The Arecibo Legacy Fast ALFA (ALFALFA) Survey
Skeptical Review Report
17 Nov 2010

M. Haynes and R. Giovanelli for the ALFALFA Collaboration

This document is intended to provide an update to the information provided in previous reports. See Appendix A for a list of websites containing relevant ALFALFA documentation.

ALFALFA overview and status:

As a local Universe ($z < 0.06$) extragalactic spectral line survey, ALFALFA (Arecibo program A2010) aims to cover ~ 7000 deg^2 of the high galactic latitude sky within the Arecibo telescope declination limits $0° <$ Dec. $< +36°$. Survey definition, strategy and goals are described in Giovanelli *et al.* (2005). As a second generation "blind" HI survey, ALFALFA is designed to improve dramatically on the predecessor HI Parkes All-Sky Survey (HIPASS; Barnes *et al.* 2001; Meyer *et al.* 2004; Wong *et al.* 2006). Exploiting the large collecting area of the 305 m antenna and its relatively small beam size, the ALFALFA survey is 8× more sensitive than HIPASS with $\sim 4\times$ better angular resolution. Furthermore, its backend spectrometer provides 3× better spectral resolution (5.3 km s^{-1} at $z = 0$) over 1.6× more bandwidth. A "minimal intrusion" drift scan observing technique yields optimal "open shutter" efficiency ($> 97\%$), baseline characteristics, flux calibration and HI signal verification. Data are flagged for RFI but retained so that memory of what channels have been discarded can inform whether a source is validly included in the dataset. Sources are identified via an automated matched filtering technique (Saintonge 2007b) but final parameters, including identification of the most probably optical counterpart, its characteristics and cross-identifications, are measured individually. In all respects, the ALFALFA survey is meeting, or exceeding, its design objectives.

Figure 1 summarizes the 2010 harvest of ALFALFA sources via a Spänhauer diagram of the current ALFALFA catalog, covering $\sim 40\%$ of the final survey area ($\sim$2900 deg^2) and thus dubbed "α.40", with superposed lines illustrating the HIPASS completeness and detection limits. With a dwell time of $t_{int} \sim 40$ sec per beam, the sensitivity of ALFALFA allows detection of 2×10^7 M$_\odot$ at the distance of the Virgo cluster and better than 10^5 M$_\odot$ within the Local Group. Since the distance a given M_{HI} is detectable increases only as $t^{1/4}$, where t is the effective integration time, the ALFALFA strategy maximizes sky coverage while still yielding almost an order of magnitude improvement in sensitivity over HIPASS. More than 2/3 of ALFALFA sources have never been previously observed in the 21 cm HI line, illustrating that the criteria for the selection of targets for past HI observations, usually based on optical parameters, missed the majority of HI rich systems. Moreover, ALFALFA detects a rich population of extremely massive HI disks which HIPASS severely undercounted. **Expectations for the detection of HI at higher redshift** by the Square Kilometre Array (SKA) and its pathfinders based on HIPASS **can now be revised upwards**.

As evident in Figure 1, ALFALFA samples much more deeply than did HIPASS; the median cz of detections is 9000 km s^{-1}, 3× higher than HIPASS, yielding **for the first time**, a sampling of the HI population over a **cosmologically fair volume**. α.40 already detects hundreds of sources with $M_{HI} < 10^8$ M$_\odot$ and thousands with $M_{HI} > 10^{10}$ M$_\odot$, with a detection density of $\sim$5.5 sources deg^{-2}, in comparison with the ~ 0.15 sources deg^{-2} of HIPASS. Most importantly, centroiding of ALFALFA HI sources is typically attained with an accuracy of $\sim 20''$, allowing unconfused identification of the optical counterpart for $> 95\%$ of the detected extragalactic sources (in a few percent, confusion due to close galaxy pairs remains unresolved; <2% are optically "dark").

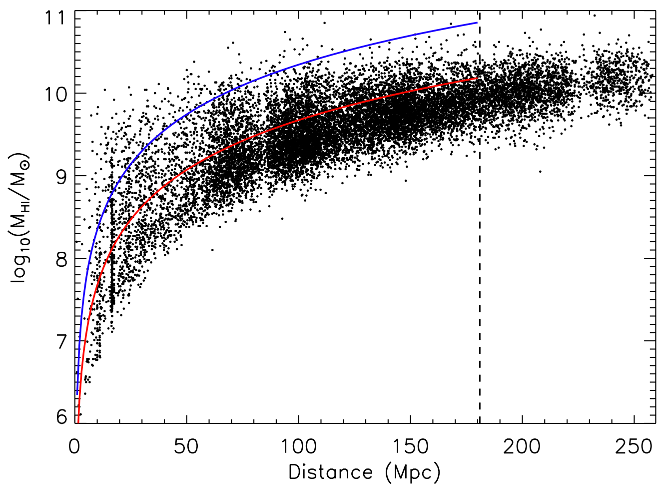

Figure 1: Spänhauer diagram of detected ALFALFA sources contained in the current ALFALFA catalog (α.40). For comparison, the HIPASS completeness limit (upper, blue line) and its absolute detection limit (lower, red line) are overplotted. The vertical dashed line marks the HIPASS low frequency bandpass edge. The "detection gap" at $cz \sim 15000$ km s^{-1} is caused by RFI from the 1350 MHz FAA radar at the San Juan airport; contaminated data are flagged and dropped so that the final 3-d data products also include a "weights" cube, thereby enabling spectral stacking analysis. Note that ALFALFA remains sensitive even at its low frequency ($z \sim .06$) bandpass limit.

Status of observing program: Figure 2 provides a graphical summary of the sky coverage of the observing program A2010 to date. Observations began in Feb 2005 and were proposed to require ∼5 years. In the practical context of time allocation at a widely used, multidisciplinary national facility like Arecibo, there have been significant delays: e.g., the telescope was out of service for all of the fall 2007 season; ALFA was serviced for use by another program for 6 weeks in fall 2009; structural issues resulted in 5 weeks down time in spring 2010 and afterward, have restricted the zenith angle movement of the Gregorian dome. As of Nov 15, 2010, 692 observing sessions have been successfully conducted, amounting to about 90% of the total required allocation of 4400 hours.

Observing team responsibilities: Each ALFALFA observing session is staffed by a member of the collaboration, mostly by remote access, who oversees and monitors the acquisition of data for ALFALFA. Immediately following each observing session, this "designated observer" posts relevant log and schedule files and converts the FITS data to IDL format, thereby verifying file integrity. In practice, the ALFALFA team provides regular feedback to the NAIC staff on receiver performance, hardware and software issues, and radio frequency interference (RFI). In September 2010, a quick complaint about excessive GPS-NUDET testing resulted in the folks at Space Command changing their test schedule (perhaps for the first time in 25 years!) to avoid contaminating A2010 data acquisition. We make available our detailed observing logs to the NAIC staff and the commensal galactic HI TOGS ("Turn-on-the-GALFA-Spectrometer"; M. Putman, P.I.) survey. The ALFALFA team bears the full observing burden for the TOGS program, including running the necessary calibration scripts. In fall 2010, the ALFALFA team agreed to accept the allocation of extra time of benefit to TOGS (1.25 hours earlier than the start of A2010); then too the ALFALFA team carried out the observations on behalf of the TOGS team.

During 2010, most of the A2010 observing was conducted when the zenith angle of the Gregorian dome was restricted. By positioning the feed arm off the meridian and rotating ALFA appropriately, we were able to conduct ALFALFA observations in the declination zone close to the Arecibo zenith. The modification was sometimes complicated by CIMA "features" (e.g., where the ZA is double

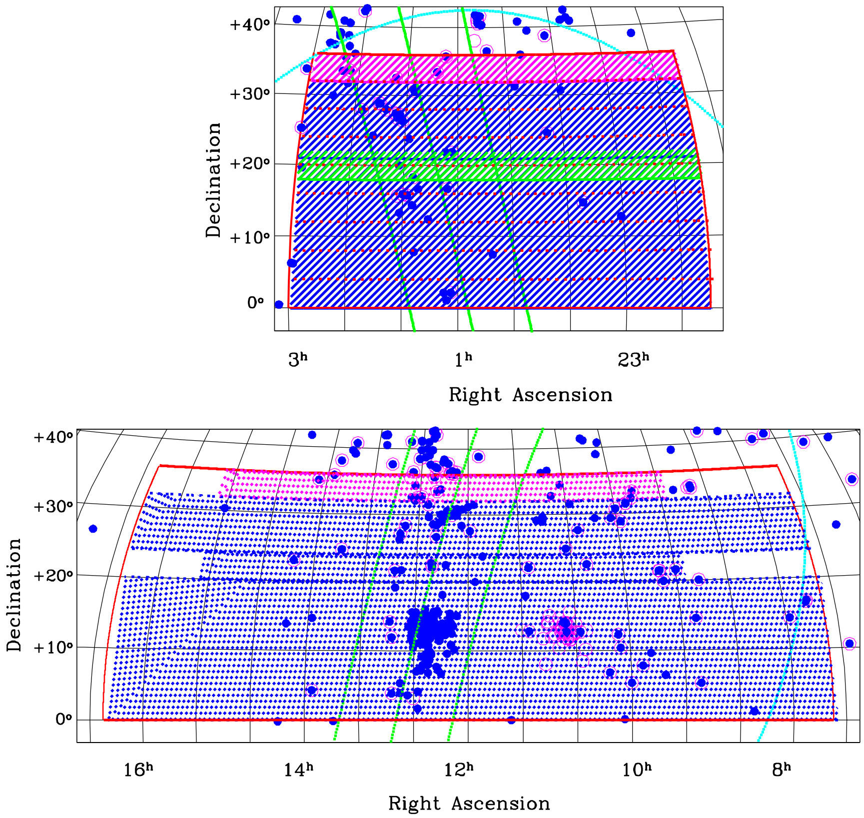

Figure 2: ALFALFA survey coverage as of 15 Nov 2010. The upper panel shows the fall sky region (anti-Virgo direction: 22^h < R.A. < 03^h; the lower one, the spring sky region (Virgo direction: 07^h30^m < R.A. < 16^h30^m; as requested; actually spring allocations have been restricted because of the conflict with the AUDS program). Blue hatched areas have been fully observed; green areas are begun but not complete; magenta shading indicates regions not yet covered. The cyan lines mark galactic latitudes b = −20° (fall sky) and +20° (spring sky) while the green lines trace SGL = −10°, 0° and +10°. We request a time allocation which allows us to complete the spring observations in 2011, but understand that it may not be possible to complete all of the required Fall observations until 2012.

valued for the same AZ, as can happen near zenith, CIMA always chooses to go to the lower ZA even when commanded to go to the higher value) and required, at times, real-time modification of the commanded position (when the telescope did not behave as expected), adding additional complication that required expert attention. However, we were happy to make good use of the telescope time, even under what were sometimes stressful observing conditions.

Collaboration: The ALFALFA project is organized as an open collaboration which continues to grow and now has more than 80 participants. New members continue to join the team "free of charge" to carry out studies which utilize any extant data sets, provided that they do not interfere with previously approved projects. Particular attention is given to protecting student projects.

Graduate student involvement: 5 Ph.D. dissertations have been based on ALFALFA data (Saintonge 2007a; Kent 2008; Stierwalt 2009; Toribio 2010; Dowell 2010). More than ten other graduate students from Cornell, Drexel, George Mason, MPIA, Tel Aviv, U North Carolina, U Washington and U Wisconsin are actively engaged in ALFALFA–based thesis research. A number of the more junior students do not have a significant background in radio astronomy, so we will hold a workshop on ALFALFA-specific radio astronomy techniques at Arecibo in February 2011.

Spin-off surveys: Several major spin-off projects have been started, at Arecibo and elsewhere.

Major programs based on ALFALFA include a GALEX survey of the Virgo Cluster (GUViCS; A. Boselli. P.I.), a WIYN/KPNO Hα survey to establish the star formation rate density at the present epoch (J. Salzer, P.I.), a San Pedro Martir Hα survey to map the star formation rate in the Coma supercluster (G. Gavazi and H. Hernandez Toledo P.I.s) and an HI synthesis mapping program with the EVLA of very low mass galaxies (J. Cannon, P.I.). ALFALFA is also contributing to the REsolved Spectroscopy Of a Local VolumE survey (RESOLVE; S. Kannappan P.I.). Of particular note, the GALEX-Arecibo-SDSS Survey (GASS; D. Schiminovich P.I.) is a targeted L-band wide program (A2335) which has itself spawned a large program on the IRAM 30 m telescope (COLD GASS; G. Kauffmann P.I.). The GASS team taps the ALFALFA dataset both to provide spectra and catalogued properties when ALFALFA detects the GASS target and to apply stacking techniques (see below) to measure ensemble averages to considerably lower thresholds.

Data processing: Processing of ALFALFA data is carried out within the IDL environment with software developed at Cornell under RG's lead by RG, MH and their students. The ALFALFA software package and its associated documentation has been freely exported to more than 27 external sites and is used regularly by other members of the ALFALFA collaboration. Calibration, bandpass subtraction and data quality inspection is performed normally within 24 hours of data acquisition; RFI flagging of 2–d datasets is done of contiguous datasets (for data needed to generate the 3–d grids) within months thereafter. The ALFALFA pipeline yields spectral data cubes and continuum maps, "bad pixel/weights" 3–d maps (where the bad pixels result most notably from man-made RFI), source spectra and catalogs. Tools also exist for cross identification with other datasets, beam cleaning and moment map production.

Data status: Five catalog releases are in the public domain (Giovanelli *et al.* 2007; Kent *et al.* 2007; Saintonge *et al.* 2008; Martin *et al.* 2009; Stierwalt *et al.* 2009). Data releases are made available on a VO-registered website upon publication; interim catalogs, datasets and followup observation target lists are made available to members of the collaboration for their use in advance of publication. For the last 1.5 years, the ALFALFA α.40 dataset covering about 2900 deg^2 has been used by the "distributed" ALFALFA team to pursue a number of science goals. The α.40 catalog contains HI masses, systemic velocities, HI line widths and, where relevant, optical identifications for $\sim$16000 detected sources and their respective cataloged properties. Because catalogs are released only after full completion in R.A. of 2°-wide Declination strips, the α.40 dataset is not representative of the current degree of completion of survey data processing; $\sim$ 60% is ready for the final stage of source extraction and measurement.

During 2010, we have delayed full public release of the α.40 dataset for several reasons: (1) We wanted to protect several graduate student projects from potentially being scooped. Particularly because of the delay in the ALFALFA observing program, the more advanced students have needed to focus on science rather than data reduction. (2) Based on experience with the early ALFALFA datasets, we revised the photometric scale and the algorithm for flux measurement. The updated version will fully replace the older datasets. (3) Because of the migration to new hardware at the Cornell Center for Advanced Computing (CAC), the dataset must be regenerated in its entirety to accommodate version changes in the server software; and (4) We wanted to add the detailed SDSS cross-identification (see below), a step which has required significant protocol development and testing, and which had to be retroactively applied to the earlier datasets. Furthermore, the CAC server has also been experiencing irregularities associated with maintenance and software issues which have not been fully diagnosed by the experts; we continue to consult with Adam Brazier. *It should be noted that, as a fully open collaboration, we encourage parties interested in pre-release ALFALFA catalogs to join the ALFALFA collaboration. Team members have full access to the pre-publication catalogs and datasets, but we are thus better able to protect proposed student projects and to keep users fully aware of dataset status and best usage practices.*

Data cross-reference: The public-access ALFALFA database is inherently multiwavelength, including direct links via VO tools to SDSS and DSS2 images. Because the cross referencing is done as a part of the ALFALFA catalog construction process, it is more robust than would be a blind positional cross match with the SDSS. Of particular note, we have added a step to the reduction procedure which adds to the ALFALFA catalog the SDSS photometric and spectroscopic identifiers (PhotoObjID and SpecObjID), as well as a comment on issues associated with the crossmatch (e.g., multiple photometric sources within a single galaxy, offsets between the spectroscopic target and the photometric center of light, location outside the SDSS area, contamination by nearby stars, background superpositions, etc). Additionally a set of notes regarding blends, confused sources and other features is included with the published source catalogs. Provision of such ancillary information adds non-trivially to the work required to produce a source catalog; however, we believe that the multiwavelength cross-linkage is vital to ALFALFA's legacy value in the broad extragalactic community.

Data stacking tools: The tracking of RFI and data quality weighting makes possible stacking of spectra to explore the statistical characteristics of an ensemble population at greater sensitivity. During two extended visits at Cornell, graduate student Silvia Fabello (MPA-Garching) developed a tool to stack the ALFALFA spectral images centered on an arbitrary set of extragalactic targets, allowing her to derive very sensitive ensemble averaged gas fractions for the early type galaxies included in GASS (Fabello *et al.* 2010), discussed further below.

Deconvolution tools: Indiana U. graduate student Jayce Dowell (now at U New Mexico) has developed a tool within the ALFALFA IDL package which performs a first order deconvolution of extended sources. The core of this tool attempts to model the complex beam pattern that arises when ALFA drift scans are combined to produce gridded data cubes. The scientific product of this work, which makes up a substantial part of his Ph.D. thesis (Dowell 2010) is the measurement of the HI diameter function, discussed further below (Dowell *et al.*, in preparation).

Testing tomography techniques: Future observations of the large scale distribution of HI at high redshift with SKA pathfinders including the HERA program endorsed by Astro 2010 promise to track the transition from the Dark Ages through the Epoch of Reionization and provide constraints on cosmological dark energy models. Such experiments are technically difficult as the cosmic signals are much weaker than the foregrounds and calibration is challenging. The ALFALFA dataset is being used by graduate student Alexander Fry (U. Washington) to test techniques for foreground removal and to probe large scale spatial density variations of HI to prepare for future HI intensity mapping experiments. A first result will be presented at the Seattle AAS meeting (Fry *et al.* 2011).

Publications: To date, 33 papers based on ALFALFA data have been submitted or published in the refereed literature; for a complete list see: *http://egg.astro.cornell.edu/alfalfa/pubs.php.* Here we list those ALFALFA-based papers which have been published in or submitted since Aug 1, 2009.

1. "*COLD GASS, an IRAM Legacy Survey of Molecular Gas in Massive Galaxies: I. Relations between H_2, HI, Stellar Content and Structural Properties*"
Saintonge, A., Kauffmann, G., Kramer, C., Tacconi, L.J., Buchbender, C., Catinella, B., Gracia-Carpio, J., Wang, J., Cortese, L., Fabello, S., Fu, J., Genzel, R., Giovanelli, R., Gui, Q., Haynes, M.P., Heckman, T.M., Krumholz, M.R., Li, Cheng, Moran, S., Rodriguez-Fernandez, N., Schiminovich, D., Schuster, K. & Sievers, A. 2010, MNRAS (submitted)
2. "*Clouds Toward the Virgo Cluster Periphery: Gas-Rich Optically Inert Galaxies*"
Kent, B.R. 2010, Astrophys. J. (in press)
3. "*HI Content and Optical Properties of Field Galaxies from the ALFALFA Survey. I. Selection of a Control Sample*"
Toribio, M.C., Solanes, J.M., Giovanelli, R., Haynes, M.P. & Masters, K.L. 2010, ApJ (submitted)
4. "*The GALEX Arecibo SDSS survey: III. Evidence for Inside-Out Formation of Galactic*

Disks"
Wang, J., Kauffmann, G., Overzier, R., Catinella, B., Schiminovich, D., Heckman, T.M., Moran, S.M., Haynes, M.P., Giovanelli, R. & Kong, X. 2010, MNRAS (in press)

5. *"Galaxies with Wide HI Profiles"*
Brosch, N., Spector, O. & Zitrin, A. 2010, MNRAS (submitted)

6. *"The ALFALFA HI Absorption Pilot Survey: A Wide-Area Blind Damped Lyman Alpha System Survey of the Local Universe"*
Darling, J., Macdonald, E.P., Haynes, M.P. & Giovanelli, R. 2010, ApJ (submitted)

7. *"ALFALFA HI Data Stacking I: Does the Bulge Quench Ongoing Star Formation in Early-Type Galaxies?"*
Fabello, S., Catinella, B., Giovanelli, R., Kauffmann, G., Haynes, M.P., Heckman, T.M. & Schiminovich, D. 2010, MNRAS (in press)

8. *"The Arecibo Legacy Fast ALFA Survey: X. The HI Mass Function and HI from the 40% ALFALFA Survey"*
Martin, A.M., Papastergis, E., Giovanelli, R., Haynes, M.P., Springob, C.M. & Stierwalt, S. 2010, Ap.J. (in press)

9. *"The GALEX Arecibo SDSS Survey II: The Star Formation Efficiency of Massive Galaxies"*
Schiminovich, D., Catinella, B., Kauffmann, G., Fabello, S., Wang, J., Hummels, C., Lemonias, J. Moran, S.M. Wu, R., Giovanelli, R., Haynes, M.P., Heckman, T.M., Basu-Zych, A.R., Blanton, M.R. Brinchmann, J. Budavari, T., Goncalves, T., Johnson, B.D., Kennicutt, R.C., Madore, B.F., Martin, C.D., Rich, M.R., Tacconi, L.J., Thilker, D.A., Wild, V. & Wyder, T.K. 2010, MNRAS 408, 919

10. *"Multiwavelength Properties of Barred Galaxies in the Local Universe: I: Virgo Cluster"*
Giordano, L., Tran, K., Moore, B. & Saintonge, A. 2010, ApJ (submitted)

11. *"Diffuse Far-infrared and Ultraviolet Emission in the NGC4435/4438 System: Tidal Stream or Galactic cirrus?"*
Cortese, L., Bendo, G.J., Isaak, K.G., Davies, J.I., & Kent, B.R. 2010, MNRAS 2010 MNRAS 403, L26

12. *"The GALEX Arecibo SDSS Survey. I. Gas Fraction Scaling Relations of Massive Galaxies and First Data Release"*
Catinella, B., Schiminovich, D., Kauffmann, G., Fabello, S., Wang, J., Hummels, C., Lemonias, J., Moran, S.M., Wu, R., Giovanelli, R., Haynes, M.P., Heckman, T.M., Basu-Zych, A.R., Blanton, M.R., Brinchmann, J., Budavri, T., Goncalves, T., Johnson, B.D., Kennicutt, R.C., Madore, B.F., Martin, C.D., Rich, M.R., Tacconi, L.J., Thilker, D.A., Wild, V. & Wyder, T.K. 2010, MNRAS 403, 683

13. *"Are Newly Discovered High Velocity Clouds Minihalos in the Local Group?"*
Giovanelli, R., Haynes, M.P., Kent, B.R., & Adams, E.A.K. 2009, ApJL 708, L22

It should be noted that the overlap with the GASS project occurs because ALFALFA provides GASS with detections for 30% of its parent sample and the two teams work closely together towards mutual science goals.

Selected ALFALFA science highlights of the last year:

Here we highlight some of the ALFALFA science harvest since the last annual report was submitted in August 2009.

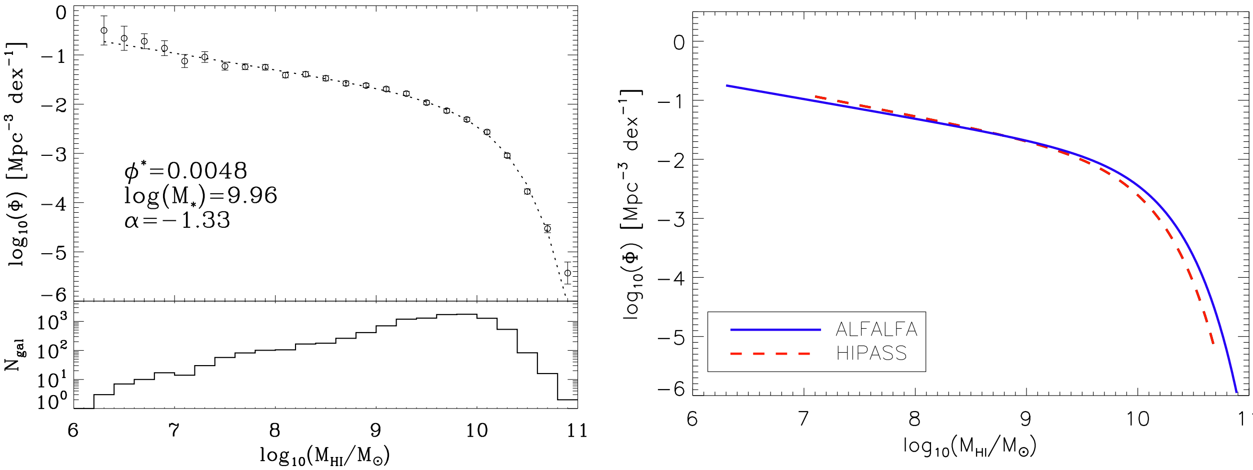

Figure 3: Left: HI mass function of α.40 galaxies from Martin *et al.* (2010). The dotted line is the Schechter function fit to the points using the 2-d stepwise maximum likelihood method; the same fit is shown as the blue solid line in the right panel. The lower histogram shows the logarithm of the number of detections in each bin used to derive the HIMF. Right: Comparison of the best fit HIMF for HIPASS (red dashed line; from Zwaan *et al.* 2005) and α.40 (blue solid line which is shown as the dotted line in the right panel). The HIPASS fit is traced only over the mass range sampled by that survey. Note that error bars are not shown but are significant for HIPASS at both the low and high HI mass ends.

The Distribution and Abundance of Gas-Rich Halos at $z \sim 0$: Our current understanding of large scale structure and the galaxy population traced by HI redshift surveys is distinctly immature relative to that derived from O/IR ones. ALFALFA is designed to provide a factor of ten improvement in the volume sampling over HIPASS, important not only for taking proper account of the impact of large scale structure and distance uncertainties among the nearest objects (Masters *et al.* 2004) but also for separating the impacts of morphological segregation from the morphological variation in HI content (Roberts & Haynes 1994). Prime objectives are to derive the HI mass function (HIMF) and its integral, the cosmic density of HI Ω_{HI}, the HI correlation function ξ_{HI}, and the gas-rich halo velocity function (VF). The current α.40 catalog samples a sufficiently large volume that first attempts are being made to determine these cosmologically relevant parameters.

The HI Mass Function: Figure 3 shows the recent determination of the HI mass function (HIMF) from a spatially contiguous subset of the current detections (Martin *et al.* 2010). Although incorporating only 35% of the final survey data, this determination clearly extends to lower HI masses than any previous one: the number of objects with $M_{HI} < 3 \times 10^7$ M$_\odot$ is 10$\times$ higher than included in the HIPASS sample (Zwaan *et al.* 2003; 2005). Martin *et al.* (2010) have carefully tested the robustness of different methods used to derive the HIMF by performing a comparison of results from a large scale structure corrected 1/V$_{max}$ analysis with those obtained from an approach using the structure-insensitive two-dimensional stepwise maximum likelihood (2DSML) method. The right panel of Figure 3 shows a comparison of the HIMF functions derived by HIPASS and ALFALFA; the HIPASS result is drawn only over the range of HI masses sampled by that survey. While the low mass slopes derived by the two surveys are comparable, the ALFALFA determination is on much more solid statistical ground. Martin *et al.* (2010) measure a cosmic HI abundance $\Omega_{HI} = 4.3 \pm 0.3 \times 10^{-4}$ h$_{70}^{-1}$, 16% larger than the 2005 HIPASS result (Zwaan *et al.* 2005), and predict an order of magnitude higher space density of the most massive HI systems. Note also that the AUDS precursor survey (Freudling *et al.* 2010) is also **severely limited** by statistics: it samples the HI mass function only over the range $2 \times 10^9 < M_{HI} < 10^{10}$ M$_\odot$. Despite its intended survey depth, AUDS has so far not uncovered the important ALFALFA result on the favorable cosmic abundance of high HI mass galaxies. **Because massive HI disks are rare, surveys that do not sample enough volume miss this important high mass population and therefore undercount their abundance in the present universe.**

Fitting the low M_{HI} slope of the HIMF requires a sample covering enough solid angle to encompass adequate local volume; ALFALFA is specifically designed to detect the gas-rich population of low mass halos, often the metal-poor "late bloomers" of cosmic star formation history. Mimicking an effect already known for luminosity functions based on optical catalogs, the low M_{HI} slope of the HIMF, $\alpha = -1.33$, is much shallower than the corresponding slope expected by ΛCDM for the low mass end of the halo mass function. Thus, the paucity of dwarf galaxies apparent in optical catalogs is not fully made up by HI rich but optically tenuous systems. It is however of interest that the low mass slope of the HIMF becomes steeper in HI-rich moderate density environments like the Leo I group (Stierwalt *et al.* 2009). With the availability of the full ALFALFA dataset, we will be able to compare results for HI-selected subsamples subdivided both by individual galaxy properties and by quantitative measures of local galaxy density to explore how the contribution of HI to the cosmic density may vary across the full range of extragalactic environments sampled by ALFALFA.

The HI Correlation Function: Like HIPASS, other surveys which do not sample adequate volume will be limited by statistics at the low and high mass ends and by cosmic variance. For example, inadequate volume sampling is the likely cause of the discrepant conclusions on ξ_{HI} for the HIPASS catalog reported by Meyer *et al.* (2007) and by Basilakos *et al.* (2007) who used different methods to interpret the HIPASS dataset in light of large scale structure. Graduate student Ann Martin is currently performing the first derivation of the ξ_{HI} from the α.40 catalog. In the future, this work will be fully developed and extended to provide a robust measurement over the full ALFALFA volume, thereby serving as the $z = 0$ benchmark for future studies of the evolution of the clustering properties of HI-rich galaxies at higher redshift.

The Velocity Function of Gas-Rich Halos: ΛCDM predicts that substructure forms first on small scales, resulting in a large population of low mass halos. The distribution of dark matter (DM) halos, described by the DM mass function (MF), is predicted to display a power law behavior at low masses, $N(M) \propto M^{\alpha}$, with $\alpha \simeq -1.8$ (Press & Schechter 1974). The MF can be extracted from simulations; its observational proxy, the VF, can be measured from spectroscopic surveys. While early-types require absorption line measures of velocity dispersion, the VF for late-type galaxies can be extracted for large samples from HI line surveys like ALFALFA. To date, the most accurate VF for gas-bearing galaxies has been derived from a subsample of 2646 HIPASS sources (Zwaan *et al.* 2010). As in the case of HI masses, the dynamic range in the HI line width sampled by ALFALFA is significantly greater than that of HIPASS at both the low and wide width ends. A preliminary α.40 result led by graduate student Manolis Papastergis measures the VF down to an observed velocity width of 20 km s^{-1}, a factor of 2 lower than the HIPASS VF. The preliminary α.40 result yields a much shallower slope at small widths than expected from the MF in simulations; it also finds a much higher density of halos with high rotational velocity than did HIPASS. Future work will clarify the relationship of the observed maximum rotational velocity to the halo circular velocity and its impact on the observationally derived halo masses as well as probe the current mismatch between observations and simulations for the gas-rich population at $z \sim 0$.

The HI Diameter Function: Characterizing the HI diameter function (HIDF) of galaxies is important not only for our understanding of galaxy formation but also to explore the connection between galaxies in the local universe and the damped Lyman α systems seen along the line of sight to quasars. As part of his Ph.D. thesis, Jayce Dowell has used the ALFALFA survey data for 304 galaxies resolved by the ALFA beam to characterize the HIDF in the local universe. In order to measure their HI diameters, Dowell developed a tool to deconvolve the maps using a CLEAN-like algorithm specifically geared to ALFA-based data. As a test of this process, Figure 4 shows a

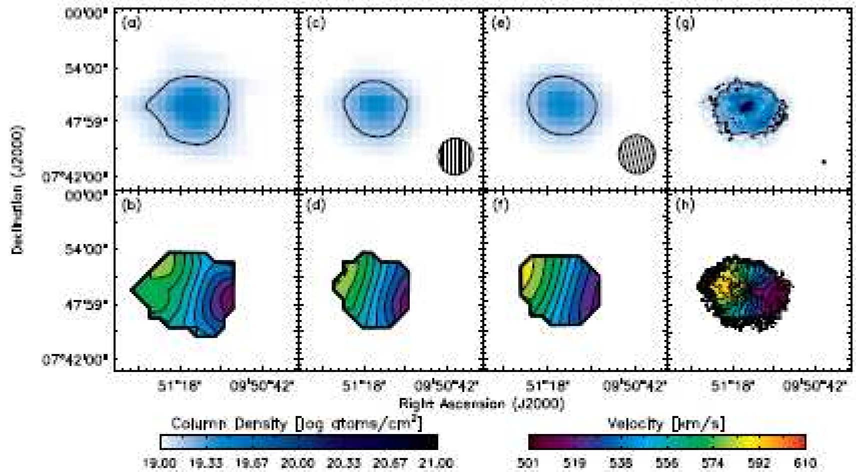

Figure 4: Comparison among the column density (upper panel series) and velocity field maps (lower panel series) for the dwarf galaxy UGC 5288, known to have a very extended HI envelope. The raw ALFALFA maps are shown in panels (a) and (b), with the deconvolved ALFALFA maps in panels (c) and (d). In addition, VLA-C data from program AZ108 are shown at a spatial resolution matching that of ALFALFA in panels (e) and (f). The full resolution VLA maps are shown in panels (g) and (h). For the deconvolved ALFALFA maps, the beam size is $4.0' \times 3.7'$. The beam size of the full resolution VLA maps is $20.7'' \times 17.5''$. From Dowell (2010).

comparison of moment maps extracted from ALFALFA and ones obtained from observations in the VLA archive for the nearby dwarf galaxy UGC 5288, known previously to have a very extended HI envelope. The deconvolved maps derived from ALFALFA are in good agreement with the resolution-matched VLA ones, both in terms of the size of the galaxy at a column density of 10^{20} atoms cm^{-2} and in the overall shape and spread of the velocity field. A more detailed comparison of the column density maps in panels (c) and (e) does not show any significant differences until column densities of about 5×10^{18} atoms cm^{-2}.

The most robust measures of the HI diameter function to date have been based on the optically-selected "Westerbork Observations of Neutral Hydrogen in Irregular and Spiral Galaxies" survey (WHISP: van der Hulst 2001). Using optical selection criteria biases such studies towards objects near the peak of the distribution functions, i.e., the most "normal" galaxies, but it does not provide insight into the extremes of the distribution in the HI sense. ALFALFA allows investigation of the full range of HI distributions, without consideration of optical properties. Dowell (2010; also Dowell *et al.* in prep.) presents the first ALFALFA based HIDF, based on 33% of the expected final catalog. The HIDF is well approximated by a Schechter function that includes a geometric conversion from mass to diameter with a power-law slope of ~ 1.3 and a characteristic HI diameter D* of ~ 60 kpc. 95% of isolated galaxies have impact parameters < 30 kpc for column densities of 10^{20} atoms cm^{-2}. In general, galaxies exist in large gas halos that are roughly two to three times the size of their high column density (log $N_{HI} > 10^{20}$ atoms cm^{-2}) disks. Further analysis of the full ALFALFA dataset will provide an avenue for comparison of the radial distribution of neutral gas distributions from simulations of galaxy formation.

The HI Gas Fraction in Early Type Galaxies: A major problem that has plagued our understanding of the cold gas in early-type galaxies is that the available HI data have been inhomogeneous and/or limited in sensitivity to the gas-poor populations. Although ALFALFA is a relatively shallow survey, it is possible to use its high quality dataset to constrain the statistical properties of a population that lack individual detections in the survey. The stacking process basi-

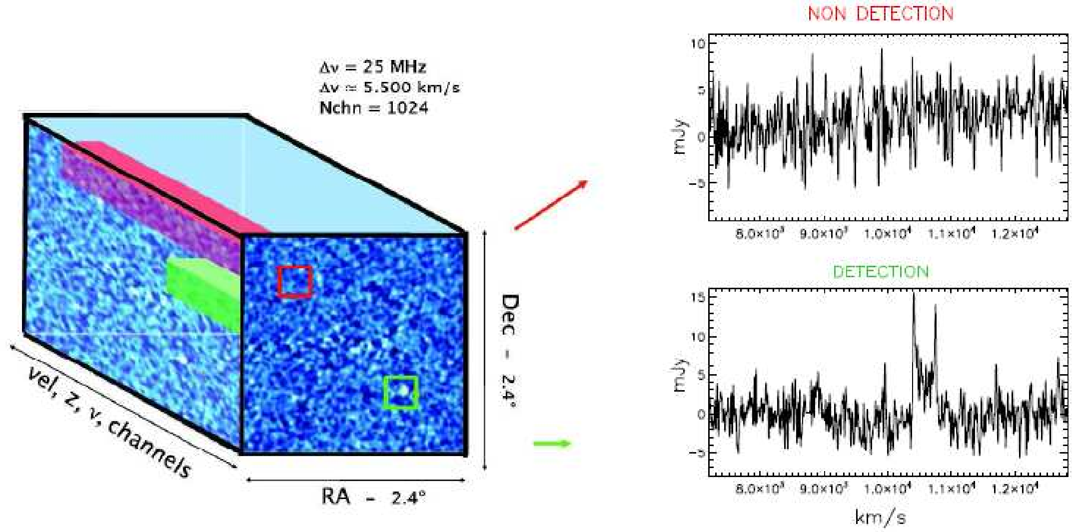

Figure 5: A schematic representation of the fully processed ALFALFA 3-d data-cube. The data cubes are $2.4° \times 2.4°$ in size and about 5500 km s^{-1} in velocity range (25 MHz in frequency). The raw spectral resolution is ~ 5.5 km s^{-1}; the angular resolution is $\sim 4'$. For each pixel, which is a point in RA, Dec and velocity, a value of flux density is recorded. For each target, a spectrum is extracted at a given position of the sky, over the velocity range of the data-cube which contains the source. Two examples of extracted spectra are shown on the right, illustrating a detection (green, bottom) and a non-detection (red,top). From Fabello *et al.* (2010).

cally consists of co-adding the signal from many objects with known sky-positions and redshifts so that the background noise is decreased. In this manner, the average HI flux of the ensemble can be recovered. Graduate student Silvia Fabello has developed a tool that allows the stacking of spectra extracted from the ALFALFA grids. The availability of the 3-d weights grid associated with each spectral cube is critical so that objects whose spectra are contaminated can be discarded and others with poor coverage or missing data can be assigned proper weight in the resultant average. Figure 5 illustrates the stacking technique. Fabello *et al.* (2010) have used this approach to examine a subsample of 1833 early type galaxies also included in the GASS sample. They find that the HI content of a galaxy is not directly influenced by its bulge, in contradiction with the "morphological quenching" scenario proposed by Martig *et al.* (2009).

Massive Gas-Rich but Optically-Underluminous Disks: The richness of the high HI mass galaxy population is perhaps the most surprising result of the ALFALFA survey. Because of its superior depth and sensitivity, ALFALFA sets on much firmer grounds the cosmic abundance of the rare, high HI mass systems which, because of Malmquist bias, are found preferentially in the more distant volumes. As evident in Figure 1, the α.40 catalog already includes thousands of galaxies (2947 in α.40) with $M_{HI} > 10^{10} M_{\odot}$; the full ALFALFA catalog will add thousands more to the count. This result is particularly important as it directly impacts, in a positive sense, estimates of the expected HI detection rate at high z with future facilities like the SKA and its pathfinders.

The widely accepted hierarchical scenario postulates that galaxies assemble their mass through mergers and the accretion of gas from the intergalactic medium (IGM). Over time, galaxies consume their gas supply as they form stars, and unless it is replenished, they presumably evolve from the blue cloud to the red sequence in the color-magnitude diagram (Baldry *et al.* 2004). In the "down-sizing" evolutionary scenario, the most massive galaxies are most efficient in consuming their gas reservoirs and, today, are red and gas-poor. Very massive galaxies with a substantial supply of cold gas are thus expected to be very rare. This expectation was corroborated by the HIMF derived

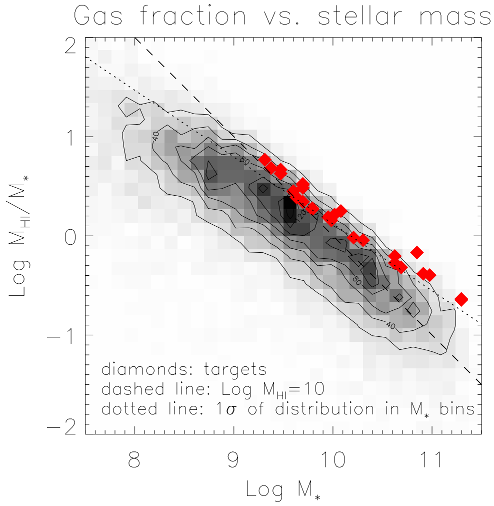

Figure 6: Gas fraction versus stellar mass for the full HI-selected ALFALFA-SDSS sample shown as greyscale with superposed contours of number density. Stellar masses have been calculated by SED-fitting the five SDSS bands (Huang *et al.* in prep.). The dashed line of slope -1 indicates a constant $M_{HI} = 10^{10} M_{\odot}$ while the dotted line denotes gas fractions 1σ above the mean for any given stellar mass bin. Diamonds are the high HI mass but optically underluminous targets for which we plan multiwavelength followup observations including GALEX UV and optical broadband R and $H\alpha$ imaging, and for a subsample, HI synthesis mapping.

from the HIPASS survey data (Zwaan *et al* 2005). However, the detection by ALFALFA of a substantial population of high mass gas-rich galaxies challenges expectations and suggests a target sample for a multiwavelength campaign to understand the origin and history of such massive HI disks.

Figure 6 shows the distribution of gas fraction (defined as M_{HI}/M_{*}, where M_{*} is the stellar mass) with increasing stellar mass for the current ALFALFA-SDSS common sample (11,938 galaxies). As part of her Ph.D. research, graduate student Shan Huang estimates stellar masses through a procedure of SED–fitting the five SDSS bands using a template which is generated following Salim *et al.* (2007; also used by the GASS MPA/JHU collaboration). Contours of number density are overplotted on the greyscale for the underlying distribution which, it should be underscored, contains only HI detections and thus reflects an already gas-rich population. The dashed line of slope -1 indicates a constant $M_{HI} = 10^{10}$ M$_{\odot}$. The general trend confirms the expectation that the gas mass fraction decreases with increasing stellar mass, i.e. massive galaxies are mostly gas poor. However, some of the ALFALFA galaxies (above the dashed line so that $M_{HI} > 10^{10}$ M$_{\odot}$) appear to contain an exceptional, and even in some cases, dominant fraction of their baryons in the form of neutral gas; at the same time, they do not meet the extreme low surface brightness criteria of objects like Malin 1.

The presence of a very large HI reservoir itself suggests an arrested stage of evolution, and we have begun a program to acquire observational evidence, via GALEX UV and optical imaging and HI synthesis mapping, that will confirm (or refute) the on-going accretion hypothesis and establish the evolutionary state of these enormous gas-rich, optically underluminous systems. An illustrative target for such a coordinated campaign is shown in Figure 7, a massive galaxy with a higher than expected gas fraction. Most of the star formation evident in the GALEX FUV image coincides with the outer disk, not the inner region; the HI disk is clearly more extended than the stellar distribution. Future GMRT, WSRT and EVLA synthesis observations will target the optically underluminous massive HI disk population. The HI synthesis maps will probe the relationship of the HI density to the locus of current star formation, and although sometimes too crude to yield a detailed rotation curve, the velocity fields will at least yield coarse grained information that will

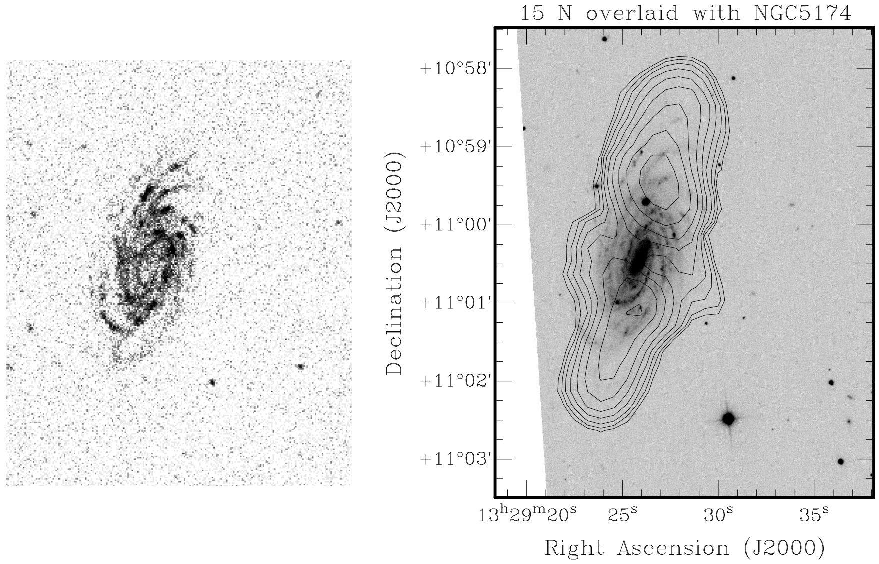

Figure 7: Left: GALEX FUV image of NGC 5174, one of the most luminous, largest and highest stellar mass objects in α.40, also exceptional for its massive HI disk. Right: Preliminary GMRT HI column density map of NGC 5174 superposed on an optical image, $6'$x$5'$. Through a collaboration with Jayaram Chengalur (NCRA), we already have GMRT maps for 3 high HI mass galaxies; graduate student Betsey Adams will observe five additional targets with the GMRT in January 2011.

help us understand the dynamics of the systems. The combination of optical SDSS photometry with newly acquired GALEX, R-band and Hα images will probe the stellar populations and star formation histories of these objects, particularly in comparison with the more common massive stellar systems targeted by GASS.

The Chase for Minihalos: In addition to providing a fair sample of the HI population over a cosmologically significant volume, ALFALFA is also designed to probe deeply the low mass end of the HIMF in the local universe. Although the ΛCDM paradigm predicts the existence of large numbers of low mass ($\leq 10^9$ M$_\odot$) halos, the cosmic census of dwarf galaxies at $z = 0$ indicates that such objects are rarer than expected from numerical simulations. While current optical identification techniques are inefficient in detecting satellites beyond the Milky Way virial radius ($\sim$ 250 kpc), the morphological segregation exhibited by the dwarf population (Grebel, Gallagher & Harbeck 2003) suggests that gas-rich dwarfs will be found broadly distributed throughout the Local Group. The recently discovered transitional dwarf galaxy Leo T is less than two virial radii away from the Milky Way, yet further than the gas-poor ultra faint dSph satellites. As a star forming galaxy with an HI mass of 2.8×10^5 M$_\odot$, an HI radius of 300 pc (at the $N_{HI} = 2 \times 10^{20}$ cm^{-2} level), an indicative dynamical mass *within the HI radius* of $\sim 3.3 \times 10^6$ M$_\odot$ and a stellar mass of $\sim 1.2 \times 10^5$ M$_\odot$ (Ryan-Weber *et al.* 2008), Leo T serves as an illustrative prototype (Ricotti 2009) of the kind of target for what we dub as our "chase for minihalos".

The precipitous drop in the cold baryon fraction for halos of mass $\leq 10^{10}$ M$_\odot$ proposed by the simulations by Hoeft *et al.* (2006) and the models of Sternberg, McKee & Wolfire (2002) can explain the shallow slopes of the low end of the optical luminosity and HI mass functions. It also suggests that the gas-bearing minihalo candidates have more extreme properties than previously thought: HI masses of $\leq 10^6$ M$_\odot$, sizes of ≤ 1 kpc and linewidths of ≤ 40 km s^{-1}. The requirements for detection of such sources are very challenging. **ALFALFA offers an opportunity to detect such HI minihalo candidates, albeit only to distances of a few Mpc**. ALFALFA detects 5×10^4 M$_\odot$ at $d_{Mpc} = 1$. Already, we (Giovanelli *et al.* 2009) have identified a first set of extremely

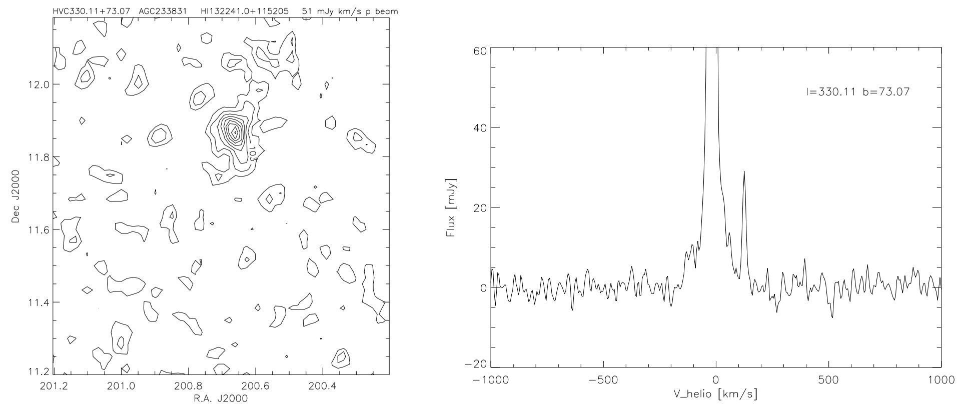

Figure 8: Left: HI cloud at $l = 330.11$, $b = +73.07$, $cz_\odot = 124$ km s^{-1}, $W = 19$ km s^{-1}, angular size of $6' \times 4'$, integrated flux of 0.60 Jy km s^{-1}, $M_{HI} = 1.4 \times 10^5 d^2_{Mpc}$ M$_\odot$, $M_{dyn}(< R_{HI}) \simeq 1.1 \times 10^7 d_{Mpc}$ M$_\odot$. Right: HI line spectrum of the cloud. The strong feature at zero velocity is the Milky Way emission; the cloud is the narrow feature at $cz_\odot = +124$ km s^{-1}. From Giovanelli *et al.* (2009)

compact clouds (some are unresolved) which are kinematically separate from the Galactic HI and spatially isolated; only $\sim$5-10$'$ in extent, these are "ultra compact HVCs" (UCHVCs). The mean properties of the clouds are as follows: HPFW linewidth $W = 34$ km s^{-1}; HI column density averaged within the half power isophote $N_{HI} = 1.3 \times 10^{19}$ cm^{-2}; if placed at $d_{Mpc} = 1$, HI radius of $R_{HI} = 0.75$ kpc; HI mass $M_{HI} = 1.8 \times 10^5$ M$_\odot$. Figure 8 shows an integrated flux map and a line profile of one of those detections (Giovanelli *et al.* 2009).

Graduate student Betsey Adams has modified the standard ALFALFA signal extractor developed by Saintonge (2007b) specifically to search for narrow-width, isolated and barely resolved (or unresolved) sources in and near the wings of the Galactic HI. To verify whether these HI minihalo candidates are indeed "galaxies", we have begun a program to conduct a deep optical search for stellar counterparts and star formation activity and to probe their velocity fields via EVLA HI synthesis mapping. Detection of a resolved stellar component would establish the distance and yield an estimate of the total baryon content of the HI minihalo candidates. If the UCHVCs are not minihalo candidates but rather, more local complexes associated with Galactic processes, they should show diffuse Hα emission as a result of ionizing photons escaping the disk of the Milky Way (e.g. Putman *et al.* 2003). A non-detection of Hα emission would thus place constraints on the lower limit for their distance. Higher resolution HI velocity fields may yield the signature of ordered rotation in the velocity field, thus providing a more reliable dynamical mass than allowed by the linewidth alone.

Broader impact with the Undergraduate ALFALFA Team:

In addition to its significant involvement of graduate students, the ALFALFA survey actively engages faculty and undergraduate students at some 15 principally undergraduate institutions, the **U**ndergraduate **A**LFALFA **T**eam (UAT), in a manner that prioritizes educational aspects, while preserving the excitement of cutting–edge research. Support from NSF funds activities of the UAT colleges and universities through grants (NSF AST-0724918, AST-0725267 and AST-090339) to R. Koopmann (Union), T. Balonek (Colgate) and S. Higdon (Georgia Southern). Workshops and telecons are organized to stimulate involvement, develop collaborative modes and discuss scientific achievements. UAT faculty members have been hosted at Cornell and Arecibo to instruct them in the use of processing software, observing procedures and data handling protocols. MH, RG and their graduate students play major roles in the engagement of the UAT, in selecting targets and providing datasets and software, and in the planning and presentation of educational materials at the annual UAT workshop in Arecibo. The 4th annual workshop is planned for January 17–19, 2011 in Arecibo.

While various members of the UAT are engaged in their own ALFALFA research projects, we have organized a collaborative team project involving faculty and students at 11 UAT schools to use ALFALFA and other public datasets to conduct a comparative study of the impact of local environment on galaxy evolution, the "UAT Groups of Galaxies project". Science projects undertaken at different UAT institutions are coordinated so that they can be pooled together across the UAT to enable a mix of individual and cooperative research experiences. These studies form the basis of numerous senior thesis/research projects while also nurturing the continued engagement of members of the UAT. The success of the UAT illustrates the efficacy of using the ALFALFA survey, its organization and followup science as a framework for engaging undergraduate students in research and training far beyond the boundaries of a single campus.

Requirements to complete ALFALFA:

To complete ALFALFA, we require 35 blocks of spring nighttime observing and 52 blocks of fall nightttime observing. We very much hope that the spring observing can be completed in 2011. We realize it may not be possible to complete the fall observing in 2011, but we request **more than 30** blocks in 2011 with the remainder slipping over to fall 2012. Bearing the observing burden requires a significant effort which we hope to redirect to data analysis after spring 2011.

The full ALFALFA observing blocks request nighttime observing as follows:

$$07^h10^m < \text{R.A.} < 16^h45^m \quad \text{in the spring}$$

and

$$21^h40^m < \text{R.A.} < 03^h15^m \quad \text{in the fall}$$

We understand that more sessions of shorter allocations may be required at the scheduler's convenience.

We also note that ALFALFA coverage has been restricted since 2009 to preclude observations in the Right Ascension range $07^h30^m <$ R.A. $< 09^h15^m$ and $15^h00^m <$ R.A. $< 16^h30^m$ to accommodate time allocation to the AUDS program. Although we would have very much liked to complete the coverage of the full 7000 deg^2 region originally proposed, it appears impractical for us to plan to cover those regions.

Justification for completing the ALFALFA survey

We believe that ALFALFA has lived up to its expectation and promise, particularly by providing a high quality, well-characterized dataset to establish a robust census of extragalactic HI sources over a cosmologically significant volume of the local universe. A major limitation of other extragalactic HI surveys is their inability to account properly for both statistical and systematic uncertainties and for cosmic variance. Through its combination of sensitivity and sky coverage, only ALFALFA probes adequate volume with adequate sensitivity. Future work on the possible variation of the HI mass function, diameter function, correlation function and velocity function with local environment or parent galaxy population requires the full statistical dataset to provide further results on near-field cosmology, thereby establishing Arecibo's legacy contribution to the cosmic census of HI bearing objects at $z = 0$.

References

Baldry, I.K., Glazebrook, K., Brinkmann, J., *et al.* 2004, Ap.J., 600, 681

Barnes, D.G., Staveley-Smith, L., de Blok, W.J.G., *et al.* 2001, Mon.Not.Roy.Astr.Soc., 322, 486

Basilakos, S., Plionis, M., Kovac, K. & Voglis, N. 2007, Mon.Not.Roy.Astr.Soc., 378, 301

Dowell, J. 2010, Ph.D. thesis, Indiana University

Fabello, S., Catinella, B., Giovanelli, R., *et al.* 2010, Mon.Not.Roy.Astr.Soc., (in press)

Freudling, W., Staveley-Smith, L., Catinella, B. *et al.* 2010, Ap.J., (in press)

Fry, A.B., Boutan, C., Carroll, P.A., Hazelton, B. & Morales, M.F. 2011, BAAS 150.03

Giovanelli, R., Haynes, M.P., Kent, B.R., *et al.* 2005, A.J., 130, 2598.

Giovanelli, R., Haynes, M.P., Kent, B.R., *et al.* 2007, A.J., 133, 2569

Giovanelli, R., Haynes, M.P., Kent, B.R. & Adams, E.K. 2009, Ap.J.(Lett), 708, L22

Grebel, E.K., Gallagher, J.S.III & Harbeck, D. 2003, A.J.125, 1926

Hoeft, M., Yepes, G., Gottlober, S. & Springer, V. 2006, Mon.Not.Roy.Astr.Soc., 371, 401

Kent, B.R. 2008, Ph.D. thesis, Cornell University

Kent, B.R. 2010, Ap.J., (in press)

Kent, B.R., Giovanelli, R., Haynes, M.P., *et al.* 2008, A.J., 2008, 136, 713

Martig, M., Bournaud, F., Teyssier, R. & Dekel, A. 2009, Ap.J., 707, 250

Martin, A.M., Giovanelli, R., Haynes, M.P., *et al.* 2009, Ap.J.Suppl, 183, 214

Martin, A.M., Papastergis, E., Giovanelli, R., *et al.* 2010, Ap.J.(in press)

Masters, K.L., Haynes, M.P. & Giovanelli, R. 2004, Ap.J.(Lett), 607, L115

Meyer, M.J., Zwaan, M.A., Webster, R.L., *et al.* 2004, Mon.Not.Roy.Astr.Soc., 350, 1195

Meyer, M.J., Zwaan, M.A., Webster, R.L., *et al.* 2007, Ap.J., 654, 702

Press, W.H. & Schechter, P. 1974, Ap.J., 187, 425

Putman, M.E., Bland-Hawthorn,. K., Veilleux, S., *et al.* 2003, ApJ, 597, 948

Ricotti, M. 2009, Mon.Not.Roy.Astr.Soc., 392, L45

Roberts, M.S. & Haynes, M.P. 1994, Ann.Rev.Astron.Ap., 32, 115

Ryan–Weber, E.V., Begum, A., Oosterloo, T., *et al.* 2008, Mon.Not.Roy.Astr.Soc., 384, 535

Saintonge, A. 2007a Ph.D. thesis, Cornell University

Saintonge, A. 2007b, A.J., 133, 2087

Saintonge, A. Giovanelli, R., Haynes, M.P., *et al.* 2008, A.J., 135, 588

Salim, S., Rich, R.M., Charlot, S., *et al.* 2007, Ap.J.Suppl, 173, 267

Sternberg, A., McKee, C.F. & Wolfire, M.G. 2002, apjs, 143, 419

Stierwalt, S. 2009, Ph.D. thesis, Cornell University

Stierwalt, S., Haynes, M.P., Giovanelli, R., *et al.* 2009, A.J., 138, 338

Toribio, M.C. 2010, Ph.D. thesis, U. Barcelona

van der Hulst, J., van Albada, T. & Sancisi, R. 2001, in *Gas and Galaxy Evolution*, ASP Conference Series Volume 240, (Astronomical Society of the Pacific; San Francisco), p. 451

Wong, O.I., Ryan-Weber, E.V., Garcia-Appadoo, D.A., *et al.* 2006, Mon.Not.Roy.Astr.Soc., 371, 1855

Zwaan, M., Staveley-Smith, L., Koribalski, B.S., *et al.* 2003, A.J., 125, 2842

Zwaan, M., Meyer, M., Staveley-Smith, L., Webster, R.L. 2005, Mon.Not.Roy.Astr.Soc., 359, 30

Zwaan, M., Meyer, M.J. & Staveley-Smith, L. 2010, Mon.Not.Roy.Astr.Soc., 403, 1969

Appendix A. Websites with supplementary material

In previous reports for the annual NAIC skeptical review, we have provided details of the observing program and strategy, data processing scheme, IDL-based analysis software, database development activities, outreach activities and science results. By nature of its broad, diverse and ever-expanding nature, the ALFALFA collaboration functions by means of well-developed websites containing copious information, documentation, cookbooks and software utilities. Rather than include here a lot of material which has already been submitted to the skeptical review committee and/or which can be found on our websites, we refer the interested reader to the URLs of websites providing links to various documents previously submitted to NAIC or containing relevant information:

- ALFALFA survey public website
  http://egg.astro.cornell.edu/alfalfa/
- ALFALFA survey publication list
  http://egg.astro.cornell.edu/alfalfa/pubs.php
- ALFALFA survey projects
  http://egg.astro.cornell.edu/alfalfa/projects.php
- ALFALFA undergraduate team website
  http://egg.astro.cornell.edu/alfalfa/ugrad.php
- ALFALFA documentation website (including the original proposal and prior annual reports)
  http://egg.astro.cornell.edu/alfalfa/docs/index.php
- ALFALFA observing team website
  http://www.naic.edu/~a2010/galaxy_a2010.html
- ALFALFA team IDL website
  http://caborojo.astro.cornell.edu/alfalfalogs/docs/idldoc.php
- Cornell HI digital archive website
  http://arecibo.tc.cornell.edu/hiarchive