn, K. A., Irwin, M., Shetrone, M. D., et al. 2004, [AJ](#), **128**, 1177
- Williams, B. F., Dalcanton, J. J., Dolphin, A. E., et al. 2015, [ApJ](#), **806**, 48
- Wojno, J., Gilbert, K. M., Kirby, E. N., et al. 2020, [ApJ](#), **895**, 78