Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY LETTERS

Mon. Not. R. Astron. Soc. **412,** L50–L54 (2011) doi:10.1111/j.1745-3933.2010.01003.x

Clustering of H I galaxies in the H I Parkes All Sky Survey and Arecibo Legacy Fast ALFA Survey

S. S. Passmoor,[1]⋆ C. M. Cress[1,2] and A. Faltenbacher[1]

[1]*Physics Department, University of the Western Cape, Bellville 7535, South Africa*
[2]*Centre for High Performance Computing, 15 Lower Hope Street, Cape Town 7700, South Africa*

Accepted 2010 December 17. Received 2010 December 15; in original form 2010 September 14

ABSTRACT

We investigate the clustering of H I-selected galaxies in the Arecibo Legacy Fast ALFA Survey (ALFALFA) and compare results with those obtained for the H I Parkes All Sky Survey (HIPASS). Measurements of the angular correlation function and the inferred 3D clustering are compared with results from direct spatial-correlation measurements. We are able to measure clustering on smaller angular scales and for galaxies with lower H I masses than was previously possible. We calculate the expected clustering of dark matter using the redshift distributions of the HIPASS and ALFALFA, and show that the ALFALFA sample is somewhat more antibiased with respect to dark matter than the HIPASS sample.

Key words: large-scale structure of Universe – radio lines: galaxies.

1 INTRODUCTION

Measurements of the clustering of galaxies allows one to investigate the relationship between dark and luminous matter. By comparing galaxies selected in different ways, one gains understanding of how different galaxies trace the underlying dark matter and also of processes at work in galaxy evolution. This information is important when using galaxies as probes of cosmological parameters.

A number of new radio telescopes, such as the MeerKAT,[1] ASKAP[2] and the SKA, are in the pipeline and they will detect huge numbers of galaxies using H I. A reliable measure of the bias of H I-selected galaxies and insight into the evolution of the bias is important for forecasting the capabilities of telescopes which will probe H I at intermediate or high redshifts. The clustering of H I-selected galaxies has been studied by Meyer et al. (2007), Basilakos et al. (2007) and Ryan-Weber (2006). They used data from the H I Parkes All Sky Survey (HIPASS, Meyer et al. 2004), a blind survey for H I of the southern sky which generated a catalogue of 4315 sources, the bulk of which have redshifts below $z \sim 0.02$. They showed that H I-selected galaxies are less clustered than galaxies selected in other ways. Meyer et al. (2007) investigated clustering of various subsamples of HIPASS galaxies, showing that galaxies with high rotation velocities are more clustered than those with lower rotation velocities. There were indications that galaxies containing more H I are also more clustered, but the differences were not as pronounced as in Basilakos et al. (2007). The latter work also measures the bias of HIPASS galaxies relative to the expected dark matter distribution.

In this Letter, we measure the clustering of H I-selected galaxies detected with the Arecibo L-band Feed Array (ALFA) and compiled in the partially completed ALFALFA (Arecibo Legacy Fast ALFA Survey, Giovanelli et al. 2005). The results are compared with those obtained for the HIPASS. Clustering measurements in the HIPASS are limited to large angular scales where the beam-size of ∼15 arcmin does not cause confusion. The ALFALFA resolution is more than four times better, allowing us to probe clustering on smaller scales. The rms noise per ALFALFA beam is about six times smaller, providing a catalogue of sources which spans a wider range of redshifts and includes galaxies with lower H I masses. We are thus able to measure clustering of H I-selected galaxies in regimes that have not yet been explored and to investigate trends seen in the HIPASS, using an independent survey.

The outline of this Letter is as follows. In Section 2, we give a short introduction to the HIPASS and ALFALFA. The computation of the angular and spatial two-point correlation functions is described in Section 3. The results are presented, discussed and compared with earlier work in Section 4. Finally, Section 5 concludes with a short summary.

2 DATA

The HIPASS covers all the southern sky with $\delta < +2^\circ$ and can detect H I with velocities in the range 300–12 700 km s^{-1}. The rms noise per beam is ∼13 mJy (Meyer et al. 2004). To exclude structure associated with the Milky Way, like high-velocity clouds and low-mass satellites, we use only sources with recessional velocities larger than 600 km s^{-1}. The average mass of H I in HIPASS galaxies is 3.24×10^9 M$_\odot$.

⋆E-mail: sean.passmoor@gmail.com
[1] www.ska.ac.za
[2] www.atnf.csiro.au/SKA/

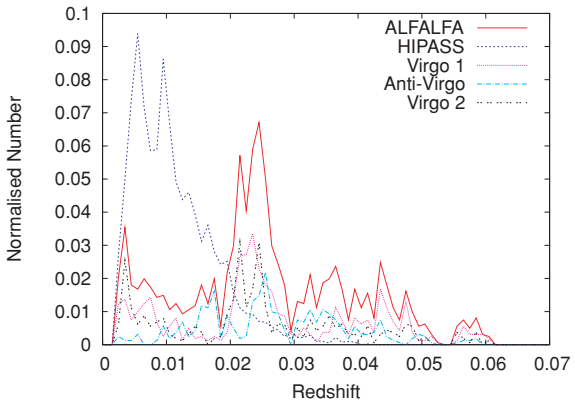

Figure 1. Plot of the normalized redshift distribution of the HIPASS and ALFALFA as well as the distribution of the ALFALFA sources in the three, spatially separated strips.

When completed, the ALFALFA (Giovanelli et al. 2005) will cover 7000 deg^2 of sky with high Galactic latitude and to a depth of $cz \sim 18\,000$ km s^{-1}. The rms noise of the survey is ∼2.2 mJy and the beam-size is ∼3.6 arcmin. Currently, the ALFALFA contains three strips covering a total area of ∼400 deg^2. These are the two strips centred on the Virgo region and the anti-Virgo strip. They contain 1796 sources with $cz > 600$ km s^{-1}. The first completed Virgo strip is defined by $11^{\rm h}44^{\rm m} < \alpha < 14^{\rm h}00^{\rm m}$ and $12^\circ < \delta < 16^\circ$ and contains 708 sources. The second Virgo strip contains 556 galaxies within $11^{\rm h}36^{\rm m} < \alpha < 13^{\rm h}52^{\rm m}$ and $8^\circ < \delta < 12^\circ$ (Kent et al. 2008). The anti-Virgo strip contains 488 sources within $22^{\rm h}00^{\rm m} < \alpha < 03^{\rm h}04^{\rm m}$ and $26^\circ < \delta < 28^\circ$ (Saintonge et al. 2008). A circular region, with radius of 1° centred on M87, has been removed from the survey area due to the interference of M87 (Giovanelli et al. 2007). The average ALFALFA H I mass is 2.48×10^9 M$_\odot$. The source density of the ALFALFA catalogue is approximately 20 times higher than that of the HIPASS.

Fig. 1 displays the normalized redshift distributions of the HIPASS and ALFALFA as well as the redshift distributions in the three separate ALFALFA regions. A redshift of $z = 0.02$ corresponds to $\sim 70\,h^{-1}$ Mpc which is roughly the distance to the Coma supercluster. The high galaxy density near the Coma cluster and Virgo cluster is evident in the redshift distributions shown for the Virgo regions. In the anti-Virgo region, the effect of the Perseus–Pisces supercluster can also be seen as a slight enhancement of galaxies at a redshift of about $z \approx 0.025$. Below, we will discuss the impact of these inhomogeneities on the determination of the two-point correlation function.

3 TWO-POINT CORRELATION FUNCTIONS

Here we review some basic properties of the angular and projected two-point correlation functions (ω and Ξ, respectively) and indicate their relations to the 3D real-space two-point correlation function, ξ. Subsequently, we introduce the estimator used here and discuss the construction of the random samples. We do not employ the weighted correlation functions used in Meyer et al. (2007) as we are interested in comparing the results of the two surveys and in comparing our results with those predicted for dark matter within a Λ cold dark matter (ΛCDM) model. The unweighted measurements suffice for this work and we are able to check our unweighted results against those of Meyer et al. (2007).

3.1 The angular two-point correlation function

The angular correlation function, ω, is a simple measure of the clustering of galaxies as a function of angular separation on the sky, θ, which does not require redshift information. It is calculated by counting galaxy pairs within a given angular separation bin and comparing this number to a corresponding figure derived from a random catalogue with the same area and shape. The angular correlation, $\omega(\theta)$, gives the excess probability, over random, of finding two galaxies separated by angle θ.

If we assume a redshift-dependent power law describes the 3D real-space correlation function, $\xi(r, z) = (r/r_0)^{-\gamma}\,(1+z)^{\gamma-(3+\epsilon)}$ (as in Peebles 1980 and Loan, Wall & Lahav 1997), then the angular correlation function is related to the spatial correlation function by the Limber equation (Limber 1954; Rubin 1954):

$$\left(\frac{\theta}{\theta_0}\right)^{1-\gamma} = \frac{\int_0^\infty \frac{N^2(z)(1+z)^{\gamma-(3+\epsilon)}\sqrt{\pi}[{\rm d}(z)\theta]^{1-\gamma}}{{\rm d}'(z) r_0^{-\gamma}} \frac{\Gamma\left(-\frac{1}{2}+\frac{1}{2}\gamma\right)}{\Gamma\left(\frac{1}{2}\gamma\right)}}{\left[\int_0^\infty N(z){\rm d}z\right]^2}\,, \qquad (1)$$

where d (z) is the comoving distance and $N(z)$ is the redshift number density distribution of the sources (cf. Fig. 1). We use $\epsilon = 0.8$, consistent with the expected clustering behaviour in linear theory, although the surveys are so shallow that the evolution of ξ could be ignored. The measured values for the logarithmic slope $a_\theta = 1 - \gamma$ and the correlation length θ_0 [$\omega(\theta_0) = 1$] can then be used to determine the 3D parameters r_0 and γ.

The errors for ω are calculated using jack-knife re-sampling (Lupton 1993). For this purpose, the data are split up into N right ascension (RA) bins and the correlation function is recalculated repeatedly each time leaving out a different bin. Thus, a set of N values $\{\omega_i, i = 1, \ldots, N\}$ for the correlation function are obtained and the jack-knife error of the mean, $\sigma_{\omega_{\rm mean}}$, is given by

$$\sigma_{\omega_{\rm mean}} = \sqrt{(N-1)\sum_{i=1}^{N}(\omega_i - \omega)^2/N}\,. \qquad (2)$$

The HIPASS sample has been divided into 24 RA bins, while for the ALFALFA catalogue, we use 12 bins such that each bin contains approximately the same area of the sky.

3.2 The projected two-point correlation function

The projected correlation function, $\Xi(\sigma)$, is determined by the number of pairs at given radial and projected separations, π and σ, and a subsequent integration along the radial direction. For that purpose, the absolute radial distance between a pair of galaxies, $\pi = |(v_i - v_j)/H_0|$, and their angular separation, θ, are converted into a projected distance, $\sigma = [(v_i + v_j)/H_0] \tan(\theta/2)$. Thus,

$$\frac{\Xi(\sigma)}{\sigma} = \frac{2}{\sigma}\int_0^{D_{\rm limit}} \xi(\sigma, \pi){\rm d}\pi, \qquad (3)$$

where $D_{\rm limit}$ is the limit where the integral converges. Here, we set $D_{\rm limit} = 25\,h^{-1}$ Mpc ≈ 2500 km s^{-1}. The projected correlation function is related to the real-space correlation function by (e.g. Davis & Peebles 1983):

$$\frac{\Xi(\sigma)}{\sigma} = \frac{2}{\sigma}\int_\sigma^\infty \xi(r)\frac{r{\rm d}r}{(r^2-\sigma^2)^{1/2}}\,. \qquad (4)$$

Assuming that the projected and real-space correlation functions follow power laws within the region of interest ($r < 10\,h^{-1}$ Mpc), the parameter for the real-space correlation function can be derived

from the projected one by the following expression:

$$\frac{\Xi(\sigma)}{\sigma} = \left(\frac{r_0}{\sigma}\right)^{\gamma} \frac{\Gamma(1/2)\,\Gamma((\gamma-1)/2)}{\Gamma(\gamma/2)}. \quad (5)$$

More specifically, we calculate r_0 and γ by fitting a power law, $\Xi(\sigma)/\sigma = (\sigma/\sigma_0)^{-a_\sigma}$, using the Levenberg–Marquardt non-linear least-squares method. The parameters of the real-space correlation function are then given by

$$r_0 = \sigma_0 \left[\frac{\Gamma(1/2)\,\Gamma((a_\sigma-1)/2)}{\Gamma(a_\sigma/2)}\right]^{-\frac{1}{a_\sigma}}, \quad (6)$$

$$\gamma = a_\sigma. \quad (7)$$

Therefore, similar to the angular correlation function, the projected correlation function can be used to determine the real-space clustering. We apply both methods to determine the real-space clustering strength based on the HIPASS and ALFALFA, and compare the results.

3.3 Estimator and random sampling

Three different estimators are commonly used to determine the two-point correlation function (Davis & Huchra 1982; Hamilton 1993; Landy & Szalay 1993). In this work, we use the Landy & Szalay (1993) estimator as it reduces errors caused by edges and holes within a given catalogue. In particular, this is important for the ALFALFA with the hole caused by M87 (Giovanelli et al. 2007) and the large edge effects due to the three strips. The estimator is of the form

$$\xi(\boldsymbol{r}) = \frac{\mathrm{DD}(\boldsymbol{r}) - 2\mathrm{DR}(\boldsymbol{r}) + \mathrm{RR}(\boldsymbol{r})}{\mathrm{RR}(\boldsymbol{r})}, \quad (8)$$

where $\boldsymbol{r}$ is the separation distance which has different meanings for the different correlation functions. For the angular correlation, it denotes the separation angle, θ, for the projected correlation function, it is the projected distance, σ, and for the real-space correlation, it indicates the real-space distance r. DD($\boldsymbol{r}$) is the number of data–data pairs, DR($\boldsymbol{r}$) is the number of data–random pairs and RR($\boldsymbol{r}$) is the number of random–random pairs all with separations $\boldsymbol{r}$.

The random catalogues were generated with uniform distributions on the sky and redshift distributions which resemble the distribution of recessional velocities in the survey smoothed using kernel density estimation (Wand & Jones 1995). Throughout this work, we use random samples that are equal in size compared to the corresponding data set. We repeat the random catalogue generation 20 times and the calculation of random pairs in order to reduce the variance from the random sampling.

3.4 The angular correlation function of dark matter

Based on Limber's equation, the redshift distributions of the HIPASS and ALFALFA, and the expression for the non-linear power spectrum discussed in Peacock & Dodds (1996), we predict the angular correlation function of dark matter using the cosmological parameters given in Komatsu et al. (2009). The bias parameter, b, at various angles is then determined by

$$b = \sqrt{\frac{\omega_{\mathrm{H\,I}}}{\omega_{\mathrm{dark\ matter}}}}. \quad (9)$$

4 RESULTS

4.1 The full HIPASS and ALFALFA samples

4.1.1 Angular correlation functions

Fig. 2 shows the angular correlation functions for HIPASS and ALFALFA data as well as the predicted correlation functions of CDM weighted with the redshift distributions of the surveys. The straight lines are power-law fits for pair separations in the range $0.^\circ1$–8° for the ALFALFA and 1°–10° for the HIPASS. The effect of source confusion in the HIPASS is evident at smaller angular scales. The corresponding parameters, θ_0 and a_θ, are given in Table 1.

The measured slopes in the two surveys agree reasonably well. As expected, the value of θ_0 is lower for the ALFALFA, since it is deeper than the HIPASS and the clustering in 3D is washed out in the 2D projection. The ALFALFA also detects galaxies with lower H I masses, which are potentially less clustered.

4.1.2 Projected correlation functions

Fig. 3 shows the projected correlation functions for the HIPASS and ALFALFA. The lines represent power-law fits and the corresponding parameters are presented in Table 2. Once again the slopes, a_σ, agree well, while the amplitude of clustering, σ_0, in the ALFALFA is lower (although the uncertainties are fairly large).

4.1.3 Inferred spatial correlations

Table 3 shows the spatial correlation function parameters inferred from the angular and projected correlation functions obtained using equations (1) and (6). The subscripts, θ and σ, indicate which

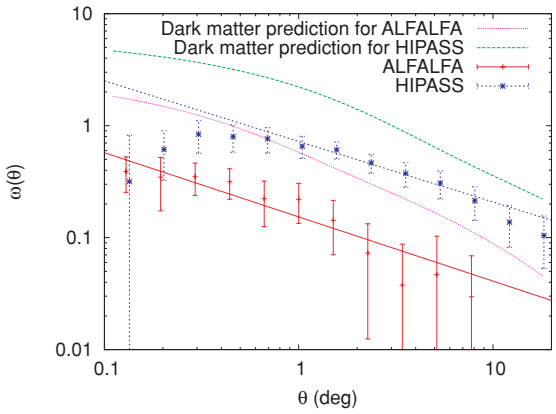

Figure 2. Angular correlation functions for the HIPASS and ALFALFA. Error bars were calculated using jack-knife sampling. The solid red line and the small-dashed blue line show the corresponding power-law fits for angles in the range $0.^\circ1$–8° for the ALFALFA and 1°–10° for the HIPASS data. The projected clustering of dark matter (in a ΛCDM model) with redshift distributions of the HIPASS and ALFALFA is shown by the green dashed line and the magenta dotted line, respectively.

Table 1. The angular clustering fitted parameters, θ_0 and a_θ.

	HIPASS	ALFALFA
θ_0	$0.^\circ603 \pm 0.^\circ04$	$0.^\circ044 \pm 0.^\circ013$
a_θ	0.56 ± 0.02	0.59 ± 0.06

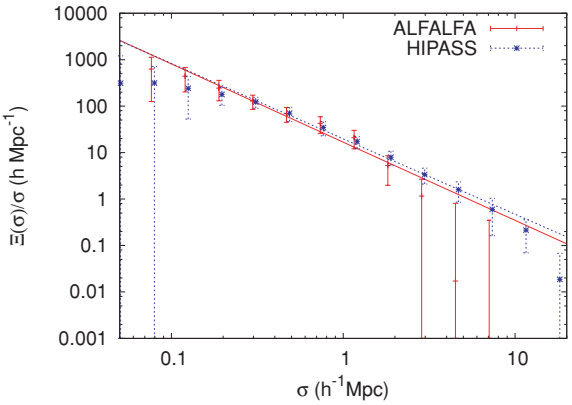

Figure 3. Projected correlation functions for the HIPASS and ALFALFA. Error bars are calculated using jack-knife sampling. Lines show the corresponding power-law fits for separations between 0.1 and 3.5 h^{-1} Mpc for the ALFALFA and 0.2 and 8 h^{-1} Mpc for the HIPASS data.

Table 2. The projected clustering fitted parameters, σ_0 and a_σ.

	HIPASS (h^{-1} Mpc)	ALFALFA (h^{-1} Mpc)
σ_0	6.29 ± 0.36	5.34 ± 1.08
a_σ	1.62 ± 0.04	1.68 ± 0.13

Table 3. The real-space clustering parameters, r_0 and γ, derived from the angular correlation function (indicated by subscript θ) and from the projected correlation function (indicated by subscript σ).

	HIPASS (h^{-1} Mpc)	ALFALFA (h^{-1} Mpc)
$r_{0,\theta}$	2.89 ± 0.08	2.00 ± 0.40
γ_θ	1.56 ± 0.02	1.59 ± 0.06
$r_{0,\sigma}$	2.51 ± 0.20	2.30 ± 0.53
γ_σ	1.62 ± 0.04	1.68 ± 0.13
	$M_{\rm H\,I} < 10^{9.25}\,h^{-2}\,{\rm M}_\odot$	
$r_{0,\sigma}$	2.26 ± 0.36	2.48 ± 0.69
γ_σ	1.60 ± 0.08	1.59 ± 0.13
	$M_{\rm H\,I} > 10^{9.25}\,h^{-2}\,{\rm M}_\odot$	
$r_{0,\sigma}$	3.32 ± 0.55	2.04 ± 0.65
γ_σ	1.50 ± 0.08	1.74 ± 0.24

correlation function has been used to derive these parameters. The two values obtained for r_0 in the HIPASS are within 2σ of each other and agree well with the unweighted value of 2.7 obtained by Meyer et al. (2004). The two ALFALFA values are consistent with each other and indicate somewhat lower clustering than the HIPASS.

4.1.4 Bias estimation

The predicted angular correlation function of dark matter is compared with our results in Fig. 2. We have calculated the bias for the two surveys at each data point in the plot. For the HIPASS, in the 1°–10° range, we find bias values ranging from 0.54 to 0.70, with an average of 0.63. This is fairly consistent with the value of 0.68 obtained by Basilakos et al. (2007). For the ALFALFA, on the same angular scales, the values range from 0.41 to 0.62, with an average of 0.52. Our results thus indicate that the ALFALFA sample is somewhat more antibiased than the HIPASS sample. This is consistent with the idea that the ALFALFA includes galaxies with lower H I mass, which are less clustered than the higher mass galaxies detected in the HIPASS. We note, however, that the lower values are found at large scales where the narrowness of the strips may affect measurements more severely.

4.2 Correlation functions of different ALFALFA subsamples

4.2.1 Flux and H I mass subsets

The ALFALFA data has been subdivided into two equivalent parts based on the flux of the sources. For these subsamples, the angular and projected correlation functions were recalculated as described in Section 3. In agreement with Meyer et al. (2007), we find that the two correlation functions compare well with each other and with the correlation function of the whole data set, indicating a negligible dependence of clustering on H I flux.

We also split the samples evenly into high- and low-H I-mass subsamples. The clustering parameters obtained are shown in Table 3. Our results for the HIPASS are consistent with those of Meyer et al. (2007), indicating that the galaxies with higher H I masses are more clustered. Interestingly, the same trend is not apparent in the ALFALFA, but the uncertainties are fairly large. We did not attempt to separate the galaxies according to their rotation velocities as this requires additional data to estimate inclinations.

4.2.2 Small field effects

The Virgo regions contain overdensities of galaxies that are associated with the Virgo and Coma clusters. There is the concern that the results will be biased by the presence of such dominant large-scale structure within the relatively small survey fields. To investigate this, the correlation functions of the three regions were calculated separately and are shown in Fig. 4. Measurements in the three regions agree to within their uncertainties, indicating that the presence of the big clusters within the Virgo regions does not affect

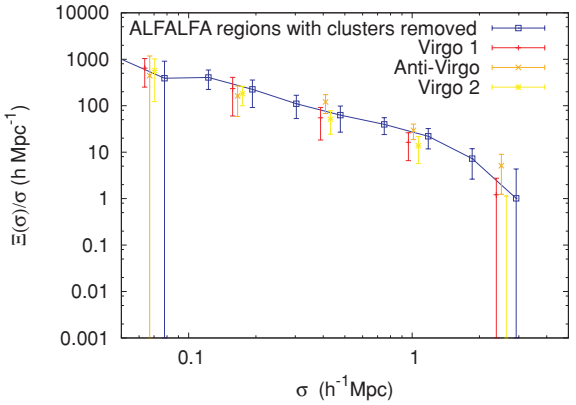

Figure 4. Comparison of the projected correlation functions of the three strips of the current ALFALFA as well as the correlation function measured when ALFALFA data between $0.02 < z < 0.03$ are excluded (i.e. when cluster galaxies in the Coma and Perseus–Pisces regions are excluded). The data points for the three strips are slightly offset to improve readability. Error bars are calculated using jack-knife sampling.

the results significantly. We note, however, that the anti-Virgo region is near the Perseus–Pisces supercluster which causes a slight overdensity in that field at a similar redshift ($z \approx 0.025$). To be sure that overdensities in all three fields at this redshift were not biasing our results, we cut the galaxies with redshifts between ∼0.02 and 0.03 out of the samples and recalculated the correlation functions. The results are also shown in Fig. 4 and it is clear that the correlation functions with and without the redshift cuts are completely consistent within the uncertainties.

As an additional check on the effect of the small fields on the measured clustering strength, we calculated the integral constraint (Peebles 1980; Ratcliffe, Shanks & Parker 1998) for a single field and obtained a value of 0.145 which indicates an effect within the uncertainty of the correlation function derived from the ALFALFA data.

5 SUMMARY AND CONCLUSIONS

We have measured the clustering of H I-selected galaxies using the ALFALFA data and compared this with results for the HIPASS. Our two methods for determining the real-space correlation function agree well and our results for the HIPASS agree with those found by Meyer et al. (2004). The real-space clustering in the ALFALFA appears to be even lower than in the HIPASS, consistent with the idea that the ALFALFA probes galaxies with lower H I masses that are less clustered than their high-mass counterparts. Our measurements of high- and low-mass subsamples in the ALFALFA do not provide evidence to support this idea, but the uncertainties on the measurements are large.

We have calculated the clustering of dark matter expected within a ΛCDM model with redshift distributions of the HIPASS and ALFALFA. We then calculated the bias of ALFALFA sources over the range 1°–10°, finding a value of 0.62 at 1° and an average value of 0.52 over the whole range. The significant antibias of galaxies with low H I mass is important to consider when estimating the signal-to-noise ratio of experiments planned for the SKA and its pathfinders.

ACKNOWLEDGMENTS

We thank the South African Square Kilometre Array Project, National Research Foundation and Centre for High Performance Computing for support. We also thank the referee for helpful comments.

REFERENCES

Basilakos S., Plionis M., Kovač K., Voglis N., 2007, MNRAS, 378, 301
Davis M., Huchra J., 1982, ApJ, 254, 437
Davis M., Peebles P. J. E., 1983, ApJ, 267, 465
Giovanelli R. et al., 2005, AJ, 130, 2598
Giovanelli R. et al., 2007, AJ, 133, 2569
Hamilton A. J. S., 1993, ApJ, 417, 19
Kent B. R. et al., 2008, AJ, 136, 713
Komatsu E. et al., 2009, ApJS, 180, 330
Landy S. D., Szalay A. S., 1993, ApJ, 412, 64
Limber D. N., 1954, ApJ, 119, 655
Loan A. J., Wall J. V., Lahav O., 1997, MNRAS, 286, 994
Lupton R., 1993, Statistics in Theory and Practice. Princeton Univ. Press, Princeton, NJ
Meyer M. J. et al., 2004, MNRAS, 350, 1195
Meyer M. J., Zwaan M. A., Webster R. L., Brown M. J. I., Staveley-Smith L., 2007, ApJ, 654, 702
Peacock J. A., Dodds S. J., 1996, MNRAS, 280, L19
Peebles P., 1980, The Large-Scale Structure of the Universe. Princeton Univ. Press, Princeton, NJ
Ratcliffe A., Shanks T., Parker Q. A., Fong R., 1998, MNRAS, 296, 173
Rubin V. C., 1954, Proc. Natl. Acad. Sci. USA, 40, 541
Ryan-Weber E. V., 2006, MNRAS, 367, 1251
Saintonge A., Giovanelli R., Haynes M. P., Hoffman G. L., Kent B. R., Martin A. M., Stierwalt S., Brosch N., 2008, AJ, 135, 588
Wand M., Jones M., 1995, Kernel Smoothing. Chapman and Hall/CRC, New York

This paper has been typeset from a TEX/LATEX file prepared by the author.