*R*esearch in
*A*stronomy and
*A*strophysics

Integrated HI emission in galaxy groups and clusters

Mei Ai[1,2,3], Ming Zhu[1,2] and Jian Fu[4]

[1] National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012, China; *aimei@nao.cas.cn*, *mz@nao.cas.cn*

[2] Key Laboratory of Radio Astronomy, Chinese Academy of Sciences, Beijing 100012, China

[3] University of Chinese Academy of Sciences, Beijing 100049, China

[4] Key Laboratory for Research in Galaxies and Cosmology, Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai 200030, China

Abstract The integrated HI emission from hierarchical structures such as groups and clusters of galaxies can be detected by FAST at intermediate redshifts. Here we propose to use FAST to study the evolution of the global HI content of clusters and groups over cosmic time by measuring their integrated HI emissions. We use the Virgo Cluster as an example to estimate the detection limit of FAST, and have estimated the integration time to detect a Virgo type cluster at different redshifts (from $z = 0.1$ to $z = 1.5$). We have also employed a semi-analytic model (SAM) to simulate the evolution of HI contents in galaxy clusters. Our simulations suggest that the HI mass of a Virgo-like cluster could be 2–3 times higher and the physical size could be more than 50% smaller when redshift increases from $z = 0.3$ to $z = 1$. Thus the integration time could be reduced significantly and gas rich clusters at intermediate redshifts can be detected by FAST in less than 2 hours of integration time. For the local Universe, we have also used SAM simulations to create mock catalogs of clusters to predict the outcomes from FAST all sky surveys. Comparing with the optically selected catalogs derived by cross matching the galaxy catalogs from the SDSS survey and the ALFALFA survey, we find that the HI mass distribution of the mock catalog with 20 s of integration time agrees well with that of observations. However, the mock catalog with 120 s of integration time predicts many more groups and clusters that contain a population of low mass HI galaxies not detected by the ALFALFA survey. A future deep HI blind sky survey with FAST would be able to test such prediction and set constraints on the numerical simulation models. The observational strategy and sample selections for future FAST observations of galaxy clusters at high redshifts are also discussed.

Key words: galaxy clusters — neutral hydrogen (HI) — galaxy evolution

1 INTRODUCTION

In the hierarchical scheme of structure formation, massive clusters of galaxies are the last objects to form, becoming prominent only at $z < 1$. Our understanding of galaxy evolution within the cluster formation scenario is dominated by two principal observational cornerstones: the strong signature of morphological segregation exhibited within rich clusters today and the Butcher-Oemler (B-O) effect, i.e., the increase in the blue cluster population with redshift. Observational tests on our understanding of the details of how environment (cluster formation) affects individual galaxies (morphology, gas content, star formation rate) are critically needed to validate the predictions of numerical simulations within the cluster formation scenario. In the local Universe, relative HI content serves as a useful comparative indicator of the future star formation potential of a galaxy. There have been many observational studies of nearby galaxy clusters, mostly in optical and X-ray wavelengths. However, study of HI selected galaxies in a cluster is rare, e.g. Freudling (1989). Its simple physics makes the HI line

transition an attractive potential tracer of galaxy evolution over cosmic time. But because of the inherent weakness of the HI line and contamination of the spectrum by terrestrial radio frequency interference (RFI), studies of redshift HI emission from individual galaxies have only recently been possible up to 0.25–0.37 (Catinella et al. 2008; Fernández et al. 2016). Beyond that, the study of HI emission from individual objects requires the resolution and collecting area of the future Square Kilometre Array (SKA).

Nowadays there are many studies on HI properties on the scales of clusters in the local Universe (Giovanelli & Haynes 1985; Solanes et al. 2001; Chung et al. 2009; Wolfinger et al. 2013) or groups (Stevens et al. 2004; Sengupta & Balasubramanyam 2006), or on the HI properties between galaxies in groups or clusters (English et al. 2010). Most of this research has focused on the influence of cluster environment on the HI content in galaxies.

Some local spiral-rich clusters, such as Virgo, Hercules, Ursa and Abell 3128, have been observed by radio telescopes, and these investigations indicated that most HI gas resides in spiral galaxies, while elliptical galaxies contain a limited amount of HI gas. Young clusters consisting of more spiral galaxies contain more HI than old clusters with more elliptical galaxies. Some mechanisms such as ram-pressure (Gunn & Gott 1972), viscous stripping (Nulsen 1982) and tidal interaction (Merritt 1983) could contribute to the process of HI decreasing during the evolution of clusters/groups. The core region of a cluster is usually hot and most HI content resides at the edge of the cluster, but recent studies show that galaxies in a cluster begin to lose their HI gas at an intermediate distance to the cluster center (Chung et al. 2009). The relative importance of the HI deficiency mechanisms in clusters/groups in the local Universe is still debated.

Now that construction of the Five-hundred-meter Aperture Spherical radio Telescope (FAST) has been completed, we propose to use FAST to study the evolution of the global HI content of clusters and groups over cosmic time by measuring the integrated HI emission from clusters. The sensitivity of FAST is high enough to detect HI emission from galaxy groups and clusters at intermediate redshifts. In this paper we discuss the observational strategy and sample selections for future FAST deep HI surveys. We also use numerical simulations to predict the global HI contents on a cluster scale at different redshifts. In all the calculations we adopt the ΛCDM cosmology parameters: $H_0 = 70 \, \mathrm{km s^{-1} \, Mpc^{-1}}$, $\Omega_\Lambda = 0.7$ and $\Omega_m = 0.3$ unless otherwise stated.

2 FAST OBSERVATIONS OF INTEGRATED HI EMISSION IN GALAXY CLUSTERS/GROUPS

2.1 Technical Specifications of FAST

Construction of major parts of the structure for FAST was completed in September 2016. This telescope has an aperture of 500 m and an illuminated aperture 300 m in diameter. It is located in Guizhou Province in southern China, with a latitude of 25.65°. FAST can reach a zenith angle of 40°, so the sky coverage of declination would be between −14° and 65°. The detailed technical specifications of FAST can be found in the paper Nan et al. (2011) and an updated status report is in Li & Pan (2016). There are seven sets of receivers being developed for FAST. The major one used for HI observations is the 19-beam feed-horn array receiver covering the frequency range of 1.05–1.45 GHz with $T_{\mathrm{sys}} = 25$ K. For galaxies at redshift higher than 0.35 and up to 1.5, a single-beam receiver covering the frequency range of 560 MHz–1.12 GHz can be used. This receiver has a better system temperature of about 20 K, which is ideal for high redshift HI observations.

FAST has very high sensitivity, but its resolution is relatively low compared to interferometers. One of the most important science goals of FAST is to investigate the HI distribution in the Universe. As the main beam of FAST is relatively large ($2.95'$), there could be more than one galaxy in the FAST beam, especially in distant clusters and close groups. Thus in this paper we will investigate detecting the total HI content in groups or clusters, without distinguishing individual galaxy members. Many groups and clusters are gravitationally bound structures, thus our study can reveal the large scale HI structure of the local Universe.

2.2 Science Goal: Evolution of HI Contents in Galaxy Groups/Clusters

The major goal for high-z HI observations is to study the evolution of HI gas contents. FAST is expected to detect HI on cluster scales at intermediate redshifts. This allows us to study the evolutionary effect of HI content in clusters at different redshifts.

Butcher and Oemler (Butcher & Oemler 1978) pointed out that there is a substantial population of blue galaxies in clusters at $z \geq 0.4$, but nearby clusters are dominated by elliptical and lenticular galaxies. The dramatic transformation of galaxies in clusters happened during redshift between $0 - 0.4$, a look back time of 4–5 billion years. However, HI observations at present are limited to the low redshift Universe. The highest redshift

of the ALFA extragalactic HI survey (ALFALFA) using the Arecibo 305 meter telescope is 0.06 (Giovanelli et al. 2005). The HIGHz Arecibo survey detected 39 galaxies at $0.16 < z < 0.25$ with

$$M_{\mathrm{HI}} = (3 - 8) \times 10^{10} \, M_\odot$$

(Catinella & Cortese 2015). About 160 galaxies in two cluster regions at $z \sim 0.2$ have been detected by the Blind Ultra Deep HI Environmental Survey (BUDHIES) using the Westerbork Synthesis Radio Telescope (WSRT), with HI mass ranging from $5 \times 10^9 \, M_\odot$ to $4 \times 10^{10} \, M_\odot$ (Jaffé et al. 2013; Verheijen et al. 2007). One galaxy (J100054.83+023126.2) at $z = 0.376$ was detected by the COSMOS HI Large Extragalactic Survey (CHILES, Fernández et al. 2016) which is the highest redshift galaxy so far observed using the HI line. The HI mass of this galaxy is $2.9 \times 10^{10} \, M_\odot$. The number of detected HI galaxy clusters at present is too small for a statistical study of the HI gas content on cluster scales.

Deep HI observations with FAST should be able to detect cluster scale HI gas at redshifts high enough to study the B-O effect. Evolution of the HI gas content in clusters as a function of redshift would be an important science goal for future FAST HI surveys. This will not only help in understanding the relative importance of mechanisms causing HI deficiency in nearby clusters, but can also shed light on the global evolution of galaxy clusters.

2.3 FAST Detection Limit for Gas Rich Clusters: a Case Study of Virgo

Virgo is one of the nearest rich galaxy clusters. It has a complex structure comprised of at least three distinct clouds, suggesting that the Virgo Cluster is in a relatively early stage of evolution. Unlike the Coma cluster where there is almost no HI detected, there is still a large amount of HI rich galaxies in the Virgo Cluster. Here we use Virgo as an example to evaluate the ability of detecting integrated HI emission at high redshifts with FAST. The Virgo Cluster region covers $130 \, \mathrm{deg}^2$ over $12^{\mathrm{h}}04^{\mathrm{m}} < \mathrm{RA} < 12^{\mathrm{h}}44^{\mathrm{m}}$, $3° < \mathrm{Dec} < 16°$, $0 < V \leq 3000 \, \mathrm{km \, s^{-1}}$. The exact boundary of the Virgo Cluster has a limited effect on our estimates.

We extract the HI sources in this region from the ALFALFA 70 (α.70) catalog (Haynes et al. 2011), which contains 70% of the survey data. We include both ALFALFA "code 1" and "code 2" sources, which have a signal-to-noise ratio S/N > 6.5 and 4.5 < S/N < 6.5, respectively. They are likely to be real because nearly all of them have a known optical counterpart (OC) at the

redshift of the HI source (Haynes et al. 2011). There are 350 HI sources from ALFALFA in the Virgo region. As shown on Figure 1, HI sources are plotted as blue circles with radius proportional to their HI mass which is varying between $10^6 \, M_\odot$ and $10^{10.07} \, M_\odot$. To estimate the integration time of FAST for observing such an object, we treat the whole Virgo Cluster as one "synthesis" HI galaxy. We put the cluster at different redshifts and calculate the integrated HI fluxes.

The luminosity distances of different redshifts are listed in Table 1. FAST beam size is $2.95'$ at $z = 0$ and varies with redshift as $2.95'(1 + z)$. The physical beam size is calculated from the angular diameter distance D_{A} multiplied by the FAST beam size in radians at the corresponding redshift where

$$D_{\mathrm{A}} = \frac{1}{1+z} \frac{c}{H_0} \int_0^z \frac{dz}{\sqrt{\Omega_\Lambda + \Omega_m (1+z)^3}}.$$

According to the theory of dark matter halo growth (White & Frenk 1991), the radius of a virialized halo at redshift z follows the relation

$$r_{\mathrm{vir}} = 0.1 H_0^{-1} (1+z)^{-3/2} v_{\mathrm{c}}.$$

Hence we can assume that the size of the Virgo Cluster at redshift z is $r_0(1+z)^{-3/2}$, where $r_0 = 1.95$ Mpc is the Virgo Cluster radius at $z = 0$. So, the cluster size should be smaller at higher redshift. As z increases, the FAST beam would cover a larger physical area of the Virgo Cluster, and more galaxies would contribute to the observed HI fluxes. At $z = 0.7$, the cluster diameter would be 1.92 Mpc which is less than the physical beam size at this redshift and all the Virgo HI galaxies would be incorporated in the beam. It should be noted that here we only consider the size evolution of the cluster and do not consider the evolution of the HI gas mass. In Section 3.2 we provide a more accurate estimate of the cluster size and HI mass based on cosmology simulations.

On Figure 1 we overlay the beam circles (black dashed circles) at different redshifts on the HI sources distributed on physical scales. The smallest black dashed circle corresponds to the physical beam size where we put the cluster at a redshift of 0.5 and the largest one corresponds to a redshift of 0.7. The redshift interval for the black dashed beam circles is 0.1. The blue filled circles are the HI galaxies that reside in the Virgo Cluster region that we extracted from the α.70 catalog. The radii of the blue filled circles are proportional to the HI mass.

To estimate the FAST detection limit, we add the HI mass of all sources covered by the beam circle, and use the co-added mass to estimate the HI flux and the

Table 1 The Observing Parameters of FAST when using the Virgo Cluster as a Sample

Redshift	M_{HI} $(10^{10} M_\odot)$	S_{peak} $(10^{-2}\mathrm{mJy})$	τ (min)	w (km s^{-1})	Beam position	Beam size (Mpc)	N_{beam}	Luminosity Distance (Mpc)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
0.1	1.30	52.0	5	500	(183.5,13)	0.36	17	460.3
0.2	2.26	19.9	35	500	(183.5,13)	0.70	38	980.1
0.3	4.20	14.8	64	500	(183.5,13)	1.02	69	1552.7
0.4	5.35	9.6	151	500	(183.5,13)	1.33	99	2172.1
0.5	9.75	7.4	258	700	(186,9.5)	1.62	249	2832.8
0.6	15.69	7.6	240	700	(186,9.5)	1.89	335	3530.1
0.7	16.20	5.4	478	700	(186,9.5)	2.15	350	4259.5
0.8	16.20	3.9	920	700	(186,9.5)	2.39	350	5017.5
0.9	16.20	2.9	1645	700	(186,9.5)	2.62	350	5801.0
1.0	16.20	2.3	2768	800	(186,9.5)	2.83	350	6607.2
1.1	16.20	1.8	4436	800	(186,9.5)	3.04	350	7433.8
1.2	16.20	1.4	6824	800	(186,9.5)	3.23	350	8278.8
1.3	16.20	1.2	10141	800	(186,9.5)	3.41	350	9140.5
1.4	16.20	1.0	14630	800	(186,9.5)	3.58	350	10017.5
1.5	16.20	0.8	20571	800	(186,9.5)	3.74	350	10908.4

Notes: Column (1) is the redshift where we put the cluster, Col. (2) is HI mass within the FAST beam, Col. (3) is the peak flux, Col. (4) is integration time, Col. (5) is the velocity width of the assumed HI source, Col. (6) is the position of the beam center, Col. (7) is physical size of the FAST beam when it is observing at different redshift, Col. (8) is the number of HI galaxies in the FAST beam and Col. (9) is luminosity distance.

required integration time of FAST

$$S_{\mathrm{peak}} = \frac{M_{\mathrm{HI}}}{2.36 \times 10^5 \cdot D_{\mathrm{Mpc}}^2 \cdot \omega_{\mathrm{km\ s}^{-1}}}, \quad (1)$$

$$S_{\mathrm{noise}} = \frac{k}{\frac{A}{T_{\mathrm{sys}}}\sqrt{\tau \cdot \delta\nu}}, \quad (2)$$

and assuming

$$S_{\mathrm{noise}} = \frac{S_{\mathrm{peak}}}{5}, \quad (3)$$

so

$$\tau = \left(\frac{k}{S_{\mathrm{noise}} \cdot \frac{A}{T_{\mathrm{sys}}}}\right)^2 \cdot \frac{1}{\delta\nu}, \quad (4)$$

where $\omega_{\mathrm{km\ s}^{-1}}$ is the estimated velocity widths which we assumed according to different redshifts, D_{Mpc} is the redshift where we put the Virgo Cluster, τ is integration time, $\frac{A}{T_{\mathrm{sys}}}$ is the sensitivity of FAST which is taken as $2000\,\mathrm{m}^2\,\mathrm{K}^{-1}$ (Nan et al. 2011), k is the Boltzmann constant and $\delta\nu$ is frequency resolution which is assumed to be $142\,\mathrm{kHz}$, corresponding to a velocity resolution of $30\,\mathrm{km\,s}^{-1}$. The results are shown in Table 1.

In Table 1, z is the redshift where we put the cluster, M_{HI} is HI mass within the FAST beam, S_{peak} is the corresponding peak flux, τ is the integration time and ω is the velocity width of the assumed HI source. For redshift from 0.1 to 0.4 we centered the beam at RA $= 12^{\mathrm{h}}14^{\mathrm{m}}$, Dec $= 13$ which is the local concentration that contains more HI galaxies than in the center. The beam size in column 7 is the physical diameter of the FAST beam when

observed at different redshifts. N_{beam} in column 8 is the number of HI galaxies in the FAST beam and the luminosity distance of each redshift is in column 9. The velocity width, $\omega_{\mathrm{km\,s}^{-1}}$, is determined based on the velocity dispersion of the galaxy group/cluster catalog extracted from the SDSS and 2MASS surveys. Its value is similar to the group velocity dispersion when N_{beam} is close to the number of member galaxies in a group.

In the above analysis we do not consider the evolutionary effect of HI content at higher redshift. Rhee et al. (2016) conclude that there is no significant evolution in cosmic HI mass density from $z = 0 \sim 0.4$. The weighted mean average of the HI mass density (Ω_{HI}) from all 21-cm measurements at redshifts $z < 0.4$ is $(0.35 \pm 0.01) \times 10^{-3}$. However, the B-O effect (Butcher & Oemler 1978, 1984) indicates that clusters at $z \geq 0.4$ have a substantial population of blue galaxies implying more HI rich galaxies, but the nearby rich clusters are very deficient in HI gas. As shown in Section 3.2, numerical simulations predict that the HI mass of a cluster could be about 2–3 times larger when z increases from 0.3 to 0.8. Thus the integration time for a high z gas rich cluster can be 1/4–1/9 times less than that listed in Table 1. If this is the case, a cluster at $z = 0.8$ can be detected by FAST in about 2 hours of integration time.

2.4 Detecting Optically Selected Clusters with FAST

To further evaluate the feasibility of detecting high redshift galaxies based on an optically selected sample (Wen

et al. 2012; Wen & Han 2015, the WHL cluster catalog), we use the WHL optically selected cluster catalog to estimate the HI mass of higher redshift clusters. We choose four clusters at redshift around 0.1, 0.2, 0.3 and 0.4. These four clusters represent the richest clusters at each redshift. We retrieve the g and r band absolute magnitudes from SDSS Casjobs (absMagG and absMagR from the Photoz table, k corrected to $z = 0$) for the WHL cluster member galaxies using their position (RA and Dec). Then we use the calculated r band luminosity and $g - r$ color to compute each member galaxy's HI mass using the following equations (Bell et al. 2003; Zhang et al. 2009):

$$\log(M_*/L) = -0.306 + 1.097(g - r) \qquad (5)$$

and

$$\log(G/M_*) = 1.09431 - 3.08207(g - r), \qquad (6)$$

where L and M_* are the r band luminosity and stellar mass of the member galaxy, respectively, and G is the calculated HI mass.

All these four clusters have a radius of about $\sim$ 2 Mpc which is larger than the FAST beam size at the corresponding redshift. We calculate the HI mass that is covered by the beam at each redshift. In each cluster we consider only the galaxies covered by the FAST beam. We add those galaxies' HI mass and apply this mass in Equations (1) to (4) to calculate the integration time needed for FAST to detect such a cluster.

The results are shown in Table 2. From Table 2 we can see that all these four clusters can be detected by FAST in less than 40 min. However, the HI masses estimated from Equation (5) and Equation (6) carry large uncertainties. The scatters of Equations (5) and (6) are 0.1 dex and 0.35 dex, respectively. Hence the integration time could be several times longer than that listed in Table 2. In the worst situation, the integration time for detecting the WHL cluster at $z = 0.4$ would be more than 4 hours. Thus target selection is crucial for detecting high redshift clusters.

3 THEORETICAL PREDICTIONS

3.1 HI Contents in Nearby Groups and Clusters

In order to understand the integrated HI contents in high redshift clusters, we need to first study the HI properties in groups and clusters at low redshifts. In the hierarchical galaxy formation model, disk galaxies form first and then fall into bigger halos. By merging, dark matter halos grow over cosmic time. Galaxies tend to appear in groups or clusters. Groups have fewer galaxy members than clusters, but can grow into clusters when more members are accumulated. Thus we include both galaxy groups and clusters in our mock catalog.

3.1.1 HI data in optically selected galaxy groups and clusters

Before creating a mock catalog of groups/clusters, we first need to choose a control sample of galaxy clusters to make a comparison with the simulation results. Berlind et al. (2006) generated a cluster catalog from the SDSS DR7 survey using an optimized algorithm of the friends of friends (FoF) method (available online at *http://lss.phy.vanderbilt.edu/groups/dr7/*). There are three galaxy samples created with different absolute magnitude limits and redshift ranges. Each one is complete to within the stated limits. Given that the highest redshift of the HI sources in ALFALFA is around 0.06 (Giovanelli et al. 2005) and the luminous HI source corresponds to less luminous optical galaxies, we choose the faintest group, Mr18, in the sample whose r band absolute magnitude is down to $M_r = -18$. Then we cross match the SDSS Mr18 cluster member galaxy catalog with the ALFALFA 70% (α.70) HI source catalog (see the ALFALFA 40% description in Haynes et al. 2011). The α.70 catalog, which contains 70% of all the ALFALFA data, is available on the website *http://egg.astro.cornell.edu/alfalfa/data/*. In this catalog, the detected HI sources with SDSS OCs available have been identified and their (RA, Dec) coordinates are listed. We cross match the Mr18 cluster member galaxies with the α.70 HI sources using the software package TOPCAT, and the angular difference between SDSS member galaxies and α.70 HI OC's coordinates is 5 arcmin. There are 1443 SDSS clusters (9868 member galaxies) in the selected region with $N \geq 3$, where N is the number of member galaxies in a cluster. However, only 2134 α.70 HI sources are cross matched with SDSS member galaxies within 977 clusters. Thus the total detection rate is only $2134/9868 = 22\%$ for the optically selected sample. High sensitivity observations with FAST can greatly improve the detection rate of HI galaxies in groups which is crucial for studying the evolution of HI contents in groups and clusters.

The final control sample of HI groups/clusters is composed of cross matched SDSS-ALFALFA clusters which fall in the ALFALFA Spring Sky region with $7^h30^m < \text{RA} < 16^h30^m$, $0° < \text{Dec} < 30°$. The redshift range of the SDSS clusters is $0.02 < z < 0.042$.

Table 2 Parameters of Optically Selected WHL Clusters

Cluster name	z_{ph}	r_{200} (Mpc)	N_{200}	Half beam size (Mpc)	N_{gal}	M_{HI} ($10^{11} M_\odot$)	M_{HI} in beam ($10^{10} M_\odot$)	$\omega_{km\,s^{-1}}$ (km s^{-1})	S_{peak} (mJy)	τ (min)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
J004629.3+202805	0.1040	1.88	121	0.18	5	7.78	1.37	1000	0.27	19
J092200.5+515521	0.2042	1.98	128	0.35	13	11.34	7.24	1000	0.31	14
J233739.7+001617	0.3042	1.93	103	0.51	26	10.25	26.46	1500	0.30	15
J081025.7+181724	0.4123	1.71	80	0.67	61	7.83	31.78	1500	0.19	40

Notes: Columns are as follows: (1) WHL cluster name; (2) cluster photometric redshift; (3) cluster radius; (4) the number of member galaxies within r_{200}; (5) radius of the FAST beam at different redshifts; (6) the number of member galaxies that are covered by the beam; (7) the HI mass of the cluster; (8) the HI mass covered by the FAST beam; (9) the velocity width; (10) peak flux of the cluster HI sources and (11) the estimated integration time for FAST.
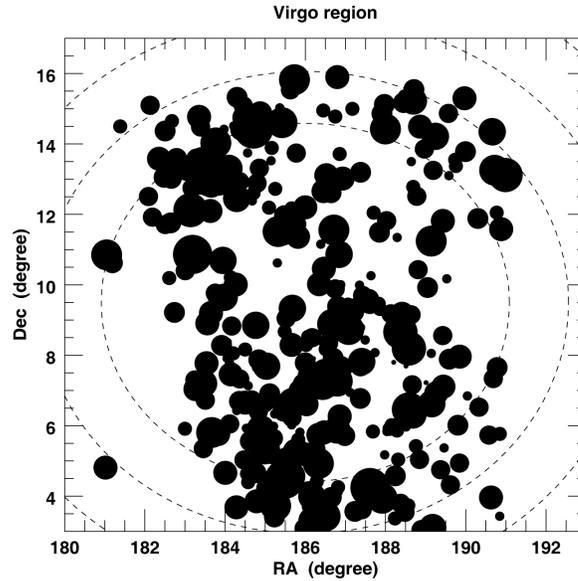
Fig. 1 HI sources within the Virgo Cluster region. Blue circles represent the HI galaxies, with their areas proportional to their HI masses. Black circles indicate the physical area covered by the FAST beam at a distance of $z = 0.5$ to $z = 0.7$. The redshift interval is 0.1.

3.1.2 A mock catalog of groups and clusters

We adopted the same method that is described in Duffy et al. (2012) to create the galaxy catalog by applying the SAM of Croton (2006), which is based on the underlying dark-mater-only MILLENNIUM SIMULATION (Springel et al. 2005). This sample of galaxies can accurately recreate the observed stellar mass function with both supernova feedbacks and feedbacks at the high-mass end due to active galactic nuclei. The cosmology used in the MILLENNIUM SIMULATION has parameters $M = 0.25$, $c = 0.75$, $b = 0.045$, $\sigma_8 = 0.9$ and Hubble parameter $h = 0.73$. Using the sensitivity parameters of FAST with the 19 beam array receiver, we make use of the CLOUD-based web application administered by the Theoretical Astrophysical Observatory (TAO) which generates mock catalogs from different cosmological simulations and galaxy models in the form of a light cone to create an all-sky galaxy catalog extending to $z < 0.35$ for the FAST HI survey. Two sets of mock HI catalogs were generated with the help of Alan Duffy (private communication). They were simulated with an integration time of 20 s and 120 s for FAST, corresponding to a sensitivity of 0.7 mJy and 0.3 mJy respectively according to Equations (1) and (2). A shallow survey using FAST with 20 s of integration per point would reach a sensitivity of 0.7 mJy, which is slightly deeper than that of the ALFALFA survey. Deeper surveys are required to detect more clusters at higher redshifts.

We run the FoF program to extract galaxy clusters from the mock HI catalog. The cluster's total HI mass is derived by summing up all the galaxies' HI masses in that cluster. We are aware of the fact that it is not easy to choose a set of appropriate linking lengths for

an HI galaxy catalog because HI galaxies tend to reside in the outer region of a cluster and some galaxies in the cluster cannot be detected by HI. To evaluate this problem, we first choose the linking length used by the SDSS Mr18 catalog (0.69 Mpc, 259 km s^{-1}) (Berlind et al. 2006). Then we try different sets of linking lengths (0.6 Mpc, 400 km s^{-1}; 0.7 Mpc, 500 km s^{-1}; 0.3 Mpc, 500 km s^{-1}) and found the extracted mock clusters are similar. Therefore, we choose the ones that are used by the SDSS Mr18 catalog (0.69 Mpc, 259 km s^{-1}) to extract HI mock clusters from the two sets of mock HI catalogs. The largest redshift of the clusters that can be extracted from the mock catalog galaxies is 0.08 because there are only a few HI galaxies at a farther distance and the distribution of those galaxies is too diffuse to extract clusters. There are 3985 mock clusters with 23 447 HI mock sources from the mock HI catalog with the integration time of 120 s and 1359 clusters with 6034 sources from the 20 s integration time catalog in the sky region identical to the control sample.

3.1.3 Comparison between the mock catalog and the SDSS catalog

Figure 2 shows a comparison between the SDSS selected clusters and the mock clusters, with number of member galaxies N greater than 3. The pink and blue bars in the upper panel are the distributions of the mock clusters as a function of cluster HI mass, which are the simulated outputs with 20 s and 120 s integration times respectively using FAST. In the lower panel of Figure 2, the pink bars show the distribution of SDSS clusters vs. HI masses, which is obtained in Section 3.1.1. We can see that the 20 s mock catalog in general looks quite similar to the SDSS-ALFALFA cross match sample, suggesting that numerical simulations can successfully reproduce the observed Universe. However, the mock catalogs (both the 20 s and 120 s ones) appear to have more groups at the low mass end. This could be due to the fact that the 20 s mock catalog has a root mean square (rms) of 0.7 mJy, which is lower than that of the ALFALFA survey (1.3 mJy), and thus it contains more low mass HI galaxies. As mentioned in Section 3.1.1, the ALFALFA HI detection rate of the SDSS member galaxies is only 22%. Thus we need to get estimates of the HI mass for a large number of SDSS galaxies not detected by ALFALFA. To do so, we use the relationship between HI gas content and optical color $g - r$ as described in Section 2.3. The $g - r$ color and absolute magnitude of each member galaxy are listed in the SDSS Mr18 cluster catalog, thus we can estimate the HI mass for the

un-matched SDSS member galaxies using Equations (5) and (6). In this way, we obtain more clusters (1443, instead of 977) and recover more HI mass for each cluster. The results are shown as blue bars in the lower panel of Figure 2. The distribution of this blue histogram agrees much better with that of the 20 s mock catalog. This suggests that a 20 s integration with FAST can achieve a much higher detection rate for nearby groups and clusters.

Longer integration times are desirable for detecting more member galaxies at higher redshifts. Hence we have also generated a mock catalog with 120 s of integration. It is remarkable that numerical simulations predict significantly more clusters in the 120 s mock catalog, as shown in the blue histogram in the upper panel of Figure 2. The 120 s mock catalog has many more HI low mass groups and clusters, compared to the control sample and the 20 s mock catalog. A plausible explanation is that SAM simulations predict a large amount of low mass HI galaxies which are optically too faint to be included in the SDSS Mr18 catalog (sources fainter than −18 mag are not included). To evaluate such an effect, we compute the minimum detectable HI mass at different distances using equation (5) of Giovanelli et al. (2005) assuming a velocity width of 200 km s^{-1} for a galaxy, but use 3σ here instead of 6σ in that equation. The 3σ detection limit of HI mass in ALFALFA with an integration time of 40 s is $2.78 \times 10^8 \, M_\odot$ at $z = 0.02$ and $1.11 \times 10^9 \, M_\odot$ at $z = 0.04$. Similarly, assuming a frequency resolution of 142 kHz (velocity width of 200 km s^{-1}), the HI mass that could be detected by FAST in 120 s of integration time is $2.33 \times 10^8 \, M_\odot$ at $z = 0.04$ ($5.83 \times 10^7 \, M_\odot$ at $z = 0.02$). Hence galaxies with HI mass in the range $2.78 \times 10^8 \, M_\odot < M_{\mathrm{HI}} < 1.11 \times 10^9 \, M_\odot$ at $z = 0.04$, for example, would be included in the 120 s mock catalog, but not in the SDSS-ALFALFA catalog.

We have further examined the stellar mass of galaxies in the 120 s mock catalog and found that the SAM simulations produce a surprisingly large amount of dwarf galaxies. Such an effect is similar to the "missing satellite" problem.

Figure 3 compares the stellar mass distributions of the mock and SDSS clusters. The upper panel of Figure 3 shows the cluster stellar mass distribution of the 120 s mock clusters and the lower panel shows the cluster stellar mass distribution of 1443 SDSS clusters. The stellar masses of the mock galaxies are generated from the SAM simulation. The stellar masses of the SDSS galaxies are calculated using Equation (5) with r band absolute magnitude and $g - r$ color which are listed in the SDSS member galaxy catalog. The cluster stellar mass is
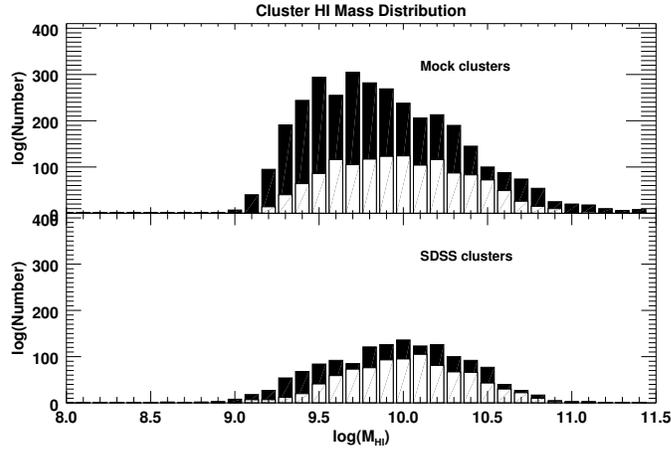
Fig. 2 The cluster HI mass distribution of two samples of clusters based on SDSS and mock galaxy catalogs. In the upper panel, the pink and blue bars represent the HI mass distribution of a mock cluster that is extracted from the 20 s and 120 s integration time catalog respectively. In the lower panel, the pink bars show the distribution of 977 SDSS clusters whose HI masses are derived from cross matching with the $\alpha.70$ catalog, and the blue bars are the distribution of all the 1443 clusters whose HI mass are derived from the cross matched HI sources plus the HI mass of the un-matched member galaxies estimated by galaxy color. See the text for details.
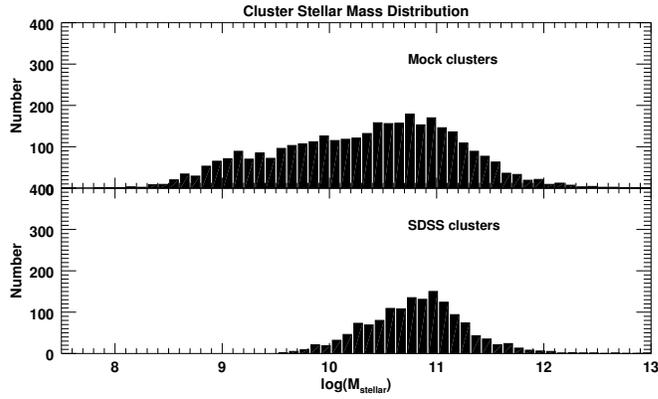
Fig. 3 Comparison of the stellar mass distribution of the SDSS clusters and the 120 s mock clusters. The upper panel is the cluster stellar mass distribution of the 120 s mock clusters and the lower panel is the distribution of the SDSS clusters.

the sum of all the member galaxy stellar masses in that cluster. It is obvious that there are too many low stellar mass galaxy groups in the 120 s mock catalog compared to the observed SDSS clusters. Many of the dwarf galaxies have an M_{HI}/M_* ratio higher than 1. Hence, the 120 s mock catalog predicts large amounts of low HI mass groups/clusters. To test such prediction, we will need to compare simulation data with observed data at both HI and optical wavelengths, with sensitivity high enough to recover all the member galaxies within the virial radius of a group or cluster. Future high sensitivity observations with FAST, as well as with optical telescopes, would be crucial for such studies and tackling the "missing satellite" problem.

3.2 Detecting High Redshift Clusters with FAST

In this part we discuss the predictions of the HI gas component in high redshift galaxy groups or clusters, which can be observed by FAST. The high-z mock catalogs of HI gas are from the outputs of L-Galaxies SAMs (Guo et al. 2011, 2013) based on both Millennium and Millennium II halos. We adopt the model versions by Fu et al. (2013) and Luo et al. (2016), which offer the results of both H_2 and HI gas in galaxies. In the models, two prescriptions are adopted to calculate the transition between H_2 and HI gas. In the Krumholz, McKee and Tumlinson (KMT) prescription (Krumholz et al. 2009), the fraction of H_2-to-HI is determined by gas surface

density and metallicity in the interstellar medium (ISM). In the Blitz and Rosolowsky (BR) prescription (Blitz & Rosolowsky 2006), the fraction of H_2-to-HI is related to the interstellar pressure (see details in sect. 2 of Fu et al. 2013). Based on model outputs, the mock catalogs are created with the light-cone algorithm by Blaizot et al. (2005) and emission line profiles by Obreschkow et al. (2009) for gas distribution on galaxy disks.

Because of the resolution limit of FAST, it is only possible to detect the integrated HI emission at high redshifts in galaxy groups or clusters. In the mock catalogs, two methods are used to identify the clustering properties of galaxies. In the first method, we define the galaxy clusters according to the merger history in the dark matter simulation, i.e. galaxies in the same subhalo are treated as a galaxy cluster. In the second method, we define the galaxy clusters in a similar way to observations, i.e. galaxies are in the same cluster if the velocity difference is lower than $1500\,\mathrm{km\,s^{-1}}$ or the redshift difference is lower than 0.005. Considering the beam size of FAST, we only add the HI mass within a galaxy cluster covered by the FAST beam of $2.95(1 + z)$ arcmin.

According to the detection limit of FAST for HI gas at redshift $z > 0.4$ in Table 1, we give the predictions for the number of groups or clusters in Table 3 and show the redshift evolution of the galaxy cluster in Figure 4. The clusters are selected with $M_{\mathrm{HI}} > 10^{11}\,M_{\odot}$ in the sky region with declination from $-10°$ to $60°$ at redshift $0.3 - 1.5$. As discussed above, the left and right panels represent the results of two different definitions of galaxy clusters in the models. In the right panel, the cluster numbers are a lot lower than those in the left panel when we use the velocity and redshift difference to define a galaxy cluster, which means many galaxies in different subhalos based on the halo merger history may be treated as one galaxy cluster or group in observations because of the projection effect and relatively low velocity difference for galaxies in a dense environment. In each panel, the red and green curves represent the model results with two different H_2-to-HI gas transition prescriptions. The model with the KMT prescription predicts more HI gas rich clusters than that of the BR prescription, which is consistent with the result that the KMT prescription gives higher HI-to-H_2 at high redshift in previous modeling work (Fu et al. 2012).

For the results at $z < 0.3$, models predict very few gas rich clusters with HI gas mass over $10^{11}\,M_{\odot}$. The KMT prescription predicts four clusters and the BR prescription predicts only two, which should be still a bit higher than in the real Universe, since Virgo is the only gas rich cluster in the nearby Universe. On the other hand, the models show that the cluster number increases a lot to about 10^5 at $z > 1$, which indicates that star formations and supernova feedbacks consume a large fraction of gas at $0.3 < z < 1$.

Figure 5 shows the redshift evolution of the size and HI mass of a Virgo-like cluster based on the SAM simulations. We can see that the HI mass increases by a factor of about 2–3 and the size decreases by a factor of about 2–3 from redshift $z = 0.3$ to $z = 1$. Hence we expect that many more HI fluxes will be included in the FAST beam at high redshift and the integration times listed in Table 1 could be reduced by a factor of 4–9.

4 OBSERVATION STRATEGY AND SAMPLE SELECTION FOR FUTURE FAST HI STUDIES

4.1 Targeted Observations of Selected Gas Rich Clusters

Although FAST has very high sensitivity, it still needs hours to detect a high z group or cluster. To make good use of telescope time, careful selection of targets is critical. We would like to select clusters that have the following properties:

a. optically selected clusters or groups that have blue colors;
b. located in an intermediate or high density environment.

The second criterion is based on the fact that on a large scale, HI follows the dark matter distribution and more HI gas is expected along cosmic web structures.

It has long been known that galaxy properties such as optical color, morphology and star formation rates are closely correlated with a galaxy's environment which is measured by the galaxy number density. For example, Dressler (1980) pointed out that the fractions of early type and S0 galaxies increase with an increasingly dense environment. Balogh et al. (1997) also found that the fraction of star forming galaxies is smaller in a cluster environment than in the field. Since spiral galaxies are normally gas rich, we expect that the HI gas fractions in blue star forming galaxies are high. Thus we need to select clusters dominated by star-forming galaxies which are likely to (still) contain significant HI masses, not having collapsed sufficiently to the state where ram-pressure stripping and dynamical interactions induce the HI deficiency seen in massive low redshift clusters like Coma. For example, Rakos & Schombert (1995) compiled a sample of 17 clusters in a study of the B-O effect, with redshifts ranging from 0 to 0.9. In this sample, the blue galaxy fraction increases from 20% at $z = 0.2$ to 80% at

Table 3 The Number of Galaxy Clusters in Different Redshift Bins Shown in Fig. 4

Redshift	0–0.3	0.3–0.6	0.6–0.9	0.9–1.2	1.2–1.5
		10^3	10^4	10^5	10^5
KMT (FoF)	175	26	36	8.2	8.6
BR (FoF)	55	14	23	6.1	6.3
KMT (velocity)	4	5.2	6.4	1.32	1.8
BR (velocity)	2	1.8	2.96	0.6	1.04
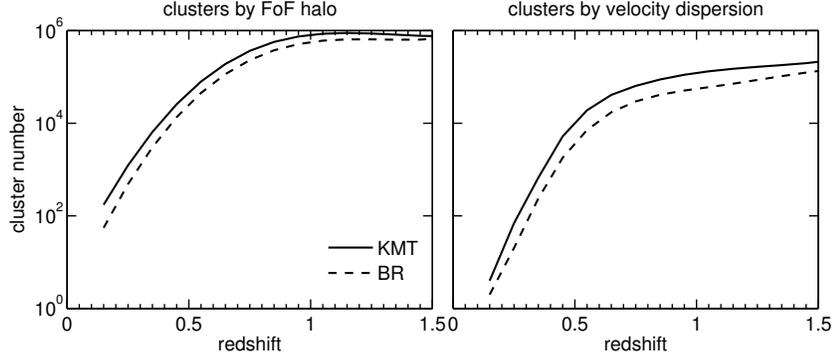
Fig. 4 The redshift evolution of galaxy cluster number with $M_{HI} > 10^{11}\, M_\odot$ based on the outputs of L-Galaxies SAMs in FAST's observable sky region. The left panel displays the results in which clusters are defined by halo merger history and the right panel shows the results in which clusters are defined by velocity and redshift difference. In each panel, the red and green curves represent model results from the KMT prescription and BR prescription respectively.
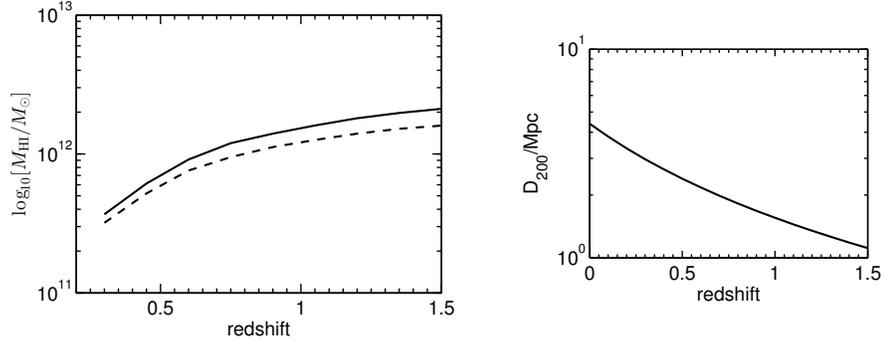
Fig. 5 The left panel is the redshift evolution of HI mass in the Virgo size cluster, in which the cluster is defined by the velocity difference. The two curves represent the results of two H_2-to-HI transition prescriptions in the models. The right panel displays the redshift evolution of size of a Virgo-like cluster, in which the cluster is still defined by the velocity dispersion with respect to the center. Because of the simple relation in the halo evolution, $r_{vir} = 0.1 H_0^{-1}(1+z)^{-3/2} v_c$, the redshift evolution between cluster size r_{200} and a fixed cluster velocity v_c is approximately proportional to $(1+z)^{-3/2}$ no matter what kind of physical prescription is used.

$z = 0.9$. Clusters such as CL1322.5+3027 ($z = 0.750$) and CL1622.5+2352 ($z = 0.927$) with a high fraction of blue galaxies could be good targets for future FAST deep HI observations.

Most clusters and big groups have $M_{HI}/M_v = 10^{-3}$ to 10^{-4}, thus for a Virgo type cluster, which has $M_v = 10^{15}\, M_\odot$, we will have $M_{HI} = 10^{11}\, M_\odot$ to $10^{12}\, M_\odot$. The extreme systems with $M_{HI} = 10^{12}\, M_\odot$, such as

those shown in Figure 5, should be detectable by FAST at $z \sim 1$, in 30–60 min.

For clusters with $z < 0.35$, the 19 beam array receiver can be used to scan the whole cluster region. When $z > 0.35$, only the $560\,MHz - 1120\,MHz$ single beam receiver is available at FAST. The system temperature of this receiver is expected to be about 20%–30% better than that of the 19-beam receiver, thus the integration

time could be reduced by about 40% compared to the values listed in Table 1. Clusters with z greater than 0.6–0.7 can be completely covered by the FAST beam, thus the single-beam receiver would be the best choice for observing high z clusters.

4.2 Blind Sky Surveys

As predicted by numerical simulations, a blind all sky HI survey using the 19 beam array receiver of FAST can detect a large number of groups and clusters with redshifts ranging from 0 to 0.35. Using drift scan mode (Qian et al. 2017 in preparation) and integrating 20 s per point, we can cover the sky area of $0^h < RA < 24^h, -10° < Dec < 60°$ in one or two years. However, deeper surveys are required to detect the major portion of member galaxies in clusters with $z > 0.4$. If we choose 120 s of integration time per point, it could take more than 10 years to survey the full FAST sky. Thus it is more practical to select particular regions for a deep blind sky survey. Higher priority should be given to regions that have high sensitivity data in optical wavelengths. The SDSS survey could miss many optically weak but HI bright galaxies at redshift $z > 0.04$, as shown in Section 3.1.3. Future surveys from the LSST may be able to provide data complementary to the FAST HI survey.

A multibeam receiver is also useful in searching for HI in clusters over a large part of the sky at intermediate redshifts. For example, if a seven-beam array receiver operating at 500–1000 MHz can be developed for FAST, a deep blind survey of HI clusters would become feasible. SAM simulations predict that there are about $10^4 - 10^5$ clusters in the sky region of $-10° < Dec < 60°$. Even if we blindly scan 1% of the sky region repeatedly, with a typical integration time of 30–60 min per point, we could detect about 100–1000 HI clusters with $z > 0.5$, and significantly improve the sample for studying the HI evolution effect in clusters.

5 SUMMARY AND CONCLUSIONS

In summary, we have used HI data from the ALFALFA survey to derive the integrated HI emission of the Virgo Cluster. We further use Virgo as an example to evaluate the feasibility of applying FAST to detect integrated HI emission at high redshifts. We have estimated the integration time to detect a Virgo type cluster at different redshifts. At a redshift of 0.7, a Virgo sized cluster could be completely covered by the FAST beam and the integration time to detect such a cluster at this redshift by FAST is about 5 hours, assuming no evolutionary effect for the HI contents. We have also employed a SAM model to simulate the evolution of HI contents in galaxy clusters. Our simulations suggest that the HI mass of a Virgo-like cluster could be 2–3 times higher and the physical size could be more than 50% smaller from $z = 0.3$ to $z = 1$. Thus the integration time could be reduced by a factor of $4 - 9$ for a cluster at $z \sim 1$. The well known B-O effect also suggests that clusters at higher redshifts contain more spiral galaxies than local clusters. Hence we conclude that it is feasible for FAST to detect the integrated HI emission at redshifts around 1. Our SAM simulations suggest that there are $10^4 - 10^5$ clusters with HI mass greater than $10^{11} M_\odot$ in the sky region visible to FAST.

For the local Universe, we have also used SAM simulations to create mock catalogs of clusters to predict the outcomes from the FAST all sky surveys. Comparing with the optically selected catalogs derived by cross matching the galaxy catalogs from the SDSS survey and the ALFALFA survey, we find that the HI mass distribution of the mock catalog with 20 s integration time agrees well with that of observations. However, the mock catalog with 120 s integration time predicts many more groups and clusters that contain a population of low mass HI galaxies not detected by the ALFALFA survey. Future deep HI blind sky surveys with FAST would be able to test such prediction and set constraints on numerical simulation models.

References

Balogh, M. L., Morris, S. L., Yee, H. K. C., Carlberg, R. G., & Ellingson, E. 1997, ApJ, 488, L75

Bell, E. F., McIntosh, D. H., Katz, N., & Weinberg, M. D. 2003, ApJS, 149, 289

Berlind, A. A., Frieman, J., Weinberg, D. H., et al. 2006, ApJS, 167, 1

Blaizot, J., Wadadekar, Y., Guiderdoni, B., et al. 2005, MNRAS, 360, 159

Blitz, L., & Rosolowsky, E. 2006, ApJ, 650, 933

Butcher, H., & Oemler, Jr., A. 1978, ApJ, 219, 18

Butcher, H., & Oemler, Jr., A. 1984, ApJ, 285, 426

Catinella, B., & Cortese, L. 2015, MNRAS, 446, 3526

Catinella, B., Haynes, M. P., Giovanelli, R., Gardner, J. P., & Connolly, A. J. 2008, ApJ, 685, L13

Chung, A., van Gorkom, J. H., Kenney, J. D. P., Crowl, H., & Vollmer, B. 2009, AJ, 138, 1741

Croton, D. J. 2006, MNRAS, 369, 1808

Dressler, A. 1980, ApJS, 42, 565

Duffy, A. R., Meyer, M. J., Staveley-Smith, L., et al. 2012, MNRAS, 426, 3385

English, J., Koribalski, B., Bland-Hawthorn, J., Freeman, K. C., & McCain, C. F. 2010, AJ, 139, 102

Fernández, X., Gim, H. B., van Gorkom, J. H., et al. 2016, ApJ, 824, L1

Freudling, W. 1989, in BAAS, 21, Bulletin of the American Astronomical Society, 1140

Fu, J., Kauffmann, G., Li, C., & Guo, Q. 2012, MNRAS, 424, 2701

Fu, J., Kauffmann, G., Huang, M.-l., et al. 2013, MNRAS, 434, 1531

Giovanelli, R., & Haynes, M. P. 1985, ApJ, 292, 404

Giovanelli, R., Haynes, M. P., Kent, B. R., et al. 2005, AJ, 130, 2598

Gunn, J. E., & Gott, III, J. R. 1972, ApJ, 176, 1

Guo, Q., White, S., Angulo, R. E., et al. 2013, MNRAS, 428, 1351

Guo, Q., White, S., Boylan-Kolchin, M., et al. 2011, MNRAS, 413, 101

Haynes, M. P., Giovanelli, R., Martin, A. M., et al. 2011, AJ, 142, 170

Jaffé, Y. L., Poggianti, B. M., Verheijen, M. A. W., Deshev, B. Z., & van Gorkom, J. H. 2013, MNRAS, 431, 2111

Krumholz, M. R., McKee, C. F., & Tumlinson, J. 2009, ApJ, 693, 216

Li, D., & Pan, Z. 2016, Radio Science, 51, 1060

Luo, Y., Kang, X., Kauffmann, G., & Fu, J. 2016, MNRAS, 458, 366

Merritt, D. 1983, ApJ, 264, 24

Nan, R., Li, D., Jin, C., et al. 2011, International Journal of Modern Physics D, 20, 989

Nulsen, P. E. J. 1982, MNRAS, 198, 1007

Obreschkow, D., Klöckner, H.-R., Heywood, I., Levrier, F., & Rawlings, S. 2009, ApJ, 703, 1890

Rakos, K. D., & Schombert, J. M. 1995, ApJ, 439, 47

Rhee, J., Lah, P., Chengalur, J. N., Briggs, F. H., & Colless, M. 2016, MNRAS, 460, 2675

Sengupta, C., & Balasubramanyam, R. 2006, MNRAS, 369, 360

Solanes, J. M., Manrique, A., García-Gómez, C., et al. 2001, ApJ, 548, 97

Springel, V., White, S. D. M., Jenkins, A., et al. 2005, Nature, 435, 629

Stevens, J. B., Webster, R. L., Barnes, D. G., Pisano, D. J., & Drinkwater, M. J. 2004, PASA, 21, 318

Verheijen, M., van Gorkom, J. H., Szomoru, A., et al. 2007, ApJ, 668, L9

Wen, Z. L., Han, J. L., & Liu, F. S. 2012, ApJS, 199, 34

Wen, Z. L., & Han, J. L. 2015, ApJ, 807, 178

White, S. D. M., & Frenk, C. S. 1991, ApJ, 379, 52

Wolfinger, K., Kilborn, V. A., Koribalski, B. S., et al. 2013, MNRAS, 428, 1790

Zhang, W., Li, C., Kauffmann, G., et al. 2009, MNRAS, 397, 1243