

## INSIGHT INTO THE FORMATION OF THE MILKY WAY THROUGH COLD HALO SUBSTRUCTURE. I. THE ECHOS OF MILKY WAY FORMATION

KEVIN C. SCHLAUFMAN<sup>1,6</sup>, CONSTANCE M. ROCKOSI<sup>1,7,8</sup>, CARLOS ALLENDE PRIETO<sup>2,9</sup>, TIMOTHY C. BEERS<sup>3</sup>, DMITRY BIZYAEV<sup>4</sup>, HOWARD BREWINGTON<sup>4</sup>, YOUNG SUN LEE<sup>3</sup>, VIKTOR MALANUSHENKO<sup>4</sup>, ELENA MALANUSHENKO<sup>4</sup>, DAN ORAVETZ<sup>4</sup>, KAIKE PAN<sup>4</sup>, AUDREY SIMMONS<sup>4</sup>, STEPHANIE SNEDDEN<sup>4</sup>, AND BRIAN YANNY<sup>5</sup>

<sup>1</sup> Astronomy and Astrophysics Department, University of California, Santa Cruz, CA 95064, USA; [kcs@ucolick.org](mailto:kcs@ucolick.org), [crockosi@ucolick.org](mailto:crockosi@ucolick.org)

<sup>2</sup> Department of Astronomy and McDonald Observatory, The University of Texas at Austin, 1 University Station, C1400, Austin, Texas 78712, USA

<sup>3</sup> Department of Physics and Astronomy, Center for the Study of Cosmic Evolution, and Joint Institute for Nuclear Astrophysics, Michigan State University, E. Lansing, MI 48824, USA; [beers@pa.msu.edu](mailto:beers@pa.msu.edu), [lee@pa.msu.edu](mailto:lee@pa.msu.edu)

<sup>4</sup> Apache Point Observatory, Sunspot, NM 88349, USA; [dmbiz@apo.nmsu.edu](mailto:dmbiz@apo.nmsu.edu), [hjbrew@apo.nmsu.edu](mailto:hjbrew@apo.nmsu.edu), [viktorm@apo.nmsu.edu](mailto:viktorm@apo.nmsu.edu), [elenam@apo.nmsu.edu](mailto:elenam@apo.nmsu.edu), [doravetz@apo.nmsu.edu](mailto:doravetz@apo.nmsu.edu), [kpan@apo.nmsu.edu](mailto:kpan@apo.nmsu.edu), [asimmons@apo.nmsu.edu](mailto:asimmons@apo.nmsu.edu), [snedden@apo.nmsu.edu](mailto:snedden@apo.nmsu.edu)

<sup>5</sup> Fermi National Accelerator Laboratory, P.O. Box 500, Batavia, IL 60510, USA; [yanny@fnal.gov](mailto:yanny@fnal.gov)

Received 2009 June 26; accepted 2009 August 10; published 2009 September 17

### ABSTRACT

We identify 10—seven for the first time—elements of cold halo substructure (ECHOS) in the volume within 17.5 kpc of the Sun in the inner halo of the Milky Way. Our result is based on the observed spatial and radial velocity distribution of metal-poor main-sequence turnoff (MPMSTO) stars in 137 Sloan Extension for Galactic Understanding and Exploration lines of sight. We point out that the observed radial velocity distribution is consistent with a smooth stellar component of the Milky Way’s inner halo overall, but disagrees significantly at the radial velocities that correspond to our detections. We show that all of our detections are statistically significant and that we expect no false positives. These ECHOS represent the observable stellar debris of ancient merger events in the stellar accretion history of the Milky Way, and we use our detections and completeness estimates to infer a formal upper limit of  $0.34^{+0.02}_{-0.02}$  on the fraction of the MPMSTO population in the inner halo that belong to ECHOS. Our detections and completeness calculations also suggest that there is a significant population of low fractional overdensity ECHOS in the inner halo, and we predict that 1/3 of the inner halo (by volume) harbors ECHOS with MPMSTO star number densities  $n \approx 15 \text{ kpc}^{-3}$ . In addition, we estimate that there are of order  $10^3$  ECHOS in the entire inner halo. ECHOS are likely older than known surface brightness substructure, so our detections provide us with a direct measure of the accretion history of the Milky Way in a region and time interval that has yet to be fully explored. In concert with previous studies, our result suggests that the level of merger activity has been roughly constant over the past few Gyr and that there has been no accretion of single stellar systems more massive than a few percent of a Milky Way mass in that interval.

*Key words:* Galaxy: formation – Galaxy: halo – Galaxy: kinematics and dynamics

*Online-only material:* color figure, machine-readable table

### 1. INTRODUCTION

The stellar halo of our Galaxy is an excellent place to study the residuals of its formation, because the timescale over which those residuals disappear is long relative to the other parts of the Galaxy. Moreover, the stellar populations in the halo of our Galaxy are predominantly old and metal-poor. This fact implies that most of the stars in the halo date to the earliest stages of the Milky Way’s formation. Indeed, Eggen et al. (1962) used the dynamical and chemical signature of high proper motion stars to deduce that metal-poor stars in the solar neighborhood are preferentially on radial orbits. They interpreted this as the result of star formation during a rapid collapse of the nascent Milky Way after its own self-gravity took over from universal expansion. Later studies of globular clusters like those of Searle & Zinn (1978) revealed that the distribution of globular cluster metallicities beyond 8 kpc from the center of the Galaxy is broad and independent of galactocentric radius. They also found that

differences in color–magnitude morphology were uncorrelated with metallicity, in contrast to expectations of a smooth collapse model. Searle & Zinn (1978) interpreted these observations as evidence of multiple episodes of star formation in the halo over an extended period. Detailed classical studies of our Galaxy as a whole (e.g., Bahcall & Soneira 1980; Gilmore & Reid 1983), and of the stellar halo in particular, showed that the halo was not well described by the same  $\rho^{1/4}$  power law that describes the bulge (e.g., Morrison 1993). Instead, the halo follows a power law  $\rho^\alpha$  with  $\alpha \approx -3.5$  (e.g., Harris 1976; Zinn 1985) and is roughly spherically symmetric at large galactocentric radii but more oblate closer in (e.g., Preston et al. 1991; Chiba & Beers 2000). These classical observations implied that stars in halo were generally, but not always (cf. Ratnatunga & Freeman 1985; Majewski 1993), smoothly distributed in both coordinate space and velocity space.

The hierarchical model of structure formation (e.g., Press & Schechter 1974; White & Rees 1978) in a  $\Lambda$ CDM universe can make predictions that match observations of large-scale structure and galaxy clustering, as well as many characteristics of individual galaxies (e.g., Bullock & Johnston 2005; Robertson et al. 2005; Springel et al. 2005; Bower et al. 2006; Croton et al. 2006; Font et al. 2006). In general, the agreement between theory and observation is best at the largest scales where the

<sup>6</sup> NSF Graduate Research Fellow.

<sup>7</sup> University of California Observatories.

<sup>8</sup> Packard Fellow.

<sup>9</sup> Current address: Mullard Space Science Laboratory, University College London, Holmbury St. Mary, Dorking, Surrey RH5 6NT, UK; [cap@mssl.ucl.ac.uk](mailto:cap@mssl.ucl.ac.uk).

physics is dominated by cosmology and dark matter. However, that agreement becomes increasingly model dependent on the smaller galactic and sub-galactic scales where baryons are important. Theorists have taken several approximate approaches to better understand the formation of Milky Way analog halos: cosmological dark matter only  $n$ -body simulations with live potentials (Diemand et al. 2007, 2008; Springel et al. 2008), cosmological dark matter only  $n$ -body simulations coupled to semi-analytic models (De Lucia & Helmi 2008), cosmological dark matter only  $n$ -body simulations without live potentials but with added resolution coupled to semi-analytic models (Harding et al. 2001; Bullock & Johnston 2005; Robertson et al. 2005; Font et al. 2006), as well as full cosmological hydrodynamic simulations (Abadi et al. 2003a, 2003b; Governato et al. 2007). While each theoretical approach has its relative advantages and disadvantages, for the time being there is no way to self-consistently track all of the necessary baryon physics important in galaxy formation over a large dynamic range in spatial scale. Nevertheless, a broad consensus has emerged: all of these approaches suggest that the inner halo formed early from the accretion of relatively massive protogalaxies into the nascent Milky Way (e.g., Bullock & Johnston 2005; Abadi et al. 2006) and that violent relaxation played a major role in producing the classically observed kinematically smooth inner halo (e.g., Diemand et al. 2005). There is also growing evidence that stars formed in the disk of the nascent Milky Way contribute to the stellar populations in the inner halo (e.g., Zolotov et al. 2009). In addition, all of these calculations predict the presence of substructure in the inner halo as a result of accretion events more recent than the last episode of violent relaxation. All of the referenced simulations are of Milky Way analogs, not the Milky Way itself, so the observations meant to verify these predictions must be statistical. We are further limited by current technology to observational comparisons of the Milky Way's halo, its satellites, and Local Group companions with these theoretical models to test their predictions for galaxy formation on small scales.

Fortunately, the halo of our own Galaxy provides us with an excellent observational example for tests of this hierarchical formation scenario (for a recent review see Helmi 2008). The halo's resolved stellar populations are bright enough to study both photometrically and spectroscopically over large fields of view with modern telescopes and instrumentation. The Sloan Digital Sky Survey (SDSS; Fukugita et al. 1996; Gunn et al. 1998; York et al. 2000; Hogg et al. 2001; Smith et al. 2002; Pier et al. 2003; Ivezić et al. 2004; Gunn et al. 2006; Tucker et al. 2006; Padmanabhan et al. 2008) is one such study that has made significant contributions to our understanding of the Galaxy. The earliest results to come out of the SDSS (Ivezić et al. 2000; Yanny et al. 2000; Chen et al. 2001) confirmed many classical results with high precision, and at same time provided more evidence that the halo was not entirely homogeneous. Those hints have been followed up with many more detailed studies. As a result of the SDSS and other modern large-scale surveys, there is now strong evidence of substructure in the halo of the Milky Way from star count maps (Totten & Irwin 1998; Totten et al. 2000; Ivezić et al. 2000; Yanny et al. 2000; Ibata et al. 2001; Odenkirchen et al. 2001; Vivas et al. 2001; Gilmore et al. 2002; Newberg et al. 2002; Rockosi et al. 2002; Majewski et al. 2003; Yanny et al. 2003; Rocha-Pinto et al. 2004; Duffau et al. 2006; Belokurov et al. 2006; Grillmair & Johnson 2006; Grillmair & Dionatos 2006; Vivas & Zinn 2006; Belokurov et al. 2007; Bell et al. 2008; Jurić et al. 2008; Grillmair 2006, 2009; Watkins

et al. 2009), kinematic information (Chiba & Yoshii 1998; Helmi et al. 1999; Chiba & Beers 2000; Kepley et al. 2007; Ivezić et al. 2008; Klement et al. 2008; Seabroke et al. 2008; Klement et al. 2009; Smith et al. 2009; Starkenburg et al. 2009), and chemical abundances (Ivezić et al. 2008). The data from these large-scale surveys have pushed the field beyond individual detections to systematic statistical searches in which both detections and non-detections are meaningful and can strongly inform theoretical models. Bell et al. (2008) used photometric overdensities in projected SDSS star counts to statistically quantify the degree of substructure in the outer halo. They showed that while the classical symmetric smooth model is a poor match to the observations, the amount of substructure in the outer halo is consistent with that expected from  $\Lambda$ CDM simulations. Likewise, Carollo et al. (2007) revealed the two-component (inner/outer) nature of the halo, a property that is naturally explained in the hierarchical model. Accordingly, now that the existence of substructure in the halo of our own Galaxy has been well established, our challenge is to use the observations of substructure to better understand the current state of the Milky Way (e.g., Helmi 2004; Johnston et al. 2005; Meza et al. 2005; Fellhauer et al. 2006; Helmi et al. 2006; Willett et al. 2009), to inform models of its formation, and to study the stellar populations and thereby the star formation processes at work in accreted protogalaxies and the disk of a nascent Milky Way.

It was recognized at least as early as Eggen et al. (1962) that searching for substructure in the halo of our own Galaxy using kinematic information would be very informative, because the long dynamical times in the halo ensure that substructure remains kinematically distinct for Gyr. This is in spite of potential degeneracies between progenitor mass, velocity dispersion, and accretion time (e.g., Johnston 1998; Helmi & White 1999; Johnston et al. 2008). Ideally, that kinematic information would include all six phase-space coordinates. Indeed, surveys of the solar neighborhood using the precise proper motions available in existing data sets have constrained both the absolute number of dynamically distinct features in the local halo as well as the fraction of the local halo population that belong to those features (e.g., Helmi & White 1999; Helmi et al. 1999; Gould 2003). More recently, Morrison et al. (2009) analyzed available local volume data and suggested that the data implied that violent relaxation were not efficient in the early Milky Way.

Unfortunately, existing proper motion catalogs extending beyond 10 kpc (e.g., USNO-B; Monet et al. 2003) can only provide 3 mas yr<sup>-1</sup> precision (Munn et al. 2004). This precision corresponds to a tangential velocity error of about 140 km s<sup>-1</sup> at 10 kpc, a value at least as large as the halo velocity dispersion at that distance. Note that this error estimate also neglects the fact that imprecise distance estimates can degrade accuracy further. At the same time, today's ground-based radial velocity surveys like the SDSS, the Sloan Extension for Galactic Understanding and Exploration (SEGUE; Yanny et al. 2009; Allende Prieto et al. 2008; Lee et al. 2008a, 2008b), and the RAdial Velocity Experiment (RAVE; Steinmetz et al. 2006; Zwitter et al. 2008) provide precise radial velocities independent of any uncertainty in distance. In the case of SEGUE, the radial velocity precision is at least an order of magnitude better at 10 kpc than the tangential velocity precision provided by USNO-B. In the future, these radial velocity surveys will be complemented by 10<sup>9</sup> space-based proper motions and 10<sup>8</sup> radial velocities from the Gaia satellite (Perryman et al. 2001; Lindegren et al. 2008). Nevertheless, a search for substructure in radial velocity is timely and beyond 10 kpc radial velocities will

remain the most precisely estimated phase-space component for the foreseeable future. Moreover, the information content of radial velocity data relative to tangential velocity data is substantial (e.g., Bovy et al. 2009).

Determining the origin of substructures detected using only radial velocity data is theoretically intractable because the properties of the substructure left behind by an accretion event observable in radial velocity are degenerate in progenitor mass, velocity dispersion, orbital properties, and time since accretion. Recently accreted, massive, and compact substructure progenitors are likely to produce broad radial velocity features, while long ago accreted, low mass, and diffuse substructure progenitors are likely to produce narrow radial velocity features. Nevertheless, the substructure left behind by a long ago accreted massive progenitor can effectively masquerade as the substructure that results from the more recent disruption of a less massive progenitor, or as the substructure left behind by the even more recent accretion of an even less massive but more diffuse progenitor. Precisely measured proper motions, parallaxes, iron metallicities, or  $\alpha$ -enhancements can potentially break these degeneracies and play an important role in connecting observations of cold substructure to the properties of their progenitors.

In this paper, we describe a systematic, statistical search for elements of cold halo substructure (ECHOS) in the inner halo using the SDSS-II Data Release 7 radial velocity data. Radial velocities are the most precisely measured property of SEGUE stars in the inner halo at distances greater than about 10 kpc. Restricting our search to radial velocities allows us to avoid the difficult-to-determine, heteroskedastic, and model-dependent systematic errors potentially associated with both proper motions and spectrophotometric parallaxes at that distance. From this point forward, ECHOS are defined as that substructure that manifests itself as an overdensity in radial velocity space. Note that ECHOS are not necessarily equivalent to the tidal streams that have been discovered as photometric overdensities; to emphasize that point, we do not use the word stream to describe any of our detections. “Cold” specifically implies that the inherent radial velocity dispersion of the substructure we seek is less than (or of order) our SEGUE radial velocity error estimates at a given S/N. In other words, cold substructure is that substructure for which the inherent velocity dispersion is unresolved (or barely resolved) in radial velocity in the SEGUE data. Our search for cold substructure has at least two advantages relative to a search for warmer substructure. First, the radial velocity scale of the search is set naturally by the observational errors and not artificially at some arbitrary value. Second, as we will subsequently show, our sensitivity to substructure is not dependent on the velocity dispersion of that substructure. Our search will not necessarily be sensitive to substructure that is now so well mixed that its constituent stars are no longer close to each other in velocity or position along their orbit, that is, we are not sensitive to the substructure that is now so dispersed that it can only be recovered by examining integrals of motion (e.g., Helmi & White 1999; Helmi et al. 1999; Klement et al. 2008, 2009). Our search will also not necessarily be sensitive to early accretion events that experienced the effects of violent relaxation. Nevertheless, our search for ECHOS still at least partially exploits the desirable property of collisionless systems that substructure remains coherent in velocity space for much longer than it remains coherent in position space. Therefore, we are likely to find in our search volume substructures that are on average older than those substructures discovered as photometric overdensities in

the same volume. Our search technique can naturally be used to estimate our completeness, and that property enables us to extrapolate our result to determine the fraction of metal-poor main-sequence turnoff (MPMSTO) stars in the inner halo that belong to ECHOS. That estimate can reveal the level of merger activity over the past few Gyr. Subsequent spectroscopic follow-up of the stars in individual ECHOS could probe through chemical abundances the physics of the high redshift universe and the star formation environments in the disk of the nascent Milky Way and in the massive protogalaxies accreted by the Milky Way over the past few Gyr. Finally, theoretical hierarchical models of Milky Way analog formation in a  $\Lambda$ CDM universe make predictions about the existence, properties, and degree of halo substructure that are testable by our search, and our results have the potential to uniquely inform models of galaxy formation.

This paper is organized as follows: in Section 2, we describe the SEGUE survey and the data we use in this analysis. In Section 3, we describe how we identify ECHOS in the observations, how we estimate our false positive rate and our completeness, and we present our detections. In Section 4, we generalize our results to the entire inner halo. In Section 5, we discuss the implications of our findings for the formation of the Milky Way. We summarize our conclusions in Section 6.

## 2. DATA

The SEGUE survey obtained 240,000 moderate-resolution ( $R \approx 1800$ ) fiber-fed spectra of Milky Way stars in the magnitude interval  $14.0 < g < 20.3$ . The spectroscopic targets were selected using SDSS photometry and additional SEGUE *ugriz* imaging data at low Galactic latitude and in the South Galactic Cap. The SEGUE spectra were taken on 212 pointings spread sparsely over the 11,663 square degrees in the combined imaging surveys. Each pointing covers a field of seven square degrees,  $3^\circ$  in diameter, and has 1180 science targets taken on two 640 fiber spectroscopic “plates.” The targets are split between a bright and faint plate in each pair at  $r = 17.8$  (for  $g - r < 0.55$ ). The instrumentation, data processing pipelines, survey strategy, along with radial velocity and atmospheric parameter accuracies are described in Yanny et al. (2009), Lee et al. (2008a, 2008b), Allende Prieto et al. (2008), and the SDSS-II DR7 paper (Abazajian et al. 2009).

The radial velocity accuracy of the SEGUE data is discussed in detail in Yanny et al. (2009). The systematic uncertainty is estimated using repeat observations. There are 20 randomly selected stars with magnitudes  $r \approx 17.8$  repeated on the bright and faint plates in each pair. The two plates in each pointing are observed independently, often on different nights, lunations, or even years; so these duplicates are a fair test of the velocity zero-point shifts. From these and an additional 12 plates with mostly stellar targets repeated in the course of the SDSS and SEGUE, we find that the mean plate-to-plate variation is zero with standard deviation  $1.8 \text{ km s}^{-1}$ . We use the duplicates and a realistic noise model applied to SEGUE spectra observed at high resolution (Allende Prieto et al. 2008) to characterize the uncertainty in the radial velocities as a function of signal to noise. We find that the uncertainty for stars at the mean  $g-r$  color of our sample and  $[\text{Fe}/\text{H}] \approx -1.5$  (typical of the inner halo) is  $5.3 \text{ km s}^{-1}$  at  $g = 18$  and rises to  $20 \text{ km s}^{-1}$  at  $g = 20.3$ .

In this analysis, we use the subset of SEGUE targets selected as metal-poor main-sequence turnoff (MPMSTO) stars. These

stars are good tracers of the inner halo because of their relatively large luminosity and high number density as compared with more evolved stars. At  $g = 20$  our MPMSTO stars have distances greater than 12 kpc, which for most Galactic latitudes is well out of the disk and in a region dominated by the halo.

We draw our MPMSTO sample from two of the SEGUE target selection categories. The first category is the “F turnoff” stars as described in Yanny et al. (2009). This is a UV excess selection in the  $ugr$  color–color diagram, and is designed to preferentially pick out blue, less line-blanketed halo stars rather than more metal-rich thick disk stars. At moderate and high Galactic latitude the thick disk stars are the majority population at the turnoff for magnitudes brighter than about  $r = 18.5$ . The selection criteria are  $g < 20.3$ ,  $0.4 < u - g < 1.4$ ,  $0.2 < g - r < 0.7$ , and  $-0.7 < P1(s) < -0.25$ , where  $P1(s)$  is defined to be a color parallel to the blue branch of the  $ugr$  stellar locus,  $0.91(u - g) + 0.415(g - r) - 1.28$  (see Helmi et al. 2003). Up to 200 fibers per pointing were allocated to objects in this selection category. If there were more than 200 candidates, as was usually the case, the spectroscopic targets were randomly selected from the candidate list with some preference for brighter and bluer stars.

The second SEGUE target category we drew from for this analysis was the red half of the “BHB” selection. Because blue horizontal branch (BHB) stars are rare and valuable tracers of the distant halo, the SEGUE selection used a generous red limit of  $g - r < 0.2$ . This has significant overlap with the main-sequence turnoff, and the reddest fraction of BHB candidates selected this way contains very few true BHBs. We add the red BHB targets with  $0.1 < g - r < 0.2$  to the F turnoff targets to create our MPMSTO sample. The selection criteria for both the BHB and F turnoff categories evolved slightly over the course of the survey, but not enough to significantly alter the global properties of our sample. In any case, the SEGUE targets were selected homogeneously along each individual line of sight. Since we conduct our search and evaluate our detection efficiency independently for each line of sight, any changes between lines of sight should not affect our analysis.

We use an M 13 globular cluster fiducial from An et al. (2008), augmented at the faint end by the M 13 fiducial in Clem et al. (2008), to obtain approximate distances to our MPMSTO stars. We do this to determine the range in distance and Galactic coordinates over which the SEGUE MPMSTO stars sample the halo. We use a cluster fiducial rather than a theoretical isochrone because the former is likely to be a better match to the data at the turnoff; we choose M 13 because it is near the mean metallicity of the halo (e.g., Ryan & Norris 1991a, 1991b; Carollo et al. 2007). We emphasize that we do not use these distances in our search for substructure. We find that 95% of the MPMSTO stars in the SEGUE sample are within 17.5 kpc of the Sun. We therefore impose a maximum distance limit of 17.5 kpc on the sample. To minimize contamination from the bulge and thick disk, we further restrict the sample to be in the inner halo. We define the inner halo as stars that are: more than 10 kpc from the center of the Galaxy, at vertical distance  $|z|$  more than 4 kpc from the Galactic plane, and at distances from the Sun less than 17.5 kpc. From this point onward, when we refer to a MPMSTO star, we mean one that is in the inner halo as defined above. This is consistent with the work of Carollo et al. (2007) who showed that at a distance of 17.5 kpc from the Sun the halo is dominated by their inner halo component. Of the 43,000 stars in our combined MPMSTO sample, 10,739 are in the inner halo as defined above. We further impose line of sight specific

distance limits that contain 95% of the photometrically selected MPMSTO candidates. We use these 95% limits to eliminate outliers that could significantly skew the distance thresholds. The median number of MPMSTO star spectra per line of sight in our final sample is 77.

The cuts on Galactocentric radius and distance from the plane effectively remove the brightest objects, with 95% of the final sample at  $r > 18.26$ . For the final inner halo sample, the effect of the color selections described above was such that 95% of the stars have  $-0.59 < P1(s) < -0.19$ ,  $0.63 < u - g < 1.1$ , and  $0.12 < g - r < 0.34$ . For the subset of our inner halo sample brighter than about  $g = 19$ , we can use the metallicities from the SEGUE stellar parameters pipeline (Lee et al. 2008a) to estimate the metallicity bias introduced by the UV excess and blue  $g-r$  selection cuts. We find that the  $[Fe/H]$  distribution is approximately Gaussian, with mean  $-1.62$  and standard deviation 0.5. These are consistent with other estimates of the halo metallicity distribution (Ryan & Norris 1991a, 1991b; Carney et al. 1996; Allende Prieto et al. 2006) from which we conclude that any metallicity bias in our sample is not serious. At magnitudes  $r > 19$ , the errors in the  $u$  magnitude increase rapidly. The average SDSS  $u$  point-spread function (psf) magnitude error for stars at  $g = 19$  at the mean color of our sample is 0.06. The increasing  $u$  error causes the UV excess selection to become inefficient at faint magnitudes, reducing any bias even further.

### 3. IDENTIFICATION OF COLD SUBSTRUCTURE

We search along each SEGUE line of sight for statistically significant differences between the observed MPMSTO star radial velocity distribution and the radial velocity distribution that would result from SEGUE observations of a smooth inner halo. In our Monte Carlo approach to the identification of ECHOS in the inner halo, we first generate a mock catalog of synthetic stars distributed according to published empirical models for the position and velocity distributions of inner halo stars. We then sample the mock catalog line of sight by line of sight in the same way that SEGUE sampled the Milky Way’s inner halo to obtain the radial velocity distribution that would result from observing a smooth inner halo. We employ this derived radial velocity distribution as our null hypothesis. We use two independent algorithms to systematically search for radial velocities at which there are statistically significant excesses of MPMSTO stars relative to the smooth model. The fact that our search is both systematic and statistical means that we can accurately quantify our false positive rate and completeness. From that information, we can derive an upper limit on the fraction of MPMSTO stars in the inner halo that belong to ECHOS, as well as an estimate for the absolute number of ECHOS like those we find in the inner halo.

#### 3.1. Simulation of the Radial Velocity Distribution of the Inner Halo’s Smooth Component as Viewed From the Sun

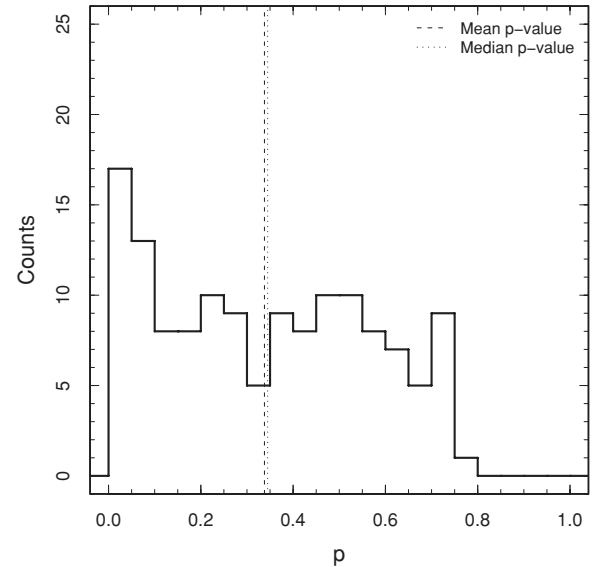
To construct our null hypothesis, we simulate the radial velocity distribution of the smooth inner halo as viewed by an observer in the plane of our Galaxy 8 kpc from its center and moving with the Sun. We include a detailed description of the procedure we use in Appendix A and briefly describe our method here. We start by creating a very large mock catalog of synthetic stars, each with its own position and velocity coordinate drawn randomly from empirically determined distributions for the galactocentric position space and velocity space structure of the

smooth inner halo. Specifically, we use a spherically symmetric distribution in galactocentric position-space  $\rho \propto r^\alpha$  with  $\alpha = -3.5$  (e.g., Morrison et al. 2000; Yanny et al. 2000; Bell et al. 2008) and a galactocentric velocity ellipsoid selected to match previous observations of the inner halo in the range of galactocentric radii we search (e.g., Sommer-Larsen et al. 1997; Sirko et al. 2004a, 2004b; Xue et al. 2008)

$$\Sigma_{r,\theta,\phi} = \begin{pmatrix} 120^2 & 0 & 0 \\ 0 & 100^2 & 0 \\ 0 & 0 & 100^2 \end{pmatrix}. \quad (1)$$

The velocity ellipsoid results from diagonalizing the symmetric stress tensor in the Jeans equations (e.g., Binney & Tremaine 1987); the numerical values in Equation (1) are in units of  $\text{km s}^{-1}$ . Our final result is insensitive to the power law exponent and the inner halo velocity ellipsoid. We tried  $-4 < \alpha < -3$  and  $100 \text{ km s}^{-1} < \sigma_r < 120 \text{ km s}^{-1}$  and arrived at the same set of detections; the exact parameters of the power law and the velocity ellipsoid are not as important as the functional forms themselves. To ensure that the smooth model is a reasonable null hypothesis for a substructure search, for each line of sight we use a two-sample Kolmogorov–Smirnov (KS) test on the observed MPMSTO radial velocity distribution and on the radial velocity distribution derived from observing our smooth model. Large  $p$ -values for a great majority of our lines of sight would demonstrate that the smooth model is not too gross an approximation of the radial velocity distribution of the inner halo. Figure 1 indicates that most lines of sight have KS  $p$ -values  $\gtrsim 0.05$  characteristic of a common parent distribution for both observations. Those lines of sight with small  $p$ -values are those lines of sight along which we subsequently identify prominent ECHOS. In other words, for the majority of our lines of sight we find that the smooth model is not obviously incorrect; for those lines of sight where it is obviously incorrect as indicated by the KS tests, we subsequently find ECHOS with more sensitive algorithms discussed in the next subsection. In any case, we do not use the results of these KS tests in our search for substructure, and we find that our observed radial velocity distributions give us no reason to reject our fiducial null hypothesis. However, we cannot reject the possibility that the inner halo is entirely made up of elements of substructure below our sensitivity thresholds that at our velocity resolution masquerade as a kinematically smooth population.

We find that we need  $n_s \sim 10^7$  synthetic stars in the mock catalog to ensure that we have at least an order of magnitude more synthetic stars than SEGUE MPMSTO stars along a given line of sight. The extra synthetic stars are necessary to robustly sample the radial velocity distribution along each line of sight. We then project the full three-dimensional velocity of each synthetic star onto the line of sight between the star and an observer in the plane of the Galaxy 8 kpc from its center and moving with the space velocity of the Sun (Dehnen & Binney 1998). As a result, we can characterize the radial velocity distribution of the smooth inner halo as viewed from Sun. We sample the radial velocity distribution of the mock catalog in exactly the same way that SEGUE sampled the radial velocity distribution of MPMSTO stars in the inner halo of the Galaxy. Everything is the same: we sample the same number of radial velocities from the mock catalog that SEGUE obtained along each line of sight and we use the line-of-sight specific heliocentric distance thresholds discussed in Section 2 to ensure that the radial velocity distribution we obtain from the mock catalog is an accurate representation of the smooth



**Figure 1.**  $p$ -value distribution from line of sight by line-of-sight KS tests comparing the observed SEGUE MPMSTO star radial velocity distributions with the radial velocity distributions derived from our mock catalog for that line of sight. Recall that  $p$ -values from the KS test give the probability that the two data sets under comparison are drawn from the same distribution. Therefore,  $p$ -values  $\gtrsim 0.05$  are usually a sign that both samples under comparison are plausibly drawn from the same parent distribution. As such, we see no reason to reject the radial velocity distribution produced by observing our mock catalog for the majority of the lines of sight we study. The pile-up at small  $p$ -values is caused by the presence of substructure along those lines of sight. One caveat is that we are not necessarily sensitive to an inner halo entirely populated by an ensemble of very diffuse substructures, as the velocity distribution of the MPMSTO sample in that scenario could very well resemble the distribution that results from a kinematically smooth model.

inner halo along that line of sight. It is important to keep in mind the fact that the radial velocity distribution observed by SEGUE in kinematically smooth halo—our null hypothesis—is not a strong function of the distance distribution of synthetic stars. As a result, our detections themselves are not sensitive to any small biases present in the selection of our halo tracer population. Therefore, any possible small bias in the SEGUE data to nearer MPMSTO stars will not affect our result. In short, the radial velocity distribution that results from sampling our mock catalog matches as closely as possible the distribution that SEGUE would have observed in a kinematically smooth inner halo. Statistically significant departures from this null hypothesis are the substructure we seek.

### 3.2. Detection Algorithms

For every SEGUE line of sight, we compare the simulated radial velocity distribution of the smooth inner halo described in Section 3.1 with the observed MPMSTO star radial velocity distribution derived for that line of sight. We analyze each  $3^\circ$  diameter SEGUE line of sight independently. To see why, recall that the Sagittarius tidal stream is about  $5^\circ$  wide and is by far the most substantial known element of high latitude substructure. The Orphan stream is about  $2^\circ$  wide and is more representative of the currently known substructure population as a whole. Thus, the  $3^\circ$  diameter SEGUE plates are likely well matched to the angular size of inner halo substructure. In addition, the Sculptor dwarf is the largest dwarf spheroidal galaxy in projected area on the sky and it subtends only about 0.4 square degrees. Individual SEGUE lines of sight are separated by  $10^\circ$ – $20^\circ$  from their nearest neighbor, and the angular correlation function between SEGUE lines of sight peaks at about  $80^\circ$ .

Therefore, the characteristic angular scale of the objects we seek is much less than the typical pairwise line-of-sight angular separation. For this reason, we assert that the short axis of any potential element of substructure is unlikely to intersect two different  $3^\circ$  diameter SEGUE lines of sight. Along the long axis of potential streams, the characteristic  $80^\circ$  angular distance between lines of sight ensures that the velocity gradient along the orbit of the stream will produce distinct velocity peaks. Without a model for the accretion history of the Milky Way that allows us to a priori associate multiple lines of sight with distinct velocity peaks, analyzing multiple lines of sight together does not provide any extra sensitivity to individual elements of substructure. Taken as whole, these numbers suggest that each SEGUE line of sight probes a distinct piece of the inner halo and that data obtained along one line of sight is unlikely to add significantly to the detectability of an element of substructure in a neighboring line of sight. At the same time, our completeness calculations in Section 3.5 indicate that along almost all individual lines of sight, we have enough MPMSTO radial velocities to both characterize the background smooth halo and to resolve ECHOS.

ECHOS in the observed radial velocity distribution manifest themselves as relative overdensities at their mean radial velocities. Our task is therefore to differentiate the radial velocity overdensities that result from the presence of ECHOS along a line of sight from those that are produced by chance through random sampling of the smooth inner halo. Note that we do not use our admittedly imprecise spectrophotometric parallaxes as we search for substructure. We only use them to determine the lower and upper bounds of the observed column along each line of sight. We use that information to sample our mock catalog to determine the radial velocity distribution of a smooth inner halo along that line of sight. That is, we only used the spectrophotometric parallaxes in the construction of our null hypothesis. We describe the two algorithms we employed to solve this problem in the following two subsections.

### 3.2.1. The Bin Algorithm

Our first approach is to compare the radial velocity histogram derived from the observed MPMSTO sample with an equivalent observation of our mock catalog. First, we compute a histogram that describes the distribution of MPMSTO star radial velocities along a given line of sight, and we use bootstrap resampling to quantify the uncertainty in the number of counts in each histogram bin. We use bins of  $20 \text{ km s}^{-1}$  width both because that is approximately our median radial velocity precision at the faint end of our sample and because that width minimizes Poisson noise while still giving us sensitivity to cold substructure. We repeatedly resample our mock catalog along that line of sight to determine the median histogram and the associated distributions for the number of counts in each bin that would result under the null hypothesis. We can then compare the two histograms and flag velocity bins that have a statistically significant overabundance of SEGUE MPMSTO stars relative to the mock catalog. Note that once we identify candidate ECHOS we could estimate the number of MPMSTO stars that belong to that element of substructure and then redo our calculation, taking into account the reduced number of MPMSTO stars that we suspect belong to the smooth background along that line of sight. We do not choose to do this because it could produce overly optimistic significance levels for our detections. We describe this algorithm in detail in Appendix B.

### 3.2.2. The Peak Algorithm

Our second approach is to compare the steepness of the cumulative distribution function (CDF) as a function of radial velocity as derived from the observed MPMSTO sample with an equivalent observation of our mock catalog. First, we compute the CDF that describes the distribution of MPMSTO radial velocities along a given line of sight. Next, we repeatedly resample our mock catalog along that line of sight to determine the average CDF that would result under the null hypothesis. Recall that the CDF  $F(x)$  of a discrete data set  $x_1, x_2, \dots, x_n$  drawn from the discrete random variable  $X$  is a monotonically increasing function that is discontinuous at each  $x_i$  and mathematically defined as

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} P(X = x_i). \quad (2)$$

Therefore, an overdensity of stars at the mean radial velocity of an ECHOS would appear as a much steeper set of steps in the observed CDF than was expected at that velocity. We can calculate line of sight and radial velocity specific significance thresholds for cold substructure by quantifying how steep we expect steps to be at that radial velocity under the null hypothesis. Then we can flag any set of steps with steepness statistic  $\Theta(v_r)$  above its local threshold as a potential ECHOS. As with the bin algorithm, we do not iteratively adjust the number of stars in the smooth background component of the inner halo for those lines of sight where we detect ECHOS. We describe this algorithm in detail in Appendix C.

### 3.3. Expected False Positive Rate

The insight into the formation of the Milky Way that we hope to gain from our search for ECHOS depends crucially on both the number of candidate ECHOS we find and on their properties. In order to be certain that our candidate ECHOS are real features in the inner halo (and not just chance projections in radial velocity space), we estimate our false positive rate with a Monte Carlo simulation. We generate two independent mock catalogs denoted  $M_1$  and  $M_2$  each with more than  $10^7$  synthetic stars and radial velocities distributed as described in Appendix A. For each SEGUE line of sight, we randomly select from  $M_2$  a sample  $S_r$  of  $n$  synthetic stars equal to the number of SEGUE MPMSTO star spectra available along that line of sight from the  $m \gg n$  synthetic stars available. We treat this sample  $S_r$  as the data and put it aside. We can then analyze the data in  $S_r$  using the algorithms described in Sections 3.2.1 and 3.2.2 with mock catalog  $M_1$  as our null hypothesis.

We examine all 137 lines of sight 10 times and count the number of detections—all of which are chance projections in radial velocity space that falsely appear to be ECHOS. We find that the expected number of false positives produced by the basic bin algorithm described in Section 3.2.1 is less than one over our entire search and therefore requires no further additions. We find that the expected number of false positives produced by the basic peak algorithm described in Section 3.2.2 is greater than one over our entire search. To understand why, recall that we analyze 137 lines of sight at each of 1,000 points in radial velocity space at which there could be a significant detection. If all of the tests are independent, the formal probability of at least one false positive is  $P(X) = 1 - (1 - 1/10^4)^{137 \times 1000} \sim 1$ . Fortunately, the tests are not independent (because of the smoothing we applied) and the peak algorithm can be modified such that the false positive rate is nearly zero. We find through our false positive

**Table 1**  
Summary of Bin Detections

ID	R.A.	Decl.	$l$	$b$	$d^a$	$v_r^b$	$n_s^c$	$N_s^d$	Figure	Comment
B-1	21.3	39.6	130	-22.8	$12.9^{+5.0}_{-1.5}$	-130	12	34	2	...
B-2	17	0	132	-62.6	$10.5^{+5.2}_{-4.5}$	-170	20	109	3a	...
B-3	39.7	28.2	150	-29	$10.1^{+2.3}_{-1.3}$	-50	17	59	3b	Mon Strm
B-4	163.8	48	162.4	59.2	$7^{+4.1}_{-1.8}$	-130	25	150	4a	GD Strm
B-5	129.6	53.9	164.3	37.2	$9.1^{+2.7}_{-2.1}$	-10	20	93	4b	...
B-6	124.5	38	183.4	32.6	$8.7^{+1.0}_{-0.6}$	30	17	83	5a	...
B-7	132.6	6.1	221.5	29.2	$10^{+2.0}_{-0.9}$	70	17	69	5b	...
B-8	134	3.2	225.2	29	$10.5^{+0.5}_{-1.4}$	90	19	74	6a	Mon Strm

**Notes.**

<sup>a</sup> Median heliocentric distance in kpc of MPMSTO stars in the significant bin.

<sup>b</sup> Central radial velocity of the significant bin in  $\text{km s}^{-1}$ .

<sup>c</sup> Number of MPMSTO star radial velocities within the significant bin.

<sup>d</sup> Number of MPMSTO star spectra obtained along that line of sight that lie in the inner halo as defined in Section 2.

**Table 2**  
Summary of Class I Peak Detections

ID	R.A.	Decl.	$l$	$b$	$d^a$	$v_r^b$	$\sigma^c$	Err <sup>d</sup>	$n_s^e$	$N_s^f$	Figure	Comment
PCI-1	214.8	56.4	100.7	56.8	$10.4^{+2.9}_{-2.8}$	-328	15.1	11.5	8	122	6b	...
PCI-2	20	31.7	130	-30.8	$12.5^{+3.9}_{-2.2}$	-125	22	17.8	20	93	7a	...
PCI-3	21.3	39.6	130	-22.8	$13^{+4.6}_{-1.6}$	-121	10.5	19.4	13	34	2	...
PCI-4	39.7	28.2	150	-29	$10.5^{+2.7}_{-1.3}$	-57	10.2	12.8	22	59	3b	Mon Strm
PCI-5	163.8	48	162.4	59.2	$6.9^{+3.6}_{-1.6}$	-132	11.7	4.6	22	150	4a	GD Strm
PCI-6	129.6	53.9	164.3	37.2	$9.4^{+2.0}_{-2.4}$	-13	10.9	9.8	20	93	4b	...
PCI-7	124.5	38	183.4	32.6	$8.7^{+1.1}_{-0.7}$	29	19	10.6	16	83	5a	...
PCI-8	132.6	6.1	221.5	29.2	$9.9^{+2.4}_{-1.0}$	71	13.2	11.4	15	69	5b	...
PCI-9	134	3.2	225.2	29	$9.7^{+1.5}_{-0.5}$	85	14.9	4	17	74	6a	Mon Strm

**Notes.**

<sup>a</sup> Median heliocentric distance in kpc of MPMSTO stars with radial velocities within one velocity resolution of the identified peak.

<sup>b</sup> Radial velocity in  $\text{km s}^{-1}$  at which  $\Theta(v_r)$  peaks.

<sup>c</sup> Velocity dispersion in  $\text{km s}^{-1}$  of the significant peak.

<sup>d</sup> Median radial velocity error in  $\text{km s}^{-1}$  for all MPMSTO radial velocities within  $12 \text{ km s}^{-1}$  of the peak in  $\Theta(v_r)$ .

<sup>e</sup> The number of MPMSTO star radial velocities within one velocity resolution of the identified peak.

<sup>f</sup> The number of MPMSTO star spectra obtained along that line of sight that lie in the inner halo as defined in Section 2.

analysis that to eliminate false positives the basic peak algorithm needs to be extended in the following ways.

1. The use of the smooth model as our null hypothesis is most likely to break down far from the median radial velocity along a line of sight (where the sampling is very sparse). We require that the radial velocity associated with a candidate ECHOS falls within an interval centered on the median radial velocity of the smooth inner halo component containing 95% of the synthetic stars from the mock catalog along that line of sight.
2. The finite size of our mock catalogs and the number of Monte Carlo iterations we can computationally afford are issues. In the limit of an infinitely large mock catalog and an infinite number of Monte Carlo iterations, the significance contours we derive would be perfectly smooth with no small-scale fluctuations. However, since both our mock catalog and Monte Carlo simulation are of finite size, we need to ensure that the small-scale fluctuations present in our significance thresholds do not lead to false positives. Therefore we impose one more requirement: we fit a Gaussian to the lower bound of our 1 in  $10^4$  significance region and increases its amplitude to ensure that it is an upper-bound for all of the small-scale structure in the

significance contours. We require that  $\Theta(v_r)$  must be above this upper-bound for a detection.

Our analysis indicates that the expected number of false positives with these properties is less than one over our entire search. We call detections that meet these criteria class I peak detections and the basic peak algorithm including these criteria the class I peak algorithm. From this point forward, we define high-confidence detections as all candidate ECHOS identified by either the bin algorithm or the class I peak algorithm. Our analysis indicates that we expect less than one false positive in our list of high confidence detections. Therefore, together the bin detections in Table 1 and class I peak detections in Table 2 provide the definitive list of promising candidate ECHOS for further study.

We note that the stringent requirements necessary to ensure an expected number of false positives less than one over our entire search also eliminates many candidates that are likely real ECHOS. While these strict requirements are necessary to isolate individual ECHOS at the highest confidence needed for follow-up investigation, these can result in a biased view of the entire population of ECHOS. Therefore, we also report in Table 3 the subset of peak detections that are formally significant at more than the 1 in  $10^4$  level but not so significant to ensure

**Table 3**  
Summary of Class II Peak Detections

ID	R.A.	Decl.	$l$	$b$	$d^a$	$v_r^b$	$\sigma^c$	Err <sup>d</sup>	$n_s^e$	$N_s^f$	Figure	Comment
PCII-1	20	31.7	130	-30.8	12.5 <sup>+3.9</sup> <sub>-2.2</sub>	-125	22	17.8	20	93	7a	...
PCII-2	20	31.7	130	-30.8	10.6 <sup>+4.2</sup> <sub>-1.5</sub>	-98	24.8	16.2	18	93	7a	...
PCII-3	21.3	39.6	130	-22.8	13 <sup>+4.6</sup> <sub>-1.6</sub>	-121	10.5	19.4	13	34	2	...
PCII-4	91.8	83.5	130	25.7	11.5 <sup>+1.5</sup> <sub>-1.8</sub>	-95	19.9	12	13	47	7b	...
PCII-5	17	0	132	-62.6	11.6 <sup>+4.2</sup> <sub>-5.2</sub>	-173	12.1	14.4	19	109	3a	...
PCII-6	38.2	25.5	150	-32	10.4 <sup>+3.0</sup> <sub>-1.6</sub>	-93	20.9	15.4	14	60	8a	...
PCII-7	38.2	25.5	150	-32	11 <sup>+1.7</sup> <sub>-1.9</sub>	-66	24.7	14.5	14	60	8a	...
PCII-8	39.7	28.2	150	-29	10.5 <sup>+2.7</sup> <sub>-1.3</sub>	-57	10.2	12.8	22	59	3b	Mon Strm
PCII-9	30	0	157	-58.3	11.9 <sup>+5.2</sup> <sub>-4.1</sub>	-177	19.1	15.9	18	173	8b	...
PCII-10	163.8	48	162.4	59.2	6.9 <sup>+3.6</sup> <sub>-1.6</sub>	-132	11.7	4.6	22	150	4a	GD Strm
PCII-11	129.6	53.9	164.3	37.2	9.4 <sup>+2.0</sup> <sub>-2.4</sub>	-13	10.9	9.8	20	93	4b	...
PCII-12	124.5	38	183.4	32.6	8.7 <sup>+1.1</sup> <sub>-0.7</sub>	29	19	10.6	16	83	5a	...
PCII-13	64.8	6.6	187	-29.5	10.3 <sup>+1.7</sup> <sub>-0.9</sub>	20	28.3	16.1	12	65	9a	...
PCII-14	64.8	6.6	187	-29.5	11.8 <sup>+2.3</sup> <sub>-1.5</sub>	44	31.3	17.6	14	65	9a	...
PCII-15	116.9	28	192.4	23.9	10.8 <sup>+2.0</sup> <sub>-0.5</sub>	44	9.1	12.1	10	35	9b	...
PCII-16	139.4	30.4	195.6	43.5	7.6 <sup>+6.0</sup> <sub>-1.1</sub>	-103	13	10.1	11	114	10a	...
PCII-17	127.7	24.4	199.8	32	13.3 <sup>+2.0</sup> <sub>-1.9</sub>	-40	18.1	17.7	10	83	10b	...
PCII-18	165.6	28.6	203.1	65.9	9.5 <sup>+6.3</sup> <sub>-4.0</sub>	-157	13.3	14.2	10	151	11a	...
PCII-19	139.9	22.2	206.6	41.9	13.7 <sup>+2.4</sup> <sub>-5.0</sub>	-55	27.3	16.3	11	102	11b	...
PCII-20	132.6	6.1	221.5	29.2	9.9 <sup>+2.4</sup> <sub>-1.0</sub>	71	13.2	11.4	15	69	5b	...
PCII-21	134	3.2	225.2	29	9.7 <sup>+1.5</sup> <sub>-0.5</sub>	85	14.9	4	17	74	6a	Mon Strm

**Notes.**

<sup>a</sup> Median heliocentric distance in kpc of MPMSTO stars with radial velocities within one velocity resolution of the identified peak.

<sup>b</sup> Radial velocity in km s<sup>-1</sup> at which  $\Theta(v_r)$  peaks.

<sup>c</sup> Velocity dispersion in km s<sup>-1</sup> of the significant peak.

<sup>d</sup> Median radial velocity error in km s<sup>-1</sup> for all MPMSTO radial velocities within 12 km s<sup>-1</sup> of the peak in  $\Theta(v_r)$ .

<sup>e</sup> The number of MPMSTO star radial velocities within one velocity resolution of the identified peak.

<sup>f</sup> The number of MPMSTO star spectra obtained along that line of sight that lie in the inner halo as defined in Section 2.

that the entire list is free from false positives; we call these class II peak detections. Our analysis indicates that false positives are almost always associated with candidates that have very few stars within an interval centered on the peak of  $\Theta(v_r)$  with width equal to our median velocity resolution—about 11.5 km s<sup>-1</sup>. For this reason, we can reduce the level of contamination in Table 3 to an acceptable level by varying the required number of MPMSTO stars within that radial velocity interval. We find that if we require more than 10 MPMSTO stars within that velocity interval, then the expected number of false positives in the 21 rows of Table 3 is less than 3. We call the basic peak algorithm with this additional criteria the class II peak algorithm. We believe that the population of candidate ECHOS in Table 3 is perhaps more representative of the ECHOS population as a whole, but that individual entries in that list are potentially unreliable.

### 3.4. An Example

To illustrate these algorithms, consider Figure 2. It shows the radial velocity data and our analyses for the line of sight along which we found ECHOS B-1 from Table 1, PCI-3 from Table 2, and PCII-3 from Table 3. The top panel shows radial velocity versus  $r$ -magnitude. We plot radial velocity errors as gray horizontal bars—the median error is approximately 11.5 km s<sup>-1</sup>—while the photometric errors are much smaller than the plotted points. The second panel shows in black a histogram derived from the observed SEGUE MPMSTO star radial velocity distribution with bin-width 20 km s<sup>-1</sup> and 95% confidence intervals from bootstrap resampling. We also plot in gray an area that corresponds to a 95% confidence region for the histogram that would result from a SEGUE observation of our mock catalog. As a result, a significant bin is one for

which the error bar on the black histogram does not intersect the gray region. The bin centered at  $v_r = -130$  km s<sup>-1</sup> hosts a significant excesses of MPMSTO stars relative to the smooth model. The third panel shows in black the CDF of the observed MPMSTO star radial velocity distribution and in gray the average CDF of the simulated radial velocity distribution obtained from our mock catalog. Note the large discrepancy between the slope of the CDF  $F'(v_r)$  of the MPMSTO star data and the “average” CDF of smooth component  $\overline{F}'_S(v_r)$ ; this indicates that there is a very significant overdensity in the MPMSTO data at  $v_r \approx -120$  km s<sup>-1</sup>. The fourth panel shows in black our  $\Theta(v_r)$  statistic. We plot the formal 1 in 10<sup>2</sup> significance region in light gray, the formal 1 in 10<sup>3</sup> significance region in medium gray, and the formal 1 in 10<sup>4</sup> region in the dark gray. The white region is significant at less than the 1 in 10<sup>2</sup> level; note that the medium and dark gray regions are nearly coincident, emphasizing the extreme non-Gaussianity of our  $\Theta(v_r)$  statistic. These significance thresholds are naively formally equivalent to 2.33 $\sigma$ , 3.09 $\sigma$ , and 3.72 $\sigma$ , though the distribution from which we derive the significance thresholds is only defined for positive real numbers and highly non-Gaussian, so this comparison is not robust. The heavy black curve is the Gaussian upper-bound that we use to ensure that the small-scale fluctuations that result from the finite size of our mock catalog and the finite number of Monte Carlo iterations we can computationally afford do not produce false positives. Our false positive analysis indicates that the probability that a smooth halo can produce peaks like that peak observed at  $v_r = -121$  km s<sup>-1</sup> is vanishingly small. We present similar plots for all of our detections in order of galactic longitude in Figures 2 through 11. The interested reader should use Tables 1–3 (also ordered by galactic longitude) as a guide to the plots in Figures 2 through 11. Note that because we



sample from the mock catalog the same number of MPMSTO stars that were observed by SEGUE along a given line of sight, any overdensity ensures a subsequent underdensity somewhere else. Those underdensities are not meaningful. Also be aware that the apparent  $r$ -magnitude inhomogeneities present along some lines of sight do not necessarily imply position-space substructure. The reason is that the transformation of apparent  $r$ -magnitude into absolute  $r$ -magnitude is difficult for individual turnoff stars because isochrones are nearly vertical in a color-magnitude diagram at the turnoff. As a result, small errors in observed color combined with uncertain metallicity can lead to large systematic uncertainties in photometric parallaxes for individual MPMSTO stars.

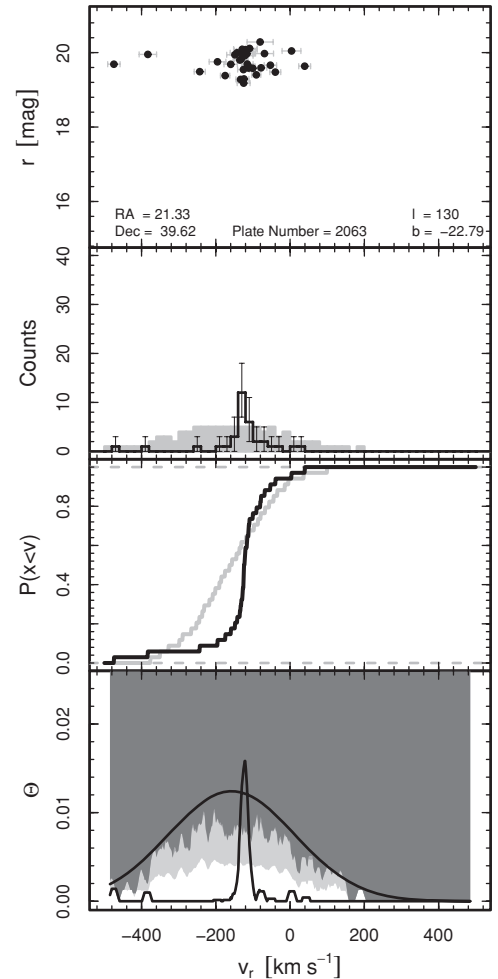
### 3.5. A Completeness Estimate

We would like to quantify the fraction of cold substructure everywhere in the inner halo of the Milky Way Galaxy, not just along those lines of sight for which we have candidate ECHOS. To that end, we need to know the properties of our detections as well as the properties of the substructure that our algorithms are not capable of separating from the smooth inner halo. Therefore, we calculate our completeness with a Monte Carlo simulation similar to that described in Section 3.3. This time, however, we replace a certain fraction of the synthetic stars in  $S_r$  with a synthetic ECHOS with known number of stars, physical extent, radial velocity, and velocity dispersion. We determine, on average, how often our algorithms find the inserted ECHOS as a function of the fraction of the total number of stars along the line of sight that are a part of the ECHOS.

#### 3.5.1. Strategy

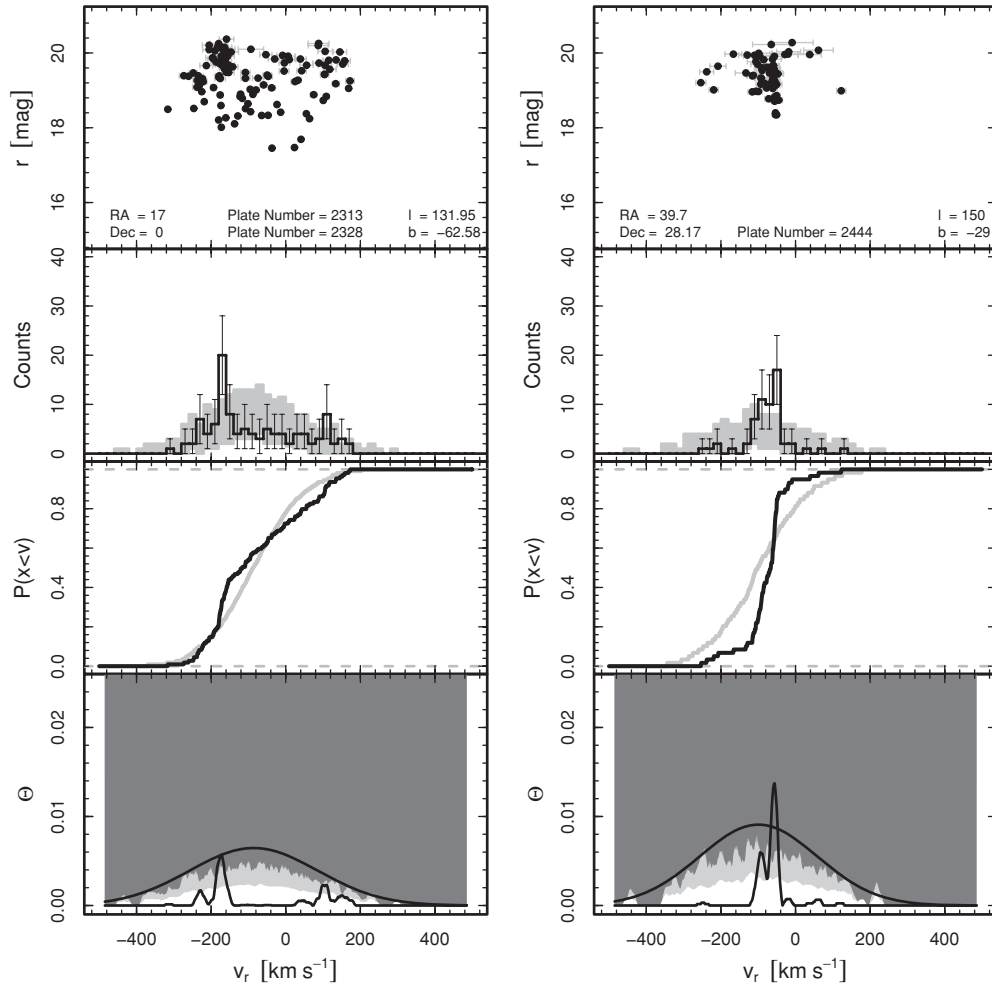
For each SEGUE line of sight we perform 1,000 simulations in each of 10 steps in the fraction of the stars along that line of sight in an ECHOS, from 10% to 100%. We use the fraction of stars in substructure instead of an absolute number because our detection probabilities also depend on the number of spectra obtained along a line of sight. For each simulation, we select the synthetic ECHOS's mean radial velocity at random and with uniform probability in the range  $v_\mu \pm 200 \text{ km s}^{-1}$  where  $v_\mu$  is the observed mean radial velocity for the smooth inner halo at that galactic longitude and latitude. We give the inserted ECHOS a three-dimensional Cartesian velocity dispersion chosen to match (at least to order of magnitude) the velocity dispersion of possible progenitors and described by a diagonal matrix in which we select the diagonal entry at random and with uniform probability in the interval  $[0, 5/\sqrt{3}] \text{ km s}^{-1}$ . We then add noise to the individual radial velocities of the stars that make up the inserted substructure at a level characteristic of our median estimated radial velocity errors at the S/N of the individual stars in our detections, about  $11.5 \text{ km s}^{-1}$ . We then use our algorithms to determine whether they detect the synthetic ECHOS.

Considering the significance contours in the bottom panel of Figure 2, our sensitivity to a given fractional overdensity depends on its mean velocity. That is, an ECHOS with mean velocity far from the mean velocity of the smooth inner halo is easier to find than a similar ECHOS with mean velocity close to the mean velocity of the smooth component. We attempt to marginalize over this effect by using a uniform distribution in the synthetic ECHOS's mean radial velocity, as described in the preceding paragraph. The uniform distribution is our attempt to minimize the amount of "prior information" we include in the completeness calculation. This should be a reasonable



**Figure 2.** Data and our analyses for the line of sight along which we found the element of cold substructure B-1 from Table 1, PCI-3 from Table 2, and PCII-3 from Table 3. In the top panel, we plot radial velocity vs.  $r$ -magnitude. In the same panel, we also plot our estimated radial velocity errors as the gray horizontal bars (the photometric errors are much smaller than the plotted points). In the second panel, we plot in black a histogram derived from the observed MPMSTO star radial velocity distribution with bin-width  $20 \text{ km s}^{-1}$  and 95% confidence intervals from bootstrap resampling. In the same panel, we also plot in gray an area that corresponds to a 95% confidence region for the histogram that would result from a SEGUE observation of our mock catalog. As a result, a significant bin is one for which the error bar on the black histogram does not intersect the gray region. In the third panel, we plot in black the CDF of the observed MPMSTO star radial velocity distribution, and we plot in gray the average CDF of the simulated radial velocity distribution obtained from our mock catalog. In the fourth panel, we plot in black our  $\Theta(v_r)$  statistic, while we plot the 1 in  $10^2$  significance region in light gray, the 1 in  $10^3$  significance region in medium gray, and the 1 in  $10^4$  region in the dark gray. The white region is significant at less than the 1 in  $10^2$  level; note that the medium and dark gray regions are nearly coincident, emphasizing the extreme non-Gaussianity of our  $\Theta(v_r)$  statistic. The black Gaussian curve is an upper-bound that we use to ensure the small-scale fluctuations in our significance contours do not lead to false positives. Note the significant feature in all panels at  $v_r \approx -121 \text{ km s}^{-1}$ .

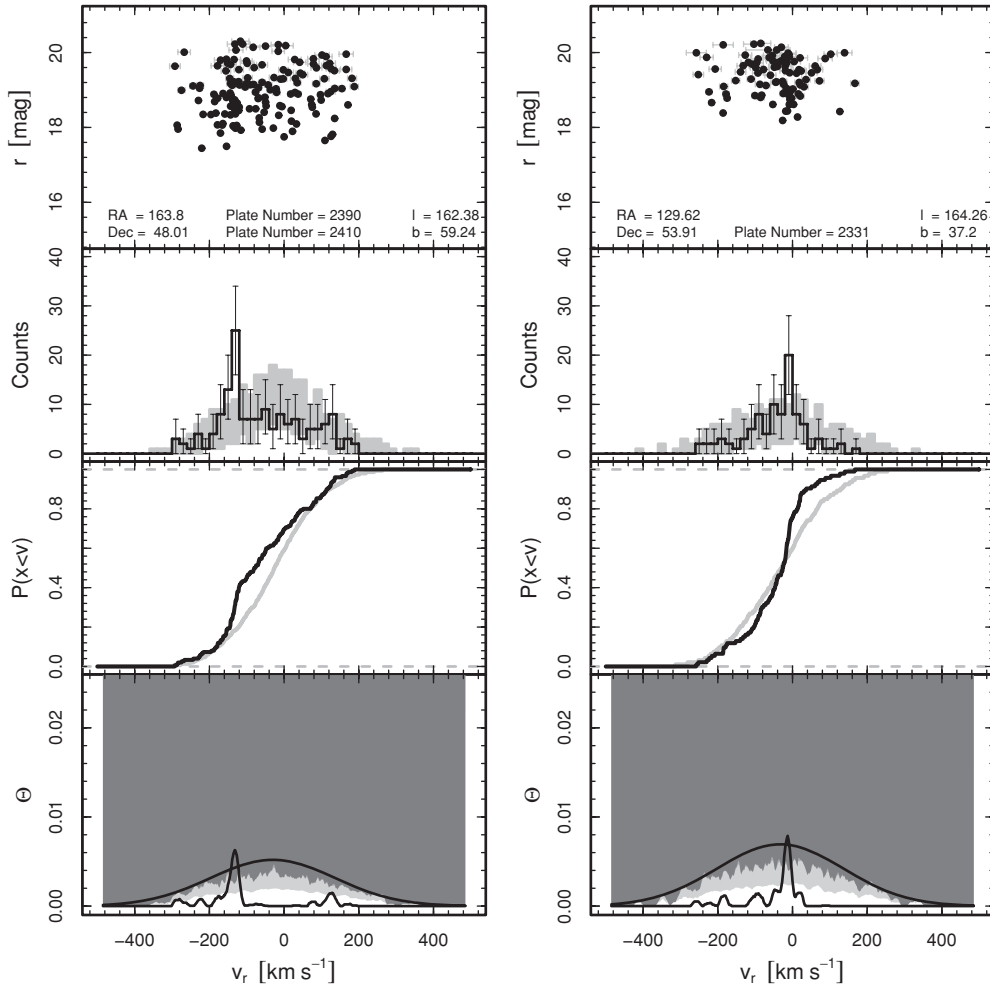
assumption if the mean velocities of the real population of ECHOS are equally likely to lie anywhere in the radial velocity range of the smooth inner halo. To test this assumption, we compare the difference between the observed mean velocities of our candidate ECHOS with the median radial velocity of the smooth component along that line of sight with velocities selected at random and with uniform probability in the range  $v_\mu \pm 200 \text{ km s}^{-1}$ . We perform a KS test on the two samples, and repeat the process 1,000 times drawing a new random set each time. We find that the median  $p$ -value from the KS



**Figure 3.** Top panel: we plot radial velocity vs.  $r$ -magnitude. In the same panel, we also plot our estimated radial velocity errors as the gray horizontal bars (the photometric errors are much smaller than the plotted points). In the second panel, we plot in black a histogram derived from the observed MPMSTO star radial velocity distribution with bin-width  $20 \text{ km s}^{-1}$  and 95% confidence intervals from bootstrap resampling. In the same panel, we also plot in gray an area that corresponds to a 95% confidence region for the histogram that would result from a SEGUE observation of our mock catalog. As a result, a significant bin is one for which the error bar on the black histogram does not intersect the gray region. In the third panel, we plot in black the CDF of the observed MPMSTO star radial velocity distribution, and we plot in gray the average CDF of the simulated radial velocity distribution obtained from our mock catalog. In the fourth panel, we plot in black our  $\Theta(v_r)$  statistic, while we plot the 1 in  $10^2$  significance region in light gray, the 1 in  $10^3$  significance region in medium gray, and the 1 in  $10^4$  region in the dark gray. The white region is significant at less than the 1 in  $10^2$  level; note that the medium and dark gray regions are nearly coincident, emphasizing the extreme non-Gaussianity of our  $\Theta(v_r)$  statistic. The black Gaussian curve is an upper-bound that we use to ensure that the small-scale fluctuations in our significance contours do not lead to false positives. Left: data and analyses for the line of sight along which we found the element of cold substructure B-3 from Table 1 and PCI-4 from Table 2. Right: data and analyses for the line of sight along which we found the element of cold substructure B-2 from Table 1, PCI-5 from Table 2, and PCII-8 from Table 3. This line of sight is also expected to intersect the Monoceros stream.

tests is 0.90; therefore our assumption that ECHOS are equally likely to be at any radial velocity in an interval about the mean radial velocity of the smooth model is a good one. This is not surprising, as the radial velocities of both ECHOS and of the smooth background is primarily determined by the relationship between the Sun’s velocity vector in its orbit about the Galaxy and the vector describing the direction along a given line of sight. The velocity dispersion we assign to each inserted ECHOS is an order of magnitude estimate of the velocity dispersion of a real element of cold substructure. Regardless, the dominant source of dispersion is random measurement error in the radial velocities—which will be larger than the velocity dispersion of any ECHOS—and we have a solid understanding of that distribution. In any case, we tested our algorithms’ abilities to detect ECHOS with velocity dispersions approaching  $50 \text{ km s}^{-1}$ —half that of the halo itself—and found no significant change in performance.

For each SEGUE line of sight we find that, averaged over the given distributions for radial velocity and velocity dispersion, the relationship between probability of detection and fraction of stars along that line of sight that are part of an ECHOS is nonlinear. As a result, we interpolate the results for each line of sight and invert the function to determine the fraction of the total number of stars observed along that line of sight that must be part of an ECHOS to ensure that our algorithms can detect it 95% of the time. We illustrate these calculations for the class II peak algorithm in Figures 12 and 13; the results for the bin algorithm and the class I peak algorithm are similar. Note that for all of the algorithms the probability of detection in the  $N = 99$  lines of sight with more than 50 stars quickly approaches unity as the fraction in substructure increases. On the other hand, for the  $N = 38$  lines of sight with fewer than 50 stars the probability of detection increases much more slowly. Lines of sight with less than about 30 stars correspond to those lines in the left



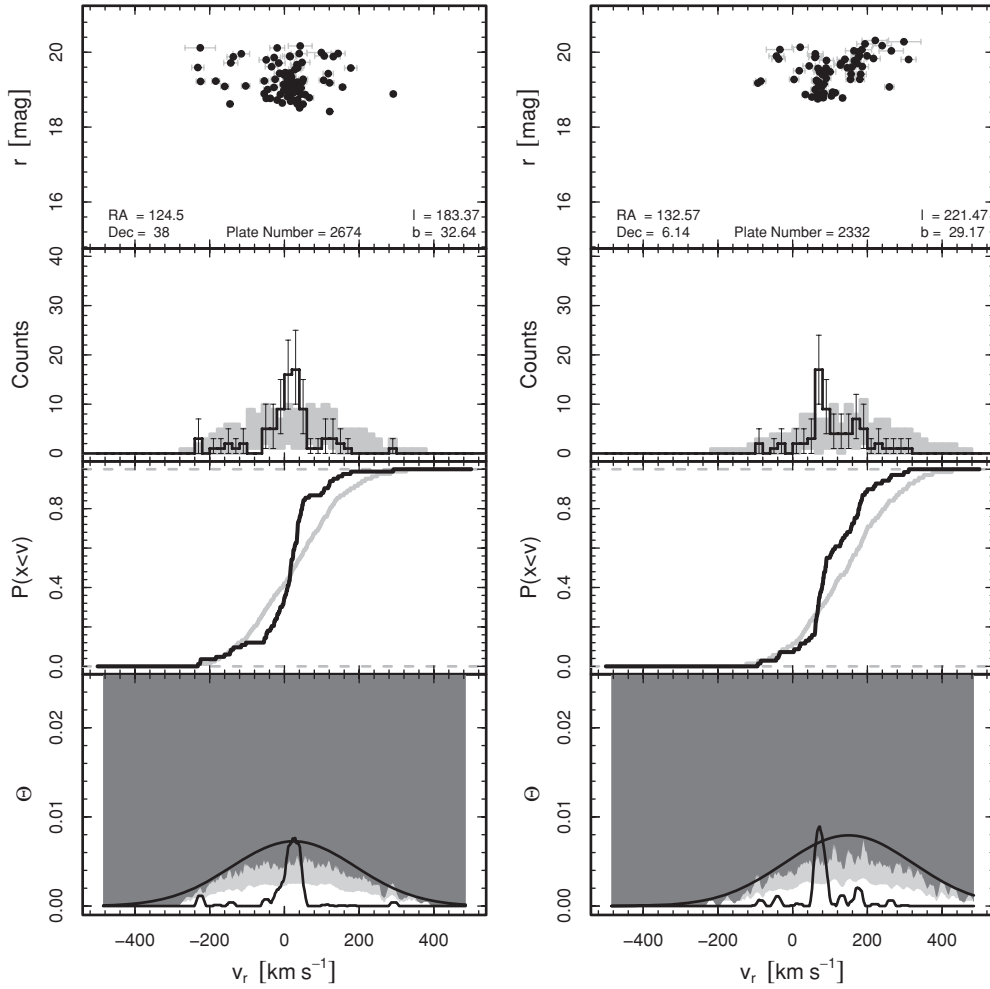
**Figure 4.** Left: data and analyses for the line of sight along which we found the element of cold substructure B-4 from Table 1, PCI-5 from Table 2, and PCII-10 from Table 3. This line of sight is also expected to intersect the Grillmair & Dionatos (2006) stream. Right: data and analyses for the line of sight along which we found the element of cold substructure B-5 from Table 1, PCI-6 from Table 2, and PCII-11 from Table 3. See the caption to Figure 3 for a detailed description of this type of figure.

hand panel of Figure 12 that have small average probability of detection, even with a large fraction in substructure. More precisely, for all lines of sight with fewer than 30 stars the average probability of detection never reaches unity even if every star observed along that line of sight is in the same ECHOS. Also note that the performance of the class II peak algorithm suffers along lines of sight with few MPMSTO star spectra because of the requirement that each class II peak detection must have 10 MPMSTO radial velocities within one velocity resolution element of the peak in our  $\Theta(v_r)$  statistic.

Table 4 lists our upper limits for each line of sight. These limits are the fraction of spectroscopically observed MPMSTO stars that could belong to single ECHOS yet still go undetected, on average, 5% of the time. The data in Table 4 indicates that our algorithms are unable to detect ECHOS at least 95% of the time along lines of sight with fewer than about 30 spectra, even if every star along that line of sight belongs to an ECHOS. As a result, we remove these sparsely sampled lines of sight from our analyses from this point on.

The fact that our completeness calculation only accounts for single ECHOS again emphasizes the point that our search for single overdensities in radial velocity space is not necessarily sensitive to an ensemble of substructures below our sensitivity thresholds. We are also potentially insensitive to diffuse, fully

phase-mixed, or violently-relaxed substructure and not necessarily sensitive to diffuse multiply wrapped substructure. We have shown, however, that our completeness is not a function of velocity dispersion and that we can detect substructure with velocity dispersion up to  $50 \text{ km s}^{-1}$  as well as we can detect cold substructure. In the end, an ECHOS is detectable if it contributes the threshold fractional overdensity determined by SEGUE sampling along that line of sight. Physically, that means that we can detect ECHOS that have not spread very far over their orbits. The fact that detectability is not sensitive to velocity dispersion implies that we can find ECHOS with a wide variety of possible orbital configurations, so long as there is a high enough number density along a single line of sight. As we will show in the discussion, in the volume we search we are likely to find older substructures than those discovered in photometric searches, but younger substructures than those discovered by six-dimensional (6D) searches in the same volume. We recognize that the time it takes for the debris of a given accretion event to spread beyond our detectability threshold depends on progenitor mass, velocity dispersion, and orbit. Still, for any given set of progenitor properties our search can delineate its debris for a longer time after progenitor disruption than photometric searches. In any case, we have very precisely calculated our sensitivity thresholds and developed a search strategy that can be meaningfully applied to



**Figure 5.** Left: data and analyses for the line of sight along which we found the element of cold substructure B-6 from Table 1, PCI-7 from Table 2, and PCII-12 from Table 3. Right: data and analyses for the line of sight along which we found the element of cold substructure B-7 from Table 1, PCI-8 from Table 2, and PCII-20 from Table 3. See the caption to Figure 3 for a detailed description of this type of figure.

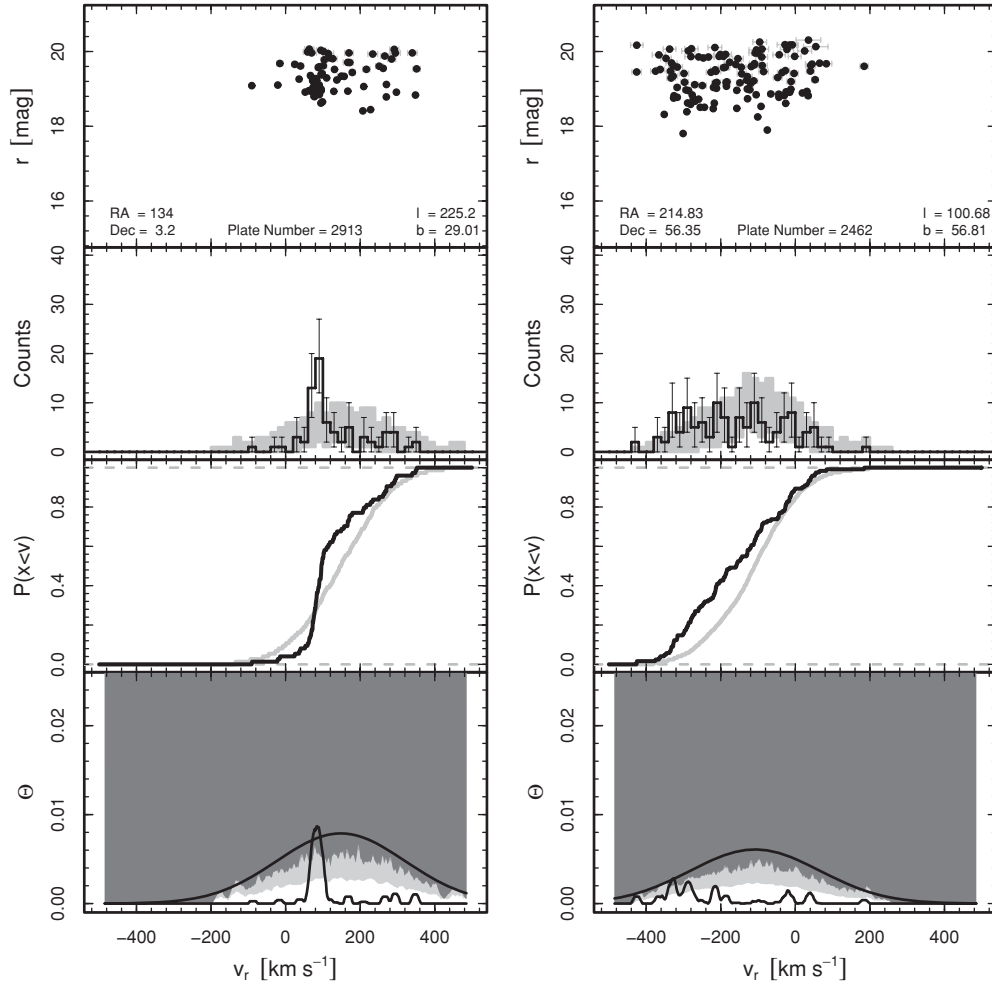
theoretical models to assess their level of agreement with our observations.

### 3.6. The Relative Merits of Each Algorithm

The bin algorithm and the class I peak algorithm are both very unlikely to produce false positive detections at the cost of a failure to detect less obvious genuine ECHOS. On the other hand, the class II peak algorithm is better at detecting diffuse or more phase-mixed ECHOS at the cost of an occasional false positive. In general the class II peak algorithm is most capable of detecting low-density ECHOS, followed by the class I peak algorithm and the bin algorithm; Table 4 quantifies this ranking. Note though that the class II peak algorithm is less capable than the other algorithms along lines of sight with relatively few stars, because of the requirement that at least 10 MPMSTO stars must be within an interval of width  $11.5 \text{ km s}^{-1}$  centered on the velocity of the detected peak. That is, the class II peak algorithm cannot detect small fractional densities along lines of sight with  $\sim 10$  spectra because a small fraction of  $\sim 10$  spectra will always be less than the required 10 MPMSTO. Recall that the limit of 10 MPMSTO stars within  $11.5 \text{ km s}^{-1}$  of the peak in  $\Theta(v_r)$  was set to minimize false positives. The peak algorithms can also self-consistently determine the velocity dispersion of an ECHOS (a characteristic lacking in the bin algorithm) and

they do not require the discretization into bins that causes the bin algorithm to sometimes split the signal of a genuine ECHOS into two neighboring bins.

Figure 14 confirms the results of our completeness calculation: the peak algorithms can discover lower fractional overdensity ECHOS along lines of sight with many stars than the bin algorithm. Figures 15 and 16 demonstrate that our detections are not all clustered at the edge of detectability. Specifically, Figure 15 demonstrates that for the peak algorithms there is no correlation between the number of stars associated with ECHOS and the total number of spectra obtained along the line of sight where the ECHOS was discovered. Recall the results of our completeness calculation (presented in Figures 12 and 13) that the detectable fractional overdensity is correlated with the number of stars per line of sight. In other words, we are most sensitive to low fractional overdensity substructure along the most well-sampled lines of sight. If there were a substantial population just below our detection thresholds, then there would be many detections at small absolute number of stars associated with elements of substructure at large total numbers of spectra in Figure 14. This is not observed, so these observations imply that the ECHOS we find are not all close to the lowest fractional overdensity we can detect along any line of sight. On the other hand, Figure 15 indicates that our detections are not just found along the lines of sight where SEGUE most densely sampled the



**Figure 6.** Left: data and analyses for the line of sight along which we found the element of cold substructure B-8 from Table 1, PCI-9 from Table 2, and PCII-21 from Table 3. This line of sight is also expected to intersect the Monoceros stream. Right: data and analyses for the line of sight along which we found the element of cold substructure PCI-1 from Table 2. See the caption to Figure 3 for a detailed description of this type of figure.

inner halo MPMSTO population. Together, these facts suggest that at our velocity resolution the inner halo is not composed of a multitude of diffuse substructures beyond our sensitivity threshold.

#### 4. EXTENSION TO THE FULL INNER HALO

We would like to generalize our search for ECHOS along 115 SEGUE lines of sight to the full inner halo. If each line of sight were targeted randomly, then the average of our upper limits on the fraction of substructure along each line of sight would be an unbiased estimator for the fraction of substructure in the entire inner halo. To ensure that our final estimate of the fraction of the inner halo in ECHOS is not biased by the lines of sight that were targeted at previously known elements of substructure, we exclude those lines of sight from the analyses in sections Sections 4.1 and 4.2. When the exceptions listed in Table 4 are excluded, the R.A. and decl. of each SEGUE line of sight center was selected without regard to the presence of substructure in the Milky Way’s inner halo. If we make the following well-justified assumptions (see Section 2):

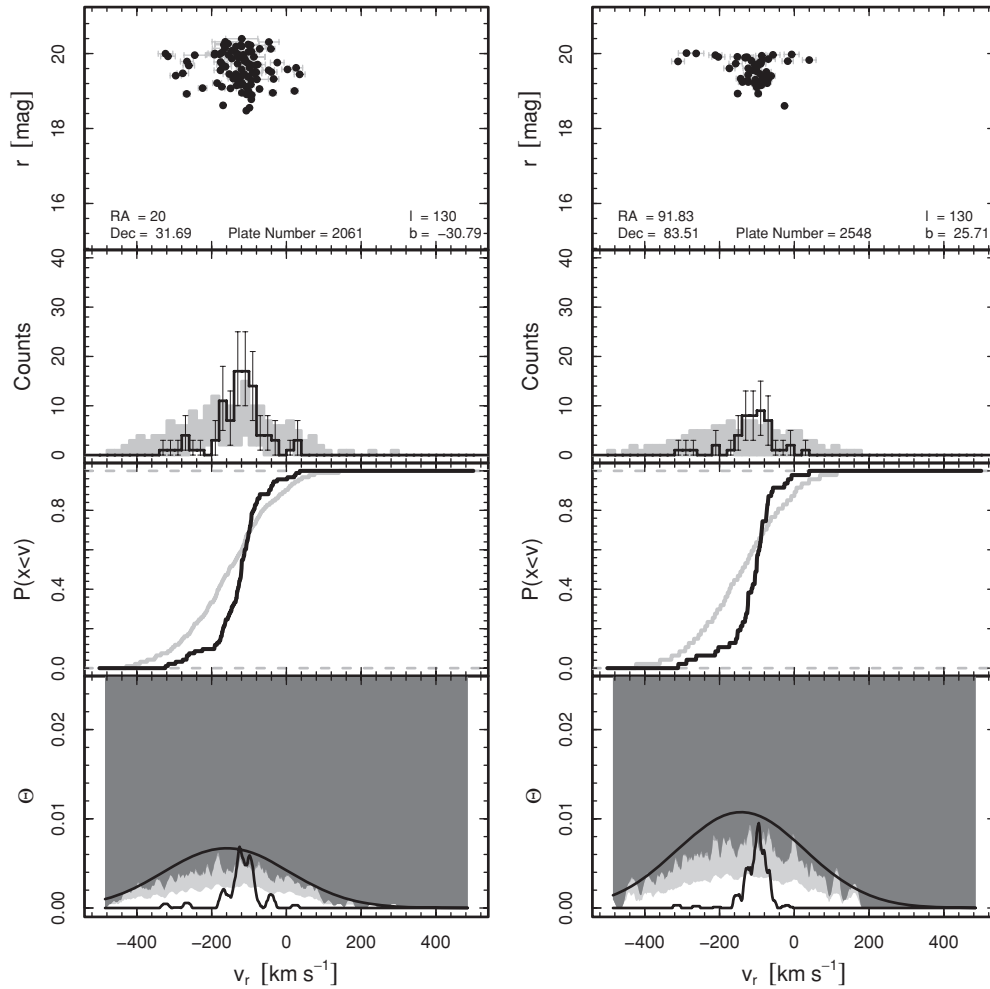
1. SDSS photometrically detected all MPMSTO stars in our sensitivity range
2. The color–color cut used to select MPMSTO candidates accurately classified them as MPMSTO stars, such that the

radial velocities belong to stars within our definition of the inner halo given in Section 2

and we recall that fibers were allocated randomly to the photometric MPMSTO candidates without prior knowledge of their radial velocities, then the SEGUE MPMSTO sample is an unbiased tracer of the inner halo. Therefore we can assert that, on average, the fraction of spectroscopically observed MPMSTO stars that reside in ECHOS along a given SEGUE line of sight is the same as the fraction of all MPMSTO stars in ECHOS in the volume searched along that line of sight. Again, by ECHOS we mean single elements of substructure that are unresolved (or barely resolved) in radial velocity in the SEGUE data.

##### 4.1. A Limiting Case

Imagine the worst-case scenario: our algorithms miss all ECHOS in the inner halo just below the 95% detection limits given in Table 4. That is, they fail to identify literally all ECHOS that they would normally detect 94.9% (or less) of the time. The probability of this occurrence is vanishingly small— $P(X) \sim 0.05^{115}$ —and the probability of missing two or more such ECHOS in each line of sight is smaller still; nevertheless, it is a useful limiting case. In that situation, any line of sight would harbor an ECHOS just below the line of sight specific 95% thresholds given in Table 4. If we imagine that every line

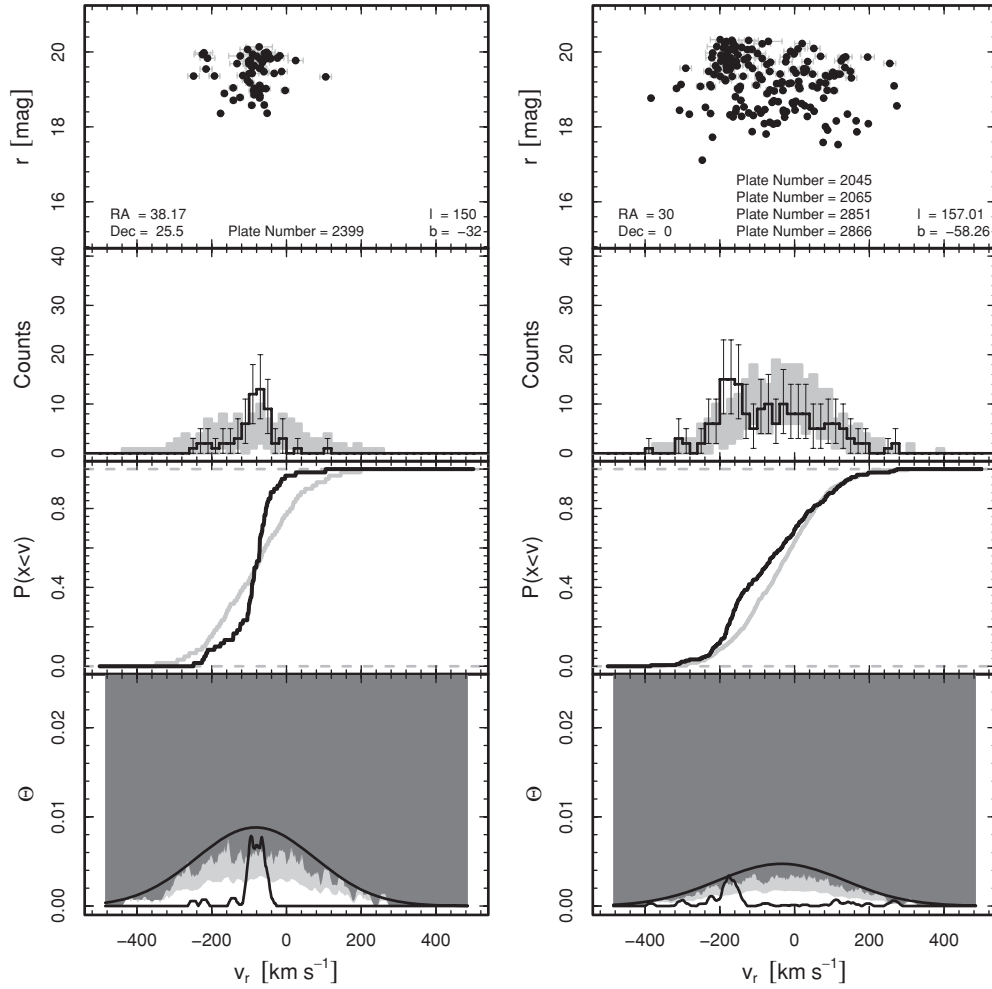


**Figure 7.** Left: data and analyses for the line of sight along which we found the element of cold substructure PCI-2 from Table 2 and PCII-1 and PCII-2 from Table 3. Right: data and analyses for the line of sight along which we found the element of cold substructure PCII-4 from Table 3. See the caption to Figure 3 for a detailed description of this type of figure.

of sight in fact intersects an ECHOS just below the threshold, then the true fraction of the MPMSTO population in ECHOS along a given line of sight would be identical to the threshold given for that line of sight in Table 4. We have 105 lines of sight that were not pointed at known elements of substructure and that possess the more than the 30 spectra necessary for a potential detection. Those 105 lines of sight provide us with 105 independent estimates of the fraction of the MPMSTO population in ECHOS. We combine these independent estimates in an average value with an appropriate weighting scheme. We can compute the volume of the inner halo (as defined in Section 2) scanned along every line of sight using Monte Carlo integration; we report that volume in Table 4. Therefore, we weight the contribution to the upper limit of each line of sight by its volume to give lines of sight that scanned a larger volume of the inner halo greater leverage in determining our average limit. Finally, for lines of sight along which we have a high-confidence detection we substitute the observed fraction in substructure from Tables 1 or 2 for the estimated upper limit given in Table 4. We use bootstrap resampling to estimate the errors on our upper limits. Ultimately, we find that the bin algorithm produces a 95% upper limit of  $0.52^{+0.04}_{-0.03}$  for the fraction of MPMSTO stars in the inner that belong to ECHOS. This slightly higher than the upper limit of  $0.42^{+0.01}_{-0.02}$  produced by the peak algorithm. We regard our limit from the

peak algorithm as more accurate because the bin algorithm is the least sensitive of our methods. If we restrict our calculation to only those 41 lines of sight with more than 100 spectra, the limits quoted above become  $0.38^{+0.02}_{-0.02}$  for the bins and  $0.34^{+0.02}_{-0.02}$  for the peaks. The discrepancy between these two is due to the fact that our volume weighting scheme does not penalize lines of sight with poor sensitivity enough in the average. That is, we do not get improved upper limits from adding more lines of sight because our sensitivity to ECHOS is a function of the number of MPMSTO spectra obtained by SEGUE along a given line of sight. As a result, the upper limit is weaker if lines of sight with poor sampling are included in the calculation. In any case, we present a much more precise calculation of the fraction of the halo in substructure in Section 4.2. Again note that our algorithms are not necessarily sensitive to the possibility that the inner halo is made up of innumerable diffuse, fully phase-mixed, or violently-relaxed elements of substructure below our sensitivity thresholds.

Equivalently, we can estimate the total number of ECHOS in the inner halo by computing the ratio of the volume scanned by SEGUE to the total volume of the inner halo as defined in Section 2. Using the same assumptions given above, that ratio should be the same as the ratio between the total number of ECHOS we identify (less the expected number of false positives) and the total number of ECHOS in the entire inner halo. We find



**Figure 8.** Left: data and analyses for the line of sight along which we found the element of cold substructure PCII-6 and PCII-7 from Table 3. Right: data and analyses for the line of sight along which we found the element of cold substructure PCII-9 from Table 3. See the caption to Figure 3 for a detailed description of this type of figure.

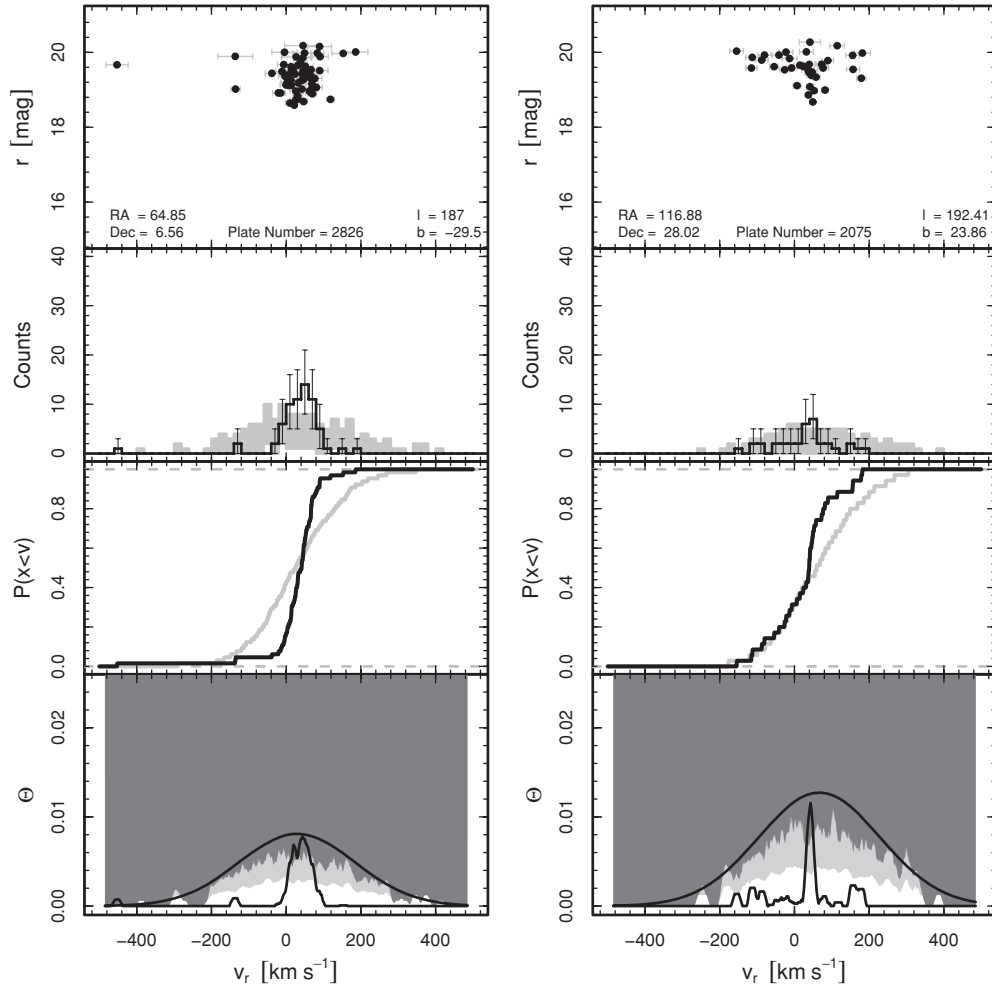
that we searched about 0.54% of the volume of the inner halo as defined in Section 2 and found seven new ECHOS. This suggests that there should be of order  $10^3$  ECHOS like those we have identified in the entire volume of the inner halo.

#### 4.2. A More General Calculation

We can also compute the total number of ECHOS we expect to miss or discover over our entire search as a function of the average fraction of the halo in ECHOS. In our previous calculation, we assumed that there was an ECHOS just below our 95% detection threshold along every line of sight and we showed that this event was extremely unlikely. That is not to say that we don't miss anything; in fact, even for ECHOS beyond our 95% thresholds it is almost certain that we miss at least one ECHOS over our entire search, as the probability of that event is  $P(X) \sim 1 - 0.95^{115}$ . To address this point, we used a Monte Carlo simulation to predict the expected number of ECHOS we both miss and discover as a function of the average fraction of the inner halo in substructure. For every line of sight and for each of the 10 steps in the fraction of the field that is a part of an ECHOS we have already computed the average detection probability for an ECHOS at that fractional overdensity. Therefore, we can model the detection process for every line of sight and for every one of those 10 fractional overdensities by drawing a random number with uniform probability from the interval  $[0, 1]$ . If

that number is greater than the average detection probability, it counts as a missed substructure; if it is less, it counts as a discovered substructure. We can repeat this process for every line of sight in our search 100 times such that we compute both the expected number of missed and discovered ECHOS over our entire search, as well as the distribution of both quantities. We plot the result of these calculations for the class II peak algorithms in Figure 17; the results for the bin algorithm and the class I peak algorithm are similar.

We can use our list of class II detections combined with the results of this more general completeness calculation to compute a prediction for the fraction of the halo in ECHOS of a given overdensity (we use our list of class II detections because that list is likely more representative of the whole inner halo ECHOS population than either of the other two lists). That is, Table 3 tells us how many ECHOS we found at each fractional overdensity and we know from Figure 17 how many ECHOS our search would have yielded if the entire halo was at a particular fraction in substructure. The ratio between the observed number of ECHOS at a given fraction in substructure and the number predicted by the completeness calculation at the fraction is an estimate of the fraction of the halo that has substructure at that level. We find that we expect about 1/3 of the halo (by volume) to have 10% of its MPMSTO population in ECHOS and about 1/6 of the halo (by volume) to have 20% of its MPMSTO



**Figure 9.** Left: data and analyses for the line of sight along which we found the element of cold substructure PCII-13 and PCII-14 from Table 3. Right: data and analyses for the line of sight along which we found the element of cold substructure PCII-15 from Table 3. See the caption to Figure 3 for a detailed description of this type of figure.

population in ECHOS; the fraction of the halo (by volume) with more than 20% of its MPMSTO population in ECHOS is very small. We plot the results of this calculation in Figure 18. We include in Figure 18 the location of ECHOS with properties like ultrafaint dwarf galaxies, known tidal streams like Monoceros and the Grillmair & Dionatos (2006) stream, and classical dwarf spheroidal galaxies and globular clusters. There are unlikely to be ECHOS like the classical dwarf spheroidal galaxies or globular clusters in the inner halo, and only a few percent of the halo volume hosts ECHOS like the Monoceros stream or the Grillmair & Dionatos (2006) stream. On the other hand, our search does not rule out the possibility of ECHOS comparable to ultrafaint dwarf galaxies.

## 5. DISCUSSION

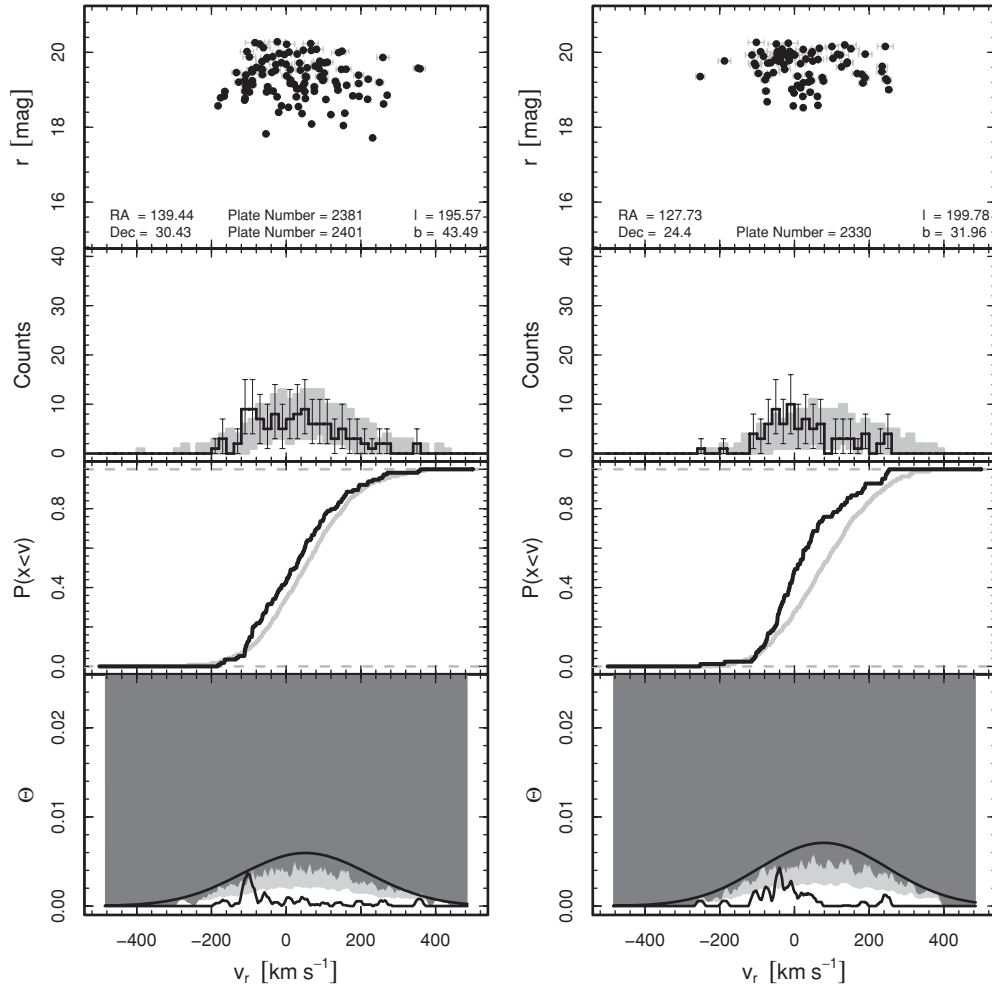
We plot our ECHOS on the sky in galactic coordinates in Figure 19, and we indicate lines of sight targeted at known elements of substructure. The distribution of ECHOS on the sky is consistent with an isotropic distribution given our completeness, simply because more stars fall within our definition of the inner halo toward the Galactic anticenter; for this reason, we are more sensitive to lower fractional density substructures in that direction. More quantitatively, imagine that the fraction of the halo in cold substructure is uniform in galactic longitude and

latitude. Under this assumption, we can use the line-of-sight specific sensitivity thresholds given in Table 4 to compute the expected  $l$  and  $b$  distribution under the assumption of isotropy. The distribution that results from that analysis is statistically indistinguishable from the distribution of our ECHOS.

We give  $r$ -magnitude histograms and approximate heliocentric distance distributions for all of our detections in Figures 20 and 21. We plot in Figure 22 a multiplot for the relevant physical properties of our class II detections given in Tables 3 and 4. We plot our class II detections because they provide the largest self-consistent sample, and we believe the population of class II detections is more representative of the inner halo ECHOS population collectively. We find no obvious non-trivial correlations. Finally, we note that the stars belonging to all of our ECHOS are spread uniformly over the solid angle sampled along the line of sight where they were discovered. In other words, our ECHOS appear to have sheet-like (as opposed to stream-like) morphologies. This observation is consistent with the prediction in Johnston et al. (2008) that substructures within 20 kpc of the Galactic center are more likely to have sheet-like “cloudy” morphologies than stream-like “great circle” morphologies.

Note in Figure 22 that we expect the 10% fractional overdensity ECHOS to both have MPMSTO star number densities  $n \approx 15 \text{ kpc}^{-3}$  and velocity dispersions no larger than the floor set by our radial velocity errors. We previously showed in





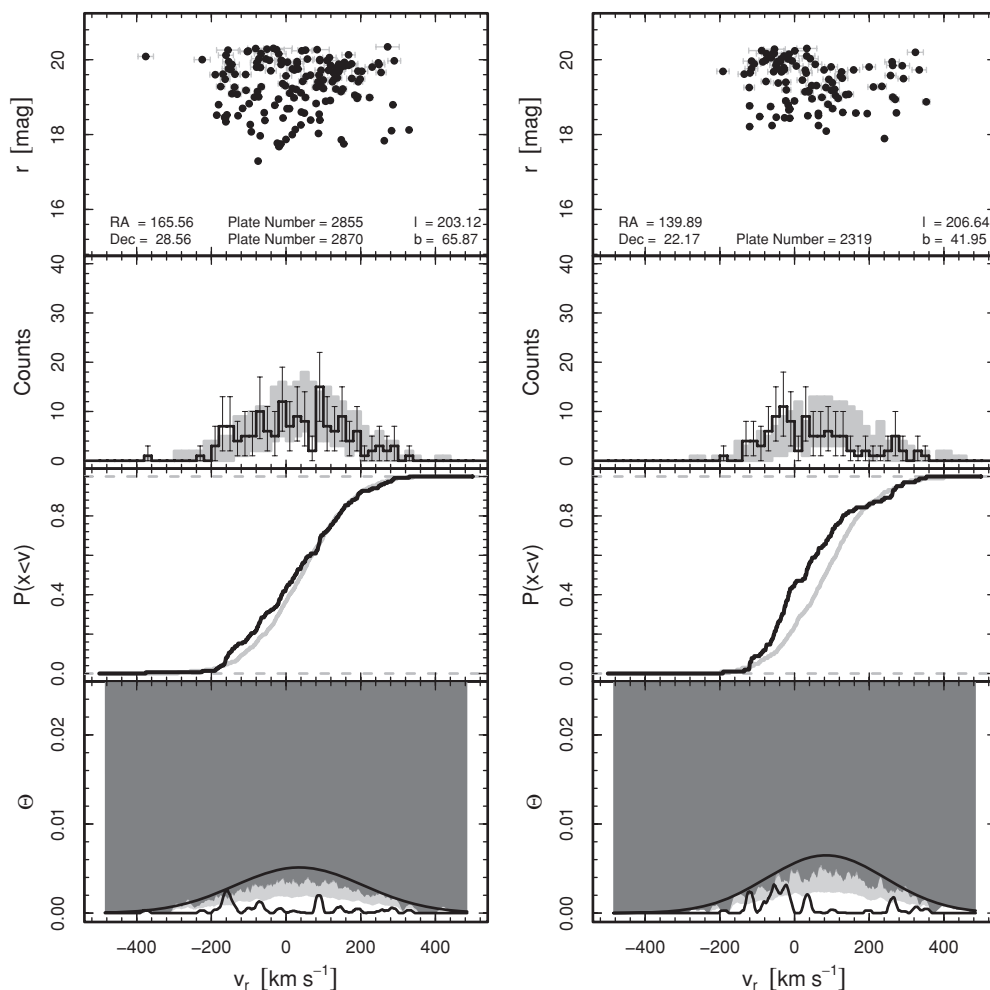
**Figure 10.** Left: data and analyses for the line of sight along which we found the element of cold substructure PCII-16 from Table 3. Right: data and analyses for the line of sight along which we found the element of cold substructure PCII-17 from Table 3. See the caption to Figure 3 for a detailed description of this type of figure.

Section 4.2 that these 10% overdensity ECHOS are likely to be found in 1/3 of the halo volume. That is, we expect a significant but not unlimited population of low density ECHOS in the inner halo of the Milky Way. For comparison, Seabroke et al. (2008) used RAVE data to show that there are no vertical streams in the solar neighborhood with total (not just MPMSTO) stellar number densities  $n \gtrsim 10^3 \text{ kpc}^{-3}$ . By itself, the fraction of the Milky Way’s halo in substructure provides an independent, qualitative measure of the intensity of the Milky Way’s stellar accretion history over the past few Gyr: more substructure indicates a more intense stellar accretion history while less substructure indicates a less intense stellar accretion history. We make a more quantitative estimate in the following subsection.

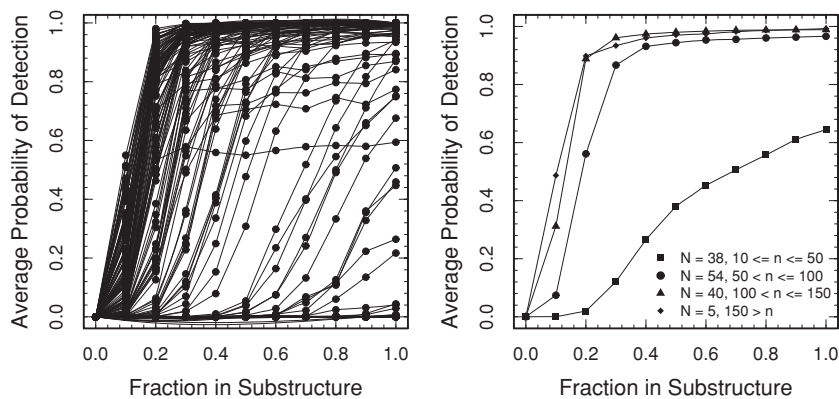
### 5.1. Comparison To Previous Studies

We identify seven new substructures as well as rediscover all known substructures in our search volume, so in that way our search is more sensitive than past studies. Moreover, our search for cold radial velocity substructure in the inner halo bridges a gap in galactocentric distance between solar neighborhood searches using 6D phase-space information and more distant in situ halo searches using surface brightness. At the same galactocentric distance, substructures discovered in surface brightness are likely younger than substructures discovered in radial velocity. Those radial velocity substructures

are themselves likely younger than substructures discovered using 6D phase-space information. For that reason, we argue that we are sensitive to older substructure than those identified by Bell et al. (2008). Note that Bell et al. (2008) statistically quantified the degree of surface brightness substructure using the same tracer population in an  $r$ -magnitude range (and therefore distance range) that partially overlaps with our search. As a result, our searches are complementary in relative look-back time—Bell et al. (2008) quantified the level of debris left behind by more recent accretion events and we have quantified the level of debris left behind by less recent accretion events. The fact that both studies find that about 30% of the MPMSTO population in the inner halo is in substructure suggests that the level of stellar accretion into the Milky Way has been relatively constant over the past few Gyr. More quantitatively, the typical velocity of stars in the inner halo of the Milky Way at  $z = 0$  is  $v \sim 200 \text{ km s}^{-1}$  and the characteristic distance from the center of the Galaxy is 10 kpc. Therefore, a crossing time is  $t_c \sim 50 \text{ Myr}$  and radial velocity substructure will likely dissolve below our limit of detectability in  $t_d \sim 100 t_c \sim 5 \text{ Gyr}$ . The nascent Milky Way was less massive than today, so the crossing time was likely longer in the past and this estimate is probably a lower limit on the look-back time to which our search is sensitive to. This observation also implies that there have been no significant mass-ratio mergers in the past  $\sim 5 \text{ Gyr}$ . This is consistent with observations of the scale height of the thin disk and the properties



**Figure 11.** Left: data and analyses for the line of sight along which we found the element of cold substructure PCII-18 from Table 3. Right: data and analyses for the line of sight along which we found the element of cold substructure PCII-19 from Table 3. See the caption to Figure 3 for a detailed description of this type of figure.

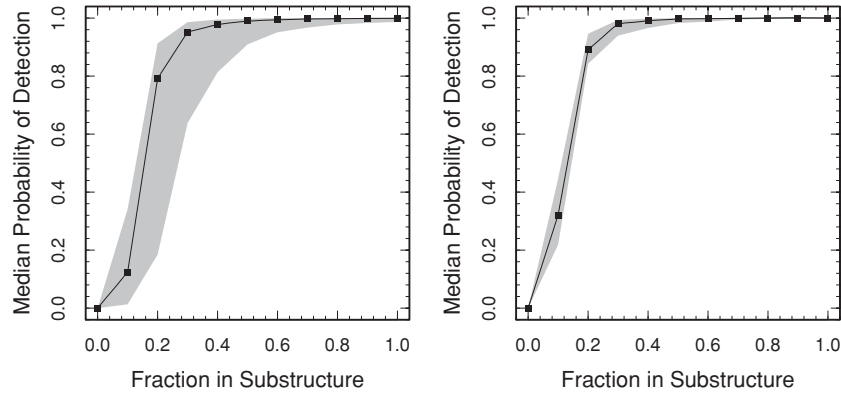


**Figure 12.** Results of our completeness calculation for the class II peak algorithm. Left: the completeness result for every line of sight. Right: the average completeness result in bins by the number of spectra obtained by SEGUE along lines of sight;  $N$  is the number of lines of sight with  $n$  spectra.

of the bulge, and suggests that the Milky Way has a less active merger history than might be expected for a halo of its mass (e.g., Kormendy & Kennicutt 2004; Stewart et al. 2008; Howard et al. 2009). In summary, the union of the Bell et al. (2008) result with our work suggests that the stellar accretion history of the Milky Way has been more or less constant with no significant mass-ratio mergers over the past  $\sim 5$  Gyr.

Direct comparison of our results with more local searches using 6D phase-space (e.g., Helmi & White 1999; Gould 2003;

Klement et al. 2008; Seabroke et al. 2008; Klement et al. 2009; Morrison et al. 2009) information is not as simple. While it is true that in a given volume, 6D phase-space searches are sensitive to the oldest extant substructures and capable of probing the accretion history farthest into the past, the fact that there is currently no proper motion database that is both sufficiently large and sufficiently precise in the volume we search precludes a 6D search. Published searches for substructure in 6D phase space are limited to the local volume—a region that does not



**Figure 13.** Median detection probability for the class II peak algorithm as a function of the fraction in substructure; the gray area is the  $1\sigma$  region. Left: for all 115 lines of sight with more than 30 spectra. Right: for all 47 lines of sight with more than 100 spectra.

**Table 4**  
Summary of Completeness Calculation

R.A.	Decl.	l	b	$N_s^a$	$N_p^b$	Volume <sup>c</sup>	$\Omega_b^d$	$\Omega_1^e$	$\Omega_{II}^f$	Comment
207.2	18.6	3.2	74.3	61	359	2.7	0.55	0.48	0.43	...
229.4	7.2	9.8	50	16	82	1.08	>1	0.97	>1	...
243.8	16.7	31.4	41.9	20	89	1.11	>1	0.82	>1	...
238.5	26.5	42.9	49.5	29	108	2.22	0.8	0.69	0.75	...
253.1	24	44	36.1	16	135	1.28	>1	>1	>1	...
320.6	-7.2	44.8	-36.7	19	101	2.16	>1	>1	>1	...
311	0	46.6	-24.8	10	437	1.06	>1	>1	>1	...
271.6	23.7	50	20	11	125	1.41	>1	>1	>1	...
266.5	25.4	50	25	12	188	1.07	>1	>1	>1	...
261.2	27	50	30	16	160	2.37	>1	>1	>1	...

#### Notes.

<sup>a</sup> The number of spectra obtained along that line of sight that lie in the inner halo as defined in Section 2.

<sup>b</sup> The number of photometrically classified MPMSTO stars in the volume scanned by SEGUE along that line of sight that lie in the inner halo as defined in Section 2.

<sup>c</sup> In  $\text{kpc}^3$ .

<sup>d</sup> The fraction of the total MPMSTO spectra sample that must belong to a single cold element of substructure for it to be classified as a bin detection 95% of the time.

<sup>e</sup> The fraction of the total MPMSTO spectra sample that must belong to a single cold element of substructure for it to be classified as a class I detection 95% of the time.

<sup>f</sup> The fraction of the total MPMSTO spectra sample that must belong to a single cold element of substructure for it to be classified as a class II detection 95% of the time.

(This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.)

overlap with the volume we searched in the inner halo. The rate at which substructure phase mixes is a function of the galactocentric distance; so constructing a consistent accretion history including both the radial velocity substructures we detect in the inner halo and phase-space substructures discovered in the local volume is problematic. While it appears that there is more substructure in 6D solar neighborhood searches, perhaps to the point that there is no smooth component as in Morrison et al. (2009), the dynamical effects of the disk on the local volume population make comparison between our result and local volume results difficult (e.g., Dehnen & Binney 1998; Bovy et al. 2009). As a result, placing the debris of the accretion activity identified in local volume samples in the same accretion timeline as our results will require more detailed modeling of the Milky Way's accretion and dynamical evolution than is

**Table 5**  
Known Substructure That Our Algorithms Recover

Stream	R.A.	Decl.	l	b	$d^a$	$v_r^b$	$\sigma^c$	Err <sup>d</sup>
Grillmair & Dionatos	163.8	48	162.4	59.2	$6.9^{+3.6}_{-1.6}$	-132	11.7	4.6
Monoceros	39.7	28.2	150	-29	$10.5^{+2.7}_{-1.3}$	-57	10.2	12.8
Monoceros	134	3.2	225.2	29	$9.7^{+1.5}_{-0.5}$	85	14.9	4

#### Notes.

<sup>a</sup> Median heliocentric distance in kpc.

<sup>b</sup> Radial velocity in  $\text{km s}^{-1}$  at which  $\Theta(v_r)$  peaks.

<sup>c</sup> Velocity dispersion in  $\text{km s}^{-1}$  of the significant peak.

<sup>d</sup> Median radial velocity error in  $\text{km s}^{-1}$  for all MPMSTO radial velocities within  $12 \text{ km s}^{-1}$  of the peak in  $\Theta(v_r)$ .

currently available. In the future, Gaia measurements of the full 6D phase-space distribution of SEGUE MPMSTO will allow us to construct a self-consistent inner halo stellar accretion timeline all the way back to the last major instance of violent relaxation.

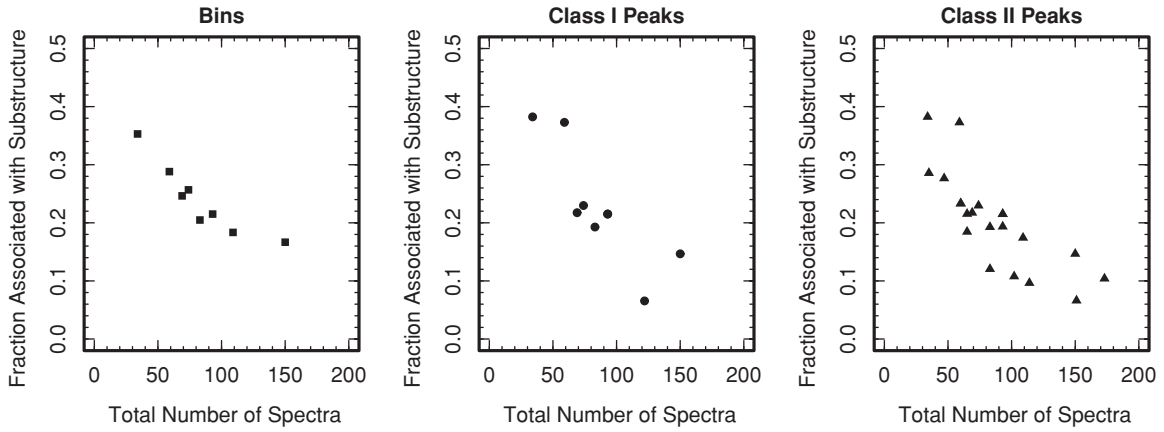
It is also difficult to assess the relative performance of our algorithm optimized to work on densely sampled in situ data with precise radial velocities but imprecise distance estimates with algorithms designed to work on very sparsely sampled in situ data with precise distance estimates. Searches like the latter are appropriate for distant, luminous, and rare tracers of the outer halo (e.g., Starkenburg et al. 2009).

### 5.2. Previously Known Substructure

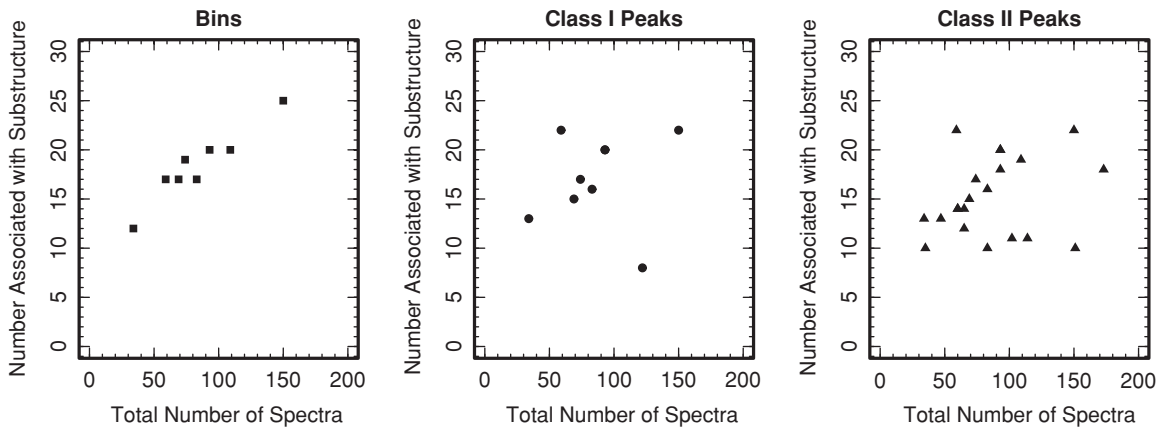
Thirteen lines of sight from the 137 total lines of sight in our sample were targeted at known substructures: the Grillmair & Dionatos (2006) stream, the Monoceros stream, the Orphan Stream, the Sagittarius stream, and the Virgo stream. Table 5 lists the radial velocities and mean heliocentric distances associated with each detection for the lines of sight targeted at known substructure for which we found an ECHOS. For the lines of sight targeted at known substructure along which we have non-detections, we list in Table 6 upper limits on the fraction of the MPMSTO star population in ECHOS along those lines of sight. We discuss each case in detail below.

#### 5.2.1. Grillmair & Dionatos (2006) Stream

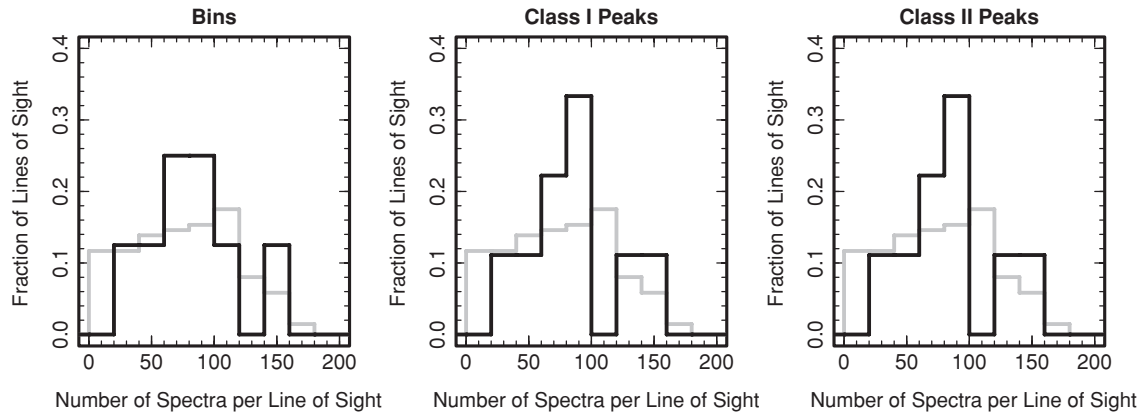
We detect the Grillmair & Dionatos (2006) stream along one line of sight listed in Table 5. Our estimate of its radial velocity is based on a line of sight with equatorial coordinates substantially different from the fiducial radial velocity point given for the best-fit model in Grillmair & Dionatos (2006). These authors used a line of sight centered at (R.A., decl.) = (202.0, 58.4)



**Figure 14.** Plots showing the fraction in substructure vs. the total number of spectra obtained along the line of sight where the substructure was detected for all detections and all methods. Note that the lower detected fractional overdensities are associated with a large number of spectra in agreement with our completeness calculation.



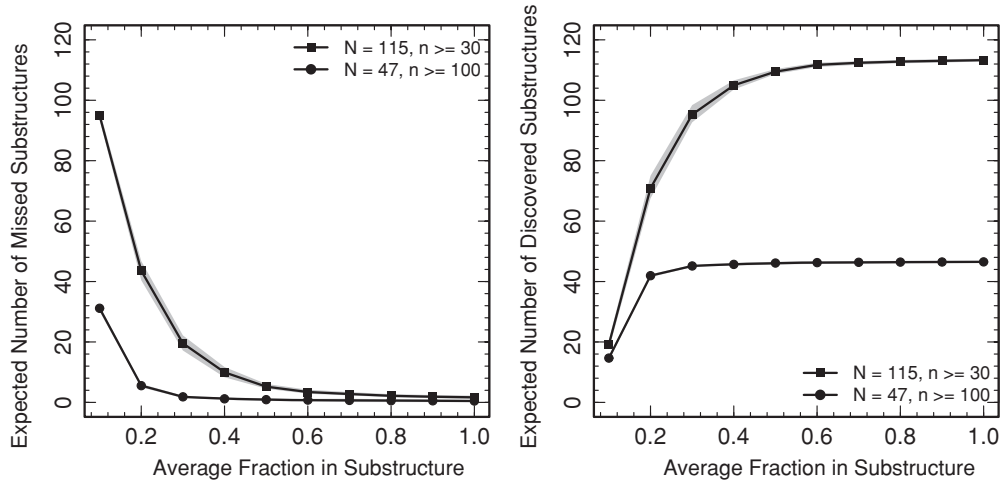
**Figure 15.** Plots showing the number of stars associated with an element of substructure vs. the total number of spectra obtained along the line of sight where the substructure was detected for all detections and all methods. Note that for the peak algorithms there is no correlation between number of stars associated with an element of substructure and the total number of spectra obtained along that line of sight. This fact combined with the result of our completeness calculation implies that our detections do not cluster just at our detection thresholds. Therefore, it is not likely that the inner halo is comprised of an array of substructures below our detection thresholds.



**Figure 16.** Histograms of the distribution of the number of spectra for each line of sight. We plot the distribution of our entire set of lines of sight in gray and the subset of lines of sight with detections in black. Note that the distribution of the total number of SEGUE spectra obtained along lines of sight with substructure detections is similar to the same distribution for all lines of sight in our sample. That is, our detections are not just found along the lines of sight where SEGUE most densely sampled the inner halo MPMSTO population. This suggests that the inner halo is not made up of a population of diffuse substructure below our sensitivity thresholds.

and found  $v_r = -208 \pm 30 \text{ km s}^{-1}$ ; more precise modeling is necessary to determine if the two observations are consistent. Our estimate of  $6.9^{+3.6}_{-1.6}$  kpc to the Grillmair & Dionatos (2006) stream is in agreement with the heliocentric distance obtained by those authors of 7.7 kpc. Willett et al. (2009) used SEGUE spectroscopic data to derive the properties of the Grillmair &

Dionatos (2006) stream along the line of sight listed in Table 5; they found  $v_r = -124 \text{ km s}^{-1}$  and a heliocentric distance of 8.8 kpc, both consistent with our measurements. We also note that we resolve the Grillmair & Dionatos (2006) stream in radial velocity and find that its radial velocity dispersion ( $\sigma = 11.7 \text{ km s}^{-1}$ ) is much hotter than the estimated radial



**Figure 17.** Left: the expected number of undetected substructures as a function of the average fraction in substructure for the class II peak algorithm; the gray area is the  $1\sigma$  region. Right: the expected number of detected substructures as a function of the average fraction in substructure for the class II peak algorithm; the gray area is the  $1\sigma$  region.

**Table 6**  
Known Substructure That Our Algorithms Do Not Recover

Stream	R.A.	Decl.	$l$	$b$	$\Omega_b^a$	$\Omega_1^b$	$\Omega_{II}^c$
Grillmair & Dionatos	217.7	58.2	100.6	54.4	0.39	0.35	0.28
Grillmair & Dionatos	158.6	44.3	171.7	57.6	0.29	0.27	0.21
Monoceros	357.3	39.3	110	-22	>1	>1	>1
Monoceros	111.3	37.6	180.9	22.4	>1	>1	>1
Monoceros	118	23.2	197.7	23.2	0.89	0.78	0.87
Orphan	152.4	25.9	205.4	53.9	0.35	0.29	0.23
Orphan	156.5	17.7	220.9	55.3	0.35	0.29	0.26
Orphan	162	0	250.3	49.8	0.28	0.26	0.19
Sagittarius	25.3	-9.4	158.8	-68.7	0.38	0.31	0.27
Virgo	186	0	288.2	62.1	0.47	0.4	0.36

#### Notes.

<sup>a</sup> The fraction of the total MPMSTO spectra sample that must belong to a single cold element of substructure for it to be classified as a bin detection 95% of the time.

<sup>b</sup> The fraction of the total MPMSTO spectra sample that must belong to a single cold element of substructure for it to be classified as a class I detection 95% of the time.

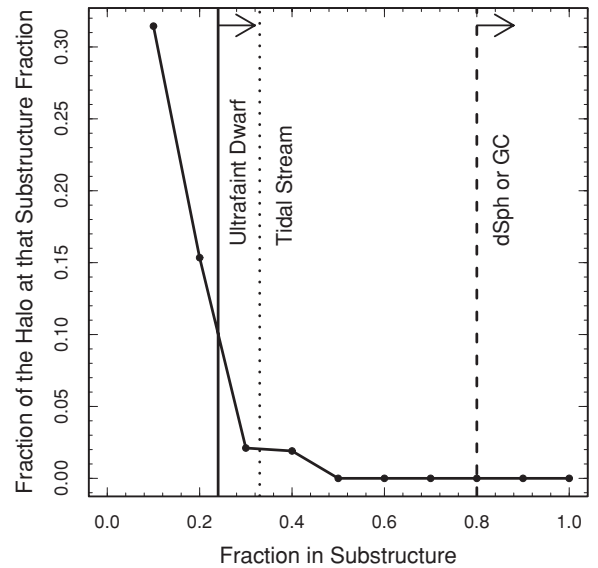
<sup>c</sup> The fraction of the total MPMSTO spectra sample that must belong to a single cold element of substructure for it to be classified as a class II detection 95% of the time.

velocity errors ( $\text{Err} = 4.6 \text{ km s}^{-1}$ ) associated with the stars in our detection.

We fail to detect the Grillmair & Dionatos (2006) stream along two lines of sight listed in Table 6. Our non-detection along the line of sight targeted at (R.A., decl.) = (217.7, 58.2) is consistent with Figure 1 of Grillmair & Dionatos (2006), as the stream is nearly invisible in their matched-filter analysis at those coordinates. We plot the data for the line of sight targeted at (R.A., decl.) = (158.6, 44.3) in Figure 23—the substructure present in the data is manifestly not cold as it has a velocity dispersion of at least  $40 \text{ km s}^{-1}$ . The large velocity dispersion is not the reason for its non-detection (as we showed in Section 3.5.1); its non-detection is due to the fact that the mean radial velocity of the apparent feature at the bright end is offset by  $40 \text{ km s}^{-1}$  from its mean radial velocity at the faint end.

#### 5.2.2. Monoceros Stream

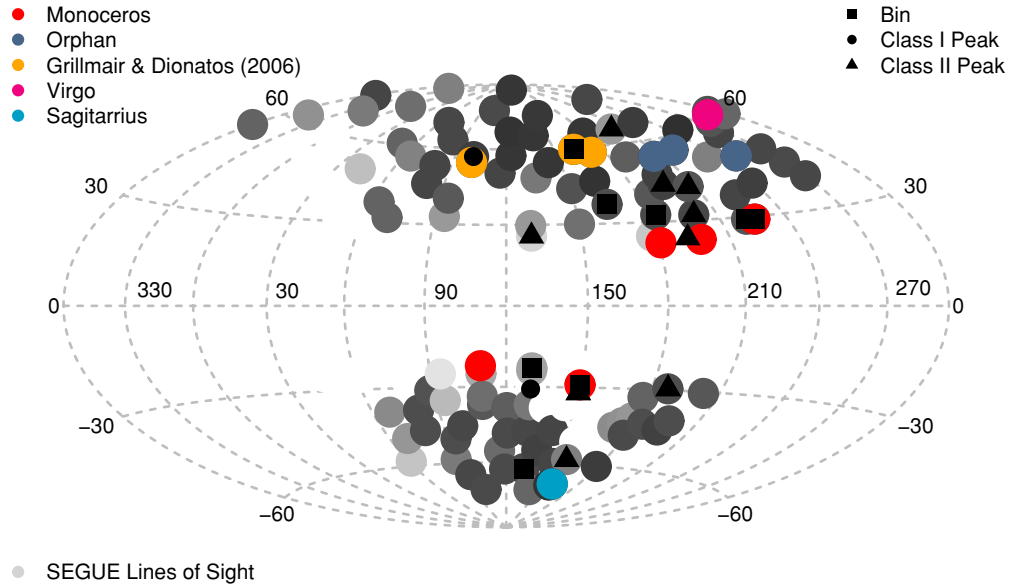
We detect the Monoceros stream along two lines of sight listed in Table 5. Our estimates for its heliocentric distance and



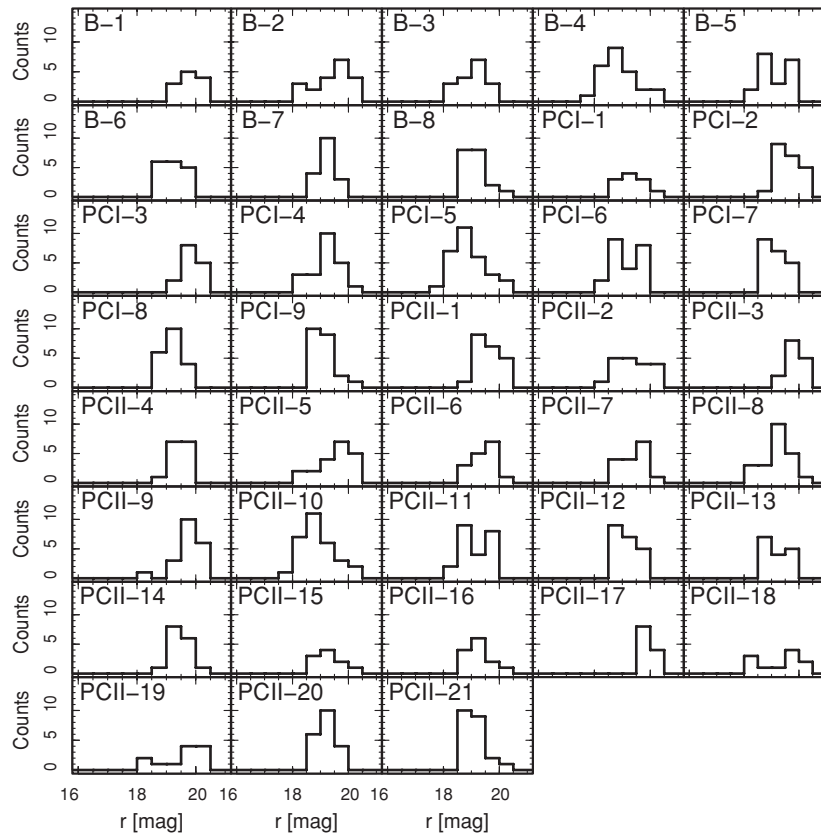
**Figure 18.** Volume fraction of the halo at a given fraction in substructure. Note that about 1/3 of the halo (by volume) has 10% of its MPMSTO population in ECHOS and about 1/6 of the halo (by volume) has 20% of its MPMSTO population in ECHOS; the fraction of the halo (by volume) with more than 20% of its MPMSTO population in ECHOS is just a few percent. We also plot the expected fraction of the halo in ECHOS with properties similar to ultrafaint dwarf galaxies, known tidal streams like Monoceros and Grillmair & Dionatos (2006), and classical dwarf spheroidal galaxies and globular clusters. There are unlikely to be ECHOS like undiscovered classical dwarf spheroidal galaxies or globular clusters in the inner halo, and only a few percent of the halo hosts ECHOS like the Monoceros or Grillmair & Dionatos (2006) tidal streams. Our search does not rule out the possibility that there could be ECHOS like ultrafaint dwarf galaxies in the inner halo.

radial velocity are consistent with the comprehensive model for the Monoceros stream given in Peñarrubia et al. (2005), as well as with previous observational results referenced therein. We resolve the Monoceros stream in radial velocity and find that its radial velocity dispersion ( $\sigma = 10.2 \text{ km s}^{-1}$ ) is much hotter than the estimated radial velocity errors ( $\text{Err} = 4.0 \text{ km s}^{-1}$ ) associated with the stars in our detection. We also note that many of our low galactic latitude detections in Tables 1–3 are plausibly related to Monoceros as well.

We fail to detect the Monoceros stream along three lines of sight listed in Table 6. Our non-detections along those lines of



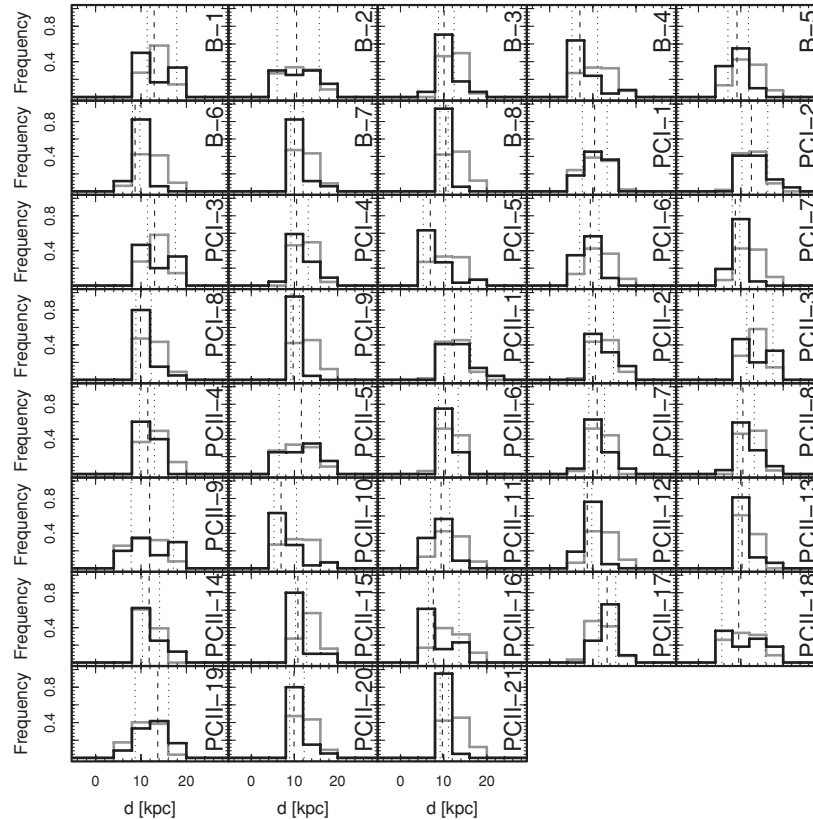
**Figure 19.** Hammer projection of our detections in galactic coordinates. We plot all 137 of the SEGUE lines of sight we analyze in a gray scale. The darker circles are those lines where we are the most complete according to Table 4 and lighter circles where we are least complete. We indicate lines of sight that were pointed at pieces of the sky expected to potentially intersect a known element of substructure by coloring that line of sight according to the legend. We plot our detections from the bin algorithm as black squares, our class I detections from the peak algorithm as black circles, and our class II detections from the peak algorithm as black triangles. The distribution of our ECHOS in galactic coordinates is consistent with an isotropic distribution given our completeness. If a piece of substructure was discovered by more than one algorithm we only plot the symbol corresponding to the most robust algorithm. In order of decreasing robustness: bins  $\approx$  class I peaks  $>$  class II peaks. (A color version of this figure is available in the online journal.)



**Figure 20.**  $r$ -magnitude histograms for stars in our detections listed in Tables 1–3—the identifier in each panel corresponds to its ID number in Tables 1–3. For the bin detections the histograms include all stars with radial velocities that place them in the significant bin while for the peak detections the histograms include all stars with radial velocities that place them within one median velocity error of the significant radial velocity peak.

sight are almost certainly because of our lack of sensitivity to substructure along those lines of sight, primarily because most of the spectra obtained by SEGUE along those lines of sight

belonged to MPMSTO stars that did not fall within the inner halo as defined in Section 2. As a result, we had too few radial velocities to find anything significant.



**Figure 21.** Approximate heliocentric distance histograms for stars in our detections listed in Tables 1–3—the identifier in each panel corresponds to its ID number in Tables 1–3. In black we plot approximate heliocentric distance distributions for all stars with radial velocities that place them in the significant bin (for bin detections) or within one median velocity error of the significant radial velocity peak (for the peak detections). The dashed vertical line denotes the median heliocentric distance and the two vertical dotted lines delimit the interval that contains 95% of the distribution. In the same panels we plot in gray the average heliocentric distance distribution of all stars in our mock catalog along that line of sight.

### 5.2.3. Orphan Stream

We fail to detect the Orphan stream along three lines of sight listed in Table 6. However, our non-detections along those lines of sight are consistent with the Belokurov et al. (2007) analysis because they find that the stream should be beyond our heliocentric distance threshold of 17.5 kpc at heliocentric distance between 20 and 30 kpc at those coordinates.

### 5.2.4. Sagittarius Stream

We fail to detect the Sagittarius stream along one line of sight listed in Table 6. However, our non-detection along that line of sight is consistent with the comprehensive model for the Sagittarius stream given in Law et al. (2005) as well as with previous observational results referenced therein. That is, Law et al. (2005) predict that the Sagittarius tidal stream should be beyond our heliocentric distance threshold of 17.5 kpc at a heliocentric distance  $\gtrsim 20$  kpc.

### 5.2.5. Virgo Stream

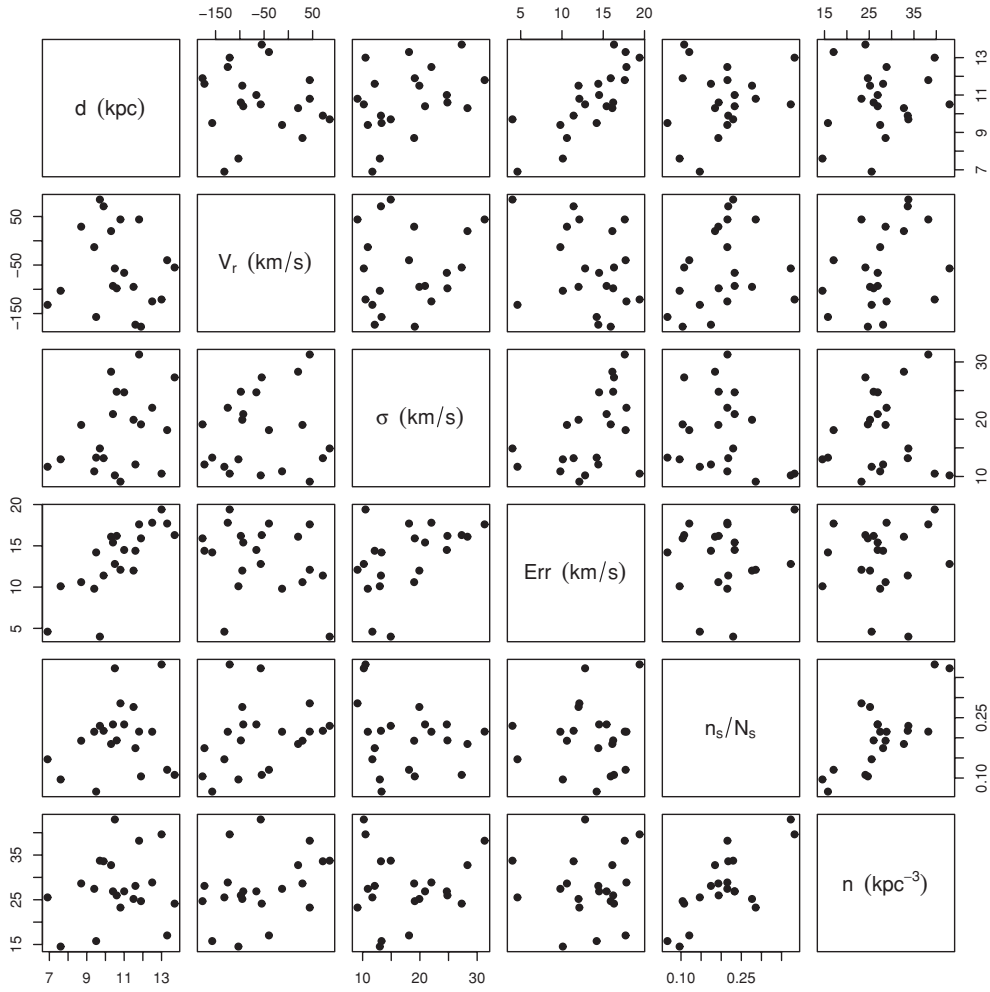
We fail to detect the Virgo stream along one line of sight listed in Table 6. However, our non-detection along that line of sight is consistent with previous analyses (e.g., Duffau et al. 2006; Martínez-Delgado et al. 2007) because their RR Lyrae were all beyond our heliocentric distance threshold of 17.5 kpc at heliocentric distances greater than 18 kpc.

## 5.3. Implications for the Formation of the Milky Way

Our seven new high-confidence ECHOS greatly expand the known number of inner halo substructures, and our detections

of previously known elements of substructure can be used to further constrain models of the substructures themselves and their progenitors. The existence of a substantial population of low density ECHOS in the inner halo provides a strong constraint that theoretical models must meet. At the same time, we note that our observed radial velocity distributions taken as a whole give us little reason to reject a smooth model for the radial velocity distribution of the inner halo. These observations are therefore consistent with inner halo formation scenarios in which relatively massive protogalaxies accrete into the nascent Milky Way early on. As a result of the massive mergers, the potential of the nascent Milky Way changes on short timescales, so violent relaxation smooths out the stellar phase-space distribution. The substructure that remains is mostly erased as the number of crossing times since accretion grows large. On the other hand, Morrison et al. (2009) examined a solar neighborhood sample with 6D phase-space information and concluded that violent relaxation was not efficient. We expect that the efficiency of violent relaxation in the Milky Way’s past will remain an active area of research.

We cannot assess the relative probability that a single accretion event produced the seven ECHOS we observe as compared to seven unique accretion events, nor can we unambiguously determine a class of progenitors. Our ECHOS could result from the disruption of globular clusters, small dwarf spheroidal galaxies, large Large Magellanic Cloud-like progenitors of the bulk of the inner halo, or from dynamical interactions of any of those three classes of objects with the stellar disk of the nascent Milky Way. Nevertheless, there is an enormous amount of information



**Figure 22.** Multiplot for the properties of our class II detections: median distance  $d$  in kiloparsecs, radial velocity  $v_r$  in  $\text{km s}^{-1}$ , velocity dispersion  $\sigma$  in  $\text{km s}^{-1}$ , median radial velocity error in  $\text{km s}^{-1}$ , the fraction in substructure  $n_s/N_s$ , and number density  $n$  in  $\text{kpc}^{-3}$ . There are no obvious trends save for those expected from instrumental limitations and basic physics. Substructures at greater distance appear to have larger velocity dispersions (and larger median errors) because of decreasing radial velocity precision for faint stars. Substructures with larger fractional overdensities also tend to have larger physical number densities.

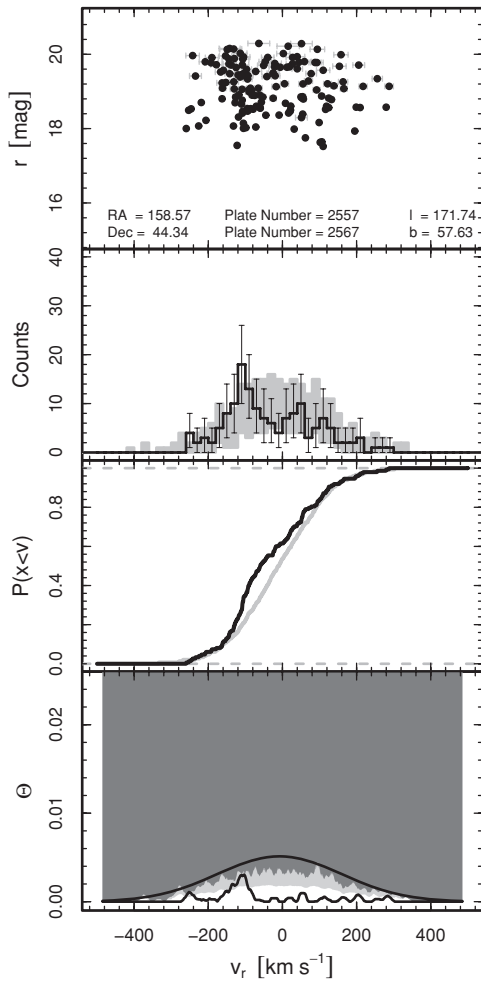
left to be extracted from our detections. Many authors (e.g., Wheeler et al. 1989; Nissen et al. 1994; Carretta et al. 2000) have observed that inner halo stars in the solar neighborhood are enriched in  $\alpha$ -elements relative to stars in surviving classical dwarf spheroidal galaxies at constant  $[\text{Fe}/\text{H}]$ . Robertson et al. (2005) and Font et al. (2006) explained this observation in the context of the hierarchical paradigm by noting that most of the stars in the inner halo were formed in a few relatively massive ( $\sim 5 \times 10^{10} M_\odot$ ) protogalaxies that merged with the nascent Milky Way  $\sim 10$  Gyr in the past. The star formation histories of those protogalaxies would therefore have been sharply truncated, resulting in enrichment mostly by Type II supernovae. On the other hand, the surviving classical dwarf spheroidals have lower mass ( $\sim 10^9 M_\odot$ ) with more sustained star formation histories that allow for chemical enrichment by Type Ia supernovae. Even moderate-resolution spectroscopic follow-up of our ECHOS using the techniques presented in Kirby et al. (2008) should reveal the degree of  $\alpha$ -enhancement in the stars in each ECHOS. This may even be feasible using the subset of existing SEGUE spectra with sufficiently high S/N (Y. S. Lee et al. 2009, in preparation). In any case, individual stellar  $[\text{Fe}/\text{H}]$  and  $[\alpha/\text{Fe}]$  measurements within an ECHOS will reveal the distribution in composition within single, massive, long-age disrupted inner halo progenitors. This information has the

potential to uniquely inform not only models of Milky Way formation, but also the physics of the high redshift universe and the star formation environments in the ancient massive stellar systems that merged with the nascent Milky Way to form the inner halo.

## 6. CONCLUSION

We used the observed spatial and radial velocity distribution of MPMSTO stars in 137 SEGUE lines of sight to identify 10—seven for the first time—high-confidence ECHOS in the inner halo of the Milky Way, none of which we expect to be false positives. We also found 21 lower confidence ECHOS of which we expect three to be false positives. ECHOS are the debris of ancient merger events, and we used our detections and completeness estimates to infer that at most  $0.34^{+0.02}_{-0.02}$  of the MPMSTO stars in the inner halo belong to ECHOS. Our result also implies that there exists a significant population of low fractional overdensity ECHOS in the inner halo; we predict that 1/3 of the inner halo (by volume) hosts low density ECHOS with number densities  $n \approx 15 \text{ kpc}^{-3}$  and that there are of order  $10^3$  ECHOS in the entire inner halo. When combined with the work of Bell et al. (2008), our result suggests that there has been a constant rate of merger activity over the past  $\sim 5$  Gyr with no





**Figure 23.** Data and analyses for the line of sight target at (R.A., decl.) = (158.6, 44.3) expected to intersect the Grillmair & Dionatos (2006) stream. Note the feature obvious to the naked eye but invisible to our detection algorithms because of the strong dependence of the mean radial velocity of the feature on  $r$ -magnitude and therefore distance. The feature has a velocity dispersion of at least  $40 \text{ km s}^{-1}$  so it cannot be described as cold. That large velocity dispersion is not the reason for its non-detection (as we showed in Section 3.5.1); its non-detection is due to the fact that the mean radial velocity of the apparent feature at the bright end is offset by  $40 \text{ km s}^{-1}$  from its mean radial velocity at the faint end.

accretion of single stellar systems with mass more than a few percent of a Milky Way mass in that time.

We thank James Bullock, Jürg Diemand, Evan Kirby, David Lai, Doug Lin, and Piero Madau for useful comments and conversation. We are especially grateful to Heather Morrison and Matthias Steinmetz for their insightful comments on an earlier draft of this paper. This research has made use of NASA's Astrophysics Data System Bibliographic Services. This material is based upon work supported under a National Science Foundation Graduate Research Fellowship. T.C.B. and Y.S.L. acknowledge partial support for this work from PHY 02-16783 and PHY 08-22648: Physics Frontiers Center/Joint Institute for Nuclear Astrophysics (JINA), awarded by the U.S. National Science Foundation. Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the U.S. Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck

Society, and the Higher Education Funding Council for England. The SDSS Web Site is <http://www.sdss.org/>.

The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The Participating Institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, University of Cambridge, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle Astrophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory, and the University of Washington.

*Facilities:* Sloan

## APPENDIX A

### PHASE-SPACE STRUCTURE OF THE SMOOTH COMPONENT OF THE INNER HALO AS VIEWED FROM THE SUN

1. We model the galactocentric position-space distribution of stars in the inner halo by a spherically symmetric power law in radius with index  $\alpha = -3.5$  (e.g., Morrison et al. 2000; Yanny et al. 2000; Bell et al. 2008)

$$\rho \propto r^\alpha. \quad (\text{A1})$$

We draw  $n_s$  radial coordinates  $r'$  from the distribution described by Equation (A1). We then draw random  $\theta'$  and  $\phi'$  coordinates such that the points are spread uniformly over  $4\pi$  steradians. Together these three coordinates define the standard spherical coordinate vector  $r'$ .

2. We model the galactocentric spherical velocity-space distribution of stars in inner halo as a multivariate normal with mean  $\mu_{r,\theta,\phi}$  and the variance–covariance matrix  $\Sigma_{r,\theta,\phi}$  (e.g., Sommer-Larsen et al. 1997; Sirko et al. 2004a, 2004b; Xue et al. 2008)

$$\mathbf{v}'_{r,\theta,\phi} \sim \mathcal{N}(\mu_{r,\theta,\phi}, \Sigma_{r,\theta,\phi}) \quad (\text{A2a})$$

$$\mu_{r,\theta,\phi} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad (\text{A2b})$$

$$\Sigma_{r,\theta,\phi} = \begin{pmatrix} 120^2 & 0 & 0 \\ 0 & 100^2 & 0 \\ 0 & 0 & 100^2 \end{pmatrix}. \quad (\text{A2c})$$

We draw  $n_s$  galactocentric spherical velocities  $\mathbf{v}'_{r,\theta,\phi}$  from the distribution described by above and associate them with the position-space distribution derived in step 1.

3. We transform the galactocentric spherical velocities into Cartesian velocities using the transformation  $\mathbf{v}'_{x,y,z} = \mathbf{A} \mathbf{v}'_{r,\theta,\phi}$  defined by the matrix  $\mathbf{A}$

$$\mathbf{A} = \begin{pmatrix} \cos \theta' \sin \phi' & -\sin \theta' & \cos \theta' \cos \phi' \\ \sin \theta' \sin \phi' & \cos \theta' & \sin \theta' \cos \phi' \\ \cos \phi' & 0 & -\sin \phi' \end{pmatrix}. \quad (\text{A3})$$

Then transform the velocity distribution into the Sun's standard of rest (e.g., Dehnen & Binney 1998)

$$v_x = v'_x - v_{\odot,x} \quad (\text{A4a})$$

$$v_y = v'_y - v_{\odot,y} \quad (\text{A4b})$$

$$v_z = v'_z - v_{\odot,z}. \quad (\text{A4c})$$

4. We transform the galactocentric spherical coordinates  $r'$  into galactocentric Cartesian coordinates  $x'$  with the usual transformation

$$x' = r' \cos \theta' \sin \phi' \quad (\text{A5a})$$

$$y' = r' \sin \theta' \sin \phi' \quad (\text{A5b})$$

$$z' = r' \cos \phi'. \quad (\text{A5c})$$

We translate the distribution such that the zero point of the  $x$ -coordinate corresponds to the position of the Sun

$$x = x' + 8 \quad (\text{A6a})$$

$$y = y' \quad (\text{A6b})$$

$$z = z'. \quad (\text{A6c})$$

We compute Sun-centered spherical coordinates  $\mathbf{r}$  from the Sun-centered Cartesian coordinates  $x, y, z$ .

5. We project the Sun-centered Cartesian velocities onto the line of sight between the synthetic star and the Sun using the transformation  $\mathbf{v}_{r,\theta,\phi} = \mathbf{B}\mathbf{v}_{x,y,z}$  defined by the matrix  $\mathbf{B}$ :

$$\mathbf{B} = \begin{pmatrix} \cos \theta \sin \phi & \sin \theta \sin \phi & \cos \phi \\ \frac{-\sin \theta}{Kr \sin \phi} & \frac{\cos \theta}{Kr \sin \phi} & 0 \\ \frac{\cos \theta \cos \phi}{Kr} & \frac{\sin \theta \cos \phi}{Kr} & \frac{-\sin \phi}{Kr} \end{pmatrix}, \quad (\text{A7})$$

where  $K$  is a constant of proportionality between kilometers and kiloparsecs. Finally, we eliminate all synthetic stars that would fall outside of our definition of the inner halo.

## APPENDIX B

### DETAILED DESCRIPTION OF THE BIN ALGORITHM

1. Consider each SEGUE line of sight in sequence and let  $n$  be the total number of MPMSTO star spectra obtained along that line of sight. We compute the histogram describing the MPMSTO star radial velocity distribution along that line of sight. We use bootstrap resampling to estimate the uncertainty in the number of counts in each bin.
2. Under the null hypothesis, the radial velocity distribution of the Milky Way's inner halo can be calculated as discussed in Section 3.1. For each SEGUE line of sight, we determine which synthetic stars from our mock catalog fall within the volume scanned by SEGUE along that line of sight. There are typically more than an order of magnitude more synthetic stars  $m$  in a given patch of sky than the number of MPMSTO star spectra  $n$  observed along the corresponding SEGUE line of sight. We select a random subsample  $S$  of  $n$  synthetic stars from the  $m$  available and compute the histogram of that subsample.

3. We repeat step 2 a large number of times. In this way, we calculate the median histogram that results from observing the mock catalog a large number of times as well as distributions for the number of counts in each bin. In this analysis, we have always repeated step 2  $10^4$  times.
4. We identify bins for which the distribution of counts estimated in step 1 is inconsistent with the distribution calculated in step 3 and flag the stars in that radial velocity bin as a potential element of cold substructure. An inconsistent bin is one for which the 95% confidence interval on the number of counts in the bin from the observed MPMSTO population from bootstrap resampling does not overlap with the 95% confidence interval for the expected number of counts in the bin from the mock catalog under the null hypothesis.

## APPENDIX C

### DETAILED DESCRIPTION OF THE PEAK ALGORITHM

1. Consider each SEGUE line of sight in sequence and let  $n$  be the total number of MPMSTO star spectra obtained along that line of sight. Let  $F(v_r)$  and  $F'(v_r)$  denote the CDF of the radial velocities observed along that line of sight and its slope, respectively.
2. Under the null hypothesis, the radial velocity distribution of the Milky Way's inner halo can be calculated as discussed in Section 3.1. For each SEGUE line of sight, we determine which synthetic stars from our mock catalog fall within the volume scanned by SEGUE along that line of sight. There are typically more than an order of magnitude more synthetic stars  $m$  in a given patch of sky than the number of MPMSTO star spectra  $n$  observed along the corresponding SEGUE line of sight. We select a random subsample  $S$  of  $n$  synthetic stars from the  $m$  available and compute the CDF  $F_S(v_r)$  and its slope  $F'_S(v_r)$  of that subsample.
3. We repeat step 2 a large number of times. In this way, we calculate the distribution of the CDF and its slope at each point in radial velocity space. Specifically, its average value  $\overline{F_S(v_r)}$  and an estimate of its average slope  $\overline{F'_S(v_r)}$ . In this analysis, we have always repeated step 2  $10^4$  times.
4. Again we select a random subsample of  $n$  stars from the  $m$  available, and compute the CDF  $F_S(v_r)$  and slope  $F'_S(v_r)$  of this subsample. We then calculate the difference  $\Theta_S(v_r) = F_S(v_r) - \overline{F_S(v_r)}$  and smooth it using a moving average kernel with its width set to  $10 \text{ km s}^{-1}$ , very close to the median velocity error of the MPMSTO sample.
5. We repeat step 4 a large number of times. As a result, we calculate the distribution of  $\Theta_S(v_r)$ , or in other words, the differences between the average value of the slope  $\overline{F'_S(v_r)}$  and a single random realization  $F'_S(v_r)$  under the null hypothesis. In particular, we compute formal significance contours that correspond to 1 in  $10^2$ , 1 in  $10^3$ , and 1 in  $10^4$  events. In this analysis, we have always repeated step 2  $10^4$  times. We emphasize that the distribution is not Gaussian; nevertheless, these significance thresholds would naively correspond to  $2.33\sigma$ ,  $3.09\sigma$ , and  $3.72\sigma$ .
6. We compute  $\Theta(v_r) = F'(v_r) - \overline{F'_S(v_r)}$ , the difference between the observed slope along a single SEGUE line of sight and the average slope under the null hypothesis, and smooth as before. Note that since we normalize the number of synthetic stars in  $S$  to the number of stars  $n$  observed along the SEGUE line of sight, every interval in

which  $\Theta(v_r) < 0$  must necessarily correspond to an interval in which  $\Theta(v_r) > 0$ ; only the intervals with  $\Theta(v_r) > 0$  correspond to an overdensity.

- We flag any radial velocity  $v_r$  at which  $\Theta(v_r)$ , the difference between the slope of the CDF of the MPMSTO radial velocity distribution and the average CDF of the smooth model, is significant at more than the 1 in  $10^4$  level as a potential element of cold substructure. We are formally limited to 1 in  $10^4$  events because of computational limits on the number of Monte Carlo iterations we can execute.

We can self-consistently estimate the radial velocity dispersion of the candidate ECHOS identified by the peak algorithm by fitting a Gaussian to the overdensity in  $\Theta(v_r)$  in a window centered on the peak of the overdensity with width 6 times our median velocity resolution, such that the window contains 99% of the signal from the detection.

## REFERENCES

- Abadi, M. G., Navarro, J. F., & Steinmetz, M. 2006, *MNRAS*, **365**, 747
- Abadi, M. G., Navarro, J. F., Steinmetz, M., & Eke, V. R. 2003a, *ApJ*, **591**, 499
- Abadi, M. G., Navarro, J. F., Steinmetz, M., & Eke, V. R. 2003b, *ApJ*, **597**, 21
- Abazajian, K. N., et al. 2009, *ApJS*, **182**, 543
- Allende Prieto, C., Beers, T. C., Wilhelm, R., Newberg, H. J., Rockosi, C. M., Yanny, B., & Lee, Y. S. 2006, *ApJ*, **636**, 804
- Allende Prieto, C., et al. 2008, *AJ*, **136**, 2070
- An, D., et al. 2008, *ApJS*, **179**, 326
- Bahcall, J. N., & Soneira, R. M. 1980, *ApJS*, **44**, 73
- Bell, E. F., et al. 2008, *ApJ*, **680**, 295
- Belokurov, V., Evans, N. W., Irwin, M. J., Hewett, P. C., & Wilkinson, M. I. 2006, *ApJ*, **637**, L29
- Belokurov, V., et al. 2007, *ApJ*, **658**, 337
- Binney, J., & Tremaine, S. 1987, *Galactic Dynamics* (Princeton, NJ: Princeton Univ. Press)
- Bovy, J., Hogg, D. W., & Roweis, S. T. 2009, *ApJ*, **700**, 1794
- Bower, R. G., Benson, A. J., Malbon, R., Helly, J. C., Frenk, C. S., Baugh, C. M., Cole, S., & Lacey, C. G. 2006, *MNRAS*, **370**, 645
- Bullock, J. S., & Johnston, K. V. 2005, *ApJ*, **635**, 931
- Carney, B. W., Laird, J. B., Latham, D. W., & Aguilar, L. A. 1996, *AJ*, **112**, 668
- Carollo, D., et al. 2007, *Nature*, **450**, 1020
- Carretta, E., Gratton, R. G., & Sneden, C. 2000, *A&A*, **356**, 238
- Chen, B., et al. 2001, *ApJ*, **553**, 184
- Chiba, M., & Beers, T. C. 2000, *AJ*, **119**, 2843
- Chiba, M., & Yoshii, Y. 1998, *AJ*, **115**, 168
- Clem, J. L., Vanden Berg, D. A., & Stetson, P. B. 2008, *AJ*, **135**, 682
- Croton, D. J., et al. 2006, *MNRAS*, **365**, 11
- De Lucia, G., & Helmi, A. 2008, *MNRAS*, **391**, 14
- Dehnen, W., & Binney, J. J. 1998, *MNRAS*, **298**, 387
- Diemand, J., Madau, P., & Moore, B. 2005, *MNRAS*, **364**, 367
- Diemand, J., Kuhlen, M., & Madau, P. 2007, *ApJ*, **667**, 859
- Diemand, J., Kuhlen, M., Madau, P., Zemp, M., Moore, B., Potter, D., & Stadel, J. 2008, *Nature*, **454**, 735
- Duffau, S., Zinn, R., Vivas, A. K., Carraro, G., Méndez, R. A., Winnick, R., & Gallart, C. 2006, *ApJ*, **636**, L97
- Eggen, O. J., Lynden-Bell, D., & Sandage, A. R. 1962, *ApJ*, **136**, 748
- Fellhauer, M., et al. 2006, *ApJ*, **651**, 167
- Font, A. S., Johnston, K. V., Bullock, J. S., & Robertson, B. E. 2006, *ApJ*, **638**, 585
- Fukugita, M., Ichikawa, T., Gunn, J. E., Doi, M., Shimasaku, K., & Schneider, D. P. 1996, *AJ*, **111**, 1748
- Gilmore, G., & Reid, N. 1983, *MNRAS*, **202**, 1025
- Gilmore, G., Wyse, R. F. G., & Norris, J. E. 2002, *ApJ*, **574**, L39
- Gould, A. 2003, *ApJ*, **592**, L63
- Governato, F., Willman, B., Mayer, L., Brooks, A., Stinson, G., Valenzuela, O., Wadsley, J., & Quinn, T. 2007, *MNRAS*, **374**, 1479
- Grillmair, C. J. 2006, *ApJ*, **645**, L37
- Grillmair, C. J. 2009, *ApJ*, **693**, 1118
- Grillmair, C. J., & Dionatos, O. 2006, *ApJ*, **643**, L17
- Grillmair, C. J., & Johnson, R. 2006, *ApJ*, **639**, L17
- Gunn, J. E., et al. 1998, *AJ*, **116**, 3040
- Gunn, J. E., et al. 2006, *AJ*, **131**, 2332
- Harding, P., Morrison, H. L., Olszewski, E. W., Arabadjis, J., Mateo, M., Dohm-Palmer, R. C., Freeman, K. C., & Norris, J. E. 2001, *AJ*, **122**, 1397
- Harris, W. E. 1976, *AJ*, **81**, 1095
- Helmi, A. 2004, *ApJ*, **610**, L97
- Helmi, A. 2008, *A&AR*, **15**, 145
- Helmi, A., Navarro, J. F., Nordström, B., Holmberg, J., Abadi, M. G., & Steinmetz, M. 2006, *MNRAS*, **365**, 1309
- Helmi, A., & White, S. D. M. 1999, *MNRAS*, **307**, 495
- Helmi, A., White, S. D. M., de Zeeuw, P. T., & Zhao, H. 1999, *Nature*, **402**, 53
- Helmi, A., et al. 2003, *ApJ*, **586**, 195
- Hogg, D. W., Finkbeiner, D. P., Schlegel, D. J., & Gunn, J. E. 2001, *AJ*, **122**, 2129
- Howard, C. D., et al. 2009, arXiv:0908.1109
- Ibata, R., Lewis, G. F., Irwin, M., Totten, E., & Quinn, T. 2001, *ApJ*, **551**, 294
- Ivezić, Ž., et al. 2000, *AJ*, **120**, 963
- Ivezić, Ž., et al. 2004, *Astron. Nachr.*, **325**, 583
- Ivezić, Ž., et al. 2008, *ApJ*, **684**, 287
- Johnston, K. V. 1998, *ApJ*, **495**, 297
- Johnston, K. V., Bullock, J. S., Sharma, S., Font, A., Robertson, B. E., & Leitner, S. N. 2008, *ApJ*, **689**, 936
- Johnston, K. V., Law, D. R., & Majewski, S. R. 2005, *ApJ*, **619**, 800
- Jurić, M., et al. 2008, *ApJ*, **673**, 864
- Kepley, A. A., et al. 2007, *AJ*, **134**, 1579
- Kirby, E. N., Guhathakurta, P., & Sneden, C. 2008, *ApJ*, **682**, 1217
- Klement, R., Fuchs, B., & Rix, H.-W. 2008, *ApJ*, **685**, 261
- Klement, R., et al. 2009, *ApJ*, **698**, 865
- Kormendy, J., & Kennicutt, R. C., Jr. 2004, *ARA&A*, **42**, 603
- Law, D. R., Johnston, K. V., & Majewski, S. R. 2005, *ApJ*, **619**, 807
- Lee, Y. S., et al. 2008a, *AJ*, **136**, 2022
- Lee, Y. S., et al. 2008b, *AJ*, **136**, 2050
- Lindgren, L., et al. 2008, *IAU Symp.*, **248**, 217
- Majewski, S. R. 1993, *ARA&A*, **31**, 575
- Majewski, S. R., Skrutskie, M. F., Weinberg, M. D., & Ostheimer, J. C. 2003, *ApJ*, **599**, 1082
- Martínez-Delgado, D., Peñarrubia, J., Jurić, M., Alfaro, E. J., & Ivezić, Ž. 2007, *ApJ*, **660**, 1264
- Meza, A., Navarro, J. F., Abadi, M. G., & Steinmetz, M. 2005, *MNRAS*, **359**, 93
- Monet, D. G., et al. 2003, *AJ*, **125**, 984
- Morrison, H. L. 1993, *AJ*, **106**, 578
- Morrison, H. L., Mateo, M., Olszewski, E. W., Harding, P., Dohm-Palmer, R. C., Freeman, K. C., Norris, J. E., & Morita, M. 2000, *AJ*, **119**, 2254
- Morrison, H. L., et al. 2009, *ApJ*, **694**, 130
- Munn, J. A., et al. 2004, *AJ*, **127**, 3034
- Newberg, H. J., et al. 2002, *ApJ*, **569**, 245
- Nissen, P. E., Gustafsson, B., Edvardsson, B., & Gilmore, G. 1994, *A&A*, **285**, 440
- Odenkirchen, M., et al. 2001, *ApJ*, **548**, L165
- Padmanabhan, N., et al. 2008, *ApJ*, **674**, 1217
- Peñarrubia, J., et al. 2005, *ApJ*, **626**, 128
- Perryman, M. A. C., et al. 2001, *A&A*, **369**, 339
- Pier, J. R., Munn, J. A., Hindsley, R. B., Hennessy, G. S., Kent, S. M., Lupton, R. H., & Ivezić, Ž. 2003, *AJ*, **125**, 1559
- Press, W. H., & Schechter, P. 1974, *ApJ*, **187**, 425
- Preston, G. W., Shectman, S. A., & Beers, T. C. 1991, *ApJ*, **375**, 121
- Ratnatunga, K. U., & Freeman, K. C. 1985, *ApJ*, **291**, 260
- Robertson, B., Bullock, J. S., Font, A. S., Johnston, K. V., & Hernquist, L. 2005, *ApJ*, **632**, 872
- Rocha-Pinto, H. J., Majewski, S. R., Skrutskie, M. F., Crane, J. D., & Patterson, R. J. 2004, *ApJ*, **615**, 732
- Rockosi, C. M., et al. 2002, *AJ*, **124**, 349
- Ryan, S. G., & Norris, J. E. 1991a, *AJ*, **101**, 1835
- Ryan, S. G., & Norris, J. E. 1991b, *AJ*, **101**, 1865
- Seabroke, G. M., et al. 2008, *MNRAS*, **384**, 11
- Searle, L., & Zinn, R. 1978, *ApJ*, **225**, 357
- Sirko, E., et al. 2004a, *AJ*, **127**, 899
- Sirko, E., et al. 2004b, *AJ*, **127**, 914
- Smith, J. A., et al. 2002, *AJ*, **123**, 2121
- Smith, M. C., et al. 2009, *ApJ*, in press (arXiv:0904.1012)
- Sommer-Larsen, J., Beers, T. C., Flynn, C., Wilhelm, R., & Christensen, P. R. 1997, *ApJ*, **481**, 775
- Springel, V., et al. 2005, *Nature*, **435**, 629
- Springel, V., et al. 2008, *MNRAS*, **391**, 1685
- Starkenburg, E., et al. 2009, *ApJ*, **698**, 567
- Steinmetz, M., et al. 2006, *AJ*, **132**, 1645

- Stewart, K. R., Bullock, J. S., Wechsler, R. H., Maller, A. H., & Zentner, A. R. 2008, *ApJ*, 683, 597
- Totten, E. J., & Irwin, M. J. 1998, *MNRAS*, 294, 1
- Totten, E. J., Irwin, M. J., & Whitelock, P. A. 2000, *MNRAS*, 314, 630
- Tucker, D. L., et al. 2006, *Astron. Nachr.*, 327, 821
- Vivas, A. K., & Zinn, R. 2006, *AJ*, 132, 714
- Vivas, A. K., et al. 2001, *ApJ*, 554, L33
- Watkins, L. L., et al. 2009, *MNRAS*, in press (arXiv:0906.0498)
- Wheeler, J. C., Sneden, C., & Truran, J. W., Jr. 1989, *ARA&A*, 27, 279
- White, S. D. M., & Rees, M. J. 1978, *MNRAS*, 183, 341
- Willett, B. A., Newberg, H. J., Zhang, H., Yanny, B., & Beers, T. C. 2009, *ApJ*, 697, 207
- Xue, X. X., et al. 2008, *ApJ*, 684, 1143
- Yanny, B., et al. 2000, *ApJ*, 540, 825
- Yanny, B., et al. 2003, *ApJ*, 588, 824
- Yanny, B., et al. 2009, *AJ*, 137, 4377
- York, D. G., et al. 2000, *AJ*, 120, 1579
- Zinn, R. 1985, *ApJ*, 293, 424
- Zolotov, A., Willman, B., Brooks, A. M., Governato, F., Brook, C. B., Hogg, D. W., Quinn, T., & Stinson, G. 2009, *ApJ*, 702, 1058
- Zwitter, T., et al. 2008, *AJ*, 136, 421