

A MeerKAT study of a neutral hydrogen rich grouping of galaxies with megaparsec-scale filamentary-like structure

G. D. Lawrie^{1,*}, R. P. Deane^{1,2}, and R. Dave^{3,4,5}

¹ Wits Centre for Astrophysics, School of Physics, University of the Witwatersrand, 1 Jan Smuts Avenue, Johannesburg, 2000, South Africa

² Department of Physics, University of Pretoria, Private Bag X20, Pretoria 0028, South Africa

³ Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK

⁴ University of the Western Cape, Department of Physics and Astronomy, Bellville, Cape Town 7535, South Africa

⁵ South African Astronomical Observatory, Observatory, Cape Town 7925, South Africa

Received 28 May 2025 / Accepted 18 July 2025

ABSTRACT

Context. Environmental effects within cosmological overdensities, such as galaxy groups and clusters, have been shown to impact galaxies and their cold gas reservoirs and thereby provide constraints on galaxy evolution models. Galaxy groups foster frequent galaxy-galaxy interactions, making them rich environments in which to study galaxy transformation.

Aims. In this work, we study a serendipitously discovered large overdensity of neutral hydrogen (HI) galaxies at $z \sim 0.04$. The galaxies appear to lie in a filamentary-like structure of megaparsec scale. Using MeerKAT's angular resolution and field of view, we were able to spatially resolve the HI galaxies while simultaneously probing large-scale structure.

Methods. The HI and sub-arcsec Dark Energy Survey (DES) imaging have revealed a large number of both interacting and disturbed galaxies in this collective group. MeerKAT data enabled us to derive HI masses and investigate interacting galaxies. We used DES and Wide-field Infrared Survey Explorer (WISE) data to quantify the star formation rates, stellar masses, and stellar morphologies of member galaxies and compared these with field scaling relations. To place this discovery and the environmental effects in context, we used the SIMBA cosmological hydrodynamical simulation to investigate the prevalence of qualitatively similar HI overdensities and their large-scale morphological properties. This enabled us to make a prediction of how frequently such structures might be serendipitously discovered with MeerKAT and SKA-Mid HI observations in comparable observation time.

Results. The combination of spatially resolved HI data and optical imaging revealed a group rich in interactions, suggesting environmental processes are already shaping galaxy properties within the structure.

Conclusions. More of these serendipitous discoveries are expected, and alongside ongoing targeted programmes, they will provide a rich, unbiased sample to study galaxy transformation and enable a MeerKAT HI perspective on large-scale structure, including filaments.

Key words. galaxies: evolution – galaxies: formation – galaxies: groups: general – galaxies: interactions – large-scale structure of Universe – radio lines: galaxies

1. Introduction

Currently, the most widely accepted cosmological structure formation paradigm is the Λ cold dark matter (Λ CDM) model, which predicts that all galaxies form and evolve within dark matter halos and that dark energy accelerates expansion at late cosmological epochs (Springel et al. 2005). Galaxy groups have been shown to form an integral part of this hierarchical evolution model (Springel et al. 2018), making the study of these groups important for our understanding of large-scale structure formation.

In galaxy clusters, environmental effects such as tidal interactions (Moore et al. 1998) and ram-pressure stripping (Gunn & Gott 1972) are commonly observed. This is due to the high galaxy number density leading to an increase in galaxy-galaxy interactions and the hot and dense intracluster medium (ICM) providing a medium for gas stripping. Galaxy groups are generally less dense and have fewer members (~ 3 –100, Catinella et al. 2013) than their larger cluster counterparts (> 100

members), with lower dark matter halo masses of $\sim 10^{12-14} M_{\odot}$ (Catinella et al. 2013; Hess & Wilcots 2013). However, it is important to note the difference between clusters and high-mass groups is not well defined and that no clear distinction exists in the literature (see Lovisari et al. 2021). Albeit less common, tidal interactions are also observed in lower halo mass group structures, both in compact groups of ~ 4 galaxies with a ~ 30 –40 kpc projected separation (e.g. Sulentic et al. 2001) as well as in larger groups of ~ 20 members (e.g. Ranchod et al. 2021). There is evidence of these interactions in group-like environments before in-fall into a cluster environment (often referred to as ‘pre-processing’), such as in the Fornax A group (Kleiner et al. 2021; Loubser et al. 2024). About half of all optically identified galaxies reside in group environments at $z = 0$ (Eke et al. 2004; Robotham et al. 2011), and 25% of the neutral hydrogen (HI) identified galaxies reside in group environments (Hess & Wilcots 2013), making the study of galaxies in this environment key to improving our understanding of galaxy evolution.

It has been shown that the interstellar medium (ISM) properties of galaxies are intimately linked to the circumgalactic

* Corresponding author.

medium (Tumlinson et al. 2017; Nielsen et al. 2020). The H I in galaxies is typically significantly more extended than the stellar component (Broeils & Rhee 1997; Leroy et al. 2008). This, together with the sensitivity of H I to tidal interactions and ram-pressure stripping, makes H I observations a powerful tool for probing the influence of environmental effects as a tracer of long dynamical timescales, as can be seen in the M81, M82, and NGC 3077 triplet (Yun et al. 1994). H I has also been used as a tool to probe cold gas replenishment from large-scale structures such as filaments (Kleiner et al. 2017). The H I mass function (HIMF; Jones et al. 2018) and the H I mass – stellar mass ($M_{\text{HI}} - M_*$) relation (Parkash et al. 2018) provide important probes of environmental effects. However, a large number (order of hundreds) of galaxies are required to ensure sufficient statistical significance in a given environment.

MeerKAT (Jonas & MeerKAT Team 2016), an SKA-mid precursor, has already made significant contributions to the exploration of the H I universe. This has been done through surveys such as Looking At the Distant Universe with the MeerKAT Array (LADUMA; Blyth et al. 2016), the H I component of the MeerKAT International GigaHertz Tiered Extragalactic Exploration survey (MIGHTEE; Jarvis et al. 2016; Maddox et al. 2021; Ranchod et al. 2021; Heywood et al. 2024), MeerKAT H I Observations of Nearby Galactic Objects: Observing Southern Emitters (MHONGOOSE; de Blok et al. 2024), and the MeerKAT Absorption Line Survey (MALS; Gupta et al. 2016; Deka et al. 2024).

The advances in both cosmological hydrodynamical simulations and H I observations over the past years provide an opportunity for comparison between empirical and theoretical results. SIMBA (Davé et al. 2019), EAGLE (Schaye et al. 2015), and ILLUSTRISTNG (Pillepich et al. 2018) have been shown to reproduce the HIMF (see Davé et al. 2020, for a comparison). However, the environmental effects on H I within large-scale structure have not been studied as well (e.g. Bulichi et al. 2024), so it is valuable to explore additional constraints towards this.

In this paper, we perform a detailed study of the serendipitous discovery of a large grouping of H I-rich galaxies in the foreground of the Abell 3365 galaxy cluster, which we first reported in Knowles et al. (2022). The grouping contains 20 detections and takes on a peculiar filamentary-like shape in projection. The galaxies span a projected ~ 3.7 Mpc from the northernmost to the southernmost galaxy and include several interacting systems. In Knowles et al. (2022), we showed a sample of two H I galaxies with radio continuum counterparts and an overview H I map of the group before group member identification. Here, we perform individual galaxy analysis of the H I galaxies detected in the grouping using multiwavelength data from the Dark Energy Survey (DES), an optical imaging survey carried out by the Dark Energy Camera (DECam; Abbott et al. 2021), and NASA’s Wide-field Infrared Survey Explorer (WISE; Wright et al. 2010), together with MeerKAT data to probe H I and stellar masses, kinematics, galaxy-galaxy interactions, and optical morphologies. We use both manual and automated source finding along with a friends-of-friends algorithm to establish robust group membership. We then explore similar structures in SIMBA to contextualise this H I overdensity.

This paper is organised as follows: In Sect. 2 we describe the MeerKAT observations as well as the calibration and imaging implemented in this work, and we summarise the ancillary data. In Sect. 3, the source-finding process is discussed, and we then briefly address the approach we took for group member selection. We also present moment 0, moment 1, DES *RGB*, WISE, and radio continuum cutouts of individual galaxies. In Sect. 4 we

Table 1. Summary of MeerKAT observations.

Experiment ID	Date	Antennas	Track length
20181114-0020	2018/11/14-15	56/64	10.05 hours
20181118-0008	2018/11/18-19	59/64	10.10 hours

discuss the morphology and kinematics of individual galaxies as well as potentially interacting systems within the grouping. The properties of the group are then compared to established scaling relations. Finally, we use the SIMBA cosmological hydrodynamical simulations to put the peculiar filamentary-like projection of the grouping in context with respect to the large-scale structure. A summary of the findings of this work is presented in Sect. 5. Throughout this paper we adopt cosmological values found in Planck Collaboration VI (2020): $\Omega_M = 0.315 \pm 0.007$ and $H_0 = 67.4 \pm 0.5 \text{ km s}^{-1} \text{ Mpc}^{-1}$. This corresponds to a scale of $0.822 \text{ kpc arcsec}^{-1}$ for sources at $z = 0.04$.

2. Data, calibration, and ancillary data

2.1. MeerKAT observations

MeerKAT is a radio interferometer located in the Northern Cape of South Africa. The mid-frequency array serves as a precursor to the Square Kilometre Array (SKA). The MeerKAT array consists of 64 antennas in a dense core layout where $\sim 70\%$ of the antennas are in the core with a minimum baseline of 29 m and a maximum baseline of ~ 1 km. Approximately 30% of the antennas are located beyond the core and provide the interferometer with a maximum baseline of ~ 8 km. Both regions are distributed such that they provide near Gaussian (u, v) coverage. All dishes are 13.5 m in diameter with offset Gregorian optics and equipped with *S*-band, *L*-band, and UHF-band receivers with frequency ranges of 1750–3500 MHz, 900–1670 MHz, and 580–1015 MHz respectively (see Jonas 2009; Jonas & MeerKAT Team 2016; Camilo et al. 2018, for detailed technical specifications).

The raw visibility data used in this paper stem from the MeerKAT Galaxy Cluster Legacy Survey (MGCLS), a SARAO Science Legacy Survey that performed deep MeerKAT *L*-band observations of 115 galaxy clusters (Knowles et al. 2022). Each cluster was typically observed for 6–10 hours each with MeerKAT in the correlator’s 4k configuration, which has 4096 channels leading to a channel width of 209 kHz (corresponding to $\sim 44 \text{ km s}^{-1}$ at $z = 0$).

The group lies in the Abell 3365 field, located at RA = 05h48m12s and Dec = $-21^\circ 56' 03''$. We note that Abell 3365 is at $z \sim 0.093$, within a frequency range affected by radio frequency interference (RFI), and is cosmologically unrelated to the galaxies studied in this work. Both observations used in this work had a correlator dump time of 8 seconds and included 3 sources: J0408-6545 (bandpass calibrator at RA = 04h08m20s, Dec = $-65^\circ 45' 07''$), J0609-1542 (complex gain calibrator at RA = 06h09m40s, Dec = $-15^\circ 42' 40''$), and Abell 3365. The observations’ properties are summarised in Table 1.

2.2. Calibration and imaging

The MeerKAT raw visibilities are reduced using OXKAT¹ (Heywood 2020) through a SINGULARITY container on ILIFU, a

¹ <https://github.com/IanHeywood/oxkat>

high-performance computing cluster developed and maintained by IDIA². OXKAT is a semi-automatic complete data reduction pipeline software that implements standard routines such as flagging, reference calibration (1GC), self-calibration (2GC), and direction-dependant calibration (3GC). OXKAT makes use of CASA³ (McMullin et al. 2007), WSCLEAN (Offringa et al. 2014), and CUBICAL⁴ (Kenyon et al. 2018), among other software.

The autoflagger TRICOLOUR⁵ is used in conjunction with manual flagging of bad antennas and (u, v) ranges to remove RFI and bad data. Cross calibration is done using OXKAT through standard CASA functions such as `gaincal` and `applycal` to correct for atmospheric phase shifts among other effects. J0609-1542 is observed for ~ 1 minute every ~ 10 minutes and used as the complex gains calibrator. We use J0408-6545 as the flux and bandpass calibrator. Self-calibration is performed to improve the fidelity and dynamic range of the image by iteratively improving the mask used by WSClean.

Continuum subtraction was performed in 3 steps, first in the visibility plane using CASA by subtracting the current sky model contained in the model data column from the corrected data using `uvsub`. Thereafter, remaining continuum emission is modelled and subtracted through fitting a 1st-order polynomial to the emission in the visibility plane at the native 8-sec resolution with `uvcontsub`. Using WSClean and a Briggs weighting scheme with a robust parameter of 2.0 we image the data cube. This is to produce the most sensitive data product, thereby complementing the source-finding-orientated science goals of this work. Finally, we use `imcontsub` to model and remove continuum emission in the image plane by fitting a 2nd-order polynomial to the cube channels. Polynomial fitting orders as high as 4th-order were investigated but did not lead to significant differences in the integrated spectra, moment maps or derived H I masses, thereby having no impact on the scientific outcomes of this paper. The final sub-cube consists of 102 channels covering a frequency range of 1.3523–1.3734 GHz, and has a restoring beam dimension of $30.6547'' \times 30.6547''$.

2.3. MeerKAT data properties

The radio continuum image used in this study is from the MGCLS raw visibilities. The continuum image is a 10240×10240 pixel image with a restoring beam of $6.87'' \times 6.87''$ with a root mean square (RMS) of $\sim 5 \mu\text{Jy beam}^{-1}$. The map is centred on RA 05h 48m 12s and Dec $-21^\circ 56' 06''$. The full continuum image is relatively free of bright, troublesome sources and obvious artefacts.

The data cube produced using the method outlined in Sect. 2.2 has a circular restoring beam of $30.65'' \times 30.65''$ and a pixel size of $2'' \times 2''$. The cube has a channel width of 209 kHz, which corresponds to a velocity width of $\sim 46 \text{ km s}^{-1}$ at $z = 0.0395$. With an RMS of $\sim 52 \mu\text{Jy beam}^{-1}$, we achieve an H I column density sensitivity of $3.146 \times 10^{18} \text{ cm}^{-2}$ at 1σ . A summary of the specifications of the reduced data is given in Table 2.

2.4. Ancillary data

For the identification of optical counterpart galaxies, we used the DES Data Release 2 (DR2). DECam is situated in Chile at

Table 2. MeerKAT H I data cube properties.

Property	Value
Restoring beam	$30.65'' \times 30.65''$
BPA	0°
Pixel size	$2'' \times 2''$
Channel width	209 kHz
	$\sim 46 \text{ km s}^{-1}$ (at $z = 0.04$)
Robust weighting	2.0
RMS	$\sim 52 \mu\text{Jy beam}^{-1}$
H I column density sensitivity (1σ)	$3.15 \times 10^{18} \text{ cm}^{-2}$

the Cerro Tololo Inter-American Observatory on the 4 m Blanco telescope. The data we used comes from the coadded image data products of the DES wide-area survey that covered $\sim 5000 \text{ deg}^2$ of the southern Galactic cap and observed in five bands: g , r , i , z , and Y at $1.11''$, $0.95''$, $0.88''$, $0.83''$, and $0.90''$, respectively (Abbott et al. 2021). DES has a median coadded catalogue depth of 24.7, 24.4, 23.8, 23.1, and 21.7 AB mag (for an aperture of $1.95''$ diameter at $S/N = 10$) for the different bands, respectively; an internal astronomical precision of $\sim 27 \text{ mas}$; and a photometric accuracy of $\sim 10 \text{ mag}$. In Fig. 3 we show the g -band cutouts in the third column.

The mid-infrared data used in this analysis is from WISE, which mapped the entire sky between 7 January 2010 and 6 August 2010. WISE observed in the following 4 bands: W1 ($3.4 \mu\text{m}$), W2 ($4.6 \mu\text{m}$), W3 ($12 \mu\text{m}$), and W4 ($22 \mu\text{m}$) that have an angular resolution of $6.1''$, $6.4''$, $6.5''$, and $12.0''$ respectively (Wright et al. 2010). WISE achieves 5σ sensitivities of 0.08, 0.11, 1.00, and 6.00 mJy in unconfused regions on the ecliptic for the bands, respectively. The W1 is used to estimate M_* for the H I detected galaxies (Jarrett et al. 2023), while W3 is used to calculate the star formation rate (SFR) of these galaxies (Cluver et al. 2017). We show a cutout of W3 in the fourth column in Fig. 3.

The aforementioned ancillary data greatly enrich the study of the H I group. Stellar properties of galaxies can be determined with the WISE data, which allows for comparisons to scaling relations. DES images have sub-arcsecond resolution, giving us a robust way to both confirm or reject potential detection candidates from source-finding, as well as assisting in differentiating sources in the relatively poor angular resolution WISE data to avoid confusion and mismatching during cross-matching.

3. Results

3.1. Source finding

The source finding software implemented in this work is SOFIA 2⁶ (SOURCE FINDING APPLICATION, Westmeier et al. 2021). This is a fully automated source finder specifically optimised for extragalactic H I spectral line source finding in 3D data cubes. It is based on SOFIA 2 (Serra et al. 2015) with significant improvements in speed and efficiency through inclusion of OPENMP to allow for parallelisation. These changes make SOFIA 2 an ideal source-finding tool for this work.

The source-finding strategy is as follows: First, the default settings are used together with the algorithm recommended by the authors (S+C or smooth and clip algorithm) to establish a

² <https://www.idia.ac.za/>

³ <https://casa.nrao.edu>

⁴ <https://github.com/ratt-ru/CubiCal>

⁵ <https://github.com/ratt-ru/tricolour>

⁶ <https://gitlab.com/SoFiA-Admin/SoFiA-2>

baseline. Thereafter, we do multiple runs of the source finder with various values for `scfind.threshold` between 3 and 5, which is the flux threshold relative to the measured noise in each smoothing iteration and the range recommended in the manual. We find that a flux threshold of 3.4 provides us with a complete catalogue when compared to a manual inspection of the cube. In this case a manual inspection refers to looking at every channel in the cube using CARTA⁷ (Comrie et al. 2021) and placing apertures on bright features, then examining the extracted spectra to classify the feature as real, noise, or continuum subtraction error. However, this rarely results in the presence of false positives in the SOFIA 2 generated catalogue, which are manually removed.

The majority of linking parameters are left as default with some minor adjustments to better suit our input data cube. We utilise the `linker.minSizeXY` and `linker.minSizeZ` to filter out noise peaks. We set `linker.minSizeXY` = 10 to approximate our beam size of ~ 30 arcsec to discard sources smaller than the synthesised beam. We set `linker.minSizeZ` = 2, which is significantly lower than the default of five, but at a channel width of 46 km s^{-1} , five channels would discard all galaxies with a full velocity width of less than $\sim 250 \text{ km s}^{-1}$. We note that `linker.minSizeZ` = 1 was tested and no significant difference was noticed in the final catalogue. Therefore, to avoid single channel noise peaks, we chose to use `linker.minSizeZ` = 2. Finally, we set `linker.radiusXY` = 2 instead of the default of 1. This was done since we observed multiple interacting systems and our aim to capture these interactions as opposed to separating the sources involved.

To limit the number of false positives due to a lower `scfind.threshold`, the reliability filter is used after being fine-tuned. We make use of the diagnostic plots output by SOFIA 2 to achieve an optimal reliability filter setup. This is done by optimising the size of the kernel used to determine the reliability of the source, (see Westmeier et al. 2021, for details). The goal is to obtain a normalised distribution of Skellam parameters that closely resembles a standard Gaussian with $\sigma = 1$. The `reliability.autoKernel` does not converge given a tolerance of `reliability.tolerance` = 0.1 and `reliability.iterations` = 100. However, through trial and error, we find that `reliability.scaleKernel` = 0.34 results in a distribution of normalised Skellam parameters with a median of $\mu = -0.083$ in the Skellam plot. This is a good indication that the kernel size is optimised for the data cube. Using the reliability plot, we set `reliability.minSNR` = 2.8. This is slightly below the default of `reliability.minSNR` = 3.0 to allow for potential detections closer to the noise level in the mean and summed flux density parameter space. The reliability threshold was established with `reliability.threshold` = 0.70 (see Westmeier et al. 2021, for details). We note that this value is significantly lower than the default of 0.9; however, in the final catalogue, only ID 17 has a reliability of < 0.99 (~ 0.8), but has been manually confirmed to be real.

We briefly list the preconditioning parameters used. The standard continuum subtraction offered by SOFIA 2 is used with `contsub.enable` = True. To correct for spectral noise variations across the cube, we used `scaleNoise.enable` = True together with `scaleNoise.mode` = spectral. To correct for spatial ripple effects, we enabled the ripple filter with `rippleFilter.enable` = True.

Lastly, we implemented manual cross-matching with DES *i,r,g*-bands to remove false positives that were missed by the reli-

ability filter. The majority of galaxies removed in this step were near bright continuum sources and are likely the result of imperfect continuum subtraction and/or variable point spread function dimensions. After these steps, we were left with a SOFIA 2 generated catalogue of 35 H I sources.

3.2. Grouping of galaxies

When grouping galaxies, there are two main approaches: halo-based finders, such as in (e.g. Yang et al. 2005, 2007, 2021; Lim et al. 2017), or friends-of-friends algorithm-based finders (e.g. Berlind et al. 2006; Crook et al. 2007; Tempel et al. 2014). Friends-of-friends finders associate different galaxies with one another using two linking lengths that are either fixed or scaled with the redshift of a galaxy. The linking lengths, R_L and V_L , define lengths in projection and velocity, respectively, by which to group galaxies (Huchra & Geller 1982). Linking lengths are often difficult to choose, as one set of parameters does not satisfy all relevant tests based on the multiplicity of groups, projected size, and velocity dispersion distributions (see the appendix in Berlind et al. 2006). It is then common practice to choose linking parameters that satisfy the specific science goals at hand, leading to a large variation in the commonly used linking lengths.

In this work, we implement a simple friends-of-friends algorithm similar to Crook et al. (2007) in an analysis of the Two Micron All-Sky Survey (2MASS). The 2MASS redshift survey (2MRS) covers 91% of the sky, is 97.6% complete at $K_s \leq 11.75$ mag, and extends to a depth of $z = 0.05$ (Huchra et al. 2012). In Crook et al. (2007) R_L is scaled to account for bias introduced by the variation in sampling the luminosity function with redshift while V_L is fixed. We choose a naive prescription where R_L and V_L are fixed parameters. The values chosen for this work are $R_L = 1.2 \text{ Mpc}$ and $V_L = 500 \text{ km s}^{-1}$. The value of V_L is similar to V_0 in Ramella et al. (1997) in an analysis of the northern CfA redshift survey, while the value of R_L is between 0.89 in Ramella et al. (1997) and 1.6 in Crook et al. (2007). Both parameters fall well within the reasonable range investigated in Crook et al. (2007). The science goals of this paper are not to establish ideal linking parameters for H I group finding, but rather to probe the environmental effects on galaxies in high-density environments such as the one reported here. Choices of R_L between 1 Mpc and 1.6 Mpc and choices of V_L between 350 km s^{-1} and 600 km s^{-1} result in a small variation in the number of galaxies included in the structure, but the structure itself maintains its projected filamentary appearance and the primary scientific results are not impacted.

In Fig. 1, the distribution of galaxies is shown for all detections produced by our SOFIA 2 strategy. For reference, the geometric mean of the largest group is indicated by a red cross. Of the 35 sources in the SOFIA 2 catalogue, 20 are grouped by our chosen linking lengths; these galaxies are shown by the circle markers in Fig. 1, scaled by H I mass. Although the remaining 15 galaxies (denoted with squares) are reliably detected, they are predominantly isolated and show no signs of interaction. As such, they are excluded from further analysis. We note that the distribution of the galaxies results in a projection of a filamentary-like structure from north to south. In the figure, all markers are coloured by redshift.

Additionally, we searched for spectroscopic redshifts of galaxies within 3σ of the redshift distribution of the largest group. We find numerous galaxies with confirmed, reliable spectroscopic redshifts from the 6dF Galaxy Survey (6dFGS, Jones et al. 2009). The filamentary structure observed in the H I identified galaxies in this work is supported by the distribution

⁷ <https://cartavis.org>

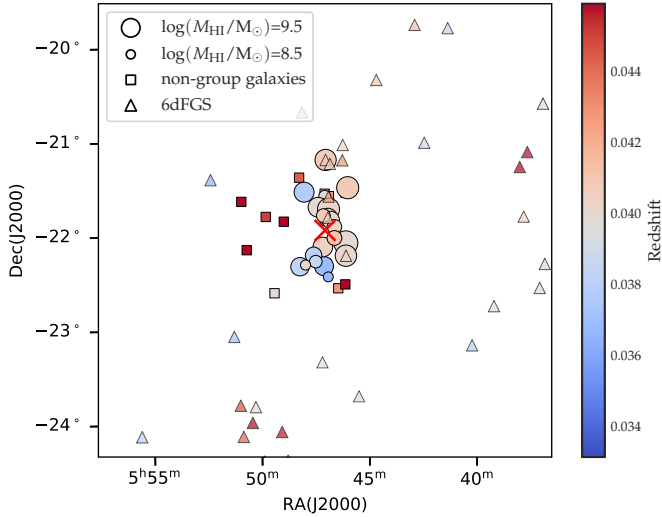


Fig. 1. Spatial distribution scatter plot in RA and Dec of all galaxies in the final SOFIA 2 catalogue. Circular markers have been linked by a friends-of-friends algorithm described in Sect. 3.2. Square markers indicate isolated galaxies or group galaxies belonging to groups with fewer members than the primary group. The size of the circular markers is scaled by H I mass with two indicative masses displayed in the legend. The red cross represents the geometric mean coordinates of the orange points. Galaxies from 6dFGS with spectroscopic redshifts are shown with triangles. All markers are coloured by redshift.

of spectroscopic redshifts from 6dFGS. With the uniform coverage from 6dFGS extending beyond the observation’s bounding box, the spectroscopic redshifts still reveal a north-to-south structure. The filamentary projection of the group is not due to sensitivity limits, as we detect a sub-Milky Way mass galaxy ($\log(M_{\text{HI}}/M_{\odot}) = 8.68$) $\sim 44.3^{\circ}$ from the pointing centre.

3.3. H I detections and properties

Focusing only on the galaxies associated with the friends-of-friends algorithm, we obtained 20 clear detections with at least 22 galaxies. In Fig. 2 we show a DES *RGB* image constructed from *i*, *r*, and *g* bands where the 20 group detections are overlaid with white contours at an H I column density of $1 \times 10^{19} \text{ cm}^{-2}$ assuming the average redshift, $z = 0.0395$. The scale bar in the bottom left is based on the average galaxy redshift. The source properties are listed in Table 3, where IDs are ordered by descending declination. Fig. 3 shows the moment 0 (integrated H I flux), moment 1 (H I velocity field), DES *g*-band (473 nm centre wavelength), WISE W1 band (3.4 μm centre wavelength), and the MeerKAT *L*-band (900–1670 MHz) continuum map for each source in that order from left to right. These plots were made in part with the SOFIA IMAGE PIPELINE⁸ (Hess et al. 2022). The integrated spectra for all galaxies and the multiwavelength images of the remaining sources are in Appendix A and B, respectively. The error bars on the integrated spectra were derived from the per channel RMS of the local noise within the bounding box of the source. The contours overlaid on the moment 0, DES, WISE, and continuum maps correspond to H I column densities stated at the bottom of the image. H I column densities were calculated as

$$N_{\text{HI}} = 2.33 \times 10^{20} (1+z)^4 \left(\frac{S}{\text{JyHz}} \right) \left(\frac{ab}{\text{arcsec}^2} \right)^{-1}, \quad (1)$$

⁸ <https://github.com/kmhess/Sofia-image-pipeline>

where z is extracted from the redshift corresponding to the central frequency of each source, S is the integrated H I flux of the source in JyHz, and a and b are the major and minor axes of the restoring beam in arcsec respectively (Meyer et al. 2017). The moment 1 map is created using pixels in the moment 0 map, excluding values below the lowest contour. The v_{sys} , W_{50} , W_{20} , and Δv contours were calculated using the relation between frequency and velocity:

$$V = c \frac{v_0 - \nu}{\nu}, \quad (2)$$

where ν_0 is the rest frequency of the spectral line, ν is the observed frequency, and c is the speed of light. v_{sys} uses the central frequency of the spectrum of the source, while W_{50} (W_{20}) describes the width of the spectral profile where the flux density has increased by more than 50 (20) % of the peak flux density moving from the outer channels of the cubelet inwards. We adopt a conservative approach to uncertainties associated with the W_{20} and W_{50} by taking an indicative uncertainty of the channel width, 46 km s^{-1} . While this represents a relatively large uncertainty when compared to typical galaxy widths, it does not affect the main scientific conclusions of this paper. The line in the moment 1 map indicates the kinematic semi-major axis, with direction pointing to the high-frequency channels. The kinematic semi-major axis of a source is determined by calculating the flux-weighted centroid in each channel where the source has emission. Plotting these points will result in a straight line in the event that the source is a regularly rotating galaxy. Fitting a straight line to these points on the sky using orthogonal regression results in a good approximation of the semi-major axis (see SOFIA 2 documentation for details). We note that kinematic analysis for single- or two-channel sources could be inaccurate. This is due to the fairly coarse spectral resolution of 209 kHz. The H I mass was calculated following the prescription in Meyer et al. (2017):

$$\left(\frac{M_{\text{HI}}}{M_{\odot}} \right) = 49.7 \left(\frac{D_L}{\text{Mpc}} \right)^2 \left(\frac{S}{\text{JyHz}} \right), \quad (3)$$

where S is the integrated H I flux of the source in JyHz. S is calculated by summing the flux densities across the source. The H I redshift calculated from the central frequency of each source is used to calculate the luminosity distance, D_L , of the source. Uncertainties associated with D_L are derived from SOFIA 2 statistical uncertainties in the centroid channel location along the frequency axis. To obtain an estimate for the uncertainty in the integrated flux S , we calculate the signal in four emission-free masked regions near the source, for this we use the unique 3D mask output by SOFIA 2 for each galaxy. We then take the mean RMS scatter of the flux values as our uncertainty (Ponomareva et al. 2021). This is then combined in quadrature with a 10% absolute flux uncertainty. Uncertainties are shown in Table 3.

The SFR and M_* for viable sources were estimated in the following way. We manually crossmatched sources using CARTA with DES *i*, *r*, *g*-bands; WISE W1, W2, and W3 bands; moment 0 cutouts produced by SOFIA; and the WISE All-Sky Source Catalog. For H I sources that have been matched to a source in the WISE catalogue, we inspected the `ext_flg`. If `ext_flg` = 0, the source is consistent with a point source, and we use the largest non-contaminated aperture (checked with flags provided in the WISE catalogue for each aperture) to acquire a magnitude for that band. If the source is extended (`ext_flg` > 0), we manually ensured that the aperture adequately covers the source.

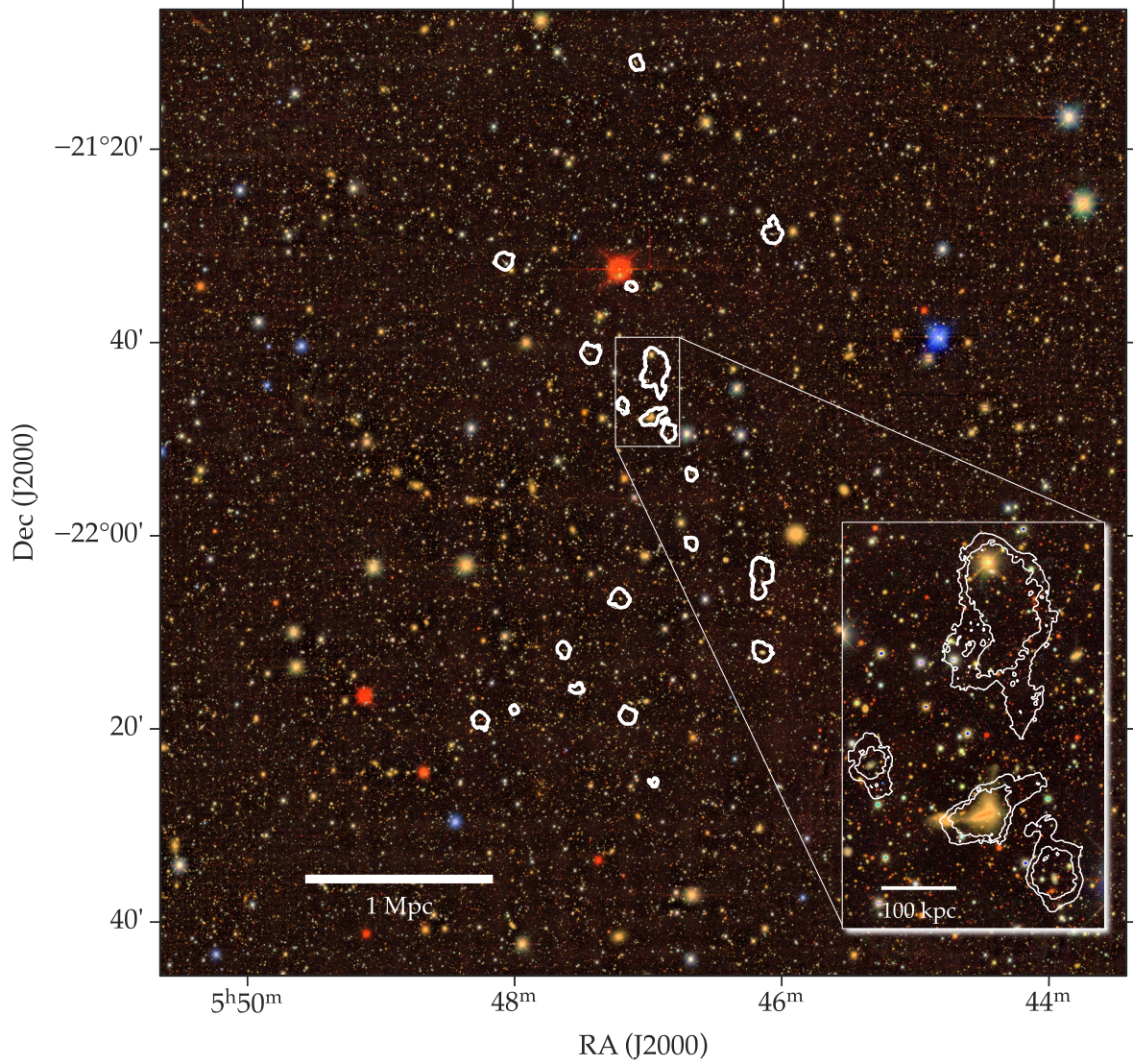


Fig. 2. Dark Energy Survey RGB image constructed from *i*, *r*, and *g* bands. MeerKAT H I contours are overlaid in white at an H I column density of $1 \times 10^{19} \text{ cm}^{-2}$ at the average galaxy redshift of $z = 0.0395$. A zoomed inset of the compact group discussed in Sect. 4.2 is shown with white H I contours at $1 \times 10^{19} \text{ cm}^{-2}$ and $1 \times 10^{20} \text{ cm}^{-2}$.

Finally, we inspected the upper limit flag supplied by WISE to determine if the magnitude quoted is a 95% upper limit. If the source is a non-detection (with no contamination present), we derived an upper limit from the appropriate WISE band sensitivity.

Light emitted from K- and M-type giant stars dominate the W1 and W2 bands, making them good tracers of the stellar mass within a galaxy by tracing the continuum emission from the evolved stellar population of a galaxy (Cluver et al. 2014). The initial relation in Cluver et al. (2014) has a large observed scatter, correspondingly derived M_* values have large uncertainties. Since then, improvements have been made on these relations (Jarrett et al. 2023). We choose to use a simplistic relation from Jarrett et al. (2023), only considering the W1 magnitude. This is motivated by the sensitivity of the W1 band in WISE, as well as eliminating the uncertainty associated with the ambiguous W2 emission of various sources. The relation is as follows:

$$\log M_* = A_0 + A_1 (\log L_{W1}) + A_2 (\log L_{W1})^2 + A_3 (\log L_{W1})^3 \quad (4)$$

with

$$L_{W1} (L_\odot) = 10^{-0.4(M - M_{\text{sun}})}, \quad (5)$$

where $M_{\text{sun}} = 3.24$ is the absolute Vega magnitude of the sun, M is the W1 absolute magnitude of the source, and the A coefficients are -12.62185 , 5.00155 , -0.43857 , and 0.01593 , respectively. We used D_L , calculated from the H I redshift, to convert the W1 Vega magnitude to an absolute magnitude for each source. Upper limits are indicated in Table 3, while contaminated sources or sources where no adequate aperture could be found have their M_* values omitted.

The WISE W3 band luminosity is dominated by warm dust and polycyclic aromatic hydrocarbons for $\sim L^*$ star-forming galaxies, with the contributions being 62.5% and 34% respectively (Cluver et al. 2017). This makes W3 a tracer of star formation. However, for these galaxies stellar continuum contributes 15.8% of the W3 light. To correct for this effect we subtract 15.8% of the flux in W1 from W3 before using the relation. The SFR values for galaxies with corresponding W3 magnitudes were calculated using Eq. (4) in Cluver et al. (2017), calibrated using the Spitzer Infrared Nearby Galaxy Survey (SINGS;

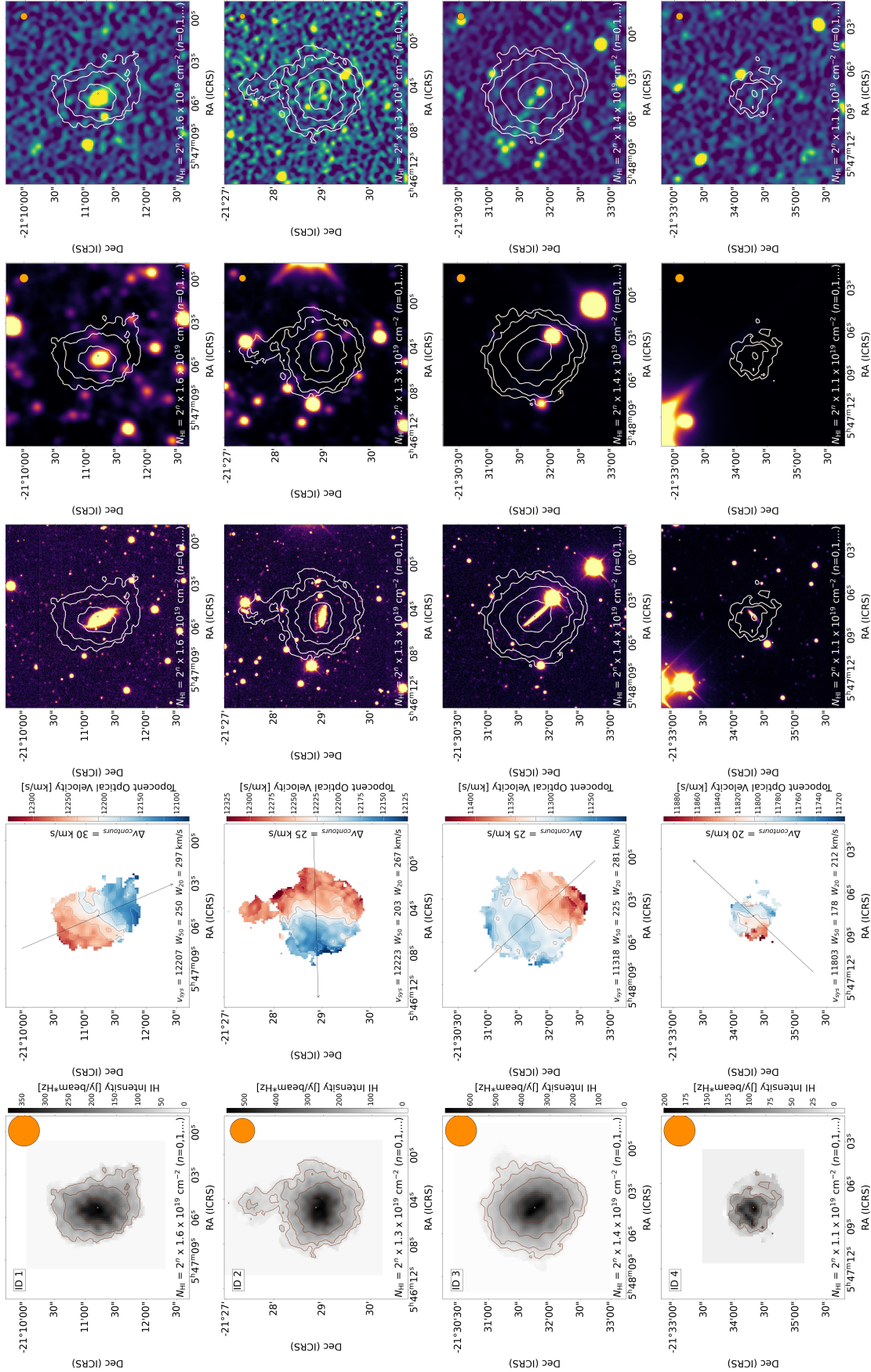


Fig. 3. From left to right (for IDs 1–4), moment 0 (integrated HI) map, moment 1 (velocity field), DES g -band, WISE W1 band, MeerKAT L -band continuum. In the moment 1 map, we show the recessional velocity of the source (v_{sys}), spectral line width at 50% and 20% of the peak flux density (W_{50} and W_{20}), as well as the velocity difference between contours ($\Delta v_{\text{contours}}$). An indication of the kinematic position angle is given by the line centred on the flux-weighted centroid of the source. In the other images, we show contours from the moment 0 map overlaid at HI column densities indicated at the bottom of the images. The moment 0 map has the HI restoring beam ($30.65'' \times 30.65''$) displayed in the top-right corner. Similarly, we show the MeerKAT continuum map restoring beam ($6.87'' \times 6.87''$). The DES g -band and WISE W1 band have a point spread function full width half maximum of $1.11''$ and $6.1''$, respectively, also shown in the top right.

Table 3. Properties of group galaxies or sources using HI and infrared data.

ID	RA (J2000) (h:m:s)	Dec (J2000) (d:m:s)	z_{HI}	D_L (Mpc)	$\log$ ($M_{\text{HI}}/M_\odot$)	W_{50} (km s^{-1})	$\log$ ($M_*/M_\odot$)	SFR ($M_\odot \text{ yr}^{-1}$)	f_{HI}^*
1	5:47:05.3	-21:11:11.89	0.0407	186.0 $\pm$ 0.12	9.832 $\pm$ 0.044	250	...	...	...
2	5:46:04.7	-21:28:51.63	0.0408	186.2 $\pm$ 0.06	9.988 $\pm$ 0.048	203	9.24	0.18	0.74
3	5:48:04.2	-21:31:44.92	0.0378	172.1 $\pm$ 0.07	9.565 $\pm$ 0.048	226	9.50	0.19	0.06
4	5:47:07.7	-21:34:18.03	0.0394	179.6 $\pm$ 0.24	8.616 $\pm$ 0.061	179	...	...	...
5	5:47:25.9	-21:41:19.86	0.0399	182.2 $\pm$ 0.05	9.593 $\pm$ 0.048	171	...	...	...
6	5:46:57.1	-21:42:33.99	0.0401	183.2 $\pm$ 0.03	9.965 $\pm$ 0.055	152	...	...	...
7	5:47:11.2	-21:46:40.14	0.0408	186.4 $\pm$ 0.10	8.843 $\pm$ 0.045	111	<8.67	<0.52	>0.17
8	5:46:58.6	-21:47:54.69	0.0405	184.7 $\pm$ 0.19	9.438 $\pm$ 0.059	542	10.98	3.27	-1.54
9	5:46:51.0	-21:49:27.23	0.0401	182.9 $\pm$ 0.08	9.266 $\pm$ 0.053	138	9.27	0.21	-0.00
10	5:46:40.6	-21:53:47.16	0.0403	183.9 $\pm$ 0.10	8.916 $\pm$ 0.055	137	8.56	<0.12	0.36
11	5:46:40.7	-22:00:53.58	0.0411	187.8 $\pm$ 0.09	8.964 $\pm$ 0.057	91	<8.68	<0.53	>0.28
12	5:46:09.3	-22:04:00.44	0.0399	182.3 $\pm$ 0.02	10.314 $\pm$ 0.045	269	9.61	0.56	0.70
13	5:47:12.5	-22:06:39.34	0.0404	184.4 $\pm$ 0.06	9.584 $\pm$ 0.043	175	9.22	0.20	0.37
14	5:47:37.9	-22:11:56.29	0.0382	173.9 $\pm$ 0.09	9.096 $\pm$ 0.063	169	8.69	<0.07	0.40
15	5:46:08.9	-22:12:09.80	0.0400	182.8 $\pm$ 0.05	9.852 $\pm$ 0.046	229	10.16	2.14	-0.30
16	5:47:31.8	-22:15:59.01	0.0386	175.8 $\pm$ 0.16	8.682 $\pm$ 0.058	76	<8.46	<0.04	>0.22
17	5:48:00.0	-22:18:09.32	0.0401	183.0 $\pm$ 0.10	8.498 $\pm$ 0.060	63	<8.66	<0.50	>-0.16
18	5:47:09.2	-22:18:50.47	0.0368	167.7 $\pm$ 0.06	9.415 $\pm$ 0.047	194	<8.58	<0.43	>0.83
19	5:48:15.2	-22:19:13.58	0.0382	174.3 $\pm$ 0.07	9.358 $\pm$ 0.058	169	...	...	...
20	5:46:57.3	-22:25:41.12	0.0367	167.3 $\pm$ 0.17	8.410 $\pm$ 0.059	58	...	...	...

Notes. Cols (1)–(3): Source IDs sorted by decreasing declination with the RA and Dec from HI flux weighted centroid produced by SOFIA 2. Cols (4)–(5): HI derived redshift and luminosity distance with uncertainties. Cols (6)–(7): HI mass with uncertainties and W_{50} (uncertainty of 46 km s^{-1}) of integrated HI line profile outlined in Sect. 3.3. Cols (8)–(10): Stellar mass (uncertainty of 0.12 dex), SFR (uncertainty of 0.15 dex), and HI gas fraction, described in Sect. 3.3.

(Kennicutt et al. 2003) and Key Insights on Nearby Galaxies: a Far-Infrared Survey with Herschel (KINGFISH; Kennicutt et al. 2011):

$$\log \text{SFR}(M_\odot \text{ yr}^{-1}) = (0.889 \pm 0.018) \log L_{12\mu\text{m}}(L_\odot) - (7.76 \pm 0.15), \quad (6)$$

where $L_{12\mu\text{m}}$ is the luminosity in the WISE W3 band. We obtain flux densities for the W3 and W1 bands by following the outline given in the WISE data processing documentation⁹:

$$F_\nu(\text{Jy}) = F_{\nu 0} \times 10^{(-m_{\text{Vega}}/2.5)}, \quad (7)$$

where values for $F_{\nu 0}$ are given in the documentation. Similarly to the M_* calculation given by Eq. (4), we omit contaminated sources and indicate WISE quoted 95% upper limits, as well as sensitivity derived upper limits in Table 3. Finally, we define $f_{\text{HI}}^* = \log_{10}(M_{\text{HI}}/M_*)$ for sources where stellar masses are calculated.

4. Discussion

Sources within the structure are classified into ‘disk-like’, ‘interacting system’, or ‘unclear’ based on visual inspection of the DES imaging of the optical counterpart to the HI detection. We find IDs 1, 3, 5, 13, 14, 15, and 19 to be disk-like galaxies. Interacting systems are apparent in IDs 2, 6, 8, and 12. For IDs 4, 7, 9, 10, 11, 16, 17, 18, and 20 their morphologies are not clear based on the data inspected. Inspection of the HI spectra of the sources finds double horn profiles for IDs 1, 4, 6, 9, 12, 14, 15,

18, and 19, though some appear to be asymmetric. We note that this approach only loosely classifies the sources, as we acknowledge the subjective nature of the methodology.

The HI mass distribution of the grouped galaxies sample is shown in Fig. 4 as a histogram. The masses are divided into 10 bins of equal size (~ 0.2 dex) spanning the range of $8.5 \leq \log(M_{\text{HI}}/M_\odot) \leq 10.5$. All 20 sources fall within $\Delta \log(M_{\text{HI}}/M_\odot) = 1.9$. A comparison to the $z = 0$ HIMF based on Arecibo Legacy Fast ALFA (ALFALFA) blind HI survey of the nearby universe provides valuable insight into the sample in question. The Schechter function fit has a ‘knee’ mass of $m_* = 9.94 \pm 0.01$ (Jones et al. 2018). ID 2, ID 6, and ID 12 all have HI masses above the ‘knee’ and are discussed in depth in Sect. 4.3.

4.1. Scaling relations

We investigated the $M_{\text{HI}}-M_*$ relation with a comparison of linked galaxies within the group structure to the stellar-HI mass relation for HI-selected galaxies reported in Parkash et al. (2018) (see Fig. 5). Of the nine sources with clear M_* values, we find that six (three) fall within the 1σ (3σ) confidence interval. All but one HI detection with stellar mass upper limits (non-detections or upper limits quoted by WISE) is within 1σ of the relation. The most notable outlier is ID 8, with the highest stellar mass of $\log(M_*/M_\odot) = 10.98$ (see Sect. 4.2 for an expanded discussion on this system). In general, we find the galaxies to be in good agreement with Parkash et al. (2018). We note that IDs 1, 4, 5, 6, 19, and 20 are all either quoted as contaminated by the WISE catalogue or clearly contaminated as per a visual inspection. These HI galaxies all have counterparts in the DES imaging and have hints of mostly contaminated counterparts in the WISE bands or catalogue.

⁹ https://wise2.ipac.caltech.edu/docs/release/allsky/expsup/sec4_4h.html

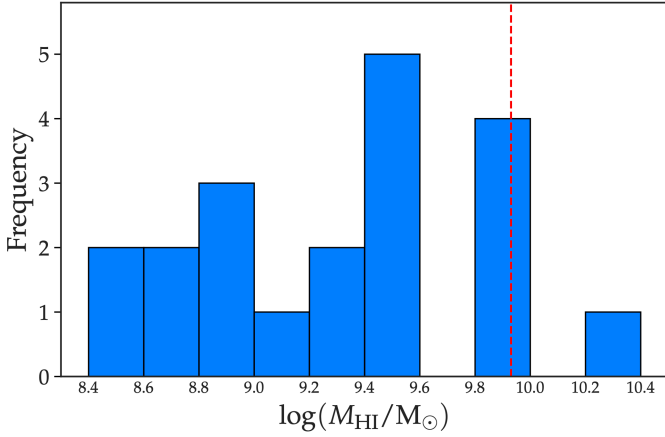


Fig. 4. H I mass distribution of galaxies in bins of 0.2 dex. The vertical dashed red line indicates the ‘knee’, M_* , of the Schechter function fit to the HIMF at $z = 0$ (Jones et al. 2018).

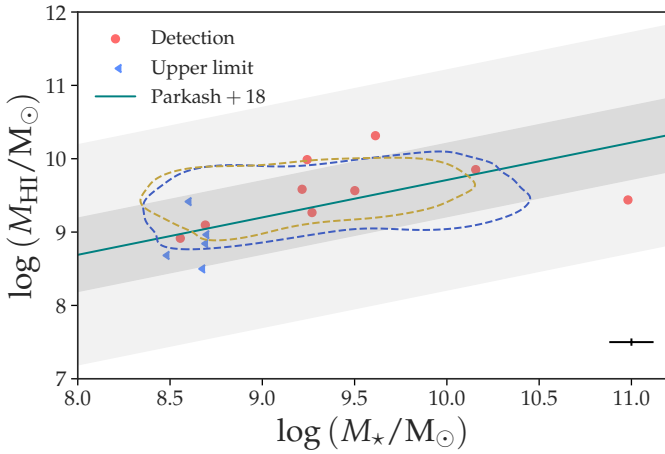


Fig. 5. Relation of $M_{\text{HI}}-M_*$ from Parkash et al. (2018) in teal. The 1σ and 3σ confidence intervals are shown with the shaded areas. Sources from this work are shown in coral (detections) and blue (upper limits). An indicative error is given in the bottom right as 0.12 dex on the stellar mass (based on the scatter of the relation (Jarrett et al. 2023)) and 0.05 dex on the H I mass (based on the uncertainty outlined in Section 3.3). The blue (yellow) dashed contours enclose 68% of the SIMBA analogue (field) galaxies (see Sect. 4.5.4).

The relationship between SFR and M_* is also investigated with a comparison to the Parkash et al. (2018) relation. The star-forming main sequence (SFMS) compared to in Fig. 6 is based on a spiral sample. Of the detections, 5 are within the 1σ confidence interval from the relation, and 2 are outside the 1σ level but within the 3σ confidence interval. For SFR upper limits, both are within 1σ of the relation. For SFR and M_* upper limits, one is within 1σ and 4 are within 3σ . ID 8 has the highest SFR ($(\text{SFR}/M_\odot \text{ yr}^{-1}) = 3.27$) and is discussed in Sect. 4.2.

We caution against over-interpretation of the comparisons with Parkash et al. (2018) due to differences in selection criteria (see Parkash et al. 2018). However, we wish to see if there are any clear correlation differences between an H I-selected sample and field galaxies despite the selection differences. We conclude that the group galaxies identified in this work are consistent with what we expect from local star forming galaxies both in the SFMS and the $M_{\text{HI}} - M_*$ relations.

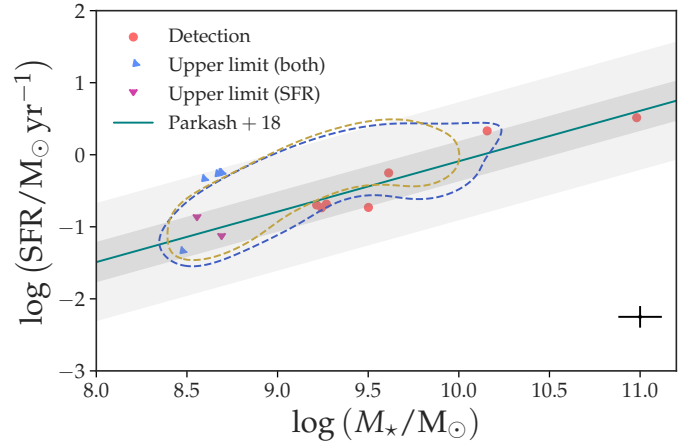


Fig. 6. Star-forming main sequence (SFR– M_* relation) from Parkash et al. (2018) in teal. The 1σ and 3σ confidence intervals are shown with shaded areas. Sources from this work are shown in coral (Detections), purple (upper limits on SFR), and blue (upper limits on SFR and stellar mass). An indicative error on the stellar mass and SFR is given in the bottom right as 0.12 dex and 0.15 dex, respectively, based on the scatter of the relations (Jarrett et al. 2023; Cluver et al. 2017). The blue (yellow) dashed contours enclose 68% of the SIMBA analogue (field) galaxies (see Sect. 4.5.4).

4.2. Compact group

Of the most interesting interactions is a compact group of 4 sources, highlighted in Fig. 2, found near the geometric centre of the structure. Source IDs 6, 7, 8, and 9. The central frequency of the compact group corresponds to $z = 0.0405$, and a luminosity distance of 184.3 Mpc. The centre channels of these sources fall within $\sim 210 \text{ km s}^{-1}$ of each other. A rectangular bounding box of the four sources is $\sim 300 \times 500 \text{ kpc}$. ID 6 shows the clearest signs of interaction, with a large emission tail extending to the south. The compressed contours to the north and the extended H I emission to the south may be indicative of ram-pressure stripping. However, the intragroup medium is typically less dense than the ICM thought to be the primary cause of ram-pressure stripping seen in clusters (Gunn & Gott 1972). Additionally, the close proximity to IDs 7, 8, and 9 suggests tidal stripping. This tail is shown at an H I column density of $1.5 \times 10^{19} \text{ cm}^{-2}$ in Fig. B.1. A foreground star contaminates all WISE bands, and no NIR emission can confidently be associated with the H I galaxy. The optical host galaxy can be seen in the DES g -band directly south of the aforementioned foreground star. ID 7 seems to be only mildly affected by the neighbouring galaxies, with a disturbed H I rotation visible in the moment 1 map, and an asymmetrical spectral profile. ID 7 also has a subtle elongation in its lowest H I contour to the southwest. Similarly to ID 6, ID 8 appears to be an influential factor in the interactions occurring. ID 8 consists of 2 galaxies, both of which can be identified in the DES g -band image. While the moment 1 map does not show abnormal rotation, the spectral profile has the largest extent ($W_{50} = 541 \text{ km s}^{-1}$) of all sources in the full sample and represents a bimodal Gaussian distribution characteristic of two interacting galaxies. We see a clear extended continuum counterpart aligned with the highest density H I contour, potentially indicating active galactic nucleus (AGN) activity or disk star formation. Finally, there is a faint H I-tail extending to the northwest. A single entry in the WISE catalogue allows us to calculate the stellar mass and SFR, and from this ID 8 is found to have the highest stellar mass at $\log(M_*/M_\odot) = 10.98$ but with the caveat

that WISE identifies only one galaxy when two are present. The high SFR seen in ID 8 is a common characteristic of similar interacting systems (Laurikainen & Moles 1989). Taking all into consideration, ID 8 is proposed to be an ongoing merging system between two galaxies with a high SFR, high stellar mass and potential AGN activity, while the H I mass of the system is low in comparison. We note that the galaxy east of the interacting pair visible in the DES *g*-band has a spectroscopic redshift of $z \sim 0.048$, and is thus unrelated to the interaction. Lastly, ID 9 has small H I clouds extending to the north of the galaxy. The rotation of ID 9, shown in the moment 1 map, seems disturbed. ID 9 does have a WISE counterpart and therefore has a corresponding SFR and stellar mass estimates. Some continuum emission seems to be present in the vicinity of the H I galaxy but it is not clear if it is associated. These 4 galaxies appear to be a compact group where the H I extent of the galaxies is proving to be useful for probing the interactions between these galaxies. We propose that ID 6 is potentially an example of a galaxy that has undergone tidal interactions with the other galaxies in the compact group, before having its gravitationally disrupted and rarefied, H I gas ram-pressure stripped by the medium present within the compact group. A combination of tidal interactions and ram-pressure stripping has been studied in detail in the Fornax cluster (Serra et al. 2023). Given the spectral and spatial resolution of this data, a deeper MeerKAT image in the 32k mode would provide a better column density sensitivity to potentially find H I clouds not detected in this work and provide superior velocity fields through finer spectral resolution, thus allowing for more accurate kinematic analysis. Furthermore, a detailed X-ray study of the medium present in this compact group is required to definitively conclude that the galaxy has undergone ram-pressure stripping.

4.3. Individual sources of interest

On the northern outskirts of the large-scale structure, we find an interacting system of two galaxies seen in source ID 2. ID 2 consists of a large disk-like galaxy in the south and a smaller disk-like galaxy in the north. The larger southern galaxy has a position angle nearly perpendicular to that of the northern galaxy. ID 2 has a relatively low stellar mass ($\log(M_*/M_\odot) = 9.24$) compared to its H I mass ($\log(M_{\text{HI}}/M_\odot) = 9.988$). This is somewhat expected as the northern galaxy does not appear in the WISE catalogue and so stellar mass and SFR are calculated based solely on the larger southern galaxy, while the H I mass takes both galaxies into account. For the aforementioned reasons, ID 2 holds the highest non-upper limit neutral gas fraction in the sample.

In source ID 12, we see an analogue of source ID 2, albeit with the primary galaxy north of the satellite galaxy. In ID 12 the inclination angles of the galaxies are not as perpendicular as was the case in ID 2. This is the most H I massive source in the sample ($\log(M_{\text{HI}}/M_\odot) = 10.314$), with the caveat that the satellite galaxy also contributes to the H I mass calculation. Similarly to ID 2, we estimate SFR and stellar mass using only the WISE entry for the main galaxy, as the infalling galaxy does not appear in the catalogue. ID 12 also has a particularly high neutral gas fraction that is expected, considering a significant amount of stellar mass is not being accounted for in our estimate. This is perhaps a system that is in the very early stages of merging as the velocity fields are relatively undisturbed, while a clear H I bridge is visible.

Source ID 15 is in and of itself an interesting spiral galaxy with hints of spiral arms visible in the DES *g*-band image. The

galaxy has a high SFR and therefore lies within the upper 3 σ confidence interval of the SFMS. The radio continuum map shows bright continuum emission associated with the galaxy, coincident with the WISE W1, DES *g*-band, and H I flux weighted centroid, potentially indicating the presence of AGN activity. However, the WISE colour properties (W1-W2 ~ 0.1 mag and W2-W3 ~ 3.7 mag) are more consistent with a starburst galaxy than AGN activity (Wright et al. 2010). The integrated H I spectrum is an undisturbed double-horn profile.

4.4. Group properties

The H I grouping consists of 20 H I detections, all of which have clear DES counterparts. The structure has a spatial extent from the northernmost to the southernmost galaxy of ~ 3.7 Mpc. This is calculated using the average group galaxy redshift of $z = 0.04$. Along the line-of-sight, the group extends 20.5 Mpc assuming only Hubble flow, with the nearest galaxy at 167.3 Mpc and the furthest at 187.8 Mpc. Clearly, this is too large for the galaxies to be gravitationally bound. However, at the redshift of $z \sim 0.04$, peculiar velocities dominate, making the line-of-sight distance uncertain and likely an upper limit. Summing the H I mass of all the group galaxies, we find a total H I mass of $\log(M_{\text{HI}}^{\text{tot}}/M_\odot) = 10.9$, while the average H I mass of a galaxy is found to be $\log(M_{\text{HI}}/M_\odot) = 9.599 \pm 0.053$.

An interesting comparison can be made with a similar H I group discovery in Ranchod et al. (2021). This serendipitous discovery of a group of 20 H I galaxies is at a similar redshift of $z \sim 0.044$. While more tightly bound along line-of-sight, the group in Ranchod et al. (2021) is also spatially extended, at ~ 3 Mpc. The average galaxy H I mass is higher at $\log(M_{\text{HI}}/M_\odot) = 9.90$. Total H I mass is given as $\log(M_{\text{HI}}^{\text{tot}}/M_\odot) = 11.20$, making the group significantly more H I massive.

More recently, Glowacki et al. (2024) has reported the discovery of 49 H I-rich galaxies in a 2.3-hour MeerKAT observation. This discovery consists of at least 3 galaxy groups, with multiple sub-groups. Due to significant differences in grouping strategies, we do not make group property comparisons here; we do however highlight the redshifts of the three groups at $z \sim 0.033$, 0.041, and 0.055. Although the groups have clear differences, this growing number of discoveries demonstrates the ability of MeerKAT to detect loosely bound H I groups within a redshift window around $z \sim 0.04$.

4.5. Large-scale structure context with SIMBA

To place the group structure in context within the evolution of the large-scale structure of the universe, we investigate distributions of galaxies using the cosmological hydrodynamical simulation SIMBA (Davé et al. 2019), the successor to MUFASA (Davé et al. 2016). We intend to provide an indicative sense of how common or extreme our group is. We use the catalogues from the flagship SIMBA¹⁰ run, which has 1024³ gas particles and 1024³ dark matter particles within a co-moving volume of 100 h^{-1} Mpc. The simulation assumes Planck Collaboration XIII (2016) cosmological parameters and runs starting from $z = 249$ to $z = 0$. We query the catalogues on ILIFU¹¹ using the CAESAR¹² package.

¹⁰ <http://simba.roe.ac.uk/>

¹¹ <https://www.ilifu.ac.za>

¹² <https://caesar.readthedocs.io/en/latest/catalog.html>

4.5.1. Detection criteria

We applied individual galaxy detection criteria on the peak H I flux of the galaxies within the cone. For each galaxy, we assumed a boxcar spectral profile and calculated the peak flux of this profile in the following way: SIMBA provides the intrinsic H I and stellar masses, respectively M_{HI} and M_* , which can be used to obtain an approximate total baryonic mass, M_{bar} (see Lelli et al. 2019):

$$M_{\text{bar}} = 1.33M_{\text{HI}} + M_*. \quad (8)$$

From the total baryonic mass, we used the Baryonic Tully-Fisher relation (Lelli et al. 2019) to estimate the mean line width of the H I profile at 20% of the peak flux density, $W_{\text{P}20}$:

$$\log\left(\frac{M_{\text{bar}}}{M_{\odot}}\right) = (3.75 \pm 0.08) \log\left(\frac{W_{\text{P}20}/2}{\text{km s}^{-1}}\right) + (1.99 \pm 0.18). \quad (9)$$

We used equation 46 from Meyer et al. (2017) to obtain rest-frame velocity integrated flux using the intrinsic H I mass for each SIMBA galaxy:

$$\left(\frac{M_{\text{HI}}}{M_{\odot}}\right) = \frac{2.35 \times 10^5}{1+z} \left(\frac{D_L}{\text{Mpc}}\right)^2 \left(\frac{S^{V_{\text{rest}}}}{\text{Jy km s}^{-1}}\right). \quad (10)$$

Under the boxcar approximation for the H I-velocity profile, the rest-frame velocity integrated flux becomes

$$S^{V_{\text{rest}}} = S_{v,\text{peak}} \Delta V_{\text{rest}}, \quad (11)$$

where ΔV_{rest} is the width and $S_{v,\text{peak}}$ is the peak of the rest-frame boxcar spectral profile.

Assuming that $\Delta V_{\text{rest}} = W_{\text{P}20}$, and with $W_{\text{P}20}$ and $S^{V_{\text{rest}}}$ in hand for each galaxy, we utilise Eq. (11) to find the peak of the boxcar profile. The detection criterion is chosen to be $S_{v,\text{peak}} > 5 \sigma_{\text{RMS}}$ where σ_{RMS} is an empirically determined RMS from our MeerKAT observation. We require that each galaxy is at least 3 channels wide, i.e. $\Delta V_{\text{rest}} > 3 \times 46 \text{ km s}^{-1}$. A qualitative comparison is made to a similar future SKA-mid observation. Here we aim only to show the potential enhancement for galaxy group detection given an SKA-mid observation of the same on-source integration time as our observations (~ 14 hours). We scale the RMS from our observation with the factor improvement gained by accounting for the increase in collecting area from adding $80 \times 15 \text{ m}$ diameter SKA-mid dishes to the array.

4.5.2. Analogue criteria

To qualitatively illustrate how likely it is for MeerKAT to discover similar structures, we defined analogous structures as groupings that meet the following requirements: $10 \leq \log(M_{\text{HI}}^{\text{tot}}/M_{\odot}) \leq 11.4$, and $1.7 \leq D_p \leq 5.7 \text{ Mpc}$, where the projected extent, D_p , is the maximum projected distance between any two group members in Megaparsec. Small variations in these selection criteria do not lead to significant changes in our findings. We then ran the same friends-of-friends algorithm used to associate galaxies in the SIMBA light cones (see Sect. 3.2). The same linking lengths of $R_L = 1.2 \text{ Mpc}$ and $V_L = 500$ were chosen for consistency.

4.5.3. Light cone generation

To simulate a MeerKAT observation, we construct light cones from SIMBA snapshots out to $z = 0.1$ following the approach taken in Lovell et al. (2021), just below the RFI corrupted region

($0.1 \lesssim z \lesssim 0.3$). This is done by combining every second snapshot with a random orientation permutation to avoid repeating large-scale structure. The opening aperture used in the first SIMBA box is taken to be $\sim 1.0^\circ \times 1.0^\circ$ at a random location on the face of the SIMBA box. The subsequent apertures are scaled by the length of the light cone appropriately, while the location of the intersection remains randomised. A catalogue of SIMBA galaxies within the cone is then created.

In Fig. 7 SIMBA light cones are shown with co-moving cosmological distance in Mpc given on the bottom x-axis and the y-axis and the corresponding redshift on the top x-axis. The grey circles represent all SIMBA galaxies present within the SIMBA light cone. Grey triangles indicate galaxies that are individually detectable by MeerKAT, but do not form part of a detectable analogous structure. The blue points show galaxies that are both detectable by MeerKAT given the aforementioned detection criteria and grouped by the friends-of-friends algorithm into structures that match our analogue criteria. The red points are analogous groups that would be detectable given the improved collecting area of SKA-mid, but with the same integration time. In Fig. 7 we show four sample light cones that occurred within our simulated observations. In the topmost light cone, we see an example of a scenario where only SKA-mid would be able to detect the analogous structure at $z \sim 0.08$. In the second example, MeerKAT (and SKA-mid) would be able to detect the analogue at $z \sim 0.077$. In the third example, we see numerous galaxies individually detectable by MeerKAT. However, neither MeerKAT nor SKA-mid detect any analogous structures. In the final example MeerKAT detects two analogues while SKA-mid improves these detections by adding members below the MeerKAT detection threshold, as well as detecting an additional analogous structure. These examples showcase that some groups may remain invisible to our instruments as a result of flux sensitivity. However, the analogue criteria also play a role as there are clearly detectable overdensities that simply do not meet the analogue criteria set in this work, highlighting the importance of group structure definitions.

In Fig. 8, which shows the projected distance between the furthest group members, D_p , as a function of the total H I mass of the group, $M_{\text{HI}}^{\text{tot}}$, the red cross shows the group studied in this work at $\log(M_{\text{HI}}^{\text{tot}}/M_{\odot}) = 10.9$ and $D_p \sim 3.7 \text{ Mpc}$. The detection reported in this work lies well within the scatter of the galaxy groups found in SIMBA, demonstrating that the group studied in this work is not extreme with respect to the simulation.

In 300 realisations of randomly generated light cones, a mean of 0.37 groupings is associated (and detectable by a MeerKAT observation with similar integration time) per pointing after the selection criteria have been imposed and the friends-of-friends algorithm implemented. This together with Fig. 7 and Fig. 8 suggests that extended group structures matching our analogue definition are fairly common within the SIMBA simulation.

4.5.4. Scaling relations of SIMBA galaxies

We investigate the position of the simulated SIMBA galaxies on the Parkash et al. (2018) scaling relations in Fig. 5 and Fig. 6. The probability density function of the respective data sets is estimated using Gaussian kernel smoothing with Scott bandwidth determination from SCIPY. Contours are then placed such that the dashed line encloses 68% of the data points based on the distribution estimate. In both figures, the analogue members are galaxies belonging to analogous structures identified within the light cones as per Section 4.5.2 and are shown in dashed-line

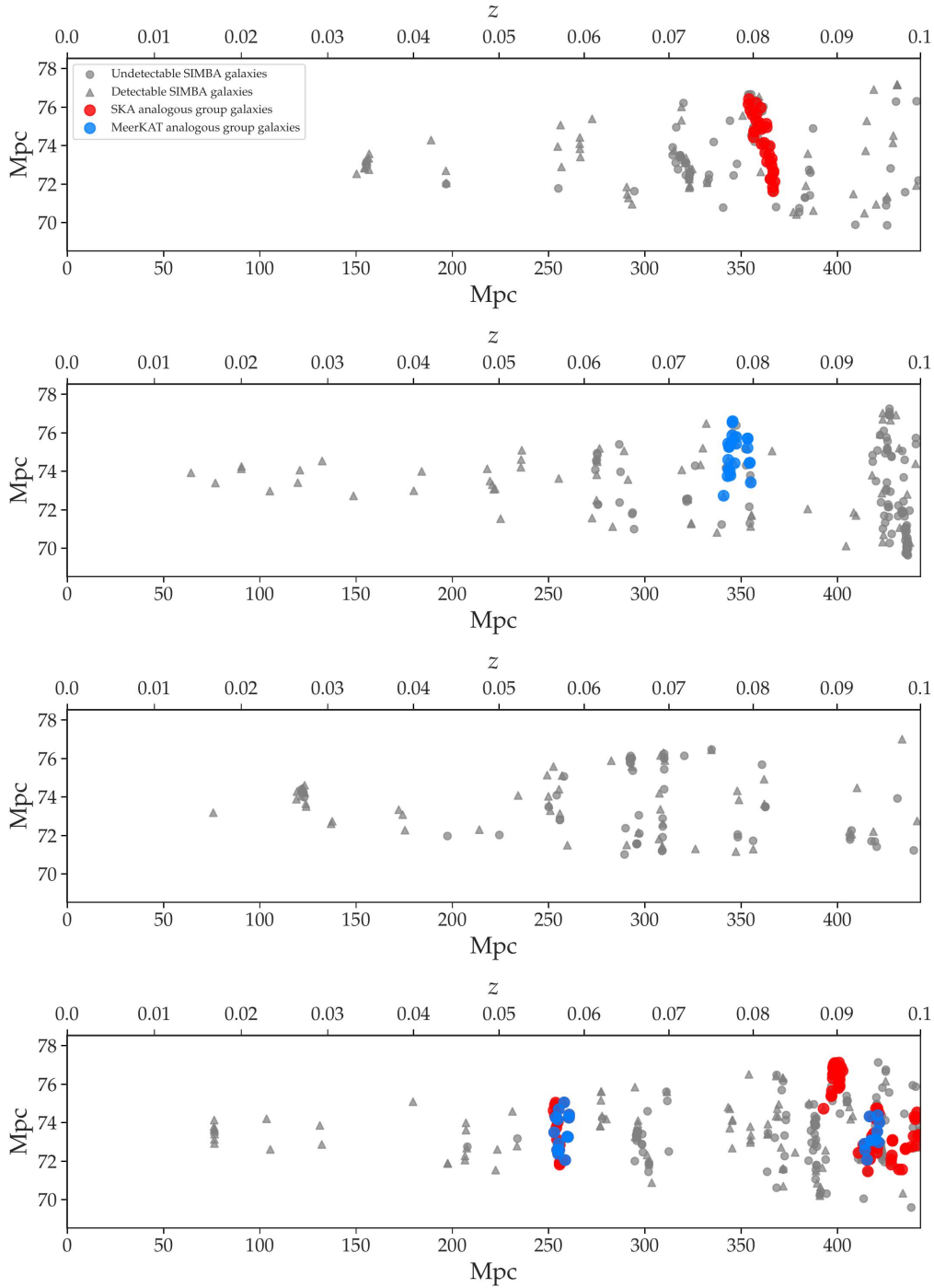


Fig. 7. Depiction of example light cones generated with SIMBA. The bottom x-axis and the y-axis refer to the cosmological co-moving distance. The top x-axis shows the corresponding redshift. Blue points are galaxies that would be detectable with a MeerKAT observation analogous to the one in this work. Red points show galaxies that would be above the detection threshold given a similar observation with the addition of 80 SKA-mid dishes. Grey circles show galaxies that do not form part of either MeerKAT or SKA-mid detectable structures, given an observation with equal integration time. Grey triangles show galaxies that are individually detectable by MeerKAT, but do not belong to a detectable analogues structure.

blue contours. The ‘field galaxies’ are randomly selected galaxies within the SIMBA simulation that meet the detection criteria in Section 4.5.1 and are shown in dashed-line yellow contours.

Generally, the subsets of SIMBA galaxies probed in this work are in agreement with the $M_{\text{HI}}-M_*$ relation from Parkash et al. (2018) as seen in Fig. 5. The most notable feature present is the lobe in the contour of the analogue members in the high- M_*

regime. This suggests that analogue members are more likely to be found above $\sim \log(M_*/M_\odot) = 10$ compared to field galaxies. Additionally, the analogue members are on average HI poor when compared to field galaxies. Similarly, in Fig. 6, the SFMS shows good general agreement between the Parkash et al. (2018) relation and the SIMBA galaxies. The same high- M_* regime lobe appears in the plot. Again, we see the field galaxies are on

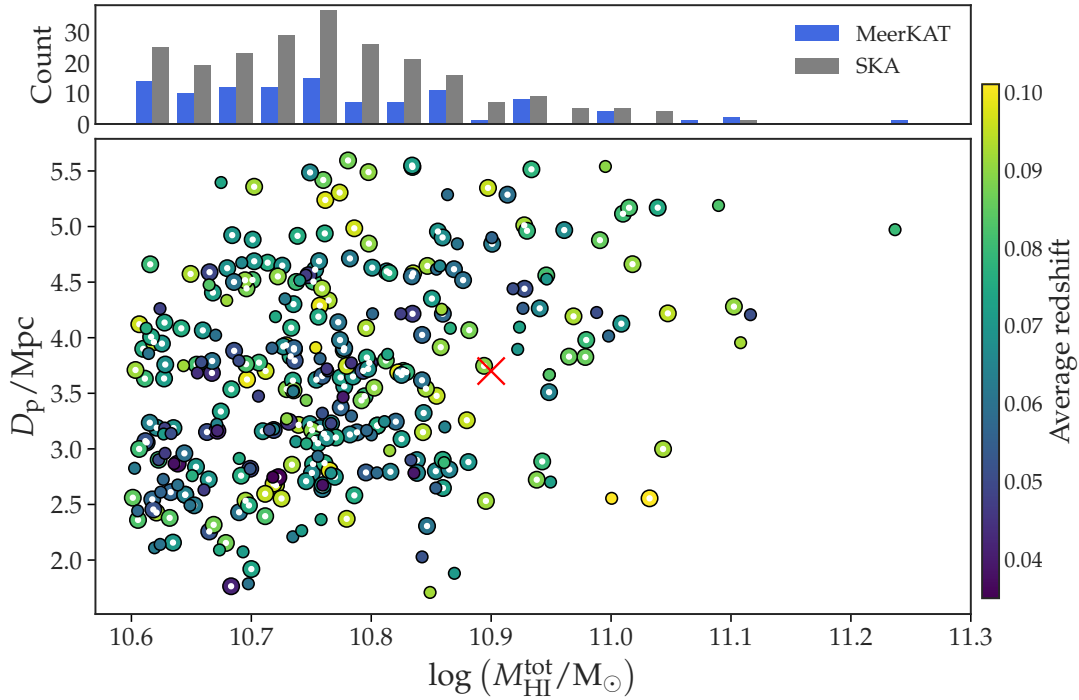


Fig. 8. Scatter plot of detectable SIMBA analogous structures found within 300 light cone realisations. Here, the projected distance between furthest group members, D_p , is plotted as a function of total group HI mass, $\log(M_{\text{HI}}^{\text{tot}})$. Open circles represent SKA-mid-detectable groups, while solid circles represent MeerKAT-detectable groups. All markers are coloured by the mean redshift of the galaxies within the group. The red cross shows the position of the MeerKAT observational structure detected in this work. A histogram is displayed above the plot showing the number of analogues detectable by SKA and MeerKAT for mass bins of 0.035 dex.

average more star-forming than the analogue members indicated by the position of the 68% contours. These trends hint towards evolutionary differences between field galaxies and analogue members, possibly in connection with the enhanced interaction rate observed in denser local environments. In both Fig. 5 and Fig. 6 the most obvious outlier from our observational sample, when compared to the SIMBA samples, is ID 8, a clearly interacting system of galaxies with an inflated stellar mass.

5. Conclusions

We have presented a multiwavelength analysis of a grouping of 20 HI-rich galaxies at $z = 0.04$ serendipitously discovered with MeerKAT. Ancillary data from DES and WISE were used to enrich the study by enabling identification of optical counterparts using high-resolution g -band imaging and derivation of stellar masses and SFRs. This, together with the wealth of HI information (HI moment maps, integrated line profiles, and masses), allowed us to make comparisons to scaling relations from (Parkash et al. 2018).

A number of interacting systems and interesting individual galaxies within the system are present in this filamentary-like structure. In particular, we have noted the presence of a compact group near the projected geometric mean of the group. Within this compact system, we highlighted galaxies that appear to be disturbed through tidal interactions with companion galaxies and a potentially ram-pressure-stripped galaxy. Furthermore, many galaxies in the full sample have clear MeerKAT L -band continuum counterparts, which are typically co-located with the galaxy nucleus, indicating potential AGN or star formation activity.

The properties of the structure as a whole were determined and compared with other recent MeerKAT group discoveries

such as Ranchod et al. (2021) and Glowacki et al. (2024). Lastly, we placed the group in context within the large-scale structure of the Universe, and we probed the serendipity of MeerKAT for HI discovery using the SIMBA cosmological hydrodynamical simulation, concluding that the discovery is not extreme in the context of the simulation. This suggests the sample size of such objects will grow substantially, allowing for statistical studies of their relation to the large-scale structure, specifically such parameters as galaxy spin orientation, atomic gas fraction, and morphology with respect to dark matter filaments.

Data availability

The raw visibilities used in this work can be found at the SARAO archive¹³. The authors may make data products available upon reasonable request.

Acknowledgements. We thank the anonymous referee for their insightful comments and constructive suggestions, which helped improve the clarity and quality of this work. We would like to thank the staff of the South African Radio Astronomy Observatory (SARAO, www.sarao.ac.za) who made these observations possible. We thank Jacques Smulders for his technical assistance with the light cone production code. We also thank Mpati Ramatsoku and Gyula (Josh) Józsa for helpful discussions. GDL and RPD acknowledge funding from the South African Radio Astronomy Observatory (SARAO), which is a facility of the National Research Foundation (NRF), an agency of the Department of Science and Innovation (DSI). RPD acknowledges funding by the South African Research Chairs Initiative of the DSI/NRF (Grant ID: 77948). The MeerKAT telescope is operated by the South African Radio Astronomy Observatory, which is a facility of the National Research Foundation, an agency of the Department of Science and Innovation. We acknowledge the use of the ILIFU (www.ilifu.ac.za) cloud computing facility, a partnership between the University of Cape Town, the University of the Western Cape, Stellenbosch University, Sol Plaatje University, the Cape Peninsula University of Technology and the South African

¹³ <https://archive.sarao.ac.za>

Radio Astronomy Observatory. The ILIFU facility is supported by contributions from the Inter-University Institute for Data Intensive Astronomy (IDIA – a partnership between the University of Cape Town, the University of Pretoria and the University of the Western Cape), the Computational Biology division at UCT and the Data Intensive Research Initiative of South Africa (DIRISA). This work has made use of the Cube Analysis and Rendering Tool for Astronomy (CARTA, Comrie et al. 2021). GDL and RPD acknowledge funding from the Wits Foundation UK in support of travel for research. This research made use of Astropy (<http://www.astropy.org>), a community-developed core Python package for Astronomy. This research makes use of the semi-automated MeerKAT data calibration pipeline, OXKAT (<https://github.com/IanHeywood/oxkat>).

References

- Abbott, T. M. C., Adamów, M., Agüena, M., et al. 2021, *ApJS*, **255**, 20
- Berlind, A. A., Frieman, J., Weinberg, D. H., et al. 2006, *ApJS*, **167**, 1
- Blyth, S., Baker, A. J., Holwerda, B., et al. 2016, in *MeerKAT Science: On the Pathway to the SKA*, 4
- Broeils, A. H., & Rhee, M. H. 1997, *A&A*, **324**, 877
- Bulichi, T.-E., Davé, R., & Kraljic, K. 2024, *MNRAS*, **529**, 2595
- Camilo, F., Scholz, P., Serylak, M., et al. 2018, *ApJ*, **856**, 180
- Catinella, B., Schiminovich, D., Cortese, L., et al. 2013, *MNRAS*, **436**, 34
- Cluver, M. E., Jarrett, T. H., Hopkins, A. M., et al. 2014, *ApJ*, **782**, 90
- Cluver, M. E., Jarrett, T. H., Dale, D. A., et al. 2017, *ApJ*, **850**, 68
- Comrie, A., Wang, K. S., Hsu, S. C., et al. 2021, <https://doi.org/10.5281/zenodo.4905459>
- Crook, A. C., Huchra, J. P., Martimbeau, N., et al. 2007, *ApJ*, **655**, 790
- Davé, R., Thompson, R., & Hopkins, P. F. 2016, *MNRAS*, **462**, 3265
- Davé, R., Anglés-Alcázar, D., Narayanan, D., et al. 2019, *MNRAS*, **486**, 2827
- Davé, R., Crain, R. A., Stevens, A. R. H., et al. 2020, *MNRAS*, **497**, 146
- de Blok, W. J. G., Healy, J., Maccagni, F. M., et al. 2024, *A&A*, **688**, A109
- Deka, P. P., Gupta, N., Jagannathan, P., et al. 2024, *ApJS*, **270**, 33
- Eke, V. R., Frenk, C. S., Baugh, C. M., et al. 2004, *MNRAS*, **355**, 769
- Glowacki, M., Albrow, L., Reynolds, T., et al. 2024, *MNRAS*, **529**, 3469
- Gunn, J. E., & Gott, J. R., III 1972, *ApJ*, **176**, 1
- Gupta, N., Srianand, R., Baan, W., et al. 2016, in *MeerKAT Science: On the Pathway to the SKA*, 14
- Hess, K. M., & Wilcots, E. M. 2013, *AJ*, **146**, 124
- Hess, K. M., Serra, P., Boschman, L., Shen, A., & Healy, J. 2022, <https://doi.org/10.5281/zenodo.6821586>
- Heywood, I. 2020, Astrophysics Source Code Library [record ascl:2009.003]
- Heywood, I., Ponomareva, A. A., Maddox, N., et al. 2024, *MNRAS*, **534**, 76
- Huchra, J. P., & Geller, M. J. 1982, *ApJ*, **257**, 423
- Huchra, J. P., Macri, L. M., Masters, K. L., et al. 2012, *ApJS*, **199**, 26
- Jarrett, T. H., Cluver, M. E., Taylor, E. N., et al. 2023, *ApJ*, **946**, 95
- Jarvis, M., Taylor, R., Agudo, I., et al. 2016, in *MeerKAT Science: On the Pathway to the SKA*, 6
- Jonas, J. L. 2009, *IEEE Proc.*, **97**, 1522
- Jonas, J., & MeerKAT Team 2016, *MeerKAT Science: On the Pathway to the SKA*, 1
- Jones, D. H., Read, M. A., Saunders, W., et al. 2009, *MNRAS*, **399**, 683
- Jones, M. G., Haynes, M. P., Giovanelli, R., & Moorman, C. 2018, *MNRAS*, **477**, 2
- Kennicutt, R. C., Jr, Armus, L., Bendo, G., et al. 2003, *PASP*, **115**, 928
- Kennicutt, R. C., Calzetti, D., Aniano, G., et al. 2011, *PASP*, **123**, 1347
- Kenyon, J. S., Smirnov, O. M., Grobler, T. L., & Perkins, S. J. 2018, Astrophysics Source Code Library [record ascl:1805.031]
- Kleiner, D., Pimblet, K. A., Jones, D. H., Koribalski, B. S., & Serra, P. 2017, *MNRAS*, **466**, 4692
- Kleiner, D., Serra, P., Maccagni, F. M., et al. 2021, *A&A*, **648**, A32
- Knowles, K., Cotton, W. D., Rudnick, L., et al. 2022, *A&A*, **657**, A56
- Laurikainen, E., & Moles, M. 1989, *ApJ*, **345**, 176
- Lelli, F., McGaugh, S. S., Schombert, J. M., Desmond, H., & Katz, H. 2019, *MNRAS*, **484**, 3267
- Leroy, A. K., Walter, F., Brinks, E., et al. 2008, *AJ*, **136**, 2782
- Lim, S. H., Mo, H. J., Lu, Y., Wang, H., & Yang, X. 2017, *MNRAS*, **470**, 2982
- Loubser, S. I., Mosia, K., Serra, P., et al. 2024, *MNRAS*, **527**, 7158
- Lovell, C. C., Geach, J. E., Davé, R., Narayanan, D., & Li, Q. 2021, *MNRAS*, **502**, 772
- Lovisari, L., Ettori, S., Gaspari, M., & Giles, P. A. 2021, *Universe*, **7**, 139
- Maddox, N., Frank, B. S., Ponomareva, A. A., et al. 2021, *A&A*, **646**, A35
- McMullin, J. P., Waters, B., Schiebel, D., Young, W., & Golap, K. 2007, *ASP Conf. Ser.*, **376**, 127
- Meyer, M., Robotham, A., Obreschkow, D., et al. 2017, *PASA*, **34**, 52
- Moore, B., Lake, G., & Katz, N. 1998, *ApJ*, **495**, 139
- Nielsen, N. M., Kacprzak, G. G., Pointon, S. K., et al. 2020, *ApJ*, **904**, 164
- Offringa, A. R., McKinley, B., Hurley-Walker, N., et al. 2014, Astrophysics Source Code Library [record ascl:1408.023]
- Parkash, V., Brown, M. J. I., Jarrett, T. H., & Bonne, N. J. 2018, *ApJ*, **864**, 40
- Pillepich, A., Springel, V., Nelson, D., et al. 2018, *MNRAS*, **473**, 4077
- Planck Collaboration VI. 2020, *A&A*, **641**, A6
- Planck Collaboration XIII. 2016, *A&A*, **594**, A13
- Ponomareva, A. A., Mulaudzi, W., Maddox, N., et al. 2021, *MNRAS*, **508**, 1195
- Ramella, M., Pisani, A., & Geller, M. J. 1997, *AJ*, **113**, 483
- Ranchod, S., Deane, R. P., Ponomareva, A. A., et al. 2021, *MNRAS*, **506**, 2753
- Robotham, A. S. G., Norberg, P., Driver, S. P., et al. 2011, *MNRAS*, **416**, 2640
- Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, *MNRAS*, **446**, 521
- Serra, P., Westmeier, T., Giese, N., et al. 2015, *MNRAS*, **448**, 1922
- Serra, P., Maccagni, F. M., Kleiner, D., et al. 2023, *A&A*, **673**, A146
- Springel, V., White, S. D. M., Jenkins, A., et al. 2005, *Nature*, **435**, 629
- Springel, V., Pakmor, R., Pillepich, A., et al. 2018, *MNRAS*, **475**, 676
- Sulentic, J. W., Rosado, M., Dultzin-Hacyan, D., et al. 2001, *AJ*, **122**, 2993
- Tempel, E., Tamm, A., Gramann, M., et al. 2014, *A&A*, **566**, A1
- Tumlinson, J., Peebles, M. S., & Werk, J. K. 2017, *ARA&A*, **55**, 389
- Westmeier, T., Kitaef, S., Pallot, D., et al. 2021, *MNRAS*, **506**, 3962
- Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, *AJ*, **140**, 1868
- Yang, X., Mo, H. J., van den Bosch, F. C., & Jing, Y. P. 2005, *MNRAS*, **356**, 1293
- Yang, X., Mo, H. J., van den Bosch, F. C., et al. 2007, *ApJ*, **671**, 153
- Yang, X., Xu, H., He, M., et al. 2021, *ApJ*, **909**, 143
- Yun, M. S., Ho, P. T. P., & Lo, K. Y. 1994, *Nature*, **372**, 530

Appendix A: H I spectra



Fig. A.1. Integrated H I spectra of all group galaxies. Error bars are derived from the per-channel RMS of the local noise within the bounding box of the source.

Appendix B: Multiwavelength imaging of group galaxies and sources

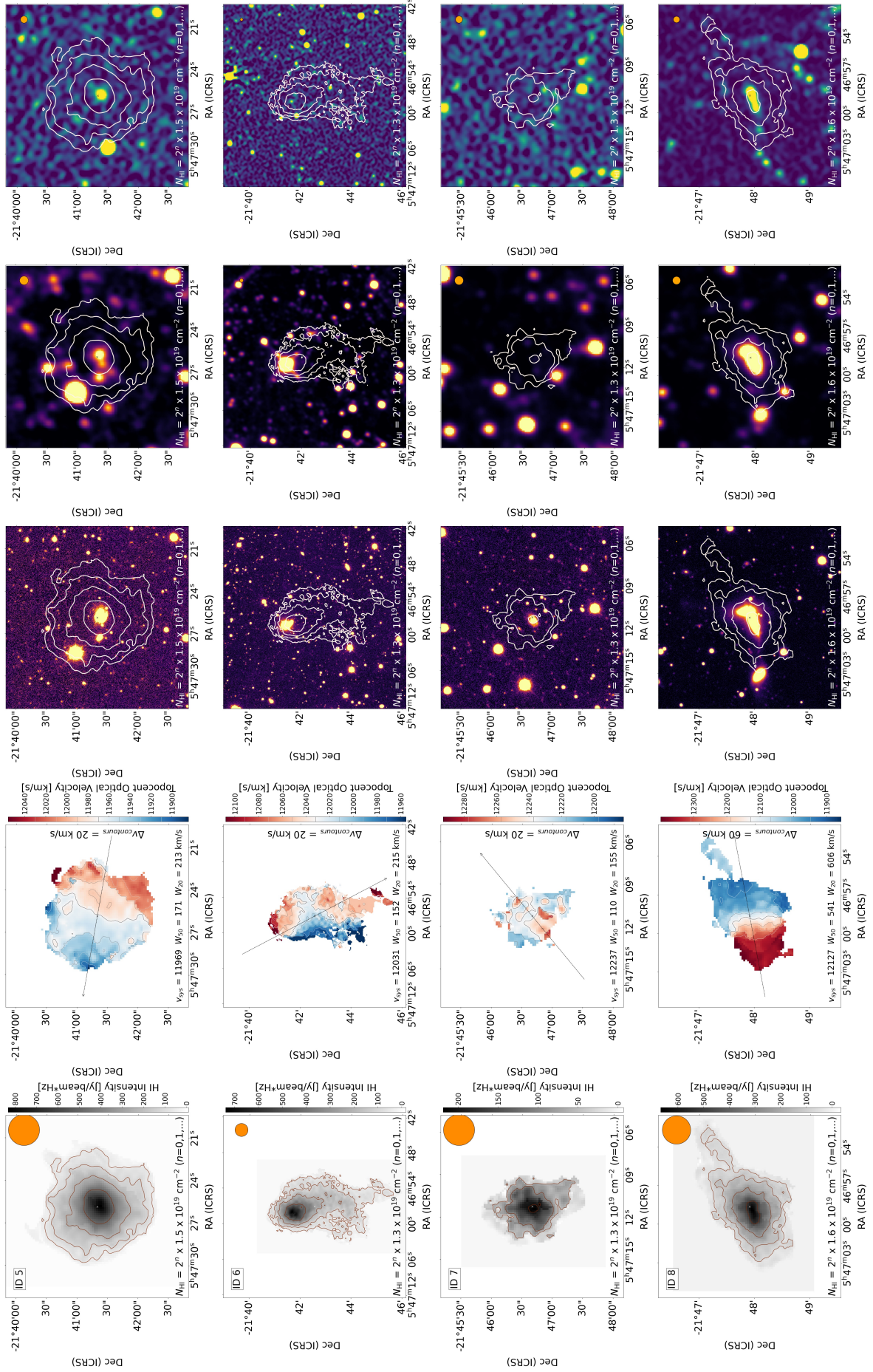


Fig. B.1. Same as Fig. 3, but for source IDs 5-8.

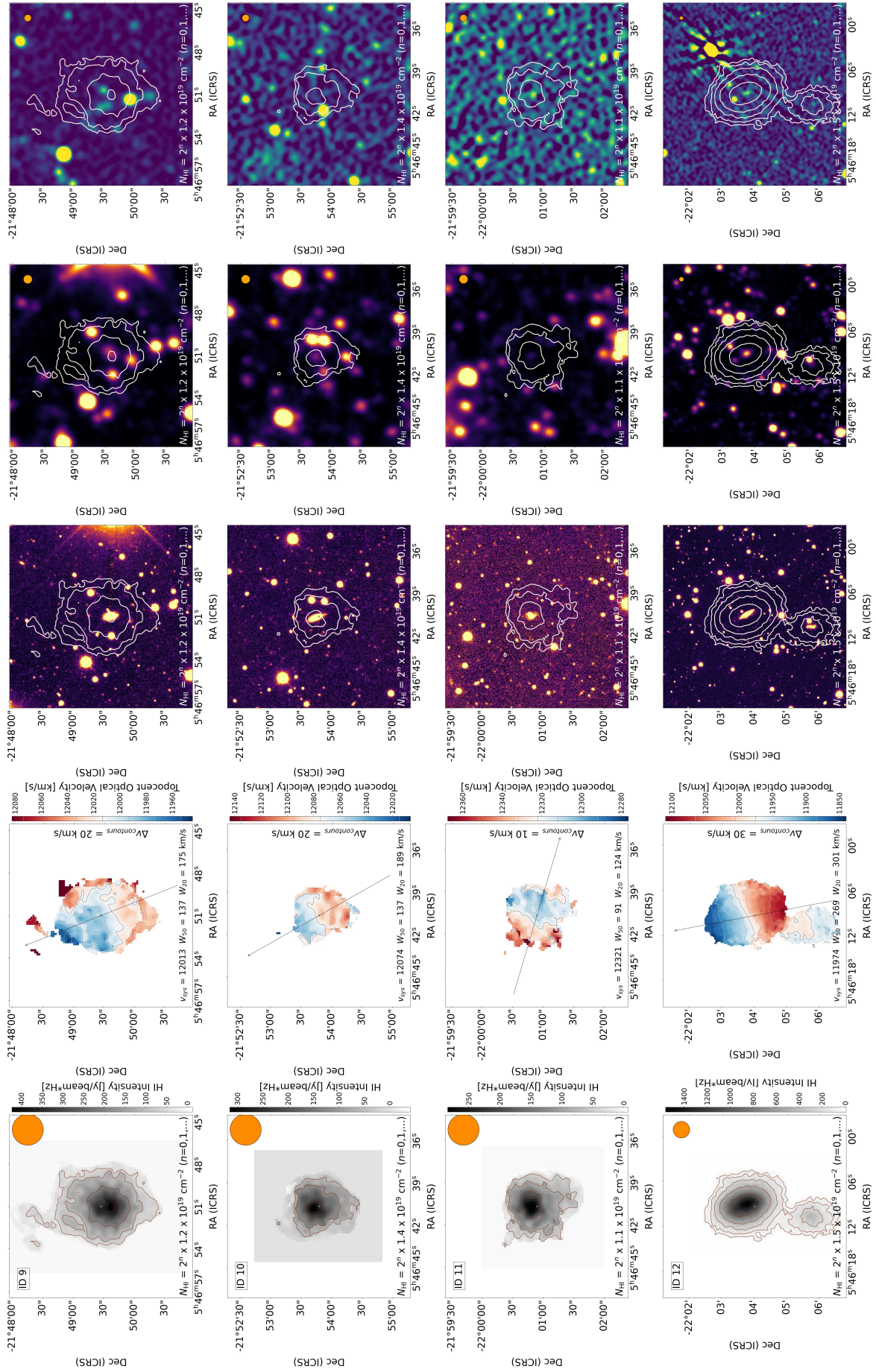


Fig. B.2. Same as Fig. 3, but for source IDs 9-12.

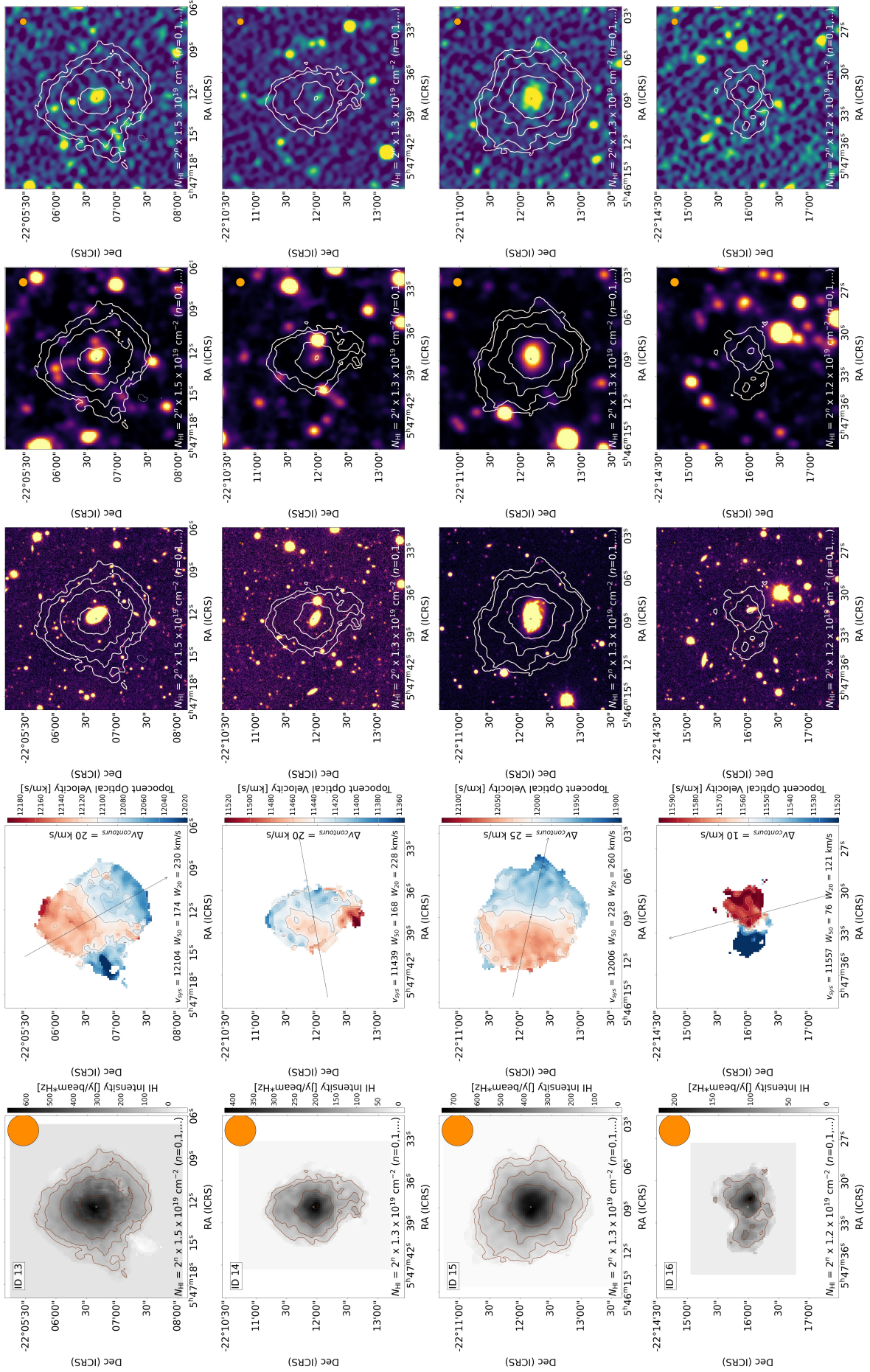


Fig. B.3. Same as Fig. 3, but for source IDs 13-16.

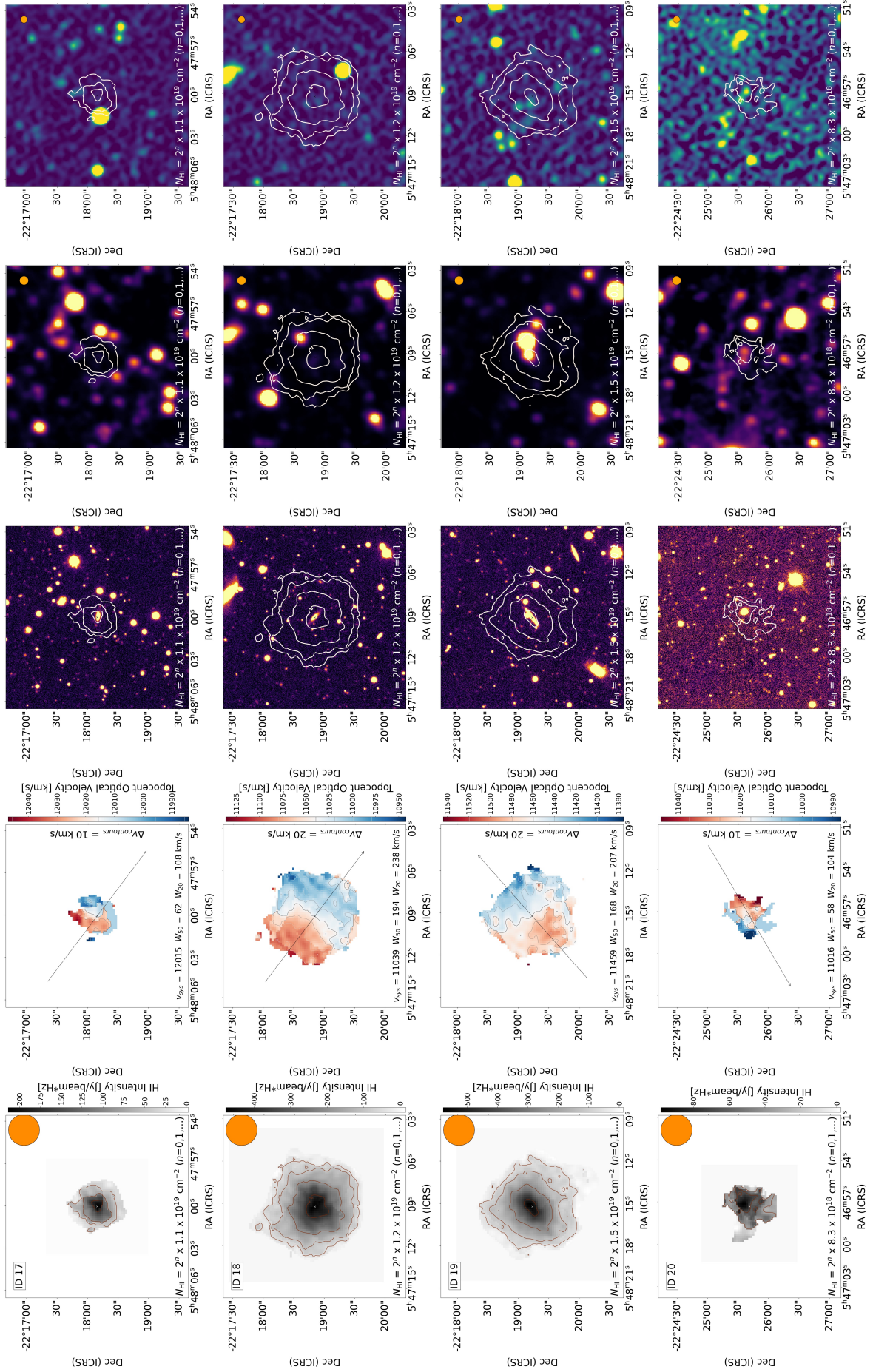


Fig. B.4. Same as Fig. 3, but for source IDs 17-20.