Rijksuniversiteit Groningen

Bachelor Thesis

Resolved HI characteristics of potential gas-bearing dark matter halo AGC 268071

Author:
A.J. Blonk

Supervisor:
Prof. dr. E.A.K. Adams
Second supervisor:
Prof. dr. T.A. Oosterloo

Abstract

This study presents insight into the nature of UCHVC AGC 268071, through resolved imaging at original 45" data and 105" resolution data smoothed in the image plane. The source was detected by ALFALFA and has a possible stellar counterpart at 1.5 kpc seperation from the ALFALFA HI detection. The possible nature scenarios discussed are: a gas bearing dark matter mini halo or a gas cloud in the Galactic halo. The interferometric data are flagged, calibrated and cleaned within masked regions. The HI morphology of AGC 268071 appears compact at both resolutions with a peak column density of 3.17 x 10^{19} N_{HI}/cm^{-2} and 9.34 x 10^{18} N_{HI}/cm^{-2} for 45" and 105" respectively. The source displays a relatively small velocity gradient across its surface; 5 km s^{-1} and 3 km s^{-1} for 45" and 105" data respectively. The velocity dispersion is roughly 6 km s^{-1} and 5.4 km s^{-1} for 45" and 105" data respectively, corresponding to a large fraction of CNM gas. The extracted HI line spectra in this research has a lower peak value and width than the ALFALFA spectrum, likely due to Galactic background emission influencing the ALFALFA data or missing source emission in the VLA cleaned datacube due to the low SNR of the data. Assuming a galactic nature of AGC 268071, the HI mass found is 1.87-2.72 x $10^4 M_\odot$ for 45" data and 1.31-1.89 x $10^4 M_\odot$ for 105" data. This result means AGC 268071 could be the lowest known HI mass galaxy in the Local Group, but should be considered a lower limit. Based on the compact HI morphology and small velocity gradient AGC 268071 is likely of galactic nature, but the gas cloud in Galactic halo scenario cannot be fully ruled out due to the low measured velocity dispersion.

Contents

1 Introduction

The study of galactic formation and evolution strives for an understanding into the processes that generate variety in galactic structures. With the development of increasingly precise observational tools, realms of extreme galactic structures are being uncovered. One of these extreme galactic structures are the so called 'dark galaxies'; low mass gas bearing dark matter halo's with a non- or barely observable optical stellar counterpart (Rhode et al. (2023)).
The ALFALFA survey observed the neutral hydrogen (HI) emission from ~31,500 sources between 2005 and 2012 (Rhode et al. (2023)). Because of its exceptional sensitivity the ALFA instrument was able to detect relatively low neutral hydrogen (HI) mass objects (~$10^4 - 10^5 M_\odot$), up to around 80 Mpc, well beyond the confines of the Local Group (Giovanelli et al. (2010)). These very low HI brightness and mass sources were labelled as Ultra-Compact High Velocity Clouds (UCHVCs), where the ALFALFA survey detected ~100 of these clouds, as identified by Adams et al. (2013). Giovanelli et al. (2010) proposes that a detection of a barely-observable stellar counterpart in these clouds could confirm or point towards the existance of dark galaxies residing in the local group.

Because these sources have a barely-observable stellar counterparts in the optical spectrum, the nature of the source can be analyzed through a different method than detecting optical starlight. This can be achieved by observing the HI spectral line component of the UCHVC, and by characterizing the HI. The UCHVC analyzed in this research is AGC268071, based on the optical search for a stellar component of 26 UCHVC's in Rhode et al. (2023), where AGC268071 was determined most likely to have an optical stellar component.

1.1 High Velocity Clouds

The first discovery of High Velocity Clouds was by Muller et al. (1963), using the Dwingeloo radio telescope. HVCs are composed of neutral hydrogen with velocities distinguishable from Galactic rotation, corresponding to optical velocities greater or equal to 90 km/s in the heliocentric frame. The origin of high velocity clouds are still not fully understood, partially due to the lack of sufficient information on the clouds distances. The leading hypotheses span from supernova shells at 100 pc, to gas expelled from the disc through galactic fountains, to remnants of the galactic formation era, to fully extra-galactic objects at 1 Mpc (Oort (1970), Bregman (1980), Wakker and Van Woerden (1997), Blitz et al. (1996)).
A couple decades after the discovery of HVCs by Oort another subgroups of HVCs was discovered and classified by Wakker and Van Woerden (1997) and Putman (1999) as Compact High Velocity Clouds (CHVCs). CHVCs are distinguished from HVCs due to their compact, isolated nature of relatively high surface brightness, as opposed to the filamentary, diffuse and extended HVCs.
Adams et al. (2013) proposed certain criteria for the sources to qualify as a (Ultra) CHVC, namely; objects should be compact (HI major axis <30'), should be relatively low mass ($\sim 10^5 - 10^6 M_\odot$), should be relatively isolated; meaning no more than three neighboring HI structures within 3 ° on the sky, should have an absolute velocity greater than 120 km s^{-1} and finally should have velocities distinguishable from Galactic High-Velocity clouds.

Certain High Velocity Clouds have resolved or narrowed down origins. For instance, Putman (1999) identified a subgroup of HVCs as HVCs related to the Magellanic stream. The origin of a large HVC named 'Chain A' was narrowed down by Van Woerden et al. (1999) to either gas expelled through galactic fountains or to a galactic formation era remnant. An HVC with a known origin is Leo T, initially identified as a gas-rich low-mass dwarf galaxy by Irwin et al.

(2007). It is located at a distance of about 420 kpc from the Milky Way and has both neutral hydrogen and a small population of stars, suggesting it is an extragalactic object and not simply a cloud of gas within the Galactic Halo. Adams et al. (2014), Adams et al. (2016) analyzed the HI content of thirteen UCHVCs, where four sources showed clear signs towards a galactic nature. Finally Rhode et al. (2023) searched for an optical stellar component on 59 UCHVCs, where six sources were found to have a potential observable optical stellar component.

1.2 Gas Bearing Dark Matter Halos

Generally galaxies are composed of a relatively large dark matter halo, a stellar component residing within the halo and interstellar media such as gas and dust. As opposed to standard galaxies, dark galaxies lack an easily observable optical stellar counterpart. According to the the Lambda Cold Dark Matter (ΛCDM) model, in the initial universe clumps of dark matter and gas condensed and moved to areas of higher density, forming the first dark matter halos. These halos merged together and collapsed under their own weight, forming proto-galaxies. However, some of the halos might have been too light, contained too little gas to collapse or underwent very little mergers, resulting in the present existence of these dark matter halos. The missing satellite problem (MSP), according to Bullock (2010), refers to the discrepancy between the observed number of low-mass dark matter halo's orbiting the Milky Way, and the number predicted through simulations. Giovanelli et al. (2010) proposed the new found UCHVCs with ALFALFA could be the missing dark matter minihalos, or atleast provide a partial solution, Lee et al. (2024). At current day the MSP is largely solved, except at the extreme realms of ultra-low-mass galaxies residing in the transition regime. The transition regime is the mass range where galaxies start to transition from being mostly dark matter halos with very few stars (if any) to having more substantial baryonic (normal matter) components. Here the characteristics of the galaxies are not fully mapped yet and require more observations to verify the number of low mass galaxies in the local Group found through models or simulations.

1.3 Instrument Description

Figure 1: Above view of the Very Large Array in D-configuration. Maximum baselength of each arm is now 1 kilometer. Image from Perley et al. (2011).

Due to the difficulty of observing an UCHVC source through optical imaging, the neutral hydrogen content can be observed and analyzed to infer the likeliness of the source being a galaxy. The transition between two energy states of neutral hydrogen emits radiation in the radio spectra known as the 21 centimeter line. Even in the radio spectra the UCHVCs are relatively compact, requiring the use of an interferometer to map an area of the sky with a high resolution. Interferometers consist of multiple telescopes working together to increase the effective aperture and resolution by combining signals from widely spaced antennas, effectively simulating a much larger single telescope (Wilson et al. (2013)).

The choice of telescope for this study is the Karl G. Jansky Very Large Array (VLA) due to it being one of the most prominent northern-hemisphere multi-dish radio telescope. The VLA is stationed in New Mexico in the United States of America. It consists of 27 individual radio telescopes arranged in a Y-shaped pattern. Each arm in the Y pattern reaches to about 21 kilometers and

R.A.	Decl.	S_{21} (Jy km s^{-1})	cz (km s^{-1})	ω_{50} (km s^{-1})	$\overline{a}$ (')	log $\overline{N}_{HI}$ (atoms cm^{-2})	logM_{HI} ($M_{\odot}$)
16:12:33.9	+14:11:59.0	2.67 ± 0.08	109	62 ± 15	9.75	19.09	5.80

Table 1: Properties of AGC 268071 taken from Rhode et al. (2023) and the observational log Cannon et al. (2023). Table contains the location of source in R.A and Decl. values, the integrated line flux from ALFALFA, the absolute velocity of the source, the FWHM value, the HI major axis, the peak column density and the HI mass.

each radio dish is 25 metres in diameter[1]. Along the arms are rails installed allowing for varying configurations of the telescopes depending on the observing goal. According to Kellermann et al. (2020), the VLA was built primarily to research the extragalactic contiuum, but has been used for varied fields of astronomy.

1.4 AGC 268071

The target AGC 268071 was first detected by ALFALFA HI survey and identified as an UCHVC by Adams et al. (2013). Amongst the 59 UCHVCs detected in that study AGC 268071 had the largest found FWHM (ω_{50}) with a value of 62 km s^{-1} (see Table 1). The absolute velocity of AGC 268071 is 109 km s^{-1}, which is smaller than the UCHVC criteria proposed by Adams et al. (2013). A decade later Rhode et al. (2023) detected an optical stellar component during an observation of AGC 268071 as depicted in figure 2. The figure shows a clear offset between the location of the HI ALFALFA observation and the optical observation of the source. The offset corresponds to a physical separation of 1.5 kpc (Rhode et al. (2023)). It therefore cannot be excluded that the latter observation was of a random stellar overdensity rather than the intended source.

Figure 3 shows the HI line spectra as observed by ALFALFA. It is apparent the detection of AGC 268071 is convoluted with emission from the Galactic neutral hydrogen. The dashed line in the figure shows the zero baseline of the flux calculation, implying a Galactic background emission of ~15 mJy.

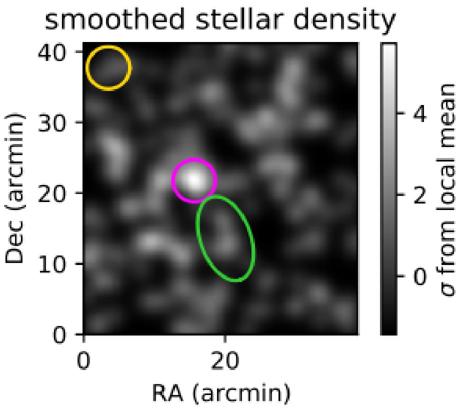


Figure 2: This figure shows the location of HI detected by ALFALFA (green ellipse) and the 3' detection circle of a stellar overdensity by Rhode et al. (2023)(magenta circle). A yellow reference circle of the same size is placed at a random location in the corner of the field. The offset between the HI detection and stellar detection is 1.5 kpc. In the 3' detection circle 11 stars were detected, whereas in the reference circle 4 stars were detected.

---

[1] https://public.nrao.edu/vla-configurations/

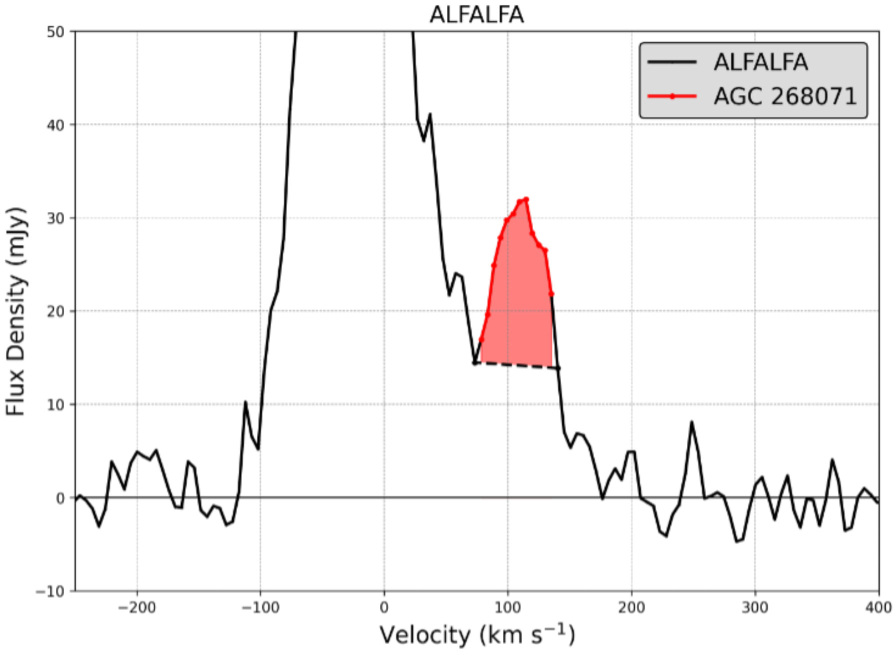


Figure 3: HI line spectrum of ALFALFA detection of AGC 268071. Figure taken from Cannon et al. (2023). The red shaded area is the source emission, whereas the large peak to the left of the source is the Galactic HI emission. A seperation of source emission and Galactic HI emission is shown by the dashed horizontal line.

2 Data Acquisition and Calibration

2.1 Observational Details

The observation of target AGC 268071 was excecuted in the 20 centimeter L-band over a frequency range from 1 to 2 GHz. The L-band receivers were used to observe the neutral hydrogen spectral line at a frequency of 1420.405 MHz, over a subset of the full band. A second band observed the 1-2 GHz continuum background, but this was later discarded and only the first band was used. As visible in Table 1, AGC 268071 moves with a heliocentric recessional velocity of 109 km s^{-1} in the optical velocity frame, meaning the HI spectral lines are redshifted. A straight-forward calculation tells us that the neutral hydrogen is redshifted to a frequency of 1419.9 MHz, as opposed to the non-redshifted HI frequency of 1420.405 MHz.
The neutral hydrogen spectral window used 8-bit samplers and implemented a 16 MHz sub-band with dual polarization products, distributed over 4096 channels. The native spectral resolution per channel was 3.9 KHz, 0.82 km s^{-1} per channel. Besides dual polarization, eight 128 MHz sub-bands with full polarization products were implemented into 128 subchannels per sub-band. The beamwidth in L-band is approximately 45 arcseconds. In total six datasets were aquired in two consecutive observation dates, each observation lasting two hours. A dump rate of 3 seconds produced a data rate of 42.4 GB/hour (Cannon et al. (2023)).

The flux calibrator used for this observation was 3C286, also known as J1331+305. 3C286 is a quasar located in the constellation Canes Venatici at redshift 0.8493, with a radial velocity of 164,137 km s^{-1} and a flux of 15.0 Jy[2]. 3C286 has a right ascension of 13:31:08.288 and a declination of +30 30' 32.9600" (Schneider et al. (2007)). The gain calibrator used for this observation was J1640+1220, also known as NRAO512. J1640+1220 is classified as a quasar of redshift 1.666, has a flux of 2.10 Jy and a UV limit of 60 kλ[2]. A UV limit corresponds to a maximum baseline length at which the source is resolved, meaning baselines longer than this length should be avoided. In this case the wavelength at which we're observing is 20 cm, which corresponds to a maximum baseline of ~12 km. This is therefore not something relevant for an observation using the D-configuration with a maximum baseline of 1km[1]. The gain calibrator has a right ascension of 16:40:29.633 and a declination of +39 46' 46" 03 (Schinzel et al. (2011)).

The data utilized in this research is requested from the National Radio Astronomy Observatory (NRAO) archive. The archive houses data from various radio telescopes as the VLA, VLBA, GBT and ALMA.
After unpacking the data, the line data was first split off from the continuum data to save storage space as the data was chosen to be manually flagged and calibrated. The software used to process the data is the Common Astronomy Software Application package (CASA)[3], the primary data processing software for the VLA. The Kapteyn computers have different versions of casa installed already, the version used here is the most recent one, version 6.6.0.
From the aforementioned six datasets created during the observation of AGC 268071, only four were used during the span of this research. The other two datasets were previously processed by Dr. J. Cannon, and were purposely left out here to combine the fully processed datasets in a later project. The datasets we continue with are eb44979672, eb44991749, eb44977689 and eb44980678. Table 2 contains an overview of information regarding the different observations.

---

[2] https://science.nrao.edu/facilities/vla/observing/callist

[3] https://casa.nrao.edu/

Observation	Date	Starting Time (UTC)	End Time (UTC)
eb44977689	31-12-'23	15:11:27	17:11:03
eb44979672	01-01-'24	14:39:30	16:20:09
eb44980678	01-01-'24	16:39:09	18:19:51
eb44991749	04-01-'24	14:10:45	16:10:21

Table 2: Table providing an overview of different execution blocks during observation of AGC 268071. Table contains name, observing date and start and beginning time of observation.

2.2 Calibration Procedures

2.2.1 Flagging and Initial Calibration

The first step in the data flagging procedures was to remove obstructed data influenced by the shadow of other antenna's covering the receivers. The data points with a zero amplitudes also needed to be filtered out, as these datapoints represent areas where the signal is absent. Furthermore, as the antenna's turn to a target they oscillate slightly before pointing perfectly to the target. The data of the initial first seconds of an observation therefore need to be flagged as well. Automatic Radio Frequency Interference (RFI) flagging was not relevant for this data, because the observation occurred in a protected band with negligible RFI.

The general calibration of the data again contained a couple procedures. Even though the VLA has standard configurations, some of the antenna's can experience a slight offset from their computed positions. The offset positions may be published a few days after the observations and stored in a table that CASA can access. The offsets can therefore be accessed and applied to the antenna positions if necessary. Furthermore, the gain of an antenna is not always constant across all elevation angles. The gain curve calibration process characterizes these variations by measuring the antenna's response at different elevation angles, and by writing these variations in a calibration table that can later be applied to the data.

An opacity correction is only relevant for high frequency observations, since atmospheric components are more effective at absorbing and scattering higher frequencies (Wilson et al. (2013)). At long frequencies such as the L-band at which we're observing, the attenuation is minimal and can be neglected.

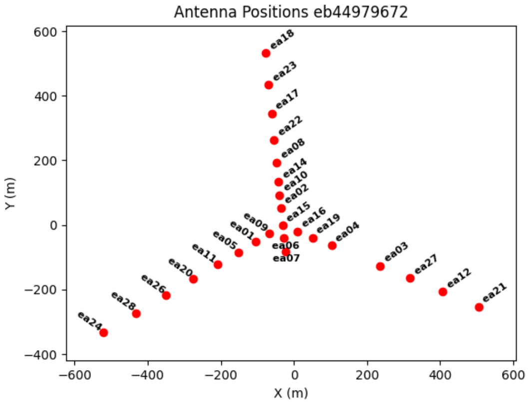


Figure 4: Antenna location of Very Large Array in D-configuration. Depicted here is only one of the four analyzed datasets, as the antenna positions are invarient over the different observations. Figure made with CASA using antpos command.

2.2.2 Setting The Flux Calibratior

As mentioned in section 2.1 the flux calibrator used throughout the calibration process is pulsar 3C268. This source is chosen because of it's high SNR and stable, well-characterized nature in the L-band. Once an appropriate recognized flux calibrator has been identified, the CASA task setjy computes the spectral flux density of the calibrator as a function of frequency. It then puts the model visibility amplitude and phase, along with a representation of clean components, into the model column of the measurement set[3]. Clean components are specific sources or regions of emission identified by the CLEAN algorithm as significant contributors to the signal, more about this is explained in the imaging section. The output of the task gave a flux density value of 15.02 Jy for 3C268.

2.2.3 Antenna Based Delays

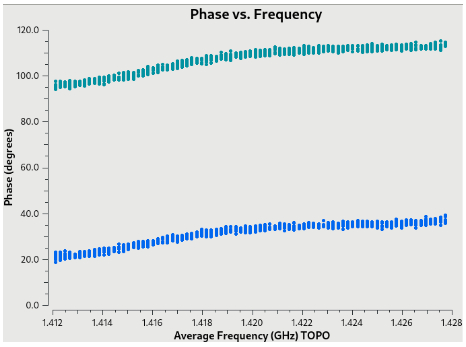


Figure 5: Plot of phase versus frequency for baseline ea05&ea08 to visualize the antenna based delays that arise between antennas and the receivers. The two colors indicate the two polarization degrees of the observation. The antenna based delay is visible as the slope in the phase versus frequency of a single baseline. Plot made while averaging 50 channels and 50 seconds.

Then moving on to more specific types of calibration, a reference antenna is required to serve as a stable phase reference point against which the phases of signals from other antenna's can be compared and corrected to. The reference antenna is chosen based on it's position within the configuration (figure 4), as intermediate baselines are favorable, and based on the antenna's range of u-v distances. However, that criteria is less applicable when observing in the D-configuration, as most antenna's will have similar u-v coverage and there are no excessively long baselines.

Working with an interferometer, delays can arise in the signal propagation path between antennas and the receiver system. Such a delay presents itself as a slope in the phase of a single baseline, as a function of frequency[4]. Figure 5 shows such a frequency dependent delay for the baseline between antennas ea05 and ea08 in dataset eb44991749. An uncorrected delay can cause the continuum signal to appear de-correlated. The correction of the delay is done using the flux calibrator due to its high signal-to-noise ratio. The stable and well-understood nature of the flux calibrator ensures that the measured delays across baselines are due to instrumental and geometric effects, rather than the variability of the source. The delays are computed by performing a fast fourier transform (FFT) on the visibility data from each baseline and by fitting the resoluting time-domain phase variations to a linear model[4]. The slope of this linear fit represents the delay, and gets stored in a calibration table to later be applied to the data.

2.2.4 Bandpass Calibration

After delay calibration the next step is to account for small impurities in the signal amplitude and phase response as a function of frequency, known as the bandpass calibration. For the bandpass calibration only the flux calibrator is utilized due to it's stable, well-characterized spectrum and it's bright nature, providing a high Signal-to-Noise Ratio (SNR) (Burke et al. (2019)). Following the VLA calibration manual[4] the phase structure over time for certain baselines is corrected for first. This step is generally useful for high frequency observations, again due to their sensitivity to atmospheric attenuation. In the case of significant phase-time variation for a small channel range this needs to be corrected for to prevent additional noise in the bandpass correction, degrading the quality of calibration (Thompson et al. (2001)). The solution of the initial phase-time calibration is found in figure 6a, where the channels are restricted to 2000~2200. Indeed 6a

---

[4] https://science.nrao.edu/facilities/vla/docs/manuals/obsguide/calibration

shows minimal phase variation over time and this was therefore not really necessary for this data. Figure 6c shows the bandpass solution of the signal amplitude over the different channels. Ideally one would see a stable amplitude of value one over all channels, meaning there are no bandpass based amplitude variations in the signal. In figure 6c the amplitudes are close to 1 between channels 800 and 3400, and quickly plummet outside of this range. During the application of this bandpass solution to the HI spectral line data, the amplitudes of the signal are divided by the bandpass solution, meaning the data would become very noisy outside of channels 800~3400. Moving on the data will therefore be restricted to channels 800~3400. Figure 6e shows the phase solution as a function of frequency. The phase solution of the signal has a clear curve, with an acceptable relatively small noise variation of 1-2 degrees. When applying the phase solution to the phase of the spectral line data, the phase gets multiplied by the phase solution to account for the frequency dependent phase variations.

2.2.5 Gain Calibration

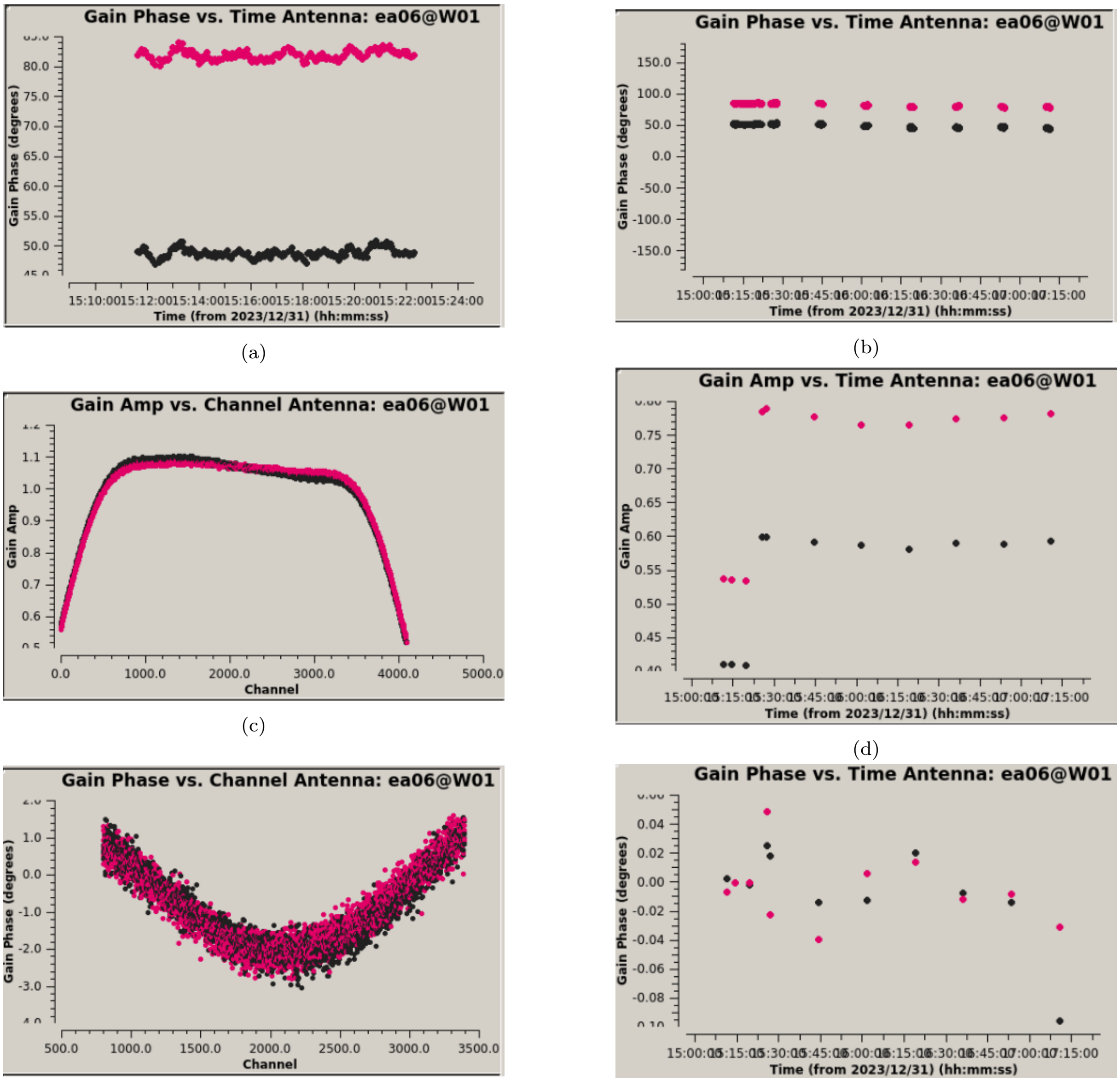
The final calibration step in the data correction process is to correct for amplitude and phase fluctuations as a function of time, known as complex gain calibration. During this step the gain calibrator is used as well next to the flux calibrator, as the gain calibrator is in close proximity to the source to ensure a similar path from source to telescope, such that applied corrections are representative of the complex gain variations affecting the source (Burke et al. (2019)). Figure 6b shows the initial phase variation with time for each individual integration. Generally doing the gain calibration for phase per integration is important for high frequency observations, due to their sensitivity for atmospheric attenuation, or for very long flux calibrator scans. This is for the same purpose as with the bandpass calibration, to prevent additional noise in the complex gain correction degrading the quality of calibration. Figure 6d shows the amplitude solution over time. The gain variation over time seems more drastic after calibration, but looking at the axes this variation is actually very minimal, and the different polarization states differ very little as well. The phase solution versus time (figure 6f) appears relatively stable over time with minor variations. The two polarizations exhibit similar stability patterns, implying that the calibration was consistent across both polarization states.

2.2.6 Flux Density Gain Calibrator

The flux density of the flux calibrator was set at the beginning of the calibration, and can now be used to derive the flux density of the other calibrators to confirm a successful calibration. For calibrators with unknown flux densities the flux density is assumed to have a value of 1.0 Jy. Due to this assumption, the gain solutions derived during the calibration will have a scaling error of the inverse of the square root of the true flux density of the source. The correction for this scaling error is done through bootstrapping[3]; an iterative method of refining an estimate by comparing it against a known standard. At the end of the correction process the mean gain amplitudes are consistent across all sources, regardless of whether the flux densities were initially known or not. The flux density for the gain calibrator was found to have a value of 2.094 ± 0.003, which perfectly compares to the expected flux of 2.1 Jy from the VLA calibrator list[5].

---

[5] https://science.nrao.edu/facilities/vla/observing/callist



(a)

(b)

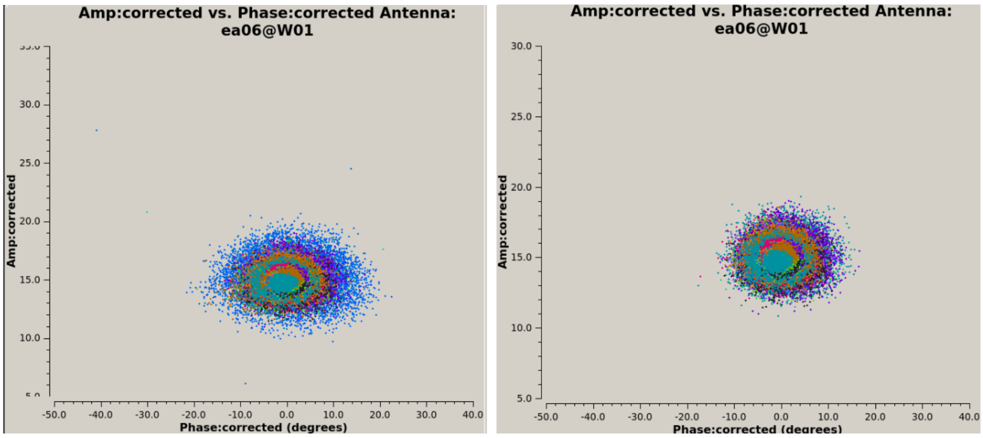
(c)

(d)

(e)

(f)

Figure 6: *All Subfigures:* This figure contains different steps of the bandpass and gain calibration for dataset eb44977689, showing antenna EA06 as an example. Color in figures represent the different degrees of polarization; RR and LL correlation. *Left Column:* Bandpass calibration. *Subfigure a:* Plot of initial phase solution for channels 2000~2200, shows minimal phase variation with time. *Subfigure c:* Plot of amplitude solution over full range of channels for bandpass calibrator. The signal has high throughput in channel range 800~3400, which is the range used for further data calibration and processing steps. *Subfigure e:* Plot of phase solution over channels 800~3400 for bandpass calibrator. The phase error is relatively small, 1~2 degrees at most. *Right Column:* Gain calibration. *Subfigure b:* Initial phase calibration for each integration, contains little variation. *Subfigure d:* Amplitude solution over time for gain calibration. The jump in amplitude is caused by the switch between flux and gain calibrators. *Subfigure f:* Phase solution of complex gain calibration. Very minimal phase variation, ~ 0.1 degrees for these antenna's, rest doesn't have much more.

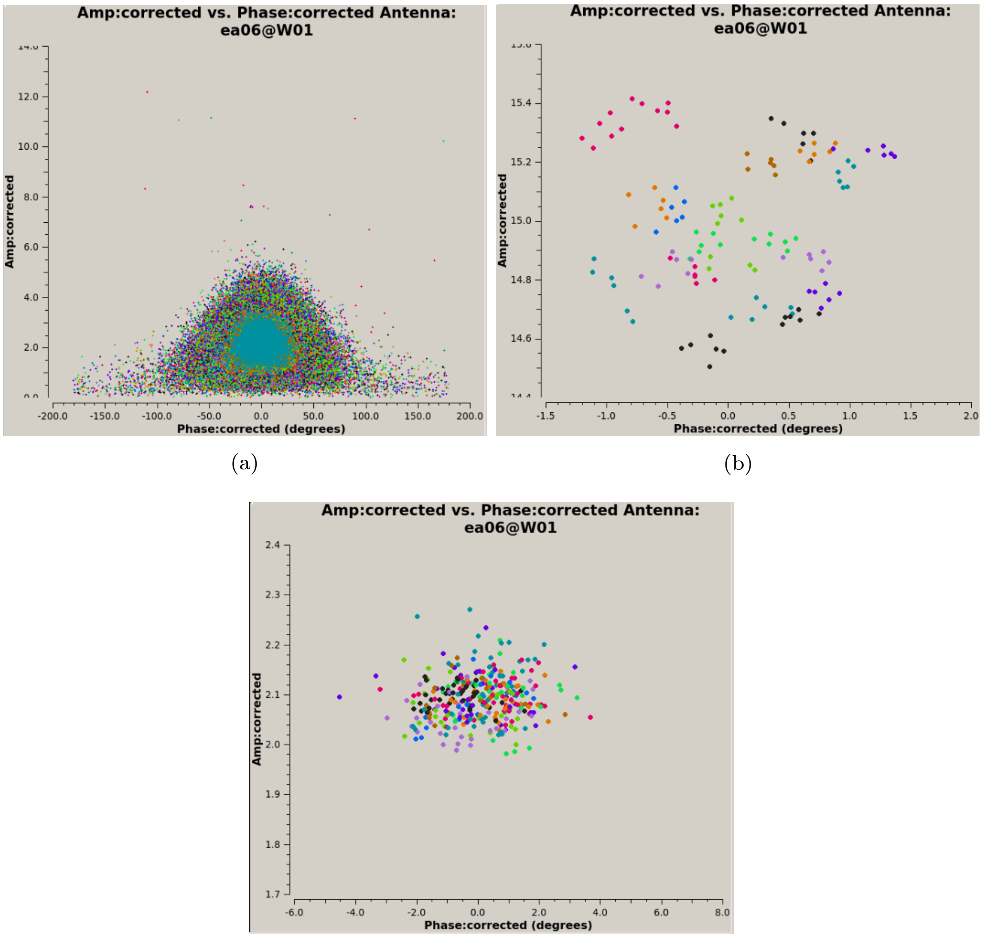


(a) (b)

Figure 7: Comparison of calibration solutions for the flux calibrator, displaying corrected amplitudes over phase using a time average of 5000 seconds. *Subfigure a* shows a lot of spread out datapoints surrounding the inner oval structure, displayed here as blue datapoints. Using CASA imaging tools, it becomes apparent the outlying datapoints correspond to short baselines. *Subfigure b* displays the same as subfigure a, with an extra restriction on short baselines by specifying a parameter uvlim='>70m'.

2.2.7 Calibration Inspection

To finish the calibration process, the previously made calibration solutions are applied to the measurement set to inspect the success of the calibration. To do a thorough inspection one could first plot the corrected amplitudes over time for both calibrators to check if the amplitudes are centered around the known flux values of the calibrators (15.0 Jy and 2.0 Jy respectively for flux and gain calibrators), and to check if the left over amplitude variations are within acceptable ranges. Furthermore the corrected phase over time for both calibrators can be analyzed to ensure phase variations over time are minimal and are centered around zero. We expect the phase to be centered at zero as this indicates that any inherent instrumental or atmospheric delays have been corrected for, and the signals are aligned in time. Another check is to plot corrected amplitude over channel for both calibrators to assess how well the bandpass calibration has been performed, ensuring the sensitivity across frequencies is uniform. Additionally the corrected phase over channel can be analyzed to check how the frequency-dependent phase delays have been accounted for. Here we will just show a quick inspection of the corrected data, where we only look at the corrected amplitudes versus corrected phases. The results of the inspection are shown in figures 7 and 8, for dataset eb44977689. Here we again try to look for very strong outliers or unexpected structures within the plots, as well as checking if the phase is centered at zero and the amplitudes are centered around the known flux values of the calibrators. Subfigure 7a shows the corrected amplitude over phase for the flux calibrator, where the blue datapoints appear quite spread out compared to the other datapoints. Subfigure 7b shows the same as subfigure 7a, with the difference of restricted short baselines (<70 meters) to visualize the importance of restricting short baselines in certain scenarios. Short baselines are more susceptible to extended source structure and atmospheric phase variations, introducing a spread in the amplitude-phase plots. Subfigure 8a shows the corrected amplitudes over phase of the gain calibrator, while again just averaging over the time. Subfigure 8b shows the corrected amplitude over gain with both the time and channels averaged together for the flux calibrator, subfigure 8c shows the same for the gain calibrator.



(a)

(b)

(c)

Figure 8: Inspection of calibration results for the flux and gain calibrators. *Subfigure a* shows the corrected amplitudes over phase of the gain calibrator, while again just averaging over the time. *Subfigure b* shows the corrected amplitude over gain with both the time and channels averaged together for the flux calibrator, *Subfigure c* shows the same for the gain calibrator.

3 Imaging Procedures

3.1 Continuum Substraction

After finishing up the calibration process, before commencing the imaging, a continuum subtraction is required when there is significant continuum emission present in what was meant as spectral line data. The source AGC 268071 is a relatively faint source, with it's peak in the line spectrum close to that of the Galactic hydrogen. Continuum emission is therefore of significance, and is filtered out of the data in fourier space by fitting a polynomial to the real and complex visibility data as a function of frequecy. The resulting fit represents a model of the continuum emission in all channels (except the main source emission channels), which is substracted from all selected channels. The substraction is done in fourier space to ensure a more accurate and robust deconvolution of the line emission[6]. Channels containing the main source emission are excluded from this process because they exhibit significantly higher emission levels compared to background emission alone. Including these channels would lead to an excessive removal of supposed background emission, resulting in an inaccurate subtraction. The continuum subtraction of AGC 268071 was done using a default polynomial of order zero, meaning the fit was linear.

3.2 Imaging Theory

The information in this section is largely taken from Thompson et al. (2001), Wilson et al. (2013) or Burke et al. (2019), unless otherwise specified.

The imaging process entails the transition from visibility measurements in fourier space to the intensity values in the standard cartesian plane. The relation between the visibilities and the specific brightness distribution is given by the Cittert-Zernike theorem. According to Schmidt et al. (2022), this theorem states that the mutual coherence function of a wave field in the far field is the Fourier Transform (FT) of the intensity distribution of the source. However, because we do not have access to an infinite number of antenna's to cover the main beam, the intensity values appear convoluted when performing just an inverse Fourier transform on the visibilities. To address this, the visibilities are first gridded onto a regular spatial grid, with each visibility weighted according to its signal-to-noise ratio and u-v coverage. This gridding and weighting process ensures that the Fourier transform accurately represents the interference pattern formed by combining data from multiple antennas. The result, known as a dirty image, appears convoluted due to incomplete spatial coverage. The dirty image can be restored to a clean image using sophisticated algorithms like CLEAN[3].

3.2.1 Weighting Schemes

There are different types of weights that can be selected to assign to the pixel values, with most used ones being 'natural', 'briggs' and 'uniform'. Generally all weighting options modify the side lobes of the dirty beam response, with varying details. All types of weighting have different pros and cons, for example the natural weighting gives equal weight to all samples, causing very large side lobes but a very low noise level with a poor resolution. The uniform weighting gives a weight inversely proportional to the sampling density function, which reduces the side lobes at expense of a higher noise ratio. The Briggs weighting is a compromise between the former two styles, giving a good trade-off between resolution and sensitivity[7]. The latter style is therefore

---

[6] https://casadocs.readthedocs.io/en/stable/index.html

[7] https://casaguides.nrao.edu/index.php?title=KarlG.JanskyVLATutorials

the one used for the imaging of AGC268071, with a robust value of 0.2.

3.2.2 CLEAN Algorithm

When creating the first dirty image of the data as mentioned above, the output is a convoluted image of the source. Such a dirty image is visible in figure 10a, which is the dirty image of joined data of the four observations of AGC 268071, smoothed in the image plane from the original 45" to 105" resolution. The dirty image is mathematically given by the product of the Fast Fourier Transform (FFT) of the true visibility function and the FFT of a function representing which values of the true visibility function are actually measured. The latter function is known as the sampling function (figure 9), or as the Point Spread Function (PSF) after performing the FFT on the sampling function (figure 10b). Having obtained the dirty image, the CLEAN algorithm tries to restore the true visibility function by attempting to remove the dirty point spread function from the dirty image using deconvolution.

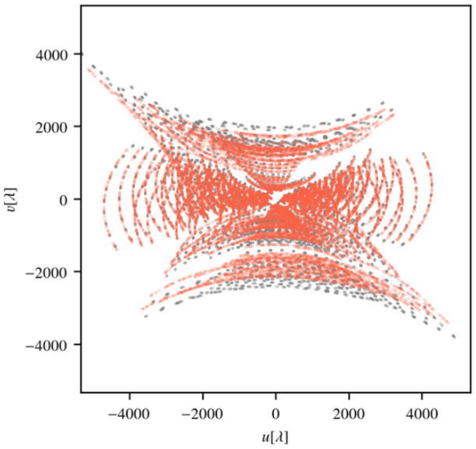
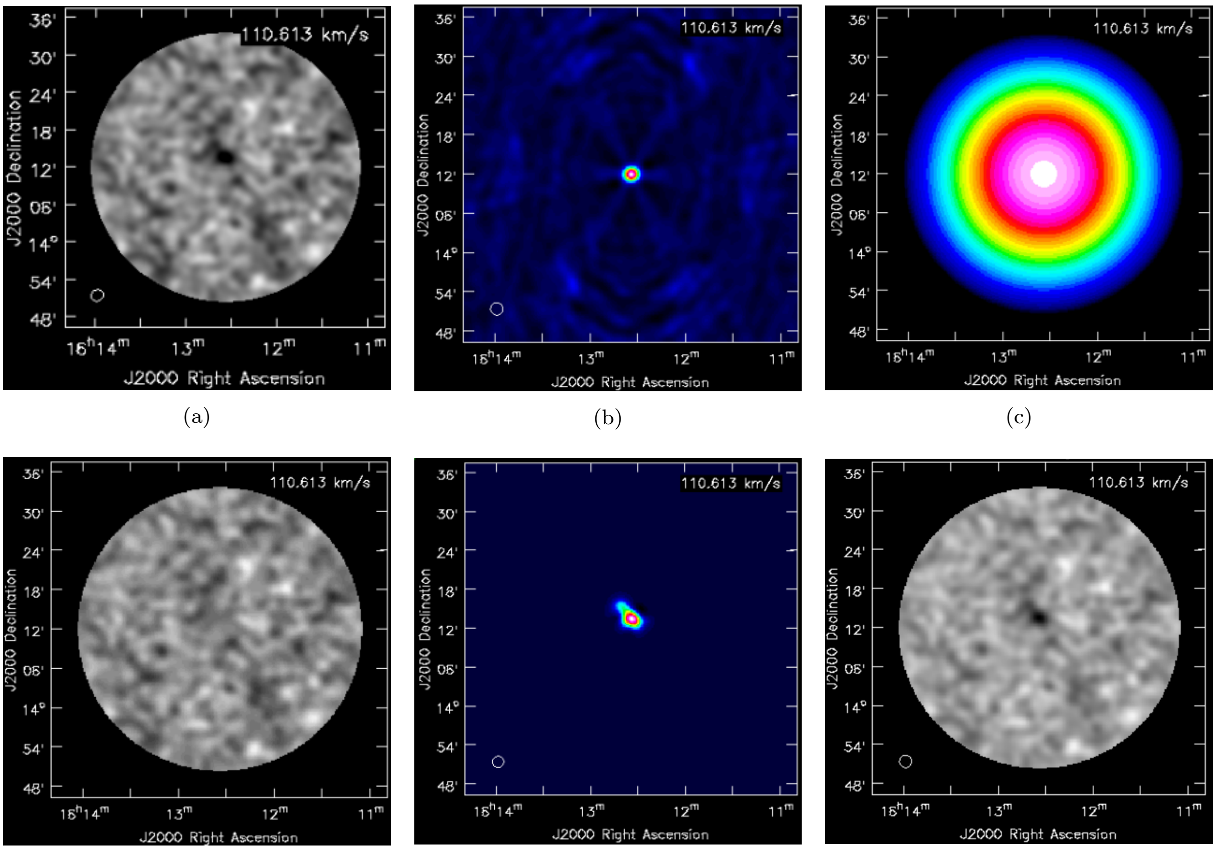


Figure 9: Sampling function of the VLA showing a random sample (30000 points) of uv-coverage of an observation with VLA from Arras et al. (2019). The gray and red points show the uv coverage of the calibration source and the science target used in that study.

Deconvolution works the following way; CLEAN finds the intensity and location of the brightest point in the dirty image and substracts from that point the corresponding dirty PSF scaled by the peak intensity and gain factor. The substraction of the PSF effectively removes the contribution of that point source and its associated sidelobes from the dirty image. This process is repeated until the algorithm runs through the number of set iterations or until the residual image reaches a set maximum value threshold. The residual image contains the leftover noise of the dirty image after removing most of the brightest sources (figure 10d). The model output contains the points that are substracted from the dirty image (figure 10e). The primary beam output represents the combined response pattern of the antennas in the array, illustrating how the strength of received signals varies across the field of view (figure 10c). This output is crucial for correcting sensitivity variations introduced by the primary beam. Without this correction, the flux densities of sources located on the outskirts of the image would be underestimated due to the reduced sensitivity in those regions. The final data product given as output to CLEAN is the clean image. The clean image consists of the reconstructed sky brightness distribution after combining the model components with the residual image.

3.3 Building a Mask

To minimize noise and limit the region over which you clean, and to enable deep cleaning without worrying about impacting noise characteristics by cleaning noise in the final cleaned image, one can try building a strong mask. The mask should contain only the main emission of the source through an iterative masking and cleaning process. The first mask has to be very strict to exclude the general noise and sidelobes, but after applying the first mask and smoothing the



(a) (b) (c)

(d) (e) (f)

Figure 10: All subfigures show the same channel, the main channel where the source resides and has the highest brightness. All images are smoothed to a resolution of 105 arcseconds to enhance certain features and reduce the overall noise. *Subfigure a:* First dirty image of AGC 268071. *Subfigure b:* Point Spread Function of the final cleaned image; FFT of final sampling function. *Subfigure c:* Primary Beam response of the final cleaned image. Blue represents low sensitivity, white represents highest sensitivity response. *Subfigure d:* Residual of the final cleaned image. Residual contains left over noise that was excluded from CLEAN algorithm due to the set threshold. *Subfigure e:* Model image of final cleaned image. Model contains everything that is not visible in residual image; the points that are substracted from dirty image. *Subfigure f:* Final cleaned image of AGC 268071. Comparing this image with subfigure a, the side lobe structure has diminished substantially, as well as the random noise structure.

partially-cleaned image to lower resolutions to reduce the noise, the mask can be expanded to contain lower intensity emission from the source. This process can be repeated until a desired accuracy and cleanliness of the source are reached.
For the data of AGC 268071, two masks were created, an initial mask and a deep cleaning mask, and deemed sufficient enough in successfully and cleanly deconvolving the source. The creation of the masks was done using the following steps:

- The dirty 45” image was first smoothed to a resolution of 150” in the image plane, as we want to make the mask based on an image with a lower resolution than the final cleaned product we want to work with. The smoothing suppresses high-frequency noise components, which results in a cleaner, more uniform background that makes it easier to distinguish actual features from noise.
- After having obtained the 150” image, a suitable mask and threshold for cleaning needed to be determined. The first clip value was taken as the minimum value of the 150” image at which no surrounding noise was included into the mask, which corresponded to 6.2 mJy.
- The mean noise of the channels without source emission in the 150” image was then determined to find a suitable threshold. The suitable threshold was taken to be twice the rms; the standard deviation of the noise, as for the initial mask only the most prominent features should be cleaned to allow for a deep second cleaning iteration.
- After applying the mask and threshold parameters in the CLEAN algorithm the output is the once masked and cleaned datacube at 45” resolution. The result of the first cleaning iteration is then smoothed again to 150” resolution in the image plane.
- The noise level of the once cleaned 150” resolution image is then determined to find a suitable clean mask. The clean mask used a high clip of ~3 rms of the 150” smoothed data (once masked), corresponding to 4.6 mJy, to include a significant amount of low source emission. Because the mask is set much lower now, to ensure as little noise as possible was going into the CLEAN command, a box restricting the pixel values of the mask to just the source was drawn around the source for the final deep clean.
- The threshold of the final cleaning was determined to be a strict ~0.5 rms of the original 45” resolution dirty image, as this is a deep clean.
- Applying the clean mask and threshold parameters in the CLEAN algorithm now results in a clean datacube as output, shown in figure 10f smoothed to 105 arcseconds resolution in the image plane.

An overview of cleaned datacube parameters are given in table 3. As visible in the table the beam of the 105" resolution data has major and minor axes values considerably larger than 105", due to smoothing with a gaussian kernel. Throughout the masking and cleaning process the original 45” datacube was consistently smoothed to 150" resolution in the image plane. This was originally done because the intent was to work with 105” resolution clean dataproducts. However as visible in the Data Products section, the cleaned data products at 45" resolution already clearly depicted the source. A higher resolution allows for better resolved details of the source, which is preferred over a lower resolution. This does however imply that the masking and cleaning process might have not been deep enough for cleaned data products of 45” resolution.

Angular Res. (arcseconds)	Beam axes (arcseconds)	Beam Position Angle (radians)	Noise (mJy)	Spectral Res. (km/s)	Robust
45”	45.83” x 41.5”	-0.136	1.22	4.8	0.2
105”	114.56” x 112.91”	-7.810	4.63	4.8	0.2

Table 3: Table providing an overview of parameters of the clean images. Table contains spatial resolution in arcseconds, beam size in arcseconds, position angle in degrees, resoluting noise in cleaned cubes in mJy, the spectral resolution used to image the clean data and robust value used to image clean data. Again here the 105” image is made by smoothing the clean 45” image in the image plane.

4 HI Content of AGC 268071

This section aims to produce higher order data products from the cleaned datacube, in order to make well-based arguments in regard to the nature of AGC 268071. For the creation of moment maps the cleaned datacube is restricted to a limited amount of channels surrounding the main source channel, depicted in figure 15 by the orange vertical lines. The channel range is restricted in an attempt to limit the high noise peaks in the moment maps.

4.1 Moment Zero Map

Moment maps are useful analysis tools to evaluate the calibrated and cleaned data cube. Mathematically, a zero-th moment in integral form is represented by the integrated intensity of the spectral line over the total velocity range of the data. The zero-th moment map thus provides the total HI intensity over the primary beam, given by the following equation:[8]

$$M_0 = \int I_v dv \tag{1}$$

The calibrated and cleaned datacube has units of Jansky $Beam^{-1}$ Km s^{-1}, which are not very intuitive when analyzing the zero moment map. Instead the units can be converted to column density ($\frac{N_{HI}}{cm^{-2}}$) using the following set of equations:[9]

$$\frac{N_{HI}}{cm^{-2}} \approx 1.823 \cdot 10^{18} \int \frac{T_B}{K} \frac{dv}{kms^{-1}} \tag{2a}$$

$$\frac{T_B}{K} \approx 605.7 \frac{S_\nu}{mJy} (\frac{\vartheta}{arcsec})^{-2} \tag{2b}$$

Here S_ν is the specific intensity per pixel in mJy/beam km/s and ϑ is the product of the major and minor beamwidth values in arcseconds. Figure 11a shows the moment zero map using the mask made in the final cleaning and imaging stage (section 3), with overlaying column density contours of [1, 1.4, 2, 3] x 10^{19} N_{HI}/cm^{-2}. The peak column density can be extracted from the image, and is found to be 3.17 x 10^{19} N_{HI}/cm^{-2}. Figure 11b shows the moment zero map of the cleaned data smoothed in image plane to 105" resolution with overlaying column density contours of [2.4, 4.0, 6.0, 9.0] x 10^{18} N_{HI}/cm^{-2}. The peak column density for 105" is found to be 9.34 x 10^{18} N_{HI}/cm^{-2}.

The intensity values at the 105" smoothed image are almost twice as large as the intensity values of the 45" image. This behaviour is expected from lower resolution images due to the increased sensitivity at the expense of resolution.

At both resolutions there appears to be a relatively dense HI core in the middle of the figure, with large surrounding noisy structure. I do not think the surrounding structure with an intensity of ~0.01 Jy at 45" resolution is part of the source structure, but instead is a result from insufficient masking and cleaning as described at the end of the imaging section (section 3.3). At 105" resolution, figure 11b shows a relatively bright side lobe to the source between 10' and 12' declination. I also believe this is not part of the source structure, based on a comparison with the 45" moment zero map. From figure 11a is it clear there are some bright noise peaks between declination 10' and 12', which gets smoothed out over a larger area in the 105" resolution image. Based on figure 11 it is apparent that this data requires further masking to isolate purely AGC 268071, which will be done using section 4.1.2.

---

[8] https://spectral-cube.readthedocs.io/en/latest/moments.html

[9] https://www.atnf.csiro.au/people/Tobias.Westmeier/toolshihelpers.php

4.1.1 Primary Beam Correction

The primary beam of an image, visible in figure 10c, represents the varying sensitivity across the beam of a telescope. The beam sensitivity is highest in the center of the beam and lowest towards to outskirts. This sensitivity needs to be accounted for when analyzing the data, especially when a source is not perfectly centered or when it is very large. A primary beam correction can be applied throughout the imaging and cleaning procedures, specified as a parameter in the CLEAN command. However, applying the primary beam correction parameter to the CLEAN command changes the noise profile of the data cube, as the noise would be elevated further from the center of the primary beam. To prevent such heightened noise the primary beam correction can be done after the masking and clipping iterations, when there will be little high noise peaks left at the outskirts of the primary beam[7]. For the case of this study the source was located fairly in the center of the beam, with no large extended source structure. The primary beam correction therefore changed visibly almost nothing compared to the non-primary beam corrected moment zero map, which is why an extra primary beam corrected moment zero map is left out here. Moving forward the first and second moment maps are made using a primary beam corrected data cube.

4.1.2 HI Significance Map

As mentioned shortly before, further masking is necessary to isolate purely AGC 268071. This can be done using a significance map of the moment zero map. A significance map serves as a tool to conveniently analyze the SNR of the HI emission in order to decide on an appropriate clipping value. When collapsing the spectral data axis of the data cube to make the moment zero map, the noise levels of the initial channels are added together using a relation from Verheijen and Sancisi (2000):

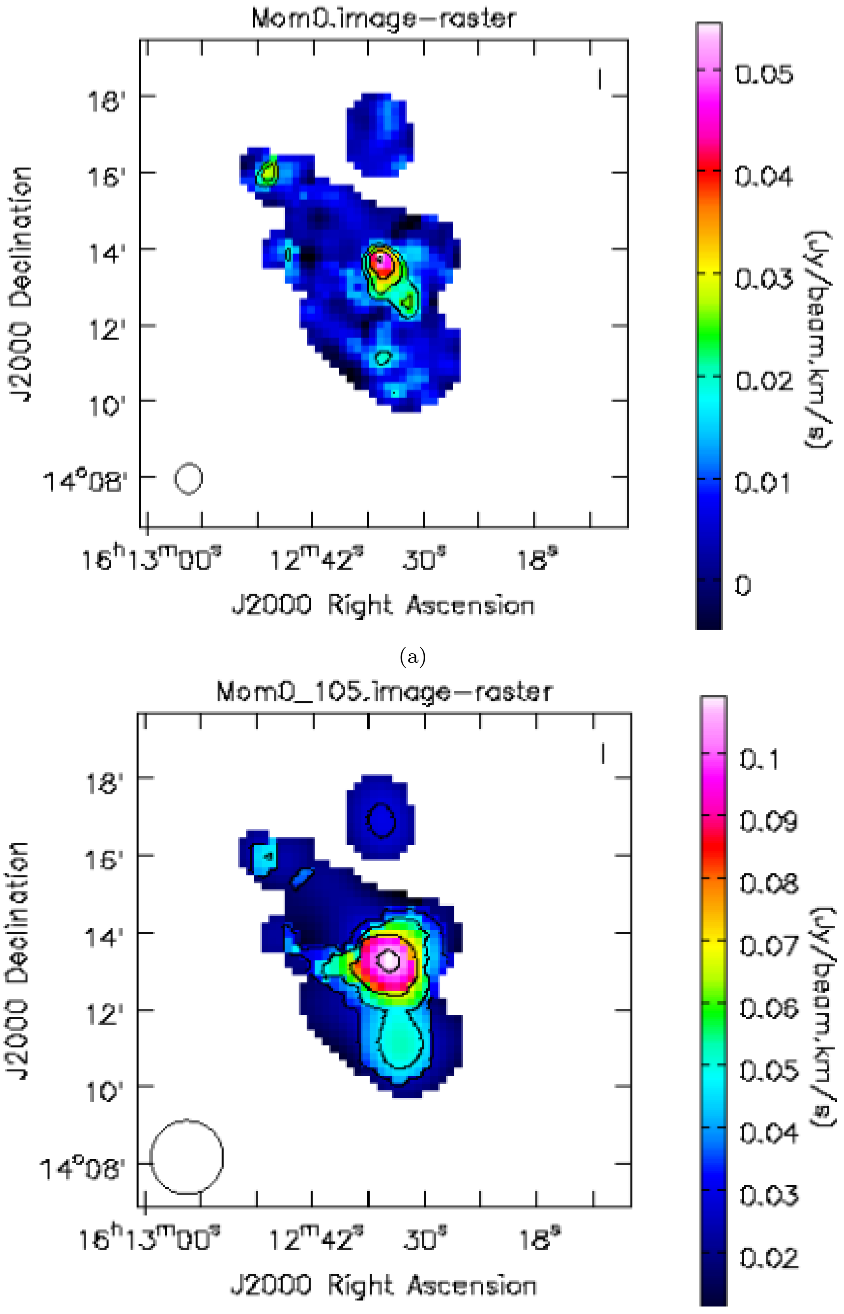
$$\sigma_N = \sigma\sqrt{N} \tag{3}$$

Where σ is the mean rms of the individual channels in the cleaned datacube used to create the moment map, N is the number of channels selected to make the moment map and σ_N is the resulting noise level of the moment map. When creating the moment zero map the clean mask is included, resulting in different number of channels contributing to different pixels. So if we take this varying channel number into account and divide the intensity per pixel by the propagated error per pixel, the result is a map showing the signal-to-noise ratio of each pixel. The significance map is mathematically given by the following equation from Hess et al. (2018):

$$\Sigma(i,j) = \frac{I(i,j)}{\sigma\sqrt{N_{chan}(i,j)}} \tag{4}$$

Where Σ(i,j) is the signal to noise ratio per pixel, I(i,j) is the intensity per pixel, σ is the mean noise per channel of the clean datacube and N_{chan}(i,j) is the number of channels used per pixel.

Figures 12a and 12b show the significance maps of AGC 268071, based on the 45" and 105" moment zero maps respectively. The 45" significance map shows a lot of noise with relatively high significance surrounding the source. The structure on the left bottom of the source largely disappears from an SNR of 3 on, and completely disappears at an SNR of 5. A SNR of 4 could be sufficiently large to be considered part the source structure, even more when comparing the 'side' structure to the 105" moment zero significance map, as there it does appear as part of the source for an SNR of 3.6. However to restrict the surrounding noise as much as possible a higher clip must be applied for the further moment maps, which will remove part of the lower

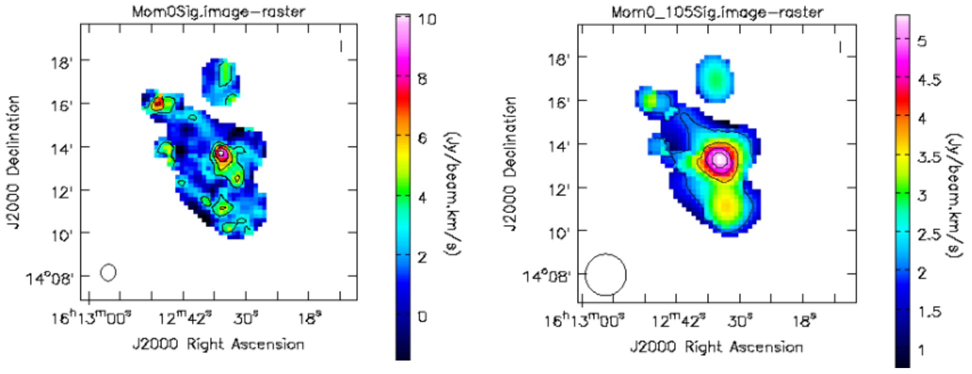


(a)

(b)

Figure 11: Moment zero maps using the final clean mask from section 3, with overlying column density contours. Axes units are given by Declination and Right Ascension in J2000. *Subfigure a:* Initial moment zero map in units Jy/Beam km/s. Column density contours made at [1, 1.4, 2, 3] x 10^{19} N_{HI}/cm^{-2}. Peak column density extracted from image is found to be 3.17 x 10^{19} N_{HI}/cm^{-2}. *Subfigure b:* Moment zero map of cleaned data smoothed in image plane to 105" resolution with overlaying column density contours of [2.4, 4.0, 6.0, 9.0] x 10^{18} N_{HI}/cm^{-2}. Peak column density extracted from image is found to be 9.34 x 10^{18} N_{HI}/cm^{-2}.

significance source emission. For the 45" image the a clip of significance 4.5 is an appropriate weigh-off between including enough source structure and excluding enough noise structure. For the 105" smoothed data a clip of 3.6 encapsulates most of the source emission while excluding the large structure below the source between 10' and 12' declination, which is considered to be noise.



(a) (b)

Figure 12: Both subfigures contain the significance map of the moment zero maps. *Left figure:* Significance map of 45" moment zero map. Contours made at [3.0, 5.0, 7.0, 9.0] x SNR *Right figure:* Significance map of 105" moment zero map, with SNR contours at [1.0, 2.0, 3.6, 4.5, 5] x SNR.

4.2 First Moment Map

The first moment represents the intensity-weighted velocity of the 21 cm line, also known as the recessional velocity, given in velocity units (km s^{-1}). Mathematically the first moment map is expressed the following way:[8]

$$M_1 = \frac{\int \nu I_\nu d\nu}{M_0} \tag{5}$$

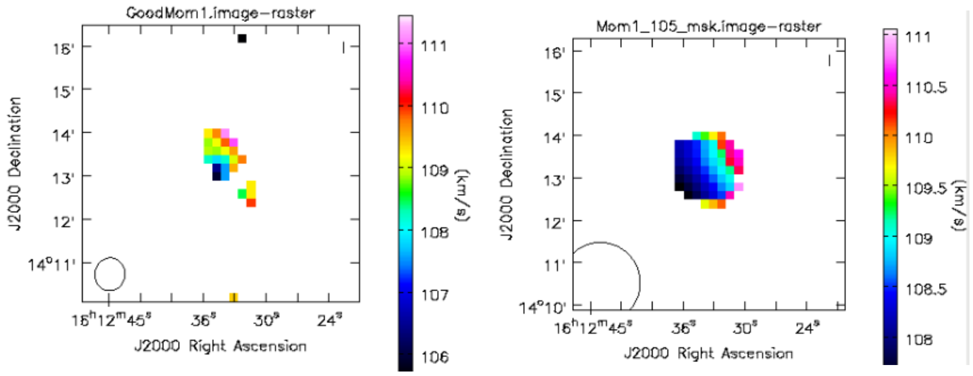
Having chosen appropriate moment zero clips, the resulting moment one maps with the applied clips are depicted in figure 13. At both resolutions the source appears to have a distinct velocity gradient accross it's surface. At 45" resolution the velocity gradient goes from ~111 km s^{-1} on the right side to ~106 km s^{-1} on the left side of the source, and at 105" resolution the gradient spans from ~111 to ~108 km s^{-1} in the same direction.

4.3 Second Moment Map

The second moment map represents the velocity dispersion of the data, referring to the range of velocities of the gas. The second moment is mathematically given by the following equation:[8]

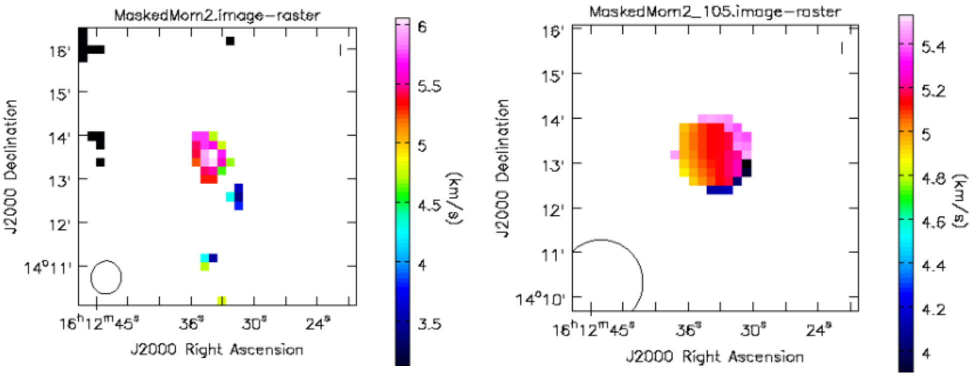
$$M_2 = \frac{\int I_\nu (\nu - M_1)^2 d\nu}{M_0} \tag{6}$$

The second moment map of the data is made using the 45" clean image with the same 45" clean mask, as well as the extra moment zero SNR mask created in the section above. Another second moment map is made at a lower resolution using the 45" clean image smoothed to 105" resolution in the image plane and the 45" clean mask, as well as the extra 105" moment zero SNR mask. Figure 14 depicts a maximum velocity dispersion of around 6 and 5.4 km s^{-1} respectively for 45" and 105" resolution moment two maps.



(a) (b)

Figure 13: Both subfigures show the moment one map made from the cleaned datacube with an extra mask based on the moment zero HI map. The 45" data is clipped to a significance of 4.5, and the 105" data is clipped to a significance of 3.6. Both figures show a velocity gradient across the source from around 111 km s^{-1} to 106-108 km s^{-1}. *Left figure:* Moment one map made with original 45" resolution data. Moment zero mask applied here is based on same 45" data. *Right figure:* Moment one map made with original 45" data smoothed to 105" resolution, using a moment zero HI mask made with a 105" moment zero significance map.



(a) (b)

Figure 14: Both figures show the second moment map made with the clean data and mask, as well as the extra moment zero HI significance mask. The second moment map shows the velocity dispersion of the gas. The 45" data is clipped to a significance of 4.5, and the 105" data is clipped to a significance of 3.6. *Left figure:* Second moment map at 45" resolution. Velocity dispersion of source has a maximum of 6 km s^{-1}. *Right figure:* Second moment map at 105" resolution, made from 105" smoothed clean image and 45" clean mask, as well as the extra 105" moment zero mask. Maximum velocity dispersion found is around 5.4 km s^{-1}.

4.4 Flux Products and Line Spectra

The total integrated line flux of the source can be found by integrating the sum of the fluxes per channel over the channel range. CASA automatically accounts for the beam of the observation and gives the flux in units Jy km s^{-1}. The total fluxes of the primary beam corrected moment zero maps are written in table 4. The difference between the computed fluxes when applying the significance mask is due to the lower amount of emission included in the computation. The flux derived by ALFALFA is 2.67 ± 0.08 Jy km s^{-1}, the possible reasons for this discrepancy are discussed in the next paragraph.

A final data product that can be extracted from the cleaned datacube is the line spectrum of AGC 268071, plotten in figure 15 versus velocity (figure made in Jupyterhub). Together with

	45” (Jy km s^{-1})	105” (Jy km s^{-1})
Clean Mask	0.33	0.23
Significance Mask	0.08	0.05

Table 4: Integrated line flux in Jy km s^{-1} for 45” and 105” resolution data, computed for both the clean mask and the significance mask.

the total flux per channel for the original 45” cleaned datacube, the total flux per channel for the smoothed 105” resolution data and the ALFALFA line spectrum (figure 3) at 210” resolution are also depicted in figure 15. Looking at the HI line spectra, the spectrum of ALFALFA appears to be shifted up from the computed VLA spectra with about 15 mJy. This can be explained when comparing figures 15 and 3, where it becomes clear the ALFALFA survey did not seperate the galactic HI emission from the source emission, which is something the VLA does account for. The galactic HI emission appears to peak around 15 mJy, exactly the difference in peak flux from the 105” data. The comparatively small integrated line flux found in this study can then be explained by the same phenomena, as well as due to the small width of the VLA spectral line. The small width of the VLA spectra compared to the ALFALFA spectra can be attributed to galactic broadening in the ALFALFA spectra.

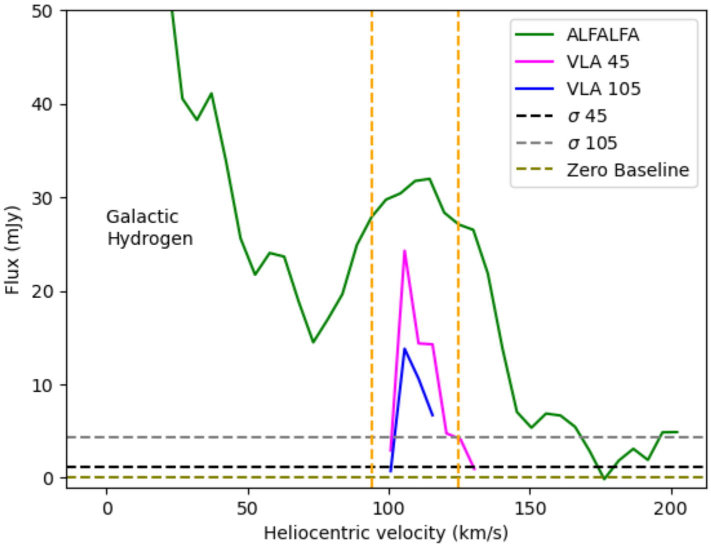


Figure 15: Figure shows line spectra’s of cleaned VLA datacube at 45” resolution and the same data smoothed to 105” resolution by the magenta and blue lines. The ALFALFA line spectrum at ~3 arcminutes resolution is given by the green line. The channel range used for moment map creation is given by the orange vertical lines, and the mean noise levels in the individual channels for the VLA data are given by the grey and black horizontal lines. Finally the zero baseline of the data is given by the olive horizontal line.

5 Nature of AGC 268071

In this section the nature of AGC 268071 is discussed for the two possible scenario's; a gas bearing dark matter halo or a cloud of gas located around the Milky Way. The discussion will explore the two scenarios by looking at different lines of evidence collected in the previous chapter.

5.1 Neutral Hydrogen Distribution

As mentioned in the introduction of section 4, the neutral hydrogen morphology of a dark matter halo is expected to appear smooth and decently homogeneous (Adams et al. (2014)), as opposed to the broken-up HI distribution of a high velocity cloud (Van Woerden et al. (1999)). When analyzing the moment zero maps (figure 11), at first glance the neutral hydrogen doesn't seem to have a very smooth distribution, especially the 45" resolution moment zero map. This can be attributed to insignificant emission being included in the masking and cleaning process, resulting in too much noise being left over surrounding the source. This is visible as the blue clumpy structure in figure 11a with a flux value around 0.01 Jy/beam km/s. When ignoring the low SNR surrounding structure the resulting source is relatively small, but does have a smooth HI morphology. The moment zero map at a lower resolution of 105" shows a similar smooth source morphology as the 45" variant, without the clumpy surrounding structure. It is important to mention the beam size here however, as the beam in figure 11 at both resolutions roughly spans the size of the source. Having a beam the same size as the source implies the source may not be fully resolved, meaning the source will lack detailed structural information. A non-resolved source can have structure at the outskirts of the source that does not appear in it's image, as well as structure within the source that gets blurred out (Wilson et al. (2013)).

Besides looking at the distribution of the neutral hydrogen, the zero moment maps can be used to analyze the peak column density of the source. Adams et al. (2016) states that dark galaxies are expected to have relatively high column density values compared to compact high velocity clouds, allowing them to be imaged at high resolutions while maintaining a smooth morphology. The peak column density at 45" resolution at the center of the source is 3.17 x 10^{19} N_{HI}/cm^{-2} and at 105" resolution is 9.34 x 10^{18} N_{HI}/cm^{-2}. As expected the column density is lower at a lower angular resolution because the signal from the dense core is spread out over a larger beam area. We can compare the determined peak column density of AGC 268071 to the column densities of twelve dark galaxy candidates researched by Adams et al. (2016) at 105" resolution, where the sources labelled as excellent galactic candidates have peak column densities of 2.9-4.9 x 10^{19} N_{HI}/cm^{-2}.

Therefore, the fact that the source is clearly detectable at a resolution of 45", with a smooth neutral hydrogen distribution and a peak column density comparable to other excellent gas-bearing dark matter mini halo candidates suggests towards this classification.

5.2 Velocity Distribution

The first and second moment maps created in section 4 are useful to analyze the velocity behaviour of the gas in AGC 268071. From figure 13 it is apparent the source has a distinct velocity gradient across its surface, measured at both 45" and 105" resolution. According to Adams et al. (2014) a distinct velocity gradient across the surface of an UCHVC is indicative of rotation, supporting the galaxy hypothesis. The velocity gradient at 45" resolution spans ~5 km s^{-1}, whereas the gradient at 105" resolution spans a smaller ~3 km s^{-1}. These velocity gradients are relatively small compared to some of the other detected velocity gradients for galaxy candidates (Adams et al. (2014), Adams et al. (2016)). The observed difference in the velocity gradient between the two resolutions is primarily due to the smoothing effects. The observed velocities are averaged over a larger area, meaning the peak velocities get convolved with surrounding velocities in the larger beam. This averaging process effectively blurs out the more extreme velocities. Apart from being unresolved spatially at both resolutions, the spectral resolution at which the clean datacubes were made as stated in table 3 is 4.8 km s^{-1}, which is more than the velocity gradient found at 105" resolution.

The velocity dispersion of our source is depicted in figure 14. At 45" resolution the maximum dispersion is ~6 km s^{-1}, whereas at 105" resolution the maximum velocity dispersion is ~5.4 km s^{-1}. The velocity dispersion in a source is caused by thermal broadening convolved with the bulk motion of the gas within the beam. Thermal broadening occurs due to the random thermal motions of the gas particles, leading to a spread in their velocities. Sources consisting of mainly Cold Neutral Medium have a HI velocity dispersion of less than 6 km s^{-1}, as cold HI has a typical velocity dispersion of ~4.5 km s^{-1} (Warren et al. (2012)). Sources consisting of mainly Warm Neutral Medium have a velocity dispersion larger than 6 km s^{-1}, with a typical HI velocity dispersion of ~10 km s^{-1} (Warren et al. (2012)). The gas particles are more energetic in WNM, but also experience turbulence, increasing the velocity dispersion. Additionally, due to the inherent rotational characteristics of the gas in a galaxy, the doppler broadening of the gas contributes to an increased velocity dispersion. Generally HVCs are composed of WNM with pockets of CNM (De Heij et al. (2002)), expected to have a lower velocity dispersion when of gas cloud in Milky Way origin than of galactic origin. The gas in a dark matter mini halo is mainly expected to be Warm Neutral Medium gas (Faerman et al. (2013)), which is partly why the source is likely inefficient at forming stars. Therefore it appears that the velocity dispersion measured for AGC 268071 corresponds closer to that expected of a cloud of gas in the Galactic halo rather than a galaxy. However, it is important to mention that velocity dispersion tends to be underestimated at a low SNR, as the Gaussian shape of the emission includes relatively large wings of low emission that are excluded from the mask. As a result, the width of the resulting masked emission is artificially decreased (Blok et al. (2024)). Comparing the velocity dispersion of AGC 268071 to that of other dwarf galaxies in the local group, the lowest measured HI velocity dispersion in McConnachie (2012) is 4.5 ±1 km s^{-1}, belonging to dwarf galaxy Wolf–Lundmark–Melotte (WLM).

The measured velocity gradient at 45" data suggest toward a gas bearing mini halo classification. The velocity dispersion of AGC268071 at both resolutions is on the low end of WNM, and implies the source contains a lot of CNM. This does not directly point away from a galactic classification, as dwarf galaxies with a large fraction of CNM have been observed in the Local Group, but also cannot be used as a strong argument towards a galactic classification.

5.3 AGC 268071 as a galaxy

The derived integrated line flux of ACG 268071 found in table 4 can be used to compute the neutral hydrogen mass of the source. The equation to compute the mass is taken from Jones et al. (2018):

$$M_{HI}(M_{\odot}) = 2.365 \cdot 10^{5} \cdot D(Mpc)^{2} \cdot S_{21} \tag{7}$$

Where the distance to the source needs to be estimated. In Rhode et al. (2023) the estimated distance for the stellar counterpart is 570 kpc, with a possible range between 490 and 590 kpc. Assuming AGC 268071 is of galactic nature, the range of HI masses it takes on at a resolution of 45" is 1.87-2.72 x $10^4 M_\odot$. At a resolution of 105" the range of HI masses the source takes on is 1.31-1.89 x $10^4 M_\odot$. McConnachie (2012) compares the masses of known dwarf galaxies in the local group, where they find no dwarf galaxies with lower HI masses than $10^5 M_\odot$. The findings of McConnachie (2012) could be an observational bias as it uses relatively aged surveys that were less sensitive to low mass HI objects than the ALFALFA survey. Currently the lowest HI mass galaxy known is Phoenix with an HI mass of 1.2 x $10^5 M_\odot$ (McConnachie (2012)), a magnitude of difference with the computed mass of AGC 268071 in this research. With a mass range of 1.87-2.72 x $10^4 M_\odot$ AGC 268071 could be the lowest HI mass galaxy known, but it is questionable whether all the flux in the VLA data is being recovered, so the found mass can be considered a lower limit. However, Adams et al. (2016) identified three UCHVCs as excellent dark galaxy candidates with HI masses ranging from 0.64-50 x $10^5 M_\odot$.

The stellar mass as measured by Rhode et al. (2023) has a value of 0.17-2.0 x $10^5 M_\odot$ for a distance of 570 kpc. Using this stellar mass and the HI mass computed in this research from the 45" integrated flux value at the same distance, the ratio of the two ranges between 0.6713 and 7.899. The $M_{HI}/M_\star$ ratio determined in Rhode et al. (2023) was between 1.21 and 12.48. This discrepancy between found $M_{HI}/M_\star$ ratios in Rhode et al. (2023) and this research once again points towards the ALFALFA survey likely including additional Galactic HI emission in its observation, or the possibility that not all flux in the VLA data is being recovered, or a combination of both.

5.4 Implications Future Candidates

Based on the compact HI morphology and the small velocity gradient found across the surface of AGC 268071, the source is likely to be a gas bearing dark matter halo. However, based on the seemingly large CNM component of the source, the gas cloud in Galactic halo hypothesis can not be fully ruled out.

This research highlights the importance of sufficient masking and cleaning to avoid noise contamination. Two masks was likely not enough to properly isolate source emission, and the masks were formed based on too low of a resolution. Furthermore, the choice of lowering the resolution in the image plane rather than in the uv plane during the image creation changed the noise characteristics of the data, making comparisons with other studies less straight-forward.

Based on the found HI mass of AGC 268071, the masses of UCHVCs could possibly reach lower values than expected. To verify this and accurately determine the HI mass limits, precise information about the distance to ACG 268071 is needed. This information can be acquired through high resolution optical detection of the stellar component of AGC 268071 to resolve some of the characteristics of the stars, which then can be compared to known stellar populations to compute the distance. To further confirm the galaxy hypothesis through HI detection, AGC 268071 should be imaged at a higher resolution to increase the SNR of the source, allowing more of the flux to be recovered. Additionally, achieving a higher SNR would reduce the likelihood of underestimating the velocity dispersion.

6 Summary

This study aimed for extra insight into the origin of ultra-compact high velocity cloud AGC 268071. The two possible natures discussed are a gas bearing dark matter halo (galaxy) or a gas cloud in the Galactic halo. The source AGC 268071 was discovered in the ALFALFA HI line survey and identified by Adams et al. (2013) as an UCHVC. A decade later Rhode et al. (2023) detected an optical stellar overdensity of 11 stars at a seperation of 1.5 kpc from the ALFALFA HI detection. To gain understanding about the nature of this source, high resolution neutral hydrogen imaging is used to characterize the HI content of the gas in AGC 268071. The resolution of the original VLA data was 45 arcseconds, which was then after the masking and cleaning process smoothed to 105 arcseconds in the image plane.

The HI content of AGC 268071 was analyzed through the use of moment maps. The moment zero map portraying the HI distribution showed a smooth and compact source morphology at both resolutions. The first moment map displaying the recessional velocity of the source showed a relatively small velocity gradient across the surface at both resolutions. The velocity gradients measured here are 5 km s^{-1} and 3 km s^{-1} for the 45" and 105" data respectively. The second moment map displaying the velocity dispersion showed a small maximum dispersion of 6 km s^{-1} and 5.4 km s^{-1} for the 45" and 105" data respectively, corresponding to a large fraction of CNM in the source. The latter velocity dispersion could be an underestimate from working with low resolution data.

The observations show a discrepancy between the VLA flux and the ALFALFA flux, which could be attributed to Galactic contamination in the ALFALFA data, missing flux in the VLA data, or a combination of both.

Assuming a galactic nature of AGC 268071, the HI mass can be computed based on the distance given by Rhode et al. (2023), which results in an HI mass of 1.87-2.72 x $10^4 M_\odot$. Using the stellar mass as measured by Rhode et al. (2023), the $M_{HI}/M_\star$ ratio ranges from 0.6713 to 7.899. Both the HI mass and the mass ratio is around an order of magnitude lower than the values computed with ALFALFA and in Rhode et al. (2023), which is directly connected to the low integrated line flux recovered in the VLA data.

From the HI content of AGC 268071 analyzed in this study, the most likely nature of the source is galactic, based on the compact morphology and small velocity gradient present across the surface. However, based on the relatively low velocity dispersion across the source, the gas cloud in Galactic halo hypothesis can not be fully ruled out. To further confirm the galaxy hypothesis through HI detection, AGC 268071 should be imaged at a higher resolution to try to increase the SNR of the source. Based on the found HI mass of AGC 268071, the masses of UCHVCs could possibly reach lower values than expected. To verify this and accurately determine the HI mass limits, precise information about the distance to AGC 268071 is needed through high resolution optical imaging of the stellar component of the source.

7 Appendix

The command log of AGC 268071 can be found through the following link:

https://www.evernote.com/shard/s578/sh/acd69ccb-a39d-1ebd-2fef-a0c1dce7da8b/v9zCDLJAsfms-RyqHTJc-ma9RI9gbAH_z6okJEhh2jXNp6poa1QmK49KGA

8 Acknowledgements

I would like to thank my supervisor Betsey Adams for her incredible support, guidance and patience throughout the whole bachelor thesis. I also thank Martin Vogelaar and the computer-group for offering much needed support with Kapteyn software malfunctions.

References

Adams, E., Faerman, Y., Janesh, W., Janowiecki, S., Oosterloo, T., Rhode, K., Giovanelli, R., Haynes, M., Salzer, J., Sternberg, A., Cannon, J., and Muñoz, R. (2014). AGC198606: A gas-bearing dark matter minihalo? *Astronomy astrophysics*, 573:L3.

Adams, E. A. K., Giovanelli, R., and Haynes, M. P. (2013). A CATALOG OF ULTRA-COMPACT HIGH VELOCITY CLOUDS FROM THE ALFALFA SURVEY: LOCAL GROUP GALAXY CANDIDATES? *Astrophysical journal/The Astrophysical journal*, 768(1):77.

Adams, E. A. K., Oosterloo, T., Cannon, J., Giovanelli, R., and Haynes, M. P. (2016). Identifying galaxy candidates in WSRT H i imaging of ultra-compact high velocity clouds. *Astronomy astrophysics*, 596:A117.

Arras, P., Frank, P., Leike, R., Westermann, R., and Enßlin, T. A. (2019). Unified radio interferometric calibration and imaging with joint uncertainty quantification. *Astronomy astrophysics*, 627:A134.

Blitz, L., Spergel, D., Teuben, P., Hartmann, D., and Burton, W. (1996). High velocity clouds : remnants of local group formation. *Bulletin of the American Astronomical Society/Bulletin - American Astronomical Society*, 189.

Blok, W. D., Healy, J., Maccagni, F., Pisano, D. J., Bosma, A., English, J., Jarrett, T., Marasco, A., Meurer, G. R., Veronese, S., Bigiel, F., Chemin, L., Fraternali, F., Holwerda, B., Kamphuis, P., Klockner, H., Kleiner, D., Leroy, A., Mogotsi, M., Oman, K. A., Schinnerer, E., Verdes-Montenegro, L., Westmeier, T., Wong, O. I., Zabel, N., Amram, P., Carignan, C., Combes, F., Brinks, E., Dettmar, R., Gibson, B., Józsa, G., Koribalski, B., McGaugh, S., Oosterloo, T., Spekkens, K., Schroder, A., Adams, E., Athanassoula, E., Bershady, M., Beswick, R., Blyth, S., Elson, E. C., Frank, B. S., Heald, G., Henning, P., Kurapati, S., Loubser, S. I., Lucero, D., Meyer, M., Namumba, B., Oh, S.-H., Sardone, A., Sheth, K., Smith, M., Sorgho, A., Walter, F., Williams, T., Woudt, P., and Zijlstra, A. (2024). Mhongoose – a meerkat nearby galaxy hi survey. *null*.

Bregman, J. N. (1980). The galactic fountain of high-velocity clouds. *Astrophysical journal/The Astrophysical journal*, 236:577.

Bullock, J. S. (2010). Notes on the Missing Satellites Problem. *arXiv (Cornell University)*.

Burke, B. F., Graham-Smith, F., and Wilkinson, P. N. (2019). *An introduction to radio astronomy*.

Cannon, J., College, M., Adams, E., for Radio Astronomy, N. I., Haynes, M., University, C., Janesh, W., University, C. W. R., Janowiecki, S., Texas at Austin, U. o., Rhode, K., University, I., Salzer, J., University, I., Smith, N., and University, I. (2023). Observing application. Technical report.

De Heij, V., Braun, R., and Burton, W. B. (2002). High–resolution imaging of compact high–velocity clouds. *Astronomy astrophysics*, 391(1):67–81.

Faerman, Y., Sternberg, A., and McKee, C. F. (2013). ULTRA-COMPACT HIGH VELOCITY CLOUDS AS MINIHALOS AND DWARF GALAXIES. *Astrophysical journal/The Astrophysical journal*, 777(2):119.

Giovanelli, R., Haynes, M. P., Kent, B. R., and Adams, E. A. K. (2010). ARE NEWLY DISCOVERED h I HIGH-VELOCITY CLOUDS MINIHALOS IN THE LOCAL GROUP? *The astrophysical journal. Letters*, 708(1):L22–L25.

Hess, K. M., Luber, N. M., Fernández, X., Gim, H. B., Van Gorkom, J. H., Momjian, E., Gross, J., Meyer, M., Popping, A., Davies, L. J. M., Hunt, L., Kreckel, K., Lucero, D., Pisano, D. J., Sanchez-Barrantes, M., Yun, M. S., Dodson, R., Vinsen, K., Wicenec, A., Wu, C., Bershady, M. A., Chung, A., Davis, J. D., Meyer, J. D., Henning, P., Maddox, N., Smith, E. T., Van Der Hulst, J. M., Verheijen, M. A. W., and Wilcots, E. M. (2018). CHILES: H imorphology and galaxy environment atz = 0.12 andz = 0.17. *Monthly Notices of the Royal Astronomical Society*, 484(2):2234–2256.

Irwin, M. J., Belokurov, V., Evans, N. W., Ryan-Weber, E. V., De Jong, J. T. A., Koposov, S., Zucker, D. B., Hodgkin, S. T., Gilmore, G., Prema, P., Hebb, L., Begum, A., Fellhauer, M., Hewett, P. C., Kennicutt, R. C., Wilkinson, M. I., Bramich, D. M., Vidrih, S., Rix, H.-w., Beers, T. C., Barentine, J. C., Brewington, H., Harvanek, M., Krzesinski, J., Long, D., Nitta, A., and Snedden, S. A. (2007). Discovery of an Unusual Dwarf Galaxy in the Outskirts of the Milky Way. *Astrophysical journal/The Astrophysical journal*, 656(1):L13–L16.

Jones, M. G., Haynes, M. P., Giovanelli, R., and Moorman, C. (2018). The ALFALFA H i mass function: a dichotomy in the low-mass slope and a locally suppressed 'knee' mass. *Monthly Notices of the Royal Astronomical Society*, 477(1):2–17.

Kellermann, K. I., Bouton, E. N., and Brandt, S. S. (2020). *The Very Large Array*, pages 319–390. Springer International Publishing, Cham.

Lee, G., Hwang, H. S., Lee, J., Shin, J., and Song, H. (2024). Understanding the formation and evolution of dark galaxies in a simulated universe. *Astrophysical journal/The Astrophysical journal*, 962(2):129.

McConnachie, A. W. (2012). THE OBSERVED PROPERTIES OF DWARF GALAXIES IN AND AROUND THE LOCAL GROUP. *The Astronomical journal*, 144(1):4.

Muller, C., Oort, J., and Raimond, E. (1963). Hydrogène neutre dans la couronne galactique? *Comptes Rendus l'Academie des Sciences*, 257.

Oort, J. (1970). The formation of galaxies and the origin of the high-velocity hydrogen. *Astronomy and Astrophysics.*

Perley, R. A., Chandler, C. J., Butler, B., and Wrobel, J. M. (2011). THE EXPANDED VERY LARGE ARRAY: a NEW TELESCOPE FOR NEW SCIENCE. *The astrophysical journal. Letters*, 739(1):L1.

Putman, M. E. (1999). High-Velocity clouds related to the Magellanic system. *arXiv (Cornell University).*

Rhode, K. L., Smith, N. J., Janesh, W., Salzer, J. J., Adams, E. A. K., Haynes, M. P., Janowiecki, S., and Cannon, J. (2023). A Search for Gas-Rich Dwarf Galaxies in the Local Universe with ALFALFA and the WIYN One Degree Imager. *The Astronomical journal*, 166(3):113.

Schinzel, F. K., Sokolovsky, K. V., D'Ammando, F., Burnett, T. H., Max-Moerbeck, W., Cheung, C. C., Fegan, S. J., Casandjian, J. M., Reyes, L. C., Villata, M., Raiteri, C. M., Agudo, I., Calle, O. J. A. B., Carosati, D., Casas, R., Gómez, J. L., Gurwell, M. A., Hsiao, H. Y., Jorstad, S. G., Kimeridze, G., Konstantinova, T. S., Kopatskaya, E. N., Koptelova, E., Kurtanidze, O. M., Kurtanidze, S. O., Larionov, V. M., Larionova, E. G., Larionova, L. V., Marscher, A. P., Morozova, D. A., Nikolashvili, M. G., Roca-Sogorb, M., Ros, J. A., Sigua, L. A., Spiridonova, O., Troitsky, I. S., Vlasyuk, V. V., Lobanov, A. P., and Zensus, J. A. (2011). Identification of-ray emission from 3C 345 and NRAO 512. *Astronomy astrophysics*, 532:A150.

Schmidt, K., Geyer, F., Fröse, S., Blomenkamp, P.-s., Brüggen, M., De Gasperin, F., Elsässer, D., and Rhode, W. (2022). Deep learning-based imaging in radio interferometry. *Astronomy astrophysics*, 664:A134.

Schneider, D. P., Hall, P. B., Richards, G. T., Strauss, M. A., Vanden Berk, D. E., Anderson, S. F., Brandt, W. N., Xiaohui, F. A., Jester, S., Gray, J. I., Gunn, J. E., SubbaRao, M. U., Thakar, A. R., Stoughton, C., Szalay, A. S., Yanny, B., York, D. G., Bahcall, N., Barentine, J., Blanton, M. R., Brewington, H., Brinkmann, J., Brunner, R. J., Castander, F. J., Csabai, I., Frieman, J. A., Fukugita, M., Harvanek, M., Hogg, D. W., Ivezić, , Kent, S. M., Kleinman, S. J., Knapp, G. R., Kron, R. G., Krzesiński, J., Long, D. C., Lupton, R. H., Nitta, A., Pier, J. R., Saxe, D. H., Shen, Y., Snedden, S. A., Weinberg, D. H., and Wu, J. (2007). The Sloan Digital Sky Survey Quasar Catalog. IV. Fifth Data Release. *The Astronomical journal*, 134(1):102–117.

Thompson, A. R., Moran, J. M., and Swenson, G. W. (2001). *Interferometry and synthesis in radio astronomy.*

Van Woerden, H., Schwarz, U. J., Peletier, R. F., Wakker, B. P., and Kalberla, P. M. W. (1999). A confirmed location in the Galactic halo for the high-velocity cloud ‘chain A’. *Nature*, 400(6740):138–141.

Verheijen, M. A. and Sancisi, R. (2000). The ursa major cluster of galaxies. iv ; hi synthesis observations. *Astronomy astrophysics.*

Wakker, B. P. and Van Woerden, H. (1997). HIGH-VELOCITY CLOUDS. *Annual review of astronomy and astrophysics*, 35(1):217–266.

Warren, S. R., Skillman, E. D., Stilp, A. M., Dalcanton, J. J., Ott, J., Walter, F., Petersen, E. A., Koribalski, B., and West, A. A. (2012). TRACING COLD H I GAS IN NEARBY, LOW-MASS GALAXIES. *Astrophysical journal/The Astrophysical journal*, 757(1):84.

Wilson, T. L., Rohlfs, K., and Hüttemeister, S. (2013). *Tools of radio astronomy.*