



HALO7D I. The Line-of-sight Velocities of Distant Main-sequence Stars in the Milky Way Halo

Emily C. Cunningham¹ , Alis J. Deason² , Constance M. Rockosi¹, Puragra Guhathakurta¹ , Zachary G. Jennings¹,
Evan N. Kirby³ , Elisa Toloba⁴ , and Guillermo Barro⁴

¹ Department of Astronomy & Astrophysics, University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA; ecunning@ucsc.edu

² Institute for Computational Cosmology, Department of Physics, University of Durham, South Road, Durham DH1 3LE, UK

³ California Institute of Technology, 1200 E. California Boulevard, MC 249-17, Pasadena, CA 91125, USA

⁴ Department of Physics, University of the Pacific, 3601 Pacific Avenue, Stockton, CA 95211, USA

Received 2018 September 3; revised 2019 March 19; accepted 2019 April 4; published 2019 May 10

Abstract

The Halo Assembly in Lambda-CDM: Observations in 7 Dimensions (HALO7D) data set consists of Keck II/DEIMOS spectroscopy and *Hubble Space Telescope*–measured proper motions of Milky Way halo main-sequence turnoff stars in the CANDELS fields. In this paper, we present the spectroscopic component of this data set and discuss target selection, observing strategy, and survey properties. We present a new method of measuring line-of-sight (LOS) velocities by combining multiple spectroscopic observations of a given star, utilizing Bayesian hierarchical modeling. We present the LOS velocity distributions of the four HALO7D fields and estimate their means and dispersions. All of the LOS distributions are dominated by the “hot halo”: none of our fields are dominated by substructure that is kinematically cold in the LOS velocity component. Our estimates of the LOS velocity dispersions are consistent across the different fields, and these estimates are consistent with studies using other types of tracers. To complement our observations, we perform mock HALO7D surveys using the synthetic survey software *Galaxia* to “observe” the Bullock & Johnston accreted stellar halos. Based on these simulated data sets, the consistent LOS velocity distributions across the four HALO7D fields indicate that the HALO7D sample is dominated by stars from the same massive (or few relatively massive) accretion event(s).

Key words: Galaxy: halo – Galaxy: kinematics and dynamics – methods: statistical – techniques: radial velocities

1. Introduction

When a dwarf galaxy falls into the Milky Way (MW) potential and is tidally disrupted, its stars become members of the MW stellar halo. The orbital timescales of these stars are long compared to the age of the Galaxy; thus, long after these debris have lost their spatial association, they remain linked by their kinematic (and chemical) properties. Six-dimensional (6D) phase-space information and chemical abundances of halo stars can therefore be used to unravel the accretion events that have contributed to the mass assembly of the MW. With the Halo Assembly in Lambda-CDM: Observations in 7 Dimensions (HALO7D) survey, we are measuring 3D kinematic information and chemical abundances (as well as constraints on 3D positions) for distant main-sequence (MS) MW halo stars.

The HALO7D data set consists of Keck II/DEIMOS spectroscopy and *Hubble Space Telescope* (*HST*)–measured proper motions (PMs) of distant ($D \sim 10$ – 100 kpc) MW main-sequence turnoff (MSTO) stars. In this paper, the first in the HALO7D series, we present the spectroscopic component of this data set. In a companion paper (Cunningham et al. 2018; hereafter Paper II), we present the PM data set and analysis of the 3D kinematic sample.

This paper is organized as follows. In Section 1.1, we motivate the HALO7D survey and place our survey in context with other MW halo studies. In Section 1.2, we introduce VELOCIRAPTOR, our hierarchical Bayesian method for measuring the LOS velocities for our faint targets. In Section 2, we describe the HALO7D fields, target selection, and observations. In Section 3, we present the details of the VELOCIRAPTOR method. In Section 4, we present the LOS velocity distributions for the four HALO7D fields, estimate their

velocity dispersions, and compare our results with those derived from other tracers. In Section 5, we compare our resulting LOS velocity distributions with predictions from simulations. We summarize our findings in Section 6.

1.1. HALO7D: A Deep, Pencil-beam Complement to Gaia

Our current picture of the stellar halo has largely been shaped by its giant population. Giants and evolved stars, such as red giant branch (RGB) stars, blue horizontal branch (BHB) stars, and RR Lyrae variables, have many advantages as halo tracers, particularly because of their bright absolute magnitudes. Giants have enabled the mapping of the stellar halo out to great distances: Slater et al. (2016) used K giants to measure the density profile out to 80 kpc, Hernitschek et al. (2018) measured the density profile of the MW stellar halo out to 150 kpc with RR Lyrae from Pan-STARRS1, and Deason et al. (2018a) used BHBs in the Hyper Suprime-Cam survey to measure the density profile out to ~ 200 kpc. These tracers have also revealed a wealth of substructure in the distant halo (see, e.g., Sesar et al. 2017; Conroy et al. 2018).

Until recently, our kinematic knowledge of the stellar halo beyond $D \sim 10$ kpc has been limited to one component of motion (the line-of-sight [LOS] velocity) for these bright tracers. While progress has been made on measuring the LOS velocity dispersion profile (e.g., Xue et al. 2008 with Sloan Digital Sky Survey [SDSS] BHBs; Cohen et al. 2017 with RR Lyrae), there has been little knowledge of the tangential motion of these stars until this year—from the second *Gaia* mission data release (*Gaia* Collaboration 2018). Increasing our knowledge of the tangential motions of stars, in order to better map our Galaxy and understand its formation and structure, is the

primary science goal of the *Gaia* mission (Perryman et al. 2001).

Giants and evolved stars represent the upper echelon of a stellar population. While they are bright, they are very rare: MS stars are the dominant population in every stellar population. In addition, it is difficult (perhaps impossible) to uniformly select giants across all age and metallicity populations in the halo. For example, RR Lyrae can only evolve in metal-poor populations, while M giants are only found in metal-rich populations (see Price-Whelan et al. 2015 for a discussion on how the relative numbers of RR Lyrae and M giants in a population can be used to constrain the metallicity of a progenitor). While MS stars are fainter than giants, they are also more numerous, and all populations, regardless of age or metallicity, contain MS stars.

While MS stars are ideal tracers thanks to their presence in all stellar populations, they are challenging to observe because they are faint. Because of its limiting magnitude of $G \sim 20$, beyond $D \sim 15$ kpc in the halo, *Gaia* will not provide PMs for distant halo MS stars. The only instrument currently capable of measuring the PMs of distant ($D > 20$ kpc) MW MS stars is *HST*. *HST* is a powerful instrument for precision astrometry, due to its stability, high spatial resolution, and well-studied geometric distortions and point-spread functions (PSFs; e.g., Anderson & King 2006). Multiepoch *HST* imaging has been exploited to make extremely accurate PM measurements of resolved stellar systems in the Local Group (LG), including the Magellanic Clouds (Kallivayalil et al. 2006a, 2006b, 2013), MW globular clusters (Sohn et al. 2018), MW dwarfs Draco and Sculptor (Sohn et al. 2017), and M31 (Sohn et al. 2012).

The first individual MW stars with measured *HST* PMs were published by Deason et al. (2013b, hereafter D13). These faint stars ($21 < m_{F606W} < 24.5$) had their PMs measured serendipitously during the Sohn et al. (2012) M31 PM study. The third component of the motion for these stars, the LOS velocity, was measured by Cunningham et al. (2016), using the DEIMOS spectrograph on the Keck II telescope, making this sample of 13 stars the first sample of stars with measured 3D kinematics outside the solar neighborhood. D13 and C16 confirmed that we can measure kinematic properties of distant MS stars with these two world-class telescopes. However, these studies were limited to only 13 stars across three *HST* pointings. More stars and lines of sight through the halo are required to use MS star kinematics to investigate the formation of the Galaxy.

The HALO7D survey aims to address the current lack of distant MS stars with measured 3D kinematics. This data set is unique even in the era of *Gaia*, measuring 6D phase-space information for MS stars as faint as $m_{F606W} \sim 24.5$. In order to obtain spectra of these stars with sufficient signal-to-noise ratio (S/N) for LOS velocity and abundance measurements, deep spectroscopy with a large telescope is required. HALO7D complements the *HST* PMs with deep spectra (8–24 hr integrations) observed with Keck II/DEIMOS. However, spectra of individual stars at this depth are unprecedented, and new techniques are required to make measurements with these data. For the interested reader, we motivate this new technique in the next subsection; for those interested in the survey details, please skip ahead to Section 2.

1.2. The Need for VELOCIRAPTOR: Challenges of Deep Slit Spectroscopy

HST can measure PMs for exceedingly faint stars. The deep spectroscopy required to observe to these same magnitudes

presents a significant challenge. To achieve the depth of our survey, targets were observed over multiple nights, and in some cases over years. In order to combine our different spectroscopic observations of a given target into a single measurement of the star’s velocity, we required a new approach that took into account the fact that different observations of the same star will have different velocities. The VELOCIRAPTOR software employs Bayesian hierarchical modeling in order to combine multiple, often noisy, observations of a star, each with different zero-point offsets, into a single posterior probability distribution for the star’s velocity.

One origin of zero-point offset is slit miscentering. Because stars are point sources, they do not fill the full width of the slit during observations (thanks to the exquisite seeing on Maunakea). If the star is not perfectly centered in the slit, the wavelength solution for the object is slightly offset from the wavelength solution given by the calibration of arc lamps (e.g., Sohn et al. 2007; Simon & Geha 2007). This wavelength solution difference corresponds to an apparent velocity shift that can be measured from the velocity of the telluric A-band absorption feature. Velocities of telluric features should be 0 km s^{-1} if the wavelength solution is correct. We refer to this velocity offset as the *A-band correction* (v_{Aband}),⁵ and it is subtracted from the raw velocity (v_{raw} , the velocity of stellar absorption features in an observed spectrum), along with the heliocentric correction (v_{helio} , due to Earth’s motion around the Sun), to yield the corrected velocity in the heliocentric frame:

$$v = v_{\text{raw}} - v_{\text{Aband}} - v_{\text{helio}}. \quad (1)$$

The offset in the slit can be due to astrometric errors or slight mask misalignment. As such, the A-band correction varies from object to object on a given mask and varies from observation to observation of the same object. For a star observed through a $1''$ slit with Keck II/DEIMOS, configured with the 600ZD grating, v_{Aband} can be up to $\approx \pm 60 \text{ km s}^{-1}$. Given that the velocity dispersion of the halo is on the order of 100 km s^{-1} and that velocity dispersions of streams and dwarfs can be less than 10 km s^{-1} , it is essential to take into account this velocity offset before combining spectra from different observations. In addition, when the spectra in question are noisy, the measurements of v_{Aband} are also noisy, and their uncertainties need to be incorporated into the ultimate measure of the corrected velocity uncertainty. In order to best leverage our signal to get accurate velocity measurements of our stars, along with correct uncertainties, we employ Bayesian hierarchical modeling to measure the velocities of individual spectroscopic measurements and their differing zero-point offsets simultaneously.

For the details of the model implemented by VELOCIRAPTOR, we refer the reader to Section 3. Fake data testing and further details are discussed in the Appendix.

⁵ While it has not been treated as such in the literature, we note that the effect of this slit miscentering is closer to a wavelength shift than a velocity shift. In this work, we measure stellar velocities using both the $\text{H}\alpha$ and Ca triplet absorption features. Because these absorption features are approximately equidistant in wavelength from the telluric A-band feature, we can safely treat the A-band correction as a velocity offset, as others have done in the past. However, if one were to measure velocities using only $\text{H}\alpha$ or Ca, or if one were to also use lines farther in the blue or red, then it is better to treat the A-band correction as a wavelength shift.

Table 1
Coordinates of the Four CANDELS Fields Studied in HALO7D

Field	R.A. (J2000)	Decl. (J2000)	l (deg)	b (deg)	Area (arcmin ²)	Catalog References
COSMOS	10:00:28	+02:12:21	236.8	42.1	288	Nayyeri et al. (2017), Muzzin et al. (2013) [†]
GOODS-N	12:36:44	+62:14:24	125.9	54.8	166	G. Barro et al. (2018, in preparation)
GOODS-S	03:32:30	-27:48:11	223.6	-54.4	160	Guo et al. (2013)
EGS	14:15:29	+52:08:19	96.4	60.4	384	Stefanon et al. (2017), Barro et al. (2011) [†]

Note. These fields were chosen for their deep, multiphase *HST* photometry. The listed field area corresponds to the field area covered with multiphase imaging. Catalogs indicated with daggers were used in the secondary target selection; see Section 2.2.1.

2. Data: The HALO7D Keck Program

In this section, we describe the properties of the HALO7D Keck II/DEIMOS spectroscopic program. We describe our choice of survey fields in Section 2.1, our target selection procedure is outlined in Section 2.2, and observations are described in Section 2.4. The extragalactic “piggyback” programs are briefly described in Section 2.3. In Section 2.5, we describe how we selected the halo star candidates used for dynamical modeling from the spectroscopically observed targets.

2.1. Survey Fields

For a 3D kinematic study of distant halo stars, we aimed to survey high-latitude fields that were characterized by many contiguous, multiphase *HST* pointings. Deep, multiphase imaging is required in order to measure PMs of distant MS stars (D13), while the large field of view of DEIMOS enables efficient spectroscopic follow-up of many contiguous *HST* pointings. The fields targeted by the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS; Grogin et al. 2011; Koekemoer et al. 2011; PIs: S. Faber, H. Ferguson) were therefore a natural choice. The *HST* footprints of the CANDELS fields were designed with spectroscopic follow-up in mind, and they have been observed by *HST* many times over the course of *HST*’s operation.

HALO7D surveyed four out of the five CANDELS fields: EGS, COSMOS, GOODS-N, and GOODS-S (the fifth field, UDS, only has one epoch of *HST* imaging). The coordinates of the four HALO7D fields are listed in Table 1; their relatively high latitudes, resulting in minimal foreground contamination from MW disk stars, makes them ideal for both extragalactic and MW halo studies. Figure 1 shows the footprints of the four HALO7D fields. Tiling patterns for one epoch of *HST* imaging are shown in gray; the HALO7D Keck/DEIMOS mask pointings are shown in purple (see Section 2.3).

2.2. Halo Star Candidate Selection

Halo star candidates were selected using u , F606W (broad V filter), and F814W (broad I filter) photometry, from the catalogs listed in Table 1. Star candidates were identified based on image morphology (using the SExtractor parameter `class_star`; Bertin & Arnouts 1996) measured in WFC3 F160W images (the images used for source detection for CANDELS; see catalog references in Table 1 and references therein for more details on source detection and photometry). In order to select as many stars as possible, we used the fairly generous stellerity cut of `class_star` > 0.5 (a more typical stellerity threshold for a study interested in including as many galaxies as possible would require that stars have

`class_star` > 0.98). We also excluded all targets that have nonzero measured redshifts.

To determine our selection boxes for optimally selecting halo stars, we used the Besançon galaxy model (Robin et al. 2003). Figure 2 shows color–magnitude diagrams (CMDs) generated from the Besançon model from a 1 deg² field centered on the coordinates of the EGS field. Green points are disk members, and halo stars are shown in magenta. In order to target as many halo star candidates as possible with minimal disk contamination, we targeted faint, blue stars. Figure 2 shows the HALO7D selection boxes in blue; our highest-priority selection boxes are shown with solid lines, and the dashed line indicates our lower-priority selection box.

Figure 3 shows the CMDs for the four HALO7D fields, with our selection boxes shown in blue. Magnitudes in the F606W and F814W bands are in the STMAG system. The u -band photometry is from ground-based imaging, and magnitudes are in the AB system. In COSMOS and EGS, we used CFHT u -band photometry; in GOODS-S, we used CTIO U -band photometry; and in GOODS-N, we used KPNO U -band photometry (see references in Table 1).

Targets were assigned a priority for selection on a scale from 1 to 4 (with 4 being highest priority) based on the selection boxes:

- Priority 4: Target falls in both solid selection boxes.
- Priority 3: Target falls in one of the solid boxes.
- Priority 2: Target falls in both dashed boxes.
- Priority 1: Target falls in one of the dashed boxes.

2.2.1. Additional Target Selection

In the COSMOS and EGS fields, the CANDELS catalogs (developed from WFC3) did not overlap the full area with multiphase ACS imaging. To select targets in these regions (which contain stars that can have measured PMs), we used additional catalogs. In EGS, we used the ACS F606W/F814W fluxes published in the Barro et al. (2011) photometric catalog and used the same prioritization scheme as described above. Sources in this catalog were identified using IRAC 3.6+4.5 μ m imaging. Stars were identified by combining eight stellerity criteria based on photometric and morphological properties; see Section 3.1 of Barro et al. (2011) for more details. We included all targets with a total sum of stellerity criteria greater than 2, meaning that it was classified as star-like by at least two of the eight stellerity criteria (greater than 3 would be typical, but we again made our selection generous in the interest of not excluding stars).

In COSMOS, we selected targets from the K_s -selected catalog of the COSMOS/UltraVISTA field from Muzzin et al. (2013). Stars are classified in this catalog using $u^* - J$, $J - K_s$ colors; see Section 3.3 and Figure 3 in Muzzin et al. (2013). To

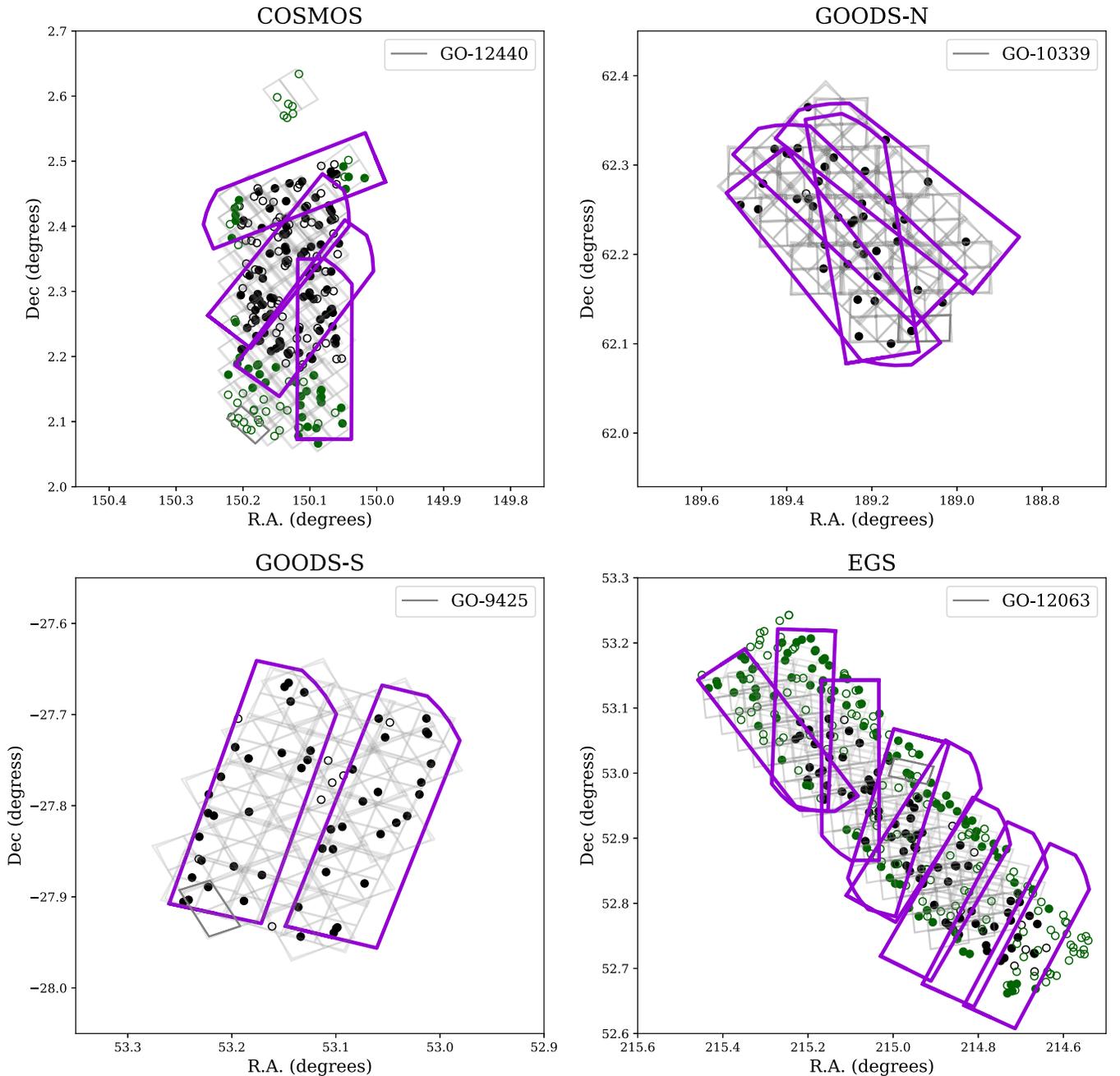


Figure 1. Footprints of the four HALO7D fields. Black points are the positions of halo star candidates selected from the CANDELS catalogs. Filled points were observed with DEIMOS, while open circles denote halo star candidates that were not observed. Green points in EGS and COSMOS are halo star candidates selected from the IRAC and UltraVista catalogs, respectively. DEIMOS mask outlines are drawn in purple. Gray squares indicate one epoch of *HST* imaging that is used for measuring PMs of these same stars (see Paper II).

select HALO7D targets, we used the CFHT *u*-band and Subaru *V*-band fluxes (Capak et al. 2007), using the same selection box as in the top panels of Figure 3.

2.3. Extragalactic Targets

Because the stellar halo is so diffuse, we typically placed only ~ 25 halo star candidates on a given DEIMOS mask. DEIMOS masks can contain up to ~ 150 slits; this provided an opportunity to obtain deep spectra for extragalactic targets as well as Galactic targets. These data have been used to study galactic winds in $z \sim 1$ (Yesuf et al. 2017; H. Yesuf et al. 2019, in preparation), quiescent galaxies at $z \sim 0.7$ (C. Conroy et al.

2019, in preparation), internal galaxy kinematics (W. Wang et al. 2019, in preparation), and dwarf galaxies (Y. Guo et al. 2019, in preparation).

2.4. Observations

Spectra were obtained on the Keck II telescope with the DEIMOS spectrograph (Faber et al. 2003). Observations took place over the course of 3 yr, beginning in 2014 March and ending in 2017 April. While this program was intended to be completed over 19 nights in three semesters of observing, due to poor weather, observations extended through four spring semesters of observations plus several fall nights.

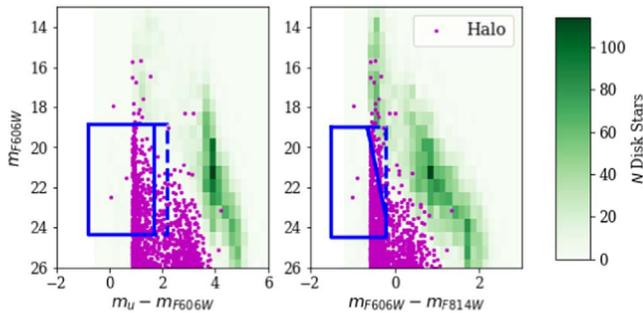


Figure 2. CMDs for stars in the Besançon galaxy model, in a 1 deg^2 field of view centered on the coordinates of EGS. The green density maps show the CMD locations of the disk stars, with the number of disk stars in each CMD bin indicated by the color bar. Halo stars are shown in magenta; only one out of five halo stars is shown for clarity. The HALO7D selection boxes are shown in blue. Stars were assigned priority based on their positions in these two CMDs: stars were assigned top priority if they fell within both solid selection boxes and lowest priority if they fell into only one of the dashed selection boxes.

Observations were conducted with the same DEIMOS configuration as described in C16, and we summarize the key details here. For HALO7D observations, DEIMOS was configured with the 600 line mm^{-1} grating centered at 7200 \AA , resulting in a typical wavelength range of $5000\text{--}9500 \text{ \AA}$. In the interest of limiting flux losses due to atmospheric dispersion, we divided our nights into observing “blocks” of 1–2 hr each and tilted the slits on our masks so that their position angles were consistent with the median parallactic angle of the observing block. Our typical exposure time was 20 minutes.

Our goal was to expose each mask for 8 “effective” hours: we sought to achieve the S/N as a function of apparent magnitude predicted by the DEIMOS exposure time calculator for 8 hr of exposure (gray dashed lines in Figure 4). S/N (computed at $H\alpha$) for each mask as a function of apparent magnitude is shown in Figure 4; in practice, we achieved a typical $\sim 5\text{--}6$ effective hours of exposure on most masks.

DEIMOS mask footprints are shown on top of *HST* pointings in each of the four fields in Figure 1. We observed eight masks in EGS, four masks in GOODS-N, two masks in GOODS-S, and four masks in COSMOS; properties of our observed masks are listed in Table 2. For one of our mask pointings in GOODS-N, we observed the same mask for twice as long as the other masks but switched the list of extragalactic targets after an effective 8 hr was reached (GN3/GN4 masks have the same pointings and MW target lists but different extragalactic targets).

The slit masks were then processed by the *spec2d* pipeline developed by the DEEP2 team at UC Berkeley (Cooper et al. 2012). Table 3 summarizes the progression of the HALO7D sample, from all CMD-identified halo star candidates to stars used for kinematic analysis in the subsequent sections. As seen in Figure 1, we were not able to observe all of our CMD-selected candidates; this is reflected in the difference between Columns (2) and (3) of Table 3. In addition, as with any observational program, we suffered the occasional loss due to errors in the reduction (e.g., the object is too near the edge of the mask, bad columns, etc.; Column (6) of Table 3). There were also targets that did not achieve sufficient S/N in their spectra for a measured velocity (Column (7) of Table 3). Finally, there were also contaminants to our sample of CMD-selected MS stars that we identified spectroscopically; we discuss these contaminants in the next subsection.

2.5. Spectroscopically Confirmed Contaminants

Following observations, all successfully reduced spectra were visually inspected in order to identify obvious contaminants in the MSTO sample. There were two sources of contamination that were identified spectroscopically and removed from the sample: extragalactic contaminants (Column (4) of Table 3) and Galactic disk contaminants (Column (5) of Table 3).

Extragalactic contaminants include quasars and emission-line galaxies; our galaxy contamination rate was the highest among targets selected from the Barro et al. (2011) catalog, for which we had the most generous stellarity cut.

Our spectroscopically confirmed Galactic disk contaminants are white dwarf (WD) stars and red disk stars. We identified two types of WDs in our sample; we found WDs with very broad Balmer features, as well as WDs with strong continua but no absorption features (these objects have disk-like PMs; see Paper II). The red disk stars contain obvious titanium oxide features in their spectra. The red stars in our sample made it into our selection boxes because they are located on the sky close to blue galaxies, which resulted in blended colors for the ground-based u -band photometry used for target selection. While we model contamination from blue disk MS stars in Section 4, we exclude the WD and red stars from our sample for dynamical modeling. In the following section, we describe in detail how we measure LOS velocities for our target spectra. To skip straight to the results, we refer the reader to Section 4.

3. Hierarchical Bayesian LOS Velocities: Velociraptor

In this section, we describe in detail the model implemented by the VELOCIRAPTOR software. As explained in Section 1.2, different observations of the same star will have different raw velocities, due to the motion of Earth around the Sun, as well as slit miscentering. We demonstrate this effect in Figure 5, which shows the $H\alpha$ region and the telluric A -band region for two spectra of a relatively bright HALO7D target ($m_{F606W} = 19.1$) observed on different nights. The raw spectra are clearly not at the same velocity: while $\sim 5 \text{ km s}^{-1}$ of this velocity offset is due to Earth’s motion around the Sun, the remaining 50 km s^{-1} offset is entirely due to the misalignment of the object in the slit.

As such, applying these corrections prior to co-adding or stacking spectra is essential in order to accurately estimate the velocity of a star. However, because the A -band correction is measured from an absorption feature, if the spectrum is faint and noisy, the estimate of the A -band correction will also be noisy.

In order to address these challenges, we present the VELOCIRAPTOR technique. VELOCIRAPTOR implements a Bayesian hierarchical model, modeling the raw velocities and A -band corrections of all observations of a star simultaneously. Standard practice would be to stack spectra and then measure a velocity, usually using a cross-correlation method (e.g., *spec1d*; Newman et al. 2013) or a maximum likelihood method (such as the Penalized Pixel Fitting method of Cappellari & Emsellem 2004). However, stacking before measuring a velocity neglects the A -band corrections (and associated uncertainties) of different observations. Bayesian hierarchical modeling provides a natural, fully probabilistic framework for incorporating all available information in the spectra while properly accounting for uncertainties.

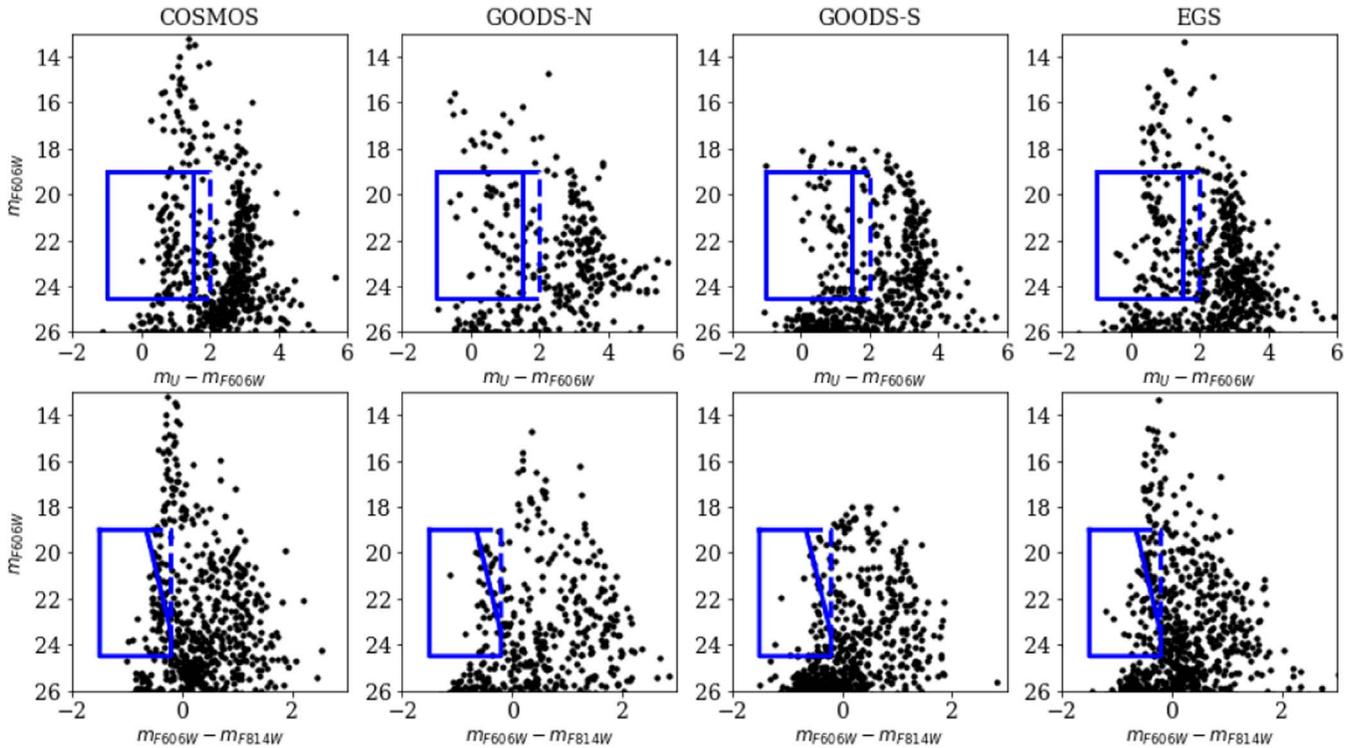


Figure 3. CMDs for stars in the four HALO7D fields. Selection boxes are shown in blue. Stars were assigned priority based on their positions in these two CMDs: stars were assigned top priority if they fell within both solid selection boxes and lowest priority if they fell into only one of the dashed selection boxes. Magnitudes in F606W and F814W are computed in the STMAG system. We note that the bright stars in GOODS-S have been masked out in the catalog used for selection.

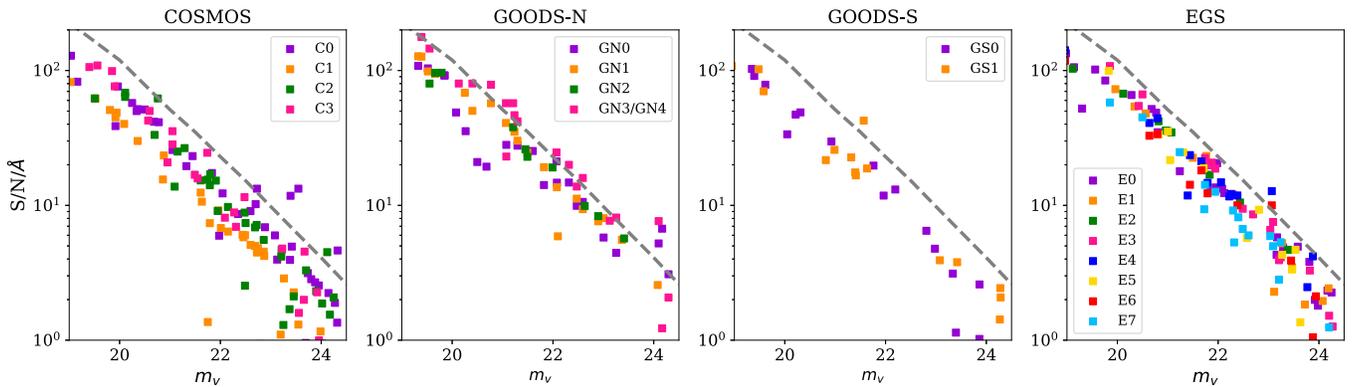


Figure 4. S/N per angstrom for all HALO7D masks, as a function of V-band magnitude. Gray dashed lines indicate the predicted S/N with 8 hr of exposure time.

In Section 3.1, we define terminology used to describe our model. We then explain how we model individual spectroscopic observations in Section 3.2. Section 3.3 describes the hierarchical model employed to measure the velocity of a star from multiple observations. Details of fake data testing, including sample trace and corner plots, can be found in the Appendix.

3.1. Definitions

We begin by explicitly defining some terminology and symbols used in the description of our method.

Spectral Regions: A segment of a spectrum, a few hundred angstroms in wavelength, centered on an absorption feature. We use the letter j to denote a specific spectral region; for example, $F_j(\lambda)$ denotes the flux as a function of wavelength for spectral region j .

Pixels: Each spectral region contains pixels $i = 1, \dots, I$. The value I_j denotes the total number of pixels in spectral region j .

$\lambda_i \rightarrow x_{ij}$: For evaluating polynomials, we rescale the wavelength array of a given spectral region j onto the range $[-1, 1]$:

$$x_{ij}(\lambda_i) = \frac{2(\lambda_i - \lambda_{\min,j})}{\lambda_{\max,j} - \lambda_{\min,j}} - 1. \quad (2)$$

Hence, x_{ij} denotes the value that λ_i takes when rescaled onto the range determined by the range of spectral region j : $[\lambda_{\min,j}, \lambda_{\max,j}]$. Using this definition, $x_{0,j} = x(\lambda_{\min,j}) = -1$, and $x_{I,j} = x(\lambda_{\max,j}) = 1$.

Observations: Each spectrum has $k = 1, \dots, K$ observations. Therefore, each observation has its own $v_{\text{raw},k}$, $v_{\text{Aband},k}$, $v_{\text{helio},k}$.

Distributions: We use standard statistical notation to express random variable distributions, which will include normally distributed ($x \sim N(\mu, \sigma^2)$) and Gamma-distributed

Table 2
Summary of the Masks Observed through HALO7D

Field	Mask Name	R.A. (J2000)	Decl. (J2000)	Mask P.A. (deg)	Semesters Observed	MW Targets	Extragalactic Targets
COSMOS							
	C0	10:00:36.50	+02:20:47.8	−38.2	2014A, 2015A	39	83
	C1	10:00:31.52	+02:16:14.6	−36.2	2017A	38	80
	C2	10:00:23.41	+02:11:54.4	−0.20	2017A	38	74
	C3	10:00:29.45	+02:26:16.0	111.8	2016A	23	81
GOODS-N							
	GN0	12:37:08.33	+62:12:44.6	−142.0	2015A, 2016A	23	95
	GN1	12:37:01.22	+62:14:05.3	46.1	2014A, 2015A	24	105
	GN2	12:36:38.82	+62:15:48.8	51.3	2016A	12	95
	GN3,GN4	12:36:58.73	+62:13:02.6	9.4	2014A, 2015A, 2016A	23, 21	104, 94
GOODS-S							
	GS0	03:32:18.81	−27:49:04.9	−17.3	2015B, 2016B	25	92, 88, 90
	GS1	03:32:47.28	−27:47:26.8	−15.6	2015B, 2016B	24	85, 85, 85
EGS							
	E0	14:20:15.98	+53:01:13.9	180.0	2014A	30	101
	E1	14:18:48.21	+52:45:18.4	−26.1	2014A	19	103
	E2	14:19:33.70	+52:49:40.4	−28.1	2014A	24	97
	E3	14:19:51.37	+52:55:04.1	−30.4	2015A	28	81
	E4	14:20:42.48	+53:05:50.4	177.8	2015A	34	80
	E5	14:19:10.83	+52:47:16.3	−25.8	2016A	22	81
	E6	14:21:03.34	+53:04:42.3	−144.7	2017A	29	87
	E7	14:19:49.96	+52:56:05.3	164.0	2017A	31	79

Note. In GOODS-N, the masks GN3 and GN4 contain the same MW targets but different extragalactic targets. In GOODS-S, three sets of extragalactic targets were observed alongside each mask of MW targets.

Table 3
Summary of the Progression of the HALO7D Sample, from CMD-selected Targets to Objects Used in Kinematic Analysis

Field	N_{CMD}	N_{Obs}	N_{Gal}	N_{WD}	N_{RedStars}	$N_{\text{ReductionErrors}}$	$N_{\text{Low S/N}}$	N Halo Star Candidates
COSMOS	101, 67	87, 36	5, 2	1	2	4	21	88
GOODS-N	48	47	1	2	4	1	6	33
GOODS-S	57	49	8	2	1	2	11	25
EGS	96, 135	84, 94	6, 28	7	7	8	25	97

Note. In COSMOS and EGS, we first indicate halo star candidates selected from the CANDELS catalogs, followed by candidates from the secondary catalogs for the first three columns of the table (see Section 2.2.1).

($x \sim \text{Gamma}(a, b)$). We also make use of inverse-gamma distributions in this work; the inverse-gamma distribution is the conjugate prior for the variance of a normal distribution. If a random variable x is distributed according to an inverse-gamma distribution with parameters a, b (i.e., $x \sim \text{Inv} - \text{Gamma}(a, b)$), the corresponding probability density function (pdf) is

$$p(x|a, b) = \frac{b^a}{\Gamma(a)} x^{-(a+1)} \exp\{-b/x\}, \quad (3)$$

where $\Gamma(a)$ denotes the gamma function.

3.2. Single Mode: Modeling a Single Spectrum

We first present our Bayesian method of estimating the velocity of a star from a single spectrum. To estimate a stellar velocity, we use the spectral regions that contain the most velocity information. We typically use three regions: the region around $\text{H}\alpha$ (6500–6650 Å) and the calcium triplet region (CaT;

8450–8700 Å) to estimate the raw velocity, and the telluric A -band region (7500–7750 Å) for the A -band correction.

To model a region of the spectrum, we first begin with a template. Our templates consist of bright velocity standards that were observed with a very similar configuration to our science spectra. The templates used in this analysis are described in detail by Toloba et al. (2016); they have high S/Ns (100–800 Å^{−1}) and span a range of spectral types (from B1 to M8) and luminosity classes (from dwarfs to supergiants).

For the HALO7D targets, we use the template HD 105546. While this star is a horizontal branch star, its color is consistent with the color range of our targets, and its spectrum has absorption in $\text{H}\alpha$ and CaT. In order to estimate the raw velocity of our template star, we use a simple model of a polynomial with inverted Gaussians for the absorption lines. Because the template star was trailed through the slit during observation, it does not suffer from slit miscentering, so its A -band correction is 0 km s^{−1}. We verify that no additional correction to the

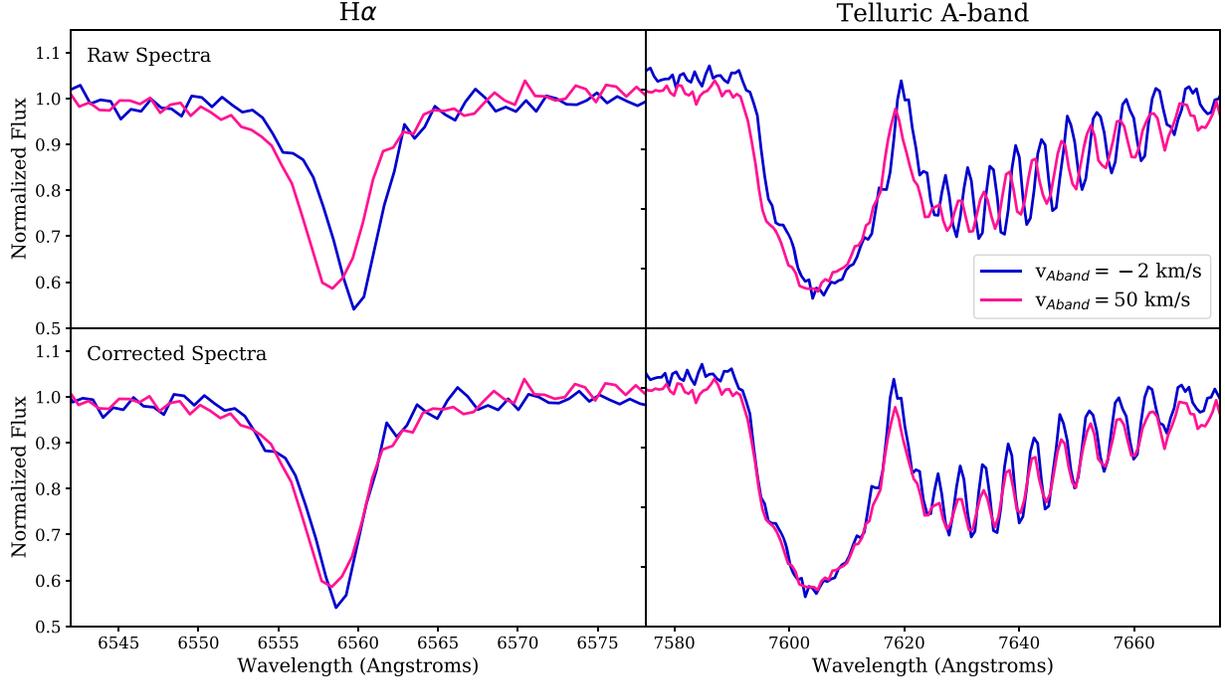


Figure 5. Illustration of the velocity offset caused by slit miscentering. Blue and pink lines show spectra of the same star taken during different observing runs; left panels show the H α region of the spectrum, and right panels show the telluric A-band region. The top panels show the raw spectra, in the observed frame, uncorrected for heliocentric motion and slit miscentering; the H α lines and telluric absorption lines are clearly misaligned. The bottom panels show the spectra with heliocentric and A-band corrections applied.

wavelength solution is required by checking the consistency of the velocities measured at H α and CaT.

We use the spectrum of HD 105546, shifted to the rest frame, to estimate the velocity of H α and CaT regions of the HALO7D target. We use the same spectrum in the observed frame (i.e., unshifted) to estimate the A-band velocities of the HALO7D targets. We model the different regions of the target spectrum separately, while demanding that the velocities at H α and CaT be the same.

To model a spectral region, we allow the velocity, absorption-line strength, and continuum level to vary. Our vector of parameters, which we denote as θ_{Line} , are the velocity v_{Line} , the absorption-line strength C , and the Legendre polynomial coefficients b_l that control the continuum level. Given that we look at narrow spectral regions, we find $l = 1$ (i.e., a straight line with varying slope and intercept) sufficient to model the continuum.

As a function of scaled wavelength x , our model $M(x, \theta_{\text{Line}})$ can be written as

$$M_{\text{Line}}(x, \theta_{\text{Line}}) = \sum_l b_l P_l(x) \times \frac{T_{v_{\text{Line}}}(x) + C}{1 + C}, \quad (4)$$

where P_l are the Legendre polynomials and $T_{v_{\text{Line}}}(x)$ is the template flux, shifted to velocity v_{Line} .

The likelihood of the observed spectral flux F_{Line} given the model parameters is thus

$$p(F_{\text{Line}}|\theta_{\text{Line}}) = \prod_{i=0}^I \text{N}(F_{\text{Line}}(x_i)|M_{\text{Line}}(x_i, \theta_{\text{Line}}), \sigma_i^2), \quad (5)$$

where, for $i = 1, \dots, I$ pixels, x_i is the rescaled wavelength value, $F_{\text{Line}}(x_i)$ is the flux at that rescaled wavelength, $M(x_i, \theta_{\text{Line}})$ is the model flux, and σ_i is the noise in that pixel (as returned by the spec2d pipeline).

We can write down the posterior probability distribution for our model parameter making use of Bayes's theorem, which gives the probability of a vector of model parameters θ given a vector of data y :

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)}, \quad (6)$$

where $p(y|\theta)$ is the likelihood of the data given the parameters, $p(\theta)$ is the prior probability of the parameters, and $p(y)$ is the probability of the data (in practice, this term serves as a normalization). In order to sample from the posterior distribution for our model parameters θ_{Line} , we must specify their prior distributions: $p(\theta_{\text{Line}}|F_{\text{Line}}) \propto p(F_{\text{Line}}|\theta_{\text{Line}})p(\theta_{\text{Line}})$.

Our prior distributions are listed in Table 4; we generally assume reference (i.e., Jeffreys) priors. For the absorption-line parameters C , we assign a Gamma distribution prior to the quantity $C + 1$. The Gamma distribution is defined over the range $x > 0$ and is a common prior choice for scale parameters. Because of our chosen parameterization, if $C < -1$, the line becomes an emission line. In addition, as C becomes large, the absorption line becomes indistinguishable from the continuum. We therefore assign a Gamma(2, 2) prior to $C + 1$ in order to constrain the possible allowed values for C .

The total posterior probability of the full vector of spectrum parameters θ_{spec} is given by the product of the posterior probabilities of the different lines used:

$$p(\theta_{\text{spec}}|\mathcal{F}) \propto \prod_{j=1}^J p(F_{\text{Line},j}|\theta_{\text{Line}})p(\theta_{\text{Line}}), \quad (7)$$

where $\theta_{\text{spec}} = (v_{\text{raw}}, v_{\text{A-band}}, C_1, \dots, C_J, b_{l,1}, \dots, b_{l,J})$ is the full vector of parameters describing the spectrum. Here we are denoting $\mathcal{F} = \{F_{\text{Line},1}, \dots, F_{\text{Line},J}\}$ as the set of fluxes over all

Table 4
Summary of Model Parameters (and Priors) for Different Levels of Our Hierarchical Model

Level	Parameters	Prior	Description
Spectral Region	θ_{Line}		
	v_{Line}	$p(v_{\text{Line}}) \propto \text{const}$	Velocity of region
	C	$C + 1 \sim \text{Gamma}(2, 2)$	Absorption-line strength parameter
	b_l	$p(b_l) \propto \text{const}$	Legendre polynomial coefficients for continuum
Single Observation with J Regions	θ_{spec}		
	v_{raw}	$p(v_{\text{raw}}) \sim \text{Unif}[-600, 600]$	Velocity of stellar absorption regions (e.g., H α)
	v_{Aband}	$p(v_{\text{Aband}}) \sim \text{Unif}[-100, 100]$	Velocity of telluric region(s)
	C_1, \dots, C_J	$C_j + 1 \sim \text{Gamma}(2, 2)$	Absorption-line strength parameters
	$b_{l,1}, \dots, b_{l,J}$	$p(b_{l,j}) \propto \text{const}$	Legendre polynomial coefficients for continuum
Star with K Observations	θ		
	v	$p(v) \sim \text{Unif}[-600, 600]$	LOS velocity of the star
	σ_v	$p(\sigma_v) \sim \text{Inv} - \text{Gamma}(7, 72)$	Dispersion of measurements; systematic error
	$v_{\text{raw},k}, v_{\text{Aband},k}$	$v_{\text{corr}} = v_{\text{raw},k} - v_{\text{Aband},k} - v_{\text{helio},k}$ $v_{\text{corr},k} \sim \mathcal{N}(v, \sigma_v^2)$	Raw and A-band velocities of individual spectra
	$C_{1,1}, \dots, C_{J,K}$	$C_{j,k} + 1 \sim \text{Gamma}(2, 2)$	Absorption-line strength parameters
	$b_{l,1,1}, \dots, b_{l,J,K}$	$p(b_{l,j,k}) \propto \text{const}$	Legendre polynomial coefficients for continuum

Note. In the first level of the model, we model a region of the spectrum containing an absorption-line feature, such as H α , the telluric A-band region, or the CaT region. In the next level, we use multiple spectral regions to estimate the corrected velocity of a star from a single spectroscopic observation. Finally, in our hierarchical model, we incorporate multiple spectroscopic observations into our estimate of the corrected velocity of the star.

spectral regions. When modeling three spectral regions, θ_{spec} contains 11 free parameters.

In order to sample from the posterior, we use `emcee` (Foreman-Mackey et al. 2013), a PYTHON implementation of the Goodman & Weare (2010) affine-invariant Markov chain Monte Carlo (MCMC) ensemble sampler. We first initialize our walkers by estimating the parameters one at a time. Results from extensive fake data testing, including sample trace and corner plots, can be found in the [Appendix](#).

3.3. Hierarchical Modeling

In order to combine spectra from different observations, we employ Bayesian hierarchical modeling. While Bayes's theorem (Equation (6)) gives the probability of a vector of model parameters θ given a vector of data y , it is often desirable for the parameters themselves to be drawn from a distribution, whose values we would like to estimate. These *hyperparameters* (φ) are incorporated into Bayes's theorem as follows:

$$p(\theta, \varphi | y) \propto p(y | \theta, \varphi) p(\theta | \varphi) p(\varphi). \quad (8)$$

Here $p(\theta | \varphi)$ is the probability of the hyperparameters given parameters θ , and $p(\varphi)$ is the hyperprior.

For our model for multiple observations of a star, we have two hyperparameters: v , which is the “true” velocity of the star, and σ_v , the dispersion of velocity measurements (this term serves to model additional uncertainty/noise not captured by the reduction pipeline). For the velocity of a star with K observations with spectra $\mathcal{F}_1, \dots, \mathcal{F}_K$, the full posterior is given by

$$p(v, \sigma_v, \theta_{\text{spec},1}, \dots, \theta_{\text{spec},K} | \mathcal{F}_1, \dots, \mathcal{F}_K) \\ = p(v_1, \dots, v_K | v, \sigma_v) \times \prod_{k=1}^K p(\theta_{\text{spec},k} | \mathcal{F}_k) p(v, \sigma_v), \quad (9)$$

where $v_k = v_{\text{raw},k} - v_{\text{Aband},k} - v_{\text{helio},k}$ is the corrected velocity for observation k and $p(v, \sigma_v)$ is the prior distribution on the hyperparameters.

For these measurements, we can consider the fact that we have substantial prior information about the extent to which these velocities should agree: we know that we are observing the same star with each observation. Therefore, it does not make sense for σ_v to be arbitrarily large, and a standard noninformative prior is not necessarily the best choice. We assign σ_v^2 to be drawn from an inverse-gamma distribution with parameters $a = 7$, $b = 72$. This prior distribution has a mean of $b/(a - 1) = 12$, a mode of $b/(a + 1) = 8$, and variance $b^2/((a - 1)^2(a - 2)) = 28.8$. This prior distribution therefore assigns highest probability to σ_v in the range of 3–4 km s $^{-1}$, but it does allow for σ_v to take on higher values if demanded by the data.

Given the complexity of our model, we use `emcee` to sample from the posterior. All our model parameters and prior distributions, for each level of the model, are listed in Table 4.

A demonstration of this technique is shown in Figure 6, for a HALO7D target with $m_{\text{F606W}} = 22.0$. This target was observed seven times, in varying conditions, over the course of spring 2015. Each of the histograms in the top panel represents the posterior distribution for the *corrected* velocity for each of these seven observations. The varying widths of these pdf's reflect the varying quality in observing conditions across the different nights of observing. These posterior samples were derived using `emcee` to sample the posterior distribution given in Equation (7). However, once we link the observations by combining them with the hierarchical model, the posterior distributions for the individual velocities converge (bottom panel). The posterior distribution for the corrected velocity of the star is shown in black: this posterior incorporates all information, as well as sources of uncertainty, from the seven observations.

The resulting velocity uncertainties for the HALO7D targets as a function of m_{F606W} apparent magnitude are plotted in Figure 7.

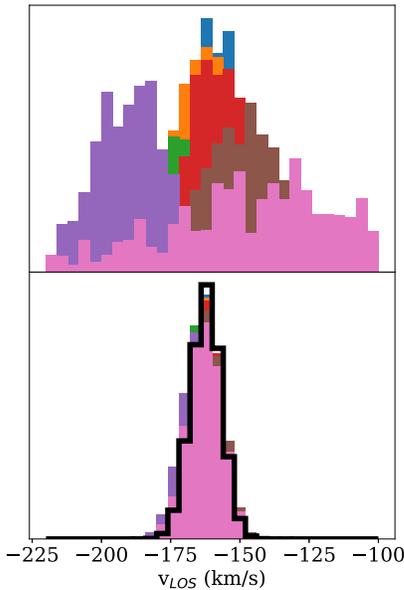


Figure 6. Histograms of posterior samples for the corrected velocity of a HALO7D target ($m_{F606W} = 22.0$) from seven observations. The top panel shows the posterior samples for the velocities when the spectra are modeled independently; note that these are the pdf’s for $v = v_{raw} - v_{Aband} - v_{helio}$. The bottom panel shows the pdf’s for the individual observations once they have been combined into the hierarchical model. The final pdf for the corrected velocity (thick black line), incorporating all observations, thus folds in all information and uncertainty from all observations of a star.

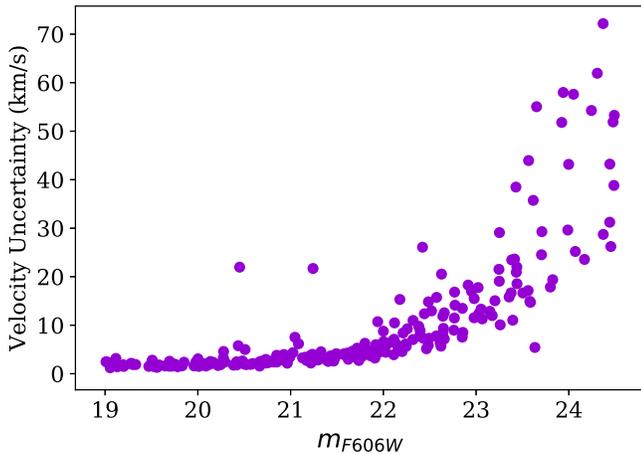


Figure 7. Velocity uncertainties for the HALO7D sample as a function of m_{F606W} apparent magnitude. Velocity uncertainties are the 16th and 84th percentiles of the posterior distributions.

Velocity errors are computed as half the difference between the 84th and 16th posterior percentiles: $v_{err} = (v_{84} - v_{16})/2$. At the bright end of our sample, our velocity uncertainties are as low as $1-2 \text{ km s}^{-1}$; velocity uncertainties remain below 10 km s^{-1} for stars brighter than $m_{F606W} = 22$. Stars at the faint end of our sample reach velocity uncertainties as high as $\sim 50 \text{ km s}^{-1}$.

For more details on testing our method on fake data, including sample trace and corner plots, please see the [Appendix](#).

4. Results

4.1. LOS Velocity Distributions

We use the velocities measured from VELOCIRAPTOR to study the LOS velocity distributions of the stellar halo.

Heliocentric LOSs are converted to the galactocentric standard of rest (GSR) frame by assuming a circular speed of 240 km s^{-1} at the position of the Sun ($R_0 = 8.5 \text{ kpc}$), with solar peculiar motion $(U, V, W) = (11.1, 12.24, 7.25) \text{ km s}^{-1}$ (Schönrich et al. 2010).

Figure 8 shows cumulative histograms for the LOS velocity distributions (in the GSR frame) for the four HALO7D fields. To capture the effects of our velocity uncertainties, we have plotted 100 realizations of the velocity cumulative distribution, each time drawing a new value for every velocity from its posterior. Therefore, the apparent thickness of a given step in the histogram is an indication of the uncertainty of that measurement. We also show traditional histograms of the LOS velocities in Figure 9.

Based on the histograms, we see that our samples across all four fields are dominated by a “hot halo” population; while there could be hints of substructure in these fields, we find that none of our fields are dominated by kinematically cold substructure, which would appear as narrow ($5-15 \text{ km s}^{-1}$) peaks in the velocity distributions. We leave the discussion of the search and characterization of potential substructure in these fields to future work, where we will also utilize PMs and abundances.

To estimate the LOS velocity dispersion of the halo, σ_{LOS} , we model the LOS velocity distributions as a two-component mixture model, with a halo component and a disk component. We model the halo velocity distribution as a normal distribution with unknown mean and variance: $v \sim N(\langle v_{LOS} \rangle, \sigma_{LOS}^2)$.

We model the disk velocity distribution along each line of sight as a skew-normal distribution, with skew parameter α , location parameter ζ , and scale parameter ω . The likelihood of an observed velocity v_i given disk parameters is given by

$$p(v_i | \alpha, \zeta, \omega) = \frac{2}{\omega} \phi\left(\frac{v_i - \zeta}{\omega}\right) \Phi\left(\alpha \left(\frac{v_i - \zeta}{\omega}\right)\right), \quad (10)$$

where $\phi(x)$ is the standard normal pdf and $\Phi(x)$ is the standard normal cumulative distribution function (CDF). We fix the parameters of the disk velocity distribution but leave the fraction of disk contamination as a free parameter. We denote our disk pdf as $g_{\text{disk}}(v_i) = p(v_i | \alpha, \zeta, \omega)$.

To determine the parameters of our disk model, we use the Besançon galaxy model (Robin et al. 2003). We use synthetic catalogs of 1 deg^2 areas centered on the coordinates of our survey fields (the larger area gives us better statistics for our simulated disk and halo populations). We then determine the velocity distribution of the (non-WD) disk contaminants within the HALO7D selection boxes and fit a skew normal to this distribution. The resulting parameters for the disk distributions can be found in Table 5; they are also plotted in Figures 8 and 9.

For this mixture model of disk and halo, the likelihood of a given LOS velocity observation v_i , with error $\sigma_{v,i}$, is given by

$$p(v_i | f_{\text{disk}}, \langle v_{LOS} \rangle, \sigma_{LOS}, g_{\text{disk}}) = f_{\text{disk}} \times g_{\text{disk}}(v_i) + (1 - f_{\text{disk}}) \times N(v_i | \langle v_{LOS} \rangle, \sigma_{LOS}^2 + \sigma_{v,i}^2). \quad (11)$$

In order to sample from the posterior distribution for our model parameters, we first must assign prior distributions. We assign a standard uniform $[0,1]$ prior on the fraction of disk contamination, and we assign the Jeffreys prior to the mean and dispersion for the halo LOS velocity distribution ($p(\langle v_{LOS} \rangle, \sigma_{LOS}) \propto 1/\sigma_{LOS}$).

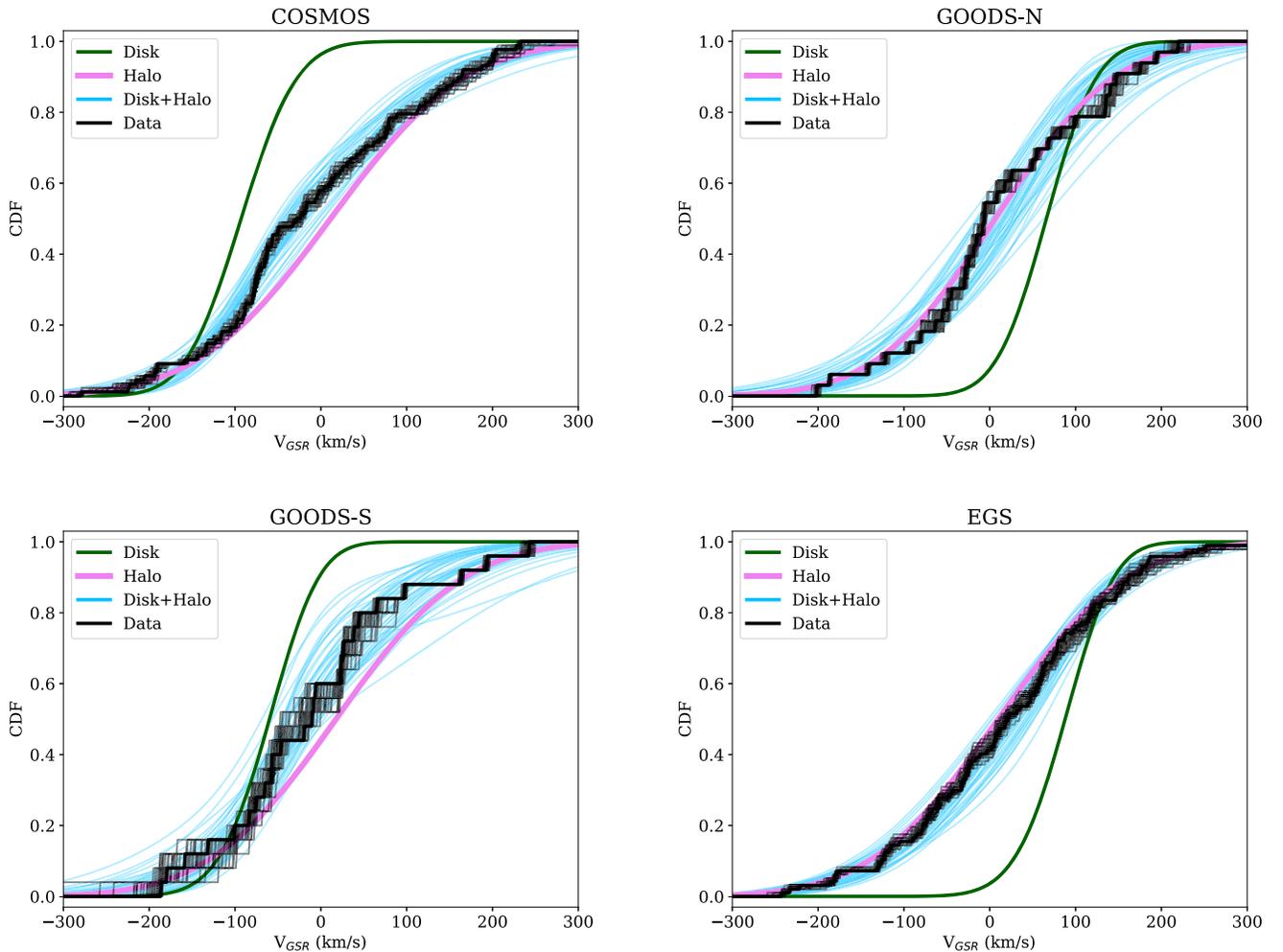


Figure 8. Cumulative LOS velocity histograms in the four HALO7D fields. Velocities are given with respect to the GSR. Black lines indicate the CDFs for the data; for each of the 100 black lines, velocity values were drawn from the posterior distributions for the measurements. The width of each step thus demonstrates the velocity uncertainty for that data point. The CDF for the disk model is shown in green, and the CDF for the halo model, using the posterior median values for $\langle v_{\text{LOS}} \rangle$, σ_{LOS} , is shown in pink. Blue lines show the CDF for the mixture of the disk and halo populations, for 100 draws from the joint posterior distribution for f_{Disk} , $\langle v_{\text{LOS}} \rangle$, σ_{LOS} .

Our posterior is thus

$$\begin{aligned}
 p(\langle v_{\text{LOS}} \rangle, \sigma_{\text{LOS}}, f_{\text{disk}} | \nu) &\propto p(\langle v_{\text{LOS}} \rangle, \sigma_{\text{LOS}}) p(f_{\text{disk}}) \\
 &\times \prod_{i=1}^{N_{\text{stars}}} p(v_i | \langle v_{\text{LOS}} \rangle, \sigma_{\text{LOS}}, f_{\text{disk}}). \quad (12)
 \end{aligned}$$

We use `emcee` to sample from this posterior. We used 500 walkers, ran the sampler for 1000 steps, and discarded the first 800 steps as burn-in. Median posterior values, along with error bars from the 16/84 percentiles, are quoted for the three model parameters in Table 5.

Posterior draws are overplotted on the histograms in Figures 8 and 9. Each pink line in Figure 8 is the CDF corresponding to a draw from the posterior for $\langle v_{\text{LOS}} \rangle$ and σ_{LOS} . In Figure 9, the amplitude of the disk pdf's reflects the uncertainty in the disk contribution. Thicker pink and green lines indicate the distributions corresponding to the median posterior values.

Histograms of posterior samples for our three free parameters are shown in Figure 10. The left panel shows the posterior distributions for $\langle v_{\text{LOS}} \rangle$; all fields have mean LOS velocity

consistent with 0 km s^{-1} . The middle panel of Figure 10 shows the posterior samples for σ_{LOS} ; posterior pdf's for σ_{LOS} are consistent across the four fields. The widths of the individual pdf's vary according to the sample size in a given field, but the pdf's substantially overlap. In the COSMOS field, we estimate $\sigma_{\text{LOS}} = 123_{-11}^{+12} \text{ km s}^{-1}$; in GOODS-N, $\sigma_{\text{LOS}} = 110_{-13}^{+16} \text{ km s}^{-1}$; for GOODS-S, $\sigma_{\text{LOS}} = 122_{-21}^{+30} \text{ km s}^{-1}$; and in EGS, we find $\sigma_{\text{LOS}} = 118_{-9}^{+11} \text{ km s}^{-1}$.

The right panel of Figure 10 shows posterior samples for f_{disk} in the four fields. We note that no color or distance information is incorporated into our estimates of the disk contamination and that this estimate is based on LOS velocities alone. Our estimates of our disk contamination will be more accurate once PM and photometric information are incorporated. EGS and GOODS-N show 0%–10% disk contamination, consistent with the predictions from Besançon (see Table 5). COSMOS, our lowest-latitude field, shows a slightly higher level of disk contamination ($\sim 25\%$, while predicted to be $\sim 11\%$). The posterior distribution for disk contamination in GOODS-S, our field with the smallest sample size, is very broad, with a

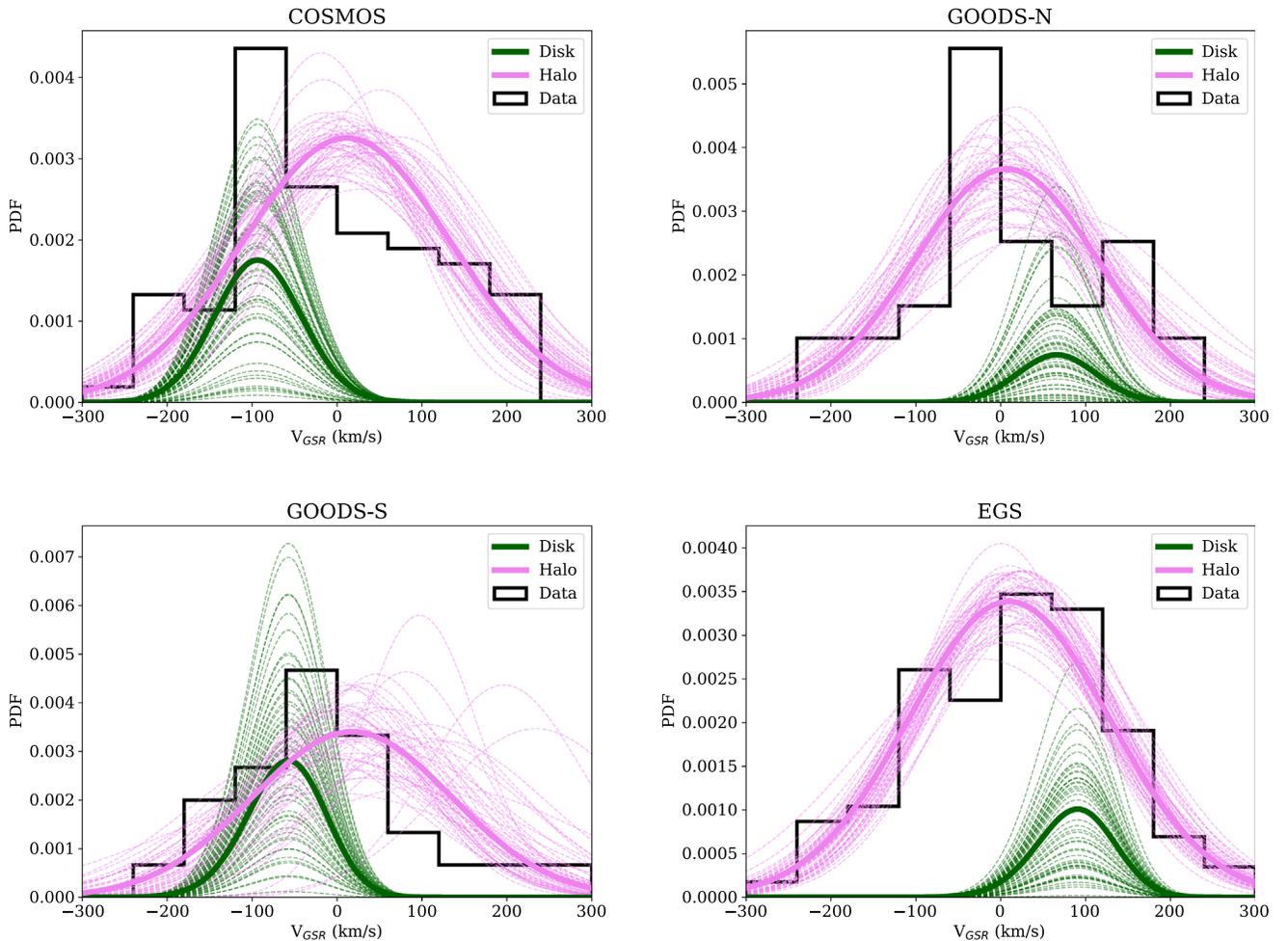


Figure 9. LOS velocity histograms in the four HALO7D fields. Shown in pink are the resulting velocity distributions from 50 draws from the posteriors for $\langle v_{\text{LOS}} \rangle$ and σ_{LOS} . The green line indicates the disk distribution. The parameters of the disk velocity distribution are fixed; only the fraction to the total contribution is allowed to vary. Bold lines show the corresponding distributions for the median posterior values of f_{disk} , $\langle v_{\text{LOS}} \rangle$, and σ_{LOS} .

Table 5
Summary of Results from the Modeling of the LOS Velocity Distributions

Field	$\langle v_{\text{LOS}} \rangle$ (km s $^{-1}$)	σ_{LOS} (km s $^{-1}$)	f_{disk}	Predicted f_{disk}	α	ζ (km s $^{-1}$)	ω (km s $^{-1}$)	$\langle D \rangle$ (kpc)
COSMOS	13_{-19}^{+23}	123_{-11}^{+12}	$0.23_{-0.11}^{+0.12}$	0.11	0.6	-115	56	21
GOODS-N	6_{-21}^{+20}	110_{-13}^{+16}	$0.11_{-0.06}^{+0.12}$	0.07	0.0	66	46	20
GOODS-S	24_{-35}^{+48}	122_{-21}^{+30}	$0.34_{-0.20}^{+0.22}$	0.07	-1.1	-27	58	23
EGS	9_{-14}^{+15}	118_{-9}^{+11}	$0.13_{-0.08}^{+0.10}$	0.07	-1.1	121	58	21

Note. Posterior medians are quoted, with error bars giving the 16/84 percentile credible regions. We list the predicted fraction of disk contamination from Besançon; because we have included no color or apparent magnitude information in our estimate of the disk contamination, we expect that our f_{Disk} estimates are higher than the actual disk contamination in our sample. We list the parameters for the disk model for each field: the skew α , the location ζ , and the scale ω . These values were derived by fitting a skew-normal distribution to the velocities of disk stars that fall in the HALO7D selection box in the Besançon galaxy model. Finally, we list the average distance to each field, as computed in Equation (13).

posterior median of 30%, much higher than the 7% predicted by Besançon. We note that this high posterior median is largely due to the small sample size and that the disk contamination in GOODS-S is poorly constrained based on LOS velocities alone.

4.2. Comparison with Other Tracers

In this section, we compare our results for the HALO7D LOS velocity distributions with other studies conducted over a

similar distance range (though using different tracer populations). In order to compare our measurements of σ_{LOS} with other studies, we first need to estimate the distance range probed by our sample. We estimate these distance distributions in a similar method to that of D13 and C16, using weighted isochrones to derive the pdf for the absolute magnitude M_{F814W} of a star given its $m_{\text{F606W}} - m_{\text{F814W}}$ color.

We weight the Vandenberg et al. (2006) isochrones according to a Salpeter initial mass function (IMF) and an age and a metallicity distribution typical of halo stars. We

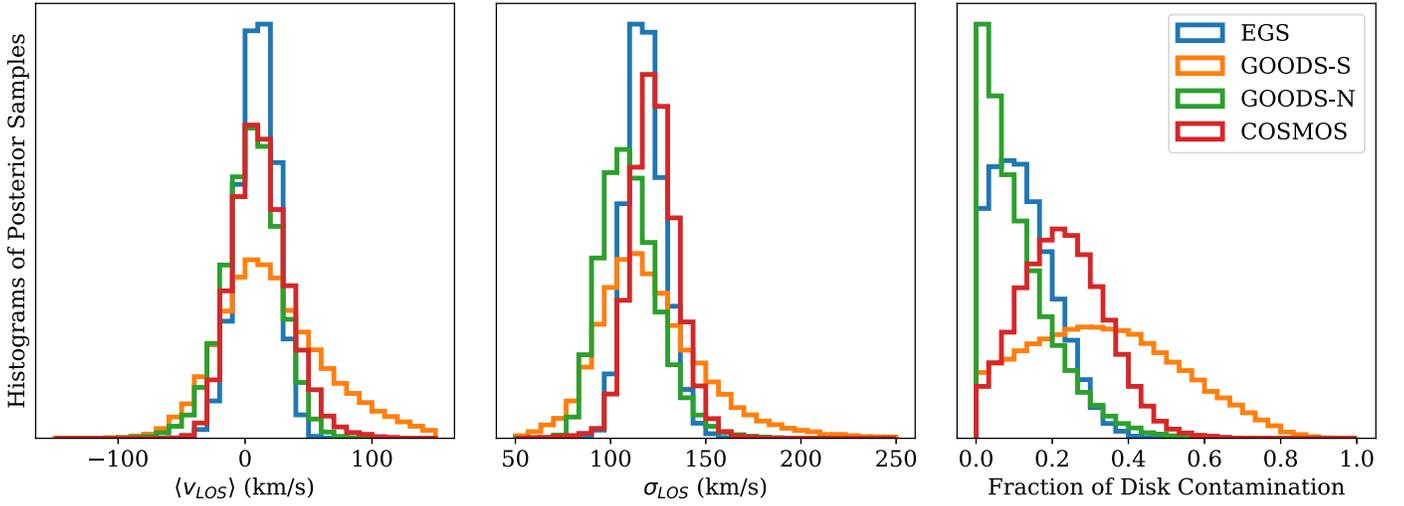


Figure 10. Posterior samples for the mean LOS velocity (left), LOS velocity dispersion (middle), and the fraction of disk contamination (right) for all four fields. All fields have mean LOS velocity consistent with 0 km s^{-1} and dispersions consistent with one another.

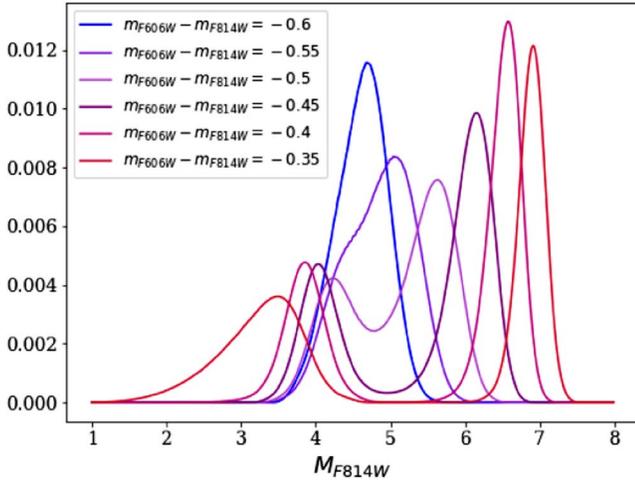


Figure 11. Normalized pdf's for absolute magnitude for six different choices of $m_{F606W} - m_{F814W}$ color. These pdf's are derived from the KDE constructed from the Vandenberg et al. (2006) isochrones, weighted by a Salpeter IMF and the approximate age and metallicity distributions of the halo.

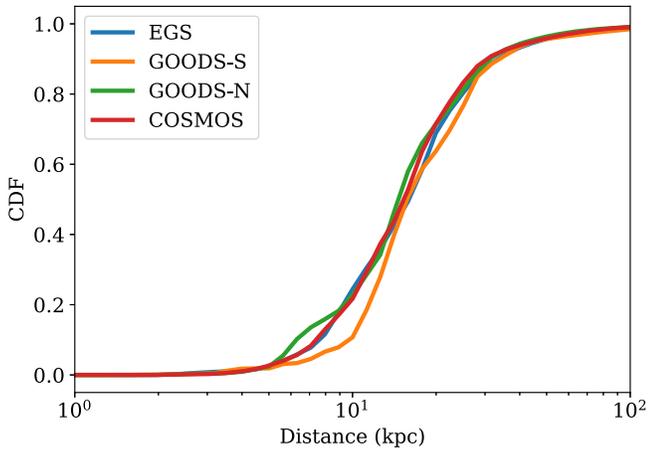


Figure 12. Cumulative distance distributions for the four HALO7D fields. Distance distributions are computed as given by Equation (13), using colors and assuming an MW stellar density profile (Deason et al. 2011). All fields have $\langle D \rangle \sim 20 \text{ kpc}$.

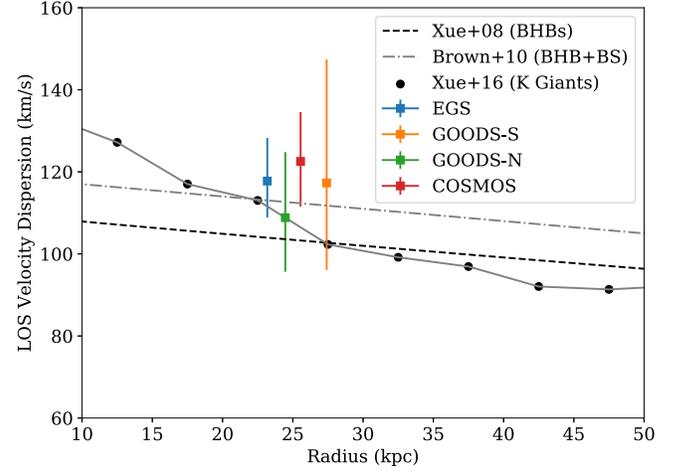


Figure 13. LOS velocity dispersions of the four HALO7D fields, plotted as a function of mean galactocentric radius. Vertical error bars show the 16%–84% quantiles of the marginalized posterior. We compare our LOS dispersions with results from other studies: the black dashed line indicates best-fit LOS dispersion profile from Xue et al. (2008), measured from BHBs in DRSS. The gray dot-dashed line indicates the best-fit dispersion profile from Brown et al. (2010), using BHB and BS stars as tracers. The connected black circles show the resulting dispersion profile from the SEGUE *K*-giant survey from Xue et al. (2016). The HALO7D dispersions are consistent with predictions from other tracers.

assume that the age and metallicity distributions are Gaussian, with $\langle T \rangle = 12 \text{ Gyr}$, $\sigma_T = 2 \text{ Gyr}$ (e.g., Kalirai 2012), $\langle [\text{Fe}/\text{H}] \rangle = -1.9$, and $\sigma_{\text{Fe}/\text{H}} = 0.5$ (e.g., Xue et al. 2008). We model the resulting weighted CMD with a kernel density estimate (KDE), using a kernel bandwidth of 0.025. The resulting pdf's for M_{F814W} , for six different colors, are shown in Figure 11.

Using the pdf for absolute magnitude as a function of color in conjunction with the halo density profile (Deason et al. 2011), we derive the pdf for the log distance distribution to our sample:

$$p(\log D | m_{F814W}, m_{F606W}, \rho) \propto p(\log D | \rho) \times \sum_{n=1}^{N_{\text{obj}}} p(M_{F814W}(\log D) | m_{F606W,n} - m_{F814W,n}), \quad (13)$$

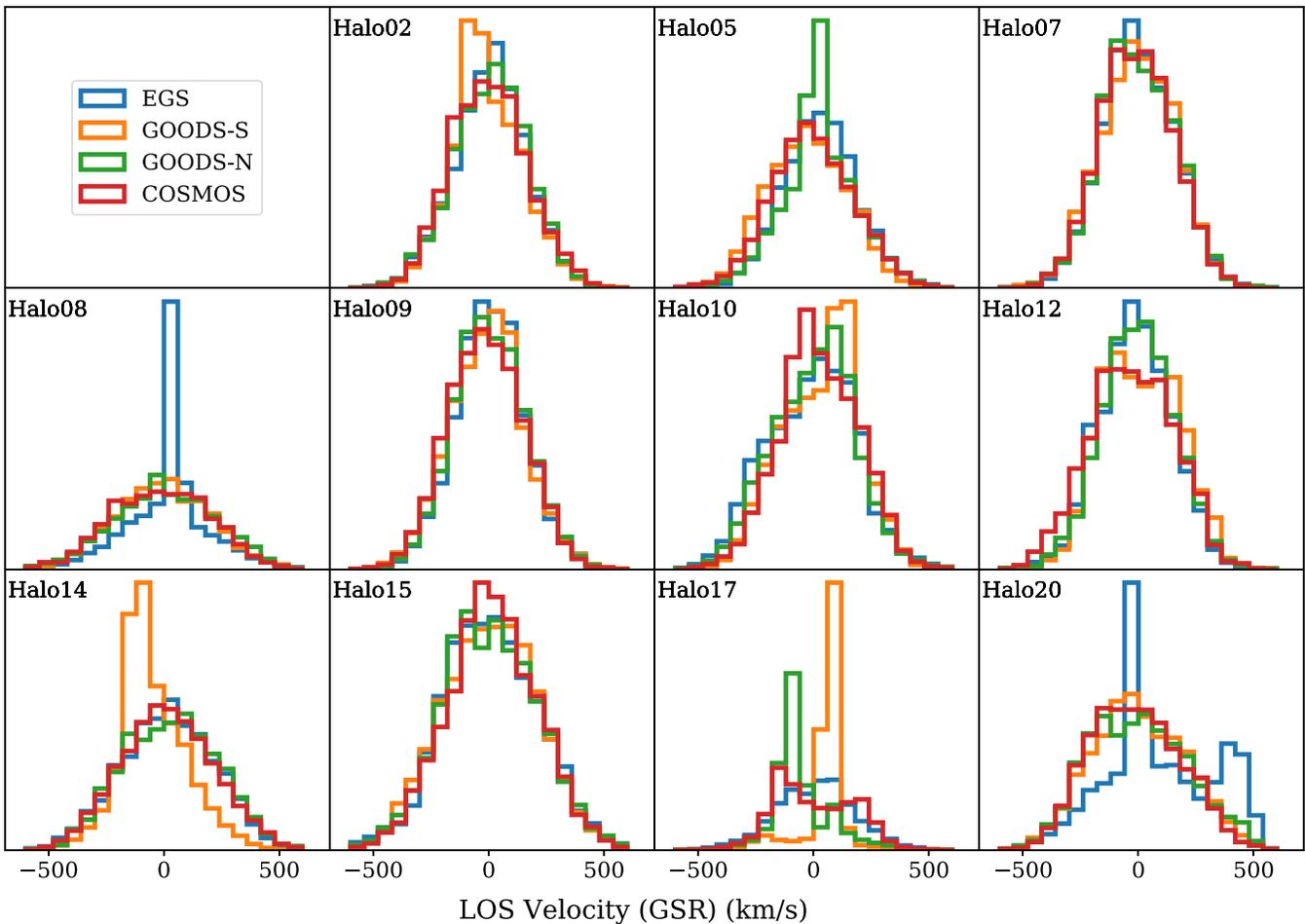


Figure 14. LOS velocity distributions from 1 deg² “observations” of the BJ05 accreted stellar halos generated with Galaxia. Different colored histograms denote the observations in the different HALO7D fields. Seven out of the 11 BJ05 halos show consistent velocity distributions across the four fields. Three halos show consistency across three fields, with one field dominated by substructure. Halo17 shows four distinct LOS distributions across the four fields.

where $p(\log D|\rho)$ is the probability of $\log D$ given the Deason et al. (2011) density profile, and $p(M_{F814W}(\log D)|m_{F606W,n} - m_{F814W,n})$ is the probability of object n having absolute magnitude $M_{F814W}(\log D)$ given its color $m_{F606W,n} - m_{F814W,n}$.⁶ We then estimate the mean distance to each field $\langle D \rangle = \int D \times p(\log D) d \log D$. Each of the four fields has an average distance $\langle D \rangle \sim 20$ kpc. Figure 12 shows the cumulative logarithmic distance pdf’s of our samples across the four fields, and average distances to each field are listed in Table 5. We note that no kinematic information is incorporated into our distance estimate and that information from 3D kinematics will improve our distance estimates in subsequent work.

Figure 13 shows the LOS dispersions of the four HALO7D fields plotted as a function of mean galactocentric radius (where we have converted heliocentric distance $\langle D \rangle$ to galactocentric radius $\langle r \rangle$). Points indicate the median of the σ_{LOS} posterior samples, and error bars indicate the 16th and 84th posterior percentiles. We compare our results to the velocity dispersion profiles measured in other studies. The

figure also shows measured velocity dispersion profiles from SDSS BHB stars (Xue et al. 2008; black dashed line), BHB and blue straggler (BS) stars from the Hypervelocity Star Survey (Brown et al. 2010; gray dot-dashed line), and the SEGUE K -giant survey (Xue et al. 2016; connected black circles). The measured LOS velocity dispersions in the HALO7D fields using MSTO stars are consistent with other studies that have measured the LOS velocity dispersion profile over our distance range.

5. Comparison with Simulations

In all four lines of sight, we see the “hot halo” population; none of our fields appear to be obviously dominated by substructure that is cold in LOS velocity. In addition, the measured LOS velocity dispersions across the four fields are all consistent with one another and are consistent with measurements made using other tracers. In this section, we interpret our LOS velocity distributions and the results from our velocity dispersion analysis in the context of simulations.

We perform mock HALO7D surveys on the 11 Bullock & Johnston (2005, hereafter BJ05) halos, using the synthetic survey software Galaxia (Sharma et al. 2011). The publicly available Bullock & Johnston simulations are high-resolution N -body simulations of accreted dwarf galaxies onto an MW-like parent galaxy. The parent galaxy has a time-dependent

⁶ In the COSMOS field, in the area where we used additional catalogs for selection, we used the Subaru B and V photometric measurements from Capak et al. (2007) and converted these to STMAG F606W as directed by Sirianni et al. (2005). We used the F814W magnitudes as published in the Leauthaud et al. (2007) catalog.

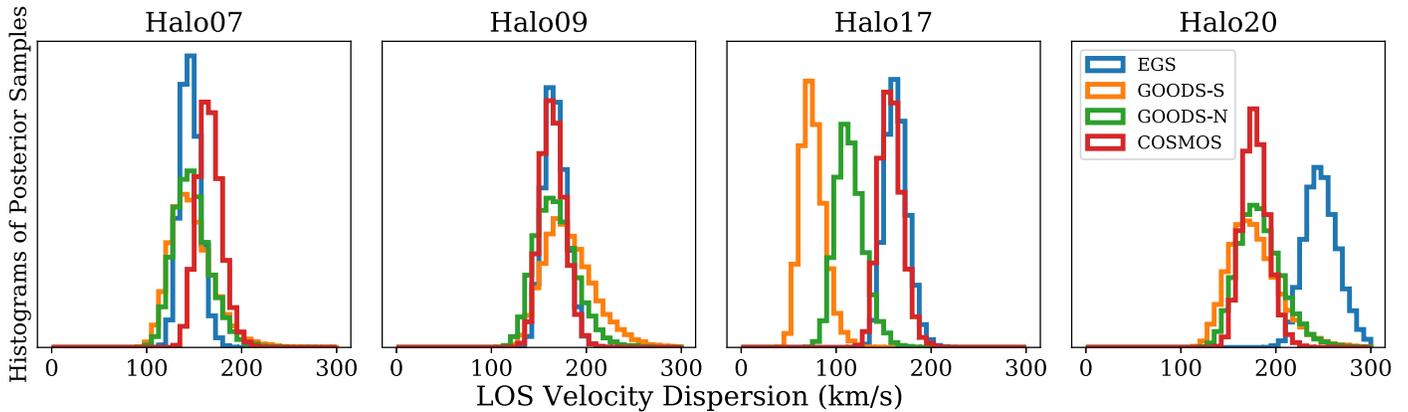


Figure 15. Examples of posterior samples from our dispersion analysis on our mock data sets. Each panel shows histograms of posterior samples for the 1D marginalized posterior distributions for the velocity dispersion (same as second panel of Figure 10) for a single mock HALO7D data set. Each colored histogram represents a different HALO7D field. The first two panels show examples of consistent estimates of the LOS velocity dispersion for Halo07 and Halo09. The third panel shows an example of posterior estimates for Halo17: here, the cold peaks in GOODS-N and GOODS-S have caused the estimates of σ_{LOS} to be lower in these fields than in the other two fields. The fourth panel shows results from a mock data set analysis for Halo20, in which the kinematically cold substructure at 400 km s^{-1} and at 0 km s^{-1} results in a higher dispersion estimate in the EGS field.

analytical potential with halo, disk, and bulge components. Because there is no stellar disk in these simulations, there is no “in situ” stellar halo component in these galaxies. Galaxia can accept N -body simulations as input, and it generates synthetic catalogs with smooth, continuous distributions of stars over any given volume.

We first use Galaxia to observe 1 deg^2 areas centered on the coordinates of the four HALO7D fields in all 11 BJ05 halos. We create catalogs of synthetic fields that are larger than our survey fields because we are interested in exploring the underlying LOS velocity distributions along these lines of sight; the larger area provides us with more samples from these distributions. We then selected stars that fell within the HALO7D CMD selection boxes, in order to study the velocity distributions over the same distance range as our survey.

Figure 14 shows the resulting LOS velocity distributions for the four HALO7D fields in the 11 BJ05 halos. Each panel represents a different BJ05 halo, while each colored histogram shows the LOS velocity distribution for a given HALO7D field. In 7 out of the 11 halos, the four LOS velocity distributions are all “hot” and appear to be consistent with one another. In three of the remaining halos, three of the fields have consistent LOS velocity distributions, with one field having a strong cold peak (Halo08, Halo14, and Halo20). The only halo with four different velocity distributions across the four LOSs is Halo17. Therefore, even if our survey had many more stars and no measurement uncertainties, based on the BJ05 velocity distributions, it would not be surprising to find that none of our four fields were dominated by kinematically cold substructure (especially given the fact that we expect more coherent substructure in the BJ05 halos than in the MW because of their assumption of a smoothly growing potential; see Section 2.5 of BJ05).

However, our survey is limited by relatively small numbers (especially in GOODS-N and GOODS-S) and large measurement uncertainties at faint apparent magnitudes. Therefore, to make a valid comparison to our observations, we need to incorporate these effects. In order to assess how these different LOS velocity distributions respond to our observations and dispersion analysis procedure, we generate mock HALO7D data sets. We draw numbers of samples from the true underlying LOS velocity distributions corresponding to the

sample sizes of our observations (e.g., we draw 88 stars from the EGS distribution, 25 stars from the GOODS-S distribution, etc.). We then assign velocity uncertainties to each star, based on a fit of the relation between apparent magnitude and LOS velocity error shown in Figure 7. We generate 30 sets of mock HALO7D observations for each BJ05 halo.

5.1. Dispersion Estimates

As can be seen in the middle panel of Figure 10, the posterior distributions for the LOS velocity dispersions in the four HALO7D fields substantially overlap, and the posterior median for each distribution lies within the 95% credible region of the other three distributions. For the purposes of this analysis, we define this result as having “consistent” estimates for the dispersion across all four fields. A specified posterior distribution is “consistent” with another posterior distribution if its posterior median lies within the 95% credible region of the other posterior distribution.

For each mock catalog, we evaluate the posterior distributions for the mean and dispersion along each sight line just as in Section 4.1 (though here we omit the disk model). Four examples of the resulting 1D marginalized posterior distributions for the velocity dispersions are shown in Figure 15. We find that the result of four consistent (as defined by our metric) posterior distributions for the velocity dispersions is relatively rare: in only 38 out of the 330 mock data sets did we find consistency across the four distributions. Out of 30 mock data sets per halo, the four distributions were consistent five times for Halo02, seven times for Halo07, three times for Halo08, six times for Halo09, two times for Halo10, eight times for Halo12, and seven times for Halo15. The first two panels of Figure 15 show examples of consistent velocity dispersion estimates (in Halo07 and Halo09). For Halo05, Halo14, Halo17, and Halo20, none of the mock HALO7D data sets yielded consistent dispersion estimates across all four fields. Example velocity dispersion posterior distributions for Halo17 and Halo20 are shown in the right panel of Figure 15.

Unsurprisingly, the halos that never achieved consistency across the four fields all contain cold peaks in their velocity distributions. It is important to note that these peaks do not necessarily have the same effects on the dispersion estimates.

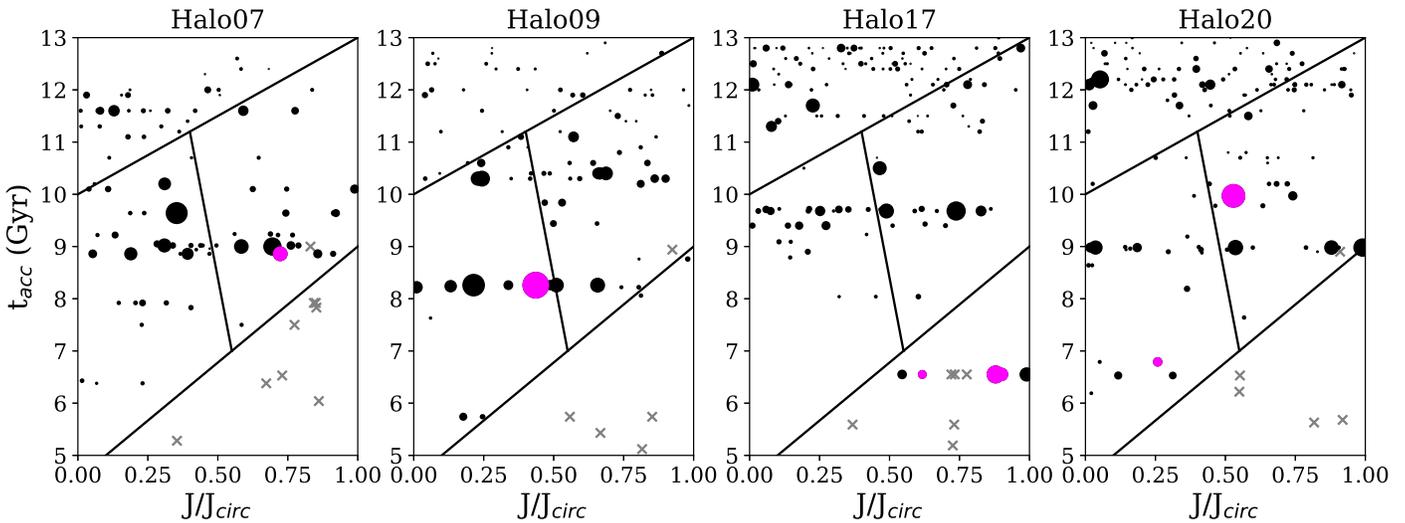


Figure 16. Age vs. orbit circularity for the accretion events making up Halo07, Halo09, Halo17, and Halo20. Points are scaled by accretion event mass. Gray crosses indicate still-bound satellites. Colored points indicate the “dominant” satellites in the mock HALO7D samples. For Halo07 and Halo09, the same satellite is dominant across all four fields. In Halo17, one satellite dominates two fields, while the other two fields are dominated by distinct satellites. For Halo20, three out of the four HALO7D fields are dominated by the most massive satellite, while the EGS field is dominated by a low-mass, recent accretion event. The recent accretion events in Halo17 and Halo20 are responsible for the variable velocity distributions, while early, massive accretion results in consistent velocity distributions for Halo07 and Halo09.

In the cases of Halo14 and Halo17, the cold peaks near 0 km s^{-1} decreased the dispersion estimates in these fields relative to those without strong peaks (as in the third panel of Figure 15). However, in the case of Halo20, which has a peak at 0 km s^{-1} as well as a peak at 400 km s^{-1} in its EGS field, the dispersion estimate in this field was consistently higher than the estimates in the other fields (as in the right panel of Figure 15).

In spite of having a strong cold peak at 0 km s^{-1} in the EGS field, Halo08’s posterior distributions were consistent for 3 out of the 30 mock data sets. This is because the stars responsible for the cold peak are all located at a far and common distance of $\sim 90 \text{ kpc}$; as a result, the velocity errors are sufficiently large for stars at these distances that the dispersion estimate is not strongly affected by the presence of the cold peak. In contrast, the two most dominant accretion events affecting Halo20’s velocity distribution in EGS have mean distances of 16 kpc (responsible for the peak at 400 km s^{-1}) and 10 kpc (responsible for the peak at 0 km s^{-1}). With our large LOS velocity errors for stars at this distance, these stars do not strongly affect the resulting dispersion estimates. Based on our dispersion analysis alone, we cannot make claims about the presence or lack thereof of substructure in our sample, and we leave the detection and characterization of potential substructure in HALO7D to future work.

5.2. Accretion Histories

We now look at four halos more closely: Halo07 and Halo09, which both achieved our metric for consistency numerous times, and Halo17 and Halo20, whose mock data sets never met our metric for consistency across the four fields. The accretion histories of these halos are shown in Figure 16, in the space of accretion time versus J/J_{circ} , with point sizes scaled by the mass of the accretion event. Lines indicate the regions of this plane dominated by the different morphological types discussed in Johnston et al. (2008); early, more radial accretion events tend to have mixed morphologies, whereas recent, circular accretion events tend to have great circle morphologies (see their Figure 3).

Accretion events that contribute the highest numbers of stars to one of the mock HALO7D fields are shown as pink points. In the case of Halo07 and Halo09, debris from the same accretion event contributes the most stars in all four of the HALO7D fields (hence only one accretion event in each halo is highlighted in pink). In contrast, in Halo17, debris from the same accretion event dominates the mock COSMOS and EGS distributions, but two different (and less massive) events contribute the most stars to the GOODS-N and GOODS-S fields. For Halo20, three out of the four HALO7D fields are all dominated by debris from the same, massive early accretion event; however, in the COSMOS field, the largest fraction of stars comes from a recent, low-mass accretion event. The total fractions of stars contributed by accretion events that contribute $>10\%$ of the stars along a given sight line are shown in Figure 17; point sizes are again scaled by mass, with the different triangle orientations indicating the different sight lines.

Neither Halo07 nor Halo09 has velocity distributions visibly dominated by kinematically cold substructure. In the case of Halo09, the most massive accretion event experienced by the halo ($m = 2.14 \times 10^{11} M_{\odot}$) contributes the most to the mock HALO7D samples along all four lines of sight (contributing 20%–40%). The second most massive accretion event, accreted at the same time, also contributes 10%–20% of the stars in each line of sight (pink points in Figure 17). These massive, coincident accretion events are responsible for the hot velocity distributions that had consistent estimates of their dispersions. Debris from the most massive satellite accreted by the halo also dominated all four lines of sight in Halo02, Halo05, and Halo12.

Halo07 also experienced several coincident accretion events, though these events occurred earlier in its accretion history ($t_{\text{acc}} \sim 9 \text{ Gyr}$) and were on average lower mass than the events experienced by Halo09. The three most massive of these events ($m = (0.5\text{--}1.5) \times 10^{11} M_{\odot}$) each contribute 10%–20% of the stars in each of the four fields (gold points in Figure 17). We therefore see that accretion from several relatively early, coincident accretion events, even of lower mass, can also result

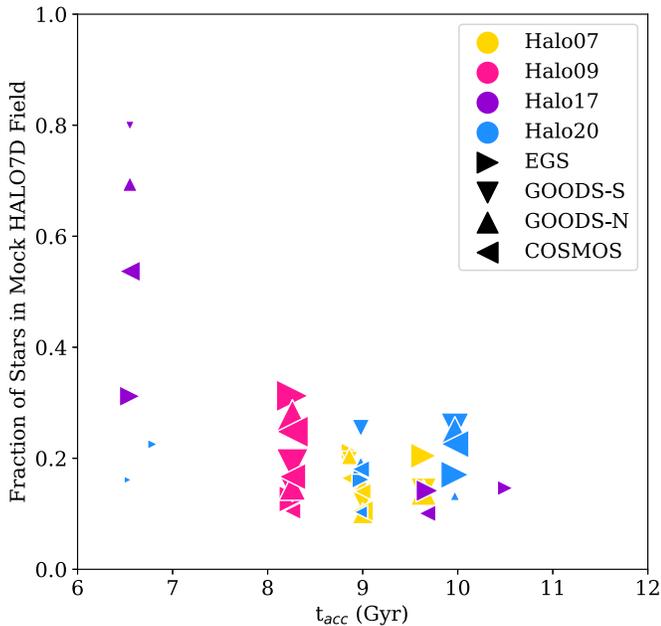


Figure 17. Fraction of stars contributed to a given line of sight as a function of accretion time, for Halo07 (gold), Halo09 (magenta), Halo17 (purple), and Halo20 (light blue). Point sizes are scaled by the mass of the accretion event. Only events that contribute $>10\%$ of the stars are shown. Different shape orientations denote different sight lines. Halo09’s most massive satellite dominates the mock HALO7D sample along all sight lines, contributing 20%–40% of the stars. The second most massive accreted satellite, which was accreted at the same time, also contributes 10%–20% along all four sight lines. In Halo07, several relatively massive accretion events that were accreted around the same time all contribute between 10% and 20% of the stars in the four sight lines. These two accretion histories give rise to consistent velocity distributions along the different sight lines. In contrast, Halo17 has experienced recent accretion of low-mass satellites. These recent accretion events dominate the four HALO7D sight lines and create cold peaks in the LOS velocity distributions. In the case of Halo20, the most massive satellite contributes strongly to the samples across all four sight lines; however, as with Halo17, two lower-mass, recent accretion events contribute $>15\%$ of the stars along the EGS sight line, creating two cold peaks in the velocity distribution.

in consistent estimates of the velocity dispersion across multiple lines of sight.

In contrast, Halo17’s velocity distributions are not consistent across the four mock HALO7D fields. Figure 16 shows that Halo17’s accretion history is characterized by several recent accretion events ($t_{\text{acc}} \sim 6.5$ Gyr) on fairly circular orbits; Figure 17 shows that these recent accretion events strongly dominate the HALO7D sight lines, with the most massive recent event dominating two sight lines ($f_{\text{sample}} \sim 30\%–50\%$) and the other two sight lines being strongly dominated by two different less massive events ($f_{\text{sample}} = 70\%–80\%$). The recent accretion experienced by this halo results in cold LOS velocity distributions along two of the four HALO7D sight lines. As a result, when we estimate the dispersion of mock observations in these fields, the cold peaks result in lower estimates of the dispersion in GOODS-N and GOODS-S compared to COSMOS and EGS.

In the case of Halo20, the most massive accretion event in its accretion history ($m = 1.65 \times 10^{11} M_{\odot}$, $t_{\text{acc}} \sim 10$ Gyr) contributes the most stars to the GOODS-N, GOODS-S, and COSMOS fields. However, in Halo20’s EGS field, two lower-mass ($m \sim (1–2) \times 10^{10} M_{\odot}$), recent ($t_{\text{acc}} \sim 6.5$ Gyr) accretion events create distinct cold peaks in the LOS velocity

distributions (represented by the small blue triangles in the lower left corner of Figure 17).

To summarize, early, massive accretion events, as well as several, early, synchronous accretion events, give rise to consistent, hot velocity distributions along different halo sight lines, which lead to consistent estimates of the velocity dispersion across multiple lines of sight. Recent accretion can lead to sight lines dominated by kinematically cold substructure, even when these accretion events are low mass. Depending on the velocity of the substructure, as well as the number of accretion events causing peaks in the velocity distribution, this can lead to lower estimates of σ_{LOS} (as is in the case for Halo17) or higher estimates of σ_{LOS} (as for Halo20) along these sight lines.

While it is challenging to distinguish between the accretion histories of Halo09 and Halo07 with kinematics alone, we note that the chemical abundances will be different for these two scenarios. An accretion history like Halo09’s should give rise to a higher average $[\text{Fe}/\text{H}]$ than Halo07’s accretion history because of the mass–metallicity relation (e.g., Kirby et al. 2013).

In their analysis of the BJ05 halo density profiles, Deason et al. (2013a) found that halos with early, massive accretion events had breaks in their density profiles (like the density profile of the MW; e.g., Watkins et al. 2009; Deason et al. 2011; Sesar et al. 2011), whereas galaxies with prolonged accretion epochs had single power-law density profiles (like M31; e.g., Gilbert et al. 2012). Recent results from *Gaia* have discovered the remnant of an early, massive accretion event, known as the “*Gaia*-Sausage” (Belokurov et al. 2018) or “*Gaia*-Enceladus” (Helmi et al. 2018), which is both relatively metal-rich and strongly radially biased in its orbital distribution. Deason et al. (2018b) find that the apocenters of these “Sausage” stars are at $r \sim 20$ kpc, coincident with the MW’s break radius; this is also the approximate mean distance to our sample. Studying the LOS velocity distributions of the simulated BJ05 halos, we find that the consistent LOS velocity distributions of the HALO7D fields provide yet another piece of evidence that the MW likely experienced a massive, early accretion event. PM information and abundances will help us determine whether our sample is dominated by *Gaia*-Sausage stars.

6. Conclusions

In this paper, we presented the spectroscopic component of the HALO7D survey, a Keck II/DEIMOS spectroscopic survey of distant, MSTO MW halo stars in the CANDELS fields. We described the survey observing strategy, mask layouts, and target selection. We also presented a new method of measuring velocities from stellar spectra from multiple observations, utilizing Bayesian hierarchical modeling. We used the measured LOS velocities to estimate the parameters of the LOS velocity distributions in the four HALO7D fields.

We summarize our conclusions as follows:

1. When performing slit spectroscopy of point sources, it is essential to consider the apparent velocity shift due to slit miscentering when measuring velocities from individual spectra or when combining multiple spectroscopic observations. The hierarchical Bayesian approach presented in this work (implemented in VELOCIRAPTOR) allows for the parameters of individual observations to be

modeled simultaneously, leveraging the available signal while properly propagating uncertainties.

2. All four HALO7D fields are dominated by the “hot” halo population and have consistent LOS velocity distributions. The estimates of the velocity dispersions are consistent with estimates derived using other tracer populations.
3. We performed mock HALO7D observations using the synthetic survey software Galaxia to observe the Bullock & Johnston (2005) halos. We found that an early, massive accretion event, or several early events, can result in consistent estimates of the velocity dispersion along the different sight lines. This consistency in the velocity dispersion estimates arises because the same satellite (or the same few satellites) dominates the halo population along all sight lines. The consistent estimates of σ_{LOS} in HALO7D therefore could indicate that the MW experienced an early, massive accretion event (or perhaps several events), whose stars are dominating the samples of all four fields. However, we emphasize that our dispersion analysis alone does not confirm or deny the presence of substructure: we intend to study substructure in HALO7D in future work when we can make use of PM and abundance information.

This paper is the first in the HALO7D series; our spectroscopy and the multiepoch *HST* imaging will enable us to measure PMs and abundances for these same stars. HALO7D is a deep complement to the *Gaia* mission: the stars in this data set will be the faintest stars with measured 3D kinematics until LSST. With upcoming PMs and abundances, we can continue to use the HALO7D data set to improve our understanding of the Galaxy’s formation.

Over the course of this work, E.C. was supported by a National Science Foundation Graduate Research Fellowship and an ARCs Foundation Fellowship, as well as NSF grant AST-1616540. Partial support for this work was provided by NASA through grants for program AR-13272 from the Space Telescope Science Institute (STScI), which is operated by the Association of Universities for Research in Astronomy (AURA), Inc., under NASA contract NAS5-26555. E.C. expresses her profound gratitude to Alexander Rudy, a conversation with whom sparked the radial velocity measurement technique outlined in this work. E.C. also thanks Tony Sohn for his help with Figure 1. A.D. is supported by a Royal Society University Research Fellowship. A.D. also acknowledges support from the STFC grant ST/P000451/1. P.G. and E.T. acknowledge support from the NSF grants AST-1010039 and AST-1412504. We thank the outstanding team at Keck Observatory for assisting us in our observations. This research made use of Astropy, a community-developed core Python package for Astronomy (Astropy Collaboration et al. 2013, 2018). This work has made use of the Rainbow Cosmological Surveys Database, which is operated by the Universidad Complutense de Madrid (UCM), partnered with the University of California Observatories at Santa Cruz (UCO/Lick, UCSC). We recognize and acknowledge the 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.

Appendix Testing Velociraptor

In this appendix, we discuss the testing of the Velociraptor technique described in Section 3 on fake data. To create fake spectra for testing, we degraded our template spectrum to a variety of S/Ns. In order to realistically reproduce the noise due to the sky background for one of the HALO7D exposures, we took the noise array for an extremely faint extragalactic target that had no visible continuum or emission lines for a 20-minute exposure.

For S/Ns of 3, 5, 10, and 30, we generated 180 individual spectra for each S/N. We tested Velociraptor in “single mode” (i.e., working with individual observations only) for 90 of these spectra for each S/N. The mean recovered velocities and the standard deviations of these velocities are shown in the top panel of Figure 18. For the spectra with S/N = 3, approximately one-third of the fake sample had failed velocity measurements (i.e., chains did not successfully converge); the error bars in Figure 18 reflect the statistics for the successful measurements.

All projections of the posterior for one of our fake spectra can be seen in the corner plot in Figure 19. The input values for the fake spectrum are shown as blue lines; the model successfully recovers the parameters of the fake data. The absorption-line coefficients C and the continuum levels for a given spectral region are covariant; this is expected because of the way in which we parameterized the absorption-line strength (see Equation (4)).

Trace plots for 20 of the *emcee* walkers for the 11 parameters of single mode are shown in Figure 20. The true parameter values that were used to generate the fake data are shown as thick black dashed lines. The traces are well mixed and converge successfully over the runtime of the sampler.

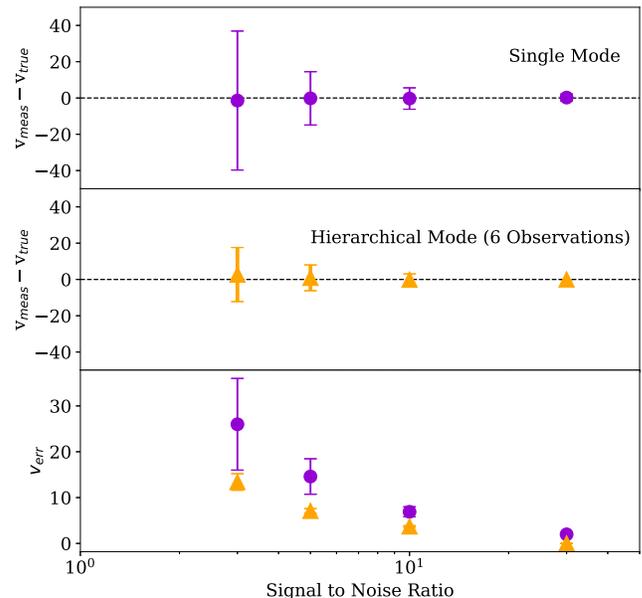


Figure 18. Results from testing Velociraptor on fake data. Top panel: error bars show the mean recovered velocity and the standard deviation of the recovered velocities for 90 runs of Velociraptor in single mode, as a function of S/N. Middle panel: resulting distributions of recovered velocities when Velociraptor is run with six observations in hierarchical mode. Note that here the x-axis refers to the S/N of a single observation. Bottom panel: velocity error (computed as half the difference between the 84th and 16th percentiles) in single mode (purple) and hierarchical mode with six observations (orange).

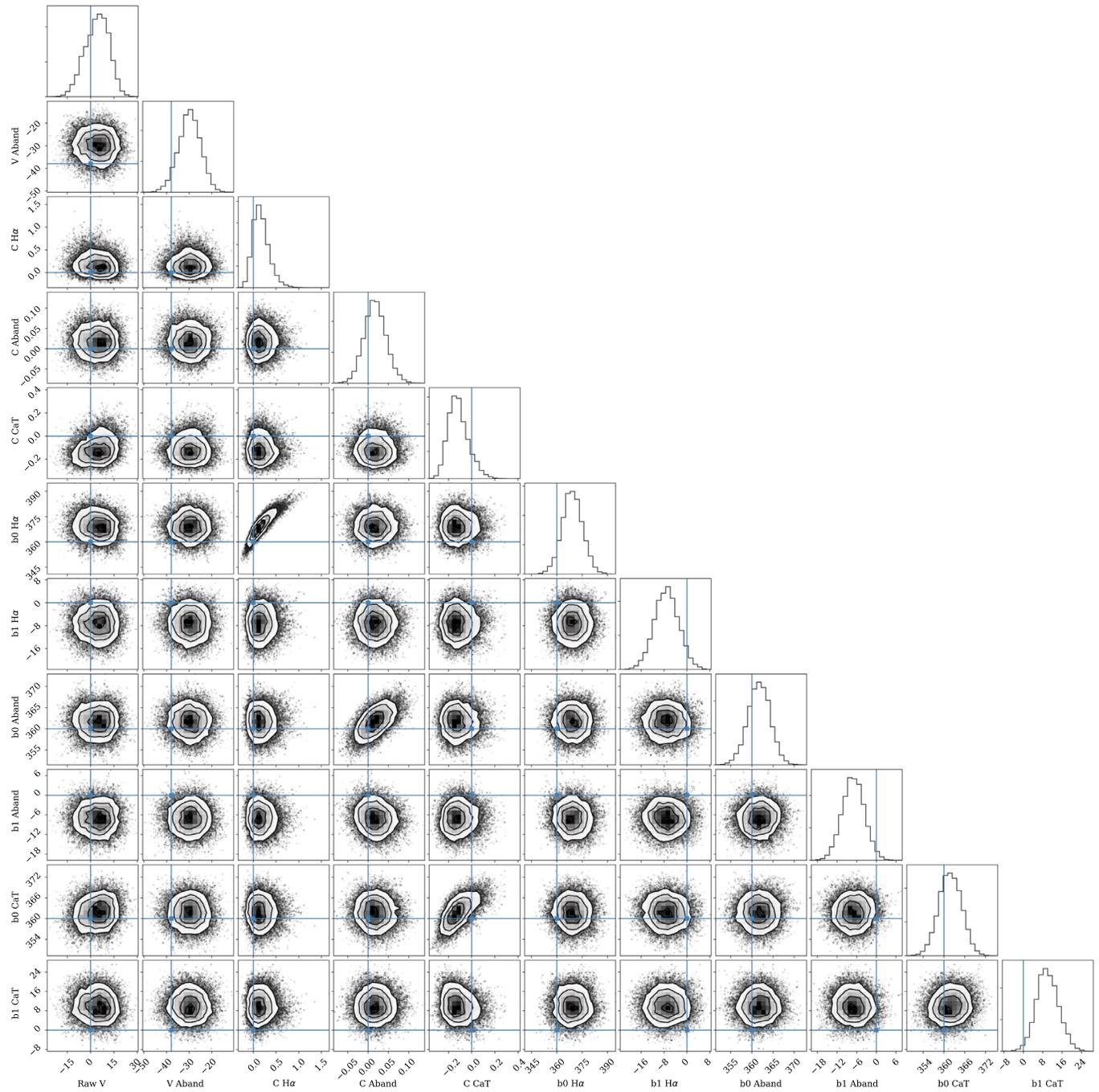


Figure 19. Full corner plot for all 11 parameters for a fake spectrum with $S/N = 10$. The true parameter values used to generate the fake spectrum are shown in blue. The absorption-line strength parameters for a given spectral region are covariant with the continuum level; this is expected based on how we have parameterized the absorption lines.

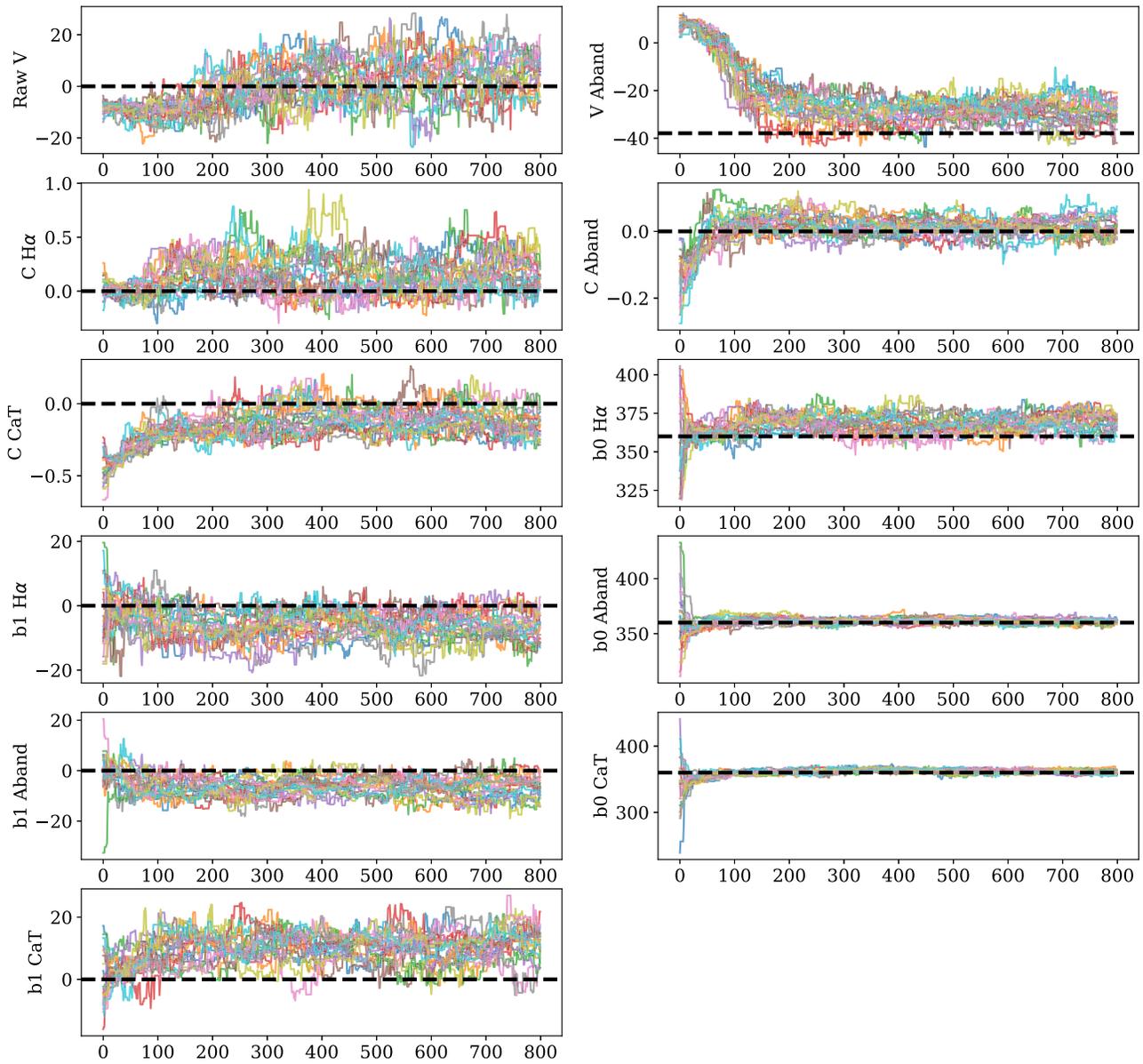


Figure 20. Traces for all 11 single-mode parameters, for a fake spectrum generated to have $S/N = 10$. For clarity, we show traces for only 20 randomly selected walkers. Black dashed lines indicate the true values of the model parameters used to generate this fake spectrum.

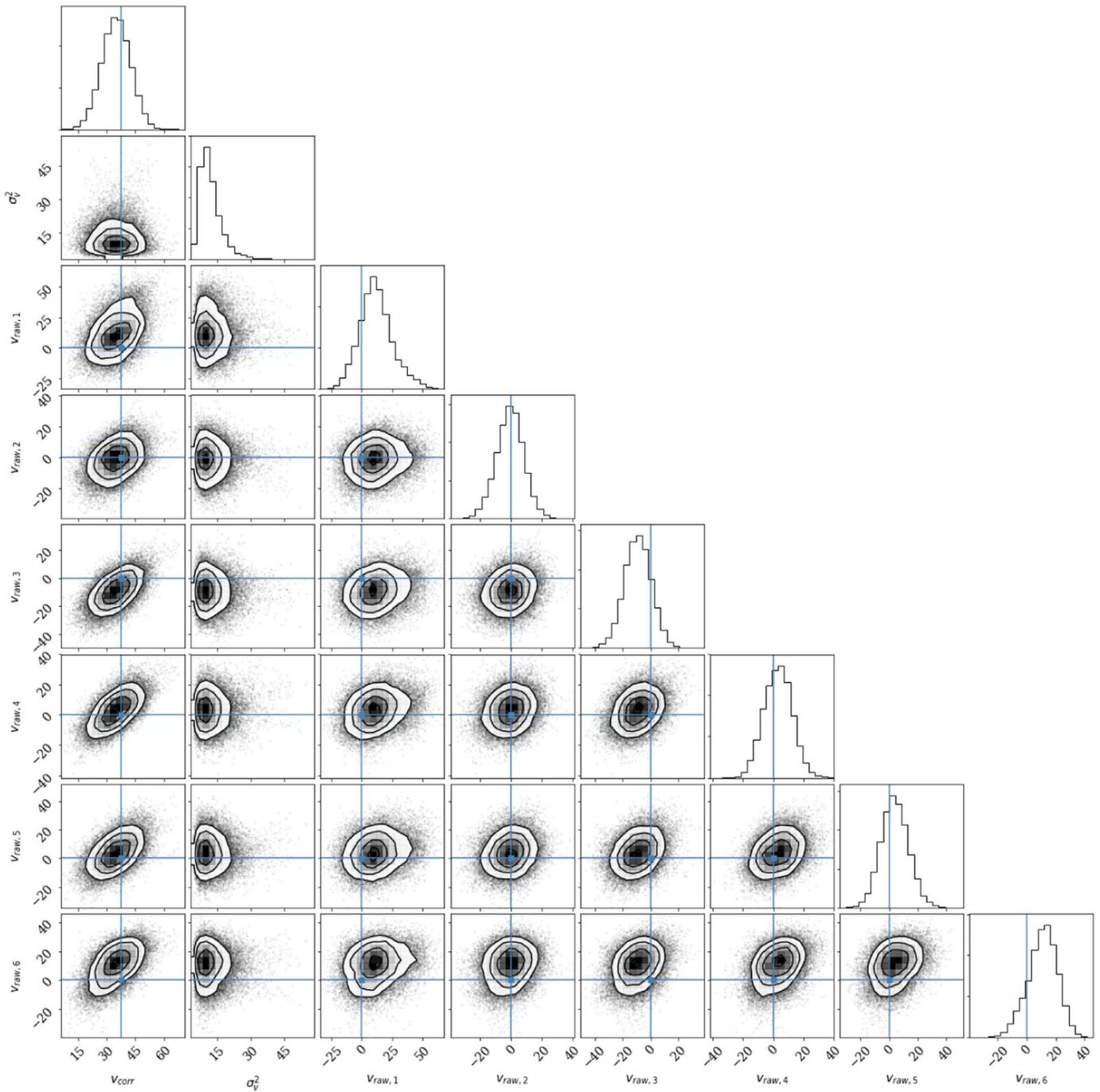


Figure 21. Corner plot for the corrected velocity, the additional uncertainty σ_v^2 , and the six raw velocities for six fake spectra that each have $S/N = 5$. Note that we are only showing projections here for 8 out of the 68 parameters in this model. This particular run of Velociraptor ran for with 800 walkers for 500 steps. For this figure, we excluded the first 3000 steps as burn-in and thinned the chain, showing every 50th sample.

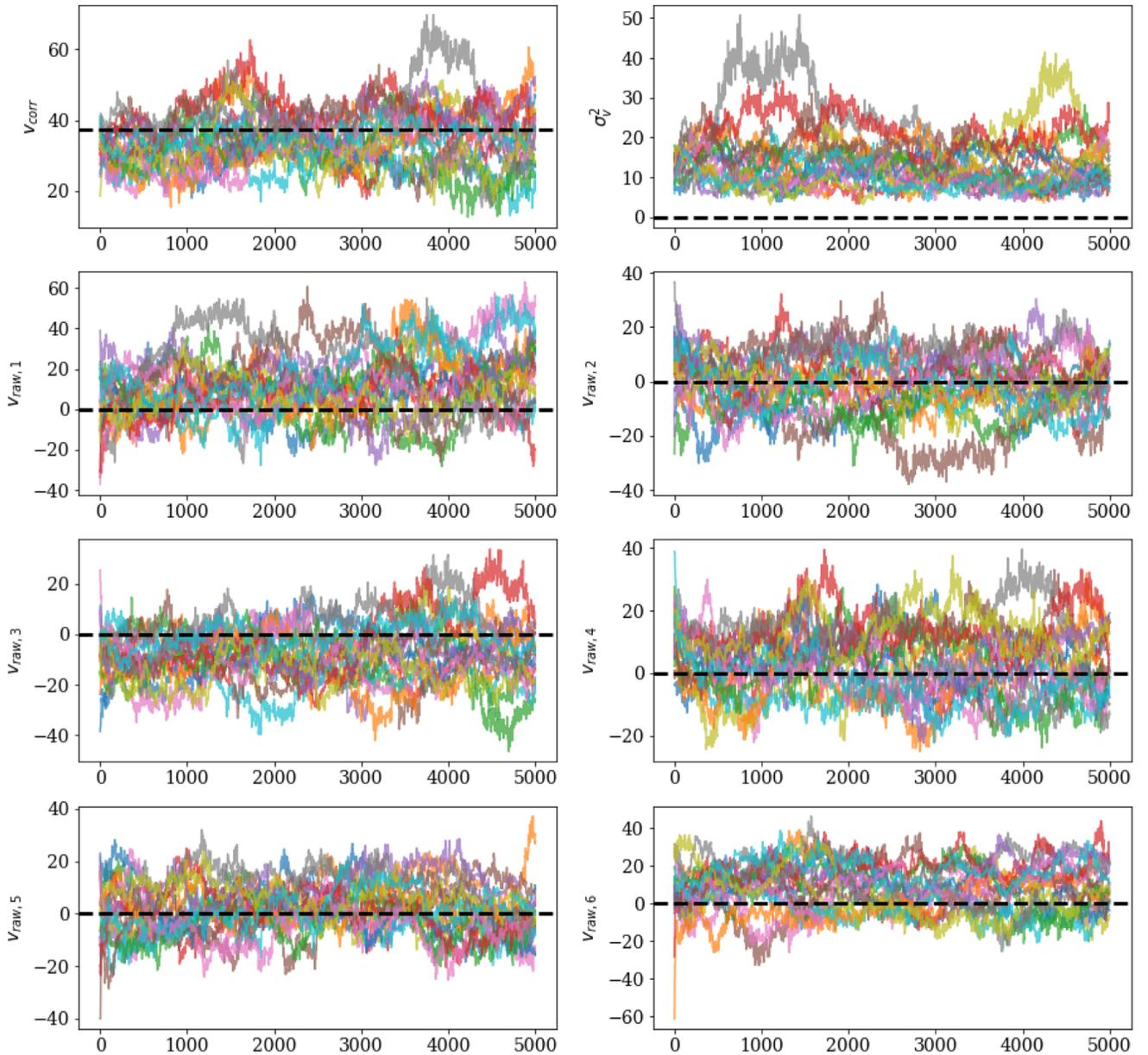


Figure 22. Traces for the corrected velocity, the additional uncertainty σ_v^2 , and the six raw velocities for six fake spectra that each have $S/N = 5$. For clarity, we show traces for only 20 randomly selected walkers. Truths are shown as black dashed lines. Because of the complexity of the model and the large number of free parameters, the chains do not mix efficiently, and the sampler needs to be run for many iterations. Note that the true value for σ_v^2 is not recovered in this case, because all of our fake spectra were generated to have exactly the same velocity.

To test Velociraptor in hierarchical mode, we combined six fake spectra at a given S/N and ran Velociraptor 30 times at each S/N . The resulting mean recovered velocities and their standard deviations are shown in the middle panel of Figure 18, where the S/N plotted on the x -axis refers to the S/N of the individual observations.

A corner plot and trace plots for the corrected velocity, the variance of velocities σ_v^2 , and the raw velocity parameters for six observations, each with $S/N = 5$, are shown in Figures 21 and 22. We run *emcee* for 5000 steps and discard the first 3000 as burn-in. For the purposes of displaying the corner plot,

we have “thinned” our chain, including every 50th sample for each walker.

ORCID iDs

Emily C. Cunningham <https://orcid.org/0000-0002-6993-0826>

Alis J. Deason <https://orcid.org/0000-0001-6146-2645>

Puragra Guhathakurta <https://orcid.org/0000-0001-8867-4234>

Evan N. Kirby <https://orcid.org/0000-0001-6196-5162>

Elisa Toloba <https://orcid.org/0000-0001-6443-5570>

Guillermo Barro <https://orcid.org/0000-0001-6813-875X>

References

- Anderson, J., & King, I. R. 2006, PSFs, Photometry, and Astronomy for the ACS/WFC, Tech. Instrument Science Report ACS-2006-01
- Astropy Collaboration, Price-Whelan, A.-M., SipHocz, B. M., et al. 2018, *AJ*, **156**, 123
- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, **558**, A33
- Barro, G., Pérez-González, P. G., Gallego, J., et al. 2011, *ApJS*, **193**, 13
- Belokurov, V., Erkal, D., Evans, N. W., Koposov, S. E., & Deason, A. J. 2018, *MNRAS*, **478**, 611
- Bertin, E., & Arnouts, S. 1996, *A&AS*, **117**, 393
- Brown, W. R., Geller, M. J., Kenyon, S. J., & Diaferio, A. 2010, *AJ*, **139**, 59
- Bullock, J. S., & Johnston, K. V. 2005, *ApJ*, **635**, 931
- Capak, P., Aussel, H., Ajiki, M., et al. 2007, *ApJS*, **172**, 99
- Cappellari, M., & Emsellem, E. 2004, *PASP*, **116**, 138
- Cohen, J. G., Sesar, B., Bahholz, S., et al. 2017, *ApJ*, **849**, 150
- Conroy, C., Bonaca, A., Naidu, R. P., et al. 2018, *ApJ*, **861**, L16
- Cooper, M. C., Newman, J. A., Davis, M., Finkbeiner, D. P., & Gerke, B. F. 2012, spec2d: DEEP2 DEIMOS Spectral Pipeline, Astrophysics Source Code Library, ascl:1203.003
- Cunningham, E. C., Deason, A. J., Guhathakurta, P., et al. 2016, *ApJ*, **820**, 18
- Cunningham, E. C., Deason, A. J., & Sanderson, R. E. 2018, arXiv:1810.12201
- Deason, A. J., Belokurov, V., & Evans, N. W. 2011, *MNRAS*, **416**, 2903
- Deason, A. J., Belokurov, V., Evans, N. W., & Johnston, K. V. 2013a, *ApJ*, **763**, 113
- Deason, A. J., Belokurov, V., & Koposov, S. E. 2018a, *ApJ*, **852**, 118
- Deason, A. J., Belokurov, V., Koposov, S. E., & Lancaster, L. 2018b, *ApJL*, **862**, L1
- Deason, A. J., Van der Marel, R. P., Guhathakurta, P., Sohn, S. T., & Brown, T. M. 2013b, *ApJ*, **766**, 24
- Faber, S. M., Phillips, A. C., Kibrick, R. I., et al. 2003, *Proc. SPIE*, **4841**, 1657
- Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP*, **125**, 306
- Gaia Collaboration 2018, yCat, I/345
- Gilbert, K. M., Guhathakurta, P., Beaton, R. L., et al. 2012, *ApJ*, **760**, 76
- Goodman, J., & Weare, J. 2010, *Communications in Applied Mathematics and Computational Science*, **5**, 65
- Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, *ApJS*, **197**, 35
- Guo, Y., Ferguson, H. C., Giavalisco, M., et al. 2013, *ApJS*, **207**, 24
- Helmi, A., Babusiaux, C., Koppelman, H. H., et al. 2018, arXiv:1806.06038
- Hernitschek, N., Cohen, J. G., Rix, H.-W., et al. 2018, *ApJ*, **859**, 31
- Johnston, K. V., Bullock, J. S., Sharma, S., et al. 2008, *ApJ*, **689**, 936
- Kalirai, J. S. 2012, *Natur*, **486**, 90
- Kallivayalil, N., van der Marel, R. P., & Alcock, C. 2006a, *ApJ*, **652**, 1213
- Kallivayalil, N., van der Marel, R. P., Alcock, C., et al. 2006b, *ApJ*, **638**, 772
- Kallivayalil, N., van der Marel, R. P., Besla, G., Anderson, J., & Alcock, C. 2013, *ApJ*, **764**, 161
- Kirby, E. N., Cohen, J. G., Guhathakurta, P., et al. 2013, *ApJ*, **779**, 102
- Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, *ApJS*, **197**, 36
- Leauthaud, A., Massey, R., Kneib, J.-P., et al. 2007, *ApJS*, **172**, 219
- Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013, *ApJS*, **206**, 8
- Nayyeri, H., Hemmati, S., Mobasher, B., et al. 2017, *ApJS*, **228**, 7
- Newman, J. A., Cooper, M. C., Davis, M., et al. 2013, *ApJS*, **208**, 5
- Perryman, M. A. C., de Boer, K. S., Gilmore, G., et al. 2001, *A&A*, **369**, 339
- Price-Whelan, A. M., Johnston, K. V., Sheffield, A. A., Laporte, C. F. P., & Sesar, B. 2015, *MNRAS*, **452**, 676
- Robin, A. C., Reylé, C., Derrière, S., & Picaud, S. 2003, *A&A*, **409**, 523
- Schönrich, R., Binney, J., & Dehnen, W. 2010, *MNRAS*, **403**, 1829
- Sesar, B., Hernitschek, N., Dierickx, M. I. P., Fardal, M. A., & Rix, H.-W. 2017, *ApJL*, **844**, L4
- Sesar, B., Jurić, M., & Ivezić, Ž. 2011, *ApJ*, **731**, 4
- Sharma, S., Bland-Hawthorn, J., Johnston, K. V., & Binney, J. 2011, *ApJ*, **730**, 3
- Simon, J. D., & Geha, M. 2007, *ApJ*, **670**, 313
- Sirianni, M., Jee, M. J., Benítez, N., et al. 2005, *PASP*, **117**, 1049
- Slater, C. T., Nidever, D. L., Munn, J. A., Bell, E. F., & Majewski, S. R. 2016, *ApJ*, **832**, 206
- Sohn, S. T., Anderson, J., & van der Marel, R. P. 2012, *ApJ*, **753**, 7
- Sohn, S. T., Majewski, S. R., Muñoz, R. R., et al. 2007, *ApJ*, **663**, 960
- Sohn, S. T., Patel, E., Besla, G., et al. 2017, *ApJ*, **849**, 93
- Sohn, S. T., Watkins, L. L., Fardal, M. A., et al. 2018, *ApJ*, **862**, 52
- Stefanon, M., Yan, H., Mobasher, B., et al. 2017, *ApJS*, **229**, 32
- Toloba, E., Li, B., Guhathakurta, P., et al. 2016, *ApJ*, **822**, 51
- VandenBerg, D. A., Bergbusch, P. A., & Dowler, P. D. 2006, *ApJS*, **162**, 375
- Watkins, L. L., Evans, N. W., Belokurov, V., et al. 2009, *MNRAS*, **398**, 1757
- Xue, X. X., Ma, Z., Rix, H. W., et al. 2016, yCat, J/ApJ/784/170
- Xue, X. X., Rix, H. W., Zhao, G., et al. 2008, *ApJ*, **684**, 1143
- Yesuf, H. M., Koo, D. C., Faber, S. M., et al. 2017, *ApJ*, **841**, 83