THE VELOCITY WIDTH FUNCTION OF GALAXIES FROM THE 40% ALFALFA SURVEY: SHEDDING LIGHT ON THE COLD DARK MATTER OVERABUNDANCE PROBLEM

Emmanouil Papastergis\textsuperscript{1}, Ann M. Martin\textsuperscript{1}, Riccardo Giovanelli\textsuperscript{1,2,3}, and Martha P. Haynes\textsuperscript{1,2,3}

\textsuperscript{1} Center for Radiophysics and Space Research, Space Sciences Building, Cornell University, Ithaca, NY 14853, USA; papastergis@astro.cornell.edu, amartin@astro.cornell.edu, riccardo@astro.cornell.edu, haynes@astro.cornell.edu

\textsuperscript{2} National Astronomy and Ionosphere Center, Cornell University, Ithaca, NY 14853, USA

Received 2010 December 21; accepted 2011 May 31; published 2011 September 6

ABSTRACT

The ongoing Arecibo Legacy Fast ALFA (ALFALFA) survey is a wide-area, extragalactic HI-line survey conducted at the Arecibo Observatory. Sources have so far been extracted over $\sim$3000 deg$^2$ of sky (40\% of its final area), resulting in the largest HI-selected sample to date. We measure the space density of HI-bearing galaxies as a function of their observed velocity width (uncorrected for inclination) down to $w = 20$ km s$^{-1}$, a factor of two lower than the previous generation HI Parkes All-Sky Survey. We confirm previous results that indicate a substantial discrepancy between the observational distribution and the theoretical one expected in a cold dark matter (CDM) universe, at low widths. In particular, a comparison with synthetic galaxy samples populating state-of-the-art CDM simulations implies a factor of $\sim$8 difference in the abundance of galaxies with $w = 50$ km s$^{-1}$ (increasing to a factor of $\sim$100 when extrapolated to the ALFALFA limit of $w = 20$ km s$^{-1}$). We furthermore identify possible solutions, including a keV warm dark matter scenario and the fact that HI disks in low-mass galaxies are usually not extended enough to probe the full amplitude of the galactic rotation curve. In this latter case, we can statistically infer the relationship between the measured HI rotational velocity of a galaxy and the mass of its host CDM halo. Observational verification of the presented relationship at low velocities would provide an important test of the validity of the established dark matter model.

Key words: dark matter – galaxies: dwarf – galaxies: luminosity function, mass function – galaxies: statistics – radio lines: galaxies – surveys

Online-only material: color figures

1. INTRODUCTION

The current “standard” $\Lambda$CDM cosmological model has been extremely successful at reproducing the bulk of the observed properties of our universe on large scales (Komatsu et al. 2011). However, given the current lack of a firm theoretical understanding of dark energy and the lack of a direct or indirect detection of the dark matter (DM) particle (The CDMS Collaboration 2009; Angle et al. 2008a, 2008b; Abdo et al. 2010a, 2010b; Adriani et al. 2009), it is important to test in detail the astrophysical implications of the established cosmological paradigm.

One of the most interesting consequences of assuming a cold dark matter (CDM) model is that substructure forms first on small scales, resulting in a present-day universe populated by a multitude of low-mass halos. More formally, the mass distribution of DM halos is described by the DM mass function (MF), which is defined as the number density of halos as a function of their virial mass; it can be analytically predicted (Press & Schechter 1974; Sheth & Tormen 2002) that the MF displays a power-law behavior at low halo masses, $n \propto M^\alpha$, with a relatively steep exponent of $\alpha \approx -1.9$ in the standard $\Lambda$CDM context. This analytical expectation, confirmed to great accuracy by N-body simulations of structure formation (Warren et al. 2006; Boylan-Kolchin et al. 2009; Klypin et al. 2010), leads to the prediction of a large number of low-mass halos for every Milky Way-sized (MW-sized) halo found in the present epoch.

This firmly established theoretical result has led to a number of observational challenges, such as the “missing satellites problem” (Klypin et al. 1999; Moore et al. 1999; Diemand et al. 2007; Strigari et al. 2007; Simon & Geha 2007), the “void phenomenon” (Peebles 2001; Tinker & Conroy 2009), as well as the discrepancy between the sizes of mini-voids observed in the local universe and those produced in CDM simulations (Tikhonov & Klypin 2009). Additional concerns, again closely related to the distribution of halo masses predicted by CDM, are raised by the flatness of the galactic luminosity function (LF; Blanton et al. 2005; Montero-Dorta & Prada 2009), HI mass function (HIMF; Martin et al. 2010), and Galactic stellar mass function (Baldry et al. 2008; Li & White 2009) at their faint/low-mass end. These observational distributions display power laws with $\alpha \approx -1.3$, much shallower than expected from the combination of a CDM universe plus a naive linear relationship between halo mass and luminosity/baryonic mass. Despite their apparent diversity, all statements described above are just different aspects of the same fundamental issue: CDM structure formation predicts large numbers of low-mass halos, seemingly in contradiction with the relative paucity of visible low-mass galaxies. Hereafter, we refer to this discrepancy as the CDM overabundance problem.

The main caveat regards the proper interpretation of these observational results. All phenomena mentioned so far rely on the measurement of quantities indirectly related to the mass of the hosting DM halo (e.g., luminosity or HI/stellar mass) and, as a result, do not provide a direct means of comparing the MF

\textsuperscript{3} The National Astronomy and Ionosphere Center is operated by Cornell University under a cooperative agreement with the National Science Foundation.

\textsuperscript{4} This statement does not aim at including a second class of potential observational challenges to CDM, related to the density profile of halos in their central regions (known as the “cusp versus core” problem).
expected for CDM with the MF realized in nature. In fact, a number of environmental and feedback effects (see Section 5.2) are expected to affect the baryonic content of halos, with low mass ones being the most impacted.

Ideally, one would need a large sample of galaxies with directly measured dynamical masses (e.g., through lensing or satellite kinematics), extending all the way to the low-mass regime. Unfortunately, current data sets are restricted to relatively massive galaxies. The best practical alternative would consist of a rich sample of resolved HI-interferometric rotation curves of galaxies, spanning a wide range in dynamical mass. Atomic hydrogen is usually the most spatially extended baryonic component in a galaxy, and therefore the best tracer of the rotation curve at large galactic radii. Such a sample could be used to determine the space density of galaxies as a function of their measured maximum rotational velocity, $v_{\text{rot}}$. This observational statistic, which is referred to as the velocity function (VF) of galaxies, is more directly related to the halo dynamics than statistics based on luminosity/baryonic mass and has a largely different set of systematic issues. However, current data sets are very limited, mostly because HI interferometry is extremely time consuming (especially for low HI-mass targets).

A more economical approach is to rely on wide-area, single-dish 21 cm surveys. Thanks to their intrinsic spectroscopic nature, HI surveys automatically obtain the spectral HI-line profile of every detected source. The velocity width of each detected galaxy, $w$, can thus be readily extracted and the associated data set can be used to measure the velocity width function (WF) of galaxies. One can furthermore apply inclination corrections to the measured widths in order to retrieve intrinsic rotational velocities ($v_{\text{rot}}$) and then estimate the galactic VF. Correcting for inclination requires however the use of external data sets, usually optical/near-IR (NIR) photometric surveys.

So far, the most accurate WF and VF for late-type galaxies have been based on 4315 and 2646 HI-selected galaxies, respectively, detected by the HI Parkes All-Sky Survey (HIPASS; Zwaan et al. 2010, hereafter Zw10). Their measurement of the WF extends over the velocity range $30 \, \text{km} \, \text{s}^{-1} < v_{\text{rot}} < 300 \, \text{km} \, \text{s}^{-1}$ and suggests a dramatic departure from the CDM expectation at low velocities ($v_{\text{rot}} \lesssim 100 \, \text{km} \, \text{s}^{-1}$). Recent determinations of the WF for massive early-type galaxies (which are mostly absent from HI-selected samples) have been obtained using the Sloan Digital Sky Survey (SDSS) and Two-degree Field Galaxy Redshift Survey (2dFGRS) data sets by Choi et al. (2007) and Chae (2010). Both the late-type and early-type distributions are needed in order to derive the “total” galactic VF, since massive early-type galaxies are the dominant population at high velocities ($250 \, \text{km} \, \text{s}^{-1} \lesssim v_{\text{rot}} \lesssim 450 \, \text{km} \, \text{s}^{-1}$) while late-types dominate the counts at lower velocities ($v_{\text{rot}} \lesssim 250 \, \text{km} \, \text{s}^{-1}$).

In this paper we present the Arecibo Legacy Fast ALFA (ALFALFA) measurement of the velocity width function of HI-bearing galaxies. The decision not to correct the measured widths for inclination is intentional, as the WF maintains all the advantages of the VF as a probe of the halo mass distribution, while featuring a number of observational advantages over the latter (see Section 3 for more details). The ALFALFA WF is based on 10,744 HI-selected galaxies (a more than twofold increase over previous data sets) and extends to widths as low as $w = 20 \, \text{km} \, \text{s}^{-1}$.

This paper is organized as follows: in Section 2 we present the ALFALFA survey and the associated data set; in Section 3 we discuss the observational advantages of the WF with respect to the inclination-corrected VF and present the ALFALFA measurement of the WF for HI-bearing galaxies; in Section 4 we address possible observational biases on the determination of the ALFALFA WF; in Section 5 we compare the ALFALFA measurement with the expectations in a CDM universe and describe the possible solutions to the observed discrepancy at low widths; in Section 6 we derive the relation between $v_{\text{rot}}$ (measured observationally) and $v_{\text{halo}}$ (calculated from N-body simulations) that would be needed to reconcile the velocity distributions of CDM halos and observed galaxies. We conclude with Section 7 by summarizing our results.

Throughout this paper we use a Hubble constant of $H_0 = 70 \, \text{km} \, \text{s}^{-1} \, \text{Mpc}^{-1}$, while $h$ refers to the Hubble constant in units of $100 \, \text{km} \, \text{s}^{-1} \, \text{Mpc}^{-1}$.

2. ALFALFA DATA SET

2.1. The Survey

The ongoing ALFALFA survey is a wide-area, blind 21 cm emission-line survey that takes advantage of the increased survey speed offered by the seven-feed Arecibo L-band Feed Array (ALFA) receiver at the Arecibo Observatory. The ALFALFA data are acquired in a minimally invasive drift-scan mode in two passes, ideally separated by several months in order to enable the discrimination between narrowband radio frequency interference (RFI) and small spectral width cosmic signals. When completed (i.e., when all data will have been processed), the survey will have detected $\gtrsim 30,000$ galaxies out over an area of $\approx 7000 \, \text{deg}^2$ of sky out to $cz \approx 18,000 \, \text{km} \, \text{s}^{-1}$. The ALFALFA survey is more sensitive than the previous generation HIPASS survey (Meyer et al. 2004; Zwaan et al. 2004), with a $5\sigma$ detection limit of 0.72 Jy km s$^{-1}$ for a source with a profile width of 200 km s$^{-1}$ as compared to a $5\sigma$ sensitivity of 5.6 Jy km s$^{-1}$ for the same source in HIPASS (GiovANELli et al. 2005). In addition to greater sensitivity, ALFALFA has a finer velocity resolution ($11.2 \, \text{km} \, \text{s}^{-1}$ versus 26.4 km s$^{-1}$ for smoothed data) and better angular resolution ($3.6$ versus $13'$ FWHM), resulting in a more accurate identification of optical counterparts.

2.2. The Sample

ALFALFA catalogs have so far been extracted (GiovANELli et al. 2007; Saintonge et al. 2008; Kent et al. 2008; Stierwalt et al. 2009; Martin et al. 2009; M. P. Haynes et al. 2011, in preparation) for a total area of 2934 deg$^2$. The current ALFALFA footprint consists of four distinct regions: two in the northern Galactic hemisphere, hereafter referred to as the Virgo direction region ($VdR: 07^h 30^m < \alpha < 16^h 30^m, 4^\circ < \delta < 16^\circ$, and $24^\circ < \delta < 28^\circ$), and two in the southern Galactic hemisphere, hereafter referred to as the anti-Virgo direction region ($aVdR: 22^h < \alpha < 03^h, 14^\circ < \delta < 16^\circ$, and $24^\circ < \delta < 32^\circ$). From this primary data set we only select extragalactic objects detected at high significance (signal-to-noise ratio (S/N) $> 6.5$, designated Code 1), and we further restrict ourselves to the redshift range $cz \lesssim 15,000 \, \text{km} \, \text{s}^{-1}$, beyond which radar interference from the nearby San Juan airport causes a significant drop of the ALFALFA detection efficiency. This final sample, corresponding to $\approx 40\%$ of the ALFALFA survey area (hereafter the $\alpha.40$ sample), contains a total of 11,086 galaxies.

Figure 1 shows the spatial distribution of the $\alpha.40$ sources in the Virgo and anti-Virgo directions, respectively, and puts in evidence the complex large-scale structure present in both
Figure 1. Spatial distribution of 5868 sources in the VdR (left panel) and 2055 sources in the aVdR (right panel). The Virgo Cluster and the “Great Wall” are the most conspicuous structures in the VdR (located at a distance of \( \approx 17 \) Mpc and \( \approx 100 \) Mpc, respectively). In the aVdR, the Pisces-Perseus Supercluster (clearly visible at \( \approx 70 \) Mpc) as well as the void in its foreground dominate the large-scale structure. Distances are assigned through a combination of a flow model for the nearby universe and Hubble distances for more distant galaxies (see Section 2.2).

Figure 2. Properties of the \( \alpha.40 \) sample: histogram (a) represents the distribution in heliocentric velocity (\( v_\odot \)), while the red solid line represents the distribution expected in a homogeneous universe according to the selection function of the survey; the complex large-scale structure in the survey volume is apparent. Histogram (b) represents the distribution of velocity width (\( w_{50} \)); note the large number of very low-width galaxies (\( w_{50} < 30 \) km s\(^{-1}\)) detected. Histogram (c) represents the distribution of galaxy HI mass (\( M_{\text{HI}} \)); again note the detections at very low HI mass (\( M_{\text{HI}} < 10^8 M_\odot \)).

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

volumes. Density fluctuations in the survey volume can be the dominant source of statistical uncertainty in surveys like ALFALFA, where the sample size ensures small counting errors. Our statistical estimator, described in Section 3.1, is chosen to minimize this structure-induced bias.

Figure 2 displays some statistical properties of the \( \alpha.40 \) sample. Histograms (a) and (b) represent the distribution of heliocentric velocity, \( v_\odot \), and of signal profile width, \( w_{50} \), which are both directly measured quantities (Giovanelli et al. 2007; Saintonge 2007). The signal profile width is measured at the 50% flux level of each of the two peaks of the typical double-horned HI profile (or at 50% of the single peak flux, for single-peaked profiles). The value of \( w_{50} \) reported in the ALFALFA catalogs is further corrected for instrumental broadening. Histogram (c) displays the distribution of galaxy HI mass, \( M_{\text{HI}} \), which is a distance dependent (and hence derived) quantity. Unlike previous HI surveys, we assign distances to nearby galaxies through a peculiar velocity flow model (Masters 2005) and use Hubble flow distances only for galaxies with \( cz > 6000 \) km s\(^{-1}\) (see Section 4.2 for a detailed discussion on the impact of distance uncertainties on ALFALFA results).

Figure 3 displays the distribution of \( \alpha.40 \) sources in the velocity width versus integrated flux (\( S_{\text{int}} \)) plane. The dashed red line is the survey completeness limit adopted in this work (\( S_{\text{int,lim}}/1 \text{ Jy km s}^{-1} = 0.06 (w_{50}/1 \text{ km s}^{-1})^{0.51} \)), which follows very closely the theoretically expected \( S_{\text{int,lim}} \propto w_{50}^{1/2} \).

(A color version of this figure is available in the online journal.)
3. THE VELOCITY WIDTH FUNCTION

We obtain rest-frame galaxy velocity widths, \( w \), by correcting the cataloged profile widths \( w_{50} \) for Doppler broadening. It is customary to apply additional inclination corrections to \( w \), in order to recover intrinsic rotational velocities, \( v_{\mathrm{rot}} \). However, since most extragalactic sources are unresolved at centimeter wavelengths, such corrections rely on external data sets (usually optical or NIR photometric surveys) for the determination of galaxy inclinations. Here, we choose to make no further corrections to \( w \) and measure the velocity WF of galaxies, denoted by \( \phi(w) \).

Even though the WF does not directly represent the distribution of any fundamental galaxy property, it is observationally superior to the (inclination-corrected) VF. In particular, it is free of the restrictions and systematics that arise from cross-matching HI and optical catalogs and correcting for galaxy inclination. For example, the HIPASS primary sample contains 4315 sources of which only 2646 have unambiguous optical counterparts (Doyle et al. 2005). Another 30% of the sources in this restricted subsample have low inclination values \((i < 45^\circ)\) and are thus subject to large inclination–correction errors. As a result, only \( \approx 43\%\) of the galaxies in the HIPASS primary sample were used for their determination of the VF. Furthermore, obtaining accurate estimates of the true orientation of irregularly shaped dwarf galaxies is challenging, and the process may introduce biases in the measurement of the low-velocity end of the VF.

Nonetheless, a measurement of the galactic WF would not be useful if it did not provide an accurate means of comparing the outcome of \( N \)-body simulations with the observed universe. Fortunately, it is relatively straightforward to project a given theoretical rotational velocity distribution and transform it into its corresponding width distribution (see Section 5.1). We conclude that the WF should be regarded as the prime observational distribution for single-dish HI surveys against which to compare theoretical expectations.

3.1. The ALFALFA Velocity Width Function

In Figure 4 we present the ALFALFA WF, based on 10,744 galaxies drawn from the \( a.40 \) sample. For the calculation of the WF we restrict ourselves to \( a.40 \) galaxies which are positioned in the portion of the flux–width plane where the ALFALFA survey is complete (i.e., above the red dashed line in Figure 3) and have profile widths broader than \( w_{50} \geq 18 \text{ km s}^{-1} \). This cut results in the elimination of \( \approx 330 \) galaxies from the calculation. An additional 13 very nearby sources are eliminated, for which the flow model assigned distances are subject to large uncertainty.

The WF is calculated in logarithmic width bins, according to the \( \Sigma 1/V_{\mathrm{eff}} \) method (Zwaan et al. 2005). The \( \Sigma 1/V_{\mathrm{eff}} \) method is a non-parametric maximum likelihood method and, as such, it is insensitive to the presence of large-scale structure in the survey volume. As its name suggests, it closely resembles the traditional \( \Sigma 1/V_{\mathrm{max}} \) method (Schmidt 1968) and consists of summing the number of detections in each width bin, weighted by the inverse of the “effective” volume available to each source. More precisely, the space density of galaxies belonging to width bin \( k \) \((k = 1, 2, \ldots, N_w)\) is

\[
\phi_k = \frac{1}{{V}_{\text{eff},i}} \int d\log w \phi^i(w) \frac{d\log w}{w} = \ln(10) \phi^i \left( \frac{w}{w_\star} \right)^{\alpha} e^{-(w/w_\star)\beta}.
\]

The least-squares parameters\(^6\) are \( \phi^i = 0.011 \pm 0.002 \), \( \log w_\star = 2.58 \pm 0.03 \), \( \alpha = -0.85 \pm 0.10 \), and \( \beta = 2.7 \pm 0.3 \) (uncertainties are statistical 1σ errors due to Poisson errors on the individual bin values). Note, however, that the final sample contains 163 sources that lack a confidently identified optical counterpart. Some of these sources correspond to tidal debris from nearby interacting galaxies and may not be

\[\text{Figure 4. ALFALFA velocity width function (WF): data points represent the space density of HI-bearing galaxies as a function of velocity width (corrected for Doppler and instrumental broadening, but uncorrected for inclination), as inferred from 10,744 galaxies detected by the 40% ALFALFA survey. The errors are 1σ Poisson errors due to galaxy counts in individual width bins. The red dotted line corresponds to a modified Schechter fit to the ALFALFA WF (see Section 3.1). The green solid line represents the fit to the HIPASS WF based on 4315 galaxies (Zwaan et al. 2010), over its measured range. (A color version of this figure is available in the online journal.)}\]
hosted by individual DM halos. Excluding these galaxies from the WF calculation leads to a somewhat shallower narrow-end slope of $\alpha = -0.68 \pm 0.11$. ALFALFA finds significantly more high-width galaxies than HIPASS (a factor of 3 at $w \approx 400 \text{ km s}^{-1}$, growing to a factor of $\sim 10$ at $w \approx 800 \text{ km s}^{-1}$), which is also evident from the marked difference in the position of the “knee” of the WF for the two surveys ($\log w_{50} = 2.58 \pm 0.03$ for ALFALFA versus $\log w_{50} = 2.21 \pm 0.10$ for HIPASS$^7$, in disagreement at the $>3\sigma$ level). Despite the fact that the nominal HIPASS volume is a factor of $\sim 5$ larger than the $\alpha = 0.40$ volume, ALFALFA detects more very broad profile galaxies. This is due to the limited sensitivity of HIPASS, which leads to a “thinning” of detections beyond $\approx 100 \text{ Mpc}$ and out to the survey boundary (area enclosed by vertical cyan lines).

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

---

7 No errors are reported for the published fit parameters to the HIPASS WF. In order to compare with ALFALFA, we derive errors by performing a least squares fit to the HIPASS WF data points.

---

4. Biases

4.1. Measurement Errors on $w_{50}$

Measurement errors on $w_{50}$ can shift galaxies among width bins, altering the bin counts and therefore the inferred space density. The $w_{50}$ value for ALFALFA sources is subject to two separate sources of error: one is statistical in nature and present for all sources, while the other is systematic and concerns only a fraction of the $\alpha = 0.40$ sample. The former is due to the distortion of the signal profile shape by noise; the latter results from the fact that the measurement of the spectral width of a signal relies on the accurate visual identification of its spectral boundaries, which is non-trivial for a number of sources (especially those found in the vicinity of RFI). The final width error reported in the ALFALFA catalogs, $\Delta w_{50}$, is the sum in quadrature of the random and systematic error terms described above. Owing to the fact that all $\alpha = 0.40$ galaxies are detected with high $S/N$ and have a clean spectral profile in the vast majority of cases, the typical $\alpha = 0.40$ width error is relatively small and its distribution well behaved. The median error is $\Delta w_{50,\text{median}} \approx 8 \text{ km s}^{-1}$ and $\sim 70\%$ of the sources have a fractional error of $\Delta w_{50}/w_{50} \lesssim 10\%$.

In order to assess the effect of $\Delta w_{50}$ on the WF, we create 50 mock galaxy samples by reassigning random widths to every galaxy $i$ in the primary ALFALFA data set according to their individual measured width ($w_{50,i}$) and error ($\Delta w_{50,i}$). Each mock sample is subject to the same cuts as the $\alpha = 0.40$ sample and a new realization of the WF is calculated (“1×” set). In order to illustrate the systematic trends introduced, we also perform an additional set of WF realizations with artificially inflated width errors (twice the reported ALFALFA width errors, “2×” set). The results are shown in Figure 6: overplotted to the original ALFALFA WF (data points and solid black line) are a modified Schechter fit to the mean WF corresponding to the $1\times$ (red solid line) and $2\times$ (dashed red line) realizations. Width errors at the ALFALFA error levels seem to only slightly affect the high-width end of the WF. As evidenced by the $2\times$ run, width errors generally lead to a rise of the high-width end, due to a net “diffusion” of galaxies from intermediate-width bins with larger number counts to high-width bins with lower number counts.

4.2. Distance Uncertainties

Since velocity width is a distance-independent quantity, galaxy counts in width bins are not altered by distance errors. However, the weights ($1/V_{\text{eff},i}$) that each galaxy contributes to its bin depend on HI-mass (see Equation (1) and discussion in Section 3.1), and therefore on the assumed distance. Masters et al. (2004) have shown that ignoring the local peculiar velocity field can lead to biased estimates of galaxy statistical distributions, especially for surveys drawing a large fraction of their sample from the Virgo direction (VdR). To avoid this bias ALFALFA uses redshift distances only for distant ($cz > 6000 \text{ km s}^{-1}$) galaxies and assigns distances to nearby galaxies through a parametric flow model developed by Masters (2005). The model includes two attractors (Virgo Cluster & Great Attractor), a dipole component (Local Group peculiar velocity),
a quadrupole component (Local Group asymmetric expansion), and a random thermal residual of $\sigma_{\text{local}} \approx 160 \text{ km s}^{-1}$. Here we assume that most of the coherent motion of nearby galaxies is correctly described by the flow model, and no significant bias results from this systematic component of galaxy peculiar velocities. Contrary to intuition however, even the random component $\sigma_{\text{local}}$ can induce a systematic bias through the “Eddington effect” (see for example Figure 6 in Zwaan et al. 2003).

In order to assess the effect of $\sigma_{\text{local}}$ on the WF, we proceed as in Section 4.1 and create 50 mock samples by adding Gaussian noise on the cataloged distance of each $\alpha.40$ galaxy. We calculate the WF corresponding to each sample realization and use the obtained average distribution to investigate the effect of distance uncertainties on the WF. We adopt the Masters (2005) value of $\sigma_{\text{local}} \approx 160 \text{ km s}^{-1}$, but we also perform simulations with double the fiducial dispersion ($\sigma_{\text{local}} \approx 320 \text{ km s}^{-1}$).

The results are displayed graphically in Figure 7, where the data points and solid line correspond to the original ALFALFA WF, and the blue solid and dotted lines correspond respectively to the results of $\sigma_{\text{local}} = 160 \text{ km s}^{-1}$ and $\sigma_{\text{local}} = 320 \text{ km s}^{-1}$ simulation sets. The largest effect is an overall increase in the amplitude of the WF at intermediate widths, which is probably due to the net transport of sources toward lower HI masses and therefore larger values of $1/V_{\text{eff}}$. Unlike in the case of the HIMF, the low-end slope $\alpha$ does not seem to be affected in any systematic way. We conclude that, apart from a mild increase in amplitude at intermediate widths, the WF is relatively insensitive to distance uncertainties due to galaxy peculiar motions.

4.3. Cosmic Variance

The WF presented in Figure 4 aspires to represent the distribution in a cosmologically representative volume. The sensitivity of the ALFALFA survey allows $\sim w_*$ and broader galaxies to be detected throughout the full $\alpha.40$ volume ($V_{\alpha.40} \approx 3.1 \cdot 10^9 h_3^2 \text{ Mpc}^3$), which ensures a cosmologically fair sampling of the MW-sized galaxy population. On the other hand, low-width galaxies tend to be faint systems that can only be detected in smaller volumes. As a result, the low-width bins of the WF are subject to increased uncertainty caused by the deviation of the galaxy distribution from homogeneity on small scales, which is referred to as cosmic variance (see Figure 8 for a graphical illustration).
In order to quantitatively assess the effects of cosmic variance on the ALFALFA WF, we jackknife resample the $\alpha_{.40}$ survey volume, by splitting it into 14 parts equally spaced in R.A. Then, we reevaluate the WF excluding each part in turn. The resulting scatter for each parameter, $\beta$, is given by

$$
\sigma^2 = \frac{N-1}{N} \sum (x - \bar{x})^2, \quad N = 14.
$$

The scatter calculated by Equation (3) would be equal to the purely statistical error if the survey volume were homogeneous, and so any excess noise results from the presence of inhomogeneities. The method described above provides a measurement of cosmic variance on linear scales smaller than those probed by the full survey and hence yields a conservative estimate of the true uncertainty (cosmic variance generally increases with decreasing scale).

The full uncertainties on the fit parameters (including cosmic variance) are $\phi_r = 0.011 \pm 0.003 \, (0.002) \, h_3^2 \, \text{Mpc}^{-3} \, \text{dex}^{-1}$, log $w_0 = 2.58 \pm 0.04 \, (0.03)$, $\alpha = -0.85 \pm 0.19 \, (0.10)$, and $\beta = 2.7 \pm 0.3 \, (0.3)$, where the term in parentheses represents the purely Poisson error reported in Section 3.1. Indeed, parameters $w_0$ and $\beta$, which dictate the shape of the WF at high widths, show a very modest increase in their uncertainty due to cosmic variance. On the other hand, the narrow-end slope $\alpha$ is significantly affected, with cosmic variance contributing a large fraction of the total error.

4.4. Beam Confusion

Beam confusion arises from the fact that the ALFA $3' \times 3'$ beam occasionally produces blends of small galactic groups at moderate distances, especially when individual galaxies are poorly separated in redshift space. The qualitative effect of such blends is to transform two or more independent sources into a single HI profile of larger $w_{50}$ than each of its constituents. We do not attempt to quantify the effect of confusion bias, but we anticipate it to be more pronounced at the high-width end of the WF. This is because galaxies with $w \gtrsim 550 \, \text{km s}^{-1}$ are preferentially found at large distances, where beam confusion is more severe. It is worth noting that this bias, even though present, cannot account for the discrepancy between the ALFALFA and HIPASS WFs at high widths, since the latter suffers from more confusion due to its larger beam size ($13' \, \text{FWHM}$).

5. COMPARISON WITH THEORY AND SIMULATIONS

The VF of halos in a CDM universe scales as $dn \propto v^{-4} dv$, where $v$ refers to the halo maximum rotational velocity. Even though a straightforward comparison of the CDM VF with the ALFALFA WF is not possible, such a steep scaling suggests a substantial discrepancy between the theoretical and observed distributions at low velocities.

In order to make a meaningful comparison between the theoretical prediction and the ALFALFA measurement, it is necessary to take into account a number of important effects.

1. DM halos exhibit significant substructure. High-resolution simulations have shown that several local density maxima (subhalos) develop within the virial radius of an underlying bound overdensity. As a result, massive halos ($M_{\text{halo}} \gtrsim 10^{13} \, \text{h}^{-1} \, M_\odot$) typically host groups or clusters of galaxies rather than a single astronomical object. In general, a one-to-one correspondence between simulated halos and visible galaxies is not always possible.

2. The collapse of baryons to the central region of DM halos affects the galactic potential and leads to a modification of the true galactic rotation curves compared to the ones obtained in dissipationless DM simulations.

3. The detectability of a galaxy in an HI survey depends on its atomic hydrogen content. Galaxies that are deficient in HI may be underrepresented in an HI-selected sample.

4. The relationship between the maximum of the rotation curve of a galaxy and its HI velocity width is non-trivial. Apart from the obvious dependence on disk inclination, the measured width depends on the spatial distribution of atomic hydrogen in the galactic potential. In particular, HI disks do not always extend far enough to sample the asymptotic outer part of the galactic rotation curve and may underestimate the mass of the host halo.

It is thus necessary to populate the DM halos of an N-body simulation with modeled galaxies and compare this virtual sample against the ALFALFA measurement. Modeling of the atomic hydrogen content of the synthetic galaxies is particularly desirable, because it greatly facilitates the comparison between theoretical and observed distributions.

Obreschkow et al. (2009, hereafter O09) have simulated the HI-line profiles for the galaxies in the De Lucia & Blaizot (2007) semi-analytic catalog, created by post-processing the Millennium N-body simulation (Springel et al. 2005). Figure 9 displays the WF (cyan solid line) resulting from projecting their modeled edge-on linewidths, assuming random galaxy inclinations. The O09 WF is in fairly good agreement with the ALFALFA measurement, but fails to display an exponential cutoff at high widths and therefore predicts too many high-width galaxies. This issue has been also pointed out in Zw10, who argue that the disagreement is caused by the fact that the O09 catalog overestimates the HI masses of massive early-type...
Figure 10. CDM overabundance problem. Data points with error bars and black solid line represent the measured ALFALFA WF (same as in Figure 4). The green lines represent the WF of a sample of synthetic galaxies modeled by Trujillo-Gomez et al. (2010, TG10), which populate the halos in the Bolshoi CDM simulation (Klypin et al. 2010). Two models were considered by TG10, one where the gravitational potential of baryons is simply superimposed on the DM potential (solid line) and one where the subsequent adiabatic contraction of the DM halo is taken into account (dash-dotted line). The blue solid line represents the WF of a modeled galaxy population corresponding to the higher resolution DM halo is taken into account (dash-dotted green line). Note that TG10 define \( v_{\text{rot}} \) as the value of the simulated rotation curve at a radius of 10 kpc. We argue that their modeling scheme and use of \( v_{10\,\text{kpc}} \) provide a good approximation of the measured velocity for galaxies with both flat and rising rotation curves.

Also plotted in Figure 10 is the WF of simulated galaxies based on the Zavala et al. (2009, hereafter Za09) constrained N-body simulation (blue solid line). Za09 perform a modest volume (\( 64\,h^{-1}\,\text{Mpc} \)) constrained simulation, designed to reproduce the large-scale structure of the local universe. Virtual galaxies are modeled according to the analytical results of Mo et al. (1998), assuming a disk-to-virial mass ratio of \( f_{\text{disk}} = M_{\text{disk}} / M_{\text{vir}} = 0.03 \) independent of halo size. Lastly, the maximum amplitude of the rotation curve (\( v_{\text{rot, max}} \)) for each galaxy is calculated by combining the disk and DM halo contributions.

Since neither model considers the distribution of the velocity field tracer (i.e., HI) in simulated galaxies, we convert rotational velocities into HI velocity widths by assuming the relationship

\[
w = 2v_{\text{rot}} \sin i + w_{\text{eff}}.
\]

Galaxies are assumed to be randomly oriented with respect to the line of sight (cos \( i \) is uniformly distributed in the \([0, 1]\) interval), while \( w_{\text{eff}} \) is a small “effective" term used to reproduce the broadening effect of turbulence and non-circular motions on HI line widths. The use of Equation (4) is only justified if the HI disk is extended enough to sample the value of \( v_{\text{rot}} \) adopted by the model under consideration (e.g., \( v_{10\,\text{kpc}} \) for TG10 and \( v_{\text{rot, max}} \) for Za09). We adopt the value \( w_{\text{eff}} = 5\,\text{km}\,\text{s}^{-1} \) for the broadening term, which is added linearly for galaxies with \( v_{\text{rot}} > 50\,\text{km}\,\text{s}^{-1} \) and in quadrature for lower velocity galaxies.

Figure 10 puts in evidence the marked departure of the theoretical distributions from the ALFALFA measurement at \( w < 200\,\text{km}\,\text{s}^{-1} \), which becomes more dramatic with decreasing width. According to the TG10 WF, the difference is approximately a factor of \( \sim 4 \) at \( w = 100\,\text{km}\,\text{s}^{-1} \), exhibiting an increasing trend. The Za09 WF\(^9\) implies a difference of a factor of \( \sim 8 \) at the lowest width where the simulation is complete (\( w \approx 50\,\text{km}\,\text{s}^{-1} \)) and displays a much steeper low-width slope than the ALFALFA measurement. An extrapolation of the Za09 WF to the ALFALFA width limit (\( w = 20\,\text{km}\,\text{s}^{-1} \)) would result in a discrepancy of a factor of \( \sim 100 \).

5.1. The CDM Overabundance Problem

CDM predictions start diverging from the observational results at low widths, and so the behavior of the theoretical WF for \( w < 200\,\text{km}\,\text{s}^{-1} \) is of great importance. Unfortunately, the very interesting work of O09 is only reliable for \( w \gtrsim 100\,\text{km}\,\text{s}^{-1} \) due to the limitations in the mass resolution of the Millennium Simulation. We employ instead two recent high-resolution CDM simulations that lack however modeling of the HI component of their virtual galaxy samples.

Figure 10 compares the ALFALFA measurement with the WF of the galaxy population corresponding to the Bolshoi simulation (Klypin et al. 2010), as modeled by Trujillo-Gomez et al. (2010, hereafter TG10). Each Bolshoi halo was assigned realistic stellar and cold gas masses, based on empirical relations. Subsequently, two models were considered, one where the gravitational potential of the baryons is simply superimposed on the DM potential (solid green line) and one where the DM halo adiabatically contracts in response to the presence of the baryons (dash-dotted green line). Note that TG10 define \( v_{\text{rot}} \) as the value of the simulated rotation curve at a radius of 10 kpc. We argue that their modeling scheme and use of \( v_{10\,\text{kpc}} \) provide a good approximation of the measured velocity for galaxies with both flat and rising rotation curves.

The ALFALFA measurement of the WF confirms the results of the HIPASS survey (Zwaan et al. 2010), which obtained its WF at lower sensitivity and velocity resolution. This fact excludes the possibility that the CDM overabundance problem is an artifact of the limited performance characteristics of past

\[^{7}\text{The value of } w_{\text{eff}} = 5\,\text{km}\,\text{s}^{-1} \text{ is derived empirically by Verheijen & Sancisi (2001), based on a sample of 22 galaxies with flat or decreasing outer rotation curves.}\]

\[^{9}\text{In order to account for the fact that the Za09 sample resides in an overdense volume (within a radius of } 20\,h^{-1}\,\text{Mpc from their simulated “Local Group”), we lower the normalization of their WF by a factor of two, as suggested in their Section 4.3.}\]
blind 21 cm surveys. The reason for the observed discrepancy can be therefore most likely attributed to one of the two following factors.

1. The inaccuracy of standard CDM simulations, presumably due to the inadequacy of the assumed DM model.
2. The improper comparison of simulated halos with observed galaxies. This could be due either to
   (a) the inadequate modeling of the baryonic counterparts hosted by DM halos, which leads to wrong predictions for maximum rotational velocities, or
   (b) the incorrect interpretation of inclination-corrected HI linewidths as maximum rotational velocities.

In what follows, we will consider these possibilities in more detail and argue about their prospects as solutions of the CDM overabundance problem.

Most large-scale simulations of cosmic structure conform to the standard ΛCDM cosmological model. In particular, they assume that all DM is cold (i.e., has negligible free streaming length), non-self-interacting, and stable (i.e., non-decaying). These properties are appropriate for a universe where DM consists of stable weakly interacting massive particles (WIMPs). WIMPs are currently the favored DM particle candidate and are expected to have masses in the GeV–TeV range and weak scale self-interaction cross sections, justifying the DM attributes most commonly assumed in cosmological N-body simulations.

However, the picture changes considerably if DM is composed of relatively light (∼keV) particles, in which case it is referred to as warm dark matter (WDM). Structure on large scales would be the same as in a CDM universe, but on small scales halo formation would be heavily suppressed due to the non-negligible free-streaming length of the light WDM particle. Za09 have considered this alternative scenario and carried out a second run of their very high-resolution simulation assuming a 1 keV WDM particle. They subsequently populate their halos with synthetic galaxies, employing the same modeling scheme as in their CDM run (Section 5.1). The result is shown by the red solid line in Figure 11, superposed on the ALFALFA WF (data points with error bars and black solid line) and the result of their CDM run (blue solid line).

Strikingly, the synthetic WF in the WDM case exhibits a shallow slope at the low-width end, in good agreement with the slope measured by ALFALFA. Such a shallow slope results from the suppressed production of low-mass halos in a WDM universe, which directly translates into a lower abundance of low-width visible galaxies. WDM could therefore provide a simple and elegant solution of the overabundance problem.

Despite its appeal in this specific context, the general prospects of WDM also depend on its overall viability as the dominant constituent of non-baryonic matter in the universe. A number of theoretical microscopic models for WDM have been proposed, most commonly involving sterile neutrinos (Dodelson & Widrow 1994; Fuller et al. 2003; Asaka et al. 2005; Kusenko 2009). Constraints on the particle’s mass can be placed by astrophysical and cosmological considerations. In particular, Lyα forest data places lower limits on the neutrino mass (a lighter particle generally results in suppression of power at larger scales), while X-ray observations can place upper mass limits (radiative decay into X-ray photons generally becomes more efficient at higher masses). The limits on the neutrino mass imposed by these observational constraints depend on the assumed neutrino production mechanism. Abazajian & Koushiappas (2006) find that non-resonantly produced neutrinos are ruled out, using a compilation of Lyα forest and X-ray data (see references therein). Boyarsky et al. (2009) have considered sterile neutrino production in the context of the vMSM (Minimal Standard Model + 3 sterile neutrinos) and argue that neutrinos with $m_\nu > 2$ keV are viable.

The second class of potential solutions attribute the disagreement between theory and observation to the process used to translate the output of simulations into actual galaxies. In particular, a number of important effects need to be taken into account (identified as items 1–4 in Section 5) to ensure a successful comparison of simulated halos with observed galactic samples.

Both theoretical works presented in Section 5 address issues 1 and 2. For example, Za09 set an explicit limit on the mass of halos hosting individual galaxies at $M_{\text{sat}} = 10^{13} h^{-1} M_\odot$. The influence of baryons on the shape of galactic rotation curves is also taken into account by both works, albeit using slightly different prescriptions and definitions of galaxy rotational velocity. Despite the use of numerous simplifying assumptions by TG10 (e.g., all baryons within 10 kpc) and Za09 (e.g., fixed disk-to-stellar mass ratio for all galaxies), their theoretical WFs are in fair agreement with the ALFALFA measurement at intermediate widths (200 km s$^{-1} < w < 500$ km s$^{-1}$).

The last two issues are related specifically to the atomic hydrogen content of galaxies, which is not modeled by either TG10 or Za09. Specifically, issue 3 concerns the detectability of a galaxy in a 21 cm survey. In principle, there exists the possibility that most of the low-mass halos predicted by CDM cosmology correspond to HI-devoid, dwarf spheroidal galaxies. In reality, a solution involving a multitude of isolated early-type dwarf systems seems rather unlikely. Direct observations (Garnett 2002; Swaters & Balcells 2002; Noordermeer et al. 2005), as well as other empirical arguments, suggest that the HI-to-stellar mass ratio grows with decreasing mass for galaxies in
the field. HI surveys should thus have an advantage, rather than a disadvantage, in detecting the baryonic counterparts hosted by low-mass DM halos. In addition, optical surveys suggest that isolated early-type dwarfs in medium/low density environments are relatively rare (Karachentsev et al. 2004). A second issue relates to the fact that satellite galaxies may be underrepresented in the α = 0.40 sample, since they are generally redder (and have presumably lower gas fractions) than central galaxies of the same luminosity (e.g., Font et al. 2008). This bias could result in a 30% underestimate of the abundance of low-width galaxies by ALFALFA (e.g., Yang et al. 2008; Klypin et al. 2010)—not nearly enough to explain the observed discrepancies.

Issue 4 regards the size and detailed spatial distribution of the atomic hydrogen component in galaxies, which determines the way in which its rotation curve is converted into an HI velocity width. In particular, $w_{HI}$ is a fair tracer of the maximum rotational velocity, only if the HI disk is extended enough to reach the flat (or decreasing) part of the galactic rotation curve. The use of Equation (4) in the derivation of the theoretical WFs implicitly assumes this situation to be true; observationally however, this is often times not the case. For example, the Catinella et al. (2006) set of template rotation curves provides evidence of the fact that lower rotational velocity galaxies tend to have steeper outer rotation curves (see their Figures 1 and 4). The dwarf galaxy samples of Spekkens et al. (2005) and Swaters et al. (2009) suggest that the effect becomes more dramatic at the lowest velocities (see Figures 3 and 4 in the respective references).

This systematic trend for lower velocity galaxies to host less extended HI disks can be understood in terms of the expected baryon depletion of low-mass halos. Results from $N$ body + hydrodynamics simulations (e.g., Hoefl et al. 2006; Ricotti 2010) indicate that halos with mass below some critical value lose a significant fraction of their cosmic share of baryonic matter, due to environmental and internal feedback processes. In particular, UV heating of the intergalactic medium after reionization is believed to lead to substantial gas removal from low-mass halos ($v_{rot} \lesssim 20–30$ km s$^{-1}$, corresponding to $M_{vir} \lesssim 10^{10–10^{13}}$ h$^{-1}$ M$_\odot$). Internal feedback processes such as supernova winds may also be important, but their efficacy is strongly model dependent.

The above considerations could lead to a solution of the overabundance problem that would not require a modification of the extremely successful ΛCDM paradigm. In simple terms, the overabundance problem would be the result of the inability of HI to trace the maximum halo rotational velocity of low-mass systems, which leads to a severe underestimate of their true mass. The same argument has been identified as a possible solution of the “mini-void size” problem by Tikhonov & Klypin (2009), while a similar effect has been proposed by Peñarrubia et al. (2008) as a solution to the “missing satellites” problem.

6. THE $V_{rot}$−$V_{halo}$ RELATION IN A CDM UNIVERSE

Assuming the CDM model to be correct, we can statistically infer the $v_{rot}$−$V_{halo}$ relationship needed to reproduce the observational galaxy WF. This can be done by abundance matching, a statistical procedure which assumes the existence of a one-to-one relationship between galaxy and halo circular velocities, $v_{rot} = f(V_{halo})$. It follows that the space density of halos with circular velocities larger than a given value, $V$, should be equal to the space density of galaxies with rotational velocities larger than the value dictated by the relationship $n(v_{halo} > V) = n(v_{rot} > f(V))$.

![Figure 12](image_url)

**Figure 12.** Thick red dash-dotted line represents the velocity function of late-type galaxies (assumed to follow a modified Schechter distribution) that best reproduces the measured ALFALFA WF (light gray data points and solid line) upon projection (thin red dash-dotted line). The green dotted line represents the velocity function of early-type galaxies determined by Chae (2010) using SDSS and 2dFGRS data. The blue solid line is a modified Schechter interpolation of the two VFs which represents a velocity function valid for all morphological types. The modified Schechter parameters for the interpolated distribution are $\phi_0 = 8.7 \times 10^{-3}$ h$^{-1}$ Mpc$^{-3}$, log $v_* = 2.49$, $\alpha = −0.81$, and $\beta = 3.35$ (see Section 6 for more details). (A color version of this figure is available in the online journal.)

Obtaining an observational velocity distribution from the ALFALFA measurement is not straightforward. First, the ALFALFA measurement regards galaxy velocity widths (uncorrected for inclination) and not intrinsic rotational velocities; second, the ALFALFA survey is biased against HI-poor massive ellipticals that dominate the counts at high velocities.

We address the first issue by searching for the velocity distribution that best reproduces the ALFALFA WF, upon projection using Equation (4). We assume that the distribution follows a modified Schechter function of the form

$$\phi(v) = \frac{dn}{d\log v} = \ln(10) \phi_0 \left( \frac{v}{v_*} \right)^\alpha e^{-\left(\frac{v}{v_*}\right)^\beta},$$

and that it corresponds to the VF of HI-rich, late-type galaxies. The set of parameters that provide the best match is identified visually, and corresponds to the values $\phi_0 = 1.2 \times 10^{-2}$ h$^{-3}$ Mpc$^{-3}$, log $v_* = 2.32$, $\alpha = −0.81$, and $\beta = 3.1$ (thick red dash-dotted line in Figure 12). In order to address the second issue (i.e., obtain a VF valid for all morphological types), we use the results of Chae (2010), who studied the velocity dispersion function of early-type galaxies in the SDSS and 2dFGRS surveys. Velocity dispersions can be transformed into rotational velocities by assuming an isothermal mass profile, in which case $v_{rot} = \sqrt{2\sigma}$. We adopt the average of the 2dFGRS and SDSS velocity distributions as a representative VF for early-type galaxies, which we plot as the green dotted line in Figure 12.

We interpolate the two distributions using a single modified Schechter function with parameters $\phi_0 = 8.7 \times 10^{-3}$ h$^{-1}$ Mpc$^{-3}$, log $v_* = 2.49$, $\alpha = −0.81$, and $\beta = 3.35$. The interpolated distribution (blue solid line in Figures 12 and 13) represents a composite galactic VF valid for all morphological types. Even though we do not formally measure errors for the derived distribution, we list below some important sources of uncertainty.
Table 1
The $v_{\text{rot}}$–$v_{\text{halo}}$ Relationship in a CDM Universe

| $v_{\text{halo}}$ (km s$^{-1}$) | $v_{\text{rot}}$ (km s$^{-1}$; $v_{\text{halo, max}} = 360$ km s$^{-1}$) | $v_{\text{rot}}$ (km s$^{-1}$; $v_{\text{halo, max}} = 290$ km s$^{-1}$) | $v_{\text{rot}}$ (km s$^{-1}$; $v_{\text{halo, max}} = 440$ km s$^{-1}$) |
|-----------------------------|---------------------------------|---------------------------------|---------------------------------|
| 40                         | 16                              | 16                              | 16                              |
| 45                         | 23                              | 23                              | 23                              |
| 50                         | 32                              | 32                              | 32                              |
| 55                         | 42                              | 42                              | 42                              |
| 60                         | 53                              | 53                              | 53                              |
| 70                         | 77                              | 78                              | 77                              |
| 80                         | 102                             | 103                             | 102                             |
| 90                         | 127                             | 127                             | 125                             |
| 100                        | 149                             | 150                             | 147                             |
| 120                        | 188                             | 190                             | 185                             |
| 140                        | 219                             | 223                             | 218                             |
| 160                        | 247                             | 252                             | 244                             |
| 180                        | 270                             | 278                             | 267                             |
| 200                        | 291                             | 303                             | 286                             |
| 220                        | 310                             | 327                             | 303                             |
| 240                        | 328                             | 353                             | 318                             |
| 260                        | 345                             | 383                             | 333                             |
| 300                        | 383                             | ...                             | 360                             |
| 340                        | 431                             | ...                             | 387                             |
| 380                        | ...                             | ...                             | 416                             |
| 420                        | ...                             | ...                             | 449                             |

First, the statistical uncertainty on the parameters of the late-type VF should be at least on the order of the errors reported in Section 3.1; the parameters of the composite VF should be expected to carry larger errors, since the determination of the interpolating distribution is subjective to some extent. More importantly, though, there are a number of sources of uncertainty related to galactic physics. For example, the assumption of isothermality of early-type galaxies is not expected to hold in detail (e.g., Dutton et al. 2010), which would affect the high-velocity end of the composite VF. Moreover, the low-velocity slope, $\alpha$, depends partly on the value of $w_{\text{eff}}$ employed in Equation (4); the value adopted here ($w_{\text{eff}} = 5$ km s$^{-1}$) has been empirically determined from a sample of relatively massive spirals (Verheijen & Sancisi 2001) and does not have to be the same for galaxies populating the low-velocity end of the VF. Also, as mentioned in Section 5.2, the inferred VF may be underestimating the true abundance of low-width galaxies by $\lesssim 30\%$, since ALFALFA is likely to miss some fraction of the satellites of massive spiral galaxies.

Next, we obtain the theoretical CDM VF from the Bolshoi simulation$^{10}$ (Klypin et al. 2010). In particular, we use the distribution of maximum halo rotational velocity, $v_{\text{halo}}$, of all simulated halos (including subhalos) at the present epoch, which is shown as the black solid line in Figure 13. Note that the simulation is run for the total matter density of the universe ($\Omega_m = \Omega_{DM} + \Omega_{\text{bar}} = 0.27$), but both DM and baryons are treated as dissipationless components. Also note that the simulation is complete only down to $v_{\text{halo}} = 50$ km s$^{-1}$, and a power-law extrapolation is used at lower velocities (which is, however, expected to hold, see for example Section 4.2 in Zavala et al. 2009).

The red thick line in Figure 14 represents the $v_{\text{rot}}$–$v_{\text{halo}}$ relation obtained by matching the CDM and galaxy velocity distributions (values listed in Table 1). We have assumed that halos with $v_{\text{halo}} > 360$ km s$^{-1}$ ($M_{\text{vir}} \gtrsim 10^{13} M_{\odot}$) do not host individual galaxies but rather groups of galaxies, and are hence excluded from the matching process. The cyan shaded region corresponds to different values for this mass cutoff, ranging from $v_{\text{halo,max}} = 290$ km s$^{-1}$ ($M_{\text{vir}} \approx 5 \times 10^{12} M_{\odot}$, upper boundary) to $v_{\text{halo,max}} = 440$ km s$^{-1}$ ($M_{\text{vir}} \approx 2 \times 10^{13} M_{\odot}$, lower boundary). The uncertainty in the value of $v_{\text{halo,max}}$ mentioned above is the only source of error considered explicitly here. There are, however, additional uncertainties involved in the determination of the presented relationship. For example, no scatter in the $v_{\text{rot}} = f(v_{\text{halo}})$ relation was considered in the

\[ h = 0.70, \Omega_{m} = 0.27, \sigma_8 = 0.82, n = 0.95. \]

\[ Ω_{m} = Ω_{DM} + Ω_{\text{bar}} = 0.27. \]

\[ v_{\text{rot}} = v_{\text{halo}}. \]
The Astrophysical Journal, 739:38 (13pp), 2011 September 20

Papastergis et al.

abundance matching process. Also, no corrections to \( v_{\text{rot}} \) for pressure support have been made in this work, even though gas thermal velocities in low-mass galaxies can be comparable with their rotational velocities.

Figure 14 shows that \( v_{\text{rot}} \) follows approximately linear relationship with \( v_{\text{halo}} \) only for intermediate-mass halos (120 km s\(^{-1}\) \( \lesssim v_{\text{halo}} \lesssim 170 \) km s\(^{-1}\)). In this range, \( v_{\text{rot}} \approx 1.5 \) \( v_{\text{halo}} \), in fair agreement with the values estimated for the MW and M31 from dynamical models (Klypin et al. 2002; diamonds) and from the kinematics of MW high velocity stars (Smith et al. 2007; triangle) and blue horizontal branch stars (Xue et al. 2008; box). However, the \( v_{\text{rot}}/v_{\text{halo}} \) ratios obtained here are significantly larger than the average values inferred by Dutton et al. (2010) from a compilation of weak lensing and satellite kinematics data sets. Note though that their results are expressed in terms of a \( v_{\text{rot}}-v_{200} \) relation, where \( v_{\text{rot}} \) is defined as the measured rotational velocity at 2.2 I-band disk scale lengths for late-type galaxies and 1.65\( \sigma \) for early-type galaxies and \( v_{200} \) refers to the virial velocity of the halo at an enclosed overdensity of 200 times the critical density. In order to display their results in Figure 14 (red and blue hatched regions), we have transformed virial velocities into halo maximum rotational velocities assuming average halo concentrations (Macciò et al. 2008).

The most important result of Figure 14 concerns the low halo velocity regime. In particular the relationship steepens continuously as we move to lower halo velocities, taking a power-law form \( v_{\text{rot}} \propto v_{\text{halo}}^\alpha \) at \( v_{\text{halo}} \lesssim 50 \text{ km s}^{-1} \). As a result, the true mass of low-velocity halos is systematically underestimated when measured by the inclination-corrected HI linewidth of the hosted galaxy; the underestimate can reach a factor of \( \approx 2.5 \) for \( v_{\text{halo}} \approx 40 \text{ km s}^{-1} \). Testing the low-velocity end of the \( v_{\text{rot}}-v_{\text{halo}} \) relation would require a sample of low-mass galaxies with directly measured dynamical masses, e.g., through weak lensing or satellite kinematics. However, some indirect observational support could come from a rich sample of HI interferometric maps of dwarf galaxies: a gradual transition from mostly flat to mostly rising rotation curves at \( v_{\text{rot}} \approx 110 \text{ km s}^{-1} \), would be required to explain the steepening of the relation at low velocities. Ultimately, observational verification of the presented relationship at low velocities would provide an important check of the validity of the CDM model.

7. CONCLUSIONS

We have measured the velocity WF of HI-bearing galaxies, based on a sample of 10,744 extragalactic sources detected in \( \approx 40\% \) of the final ALFALFA survey area. The ALFALFA measurement extends to widths (uncorrected for inclination) as low as \( w = 20 \text{ km s}^{-1} \) and results in a robust measurement of the low-width logarithmic slope \( \alpha = -0.85 \pm 0.19 \) (1\( \sigma \) statistical error including the effect of cosmic variance). This result suggests a significant incompatibility of the observational distribution with the much steeper distribution expected in a CDM universe.

We compare the ALFALFA result with the WFs of two modeled galaxy populations, one populating the Bolshoi CDM simulation halos (Trujillo-Gomez et al. 2010) and the other populating the halos of the very high resolution CDM simulation of Zavala et al. (2009). Indeed, the simulated WFs start diverging from the ALFALFA measurement at widths \( w \lesssim 200 \text{ km s}^{-1} \).
The difference in abundance is a factor of $\sim 8$ at $w = 50$ km s$^{-1}$ (which corresponds to the resolution limit of the Za09 simulation) and implies a difference of a factor of $\sim 100$ when extrapolated to the ALFALFA low-width limit ($w = 20$ km s$^{-1}$). This discrepancy is closely related to a number of other observational challenges to CDM (e.g., “missing satellites problem,” “mini-vide size problem,” etc.), which we collectively refer to as the CDM overabundance problem.

We further identify the two most promising solutions to the problem: the first involves the suppression of low-mass halo formation, which is best accomplished by assuming an $\sim$keV WDM particle; the second solution does not require a modification of the extremely successful CDM model and relies on the fact that HI disks in dwarf galaxies are frequently not extended enough to probe the full amplitude of the galactic rotation curve. The latter solution, supported by currently limited observational evidence, implies that galaxy rotational velocities derived from inclination-corrected HI linewidths ($v_{\text{rot}}$) systematically underestimate the maximum rotational velocity of their host DM halo ($v_{\text{halo}}$), below $v_{\text{rot}} \approx 110$ km s$^{-1}$.

We furthermore use an abundance matching procedure to statistically infer the $v_{\text{rot}}-v_{\text{halo}}$ relationship needed to reconcile the CDM and galactic velocity distributions. We find that for MW-sized galaxies $v_{\text{rot}} \approx 1.5 \, v_{\text{halo}}$, while at low velocities $v_{\text{rot}}$ underestimates significantly the true maximum rotational velocity of the host halo.

Determining the correct solution to the CDM overabundance problem rests both on the general prospects of WDM as a viable DM model as well as on observational verification of the $v_{\text{rot}}-v_{\text{halo}}$ relationship predicted for CDM. The latter goal could be best accomplished through a rich sample of low-mass galaxies with directly measured dynamical masses.

The authors acknowledge the work of the entire ALFALFA Collaboration in observing, flagging, and extracting the catalog of galaxies used in this work. This work was supported by NSF grant AST-0607007 and by grants from the Brinson Foundation. We also thank an anonymous referee for careful reading and very helpful comments and suggestions.