University of Cape Town

Department of Astronomy

Neutral Hydrogen Mass Distributions in Galaxies Uncovered by Three-Dimensional Stacking

---

Dissertation presented for the degree of
Master of Science in the Department of
Astronomy

Student:
Andrew Leon Firth

Supervisor:
Professor Matthew Bershady

Date: 11 February 2024

Plagiarism declaration

I, Andrew Leon Firth, know the meaning of plagiarism and declare that all of the work in the document, save for that which is properly acknowledged, is my own.

Abstract

Atomic hydrogen (HI) serves as the raw fuel from which the star-forming molecular hydrogen forms - making it an important tracer of galaxy evolution. Due to the intrinsic faintness of the HI emission line (observed at rest at 21 cm), detecting HI emission in galaxies beyond a z = 0.1 becomes difficult, thus, many statistical techniques have since been developed in order to extend the capabilities of already rapidly improving radio observatories.

HI stacking is one such technique, involving the co-addition of the HI spectra of a sample of galaxies, exploiting statistical properties of random noise to achieve a high signal-to-noise measure of their average HI content. While this technique focuses on extracted one-dimensional spectra, research extending this concept to the parent data-cube of the one-dimensional spectra is a growing area of research. Studies such as those by (Chen et al., 2021; Sinigaglia et al., 2022) focus on stacking in three-dimensions to improve the point-spread function, HI signal detectability, and hence improve HI mass-measurements.

We present an initial exploration of a novel approach to stacking, working at the level of imaged radio cubelets, i.e., in three dimensions (3D). Unlike previous stacking works focusing on the UV-plane, and extracts a one-dimensional profile to recover the average HI-mass of the stack, we attempt to recover fine structure within the stack. While our sample sizes prove to be the reason for the limited results, we demonstrate the potential increase in signal-to-noise by assessing the effect sigma-clipping a of stacking on galaxy surface-brightness profiles; comparing one-dimensional and three-dimensionally co-added techniques simultaneously.

Using full-resolution data cubes from the Noordermeer et al. (2005) sub-sample of Westerbork observations of neutral Hydrogen in Irregular and Spiral galaxies (WHISP) survey (van der Hulst et al., 2001), we make use of the homogeneous data reduction pipeline, well-documented global properties, and mass-distribution studies to calibrate our stacking suite.

In developing this pipeline, we demonstrate the concept of aligning signal through various geometric transformations (rotation, stretching, and spectral-inversion) as a first approach to a 'Comprehensive Unearthing of Baryons in Extended-regions using a Stacker' (Cube Stacker). Developed in Python 3.8, this pipeline bridges the gap of stacking techniques targeted at resolved targets, preserving structure with the aim of utilising an improved SNR to uncover faint HI features. While our stacking is based on 21cm measurements of HI distributions, we explore the possibility of utilising galaxy scaling-relations - incorporating the high-fidelity optical measurements of galaxies to be observed with high-sensitivty facilities such as MeerKAT. Thus, we present measures of central concentration (i.e. C80,20) for the 33 NSA-selected galaxies, and discuss the merits of an optically-based stacking-pipeline in uncovering faint HI in existing or upcoming blind surveys.

Acknowledgement

When registering to undertake this MSc in February of 2020, no one could have predicted what would unfold merely a month later. This thesis has followed me through four years, a global pandemic, eight countries, two terabytes of storage, and four astronomical observatories. I am filled with immense gratitude for those who helped bring this work across the line.

To Matt, my supervisor, whose guidance and commitment to regular meetings enabled me to continue to work and grow in a career which has tested me and moulded me grately. I am eternally grateful for your kindness and the opportunities you have given me.

To those at the South African Astronomical Observatory, Sabyasachi, Waanie, Vee, Linda, Patricia, and Retha - your direct support and guidance have been invaluable and I am eternally indebted.

To my office mates through the years, Brian, Antoine, Mikhail, Lisa, Thobekile, Gideon, JC, Kyle, Gilbert, and Dilbert - thank you for keeping me sane, mostly.

Sriram, FUB. Pigeon noises.

To those who let me forget the stress (occasionally), Duncan, Bret, the guys at fives, Jam, and Chanté - thank you for a memorable four years.

To my mom and dad, for their endless (although tested) support of this career choice.

Finally, to Snappy and Snippy, the ginger-est of kitties who often visited my room while I worked.

Ubuntu. E pluribus unum (unus mundus).

Enkosi. Dankie. Dank u well. Dank je. Danke schön. Merci. Gracias. Shukran. Nandi.

Thank you.

The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily attributed to the NRF.

Contents

LIST OF FIGURES

LIST OF TABLES

Chapter 1

Introduction

1.1 HI in the Local Universe

The release of a 21cm (1420 MHz) photon is attributed to the hyperfine transition in the ground state of a neutral hydrogen atom, or HI. Also known as atomic hydrogen, the high abundance of HI in the Universe compensates for the relatively low transition probability of the flip in the spin state of an electron to a lower energy state.

HI is a fundamental component of the interstellar medium (ISM) of late-type galaxies, which include spiral and irregular galaxies. These galaxies are characterized by their substantial gas content and ongoing star formation processes. The HI content is particularly significant as it provides the raw material for star formation. Thus, an investigation into the evolution of galaxies is an investigation into the evolution of gas within galaxies.

Observational studies, such as those conducted by the Arecibo Legacy Fast ALFA (ALFALFA, Giovanelli et al. (2005)) survey, have shown that HI is typically distributed in a thin disk that usually extends beyond the visible stellar disk, often forming a halo around the galaxy (Guo et al., 2017). These gaseous haloes are also known as the Circumgalactic Medium (CGM) whose extent is determined by the dark matter potential (Donahue & Voit, 2022). In terms of specific mass fractions, studies show that the HI-to-stellar mass-ratio (M_{HI}/M_{*}) can vary widely among different galaxy types. For example, massive galaxies often exhibit lower gas fractions compared to their less massive counterparts, with the gas fraction decreasing with increasing stellar mass (Catinella et al., 2010). The HI mass functions obtained from the ALFALFA survey and the GALEX Arecibo SDSS Survey (GASS, Catinella et al. (2010)) reveal that while most galaxies contain some HI, a significant number of low-mass galaxies can have high HI mass fractions, sometimes exceeding 50% of their total baryonic mass.
This CGM forms a reservoir that regulates the flow of gas due to a multitude of internal processes within the galactic disc, as well as external environmental influences. The acquisition of "new" HI in galaxies is a complex process. In broader contexts, numerous galaxy properties depend

on environment. The morphology-density relation relates the observed fraction of elliptical galaxies to spirals in the centres of galaxy clusters – regions of higher galaxy concentration (Dressler, 1980). Galaxy interactions lead to the removal of star-birthing gasses from galaxies in these dense environments thus making environment a critical consideration when studying the evolution of galaxies Leclercq et al. (2022).

In the absence of interactions, high-redshift galaxies obtain their HI from the environment through cold-mode accretion seen as streams of filaments within the CGM (Kereš et al., 2005; Kereš & Hernquist, 2009) or through hot-mode accretion from instability-driven precipitation (Donahue & Voit, 2022) or radiative cooling (Stern et al., 2019).

Gas accretion in the local Universe, however, occurs mainly via galactic fountain flows, the CGM, and (once more) the environment in which the galaxy resides (Kamphuis et al., 2022; Zhang et al., 2023). Galaxies seldom exist in isolation with more than 50% of galaxies residing in groups of at least two gravitationally bound galaxies (Eke et al., 2005; Tempel et al., 2014) This makes the role of interactions very important in the context of studying galaxy evolution.

Typical column densities of extra-planar gas clouds around galaxies are generally found within the range of $10^{19}\,\mathrm{cm}^{-2}$ to $10^{22}\,\mathrm{cm}^{-2}$ (Heald et al., 2007; Marasco, A. & Fraternali, F., 2011). The low column densities of HI means that highly sensitive instruments are required for the detection of HI beyond the disc of galaxies.

Radio observations usually take one of two forms, namely: large, single-dish surveys resulting in high-sensitivity (due to a larger collecting area) observations, typically fixed on a large and single region of the sky, of single-to-several spatial elements on the sky, or synthesis telescopes (radio interferometers) performing higher-resolution HI studies with a larger simultaneous field-of-view (FoV) due to the typically smaller individual dishes within the telescope array.

Single-dish telescopes typically take two forms, namely: fixed or steerable. Fixed telescopes such as Arecibo Observatory (now decommissioned) in Puerto Rico, have a stationary dish that is usually placed in a natural depression or a purpose-built structure. These telescopes perform transit observations, limited to observing the part of the sky directly above them and cannot be pointed elsewhere. In contrast, steerable single-dish telescopes, such as the Green Bank Telescope (GBT) in West Virginia, have a movable dish that can be directed to different parts of the sky. Despite having a smaller collecting area than their fixed counterparts, steerable telescopes offer much greater coverage of the sky, with a resolution trade-off.

Examples of single-dish surveys and instruments, limited to fixed telescopes for brevity, include the Arecibo Legacy Fast ALFA (ALFALFA, Giovanelli et al. (2005)) survey, a significant observational project conducted using the Arecibo Observatory in Puerto Rico, between 2005 and 2011. The primary goal of ALFALFA was to map the distribution and properties of HI within $z < 0.06$, aiming to create a detailed three-dimensional map of HI gas in galaxies. Despite the decommissioning of ARECIBO, ALFALFA-styled extragalactic surveys now continue on the Five-hundred-meter Aperture Spherical Telescope (FAST). To date, global parameters such as HI-velocity (cz), HI-mass (M_{HI}), and line-width (W_{50}) were obtained for over 500 galaxies (Kang et al., 2022).

On the other hand, major radio astronomy facilities performing HI aperture synthesis have the spatial resolution necessary to study the spatially resolved kinematics and distribution of HI within nearby galaxies. These include the Very Large Array (VLA) in the United States of America, *Meer* Karroo Array Telescope (MeerKAT) in South Africa, the Australian Square Kilometre Array Pathfinder (ASKAP) telescope, and the Australia Telescope Compact Array (ATCA) in Australia. By combining signals from multiple antennas, these facilities create high-resolution radio images, allowing astronomers to investigate the distribution, kinematics, and dynamics of HI gas in galaxies.

1.1.1 The Westerbork Synthesis Radio Telescope (WSRT)

Relevant to the work in this thesis, and adding to the list of radio interferometers in the previous paragraph, the Westerbork Synthesis Radio Telescope (WSRT) in the Netherlands uses aperture synthesis to combine signals from a linear array of 14 telescopes, each with a diameter of 25m. WSRT also forms part of the European VLBI Network (EVN). We consider the functionality of the WSRT in the years pertaining to the WHISP data-collection period (1992-1996). Due to the East-West orientation of the WSRT array, higher resolution is achieved in right-ascension; the longest baseline achieving a full resolution of 14" $\times$ 14"$/sin\delta$, where δ is the declination.

With the need for higher-resolution studies of the kinematic and spatial distribution of HI in galaxies, the demand for synthesis radio telescopes has increased. Most studies focused on singular galaxies in great detail, but at the time very few large-scale surveys were conducted in and around the 21cm line. In addition to the increased resolution due to the increased baseline lengths, aperture synthesis telescopes have the additional benefit of increased field-of-view over their single-dish counterparts. This is due to the combination of multiple beams on the sky as compared to the single beam of a single antenna.

Additional HI studies focused on the distribution of neutral Hydrogen in cluster galaxies, with notable contributions by Verheijen & Sancisi (2001) studying the Ursa Major cluster. The HI Nearby Galaxy Survey (THINGS; Walter et al. (2008a)) investigated the spatially resolved HI in a sample of 30 nearby galaxies, utilizing the Very Large Array in the USA. Existing technological constraints, which demand prolonged observation periods, have traditionally confined such surveys to distances within D < 15 Mpc. Since the completion of the MeerKAT array, with greater sensitivity, resolution, and field-of-view than the VLA, data from the MeerKAT HI Observations of Nearby Galactic Objects; Observing Southern Emitters (MHONGOOSE, de Blok et al. (2016)) survey is on hand and currently being investigated. As a result, MHONGOOSE is a deep survey of the distribution of HI in a sample of 30 disk and dwarf galaxies ranging in mass from 106 $M_{\odot}$ to 1011 $M_{\odot}$ in the local universe. MeerKAT International Giga-Hertz Tiered

Extragalactic Exploration (MIGHTEE, Maddox et al. (2021)), Looking At the Distant Universe with the MeerKAT Array survey (LADUMA, Baker et al. (2018)), and the MeerKAT Fornax Survey (Serra et al., 2016). Meanwhile, surveys with the Australian SKA Pathfinder (ASKAP) include the Widefield ASKAP L-band Legacy All-sky Blind Survey (WALLABY, Koribalski et al. (2020a)), and the Deep Investigation of Neutral Gas Origins survey (DINGO, Meyer (2009)) – promise to take a significant step forward in our understanding of the role of HI in galaxy evolution out to $z \geq 1$.

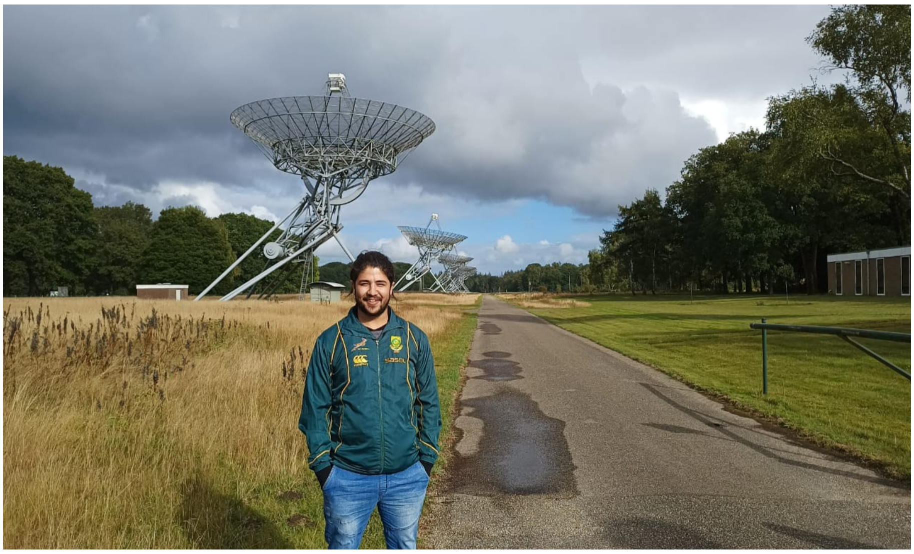


Figure 1.1: **The Westerbork Synthesis Radio Telescope, located in Hooghalen, Netherlands.**

1.1.2 Westerbork observations of neutral Hydrogen in Irregular and SPiral galaxies (WHISP)

In addition to the surveys listed in the previous sub-section, the Westerbork HI Survey of Spiral and Irregular Galaxies (WHISP, van der Hulst et al. (2001);Swaters et al. (2012)) forms the parent sample of galaxies we have selected (see Chapter 2) to develop the *Cube Stacker* pipeline. The WHISP project consists of two main aspects, namely:

1. A systematic study of the HI component of galaxies, covering distribution and kinematic investigations

2. Aggregating a large data set for secondary studies not related to the HI component, but those for which HI is a great probe or proxy, such as the study of rotation curves and dark matter probing.

WHISP galaxies were selected mainly from the Uppsala General Catalogue of galaxies (UGC, Nilson (1973)), a comprehensive catalogue of relatively bright galaxies with measurements such as positions, apparent sizes, and apparent magnitudes. Thus, the galaxies had to conform to the following selection criteria:

1. B-band diameters larger than 1.5’

2. declination north of 20 °

This was due in large part to the observational constraints of WSRT. The declination limit avoids the degradation of the synthesised beam due to the East-West orientation of the WSRT array, while the size-constraint is to ensure enough linear resolution in order to separate the fine HI-structure and velocity fields (van der Hulst et al., 2001). The resultant sample of 3148 galaxies was then further restricted to have flux densities larger than 100mJy, reducing the primary WHISP sample to 409 galaxies, ranging in Morphological type.

Primary publications emerging from the WHISP survey include Swaters & Balcells (2002) who conducted a pioneering study providing comprehensive analyses of the HI properties of galaxies observed in the WHISP survey. Similarly, Verheijen & Sancisi (2001) research the Tully-Fisher relation in the context of the WHISP survey, revealing intrinsic connections between galaxy luminosity and rotation velocity. Additionally, van der Hulst & Sancisi (2005) explored the rotation curves and mass distributions of WHISP galaxies, while research by de Blok et al. (2008) and Noordermeer et al. (2005) delved into the kinematics of early-type disk galaxies.

1.2 Extra-planar Gas in Galaxies

Cold gas ($< 10^4 K$) in galaxies, largely in the form of neutral hydrogen (HI), is supplied through various accretion, merger, and feedback processes (Kereš et al., 2005). This gas cools and condenses to form molecular gas as the precursor to gravitational collapse which ignites nuclear fusion during the process of star formation. Thus, as the raw fuel which drives star formation, HI serves as a highly sensitive tracer for galaxy interactions and minor merger events (Holwerda et al., 2011). Without gas accretion and the ability to cool the gas to the molecular phase, star formation ceases, resulting in passive stellar evolution within a galaxy.

With the advent of radio interferometry and HI-imaging, it is possible to study the distribution and kinematics of the HI in detail. Not surprisingly, most of the HI is found in star-forming galaxy discs, where the HI is in a thin layer of size on the order of 100 parsecs (van der Kruit & Freeman, 2011). However, it is clear that HI also resides in thicker disc distributions in some cases (e.g., NGC 891), and in general, it has a larger radial distribution than the stars (Oosterloo et al., 2007; Bigiel, 2010).

Given the gas consumption time-scales for the HI in galaxy discs are relatively short compared to the star-formation rate (mass into stars per unit time) (Kennicutt, 1998), there is the question of how HI in discs is replenished. Over the Milky Way’s lifetime, for example, the observed star-formation rates (SFR) in the solar neighbourhood have been constant (Binney et al., 2000). However, this implies that the gas content is not being exhausted – again suggesting a replenishing source. It is believed this source takes the form of in-falling gas along filaments in the inter-galactic medium (Sancisi et al., 2008), in a process known as accretion. This leads naturally to an exploration of extra-planar gas both as a signpost of possible gas inflow and also as a potential signature of outflow from stellar and AGN feedback (Taylor & Kobayashi, 2015).

The evolution of the stellar content is well constrained, thanks to multiwavelength studies using, e.g., Spitzer (Werner et al., 2004) in the infrared, the Sloan Digital Sky Survey (York et al., 2000) in the optical, and the Galaxy Evolution Explorer (Bianchi & GALEX Team, 2000)) in the ultraviolet. In contrast, the accretion of the cold gas component of galaxies, and its evolution, is less well-constrained.

Direct observational evidence of gas accretion has been difficult to obtain, with the intrinsically weak HI signal requiring long integration times to detect the faint accreted HI clouds. For instance, the WHISP observations consisted of single 12-hour observations on the WSRT (Noordermeer et al., 2005). Evidence for cold accretion takes the form of HI structures around galaxies, warps, and extraplanar gas or halo gas.

High-sensitivity HI observations have often focused on the vertical structure and the detailed kinematics of HI disks. For example, in their investigation of NGC 2403, Fraternali et al. (2002) uncovered an extended, differentially rotating anomalous HI layer akin to the high-velocity clouds in the Milky Way. VLA observations characterise an anomalous HI cloud trailing NGC 3521 as a slow-rotating halo gas component, consistent with similar findings for other nearby galaxies (Elson, 2013). In general, the source of these anomalous gas clouds is still unclear, however, they are believed to be a remnant of a galactic fountain mechanism wherein the gas is ejected out of the plane of the galaxy due to stellar winds or supernova events. The process by which gas, more specifically HI, enters and is expelled via star-formation processes from the inter-galactic medium is known as the gas cycle.

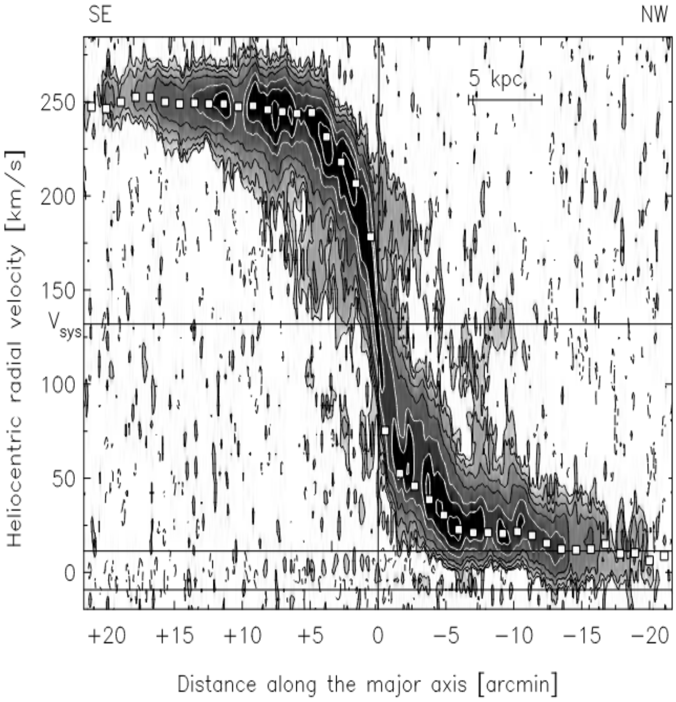


Figure 1.2: Position-Velocity diagram (PVD) extracted along the major axis of NGC 2403, with white squares marking the rotation curve of each end of the galaxy. The anomalous gas is most prominent at (d,v) (+5,+175) (From Holwerda et al. (2011))

1.3 Stacking

The direct detection of the hyperfine transition, observed at 21cm wavelengths, is the preferred method for tracing HI in galaxies in the local Universe. Detecting faint hydrogen clouds within galaxies is a persistent challenge. While the 21cm wavelength detection of the hyperfine transition remains the primary method for tracing HI in the local Universe, sensitivity-resolution trade-offs persist. Even in the age of MeerKAT, with an expected count of 10,000 galaxies in the Looking At the Distant Universe with the MeerKAT Array survey (LADUMA, Baker et al., 2018) survey volume, it is anticipated that only ∼3500 of the galaxies will have 5σ HI detections (Baker et al., 2018).

By employing radio interferometers, these signals can be synthesised to obtain both spatial and spectral information about the emitters. The wide instantaneous field-of-view provided by large radio telescope arrays such as Westerbork Synthesis Radio Telescope (WSRT) and the Square Kilometre Array (SKA) precursor MeerKAT make blind all-sky surveys feasible.

Modern interferometers have benefited from increased computational power as well as technological advances. Upgrades to the Giant Meterewave Radio Telescope (uGMRT) have produced direct HI detections at z = 0.3, and has the frequency coverage capable of reaching z = 1 with the aid of statistical techniques such as stacking. Upgrades to the Westerbork Synthesis Radio Telescope, after collection of the WHISP data used in this work, have increased the field of view (FoV) and survey speed. This is achieved through the installation of the APERture Tile in Focus (Apertif, Oosterloo et al. (2009)) phased array feed system. WSRT now has a wide- and medium-deep tier, capable of covering 3500 and 350 deg^2, respectively, along with a spectral sensitivity to HI out to $z = 0.26$. The increased speed, FoV, and sensitivity make WSRT-Apertif an excellent instrument for performing deep-field surveys.

An important distinction here is the correlator bandwidth and the telescope sensitivity. Telescopes like the VLA have frequency limits that can constrain surveys such as COSMOS HI Large Extragalactic Survey (CHILES, Fernandez et al. (2015)), which focus on HI observations up to a redshift of z=0.5. However, instruments like MeerKAT used in LADUMA extend the ultra-high frequency (UHF) range, where HI line emission is redshifted for high redshift galaxies, enabling observations of HI up to z=1. While these frequency limits define the redshift range that each telescope can effectively observe, sensitivity variations between telescopes also play a critical role in detecting faint signals within those ranges.

Despite improved technology, studying atomic hydrogen is still a costly exercise in terms of observational hours. Due to the intrinsic faintness of the HI emission, an upper limit to the redshift (with current telescopes) at which HI remains detectable, $z \sim 0.45$, is quickly reached. To date, the most distant direct detection remains that of the luminous infrared galaxy COSMOS J100054.83+023126.2 at z = 0.376, derived from the first 178 hr of CHILES data (Fernández et al., 2016). As a countermeasure against the intrinsically weak HI signal, researchers have been using statistical techniques to bolster the HI signal - pushing beyond the limitations of direct studies. One such technique involves the co-addition of many different HI spectra – thereby achieving an average HI estimate across the population. This statistical technique is known as HI stacking.

Science applications of HI stacking cover a vast number of areas, including the HI content of galaxies in dense environments and how the gas content relates to other observables (J. N Chengalur et al., 2001) gas content of active galaxies (Geréb et al., 2013); measurement of Ω_{HI} at low to intermediate (z < 0.4) redshifts (Lah et al., 2007; Delhaize et al., 2013; Rhee et al., 2013) and using stacking to study the relations between HI and various stellar mass/star formation indicators (Fabello, 2012; Gereb et al., 2015).

Conventional stacking experiments are performed on a one-dimensional signal - often the extracted HI global profile in the case of HI-astronomy. An example of an algorithm developed for the stacking of one-dimensional HI spectra is the HI Stacking Software (HISS, Healy et al., 2019). HISS formalises the process followed commonly for one-dimensional stacking. Taking a list of 1D spectra and associated redshifts for the sample of galaxies to be stacked as input, HISS aligns the signal by spectrally shifting and scaling the HI spectra. As a final step, HISS applies a weighting scheme before co-adding the spectra.

The 1D spectra are usually obtained by selecting, within the containing data cube, a spatial region over which to integrate the flux. The integrated flux within the aperture is matched to the frequency of each channel, producing the characteristic "double-horn" profile typical for HI in galaxies. Conventional one-dimensional stacking is the co-addition of these "double-horn" profiles in order to assess the wings on either the approaching end or the receding end for signal recovery within the noise. The final stacked spectrum is used to derive the average properties of the galaxy sample. These properties include flux, line width, and HI mass.

During the development of a three-dimensional cube stacking software suite for this thesis work, several experiments of stacking in three dimensions have been published. An approach taken by Chen et al. (2021) examined the effects of stacking in three dimensions as opposed to the traditional stacking of 1D spectra. The authors cite two primary shortcomings in the latter 1D approach. The first is low SNR signals from the faint HI from individual galaxy observations. This makes the galaxies too weak to detect directly and as such, deconvolution is not possible. A sidelobe contamination is present in the (one-dimensionally) stacked spectrum of the resolved galaxies, leading to incorrect estimates of quantities such as mass, extent, dynamics, and structure. Unlike our approach, the authors have focused their efforts on unresolved sources, since the experiment attempted to recover the predicted (average) HI mass.

The approach taken by Chen et al. (2021) to stack the dirty 3D cubelets, can be summarised as follows: Taking the dirty cubes and their associated point-spread function cubelet (PSF-cubelet) several transformations are performed simultaneously. The dirty cubes are blueshifted spectrally to the rest frame of the target (using a rest-frequency of $\nu_0 = 1420.41 MHz$). Next, a linear interpolation is performed on the spectral axis of both the dirty cubes and the PSF-cubelets. This is to ensure a standard spectral sampling for both the dirty cubelet and PSF-cubelet. A primary beam correction is performed on the cubelet. Next, each galaxy's PSF-cubelet and dirty-cubelet are stacked, respectively, yielding a stacked PSF-cubelet and a stacked dirty-cubelet. The authors apply an inverse-variance weighting to both stacks. Finally, the stacked dirty-cubelet is *CLEANed* using the Högbom (1974) algorithm, using the stacked-PSF cubelet.

The above techniques were developed with the challenges of blind, unresolved HI surveys in mind, such that measurements of bulk properties such as HI-mass are obtained.

In this dissertation, we explore three-dimensional stacking of already imaged cubes, in the context of *resolved* nearby galaxies ($z \sim 0.035$). Unlike the stacking of the dirty cubes by Chen et al. (2021) our approach attempts to make use of the non-uniform three-dimensional distribution of HI in galaxies. Previous 3D stacking work focused still on the extracted 1D line as well as improved HI-mass estimates. Our aim, however, is in characterising (in a statistical sense) the distribution of faint HI sources external to the galaxy discs. This will be done by performing a high-quality signal alignment in 3D before assessing any features (such as beards within the PVD and streams in the moment 0 maps) uncovered within the noise. Due to the prospect of stacking blind targets, the secondary aim was to develop a HI signal-alignment routine based on optical parameters of galaxies.

In lieu of developing this optically-based stacking routine, utilising the higher-resolution observations of galaxies beyond and including the redshift limits of current radio telescope arrays, the ability to predict the distribution of HI within a galaxy unobserved in the radio is pivotal. This is especially true for stacking performed on blind-survey galaxies.

1.4 Galaxy Scaling Relations

The ability to predict the distribution of HI within galaxies, using observations in non-radio wavelength regimes will extend the limits on detection of the 21cm line. In line with this goal, quantities of physical scale such as mass, luminosity, radius, and rotational speed within a galaxy show strong correlations. Some of these correlations pertain to aspects of the galaxies' stellar components, such as the connection between stellar mass and metallicity, or the correlation between luminosity (mass) and radius (size). Others establish links between stellar properties and dark matter characteristics, such as the Tully-Fisher relation, which relates the stellar luminosity (or mass) to the galaxy's rotation speed, and scaling relations that connect the luminosity of the stellar component to the central density of dark matter halos. In some instances, the correlations transcend wavelength regimes - providing a useful proxy for predicting galaxy morphology prior to direct observation.

HI Size-Mass to Optical Luminosity Scaling-Relations

Notable works by Broeils & Rhee (1997) demonstrate B-band relations of scale between optical and HI diameters. Using short 21-cm observations of a sample of spiral and irregular galaxies, the authors note a strong correlation between the HI diameter (D_{HI}), measured at a surface density level of $1 M_{\odot} pc^2$, and the inclination and absorption-corrected 25^{th} isophotal diameter ($D_{25}^{b,i}$), such that:

$$\log D_{HI} = (1.00 \pm 0.03) \log D_{25} + (0.23 \pm 0.04). \tag{1.1}$$

In the context of stacking, scaling laws have been used to simulate the expected HI distribution of galaxies. Relating specifically to three-dimensional stacking, Chen et al. (2021) use optical scaling relations to simulate the HI content of the observed galaxy sample from which they try to recover the simulated (stacked) HI mass.

Relating optical properties to HI mass-density units, Chen et al. (2021) make use of g- and r-band SDSS photometry to obtain a measure of Johnson's B-band absolute magnitudes of the simulated galaxy. Another scaling relation particularly of interest in stacking experiments is one relating the B-band absolute-magnitude M(B) and HI-mass (M_{HI}) by Dénes et al. (2014):

$$\log M_{HI} = 2.89 - 0.34 M(B) \tag{1.2}$$

with M_{HI} the HI-mass (in units of solar-mass), and M(B) in magnitude units. The authors then use Equation 1.1 to map out the distribution of this HI-mass, spatially.

Investigating scaling relations across WHISP galaxies in the Wide-field Infrared Survey

Explorer (WISE), Naluminsa et al. (2021) present the global scaling relations between radio (21cm) and mid-infrared ($3.4\mu m - 11.6\mu m$) galaxy parameters. In their findings, they confirm the HI mass-size relation first published in (Broeils & Rhee, 1997) as:

$$\log D_{HI} = (0.506 \pm 0.003) \log M_{HI} - (3.293 \pm 0.009) \tag{1.3}$$

with M_{HI} the HI-mass (again in units of solar-mass), and D_{HI} the HI-diameter (in units of kpc).

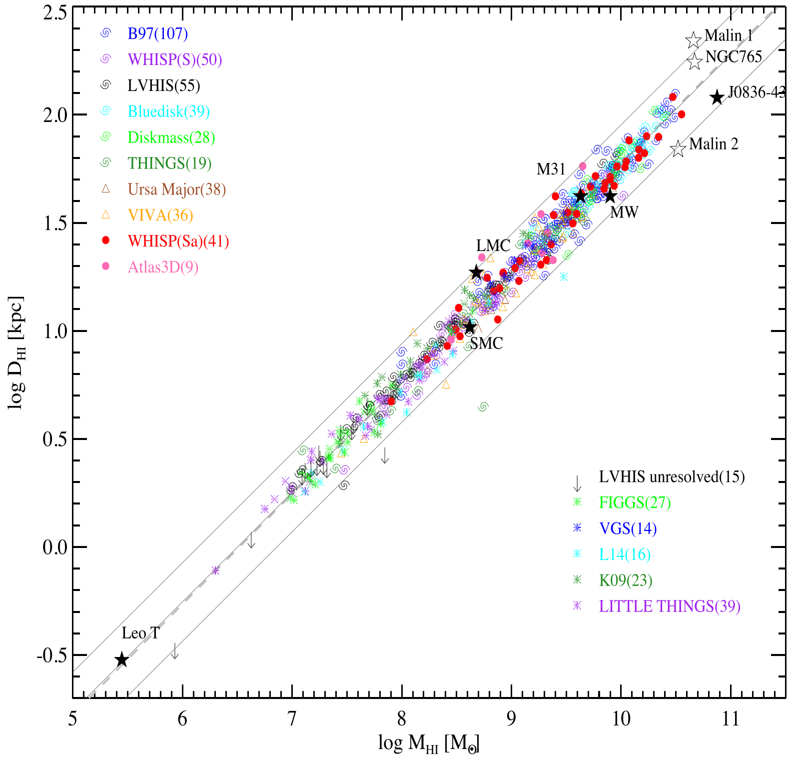


Figure 1.3: Summary of D_{HI}–M_{HI} relations of 562 galaxies from 15 interferometric data sets as presented by Wang et al. (2016). B97, and the dashed line, represents the measurements and best fit obtained by Broeils & Rhee (1997). The solid lines represent the 3σ scatter.

In addition to this Naluminsa et al. (2021) investigated the correlation between HI-mass and stellar luminosity. Defining the HI-mass as the gas contained within the 1σ isophote in the W1-band image from WISE photometry, corresponding to a surface brightness 23 mag arcsec2 (Vega units), the authors note a strong negative correlation (r = -0.91) between the HI-mass and absolute magnitude at 3.4 µm.

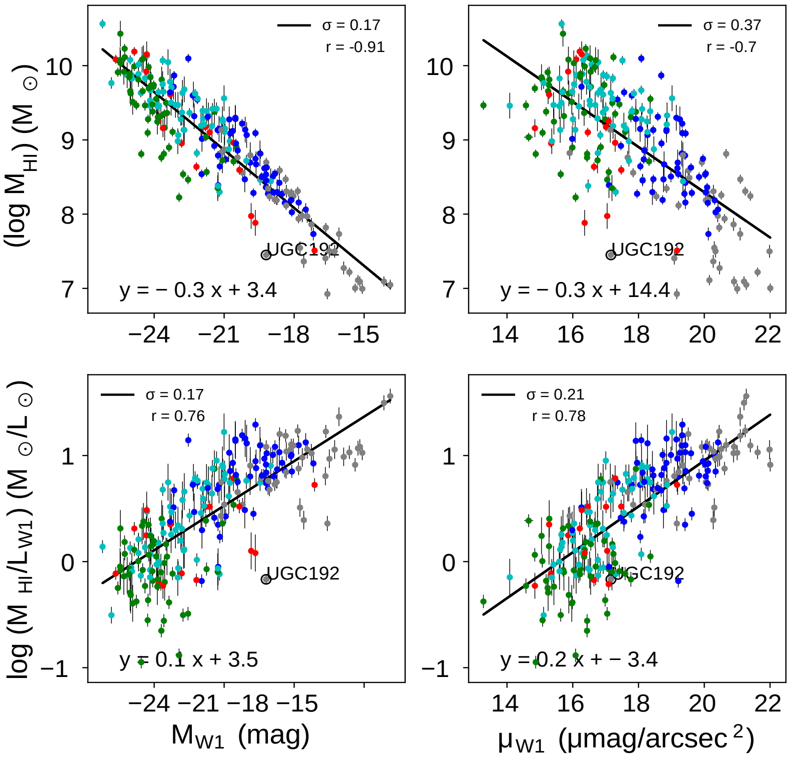


Figure 1.4: A summary of the results from (Naluminsa et al., 2021) of the scaling-relations between 21cm and 3.4 µm galaxy parameters. The plots are color-coded by morphological type, where red represents lenticulars (S0-S0/a), green represents early-type spirals(Sa-Sb), cyan represents Sbc - Scd, blue represents late-type spirals (Sd-Sm), and gray the late-type dwarfs.

Line-Width-Lumiunosity Scaling-Relations

Another broadly used scaling relation has become known as the Tully-Fisher Relation (TFR), published originally by Tully & Fisher (1977) as a relation between the B-band absolute magnitude, M(B), and the line width of the integrated HI profile (W_R) at 20% of the peak intensity. Initially, TFR was designed to be used to estimate distances to galaxies. In the past several decades the TF-relation has since benefited from a lot of work (Sanders, 1996; Verheijen & Sancisi, 2001; de Blok et al., 2008; Cannon et al., 2011; Lelli et al., 2014) to constrain the slope and scatter of the relation as a function of optical and near-infrared wavelength for the exploration of the baryonic properties, stellar population mass-to-light ratio variations, and scaling relations with dark matter.

For our purposes, TFR makes it possible to predict the kinematic extent of the HI gas in a galaxy based on its optical properties. For example, using a sample of 409 galaxies from the HI Parkes All-Sky Survey (HIPASS), Meyer et al. (2016) obtain a B-band TFR such that:

$$M(B) = (-8.1 \pm 0.3) \log_{10} V_{rot} + (-2.8 \pm 0.6) \tag{1.4}$$

where $V_{rot} = W^c_{50}/2$, the inclination-corrected line width at 50% peak-intensity, in km/s.

Rotation Curve Shapes

Galaxy rotation-curve shapes and amplitude relate directly to the gravitational field in the mid-plane of its disc (Noordermeer & van der Hulst, 2007). Thus, an investigation into rotation-curve shape is an investigation of the mass distribution of its main components.

A common approach in mass modelling, as taken by (Swaters et al., 2012), is to adopt an axisymmetric, three-component (star, gas, and dark-matter) model in equilibrium.

$$V_c^2 = V_{HI}^2 + V_{star}^2 + V_{halo}^2 \tag{1.5}$$

where V_c is the circular velocity, the contribution of the stars to the rotation curve is given by V_{star}^2, V_{HI}^2 represents the contribution of the gas, and V_{halo}^2 that of the dark halo.

It is well understood that over the bright optical part of spiral galaxy discs, observed rotation curves can usually be explained by the contribution of the stars alone (e.g. Kalnajs (1983); Kent (1986)). The minimised contribution to the rotation-curve by the dark-matter component yields a dominant stellar disk component and has been termed the "maximum disk hypothesis" (e.g., van Albada et al., 1985; Begeman, 1987; Broeils, 1992).

Analysis of the high-resolution HI-maps of galaxies in The HI Nearby Galaxies Survey (THINGS, Walter et al. (2008b)) have been compared with data from the *Spitzer* Infrared Nearby Galaxies Survey (SINGS, Dale et al. (2009)), with the rotation-curves showing correlations with the absolute-magnitude of the galaxy (Broeils, 1992). Luminous galaxies have rotation curves that rise steeply, followed by a decline and an asymptotic approach to the flat outer part of the curve; low-luminosity galaxies show a more gradual increase, never quite reaching the flat part of the curve over the extent of their HI disks.

While not *strictly* a scaling relation, this (apparent) coupling between the shape of rotation curves and the luminous matter works not only for the stellar disc but also for the HI disc. Work by Hoekstra et al. (2001) tested the applicability of a HI-scaling factor (f_{HI}) at larger radii in HI disks. Using a large sample of galaxies presented in Broeils (1992), they found that a scaling factor of around 10 to the HI contribution, in combination with the requirement of a maximal optical disc, explained the rotation curves well. The HI-scaling factor was defined as:

$$f_{HI} = \frac{\Sigma_{stars,0}}{\Sigma_{gas,0}} \frac{h_{stars}}{h_{gas}} \tag{1.6}$$

where $\Sigma_{stars,0}$ and $\Sigma_{gas,0}$ are the central surface mass densities of the stars and the gas, respectively. While h_{stars} denotes the exponential scale length of the stars, and h_{gas} is the gas exponential scale length.

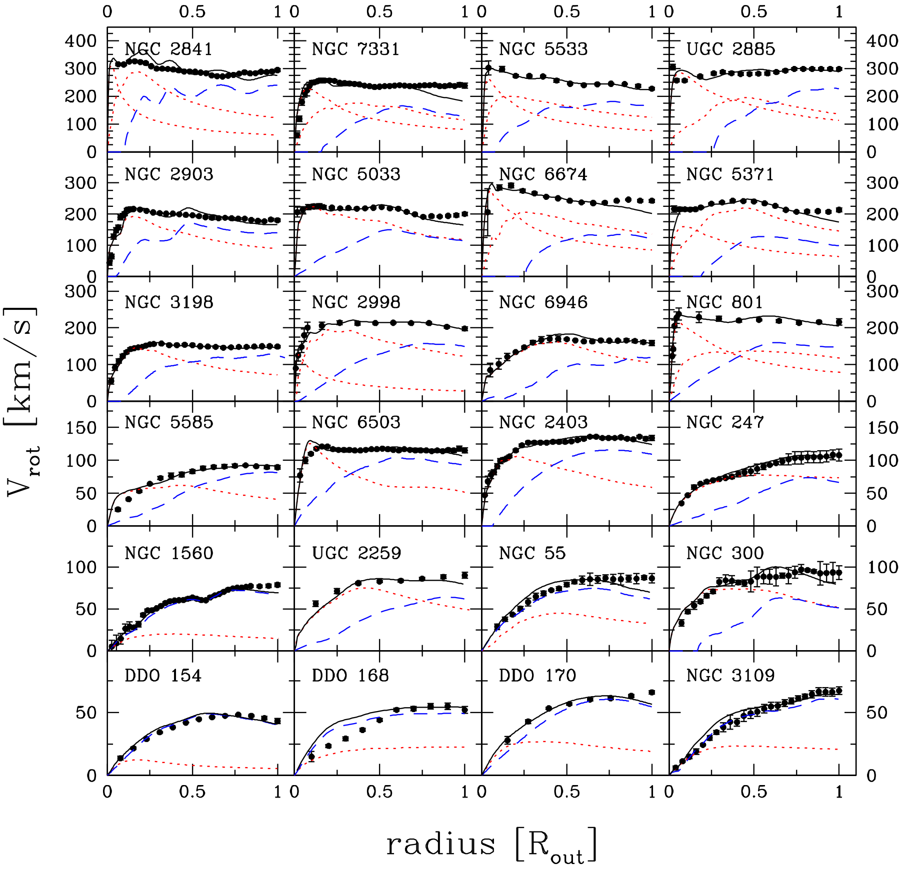


Figure 1.5: Demonstration of HI-scaling factor in predicting rotation-curve shape as presented in Hoekstra et al. (2001). The model fits to these rotation curves are indicated by the black solid lines. The red dotted lines denote the disc- and bulge-components. The blue dashed lines correspond to the scaled HI contribution. The curves are plotted against the radius expressed in intervals of the outermost point of the rotation curve (R_{out}).

Assuming a stellar mass-to-light ratio of unity, the HI scale factor (f_{HI}) can be obtained by fitting the following equation to the observed rotation curve of a galaxy:

$$V_c^2 = (f_{HI} V_{HI})^2 + (\Upsilon_b V_b)^2 + (\Upsilon_d V_d)^2 \tag{1.7}$$

where V_b and V_d are the contribution of the bulge- and stellar-disc components, and V_{HI} is that of HI only. This approach allows for dark matter as well, but it requires the rotation curve of dark matter to be a linear combination of the rotation curves of the visible components.

Furthermore, Swaters et al. (2012) describes the relationship between rotation-curve shape and the baryonic mass (stellar and gas) distribution. In what they call the ‘baryonic scaling model’, Swaters et al. (2012) note the link between the rotation-curve shape and the galactic discs extends to all the baryons in the disc. They find that for baryonic scaling to work, as in HI-scaling, a maximal disc is required - confirming Hoekstra et al. (2001).

1.5 Aims: Three-Dimensional Stacking

The primary objective of this thesis is to create, implement, and evaluate an efficient pipeline for stacking three-dimensional HI spectra, specifically, the stacking of resolved HI distributions in data cubes. The pipeline will be developed using Python 3.8 for greater accessibility while adding to existing radio-astronomy packages.

The testing phase of the pipeline forms the foundation for the subsequent goals of this thesis. By utilizing the Cube-Stacker pipeline, we will analyze the surface-brightness profiles of individual galaxies, stacklings, and the resulting stack. This analysis aims to investigate potential enhancements in the signal-to-noise ratio (SNR) of the data and quantify any signals at larger radii.

We also present measurements derived from processes complementary (see Chapter 3.2) to the Cube Stacker pipeline. These include V_{rot} and h_{rot} from the rotation-curve fitting for the 33 galaxies of the WHISP survey with NSA photometry. To this end, we have developed an envelope-fitting package in Python 3.8 which extracts the rotation curve from a galaxy with a position-angle of 0°, and implements a novel set of routines for extracting envelopes as first proposed by Courteau (1997). See Chapter 3.2.3 for a summary of this envelope fitting routine.

While time constraints have curtailed the investigation of additional scaling laws to better align signal, we present our initial forays toward an optically-based stacking pipeline in Chapter 5.2.1. We present the development of a novel cube-stacking pipeline which makes use of the HI-properties (R_{HI} and V_{rot}) and discuss ways in which these direct measurements of HI distribution can be predicted and then replaced by the scaling relations in Chapter 1.4. Finally, we present and leave the use of our measurements of central concentration in Table 7.3 for further investigation of predicting HI distributions from optical proxies.

This dissertation is structured as follows: Chapter 2 presents our sample selection process, including our stack binning methodology. Chapter 3 provides an overview of the Cube Stacker pipeline methodology, including parallel processes developed to achieve optimal signal alignment. Chapter 4 presents our utilization of surface-brightness profiles to quantify the advantages and improvements of three-dimensional stacking compared to traditional one-dimensional stacking. We then discuss these results, potential pipeline enhancements, and future analysis in Chapter 5. Finally, we present an Appendix which supplements the Chapters 1-5 with tables and figures that provide technical elaboration on discussed topics, specifically those of the rotation-curve measures and photometry measures related to central concentration.

Chapter 2

Sample Selection

2.1 Noordermeer 2005 Sample

As outlined in Chapter 1.1, all galaxies in the WHISP sample with early morphological types between S0 and Sab were grouped into a sub-sample of 68 galaxies (Noordermeer et al., 2005) - intended for use to study the rotation-curves and dark matter content of these early-type spiral galaxies, containing a study of their optical characteristics.

Within this selection criteria, a fair distribution of types is observed – consisting mostly of Sa and Sab galaxies. The under-representation of S0 correlates with their intrinsic low HI-content and is thus a remnant of observational bias. Furthermore, the diameter and HI flux criteria of the WHISP sample lead to a sample of nearby galaxies.

A motivating reason for the down-selection to the Noordermeer et al. (2005) sub-sample of WHISP galaxies was to satisfy the following criteria:

1. Measured global properties (PA,i,HI-radius, W20);
2. Consistent data reduction pipeline;
3. Existing rotation curve studies;
4. Existing Mass concentration studies.

The Noordermeer et al. (2005) sample and associated publication satisfy points 1-3 with a comprehensive study of HI surface density maps, velocity fields, global profiles and radial surface density distributions - for all 68 galaxies. Furthermore, the additional photometry work provided by studies such as Noordermeer & van der Hulst (2007) and Korsaga et al. (2018) provide a sanity check for our photometry. Thus, the Noordermeer et al. (2005) provides a reliable test bed to develop a stacking algorithm due to existing measurements which may be used to test the output of our pipeline.

2.2 WHISP Data Cubes

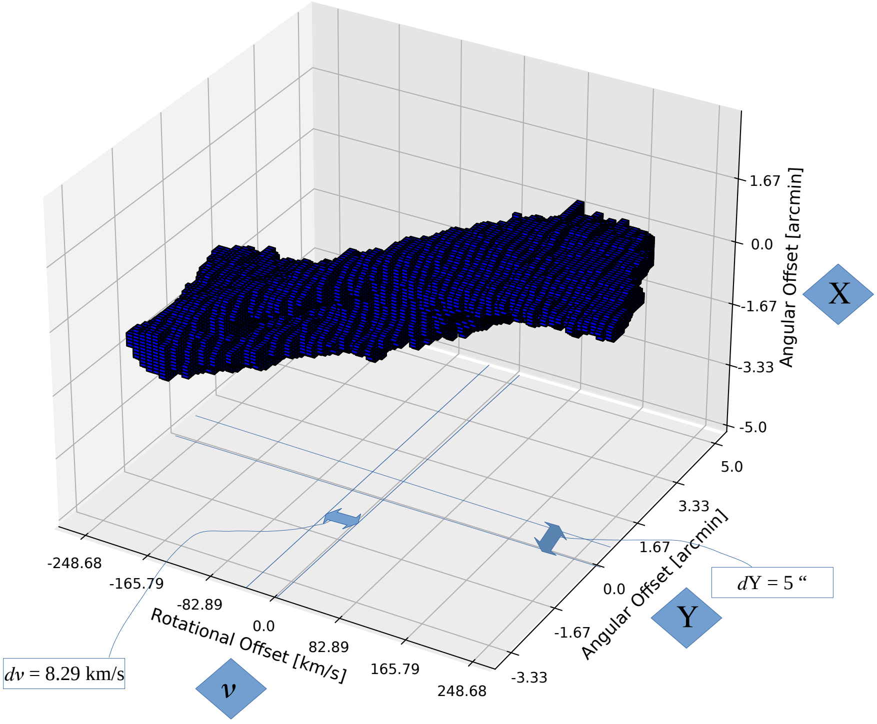


Figure 2.1: Three-dimensional rendering of the WHISP data-cube containing UGC 4605. Annotated are the two spatial-axis (X and Y) as well as the spectral-axis ν. Along the axis, the pixel-separation for WHISP data cubes is shown.

The noise properties of the resulting imaged cube are largely due to the calibration and imaging algorithms applied in the image reconstruction (Arras et al., 2019). Thus, to ensure uniformity, we make use of the full-resolution data cubes which are produced by the standard WHISP pipeline Noordermeer et al. (2005). As a result, and shown in Figure 2.1, the final full-resolution WHISP data cubes have a spatial resolution of $dY = dX = 5$ arcseconds. The WHISP pipeline follows the standard steps of reduction such as flagging, deconvolution, and continuum subtraction.

An important note is that the CLEANing algorithm is cited as Schwarz (1978). This alludes to the Högbom (1974) algorithm which, according to Cornwell, T. J. (2009), iteratively identifies and removes *point sources* in the Fourier inverse image (dirty image). The calculated point spread function (PSF) is used to update the dirty image in this removal process. The CLEANing was performed down to 0.5 times the RMS per channel-map. We discuss possible limitations this CLEANing algorithm imposes on a stacking experiment in Chapter 5.1.2.

2.3 Noordermeer 2005 in the NASA-Sloan Atlas (NSA)

A multitude of parameters of scale were required in order to align galaxy-signal of galaxies since galaxies are naturally observed in a variety of distributions and orientations. Among the scale parameters used are the apparent HI-radius (R_{HI}) for scaling all galaxies to the same apparent size, the width of the HI line profile at 20% intensity (W_{20}) to match spectral-extent, as well as position-angle of the major axis (PA) and galaxy inclination to address orientation and projection-effects, respectively.

However, as outlined in Chapter 1.4 relying on HI scale-parameters to align spectra for stacking requires already-existing radio-observations - creating a restrictive upper limit on resolution and redshift at which this algorithm is applicable. A shift to using scale parameters from wavelengths with greater resolution at greater redshifts (as proxies for HI parameters) is required to overcome these limitations.

To this end, the rest-frame optical window (approximately 300 nm - 1100 nm) was selected due to the availability of high-resolution images, as well as an abundance of studies using these images, particularly from the well-documented Sloan Digital Sky Survey (SDSS) and the surface-photometry available from the NASA-Sloan Atlas (Blanton et al., 2011). For example, measurements such as PA and inclination are readily available from the SDSS and easily incorporated into the algorithm. While optical parameters of spatial distribution (such as PA) are not guaranteed to match that of the HI distribution, we have elected to take this approach following a visual inspection of the moment 0 maps. The PA measures were obtained kinematically by Noordermeer et al. (2005).

Thus, we perform a cross-matching of the (Noordermeer et al. (2005)) galaxies with the NASA-Sloan Atlas (NSA) of galaxies. The NSA is a catalogue of 140,000 images and parameters of local galaxies in the SDSS. We use the galaxies in the "nsa_v0_1_2.fits" catalogue, released as part of the Data Release 8 including all galaxies z<0.05 (Blanton et al., 2011). In Figure 2.2, we demonstrate the location of the Noordermeer et al. (2005) sample of 44 WHISP galaxies present in the footprint of the NSA.

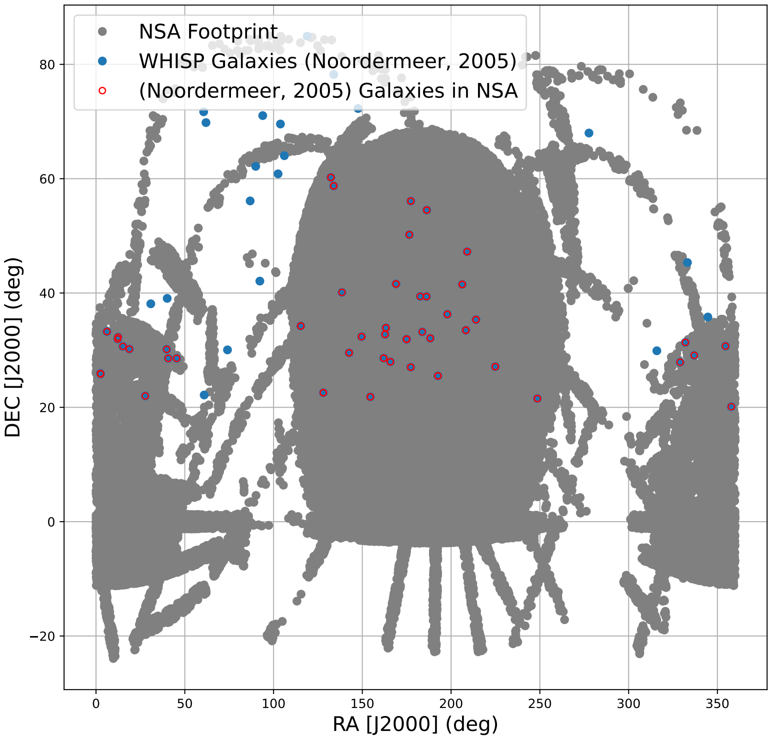


Figure 2.2: Visualisation of (Noordermeer et al., 2005) galaxies in the NASA-Sloan Atlas footprint. The blue dots denote the 68 WHISP targets from the 2005 study, with those falling in the MaNGA footprint receiving an additional red open circle. This results in 44 galaxies observed both by (Noordermeer et al., 2005) in the radio, and the NSA in the optical.

From these lists we obtain the equatorial coordinates from the NASA Extragalactic Database (NED), to ensure each of the targets has RA and DEC coordinates extracted as reliably as possible. We then extracted a list of the WHISP catalogue IDs from the WHISP website, obtaining equatorial coordinates from NED as well. From these tables, now populated with ID, RA, and DEC information, we perform sky-matching tasks in TOPCAT. This returns a combination of desired table matches, including:

1. Radio (Noordermeer et al., 2005) literature covering WHISP Galaxies

2. Optical (NSA) literature covering WHISP Galaxies

3. WHISP galaxies with both radio and optical literature

Source	WHISP Sample	Noordermeer et al. (2005) Sample	NSA Sample
WHISP Sample	409	68	177
Noordermeer et al. (2005) Sample	68	68	44
NSA Sample	177	44	140,000

Table 2.1: Summary of the results obtained from the TOPCAT sky-match process. Each intersecting row-column pair displays the number of galaxies in both literature sources. The intersection between optical and radio literature assumes all galaxies are in the WHISP catalogue.

As a result, our sub-sample consists of 44 galaxies, satisfying the following conditions:

1. Exists in the Noordermeer 2005 Sample
2. Exists in the NSA DR8 Catalogue

2.4 Flagging Warped PVDs

As a final step, we assess and flag the quality of the position-velocity diagrams by eye, thereby eliminating galaxies with disrupted PVD profiles such as counter-rotating inner regions etc. The PVDs are shown in Figure 2.3. Summarised in Table 7.5 is an assembly of data from Noordermeer et al. (2005), containing only the 33 galaxies relevant to this work.

We exclude U499 on the basis that it has a bad inclination from Noordermeer et al. (2005) Table 1, analysis of the Moment 0 shows that it is closer to face-on. The rapid decline of rotational velocity with the radial extent in U3965 means this galaxy is unable to be characterised by a Tanh curve and as such has been excluded from this stacking experiment. The empty data cubes of U5559 and U9133 are the reason for the exclusion of these galaxies, as substitute data cubes could not obtained. The low SNR of the data cubes obtained for U7506, U8805, and U12713 led to difficulty in envelope-fitting. As such, we elect to exclude these galaxies until our envelope-fitting routine is optimised to handle low-quality signals. This may take the form of a smoothing procedure for SNR below a specific threshold to smooth the noise. Finally, the truncated spectra of U10448 qualify its exclusion from this stacking experiment. Finally, the data cubes for U7489, U6621, and U6623 could not be obtained. However, U6621 and U6623 are an interacting pair (Noordermeer et al., 2005) - and are also excluded.

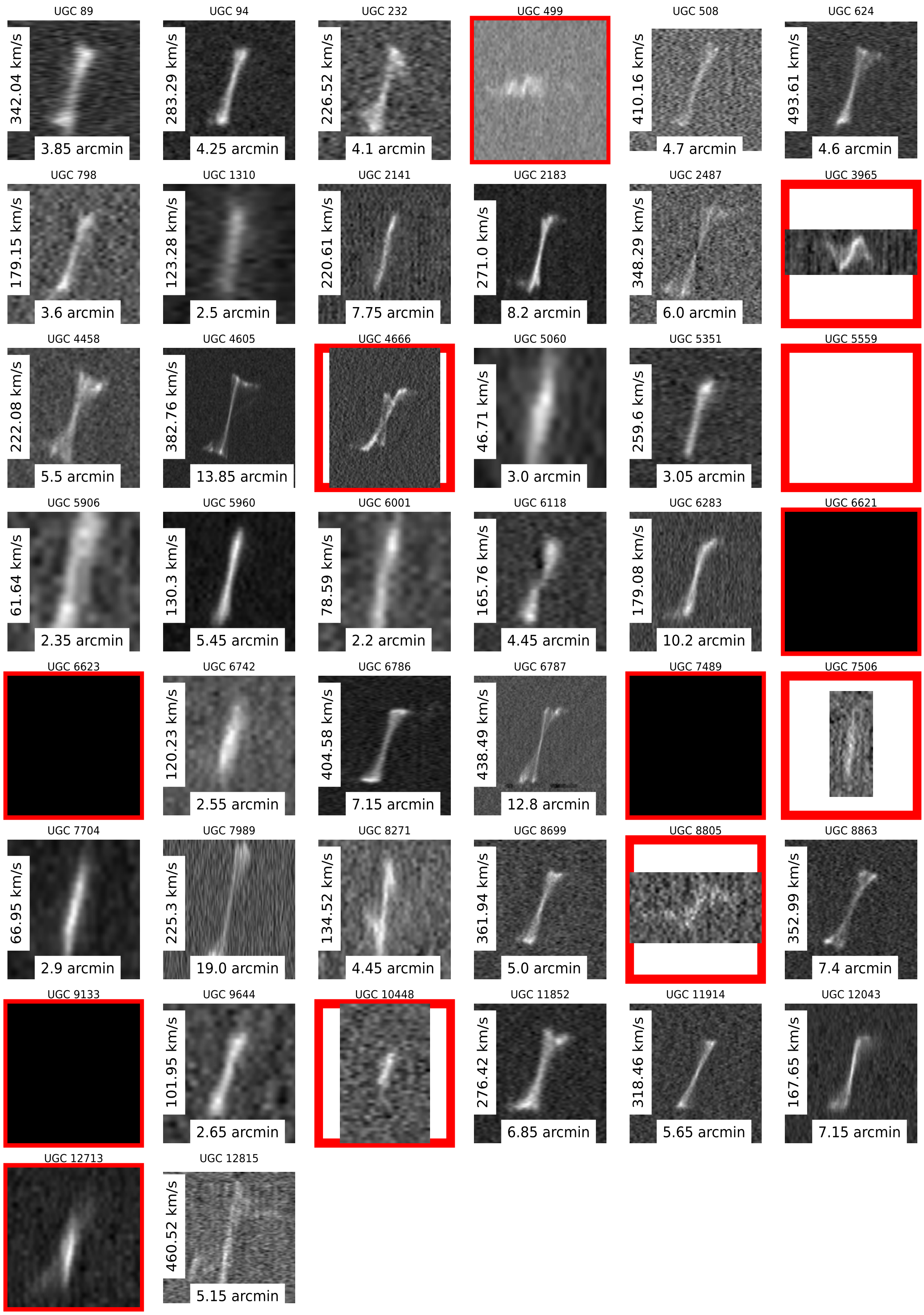


Figure 2.3: Visualisation of the 44 Noordermeer et al. (2005) galaxies with corresponding NSA photometry. The galaxies excluded based on atypical HI distributions are presented with a red border around the PVD. The 4 galaxies for which we have not obtained a data cube, presented as a blank (black or white) subplot, are U5559, U7489, U9133, and the interacting pair U6621 and U6623.

2.5 Stack Sub-Samples

With our sub-sample of 33 *usable* galaxies from the Noordermeer et al. (2005) WHISP sample, with both HI and optical data, we produce a final collection of 5 groups to pass through our cube stacking algorithm.

Binning Galaxies by Mass and Inclination

To construct these 5 groups, we bin by the HI-masses followed by a sub-binning by inclination. The HI-masses and inclinations used are summarised in Table 7.5. By binning by mass, this minimises the amount of spatial scaling required, due to the HI size-mass relation. As discussed in Chapter 1.4, the relation between HI mass (M_{HI} - in units $M_{\odot}$) and the diameter of the HI disc (D_{HI} - in units of kpc) was investigated by Broeils & Rhee (1997) and is described Equation 1.3. For our sample of galaxies, this relation holds well in Figure 2.4. We mark the mass quartiles Q1 and Q3 on the mass-radius relation for our sample of 33 galaxies, subdividing the sample into 3 mass bins containing the lower 25%, the middle 50% and the upper 25% of the sample by mass:

1. Low Mass: $M_{HI} \leq 10^{9} M_{\odot}$

2. Moderate Mass: $10^{9} M_{\odot} < M_{HI} \leq 10^{9.75} M_{\odot}$

3. High Mass: $M_{HI} > 10^{9.75} M_{\odot}$

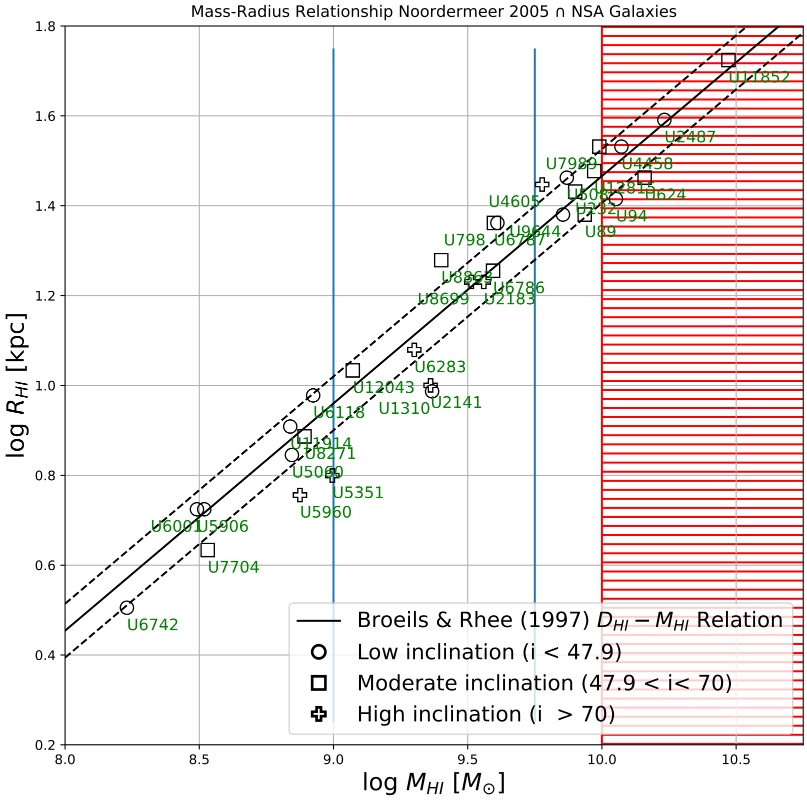


Figure 2.4: HI Size-Mass Relation for the 33 galaxies in the WHISP sample used in our stacking analysis. Three mass bins, as defined in the text, are delineated by vertical lines. Three inclination bins are demarcated with differing symbols as outlined in the legend. The shaded red region denotes the mass ranges accidentally excluded from our sample, resulting in the loss of 5 galaxies. The solid line represents the mass-radius relation as obtained by (Wang et al., 2016), with an rms scatter around the relation of 0.06 dex (14 per cent) denoted by the dashed line.

Within each mass bin we apply a second criterion of galaxy inclination. The inclination ranges were selected at intervals defined by creating equal bins of b/a. Thus, for $b/a = cos(i)$, and taking equally spaced intervals of b/a at cos(i) =[1, 0.67, 0.33, 0] corresponding to inclination bins of:

1. Low inclination: $i < 47.9°$
2. Low inclination: $47.9° \leq i < 70°$
3. High inclination: $i \geq 70°$

The inclination values were obtained optically where estimates from tilted-ring modelling of velocity fields were not readily available.

Finally, we ensure stack groups of at least 5 galaxies, obtaining a final set of stack samples:

Group 1	Group 3	Group 6	Group 9	Group 10
Low Mass Low inclination (i≤47.9)	Low Mass Low-Moderate inclination (i<70)	Moderate Mass Low-Moderate inclination (i<70)	High Mass Low-Moderate inclination (i<70)	All Mass High inclination (i≥70)
UGC 5060	UGC 5060	UGC 798	UGC 89	UGC 2141
UGC 5906	UGC 5906	UGC 1310	UGC 232	UGC 4605
UGC 6118	UGC 11914	UGC 6786	UGC 508	UGC 5351
UGC 6742	UGC 6118	UGC 6787	*UGC 624	UGC 5960
UGC 11914	UGC 6742	UGC 8863	*UGC 2487	UGC 6283
-	UGC 7704	UGC 12043	*UGC 4458	UGC 8699
-	UGC 8271	-	UGC 7989	-
-	-	-	UGC 9644	-
-	-	-	*UGC 11852	-
-	-	-	UGC 12815	-

Table 2.2: Summary of the stack sub-samples, binned by mass and (or) inclination. The galaxies accidentally excluded from the high-mass bins are marked with an asterisk, and are not present in the final stacks.

Finally, we define a tenth group containing all highly-inclined systems ($i \geq 70°$). As mentioned in Chapter 1.2, HI studies are prone to projection effects - especially at high inclination. Thus, we construct Group 10 as the stack of all highly inclined galaxies (across the full range of mass for our 33 galaxies) and alter the fiducial inclination to be $i_{fiducial} = 90°$.

A vestigial bug from previous selection iterations has led to the exclusion of the highest-mass galaxies U94, U624, U2487, U4458, and U11852. The addition of these galaxies would have led to the addition of two high-mass stack groups; those high-mass galaxies with low inclination (Group 7) and those high-mass galaxies with moderate inclinations (Group 8). Additionally, Group 9 will have benefited from the addition four of the five galaxies, becoming a stack of 10 galaxies. Due to time constraints, we were unable to recompile the stacks, however, we present the stack groups rejected based on having fewer than 5 galaxies in total in Table 2.3 for completeness.

Group 2	Group 4	Group 5	Group 7	Group 8
Low Mass Moderate inclination (47.9<i<70)	Moderate Mass Low inclination (i≤47.9)	Moderate Mass Moderate inclination (47.9<i<70)	High Mass Low inclination (i≤47.9)	High Mass Moderate inclination (47.9<i<70)
UGC 7704	UGC 798	UGC 6786	*UGC 94	UGC 89
UGC 8271	UGC 1310	UGC 6787	UGC 508	UGC 232
-	-	UGC 8863	*UGC 2487	*UGC 624
-	-	UGC 12043	*UGC 4458	UGC 7989
-	-	-	UGC 9644	*UGC 11852
-	-	-	-	UGC 12815

Table 2.3: Summary of the stack sub-samples, binned by mass and inclination, rejected due to an insufficient bin size. The galaxies accidentally excluded from the high-mass bins are marked with an asterisk.

Chapter 3

Cube Stacker

Unlike in the works of Chen et al. (2021) we focus our stacks and analysis on already imaged and cleaned data cubes, foregoing the benefits of signal enhancement of stacking in the UV-plane since the WHISP pipeline already performs CLEANing down to 0.5 times the RMS.

Our approach takes the familiar approach of *shift-and-scale* performed when standardising one-dimensional profiles, and extends the process to three dimensions. As such we make use of several mappings in order to carefully align spectra. In this chapter, we briefly discuss the properties of the data cubes which form the bedrock of our experiment. Then we discuss the two maps we use to align and assess our stacking. These are the position-velocity diagrams and zeroth-moment maps, respectively.

3.1 Cube Stacker Pipeline

3.1.1 3D Stacking Concepts

Introduced in Chapter 1.3, conventional stacking experiments performed on one-dimensional spectra often take a list of 1D spectra and associated redshifts for the sample of galaxies to be stacked as input aligns the signal by spectrally shifting and scaling the HI spectra and finally applying a weighting scheme before co-adding the spectra. The 1D spectra are usually obtained by selecting, within the containing data cube, a spatial region over which to integrate (by summing or averaging) the flux. This process aims to align as much of the 1D spectra as possible, registering the spectra across a suitable spectral extent. Aligning spectra in 3D has the same objective.

Due to the rotations performed on the data cubes, see the detailed discussion in Chapter 3.5, the conventional equatorial coordinates are an inefficient descriptor for the galaxy distribution within the data cube. In Figure 3.1 we present the two coordinate systems used to describe the data cubes before and after the cube rotation. All positions prior to the cube rotation are with respect to the conventional equatorial coordinates (α,δ,ν) with α the right-ascension (RA) along the horizontal axis, δ the declination (DEC) on the vertical axis, and spectral axis ν orthogonal to the aforementioned spatial axes. Alternatively, all coordinates following the cube rotation are replaced by (X,Y,Z) or primed equatorial coordinates (α',δ',ν'), respectively.

The galaxy registration process can be summarised into two parts, namely, spatial registration and spectral registration. In the spatial registration, spatial centring of position on the sky (α_0,δ_0) is required. Here, α_0 is the RA of the galaxy and δ_0 is the DEC of the galaxy as listed in Table 7.5.

Next, a spectral registration is required. We adopt the systemic-velocity (V_{sys}) in km/s as recorded in Table 2 in Noordermeer et al. (2005), and summarised in Table 7.5, to find the redshifted the 21cm line. This results in an initial spectral extraction point (ν_0) of:

$$\nu_0 = \nu_{21cm}\left(1 - \frac{V^{sys}}{c}\right) \tag{3.1}$$

where ν_{21cm} is the rest-frequency of atomic hydrogen spin-flip transition and c is the speed of light in a vacuum.

Once the initial galaxy extraction is performed, standardisation of the galaxy position-angle (PA), galaxy inclination (i), and axial ratio (b/a) follows. In so doing, the distribution of the galaxy major axis and minor axis relative to this initial extraction point (α_0,δ_0,ν_0) is standardised. As a result of the rotation step we apply a second galaxy registration step to standardise the barycentre within the rotated coordinate system.

In this iteration of the *Cube Stacker* pipeline, we further constrain the declination (δ'_0) and frequency (ν'_0) of the rotated data cube - adopting the spatial centre from a Tanh model fit to centroids extracted from the galaxy position-velocity diagram (PVD). See Figure 3.2 for a demonstration of the centroid extraction.

Since the rotation is not *exactly* through the initial spatial extraction point (α_0, δ_0), a secondary constraint of the *rotated* right-ascension ($X = \alpha'_0$) is also required. However, since we adopt a two-dimensional model to align the spectra within the cubes, this alignment is beyond the scope of this thesis.

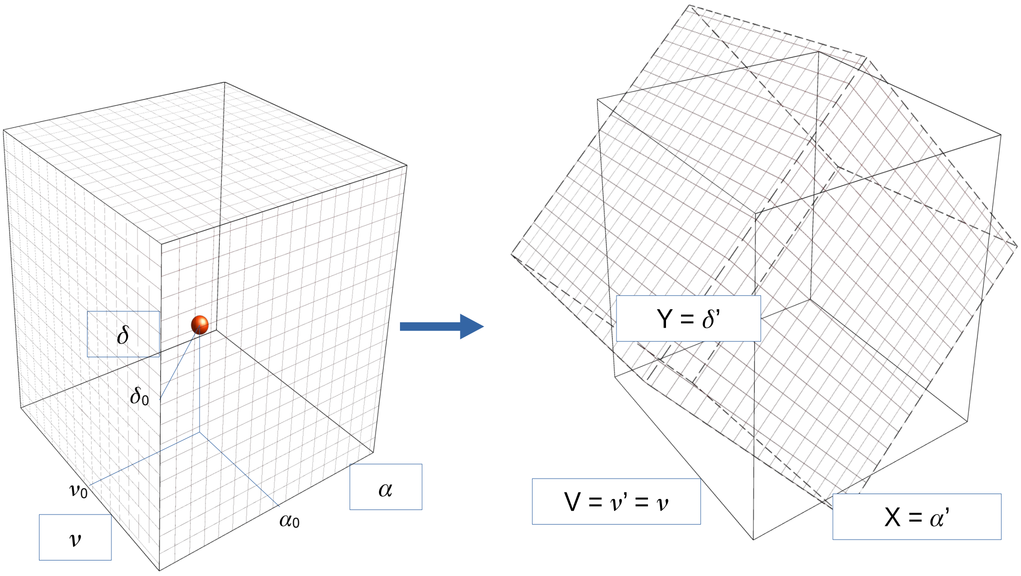


Figure 3.1: Visualisation of the coordinate frames used to describe the HI distribution before and after the rotation stage as outlined in Chapter 3.1.3.

Individual galaxies, however, are intrinsically sized differently. As a result, we apply several scaling procedures to standardise the spatial and spectral extents of the galaxies. For the spatial extent, we scale all galaxies to the largest HI radius (R_{HI}) in the stack group, as well as reproject all galaxies to the same inclination. See our chapter on spatial scaling for a detailed description. To standardise the spectral extent, we make use of the velocity amplitude of the Tanh model V_{rot} as a fiducial, this time fitting to a set of envelopes extracted from the PVD. See Figure 3.2 for a demonstration of the envelope and rotation curve fitting. In our application V_{rot} is the projected rotation amplitude. See Chapter 3.1.4.

The above approach, termed the *shift-and-scale* approach, ensures the galaxies occupy the same number of voxels, while centred on a common voxel in a like-sized array.

As a result of these cube transformations and scalings, both the HI distribution as well as the relative data cube fidelity are altered. By data cube fidelity we are referring to the spatial and spectral resolutions of the data. In the chapter titled Cube Re-convolution we discuss the measures the *Cube Stacker* algorithm accounts for in order to standardise the radio beam. Finally, following the successful alignment of HI spectra within the data cube, we provide a summary for various methods of cube co-addition.

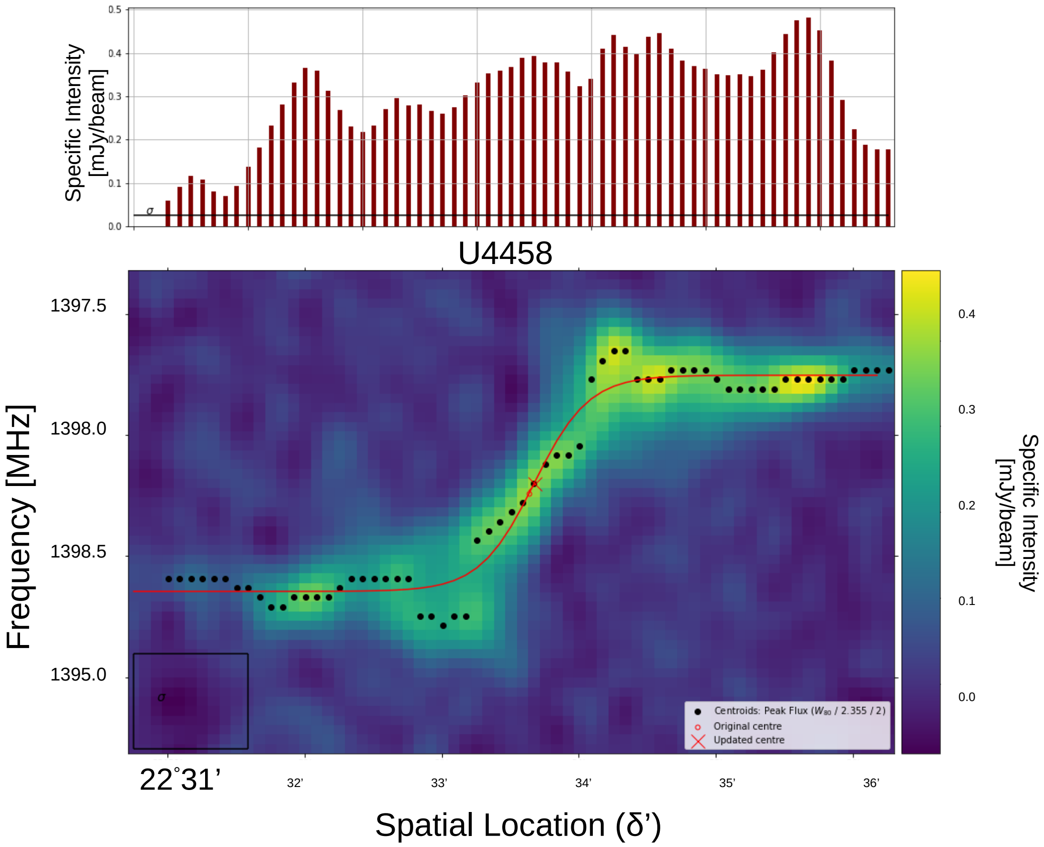


Figure 3.2: Graphical demonstration of the peak-intensity mapping performed for PVD of UGC 4458. The centroids obtained after spectral smoothing are fit by a Tanh model rotation curve and an updated centre for the spectrum is obtained.

3.1.2 Overview

We have developed a two-part algorithm for stacking three-dimensional HI data cubes. Developed in *Python 3.8*, our algorithm takes in a set of reduced and 'CLEAN'ed *FITS* spectral cubes and passes them through, broadly, two processes when producing a stack. The overall goal of this pipeline is to Comprehensively Unearth Baryons in Extended-regions using a Stacker; we dub this pipeline *Cube Stacker*.

As summarised in Figure 3.5 the two processes are:

1. Data Preparation
2. Stacking

We adopt this dual-step approach to account for the variety of conventions used across various HI surveys. With this in mind, a "Data Preparation" step performs a two-part iteration

Header Key	Description
'EPOCH'	Appears as either J2000.0 or B1950.0 in the WHISP data. We chose J2000.0
'CRVAL1'	The reference RA coordinate was converted to J2000.0 (if in B1950.0)
'CRVAL2'	The reference DEC coordinate was converted to J2000.0 (if in B1950.0)
'CUNIT3'	Standardized to 'Hz' (from 'HZ' or 'kms^{-1})
'BUNIT'	Standardized to mJy/beam (from W.U = 5 mJy/beam or 'JY/BEAM')
'FREQ'	Appears as "FREQ-OHEL' - incompatible with the spectral-cube package.
'FREQ0'	Appears as 'FREQR' or 'RESTFREQ', or 'FREQ0'.
'BMIN'	Appears as 'BMMIN' - incompatible with the spectral-cube package
'BMAJ'	Appears as 'BMMAJ' - incompatible with the spectral-cube package
'BPA'	Appears as 'BMPA' - incompatible with the spectral-cube package
'PA'	New addition. The galaxy's position-angle (PA)
'I'	New addition. The galaxy's inclination (I)
'RHIARC'	New addition. The galaxy's HI-radius in arcseconds
'W20'	New addition. The galaxy's line width at 20% peak intensity.
'POINTN'	New addition. The number of sources in the cube.
'SEPRTN'	New addition. Distance (Separation) between sources in the cube.

Table 3.1: Summary of *FITS* keys amended and appended to each data-cube. This step highlights the need for a standardized set of keys accessible by the pipeline to ensure efficient processing of multiple cubes at run-time.

when standardising all the cubes of interest prior to stacking.

3.1.3 Data Preparation

Our algorithm relies on an input table of properties for the galaxies intended for stacking. The currently supported format for the table is *.csv*. The properties listed in this table are used to align the 3D galaxy spectra within the respective data cubes, producing *stacklings*. Thus, the algorithm's main purpose is to align the spectra within the cubes prior to co-addition. To this end, the required properties are stored in the *FITS* header of the cube, streamlining the stacking process.

In addition to the varying header conventions, WHISP data cubes sometimes have multiple galaxies (or sources) in the pointing. We ensure that every data cube included in a stack contains only a single source, and as such the "Pass 1 Extraction" script, summarised in Figure 3.3, extracts the largest possible cube centred on an initial estimate for the position of the HI spectrum in the cube. We use the coordinates from (Noordermeer et al., 2005) for the RA, Dec, and redshifted HI rest-frequency ν_0 in Equation 3.1 ($\alpha_0, \delta_0, \nu_0$) to perform the initial extraction from Table 7.5.

In particular, the Noordermeer et al. (2005) full-resolution data-cube headers do not maintain conventions for the epoch of observation ('EPOCH'), flux-unit ('BUNIT'), and frequency-unit ('CUNIT3'). In Table 3.1 we summarise a subset of the most important header keys we update and/or standardise in the WHISP *FITS* data cubes.

To achieve a uniform epoch for our coordinate system across all data cubes, we make use of the *astropy.coords.SkyCoord* package with associated function *.transform-to()*. Here we pass the 'CRVAL1' and 'CRVAL2' and their associated epoch of observation 'EPOCH'. After specifying our desired epoch of observation, namely J2000.0, we receive an updated reference value for the reference pixel from the *.transform-to()* function.

Next, WHISP datacubes are presented in flux-units of 'W.U', where $1W.U. = 5mJy/beam$. In order to process these cubes in *SpectralCube*, we scale the flux within each datacube by a factor of 5, while reassigning 'BUNIT' to "$mJy/beam$".

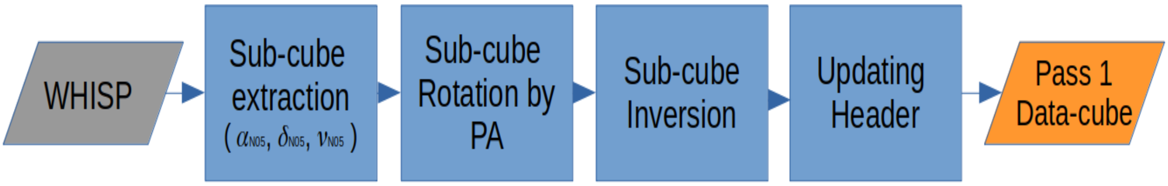


Figure 3.3: Summary of the 'Pass 1 Extraction' routine of the Data Preparation stage of the Cube-Stacker algorithm. This stage produces a set of cubelets containing a single galaxy, rotated and spectrally inverted in order to standardise the HI distribution with respect to the cubelet centre.

Next, and still part of the "Pass 1 Extraction" process, we ensure that all galaxies in their respective cubes have the same position angle (PA). We rely on the kinematic PA determined visually by Noordermeer et al. (2005) and perform a counter-rotation by the PA to set all galaxies to a fiducial $PA = 0°$. We confirm the reliability of the PAs by visual of the velocity fields produced by (Noordermeer et al., 2005). To perform the rotation we use the *scipy* function *ndimage.rotate*. We choose a spline interpolation of order $k = 1$ denoting a linear interpolation. The *FITS* header is used once more to record the position-angle of the rotated galaxy ('PA') and, subsequently, the beam ('BPA').

See Chapter 3.1.5 for a discussion of flux conservation.

An important note here is the cube world coordinates no longer align with the standard equatorial coordinates. Although the pixel separation remains the same, we re-define the coordinate system to the rotated coordinate system, using the mapping:

$$R\begin{pmatrix}\hat{\alpha}\\ \hat{\delta}\end{pmatrix} \longrightarrow \begin{pmatrix}\hat{\alpha}'\\ \hat{\delta}'\end{pmatrix} \tag{3.2}$$

A further technicality emerges; since the conversion between pixel- and world-coordinate systems relies on a known coordinate of a reference pixel (as well as a known pixel separation along each axis), it was found that the reference pixels defined in the WHISP headers (α,δ)

were often located non-centrally in the central pixel. Since the centre of the central pixel is the point-of-rotation in the *ndimage.rotate* function, the reference pixel will also have undergone a rotation.

We thus make use of a rotation matrix centred on the central pixel to find the coordinates of the new reference pixels. Setting $R = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, and rearranging Equation 3.3 we find:

$$\begin{pmatrix} \alpha'_0 \\ \delta'_0 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \delta_0 \end{pmatrix} \tag{3.3}$$

In the *FITS* convention, $\alpha_0 = CRPIX1$, and $\delta_0 = CRPIX2$.

Next, we account for the varying directions of rotation galaxies would have with respect to the line-of-sight. We opt to standardise this by having a receding major-axis (MA) at rotated North ($\hat{\delta}'_N$) on the sky. By inverting the order of spectral channels within the sub-cubes, we achieve a spectral-inversion of the data-cube, thus, standardising the direction of rotation of the galaxies. This is denoted by the "Sub-cube Inversion" step.

The final step, as summarised in Figure 3.4, in the "Pass 1 Extraction" process involves updating the *FITS* header for the updated reference pixels.

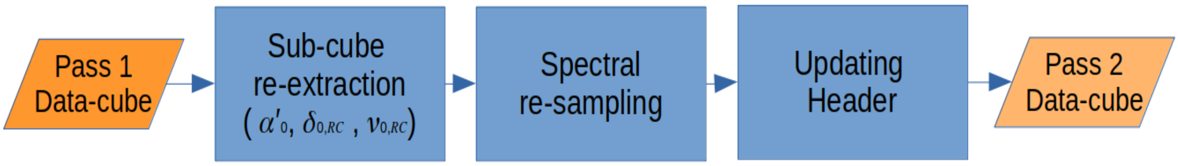


Figure 3.4: Summary of the ’Pass 2 Extraction’ routine of the Data Preparation stage of the Cube-Stacker algorithm. All cubelets now have fixed spectral- and spatial-binning, centralised galaxies with standardised orientations, and header structure.

Following the velocity-tracing and rotation-curve fitting (see Chapters 3.2.3 and 3.2.4), updated centre coordinates are found along the $\hat{\delta}'$, and $\hat{v}'$ axes. This is referred to as the "RC-fitting" process in Figure 3.5 which runs parallel to the Cube-Stacker pipeline. These values ($\delta_{0,RC}$, $v_{0,RC}$) are obtained from the rotation-curve centre, and a secondary extraction is performed using the updated coordinates (α'_0, $\delta_{0,RC}$, $v_{0,RC}$).

It was noted that the spectral sampling was inconsistent across the WHISP data cubes. This would affect the ’Stacking’ stage (see Figure 3.5) of the pipeline whereby the spectral extents need to be matched by means of scaling. In order for the HI spectra to maintain the same physical scale, the channel separations (CDELT3, in *FITS* standard) need to be equal. Following the "sub-cube re-extraction" process, we scale the spectral axes across all cubes to a fiducial spectral sampling ('CDELT3') of -39062.50 Hz. This was achieved by taking the ratio of the existing spectral sampling to fiducial and scaling the spectral axis accordingly. See Table 7.4 for a summary table of scaling factors. We perform this scaling using the Python function *skimage.transform.rescale*.

See Chapter 3.1.5 for a discussion of flux conservation.

Finally, the *FITS* header is appended, with the RC coordinates now the new reference coordinates along the $\hat{\delta_0}'$ and $\hat{\nu_0}'$ axes ($CRPIX2 = \delta_{0,RC}$, and $CRPIX3 = \nu_{0,RC}$), as well as the spectral-sampling key ('CDELT3').

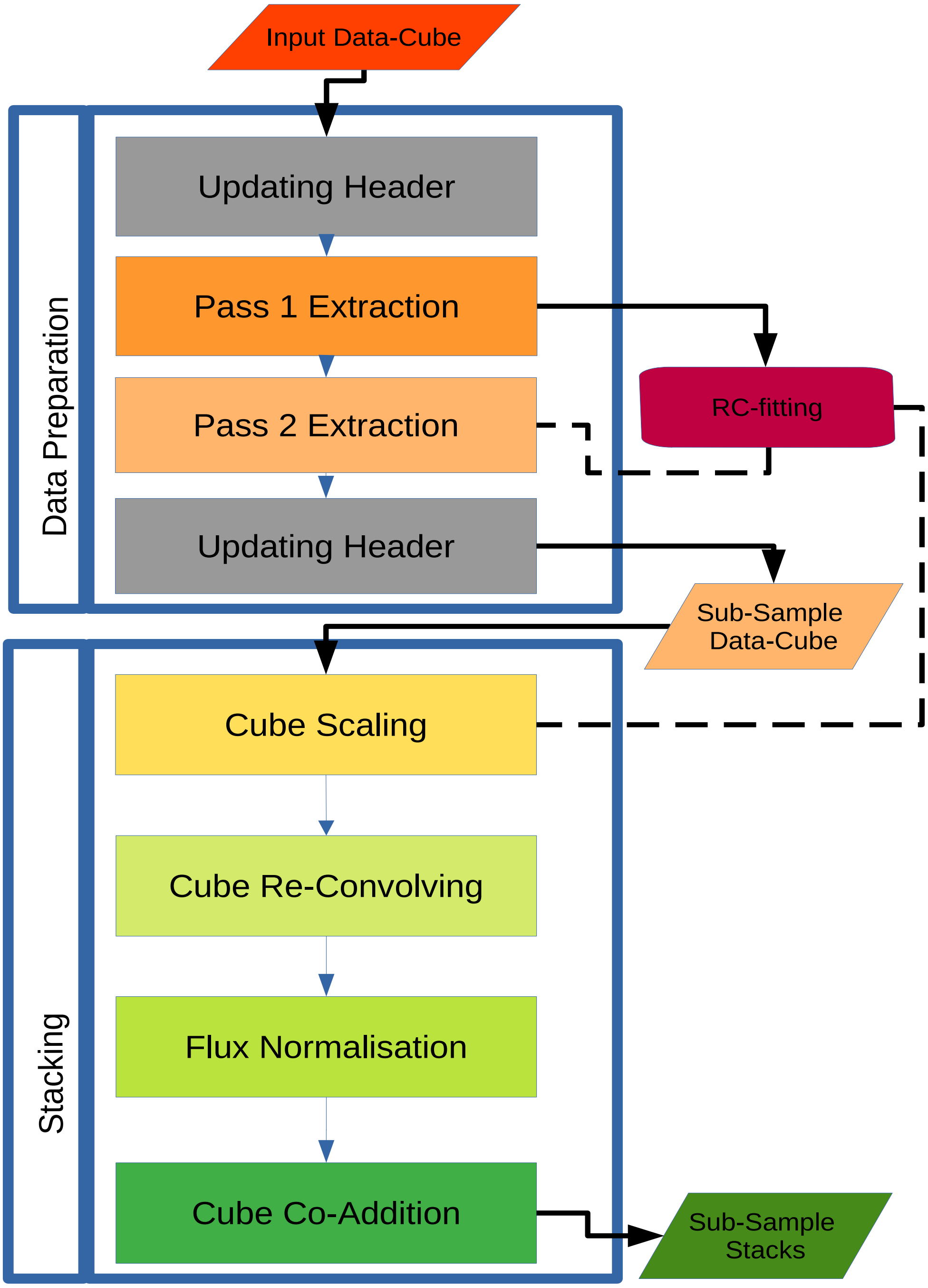


Figure 3.5: This flow diagram shows the general overview of the two-part process for generating a stack of HI data cubes. All raw data cubes (here the WHISP data products) are passed through a series of processes whereby the data format is standardised, namely the "Data Preparation" stage. Then, following a sub-sample selection procedure, individually standardised data cubes are then grouped and passed through the 'Stacking' stage. Supplementary processes (such as "RC-fitting") are shown alongside the main pipeline enclosed in the blue borders.

3.1.4 Stacking

With all the WHISP data cubes and respective galaxies standardised, stacking may thus begin.

Several sub-samples were defined (see Chapter 2.5), with each sub-sample passed through the 'Stacking' stage of the pipeline separately. We divide the 'Stacking' stage into four sub-processes, namely: Cube-Scaling, Cube-Regridding, Cube-Reconvolving, and Cube-Coaddition. See Figure 3.5.

The 'Cube-Scaling' section forms the basis of ensuring each galaxy in the stack (referred to as 'stackling' hereafter) occupies the same number of pixels within the data cube. To ensure this, we have opted to match the two spatial and one spectral extent of the stacklings on a per-stack basis. We take the largest stackling per axis in the stack and upscale all the other stacklings to have the same dimension along that axis. The 'Cube Scaling' process is summarised in Figure 3.6.

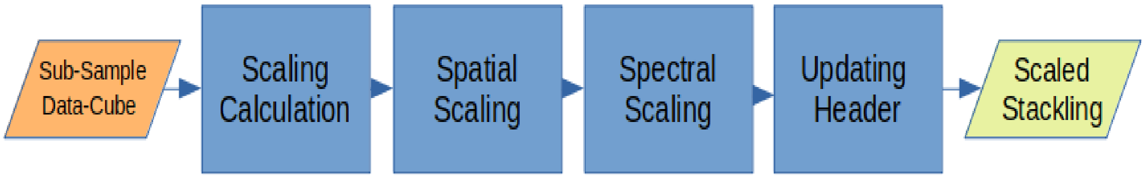


Figure 3.6: Graphical overview of the 'Cube Scaling' process, of the 'Stacking' stage of the Cube-Stacker pipeline. Here HI signal is aligned by standardising spatial- and spectral-extent, as well as inclination.

We use the HI-radius in conjunction with inclination-induced projection effects to standardise spatial extent. In the rotated frame, we adopt a simplifying label convention to track the scaling required along each axis, namely:

$$\begin{pmatrix} \hat{\alpha}' \\ \hat{\delta}' \\ \hat{\nu}' \end{pmatrix} \longrightarrow \begin{pmatrix} \hat{X} \\ \hat{Y} \\ \hat{V} \end{pmatrix} \tag{3.4}$$

Spatial Scaling

Considering the spatial plane $(\hat{X}, \hat{Y})$, as visualised in Figure 3.1, two distinct scalings are required. Both axes share a scaling to a standard HI-radius (R_{HI}). Scaling all stacklings to the largest HI-radius in the stack ($a_{max,sample}$), the scaling required to match spatial extent is:

$$a_S = \frac{a_{max,sample}}{a_{stackling}} \tag{3.5}$$

However, an extra scaling is required along the $\hat{X}$ axis to correct for the effects of inclination-induced projection effects. To do this, we set all inclinations to a fiducial of $i = 60°$. We use the axial-ratio of the semi-minor (b) and semi-major (a) axis to achieve this, taking the functional form:

$$(\frac{b}{a})^2 = (1 - q_0^2)\cos^2(i) + q_0^2 \tag{3.6}$$

with $q_0 = 0.2$ and $a = R_{HI}$, the inclination-correction factor (i_s):

$$i_S = \frac{(\frac{b}{a})_{i=60°}}{(\frac{b}{a})_{i,stackling}} \tag{3.7}$$

Then, the final scaling along the $\hat{X}$ and $\hat{Y}$ axes is:

$$\begin{pmatrix} S_x \\ S_y \end{pmatrix} = \begin{pmatrix} a_S \cdot i_S \\ a_S \end{pmatrix} \tag{3.8}$$

Spectral Scaling

Spectrally, we correct all stacklings for the amplitude of the rotational velocity (obtained from the rotation-curve fitting process in Chapter 3.2.4). Again, matching for the largest rotational amplitude across all stacklings in the stack, the scaling each stackling undergoes along the spectral axis $\hat{V}$ is:

$$v_S = \frac{V^{amp}_{max,sample}}{V^{amp}_{stackling}} \tag{3.9}$$

leading to a final scaling per axis per stackling of:

$$\begin{pmatrix} S_x \\ S_y \\ S_v \end{pmatrix} = \begin{pmatrix} a_S \cdot i_S \\ a_S \\ v_S \end{pmatrix} \tag{3.10}$$

Cube Re-convolution

As discussed in Chapter 7.1, the various transformations applied to the data in the data cubes also apply to the beam resolving this data. Thus, the standard WHISP beam is modulated by the varying values required to bring all the galaxies (or stacklings, following the scaling operations) to a standardised apparent radius (R_{HI}).

To sensibly stack these stacklings, their restoring beams need to be identical; it is to this end that we make use of the scipy.signal.convolve2d function in order to smooth the data to a common resolution.

While smoothing functions do exist in the astropy.convolution and spectral_cube packages, for our purposes we needed our function to fulfil the following requirements:

1. Convolves TO a required resolution, spatially, and

2. Conserves Total Flux

We have developed a function for circularising the beam of a radio observation to a target, circular, beam. This function is called standardise_beam(), following Pythonic naming conventions. Taking a two-step approach, standardise_beam() first convolves the data with a two-dimensional Gaussian kernel, normalised to an amplitude A = 1 in order to conserve total-flux, with axes identical to those of the stackling's beam except for the position-angle - which is made to be orthogonal. The process flow is summarised in Figure 3.7.

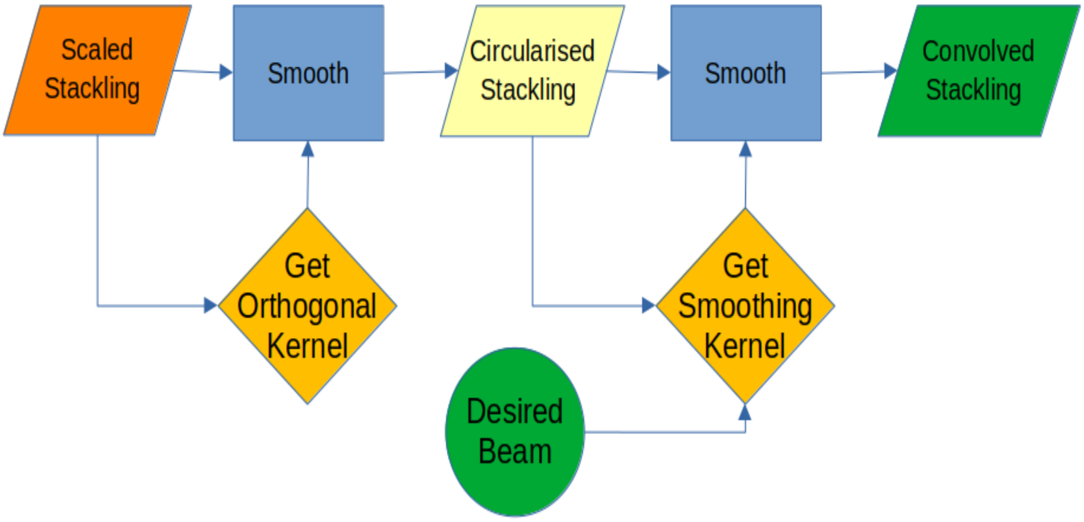


Figure 3.7: The flow Process of the standardise_beam() function developed for the convolution of a beam to a desired beam. This function ensures convolution is performed in a flux-conserving manner.

The outcome of this convolution is a circular beam with major and minor axes equal to the major axis of the stackling's beam.

With this new circular beam, calculating the required kernel to convolve *to* a new (and larger) circular beam is simple. We use the following relation to calculate the required *circular* smoothing kernel's dimensions:

$$\sigma_{kernel,circular} = \sqrt{\sigma^2_{target,major} - \sigma^2_{current,major}} \tag{3.11}$$

We define the desired beam as the smallest circular beam which contains all the beams of the stacklings in the stack - thus changes across the stack groups defined in Table 2.2.

In addition to tracking the effects of scaling and rotation on the beams in our sample, before we co-add our data cubes, we standardise the beams through a convolution process. This process is outlined in Chapter 3.1.4.

In this section we demonstrate the performance of our *standardise_beam function*, convolving a two-dimensional elliptical-gaussian distribution of fluxes.

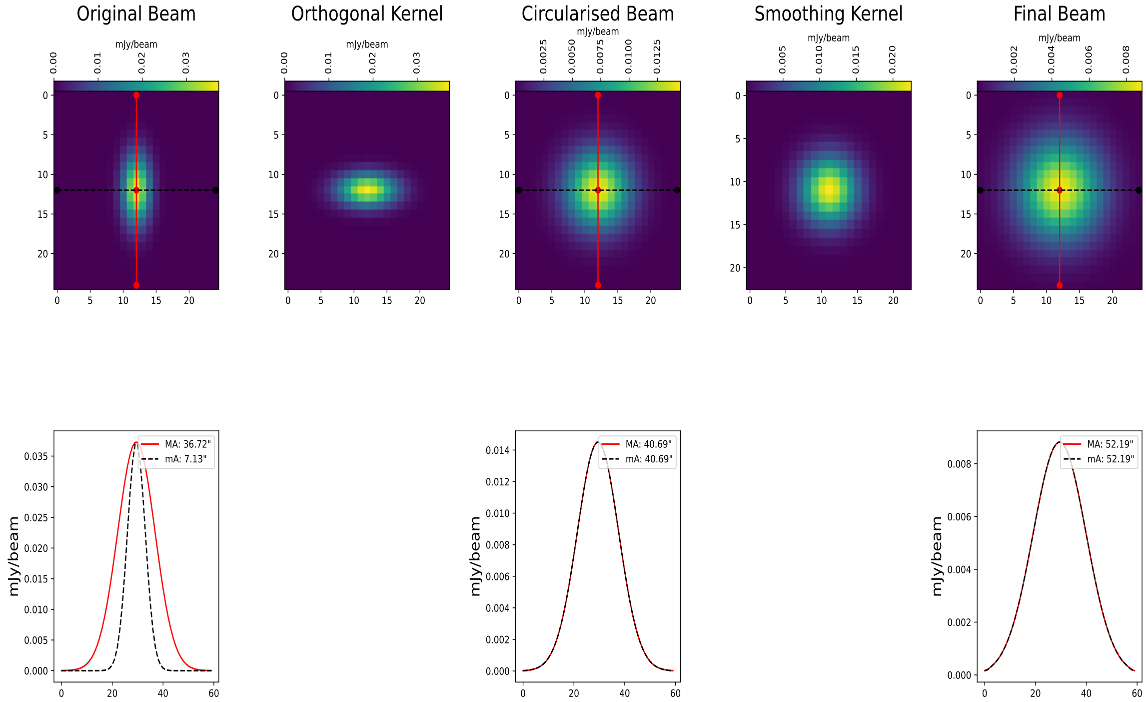


Figure 3.8: A sequenced demonstration of an elliptical-gaussian distribution, simulating the shape of a WHISP beam, undergoing a convolution under the *standardise_beam* function developed for the *Cube Stacker* algorithm.

In Figure 3.8, we start with a 36.72” × 7.13” beam, with the aim of convolving it **to** a beam with a major-axis of 50". and convolve it with a kernel of equal size, rotated 90°. We refer to this initial kernel as the "Orthogonal Kernel". We compute the product of the convolution of these two Gaussians using the astropy.convolution.convolve function, and using a normalised kernel to ensure flux conservation. This produces a circularised beam.

Finally, we compute the required kernel dimensions for convolving the current (circular) beam to our larger, desired beam. This calculation is contained in Equation 3.11. Following this transformation, we obtain a beam size of 52.19” × 52.19”.

Cube Co-Addition

With the WHISP data cubes now aligned and scaled in 3D, as well as standardised to a common beam, we proceed with the process of stacking the data.

Due to the varying noise properties of the WHISP data cubes, we calculate the weighted average ($\bar{S}_{stack}$) on the stacklings (S_i) in the stack groups. Expressed mathematically:

$$\bar{S}_{stack} = \frac{\sum_{i=1}^{n} w_i S_i}{\sum_{i=1}^{n} w_i} \tag{3.12}$$

where we adopt an inverse-variance weighting (w_i) of:

$$w_i = \frac{1}{\sigma_i^2}. \tag{3.13}$$

In the above equations, i is the index of the data cube to be included in the stack, and N is the total number of stacklings in the sample. We estimate the noise (σ_i) by sampling a flux-free region within the data cube (designated at the four corners spatially) and assessing the standard deviation of the flux across the spectral axis. We then take the average of the σ across the four regions. All voxels in cube i (S_i) receive a weighting of w_i before co-addition.

To compare the utility of applying an inverse-variance weighting we provide a comparison to an equal-weighting $w_i = 1$ approach in Chapter 4.7, specifically demonstrating the improved signal-to-noise ratio (SNR) achieved by applying a weighting scheme in Figure 4.9.

3.1.5 Flux Conservation Considerations

Due to the stacking of different sources, we elect to bin by mass and inclination to minimise the effects of the spatial and spectral scaling on the flux distribution within the data cubes. Although we ensure flux conservation for transformations where necessary, the final transformation normalises for total flux within the stack group.

Several transformations are performed on the WHISP data cubes, altering the flux distribution within the data cubes. As such, particular attention is paid to the transformations so that the total integrated flux is conserved, where appropriate. We discuss, and present our flux conservation and budget for those scenarios here. Naturally, flux is required to be conserved during the cube extraction and re-extraction steps in Figure 3.5, since these steps simply exclude noise voxels in extended regions - the sum of which should fluctuate around zero.

Furthermore, the alignment of major axes by rotating the spatial channels by the contained galaxy's position-angle (PA) is also expected to conserve flux. Similarly, the inclination spatial scaling - where galaxies are scaled to have the same axial ratio (B/A) - is also expected to conserve flux, since these are simple changes in apparent orientation. In order to preserve flux values, we perform the rotation using the *rotate* function from the *ndimage* module of the *numpy* library.

Finally, the spectral-scaling step (S_v) is required to conserve flux as we stretch the galaxy along the spectral axis. However, several steps during the signal alignment stage are not expected to conserve flux.

As we resize galaxies within stack groups to occupy the same number and configuration of voxels within the data cube, analogous to aligning the 1D spectra, we are effectively moving the galaxy's apparent distance. As such, we should observe an increase in measured flux from the galaxy proportional to the scaling applied. It is for this reason that the spatial scaling matching the angular extent of the galaxies (R_{HI}) is purposefully performed in a way which does not conserve flux.

These scalings are achieved by replicating (or combining - for scalings below a factor of 1) a three-dimensional array. To achieve this, we utilise the *rescale* function from the *transform* module of the *skimage* (scikit-image) library. In cases where flux conservation is expected, we divide the resulting scaled array by the product of the applied scaling factors (i_S and S_v), which are anticipated to conserve flux. In cases where the flux is not expected to be conserved following a scaling, we do not apply this correction.

Step	**Description**	**Flux Conserved**
Galaxy Extraction	Extracting the galaxy within the larger WHISP data cube.	Yes
Galaxy Rotation	Rotating each spectral channel to align the galaxy's major axis with the y-axis.	Yes
Spatial Scaling	Resizing the cube spatially (matching R_{HI} and inclination i)	No
Spectral Scaling	Resizing the cube spectrally (matching V_{rot})	Yes
Re-extraction	Recentering and re-extracting the scaled galaxy within the cube	Yes
Convolution	Convolving all scaled galaxy cubes to the same beam size	Yes
Flux Normalising	Normalising all galaxies to have the same total HI-mass	No

Table 3.2: Summary of transformation steps and the flux-conservation status.

As a consequence of our electing to stack by mass bins, the final step in the *Cube Stacker* pipeline is a mass-normalising step. In this step, the area beneath the surface-brightness profile is calculated, which is proportional to the HI-mass, for each of the scaled stacklings. By standardising the masses of all the stacklings in the group, before stacking, any measured quantities should be representative of a galaxy with a given HI-mass, in this case, the upper end of the mass bin, as we scale all stacklings to have the HI-mass of the stackling whose SBP has the largest area. Naturally, this scaling is not flux-conserving.

In Figure 3.9 we present the moment 0 maps for all stages to demonstrate the effects of the various transformations in the *Cube Stacker* pipeline. U7704 undergoes a spectral scaling of $v_v = 2$, while the spatial axes are scaled summarised in Equation 3.8. For ease of representation, we adopt values of $a_S = 2$ and $i_S = 0.5$. We draw attention to the conserved relative size of the beam to the size of the galaxy following the spatial scaling. As a result, the number of pixels per beam increases, ensuring that our approach of scaling up galaxies does not lead to an artificial increase in resolution. Furthermore, as predicted in Table 3.2, we note the factor of 2 increase in the peal flux as a result of the additional inclination of a factor of (0.5) correction along the minor axis.

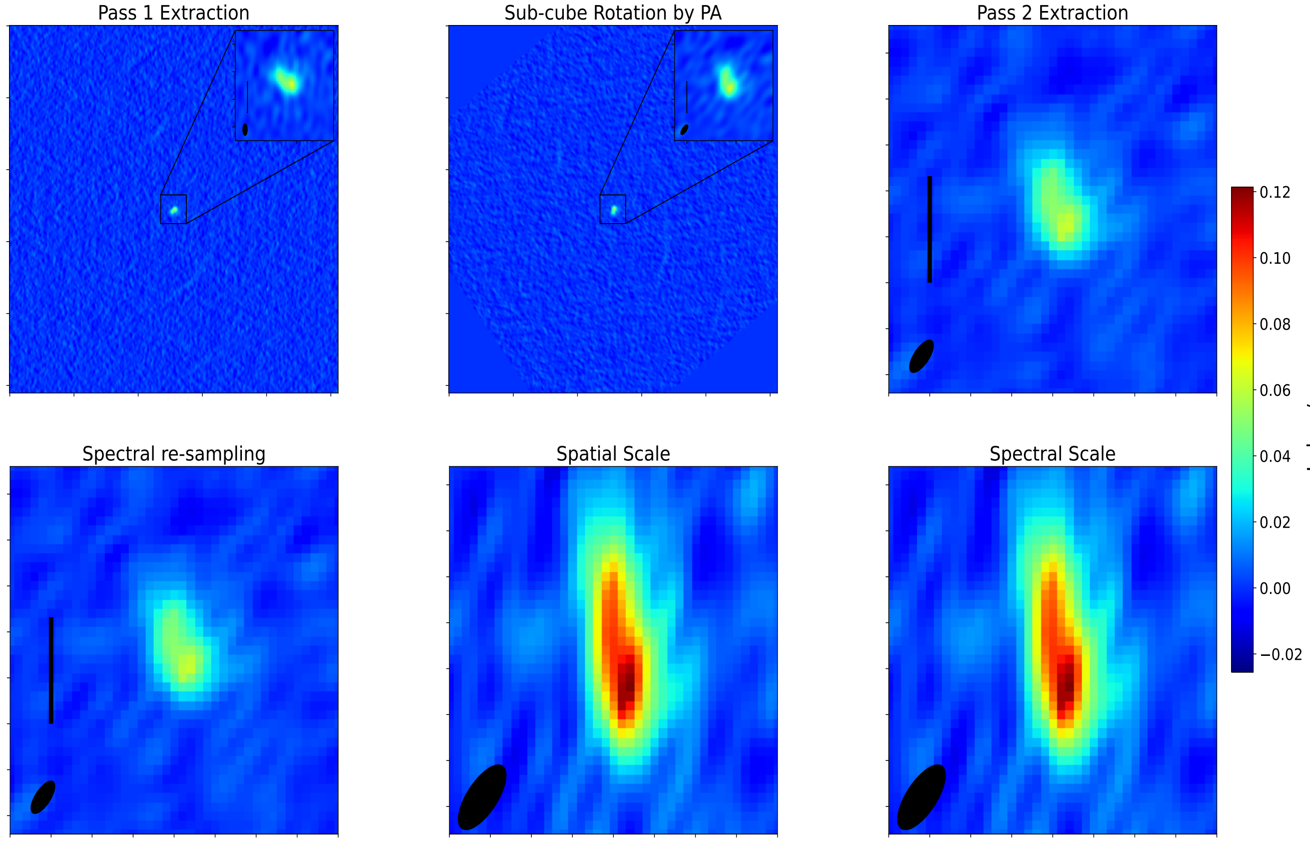


Figure 3.9: Demonstration of the effects of the *Cube Stacker* pipeline on the resulting integrated flux (Moment 0) of the resulting transformed cubes of UGC 7704. The vertical black line denotes the apparent HI-radius ($R_{HI} = 58 arcseconds$) of the original galaxy. The elliptical patch represents the beam-size following the displayed transformation above each subplot.

3.2 Constraining HI Distributions

As outlined in Figure 3.5, several complementary processes run parallel to the *Cube-Stacker* pipeline.

In order to produce meaningful stacks, the galaxies must be centred and aligned uniformly in their respective data cubes. Following the rotation of the galaxy to a fiducial $PA = 0°$, a re-centring is required to uniformly distribute the three-dimensional spectra. Thus, critical information about HI distribution - such as central-voxel, spatial-extent (R_{HI}), spectral-extent

We attempt to constrain and centre the HI both spatially and spectrally. One spatial and the spectral axis are constrained by focusing on the HI distribution within a position-velocity diagram (PVD). This is outlined in Chapter 3.2.3. Here, we begin by describing what a PVD is, as it is central to both the spectral alignment. We then describe an implementation of envelope-fitting as described in Chapter 3.2.3, the results of which are presented in Table 7.1 of Chapter 7.2.

3.2.1 Position-Velocity Diagrams (PVD)

Position-velocity diagrams (or PVDs) offer insight into the distribution of HI both spatially and spectrally. As demonstrated in Figure 3.10, a major-axis PVD is extracted by collapsing all planes falling within a slice, centred on the galaxy's position, of width dW along the galaxy's minor-axis.

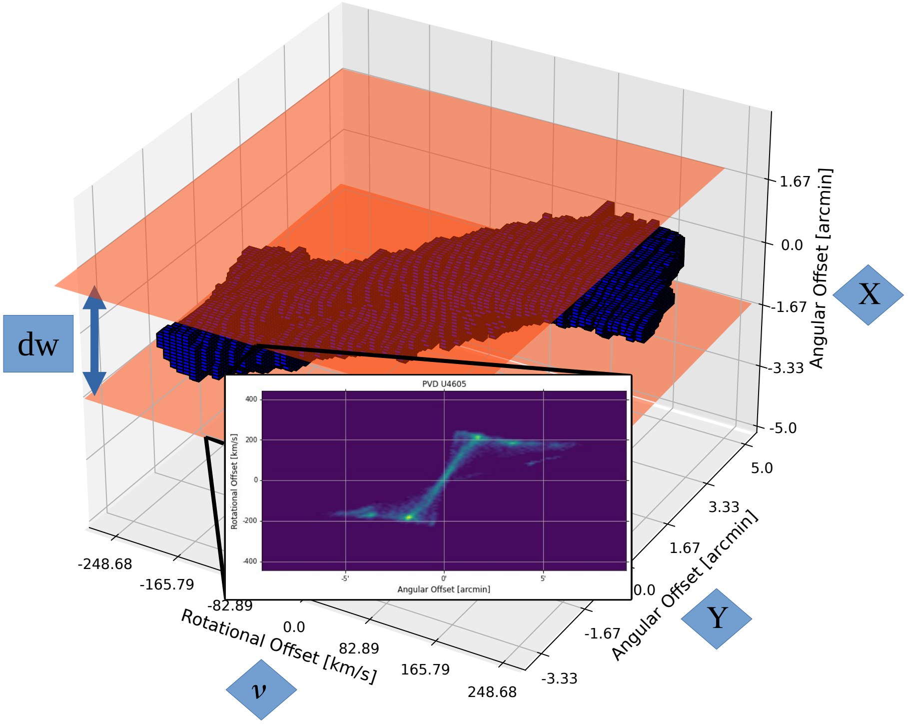


Figure 3.10: Demonstration of the extraction of a PVD. As the width of extraction chosen (dW) increases, more of the galaxy's HI is sampled. However, the contribution due to the noise also increases once dW exceeds the minor-axis width.

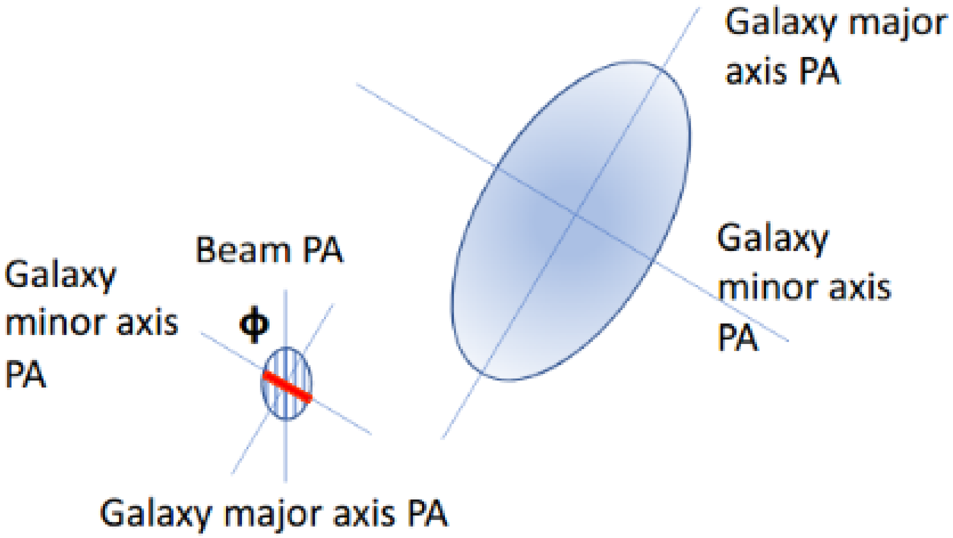


Figure 3.11: Illustration of the width used to extract a PVD from the data cube. The red line drawn across the ellipse representing the beam has a length equalling that of a line parallel to the galaxy minor axis projected along the beam. This length is used to approximate the resolution of the beam along the minor axis of the galaxy. This is important when dealing with the effects of beam-smearing near the centre of the galaxies where the radii are comparable to the beam width.

The width of the slice is significant, as it defines how much of the planar gas is included in the resulting PVD. A slice which is too narrow ($dW << R_{HI} \cos i$; the galaxy's minor-axis length), is guaranteed to be extracting mostly planar-gas.

When trying to map the HI distribution within the position-velocity frame, otherwise rotation curve fitting, the larger the aperture width, the more planar and extra-planar gas gets reprojected onto the tangent point along the line of sight. This has the effect of broadening the number of envelopes present in Figure 3.12c. In contrast, then, an extraction width which is too narrow, say the width of the resolution along the minor axis, we lose both flux and do not improve on constraining the position of the "actual" tangent point at a radial position r. The purpose of this interplay is to optimise the gas at this tangent point.

Additionally, making the aperture too large ($dW >> R_{HI} \cos i$) results in the addition of noise voxels which dilute the signal in the projected image.

As demonstrated in Figure 3.11, adopting a aperture equal to the beam-width parallel to the galaxy's minor axis decreases the effects of beam-smearing in the inner regions of the PVD since we would not be extracting data sampled by the same beam.

In the context of signal alignment, we use PVDs as a basis for standardising the two-dimensional distribution of the HI spectra prior to stacking. After blueshifting the systems, the spectral extent (later referred to as V_{rot}) is obtained by means of measuring envelopes and fitting rotation curves.

3.2.2 Peak-Intensity Mapping

We have taken the approach of mapping the peak intensity pixels in the PVD across spatial apertures drawn across the galaxy spectrum. The line profiles are assumed to be symmetric about the peak-intensity value. This assumption is most applicable at the outer radii (Mathewson et al., 1992; Mathewson & Ford, 1996). At the inner radii, this assumption fails due to the line profiles being a superposition of all radii sampled along the line of sight. At these radii, the envelope-tracing velocity method performs better. In order to increase sensitivity, the individual stacklings are spectrally smoothed using a one-dimensional Gaussian kernel with full-width at half-maximum (FWHM) of:

$$FWHM = 2.355\sigma \tag{3.14}$$

We anticipated the best results when the FWHM is comparable to the width of the flux distribution. In a one-dimensional sense this would correspond:

$$FWHM = W_{20}, \tag{3.15}$$

with W_{20} as tabulated by Noordermeer et al. (2005).

Upon further reflection, we have concluded that this was not an appropriate approach . In any given aperture, the *expected* FWHM of an individual HI cloud corresponds to a value of $\sigma_{HI} = 11.9 \pm 3.1 kms^{-1}$ (Caldú-Primo et al., 2013). Whereas W_{20} describes the width of the *integrated* HI line profile. As a result, our aggressive smoothing has resulted in the *blending* of different velocity envelopes - increasing the error in an ultimate estimate of V_{rot} and h_{rot} in the later fit of the tanh model to these envelopes. Due to time constraints, we were unable to rectify this, however, we advise a correction in future iterations of a *Cube Stacker* pipeline.

Combining Equation 3.14 and Equation 3.15, our adopted standard-deviation becomes:

$$\sigma = \frac{W_{20}}{2.355} \tag{3.16}$$

Our smoothing kernel takes the form:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp{(-\frac{x^2}{2\sigma^2})} \tag{3.17}$$

We make use of the *astropy.convolution.Gaussian1DKernel* function to generate this convolution kernel, specifying only the standard deviation (σ) from Equation 3.16. This defaults to a mean of $x_0 = 0$.

Following the spectral smoothing, we map out the brightest pixels along each column (looping through the spatial axis) of the PVD. To ensure that the peak value in the column corresponds to the signal, we impose an SNR threshold on the signal.

With a set of spatial and spectral coordinates for the peak-intensity pixels in the spectra tabulated, the central coordinates were ready to be determined by fitting a rotation curve (RC). This process is outlined in Chapter 3.2.4.

3.2.3 Envelope Fitting

Modelling a galaxy as a rotating disc, when viewing any galaxy with moderate-to-high inclination, the line of sight intersects with multiple tangent points - all located at different distances, and at varying rotational speeds. This creates the non-Gaussian shape characteristic of one-dimensional profiles of galaxies at high inclinations. A version of the envelope-tracing technique is employed by Kregel & van der Kruit (2004). The technique entails fitting a Gaussian of constant width to the high-velocity end of the peak, at fractional values of the peak intensity of the emission line, as demonstrated in Figure 3.12c.

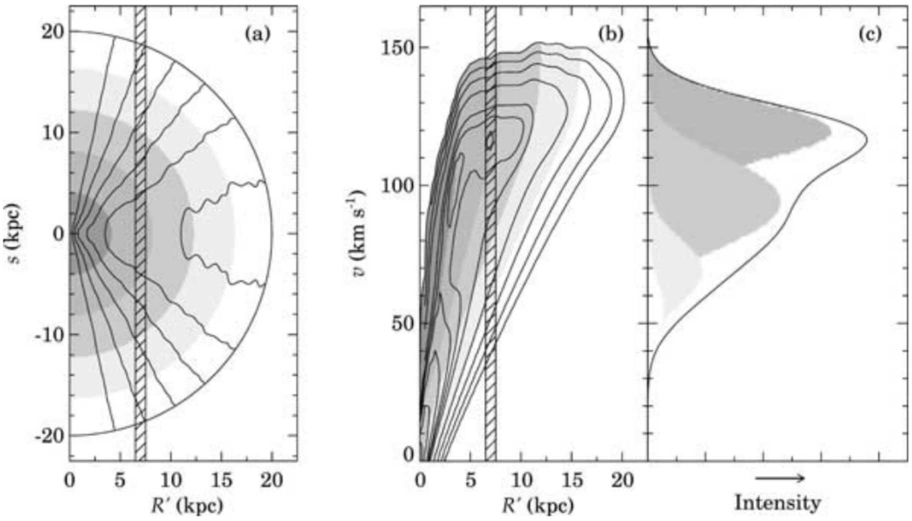


Figure 3.12: Illustration of the line-of-sight effect by Kregel & van der Kruit (2004), using a simulated edge-on view of the HI in NGC 2403 (adopted distance 3.2 Mpc). (a) A spider diagram showing the line-of-sight velocities on the receding half of the galaxy, the vertical ribbon represents the line-of-sight, along which the velocity measurement is made. (b) The vertical ribbon overlaid on the integrated major-axis position-velocity diagram of the receding end. The grey scale indicates the rings in the disc plane from which the HI originates. (c) The one-dimensional velocity profile is extracted in the vertical ribbon. The non-Gaussian profile is visualised as the product of the superposition of Gaussians centred on differing velocities.

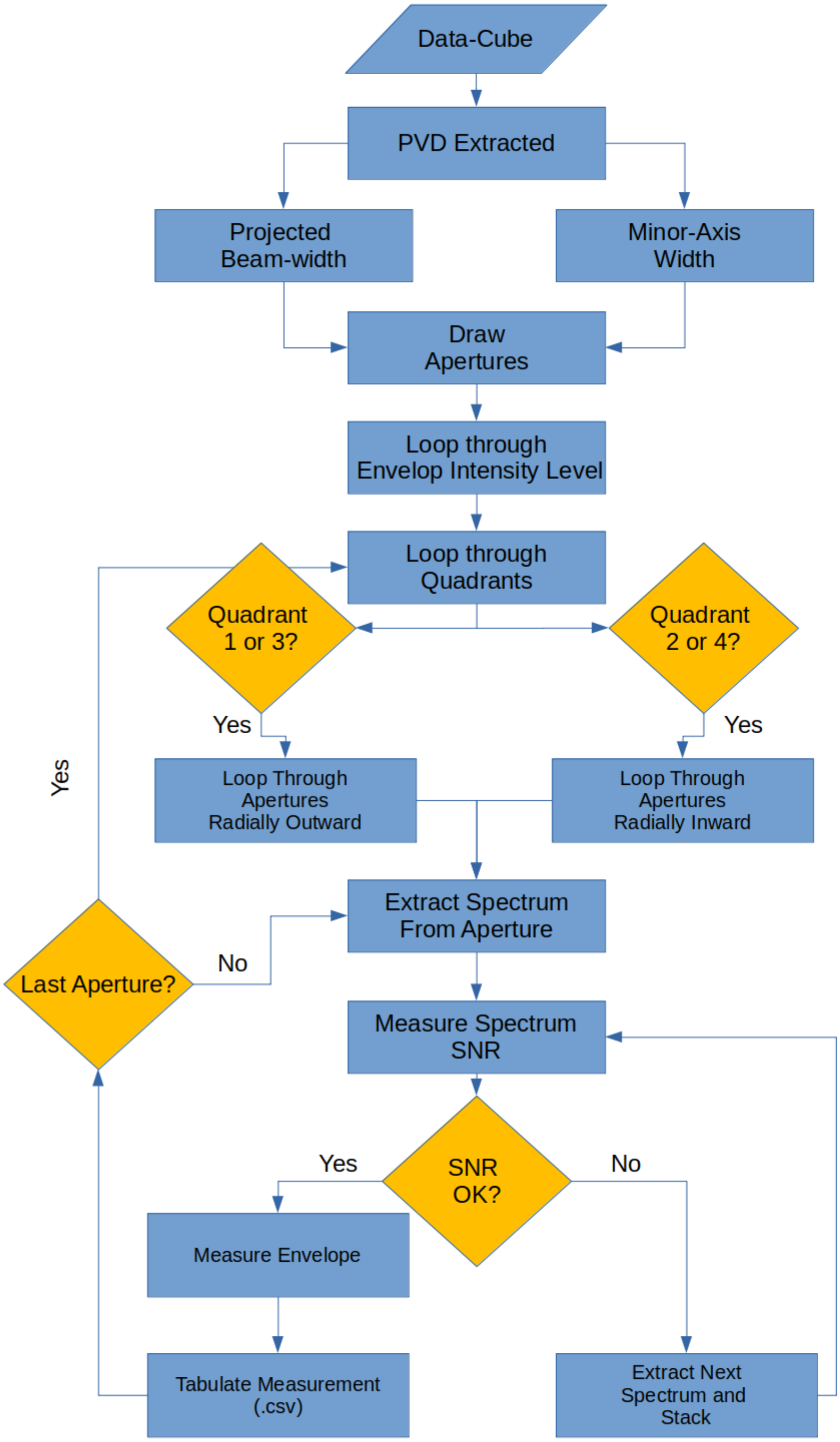


Figure 3.13: This flow diagram shows the general overview of the envelope-fitting algorithm we have developed in Python 3.8. All data cubes (standardised by the Cube-Stacker algorithm) are passed through this pipeline, from which a series of intensity envelopes are mapped out on the derived PVD. These envelopes are later used for centralising the spectrum in the data cube, as well as obtaining a measure of spectral- and spatial extent.

Summarised in Figure 3.13, we apply an envelope-tracing technique to our Noordermeer et al. (2005) sub-sample, using the following procedure:

Generating Position-Velocity Diagrams (PVD)

First, a major-axis PVD is generated by passing in a data cube whose central pixel contains the right ascension, declination, and redshifted HI frequency corresponding to the galaxy's velocity. Thus, we ensure the galaxy's measured coordinates are in the centre of the extracted PVD. This cube is referred to as a "Pass 1 Data-Cube" in Chapter 3.1.3. Two aperture widths were considered when extracting a PVD as discussed in Chapter 3.2.1. In this work, with the extraction of envelopes in mind, we opt to utilise the length of the galaxy minor axis as the value for the aperture width (dW in Figure 3.10).

Determining PVD Sampling Widths

Next, we generate a set of apertures along which spectra (one-dimensional velocity profile in Figure 3.12) will be extracted from the PVD. We have adopted an axis-symmetric approach, thus having the central pixel contain the centre of the galaxy is vital.

Each aperture has a width equal to that of a single pixel.

With the central voxel determined, we define four quadrants containing apertures corresponding to the outer-left (quadrant 1 or Q1), left-disk (Q2), right-disk (Q3), and outer-right (Q4) regions of a galaxy. The reason for defining these regions is to minimise the stacking of apertures with a high SNR with those of low SNR (or indeed apertures containing only noise).

Defining regions of expected high-SNR is not as simple as defining an outer radius (for instance, the half-light radius). Due to the existence of HI-holes, an inner- and outer-limit on the radius is required. For the inner radius (and thus the boundary between Q1 and Q2, as well as Q3 and Q4) we use $R_{inner} = 0.3R_{HI}$. For the outer radius (and thus the radial extent of the apertures) we use $R_{outer} = 1.2R_{HI}$. These values were determined through visual inspection of the 33 galaxies in our sample to ensure that we minimise the occurrence of the co-addition of high-SNR apertures with low-SNR apertures.

When iterating through the apertures and extracting spectra, we vary the direction of iteration. In an absolute sense, for $|R| < R_{inner}$, we loop towards the PVD centre to preserve the higher SNR apertures as we stack into any HI-holes as $R \longrightarrow 0''$. For $|R| > R_{inner}$, we loop radially away from the PVD centre as we progress further into the circumgalactic medium and begin to probe extraplanar regions.

As shown in Figure 3.14, we construct the boundaries between apertures, representing the location of the apertures by plotting vertical lines at the location of the boundaries, thus enclosing the region where a spectrum is extracted. The extent of the apertures is set at $2.4 \times R_{HI}$. In each aperture, a 1D spectrum is extracted, and if the SNR exceeds 3σ, a Gaussian is fit to the high-velocity end of the profile, as demonstrated in Figure 3.12.

The spectrum is thus easily extracted using only two parameters, corresponding to the left- and right-boundary enclosing the region of extraction.

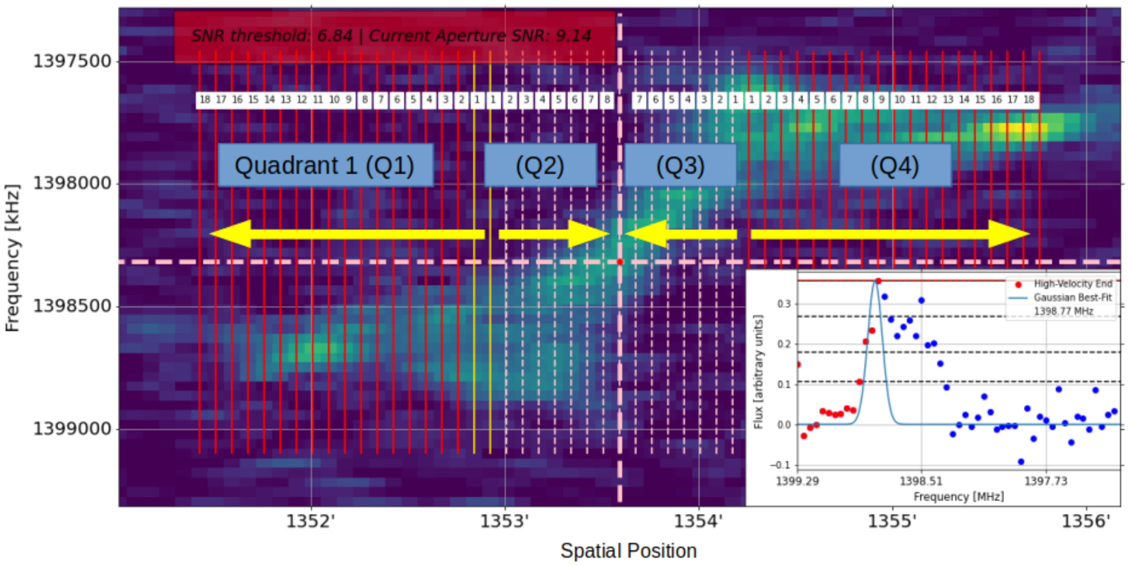


Figure 3.14: Graphic demonstration of the envelope-fitting process. Apertures (vertical lines) are used to extract spectra along a fixed Spatial Position. The apertures are grouped in quadrants as labelled in the figure, with each quadrant specifying a *direction of extraction* for the spectra (shown by the yellow arrows). An SNR criterion is set to ensure the quality of measurement, and in the case of U4458 here, the criterion is met. Following this, a Gaussian of fixed width is fit to the high-velocity end of the spectrum, with the rest of the data masked out.

Spectral Extraction and SNR Calculation

Following the spectral extraction, we notice the varying signal-to-noise (SNR) of the individual spectra. To ensure high fidelity of measurement, we impose an SNR threshold before performing any fits, initially excluding any apertures containing poor signal. This threshold is altered on a per-galaxy basis and informed by the mean of SNR across all apertures.

To calculate the SNR across all apertures, we take the ratio for the average signal within the voxels bound in the central regions of the cube (marked 'S' in Figure 3.15) to the average RMS of the voxels enclosed by the volume defined by the regions marked 1-4 in Figure 3.15. In the latter case, we use this measure as a measure of the RMS of the cube (σ_{RMS}) in subsequent envelope fitting as well as uncertainty estimates in moment 0 maps (see Equation 4.8) and the associated surface-brightness profiles (SBPs; see Equation 4.7), and finally the variance applied to the inverse-variance weighting when stacking in 3D thereafter (see Equation 3.12).

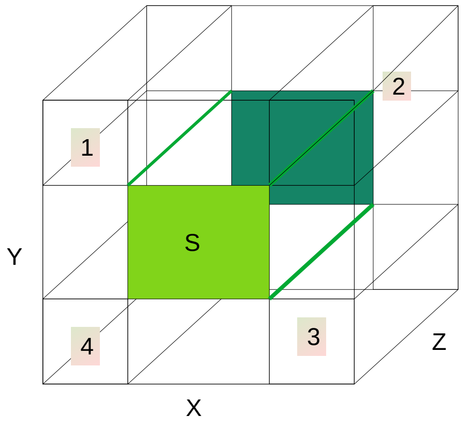


Figure 3.15: Visualisation of the measurement of the cube signal-to-noise ratio (SNR). The cube noise (σ_{RMS}) is taken as the average of the individual RMS values across voxels in regions 1-4. An estimate of the signal within the cube is taken as the average of the voxels within the region 'S'. This method is applied to data cubes whereby the galaxy is known to be centrally contained. Here, X-Y is the spatial axes, while Z is the spectral axis.

Should an aperture fall below the SNR threshold set for the current quadrant, the aperture width is broadened to include the aperture which was next in line for extraction. Depending on the quadrant in question, this is the neighbouring aperture radially internal to the current aperture (for quadrants 2 and 3), and radially external in the case of quadrants 1 and 4.

$$SNR = \frac{S}{\sigma_{RMS}} \tag{3.18}$$

To optimise the number of envelopes across the extent of the galaxy PVD, we employ differential SNR criteria. In regions where there is expected to be sufficient signal (quadrants 2 and 3), we impose a higher SNR criterion. We iteratively gauge the required SNR per quadrant by inspecting each extracted aperture within each quadrant. Figure 3.14 demonstrates the signal extraction process, with the quadrants in which the differential SNR criterion are measured and applied.

Spectral Fitting

For apertures containing sufficient signal, a Gaussian of constant width is fitted to the high-velocity end (in an absolute sense). This is achieved by clipping any data on the low-velocity end following the last data point meeting the intensity threshold.

Unlike the two intensity thresholds in Kregel & van der Kruit (2004) used to bound the value of the tangential velocity at a specific radius, we fit the peak intensity level (100%). We justify this by noting the variety of inclinations of our galaxies, and that the adoption of a chi-squared confidence interval is more appropriate - see Chapter 3.2.3. Next, we clip all data points in the extracted spectral profile at a lower rotational velocity (relative to the barycentre) than the velocity at which the spectral profile becomes maximal. The remaining, unclipped spectral profile is then fit with a Gaussian of fixed width. The width of the Gaussian widths is defined by the following relation:

$$\sigma_{gw} = \sqrt{(\sigma_{velocity-dispersion})^2 + (\sigma_{WSRT-resolution})^2} \tag{3.19}$$

where the contributions to the spectral sampling, such as the velocity dispersion and the resolution of the WSRT, are added in quadrature.

During the process of this thesis, we explored the inner workings of common chi-square fitting routines - notable among which is the *scipy.optimize.curve_fit* method. To do this, we implement our own chi-squared minimisation routine, using Python 3.8.

For our two-parameter model (a one-dimensional Gaussian with fixed width), we adopt a confidence level of CL = 68%. This corresponds to a critical level of $d\chi^2_{CL} = 2.355$. Our chi-squared minimisation obtains a χ^2_{model} for each combination of amplitude and spectral-centre fits (A_{fit},v_{fit}), storing it to a list of best-fit chi-squared values ($\chi^2_{best\_fit}$). Mathematically, the model and χ^2_{model} values are calculated as:

$$y_{model} = A_{fit} \exp \frac{-(v - v_{fit})^2}{2\sigma_{gw}} \tag{3.20}$$

$$\chi^2_{model} = \sum_{i=1}^{N} \frac{(y_{data,i} - y_{model,i})^2}{\sigma^2_{data}}, \tag{3.21}$$

where σ^2_{model} is a measure of the variance of the data in a high-velocity (absolute) region, containing little or no signal.

Envelope Error Estimation

We then constrain our error on the spectral position of the envelope using the following chi-squared analysis: We find, and store to list, all the centroid fits with $\chi^2_{best\_fit}$ values falling within the interval

$$\chi_{confidence} = \{\chi^2_{fit} \leq \chi^2_{best-fit} + d\chi^2_{CL}\} \tag{3.22}$$

$$v_{conf} = \begin{cases} v_{conf,low} = v_{fit}[MIN(\chi_{confidence})] \\ v_{conf,high} = v_{fit}[MAX(\chi_{confidence})]. \end{cases} \tag{3.23}$$

This interval, often asymmetrically distributed about $\chi^2_{best-fit}$, is next used to determine upper- and lower-limit errors as follows:

$$dv = \begin{cases} dv_{lower} = v_{best\_fit} - MAX(0, v_{conf,low}), \\ dv_{upper} = v_{best\_fit} + MAX(0, v_{conf,high}). \end{cases} \tag{3.24}$$

3.2.4 Fitting Rotation Curves

Rotation curves serve as great tools for studying galaxy kinematics; inferring evolutionary histories of galaxies; studying the effects of interactions between galaxies; and probing the matter distribution (both luminous (Swaters et al., 2012) and dark (Hoekstra et al., 2001)) responsible for the nature of galaxy rotation.

Rotation curves which are derived from emission lines such as HI (atomic gas), Hα (ionised gas), and CO (molecular gas) have added utility when deriving mass distributions. When observed in late-type galaxies, they are sensitive to the smaller velocity dispersion (5-10 kms^{-1}) of interstellar gasses, serving as excellent tracers.

HI has proven particularly useful, with early radio observations of external galaxies showing a slowly falling rotation curve for M31 (van de Hulst et al., 1957) and a flat rotation curve for M33 (Volders, 1959). Compared to the low surface brightness of the molecular gas corresponding to the stellar emission, the greater spatial extent of HI provides a deeper probe into the kinematics, and as a result greater constraints on the mass in a galaxy responsible for the observed kinematics of the galaxy.

With the galaxy kinematics traced out, as discussed in Chapter 3.2.2 and Chapter 3.2.3, rotation-curve models were fit to the data. We use a Tanh model rotation curve to describe the PV distribution of the HI spectra toward two ends. First, an initial fit is used for standardising the defined barycentre along the rotated declination ($\hat{Y}$ or $\hat{\delta}$') and spectral ($\hat{V}$ or $\hat{v}$') axes. Secondly, we fit the model, with the updated centering, and use the model to describe the HI distribution in the PV domain, thereby having a standard to perform signal alignment prior to the stacking stage. We describe these methods in detail in the following subsections.

Tanh-Model Rotation Curve

Andersen & Bershady (2013) utilize a two-parameter inclined hyperbolic tangent function (tanh-model, hereafter) with steady rise (characterised by a scale length h_{rot}) to a projected asymptotic velocity (described by V_{rot}). Thus, the model takes the functional form described in Equation 3.25. The variability in rotation-curve shapes observed in our WHISP-NSA sample makes the flattened tanh profile a neutral initial choice for characterising the spatial-spectral distribution of galaxies. In Figure 3.16 we demonstrate the effects of different values of V_{rot} and h_{rot} on the shape of the tanh-model.

$$V_{tanh}(R) = V_{rot} \tanh\left(\frac{R - \delta'_0}{h_{rot}}\right) + v'_0 \tag{3.25}$$

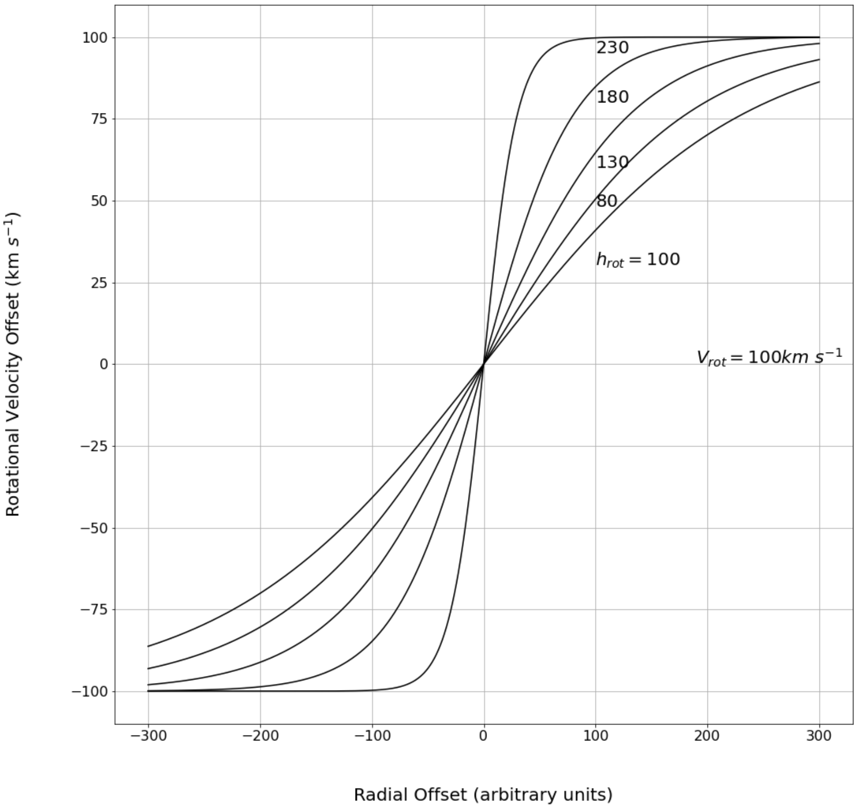


Figure 3.16: Example of Tanh models for the fitting of rotation velocities. Different values of the rise scale length, h_{rot}, are shown for five models with a rotational velocity of $V_{rot} = 100kms^{-1}$.

Improved Centering

With a set of coordinates mapping out the peak-intensity pixels in the PVD (discussed in Chapter 3.2.2), we fit a Tanh model rotation curve. Taking in the functional form of Equation 3.25 discussed in Chapter 3.2.4, initial estimates for the amplitude (V_{rot}), rising-scale (h_{rot}) and centre (δ'_0, v'_0) are made and passed in the *curve_fit* function of the *scipy.optimize* module.

Using the Levenberg-Marquardt algorithm as the least-squares optimization, *curve_fit* returns the parameters which minimise the sum of the squared residuals between the model and data.

Since the peak-intensity pixel coordinates are derived from discrete data (pixels) we apply an equal weighting for each data point.

The output of this procedure is the acquisition of a standardised measure of the barycentre across stackling datacubes. Thus, we use the fit obtained for the centre (δ_0', v_0') of Equation 3.25 as the new point around which to re-extract our stackling, using the original RA from (Noordermeer et al., 2005).

HI PV-Distribution

With our newly obtained PV barycentre, we next make use of the coordinate mapping of the HI distribution within the re-extracted data cube. The mapping is obtained using a novel envelope tracing routine developed specifically for this *Cube Stacker* pipeline. See Figure 3.13 for a summarising flow diagram.

Now constraining the centre of the Tanh model, Equation 3.25 becomes:

$$V_{tanh}(R) = V_{rot} \tanh\left(\frac{R}{h_{rot}}\right) \tag{3.26}$$

Using the parameters of scale such as V_{rot} and h_{rot}, obtained by a least-squares fit of Equation 3.26 to the envelopes from Chapter 3.2.3, we are able to describe the distribution of HI in Position-Velocity space along the kinematic major axis. We consider the spatially folded quantity, known as the rotation curve. We summarise the result of the rotation curve fits to this data in Table 7.1, as well as the individual rotation-curve, fits relative to the PVD in Figures 7.4-7.8.

Rotation Curve Errors

There is no consensus on the method of estimating the uncertainty associated with a rotation-curve de Blok et al. (2008). Application of a χ^2-fit uncertainty often produces values smaller than the velocity dispersion of the gas in the galaxy. With this in mind, we adopt the typically used technique of fitting a Tanh model to the approaching and receding ends of the galaxy separately. By taking the difference in amplitude between the two ends, we obtain an estimate of the uncertainty of the rotation-curve amplitude (V_{rot}) measures.

In this iteration of the Cube Stacker pipeline, we make use of V_{rot} to describe the spectral extent of the galaxies within their respective cubes. Thus, replacing the previous measure of spectral extent of W_{20} in the previous iteration of our pipeline. We present the results from the implementation of this section in Chapter 7.2.

Chapter 4

Results

In this chapter, we present the results from measurements performed on our stacked spectra. As mentioned in Chapter 2.5, we use a combination of HI mass and inclination binning to subdivide our 33 galaxy subsample of the Noordermeer et al. (2005) sample of WHISP galaxies. This resulted in 5 stack groups. Table 2.2 summarises these groups, with a lower limit on the number of galaxies per bin of 5 enforced.

The stacked spectra were obtained by running the Cube Stacker pipeline outlined in Chapter 3. Here we assess (i) any improvements in stacking in three dimensions compared to conventional one-dimensional stacking; and (ii) characterise any faint signal uncovered in the stacked data. In § 4.1 we describe how we perform the measurement of the HI radial surface mass-density from the data cubes. The results for each of the stack groups are then found in § 4.2 (Group 1; low mass, low inclination), § 4.3 (Group 3; low mass, low to moderate inclination), § 4.4 (Group 6; intermediate mass, low to moderate inclination), § 4.5 (Group 9; high mass, low to moderate inclination), § 4.6 (Group 10; all mass, high inclination). For ease of reference, boundaries between low, intermediate and high HI mass are 10^9 and $10^{9.75}$ $M_{\odot}$, while the boundaries between low, moderate, and high inclination are 47.9° and 70°. In § 4.7 we compare the profile results for these different bins in mass and inclination.

Finally, we summarise the signal alignment obtained by utilising a tanh model as the standard across all stacks in § 5.2.1. We direct readers to Chapter 7.2 for the rotation curve measurements used in the signal alignment process.

4.1 Measured Quantities from Stacked Spectra

4.1.1 Surface Brightness Profiles (SBP)

To quantify the effect on the signal of our three-dimensional stacking pipeline, we make use of surface brightness profiles (SBPs, hereafter) to chart the structure and properties of our

WHISP-galaxies - quantifying how the signal intensity changes as a function of distance from the galactic centre. These profiles are rendered, ultimately in units of $\mathrm{M}_\odot\ \mathrm{pc}^{-2}$.

While the cubes present the signal in units of flux density (mJy/beam), we choose to assess the radial HI distribution in the zeroth-moment (Moment 0). Starting in cube units of Jy/beam, the flux units of the data cubes were converted from mJy/beam to Jy/pixel by dividing by the ratio of the beam area to the pixel area. Next, the spectral axis is converted to units of km/s, corresponding to a spectral resolution of 8.49 km/s for these WHISP data cubes at full resolution. Following the flux density conversion, we utilise the zeroth moment map, obtained by integrating the data cube along the spectral axis, which represents the total intensity or flux distribution, providing a map of the spatial distribution of the observed emission. Expressed mathematically:

$$\mathrm{M}_0 = \int I_\nu d\nu \tag{4.1}$$

where $d\nu$ denotes the spectral sampling (in units of either Hz or kms^{-1}), and $I_\nu(x, y)$ denotes the intensity at each pixel position (x,y). we generate a Moment 0 map, using `spectral_cube`'s built-in function. A further flux-conversion is performed in the Moment 0 map, multiplying by a conversion unit to obtain the map's units in $(M_\odot/pc^2) \times (\mathrm{arcsec}^2)$.

Next, annuli (elliptical or rectangular, as discussed later) are generated. The total flux passing through each annulus is extracted, and the flux density in each annulus is calculated by dividing the flux by the annulus area. To correct for inclination effects, the flux density is multiplied by the cosine of the inclination angle to re-project the values to a face-on perspective. This process is described mathematically by:

$$\left(\frac{\Sigma_{\mathrm{HI}}}{\mathrm{M}_\odot \cdot \mathrm{pc}^{-2}}\right) = 1.024 \times 10^4 (1+\mathrm{z})^3 \left(\frac{\mathrm{M}_0 \cos \mathrm{i}}{\mathrm{Jy} \cdot \mathrm{kms}^{-1}}\right)\left(\frac{\Omega_{\mathrm{ann}}}{\mathrm{arcsec}^2}\right)^{-1} \tag{4.2}$$

where Ω_{ann} is the annular area (Meyer et al., 2017).

Due to the range of inclinations of the galaxies in our sample, we adopt two aperture regimes, namely: elliptical apertures for moderate to low-inclination galaxies, and rectangular apertures for highly inclined galaxies. See Figure 4.1 for a summary.

4.1.2 Elliptical Annuli

Since our sample of galaxies has been rotated to have a PA=0, we construct a set of concentric ellipses centred on the galaxy's centre. We select a range of semi-major axes (a_j) in 10% increments of the HI-radius, out to a radial extent of $10R_{HI}$. Subsequently, the semi-minor axes are determined by $b_j = a_j \cos i$. As such, the annuli have a fixed width, and the resulting area for each annulus is calculated in units of arcsec^2. In the resulting calculation of surface-brightness and error calculation in Equations 4.2 and 4.7, the annulus parameters in the j^{th} radial bin are defined as:

$$\Omega_{ann,j} = \pi(a_j b_j - a_{j-1} b_{j-1}), \tag{4.3}$$

$$N_{ann,j} = \frac{\Omega_{ann,j}}{dX^2} \tag{4.4}$$

where $N_{ann,j}$ are the number of pixels in the annulus. Here, dX refers to the spatial pixel-separation which is equal along both spatial directions in the data cube, at a value of $dX = 4.96\text{''}/pixel$.

4.1.3 Rectangular Annuli

For highly inclined systems we extract rectangular apertures along the major axis instead. Keeping the apertures axisymmetric, we define the parameters of the apertures to occupy a width along the major axis of $W = 0.1R_{HI}$. The vertical scale height (H) is chosen by visual inspection such that the aperture encloses all of the flux within the contour enclosing $\Sigma_{HI} = 1M_{\odot} \cdot pc^{-2}$. In the case of rectangular apertures, the resulting calculation of surface brightness and error becomes:

$$\Omega_{ann,j} = W_j H_j \tag{4.5}$$

$$N_{ann,j} = \frac{\Omega_{ann,j}}{dX^2} \tag{4.6}$$

for dX = 4.96”/pixel.

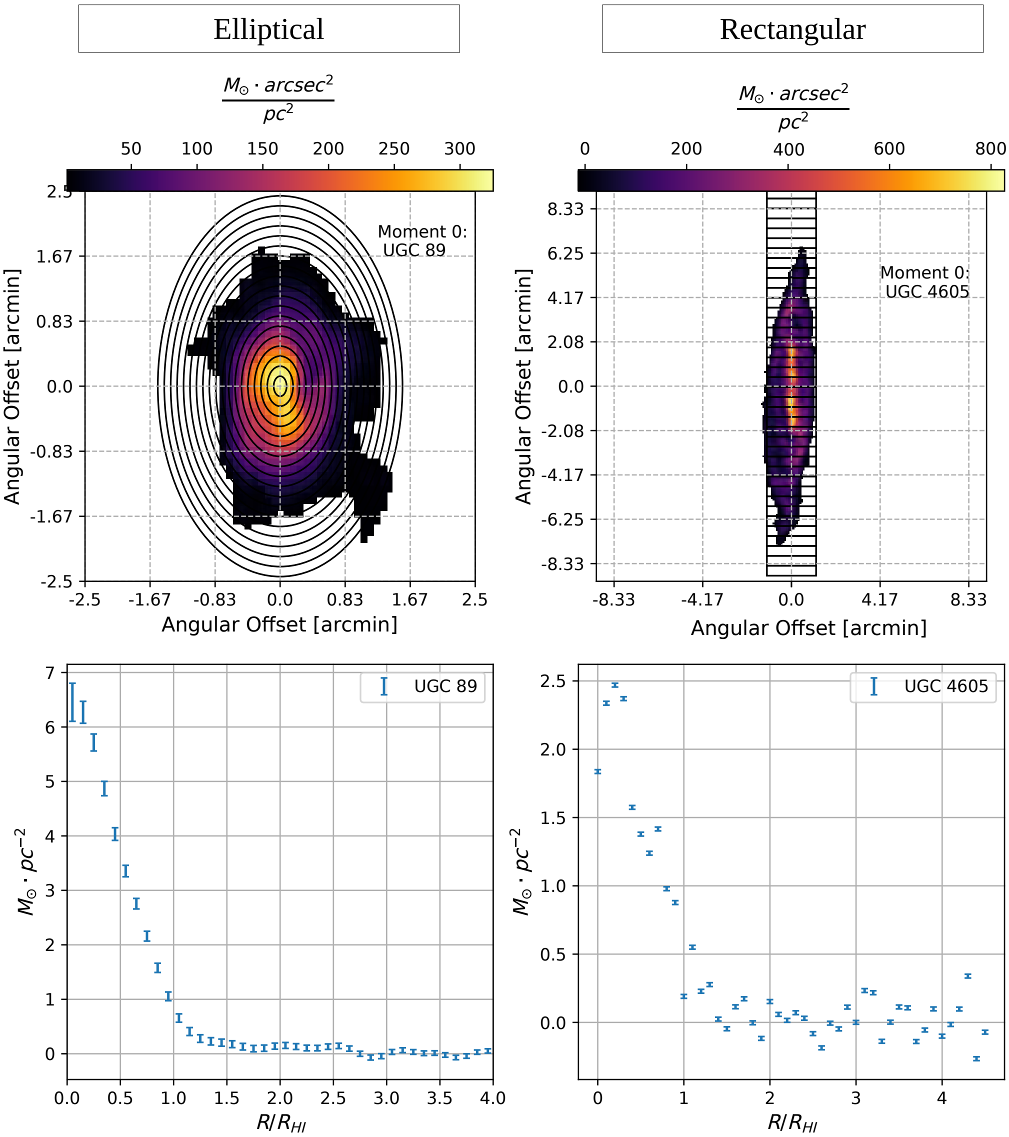


Figure 4.1: Comparison of elliptical and rectangular apertures used for deriving surface-brightness profiles (SBPs) for analyzing the radial HI content of galaxies. The top row shows the aperture types, while the bottom row displays the resulting SBPs. Our approach, using Equation 4.2 accurately determines a surface-brightness of $\Sigma_{HI} = 1 M_\odot \cdot pc^{-2}$ at $r = R_{HI}$, as defined.

4.1.4 Comparison of One-Dimensional and Three-Dimensional Stacking

To compare any potential benefits of three-dimensional stacking over one-dimensional stacking, we begin by extracting SBPs by applying Equation 4.2. We apply this process to each of the individual stacklings (transformed galaxies). We then create a weighted mean profile from all of the individual stackling profiles, as defined below; we refer to the resulting profile as the result of a 1D stack. Finally, we perform an SBP extraction on the stacked data cube for each group of galaxies; we refer to this profile as the result of a 3D stack. We overlay these curves on the measurements of the individual SBPs.

In the ensuing presentation of the surface-brightness profiles for each group, we employ a logarithmic scale to re-plot the measurements. By doing so, we intentionally amplify the differences between these two approaches, allowing for a more comprehensive evaluation of the potential gains obtained through 3D stacking. While the conversion to a log scale is simple for the measurement, to calculate the errors in the log, we use the following approach:

$$\delta log_{10} f = \begin{cases} \log(1 + R_f) & \text{upper error ,} \\ \log(1 - R_f) & \text{lower error ,} \end{cases}$$

where $R_f = \frac{\delta f}{f}$, the ratio of the flux density measurement f to its error δf. We describe how we determine the errors for the 1D and 3D stack profiles in the next subsection.

Flux density has been discussed in several forms in the preceding text, however, in both the 1D and 3D cases, this refers to the surface-brightness. In the 3D case, f is simply the extracted surface brightness profile of the stacked data cubes (weighted or unweighted) according to Equation 4.2. In the 1D case, f refers to the stacked (weighted or unweighted) *SBPs* of the individual stacklings according to Equation 4.9. A reversal in order.

However, δf differs substantially between the 3D and 1D cases. In the 3D case, δf is obtained directly during the SBP extraction process via Equation 4.7. In the 1D case, δf is obtained via Equation 4.10.

In the case of stacklings (data cubes) with identical statistical properties, Equation 4.10 would be an estimate of the random noise across the stacklings, and the resulting error bars would be identical to the 3D case. Since we are stacking the SBPs of *different* sources, the uncertainty in the 1D stack obtained via Equation 4.10 instead is a measure of the variance of the signal (SBPs) between galaxies.

Surface-brightness Profile Errors in 3D Stacks

To estimate the error on the surface brightness measurement through each aperture, we use the standard deviation of the noise through each aperture:

$$\left(\frac{\delta\Sigma_{\mathrm{HI}}}{\mathrm{M}_{\odot}\cdot\mathrm{pc}^{-2}}\right) = 1.024\times10^{4}\left(\frac{\sigma_{\mathrm{M}_0}}{\mathrm{Jy}\cdot\mathrm{kms}^{-1}}\right)\left(\frac{\Omega_{\mathrm{ann}}}{\mathrm{arcsec}^{2}}\right)^{-1}\mathrm{N}_{\mathrm{ann}}^{-\frac{1}{2}}. \tag{4.7}$$

Here, σ_{M_0} is the propagated uncertainty of the measured M_0 in each pixel of the moment-0 map based on the voxel RMS of the progenitor data cube (σ_{RMS}, outlined in Figure 3.15) - assuming spectrally independent pixels. Expressed mathematically:

$$\sigma_{\mathrm{M}_0} = \sqrt{\sum_{i}^{N} \sigma_{RMS}^2 \Delta v_i^2} \tag{4.8}$$

As discussed in Teague (2019), this provides an adequate estimation for the noise in a moment 0 map. However, existing spatial correlations present in interferometric data still need to be accounted for. This consideration is paramount when extracting radial profiles such as the surface brightness profiles we extract from these moment 0 maps (Loomis et al., 2018).

Surface-brightness Profiles and Profile Errors in 1D Stacks

To simulate the expected results from one-dimensional stacking, we stack the individual SBPs of the stacklings, using inverse-variance weighting defined in Equation 3.13. Thus, our one-dimensional weighted mean is formulated as:

$$(S^{\bar{B}P})_{stack} = \frac{\sum_{i=1}^{n} w_i (S^{BP})_i}{\sum_{i=1}^{n} w_i} \tag{4.9}$$

In this one-dimensional case, the weighting is taken as the inverse variance of the noise within the moment 0 map $w_i = \frac{1}{\sigma^2_{M_0}}$. In the case of an *unweighted* mean, the weighting is set to $w_i = 1$.

To estimate the error on this mean, we calculate the weighted standard deviation of the individual SBPs with respect to the mean SBP. This is because our weighted mean is drawn from a sample with non-identical means. As such, we compute the error using the following relation:

$$\bar{\sigma_w}^2 = \frac{N'}{N'-1} \frac{\sum_{i=1}^{N} w_i (x_i - \bar{x_w})^2}{\sum_{i=1}^{N} w_i}, \tag{4.10}$$

where w_i is the weight applied for cube x_i, N' is the effective sample size of non-zero weights, and $\bar{x_w}$ is the weighted mean of the stack.

4.1.5 Uncovering Extended HI Emission in Surface-Brightness Profiles

By treating the individual stackling surface brightness profiles as an analogue to the one-dimensional HI line, we demonstrate the improved capacity to uncover extended features brought on by three-dimensional stacking. We accomplish this by applying the following framework in subsequent subsections within this chapter. We apply sigma-clipping to the stackling as well as the stackling SBP. Next, we apply the sigma-clipping to the stack (three-dimensional) and the stack SBP. Finally, we co-add the individual masked (sigma-clipped) stackling surface brightness profiles - calling this the one-dimensional stack. In this 1D stack, too, we produce a weighted and unweighted stack - with a comparison to the three-dimensional stack provided in Figure 4.9.

As a reminder to the reader, we calculate the weighted average ($\bar{S}_{stack}$) on the stacklings (S_i) in the stack groups. Expressed mathematically:

$$\bar{S}_{stack} = \frac{\sum_{i=1}^{n} w_i S_i}{\sum_{i=1}^{n} w_i} \tag{4.11}$$

where we adopt an inverse-variance weighting (w_i) of:

$$w_i = \frac{1}{\sigma_i^2}. \tag{4.12}$$

Here, i is the index of the data cube to be included in the stack, and N is the total number of stacklings in the sample. We estimate the noise ($\sigma_i = \sigma_{RMS}$, from Chapter 3.2.3) by sampling a flux-free region within the data cube (designated at the four corners spatially) and assessing the standard deviation of the flux across the spectral axis. We then take the average of the σ across the four regions. See Figure 3.15 for a graphical overview. All voxels in cube i (S_i) receive a weighting of w_i before co-addition.

In what we term an *unweighted* average, a value of $w_i = 1$ is adopted in Equation 4.16 for the 3D stack, and Equation 4.9 for the 1D stack. We demonstrate the process of stacking one-dimensional profiles graphically in Figure 4.2 below.

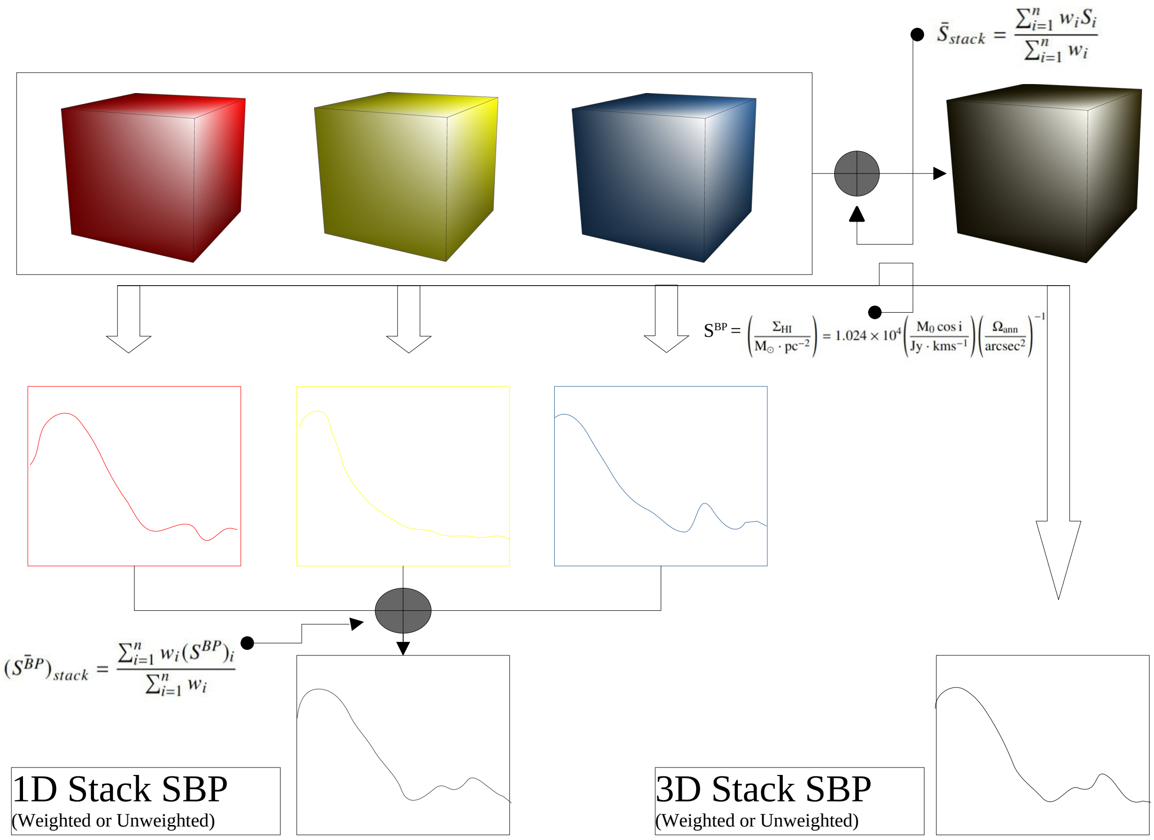


Figure 4.2: Visualisation of the production of a one-dimensional and three-dimensional stack surface brightness profile (SBP). In the former, Equation 4.9 for $S_{stack}^{B\bar{P}}$ is used to average the individual stackling SBPs. In the latter, Equation 4.16 describes the weighted co-addition (symbolised by the cross-hairs atop a grey background) of the cubes, *after* which is the extraction of the SBP. In both cases, the SBPs are extracted (denoted by the hollowed black arrows) using Equation 4.2.

We make use of the process of signal clipping to demonstrate an improved SNR. By fixing the interval of clipping (1, 2, or 3 times the noise amplitude), and assessing the remaining profile of the stack and stackling SPBs, we hope to demonstrate the expected improved SNR since, by virtue of a dampening of the noise through the co-addition of voxels, the only free parameter is the noise itself (σ_{RMS}). Thus, with a lower overall noise level, a greater amount of signal is expected to survive in a data cube (in this case the stack) following sigma-clip masking as opposed to the contributory data cube (in this case the stackling) at the *same* clipping threshold.

Thus, hereafter, we refer to a sigma-clipped map as a *masked* map.

Utilising this difference in noise properties in data cubes, we also assess the effect of an inverse-variance weighting against an unweighted scheme. In Figure 7.3 we present the stacking groups, where we compare the benefits of applying a weighting, dictated by the noise levels of the individual stacklings, to the co-addition stage of the *Cube Stacker* pipeline. This we demonstrate by the amount of signal which survives at greater and greater radial extents. Each stack group contains a discussion of any features surviving each masking threshold (1σ, 2σ, 3σ).

As shown in Figure 4.3, by comparing the SBP of this sigma-clipped data to that of the original, any differences in amplitude and radial extent of the SBP can be attributed to an inherently lower noise level in the stacked cube compared to the individual profiles averaged in the one-dimensional stack.

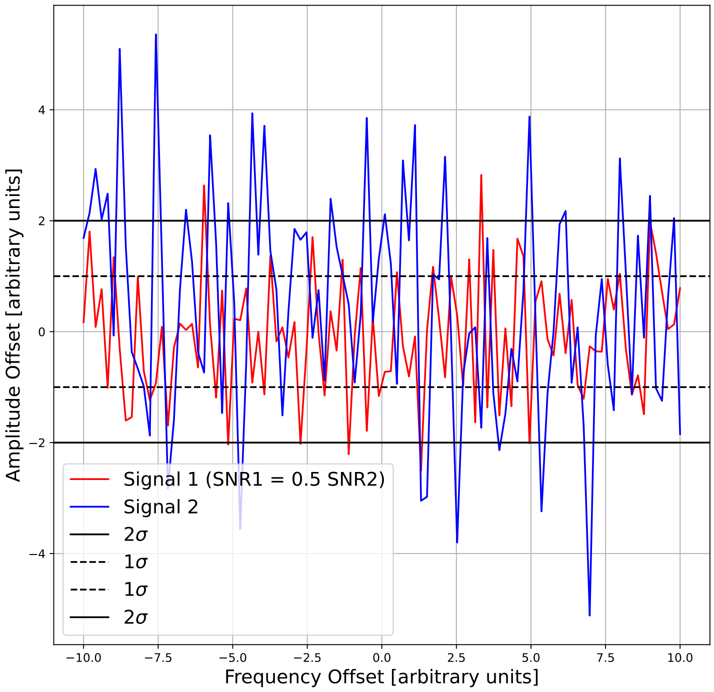


Figure 4.3: Demonstration of the effect of sigma clipping at identical thresholds on a signal with varying SNR. Signal 2 has twice the SNR of Signal 1. As such, any threshold applied to both signals will lead to less clipping in the signal with a higher SNR. Thus, we demonstrate an improved SNR in our stacked products over the individual stacklings.

4.2 Stack Group 1

In Group 1 we stack all low mass ($M_{HI} \leq 10^{8.9} M_{\odot}$) galaxies with low inclination ($i \leq 47.9$).

In Table 4.1, the various scaling parameters (X-scaling, Y-scaling, and V-scaling) are obtained following the alignment to a fiducial measure of size for the sample. In Chapter 3.1.4, we summarise the spatial scaling required to redistribute the signal containing voxels of each to occupy a similar distribution of voxels within the cube. As a reminder, the spatial scaling X-scaling (S_x) and Y-scaling (S_y), and V-scaling (S_v) are calculated as follows:

$$\begin{pmatrix} S_x \\ \\ S_y \\ \\ S_v \end{pmatrix} = \begin{pmatrix} \frac{a_{max,sample}}{a_{stackling}} \cdot i_S \\ \\ \frac{a_{max,sample}}{a_{stackling}} \\ \\ \frac{V^{rot}_{max,sample}}{V^{rot}_{stackling}} \end{pmatrix}. \tag{4.13}$$

Here, $a_{max,sample}$ is the largest major axis (R_{HI}) in the sample. The fiducial HI radius belongs to U11914 and all other galaxies are scaled spatially to have this radius by taking the ratio of the existing HI radius to the fiducial of 113 arcseconds. The X-scaling includes an additional inclination correction (i_S) to a standard inclination of 60 degrees (see Equation 3.7), resulting in a smaller scaling along the X-axis. Added to this U11914 also has the largest spectral extent, with the fiducial $V^{rot}_{max,sample}$ for the stack set accordingly.

The largest weighting in this stack is given to U5060 at approximately 50.4% of the flux in the weighted stack. In Chapter 4.7 we compare the effects of the weighted against the unweighted stack.

UGC	M_{HI} $10^9 M_\odot$	i °	R_{HI} "	V_{rot} kms^{-1}	X Scaling	Y Scaling	V Scaling	σ mJy beam^{-1}	Weighting
5060	0.7	24	60	34.89	0.09	1.88	4.69	3.04	0.11
5906	0.31	29	47	46.5	0.12	2.4	3.52	6.6	0.02
6118	0.84	24	89	80.78	0.06	1.27	2.02	6.28	0.03
6742	0.17	36	51	46.01	0.13	2.22	3.56	4.25	0.06
11914	0.69	31	113	163.55	0.05	1	1	18.58	0.003

Table 4.1: Group 1 Properties: Summary of relevant physical measures used for spectral alignment; from left to right, HI mass, inclination, HI radius, rotational amplitude. Cube scaling values along the X-axis, Y-axis, and spectral axis (V-axis). Finally, a measure of the noise within the cube following the application of scaling. In this stack, the dominant weighting is given to U5060, representing half of the stack. However, the fiducial galaxy is U11914 with the largest spatial and spectral distribution.

The surface brightness profiles (SBPs) in Figure 4.4 of the stacklings overlap well, intersecting at $1 M_\odot \cdot pc^{-2}$ at $r = R_{HI}$ as expected.

As shown in Figure 7.14, a deficiency in flux occurs in the Moment 0 of U6118 and U11914. According to Noordermeer et al. (2005), and unlike U11914, the HI hole of U6118 is not due to an absence of gas, but rather an artefact caused by HI absorption. This feature corresponds to the rising inner regions of the SBP of U11914 and U6118 in Figure 4.4, leading to a lowering of the overall stacked SBPs, both in 1D and 3D.

4.2.1 Comparison of One-Dimensional and Three-Dimensional Stacking

The divergence between the observed SBP of the weighted three-dimensional stack and the one-dimensional weighted mean of the individual stackling SBPs is minimal. A slight increase in the 3D stack is observed within $r \leq 0.5 R_{HI}$, with near identical overlap between the 1D and 3D stack seen in Region A ($1.5 R_{HI} \leq r \leq 2.5 R_{HI}$). However, this divergence still falls within the uncertainty of the measured uncertainty profile of 3D Stack's SBP, requiring greater number statistics to determine the cause.

When comparing the effects of the weighted and un-weighted stacking performed on the galaxies of Group 1 we see a much greater improvement in the SNR in the case of the weighted stack. . In the absence of any weighting, the highly noisy data cube enclosing U11914 correlates with the noise features observed in the stacks. However, despite this, suppression of the noise within the voxels when stacking in 3D leads to a lower noise level once the surface brightness profile is extracted. In the radial ranges displayed, we observe a small region around $r = 3.9R_{HI}$ where the one-dimensional stacking results in lower noise profiles. This feature occurs in both the weighted and unweighted regimes. However, the mean profiles (solid lines atop the shaded regions) remain identical when no weighting is applied. We re-present the derivation of the surface brightness from the Moment 0 map here:

$$\left(\frac{S^{BP}}{M_\odot \cdot pc^{-2}}\right) = 1.024 \times 10^4 \left(\frac{M_0 \cos i}{Jy \cdot kms^{-1}}\right)\left(\frac{\Omega_{ann}}{arcsec^2}\right)^{-1} \tag{4.14}$$

where $M_0 = \int I_\nu d\nu$, as expressed in Equation 4.1.

Identical (unweighted) mean profiles are to be expected since unweighted stacking is technically an equal weighting where $w_i = 1$ for both 1D and 3D stacking in Equation 4.15 below. This means that regardless of the actual effect on the noise properties of the data cube, the resulting (mean) profile will be identical. Reviewing the equations introduced in Chapter 4.1.4, the two profiles for the unweighted and weighted means can be expressed, respectively, as:

$$\bar{S}^{BP}_{stack,\bar{w}} = \frac{\sum_{i=1}^{n}(S_i^{BP})}{n}, \tag{4.15}$$

and

$$\bar{S}^{BP}_{stack,w} = \frac{\sum_{i=1}^{n} w_i(S_i^{BP})}{\sum_{i=1}^{n} w_i} \tag{4.16}$$

where n is the stack size, S_i^{BP} the individual SBP per galaxy in the stack, w_i the weighting applied to the SBP.

For the unweighted case, inserting Equation 4.1 into Equation 4.15, and collecting constants results in:

$$\frac{\sum_{i=1}^{n}(S_i^{BP})}{n} = \alpha \sum_{i=1}^{n} \int I_{i,\nu} d\nu = \alpha \int \sum_{i=1}^{n} I_{i,\nu} d\nu \tag{4.17}$$

where $\alpha = \frac{1.024\times10^4 \cos i}{n\Omega_{ann}}$. Equation 4.17 demonstrates that the order of stacking is irrelevant; the same mean SBP is produced whether stacking at the cube or stacking at the SBP level in both a weighted and unweighted sense.

Once inverse-variance weighting is applied, the noisy profile of U11914 is further diminished from the stack, and the overall noise (both one-dimensionally and three-dimensionally) is lowered. The (weighted-) mean profiles differ significantly, as opposed to the unweighted case due to this reason. The weighting applied in the three-dimensional case is the inverse variance of the *cube* of the individual stacklings. In the one-dimensional case, the weighting is the inverse variance of the *Moment 0*. Thus, the application of a weighted mean is shown to reduce the impact of noisy data cubes in a stack, resulting in an added increase in SNR for the stack. This is exemplified by the consistently higher amplitude observed in the SBP of the 3D weighted mean over the 1D weighted mean across masking thresholds applied in Figure 7.9.

Of particular interest are the error bars for the 1D and 3D stack uncertainty profiles shaded in gray and green, respectively. Although the mean surface brightness for both the (unweighted) 1D and 3D stacks are identical, we do not see an expected identical uncertainty profile. As discussed in Chapter 4.1.4, for stacklings with identical noise properties it is expected that the 1D and 3D stack would yield identical variance, since the uncertainty would be dominated by the *random* noise. However, as demonstrated in Figure 4.2, the 1D uncertainty is dominated by the astrophysical variance of the stackling SBPs, as well as receiving a contribution from the random noise.

Thus, we conclude that the uncertainty in the SBP of the 3D stack as calculated from Equation 4.7 provides a better constraint on the SBP of the stack as opposed to the dispersion in the individual stackiling SBPs from the mean SBP.

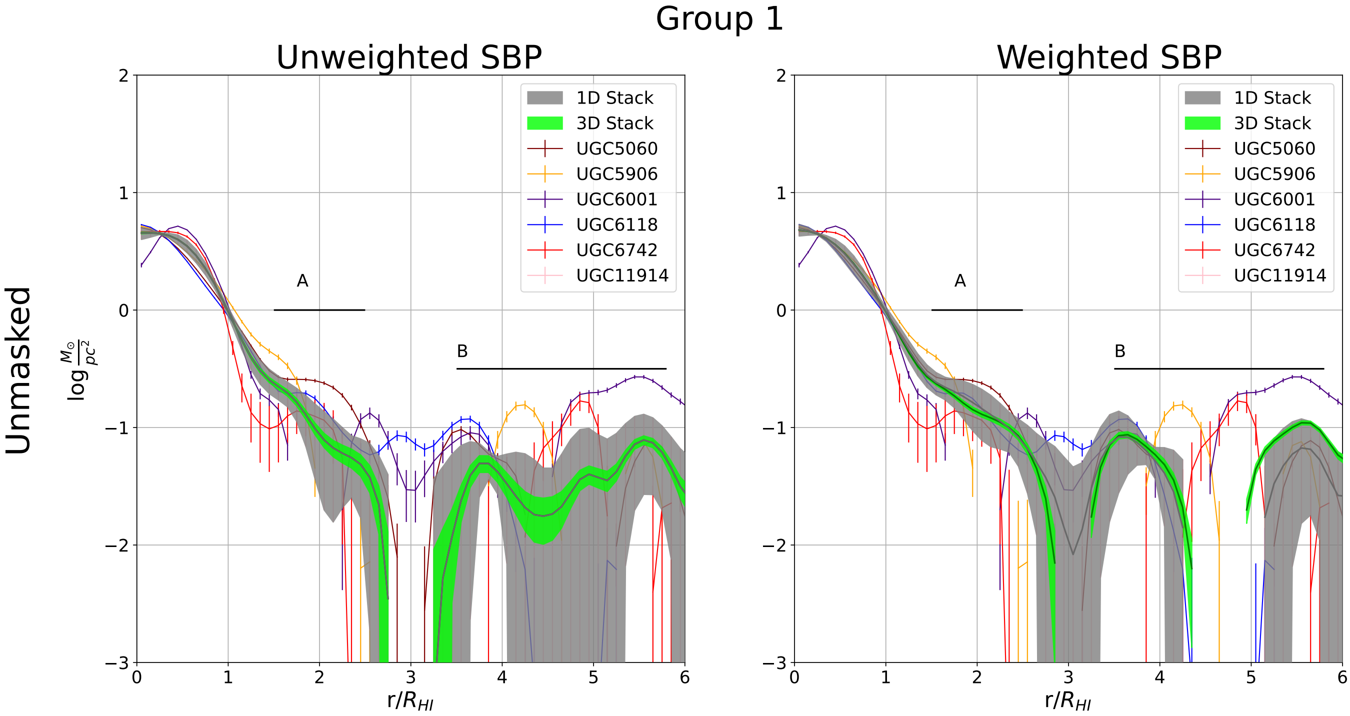


Figure 4.4: Stack Group 1: Low Mass with Low-Moderate inclination (i <70). U11914 sets the standard for the HI radius, while also serving as the fiducial for spectral extent with the largest rotation velocity. The largest weightings are assigned to U5060 (with U7704 closely following) due to their lowest average RMS across all channel maps, making them key for investigating any features. All galaxies are adjusted to a 60-degree inclination. Region A denotes the location of an initial break beyond $r = R_{HI}$ where the extended features seen in U5060 of Figure 7.14 is credited. Region B is characterised by the emission of a neighbouring galaxy U6128 within the SBP U6118 - demonstrating the need for the masking of background sources.

4.2.2 Uncovering Extended HI Emission in Surface Brightness Profiles

Beyond the initial break at $r = 3.25R_{HI}$, an increase in flux with radius occurs in the radial range of $3.25 \leq r \leq 4.25R_{HI}$. By enforcing several noise thresholds on the initial moment 0 map, we uncover the source of this feature. At these radii, U5060 and U6742 both have similar features. Exploring the moment 0 map of, a scattering of low-level flux ($< 500M_{\odot}arcsec^2pc^{-2}$) survives in the apertures within these radial ranges. In U5060 this occurs as a splutter of low column density clouds in the 3σ clipped Moment 0 in Figure 7.14, south-west of the galaxy in the rotated frame. The moment 0 of U6742 similar spluttering occurs, at flux densities of approximately $250M_{\odot}\mathrm{arcsec}^2\mathrm{pc}^{-2}$, almost evenly distributed about the annuli in the radial range of $3.25 \leq r \leq 4.25R_{HI}$.

A steep decline to negative flux follows this. This trend follows the profile of U5060, displaced largely by the influence of U7704. With an increase in sample size, the gap observed at radial range $4.5R_{HI} \leq r \leq 5.2R_{HI}$ within Region B will begin to diminish as the natural flux-variations around 0 will cancel.

At radial ranges $r \geq 5R_{HI}$, the presence of the nearby galaxy U6128 also within the cube centred on U6118 shows up as beyond a 1σ clipping across both the 3D- and 1D-stack SBPs. Since U6118 has the third highest weighting in the stack, coupled with the negative flux in U5060 and U6742 in those regions, U6118 dominates the stack profiles at these radii. A masking of the galaxy is required to remove this dominant feature from the subsequent SBPs. In Figure 7.14 we mark this galaxy with a purple ellipse.

4.3 Stack Group 3

Next, we stack all galaxies from 4.2 (galaxies of low-mass with low-to-moderate inclination ($i \leq 70°$)), including moderately-inclined systems this time around. Compared to Group 1, Group 3 has the addition of galaxies U7704 and U8271.

As in Group 1, the largest weighting in this stack is given to U5060 at 30% of the flux in the stack, however, the addition of U7704 now takes up 28% of the weighting. Furthermore, the fiducial HI-radius still belongs to U11914 and all other galaxies are scaled spatially to have this radius. Added to this, U11914 also has the largest spectral distribution, with the fiducial V_{rot} set accordingly. A further spatial scaling along the minor axes is applied to set all galaxies to a standard inclination of 60 degrees. A summary table is provided for the relevant galaxy properties used in the signal-alignment process in Table 4.2.

UGC	M_{HI} $10^9 M_\odot$	i °	R_{HI} "	V_{rot} kms^{-1}	X Scaling	Y Scaling	V Scaling	σ mJy $beam^{-1}$	Weighting
5060	0.7	24	60	34.89	0.09	1.88	4.69	3.04	0.11
5906	0.31	29	47	46.5	0.12	2.4	3.52	6.6	0.02
6118	0.84	24	89	80.78	0.06	1.27	2.02	6.28	0.03
6742	0.17	36	51	46.01	0.13	2.22	3.56	4.25	0.06
7704	0.34	60	58	37.28	0.28	1.95	4.39	3.16	0.1
8271	0.78	60	89	59.93	0.18	1.27	2.73	5.11	0.04
11914	0.69	31	113	163.55	0.05	1	1	18.58	0.003

Table 4.2: Group 3 Properties: Summary of relevant physical measures used for spectral alignment; from left to right, HI-mass, inclination, HI-radius, rotational-amplitude. Cube scaling values along the X-axis, Y-axis, and spectral-axis (V-axis). Finally, a measure of the noise within the cube following the application of scaling. In this stacked group, U11914 sets the standard for HI-radius, while also serving as the fiducial for spectral extent with the largest rotation velocity. The largest weightings are assigned to U5060 (with U7704 closely following) due to their lowest average RMS across all channel maps, making them key for investigating any features. All galaxies are adjusted to a 60-degree inclination

The SBPs of the stacklings overlap well, intersecting at $1 M_\odot \cdot pc^{-2}$ at $r = R_{HI}$ as expected. Looking at the moment 0 maps for the galaxies in this stack in Figure 7.16-7.17 we again have the neighbouring source of U6742 within the cube of U6118. This feature shows up within and corresponds to, the region marked as "Region B" occurring at $3.5 R_{HI} \geq r \geq 6.5 R_{HI}$. Features within the the additional galaxies U7704 and U8271 we notice extended HI in both galaxies off the major axes as measured by (Noordermeer et al., 2005). In U7704 this occurs northwest of the galaxy in the rotated sky, within the annuli at $1 R_{HI} \geq r \geq 2 R_{HI}$. In U8271, this is seen southeast of the rotated galaxy within apertures at radial location $2 R_{HI} \geq r \geq 3 R_{HI}$. We mark these features with a black circle.

Further investigation of the inclination of these two galaxies is required. While inclination of U7704 and U8271 are quoted as 60° in Table 2 of (Noordermeer et al., 2005), the WHISP optical overview quote values of $i = 47°$ and 70°, respectively. If true, this would disqualify U8271 as a moderately inclined galaxy, and require a move to Group 10. Despite the possible erroneous inclination measures, U7704 still shows an elongation off the expected North-South line. Suggesting that careful inspection of the position-angle is required. According to the Hyper Lyon-Meudon Extragalactic Database (HyperLEDA, Makarov et al. (2014)), the position angle is measured to be $PA = 161°$[1]. In contrast, Noordermeer et al. (2005) obtain a value of 37°. The possible bad position angle is further confirmed when an investigation of the PV diagram of U7704 in Figure 7.7 is made. There is no clear rollover at large radii, suggesting that the extracted PVD is not a major axis PVD. As a result, contributions from U7704, constituting 28% of the weighting need to be ignored.

---

[1] http://leda.univ-lyon1.fr/

4.3.1 Comparison of One-Dimensional and Three-Dimensional Stacking

Again, the observed divergence between the weighted 3D stack's SBP and the weighted mean of the stackling SBPs are minimal. Apart from the divergence within the region bounded by $r \leq 0.5R_{HI}$, a near-identical overlap between the 1D and 3D stack is seen $0.5R_{HI} \leq r \leq 2.5R_{HI}$. An interesting observation follows. The 1D stack exceeds the 3D stack in amplitude briefly between $2.50R_{HI} \leq r \leq 3.25R_{HI}$. We attribute this apparent improvement to the absence of negative flux contribution in this region from U5906 and U11914 which are absent from the log-scale SBP.

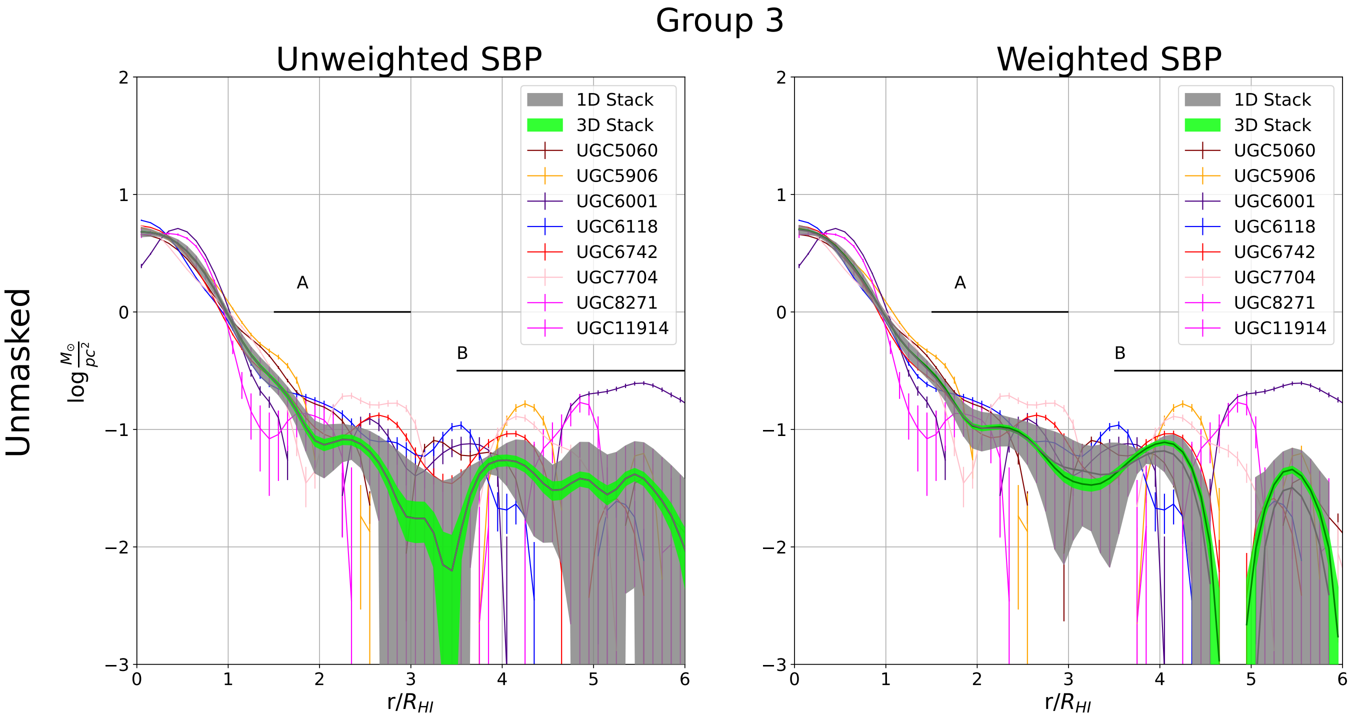


Figure 4.5: Stack Group 3: Low Mass with Low-Moderate inclination (i <70). No new HI is uncovered, and the increase in flux with radius occurring in the radial range of $3.25 \leq r \leq 4.25R_{HI}$ is attributed to U5060. Region A demonstrates how the inclusion of a misaligned galaxy (such as U7704) will appear as extraplanar gas in the SBP - strongly influencing the 1D and 3D stack SBP in the process. Region B demonstrates the importance of masking neighbouring sources within the data cube, as their influence in the SBP of interest leads to the contamination of the radial flux distribution.

4.3.2 Uncovering Extended HI Emission in Surface-Brightness Profiles

Beyond the initial break at $r = 3.25R_{HI}$, an increase in flux with radius occurs in the radial range of $3.25 \leq r \leq 4.25R_{HI}$. This is followed by a steep decline to negative flux. This trend follows the profile of U5060, displaced largely by the influence of U7704. The gap observed at radial range $4.5R_{HI} \leq r \leq 5.2R_{HI}$ will begin to diminish as the natural flux-variations will average around 0 following an increase in sample size. In the current case, the profiles average to a negative flux, which disappears in the log scale.

In U6118 a feature is seen extending beyond radial range $r \geq 4.8R_{HI}$. Inspecting the moment 0 map in Figure 7.16, we see a companion galaxy in between apertures at $4R_{HI} \leq r \leq 6R_{HI}$, lying to the 'west' of U6118. We emphasise that *west* is not true west on the sky, as our spatial coordinates are rotated.

When investigating the effects of applying masking thresholds, see Figure 7.10 we find that independent of the weighting applied, the 3D stack survives to a greater radial extent.

4.4 Stack Group 6

Group 6 is the only bin of galaxies with moderate mass ($10^{8.9} \leq M_{HI} \leq 10^{9.75} M_{\odot}$). Summarised in Table 4.3, we see the standard for HI-radius is U6787. Similarly, U6787 is the fiducial for spectral extent - having the largest V_{rot}.

Also in Table 4.3 the highest weighting in the stack is given to U12043 (with U1310 following closely) - having the lowest average RMS across all the channel maps. As such, any extra-radial features appearing in the stack should be investigated in these galaxies.

As usual, all galaxies are re-projected to an inclination of 60 degrees.

UGC	M_{HI} $10^9 M_{\odot}$	i °	R_{HI} "	V_{rot} kms^{-1}	X Scaling	Y Scaling	V Scaling	σ mJy $beam^{-1}$	Weighting
798	4.08	40	72	101.51	0.24	3.56	2.16	6.86	0.02
1310	2.33	47	50	61.64	0.42	5.12	3.56	2.16	0.21
6786	3.93	65	143	202.29	0.34	1.79	1.08	2.59	0.15
6787	3.96	67	256	219.25	0.21	1	1	4.07	0.06
8863	2.52	51	148	176.49	0.16	1.73	1.24	3.92	0.07
12043	1.18	67	143	167.29	0.38	1.79	1.31	2.14	0.22

Table 4.3: Group 6 Properties: Summary of relevant physical measures used for spectral alignment; from left to right, HI-mass, inclination, HI-radius, rotational-amplitude. Cube scaling values along the X-axis, Y-axis, and spectral-axis (V-axis). Finally, a measure of the noise within the cube following the application of scaling. In this stacked group, U6787 sets the standard for HI-radius, while also serving as the fiducial for spectral extent with the largest rotation velocity (V_{rot}). The largest weightings are assigned to U12043 (with U1310 closely following) due to their lowest average RMS across all channel maps, making them key for investigating any features beyond $r = R_{HI}$. All galaxies are adjusted to a 60-degree inclination.

In Figure 4.7, the individual stackling exhibit similar flux variations beyond $r = 2R_{HI}$. What is concerning is that U1310 and U798 have significant regions of negative flux, specifically at intermediate radial ranges of $1.2R_{HI} \leq r \leq 3.5R_{HI}$. In the case of U1310, specifically, the high weighting mentioned earlier will prove problematic as we search for extra-radial emission within the stacks since 28% of the contributing signal is lost.

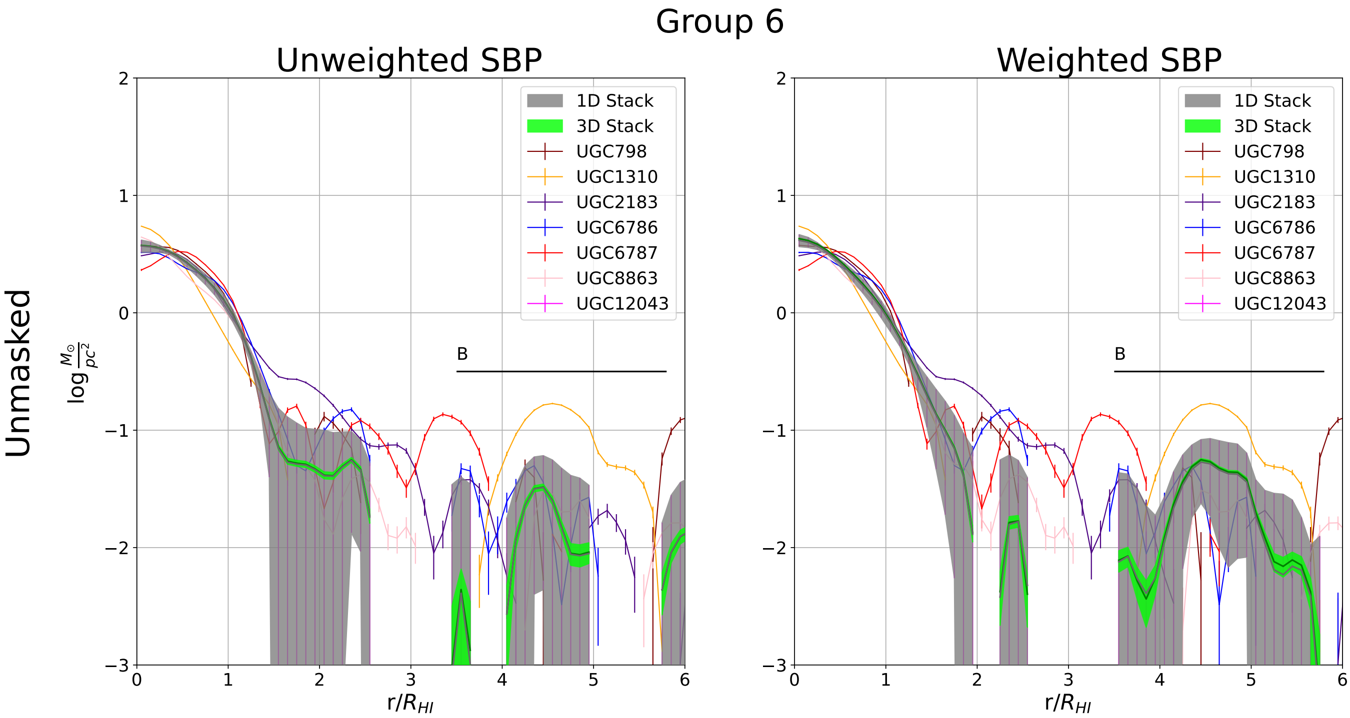


Figure 4.6: Stack Group 6: Low Mass with Low-Moderate inclination (i <70). A "bed-sheet- ghost" feature marked by Region B (centred at $r \approx 4.5R_{HI}$) is attributed to the SBP of U1310.

4.4.1 Comparison of One-Dimensional and Three-Dimensional Stacking

As higher and higher thresholds are imposed on both contributing signals, say 2σ, the increased SNR of the 3D stack leads to less signal being clipped from the resulting SBP. A "bed-sheet-ghost" feature in Region B (centred at $r \approx 4.5R_{HI}$) is initially visible for both 3D and 1D stacks, and is entirely attributable to the low-mass HI clouds South of the rotated plane of U1310 (with contributing annulii highlighted yellow in Figure 7.18), contributing 28% of the flux in the weighted stack. Once a sigma-clipping of $2\sigma_{gal}$ is applied the signal from all individual galaxies in the stack is eliminated - and subsequently the 1D stack. However, for the 3D stack, the same threshold of $2\sigma_{stack}$ has minimal effect on the resulting SBP of the stack.

Intuitively, then, it must mean that $\sigma_{gal} > \sigma_{stack}$, since the profiles are not significantly different. In other words, stacking in 3D decreases the overall noise fluctuations within the data cube.

4.4.2 Uncovering Extended HI Emission in Surface-Brightness Profiles

Analysis at small radii in Group 6 shows little signal divergence, except for the inner-most radial bins ($r \leq 0.5R_{HI}$). Immediate impressions at higher radii are that of the "bed-sheet-ghost" feature at $3.5R_{HI} \leq r \leq 5.5R_{HI}$. However, an examination of the SBP of U1310 (accounting for nearly 30% of the weighting) shows a similar feature at this radial range.

Visible beyond a $1\sigma_{gal}$ clipping, the offset from the 3D and 1D trend is due to the effects of the non-existence of flux within the galaxies contributing the remaining 70% of the stack.

However, when comparing the amplitudes of the 1D and 3D SBPs within this region, we still see that the 3D stack lies mostly above the 1D stack. We attribute this result to the increased SNR - very likely due to the suppression of the noise prior to the extraction of an SBP.

We do not observe any new signals appearing within this stacked group of moderately massed galaxies.

4.5 Stack Group 9

Due to an overlooked bug in the binning code, see Chapter 2.5, Group 9 contains galaxies with mass ($10^{9.75}M_{\odot} \leq M_{HI} \leq 10^{10}M_{\odot}$).

As before, we present a summary of the relevant stacking information used by our pipeline in Table 4.4.

With our small sample size, most features in the weighted stack (both 3D and 1D) will be contributions from U7989 (79% weighting), with the next greatest contribution coming from U89 at 10% weighting. The smallest from U508 (0.17% weighting). The most extended galaxy in the stack (by HI-radius) is U7989, and as such is the spatial fiducial.

The largest spectral distribution (V_{rot}) belongs to U12815. Unsurprisingly, U12815 also has one of the largest HI-masses in the group. As a result, all other galaxies were stretched spectrally to have this value.

UGC	M_{HI} $10^9 M_{\odot}$	i °	R_{HI} "	V_{rot} kms^{-1}	X Scaling	Y Scaling	V Scaling	σ mJy $beam^{-1}$	Weighting
89	8.63	50	77	171.94	0.45	4.94	1.54	12.98	0.006
232	7.95	48	82	113.26	0.39	4.63	2.34	19.63	0.003
508	7.4	25	94	205.08	0.2	4.04	1.29	85.7	0.0001
7989	9.78	51	380	44.73	0.1	1	5.93	4.68	0.046
9644	7.16	19	53	88.23	0.32	7.17	3.01	20.87	0.002
12815	9.36	61	103	265.35	0.56	3.69	1	43.57	0.001

Table 4.4: Group 9 Properties: Summary of relevant physical measures used for spectral alignment, from left to right, HI-mass, inclination, HI-radius, rotational-amplitude. Cube scaling values along the X-axis, Y-axis, and spectral-axis (V-axis). Finally, a measure of the noise within the cube following the application of scaling. All the galaxies in this group were adjusted to have the same tilt of 60 degrees. Observing the average RMS noise across all channel maps (σ), most of the features we see in the stacked data (both 3D and 1D) come from U7989, with U508 contributing the least. U7989 is the largest galaxy in terms of its size in the data, so we use it as a reference for the spatial measurements. U12815 has the widest range of speeds (Vrot) in its spectrum, and it also has one of the highest amounts of hydrogen gas (HI mass) in the group. Therefore, we adjusted the spectra of the other galaxies to match the same range as U12815.

4.5.1 Comparison of One-Dimensional and Three-Dimensional Stacking

Analysis of the Figure 4.7 comparing the 1D and 3D stack results shows marginal divergence in SBP amplitude in the innermost radii. This corresponds to the expected dampening of noise within the voxels of the stack, prior to extracting a SBP, and subsequently, less signal is lost due to the effects of negative flux variations within noisy voxels. This effect is amplified in the log scale, demonstrating the difference in amplitude across radial-offset, with an increase in SNR achieved by the 3D stacked data than the stack of the one-dimensional SBP.

Comparing the effects of sigma-clipping in Figure 7.12, as discussed in Chapter 4.1.5, the increased SNR (although marginal) is immediately apparent.

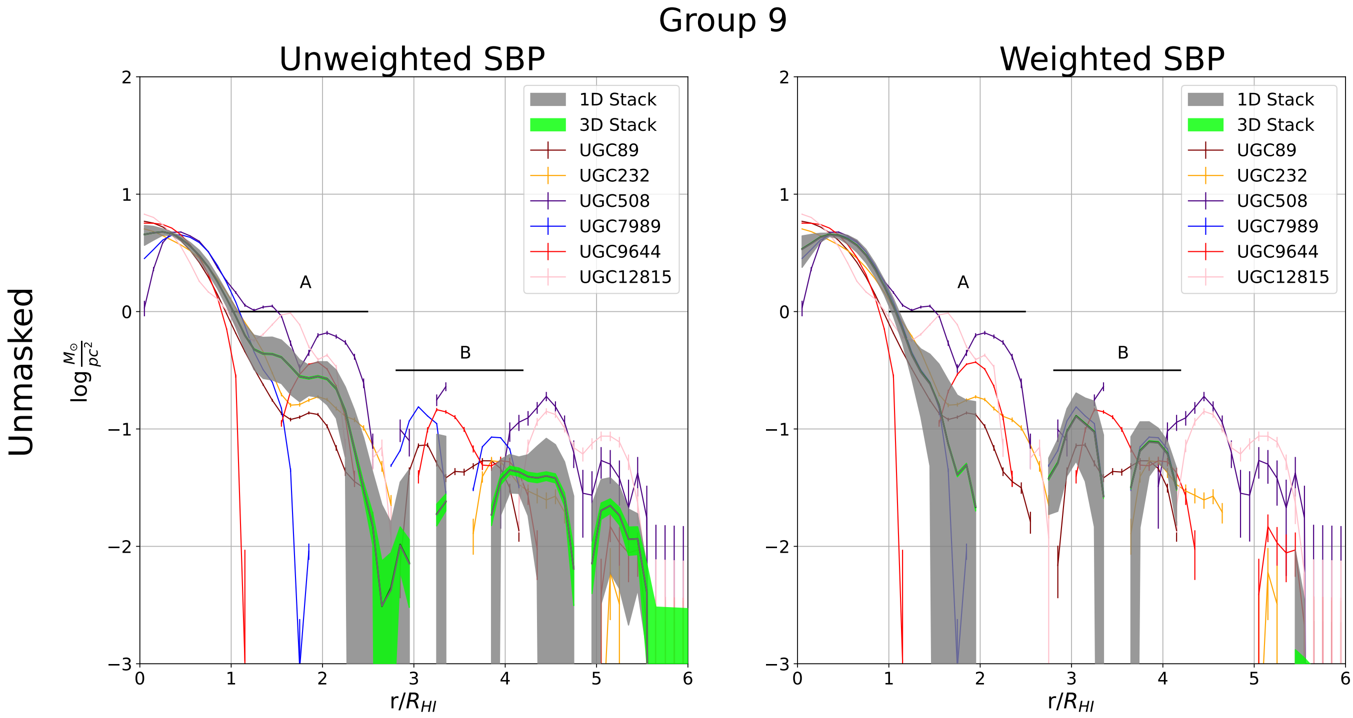


Figure 4.7: Stack Group 9: Low Mass with Low-Moderate inclination (i <70). Comparing 1D and 3D stack results, we observe slight amplitude divergence in the innermost radii in 3D. Noise reduction in stack voxels before extracting Surface Brightness Profiles (SBPs) minimizes signal loss from negative flux variations when extracting an SBP. The log-scale demonstrates higher SNR in 3D stack as sigma-clipping is applied (see Figure 7.12). Features from the dominant galaxy in the stack, U7989, emerge in the stack as extended features within Region A, with a prominent "double-bump" break emerging at Region B.

4.5.2 Uncovering Extended HI Emission in Surface-Brightness Profiles

Returning to our stack sample of high-mass galaxies with low to moderate inclinations, and taking advantage of the improved SNR, we uncover several features in the SBP of the Group 9 stack. Both in 1D and 3D.

As shown in Figure 7.12, the dominant galaxy in this stack is U7989, with the better quality data, demonstrated by smaller error bars, dominates Region A ($1R_{HI} \le r \le 2.5R_{HI}$). Here, the prevalence of extended features within the annuli in these ranges, as the flux levels approach the background noise level, closely follows the SBP of U7989, while minimal contribution is received by the noisier data. Thus, our choice of inverse-variance weighting is vindicated – leading to fewer spurious detections due to noisy data.

Extending our investigation to larger radii, two prominent features begin to emerge as we apply several signal thresholds as discussed in Chapter 4.1. Region B ($2.8R_{HI} \leq r \leq 4.2R_{HI}$) contains the existence of a “double-bump” break that emerges quite strongly – tracing similar features in U89 and U7989.

Looking at Figure 7.20, we observe the contributing HI features in U7989, further modelling would be required to accurately determine the source of the gas and its relation to the U7989 system.

In order to determine the relationship of this feature to the system of U9644, spectral analysis and 3D modelling must be performed.

4.6 Stack Group 10

Our final sub-sample of galaxies is classified as highly-inclined systems ($i \geq 70°$).

As before, we present a summary of the relevant stacking information used by our pipeline in Table 4.5. Unlike the previous stack groups, the galaxies in this bin are re-projected to an inclination of 90 degrees due to the edge-on galaxies present.

As tabulated in Table 4.5, most features in the stacked SBP (both 3D and 1D) will be contributions from U5690 and U6283, with a stack weighting of 44% and 26%, respectively.

The largest spatial (R_{HI}) and spectral extent (V_{rot}) belongs to U4605 which also has the largest HI-mass in the group. As a result, all other galaxies were stretched spectrally to have this value. Furthermore, all galaxies are scaled spatially to have the same apparent HI-size (R_{HI}) as U4605.

UGC	M_{HI} $10^9 M_\odot$	i °	R_{HI} "	V_{rot} kms^{-1}	X Scaling	Y Scaling	V Scaling	σ mJy $beam^{-1}$	Weighting
2141	2.3	90	155	115.61	1.79	1.79	1.63	9.67	0.011
4605	5.99	90	277	187.95	1	1	1	6.14	0.027
5351	0.99	70	61	129.56	1.19	4.54	1.45	12.31	0.007
5960	0.75	77	109	85.78	1.15	2.54	2.19	3.85	0.067
6283	2	90	204	97.17	1.36	1.36	1.93	5.05	0.039
8699	3.26	72	100	175.45	0.84	2.77	1.07	33.17	0.001

Table 4.5: Group 10 Properties: Summary of relevant physical measures used for spectral alignment, from left to right, HI-mass, inclination, HI-radius, rotational-amplitude. Cube scaling values along the X-axis, Y-axis, and spectral-axis (V-axis). Finally, a measure of the noise within the cube following the application of scaling.

Compared to the SBPs extracted from the elliptical apertures in the previous groups, these SBPs are derived from rectangular apertures. See Figure 4.1.

4.6.1 Comparison of One-Dimensional and Three-Dimensional Stacking

Analysis of the Figure 4.8 comparing the unmasked 1D and 3D stack results shows no improvement in amplitude in the innermost radii. As described in Equation 4.17, the order of stacking does not matter in the unweighted case. In Figure 4.6, the SBP of the 1D and 3D stacks are identical at all radii for the unweighted means.

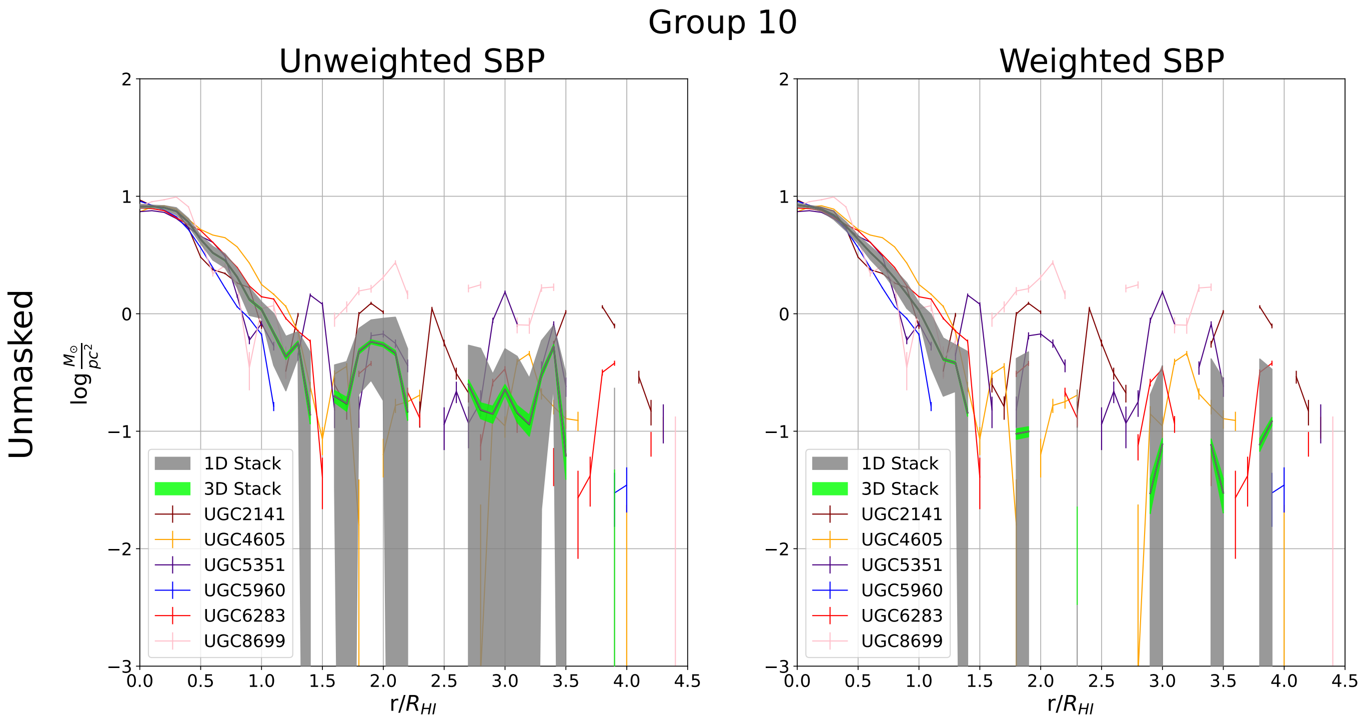


Figure 4.8: Stack Group 10: High inclination (i >70).

4.6.2 Uncovering Extended HI Emission in Surface-Brightness Profiles

In the unweighted mean, the presence of a bump approaching $r = 2R_{HI}$ correlates with the profiles of U5351, U8699, and U2141. In Figure 7.22 there occurs some scattered low flux HI clouds extending along the midplanes of U5351 and U2141. This feature is absent in the weighted scheme, replaced instead by a gently rising profile at a lower amplitude. We attribute this diminished profile to an absence in signal measured for the highest weighted galaxies in the stack in this region. The contributions of U5960 and U6283 constitute 70% of the weighting, meaning the flux in this region only has contributions from the remaining 30% of the stack group. Since U4605 constitutes 18% of the remaining 30% in the stack, it is not surprising that the gently rising profile at $r = 2R_{HI}$ in the weighted stack resembles the profile of U4605 at this position, as denoted by the orange curve.

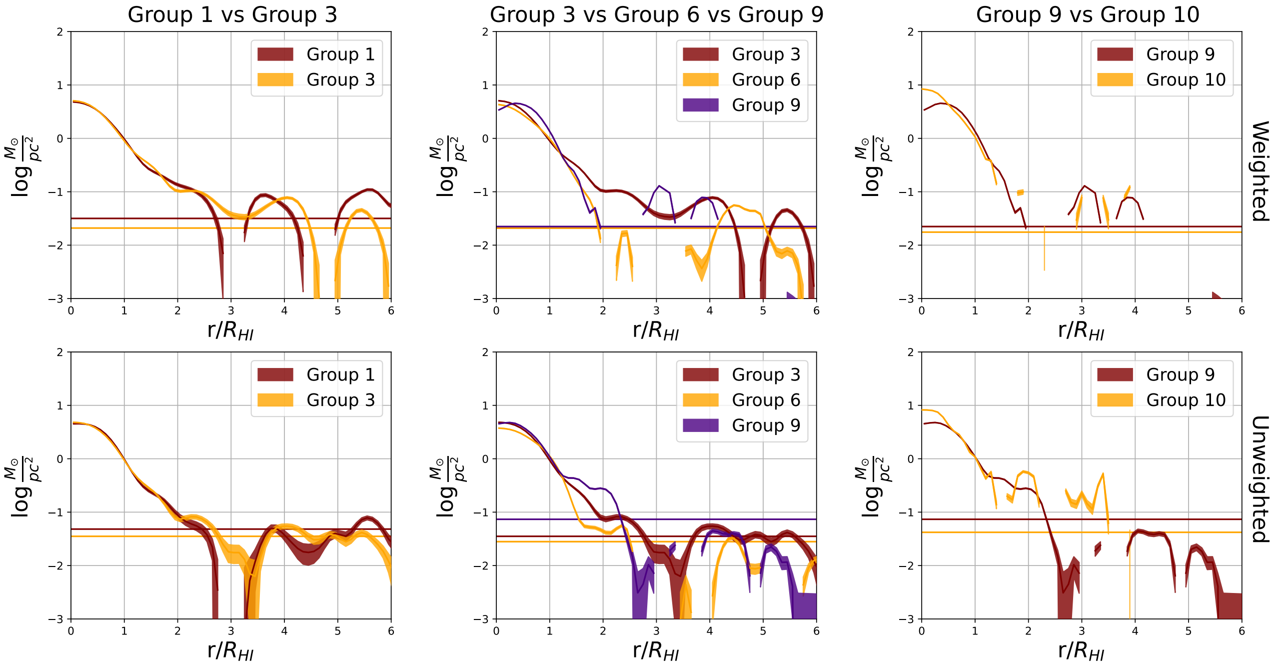


Figure 4.9: Comparison of surface-brightness profiles for different stack groups, with and without weighting. Column 1 compares the low-mass regime and the effect of the correction of low- and intermediate-inclination to a 60° fiducial inclination. Column 2 compares the intermediate mass regimes and the effect of a shift to a 60° fiducial inclination. Finally, the third column compares the effect of a correction to a 90° fiducial inclination. Horizontal lines indicate the $1\sigma_{M_0}$ noise level (per pixel), calculated via Equation 4.8 and converted to units of log mass surface-density, in the stacked cubes' moment-zero maps.

4.7 Comparing Stacking Bins

In Figure 4.9, we compare the surface-brightness profiles for different stack groups to illustrate mass and inclination dependencies.

Immediately apparent across all the stacking groups, is the order of magnitude decline (1 to 0.1) in the SBP in log-scale at radial range $R_{HI} \leq r \leq 2R_{HI}$. With the slope of the decrease seemingly correlating with mass. In the context of an optically based stacking profile, this may guide efforts to exploit the central light concentration (as we present in Chapter 7.5.2) to predict the HI surface brightness radial profile, and by extension, the 3D HI distribution. This drop coincides with the idea of a "Universal" behaviour exhibited by HI discs - an exponentially declining profile from 0.75 RHI to 1.3RHI. In comparison, our sample of galaxies (grouped by mass and inclination) which have subsequently been corrected to a face-on inclination, have SBPs exhibiting a similar exponential drop at these radii – obeying this Universal HI profile (Wang et al., 2014).

When aligning galaxies in three dimensions, subtleties in projection become critical. These projection effects are important considerations when searching for faint extraplanar gas in galaxies.

When investigating the effects of the inclination corrections to a fiducial of 60°, we note little difference between the low (Group 1) and low-moderate (Group 3) inclination groups at any radius for the low HI mass bin. The small bin sizes are likely a reason for this, as Group 3 shares all galaxies with Group 1, with the addition of U7704 and U8271. It is with these additional galaxies that we demonstrate the improved SNR in the form of the per-pixel $1\sigma_{M_0}$ noise level represented by the solid horizontal lines.

The horizontal lines denote the RMS level per-pixel of the moment 0 map from which the SBPs are extracted and propagated using Equation 4.8. The moment 0 RMS value is then converted to units of mass-density using Equation 4.2. We note that the per-pixel RMS within the moment 0 map corresponds with the observed plateau (or break) beyond $r = 2R_{HI}$ across all stack groups both weighted and unweighted. The spread in RMS mass-densities observed suggests a noise floor between -1.13dex to -1.55dex for the unweighted stacks, and -1.49dex to -1.75dex for the weighted stacks.

In terms of column densities, these correspond to an estimated column density of $9 \times 10^{18} cm^{-2}$ and $3.5 \times 10^{18} cm^{-2}$ for the unweighted stacks, respectively, while the weighted stack achieves a probing in the range of $3.87 \times 10^{18} cm^{-2}$ to $2.12 \times 10^{18} cm^{-2}$. In comparison to the column densities obtained from (Noordermeer et al., 2005) summarised in Table 7.5 push below the $1 \times 10^{19} cm^{-2}$ column density limit shared by all but one galaxy, namely U8271.

As a result of the weighting scheme applied, U7704 and U8271 are given weights of 27.55% and 11.02% of the total stack, respectively, leading to significant changes in the stacked profile. Among the improvements is the minimization of the negative fluxes which appear as gaps in these log surface-brightness profiles.

At high HI mass, a significant difference is observed at small radii. The high-inclination stack (Group 10) has a steeper rise at the very centre. We suspect that the profile of the low-moderate inclination stack (Group 9) is a more accurate assessment of the radial profile at smaller radii, as HI deficiency in the central regions is largely expected in high-mass galaxies. While the origin of the HI-holes is still an open question, one of the proposed mechanisms is the so-called supernova (SN) hypothesis, according to which a gas void is left behind following high-energy supernovae (SNe) activity. The increased prevalence of these activities in high-density environments (Rhode et al., 1999) may contribute to this phenomenon. When such high-mass galaxies are observed edge-on (as in the case of Group 10), that hole gets obscured by the projection of gas at larger radii along the minor axis.

Chapter 5

Discussion

5.1 Inherent Experiment Challenges

5.1.1 Small Sample Size

Of primary concern to any stacking project is sample size. While typical stacking experiments have stackling sample sizes numbering in the thousands, our largest stack has 9 stacklings. Several experimental constraints have resulted in a stacking sample of 33 galaxies, from an initial radio sample of 68 galaxies from Noordermeer et al. (2005). In addition to survey constraints, we perform a visual inspection, eliminating asymmetric or lopsided galaxies from the sample. We provide a summary in Figure 2.3. However, adopting a more sophisticated rotation-curve model may prove to limit the number of galaxies excluded by virtue of atypical HI distributions.

In addition to broadening the scope of our selection criteria, starting with a larger initial survey would also yield a larger stacking sample. The Widefield ASKAP L-band Legacy All-sky Blind surveY (or WALLABY) is one of a number of surveys that are now running on the Australian SKA Pathfinder (ASKAP, Koribalski et al. (2020b)). With spatial resolution similar to WHISP at 30 arcseconds, the expected survey size of 500,000 galaxies makes WALLABY a suitable test ground for future iterations of *Cube Stacker*.

5.1.2 WHISP Dataset

Despite consisting of spatially resolved observations of the largest galaxies ($D_{M_B} > 1.5$ arcmin) within 160 Mpc, the data-reduction, and specifically CLEANing step, when constructing the imaging cubes provides a restriction to stacking experiments.

As referenced in Chapter 2.2, the CLEANing algorithm used in the WHISP pipeline to the algorithm iteratively identifies and removes point sources (Högbom, 1974), down to 0.5 times the RMS per channel-map (van der Hulst et al., 2001). As a consequence, a lot of faint structures are not added to the cleaned image during the deconvolution step.

The choice of cleaning threshold imposes an upper limit on expected returns from any stacking experiment and care should be taken when selecting a survey (such as WHISP) to perform stacking.

Furthermore, and in the case of The HI Nearby Galaxy Survey (THINGS, Walter et al. (2008b)), a subtle yet critical limitation of the CLEAN algorithm occurs when producing the restored image. Two flux-scales are combined: a residual map with (with units $Jy \cdot beam_{dirty}^{-1}$) and a clean component map (with units $Jy \cdot beam_{clean}^{-1}$), leading to an overestimation of the flux in the residual as the clean beam is inherently smaller than the dirty beam.

THINGS try to mitigate the overestimation of the total flux which would be produced by the clean beam (increased sampling rate) by applying a correction factor. This acts to ensure accurate flux measurements in regions expected to contain high signal. Consequently, noisier regions are also subject to this flux attenuation, leading to flux values not reflecting the true noise properties and flux variations in regions of interest to stacking projects.

As a result, careful consideration of the HI data reduction process of the dataset needs to be made. For experiments which are heavily reliant on the noise properties (such as the application of a 3D stacking algorithm), data cubes where Högbom (1974) CLEANing has been performed will result in a larger uncertainty in the total HI mass measured.

In our case, WHISP does not make mention of the flux-scaling procedure performed in THINGS. In addition, their approach of masking regions beyond the search areas, even if this scaling were applied, would not affect the noise properties of the masked regions. However, the problem of faint emission being clipped in the original data cubes still remains in the Noordermeer et al. (2005) survey.

A possible solution is to employ more robust deconvolution techniques such as multi-scale CLEAN (MSCLEAN) developed by Cornwell (2008). According to Rich et al. (2008), MSCLEAN improves on the classical clean algorithm by modelling the sky as containing emission at differing spatial scales.

5.2 Further Work

5.2.1 Signal Alignment

The results for the envelope-fitting and subsequent rotation curve fitting performed for our sample of 33 galaxies from the Noordermeer et al. (2005) WHISP survey of late-type galaxies are summarised in Table 7.1. As discussed in Chapter 3.2.3, we develop and construct a pipeline (see Figure 3.5) which runs parallel to the main Cube Stacker pipeline. We do this to map the two-dimensional distribution of the galaxies in the cubelets.

Outlined in Chapter 3.13, this distribution is later used as a means to standardise and align all the spectra in their respective data cubes. To do this we employ the use of a simple Tanh function to model this distribution, making use of the parameter V_{rot} to standardise the spectral extent of the galaxies. While existing measures of spectral scale such as W50 and W20 may provide a more direct means to align the galaxy spectra, this work explores the possibility of a model which correlates well with optical measures, to align the HI spectra.

In Figures 7.4 - 7.8, we present a compilation of PVDs for all galaxies tabulated in Table 7.1. Overlaid are the envelopes obtained from the envelope-fitting routine. The solid curve models this distribution of envelopes as a tanh function fit to the approaching and receding ends, simultaneously. The dashed curve is a Tanh model fit to the set of envelopes on the approaching side alone, while the dotted curve is a third Tanh model fit to the receding end. We adopt the area enclosed by the dotted- and dashed-curve as an estimate of the error on the rotational amplitude (V_{rot}) as summarised in Table 7.1.

The success of this stacking routine is shown by the degree of alignment of the scaled Tanh curves of the individual stacklings. Figure 5.1 presents a compilation of all rotation curves (solid curves in Figures 7.4 - 7.8), stretched to the same spatial-extent (R_{HI}) and spectral-extent (V_{rot}).

Immediately apparent is the misalignment in the inner regions of the rotation curves. In this initial work, we have opted to match for spatial-extent (R_{HI}) and spectral-extent (V_{rot}) alone. In the search for extended rotation curves, disregarding the rising rates (hrot) of the tanh-curve in the inner-most regions is sufficient, since the mismatch occurs outside the radial ranges of our interest (r>RHI). However, due to the prevalence of projection effects in HI astronomy, and 3D alignment in general, this becomes an important consideration particularly as inclinations tend to face-on.

Assessing the alignment of the Group 1 galaxies, we see this misalignment of the interior radii (r<RHI). Following visual inspection, the Tanh curves all converge spatially at around r = 120 arcsecs, agreeing with the fiducial spatial extent of RHI = 113 arcsec of galaxy U11914.

The correlation between the rotation curve shape and the spread in the three rotation curve fits is particularly interesting. For instance, U5060 (Figure 7.5) and U6742 (Figure 7.6) both have a rising rotation curve shape, characterised by the absence of a clear rollover, resulting in a greater spread between the simultaneous model fit (bold line) and the approaching- and receding fits. As shown in Table 7.1, this corresponds to the large errors in spectral amplitude (Vrot), furthering the need for a more robust signal-alignment model. The current iteration of Cube Stacker is thus susceptible to spectral smearing as different rotation velocities are co-added, leading to the potential of any misaligned planar-gas signal being characterised as anomalous (e.g. extraplanar) gas.

Similarly, Group 3 galaxies are identical to Group 1 with the addition of U7704 and U8271. While the envelopes obtained for U7704 (Figure 7.7) are well described by the Tanh model, U8271 (Figure 7.7) performs poorly. Visual inspection indicates a declining rotation curve, which after co-addition may present as a "beard" feature as described in Figure 1.2. Despite the inclination already being at the fiducial 60 degrees, with no reprojection required, spectral smearing would still occur. Thus, a multi-parameter model must then form the basis for signal alignment in future iterations of the Cube-Stacker pipeline.

Galaxies in Group 6, with moderate mass and low-moderate inclination (i <70), have inclinations distributed closely to the fiducial inclination of 60 degrees – requiring little reprojection. Thus, spectral-smearing is minimal.

Despite the little reprojection required, the Tanh model only describes the envelope distributions of U6786. While U6787 (Figure 7.7) and U8863 (Figure 7.7) show declining profiles, U798 (Figure 7.4) and U6787 show a rising profile. Thus, despite the minimal smearing effects due to inclination reprojection, a mismatch in rotation curve shapes is present in Group 6. Detection of any 'beard' features in the stack must then be regarded with suspicion.

In the stack, the collection of model fits shows similar rising rates (hrot) for all galaxies except U1310.

Group 9 consists of all galaxies with high Mass and low-moderate inclination (i <70). Similar to Group 6, the galaxies in Group 9 (with the exception of U508 and U9644) have inclinations distributed around the fiducial i = 60 degrees. This leads to reprojection smearing effects being minimal. Despite this, the Tanh model used as the basis for signal alignment performs poorly for all but one galaxy (U89) in the sample. U232 (Figure 7.4), U7989 (Figure 7.6), and U12815 (Figure 7.7) have declining rotation curve shapes as denoted by the envelopes fit. Meanwhile, U508 (Figure 7.4), U9644 (Figure 7.7) have increasing rotation curve shapes.

Finally, Group 10 consists of all galaxies with high inclination (i>70 degrees). In this instance, the fiducial inclination is set to i = 90 degrees, where half the galaxies in the sample (U2141, U2183, U4605, and U6283) are already edge-on. Thus, U5351, U5960, U8699, and U12713 are all reprojected and any extraplanar features found in the stack should be cross-referenced with these galaxies.

In terms of the performance of the Tanh model, the envelopes for U8699 (Figure 7.6) are the only set well described by a Tanh curve. The rising curves of U2141 (Figure 7.5), U5960 (Figure 7.6), and U6283 (Figure 7.6), and declining curves of U2183 (Figure 7.5) and U4605 (Figure 7.5) are best described with the multi-parameter function (MPF). Assessment of the stack, and the scaled model-fits, we see the curves overlap at V = 190km/s, in agreement with the sample fiducial of Vamp =187.95 km/s. With the exception of U6283 and U5351, the rising rates of the tanh-curves are comparable.

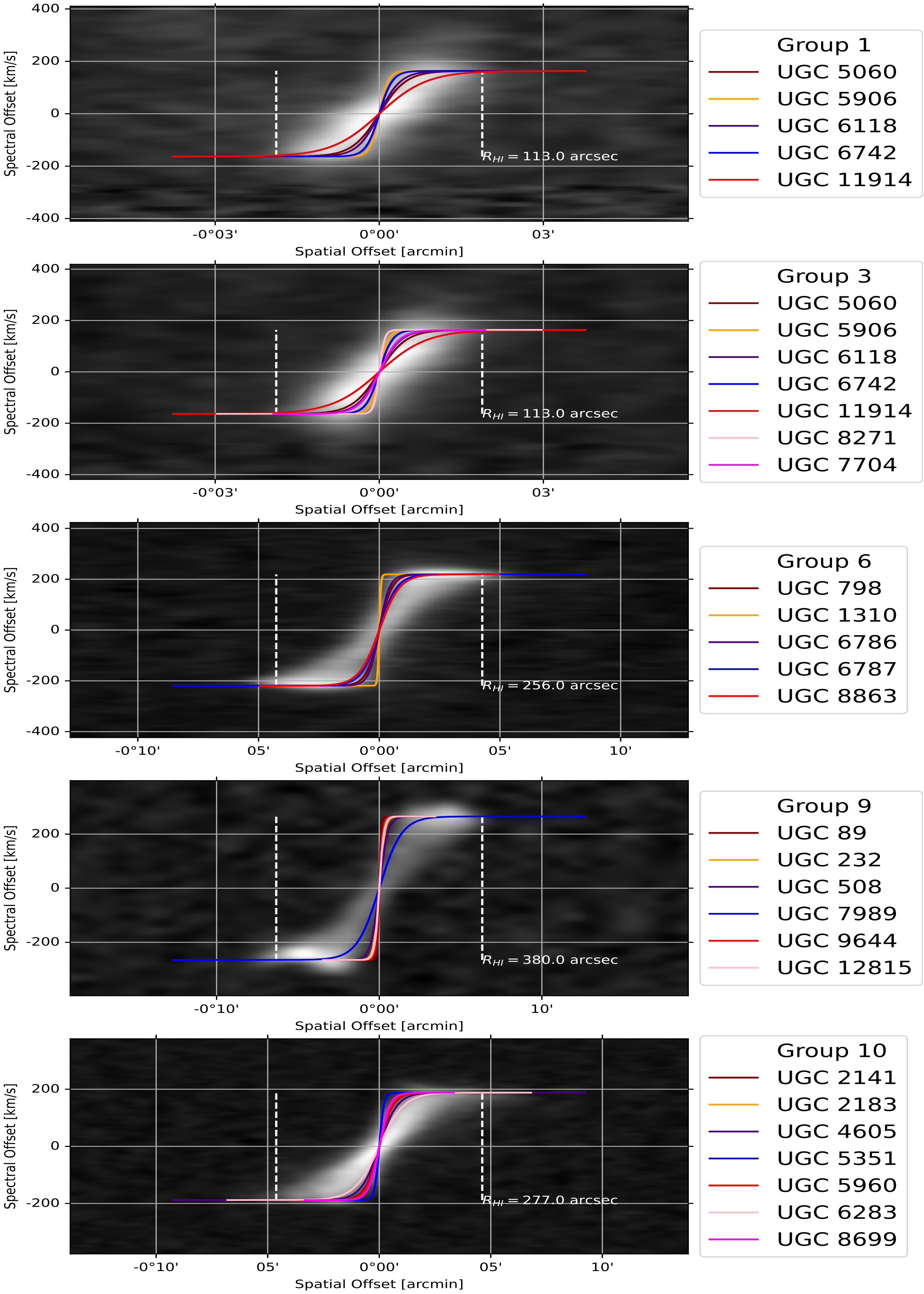


Figure 5.1: Demonstration of the scaled Tanh model rotation curves of the stacklings atop the weighted stack's PVD. Assessment of the degree of alignment of the scaled rotation curves of the stacklings provides insight into the success of the *Cube Stacker* algorithm's ability to align galaxy signals in two dimensions. A feature consistent across all of the stacking groups is the misalignment of the rotation curves in the inner regions ($r \leq R_{HI}$) - demonstrating the absence of a means to standardise rising rates of galaxy PVDs. Where R_{HI} is denoted by the white dashed line across the PVD.

Nonetheless, our decision to scale all galaxies to the largest parameter in the sample has the potential to introduce inter-pixel dependencies rendering assumptions such as those made in the adoption of Equation 4.8 erroneous, since we would duplicate spectral channels, leaving spectrally *dependent* pixels. A more conservative approach would be to scale down to the smallest galaxy, and thus merge voxels such that the pixels remain independent, ensuring a better constraint on the HI mass, and better tool overall for the investigation of excess HI around galaxies.

We note also, our decision to initially scale the spectral axes of all datacubes in our subsample of 44 galaxies to have a fiducial spectral resolution of -39062.5 Hz since a galaxy at lower resolution will need to have spectral channels replicated, however, this approach is vulnerable to introducing artificial data, which may appear as excess HI once spectral scaling is performed. Fortunately, only U7989 has been affected by this and as a result, only Group 9 results have been affected. Once more, we suggest scaling down to the most coarse spectral resolution in the stack to avoid problems of artificial data.

Adopting Better RC-Models

In this work, we made a simplifying assumption that the rotation curves of our Noordermeer et al. (2005) WHISP sub-sample of galaxies followed a Tanh distribution. That is, the disk rotation has a steadily rising rate (h_{rot}) and ultimately plateaus to a peak rotational amplitude (V_{rot}) at large radii. We omit consideration of any other rotation curve shapes.

For the Noordermeer et al. (2005) sub-sample of WHISP galaxies, the above assumption is largely inadequate. After performing a visual inspection of the full set of PVDs, for our NSA-selected sample of Noordermeer et al. (2005) galaxies, several galaxies have non-asymptotic rotation-curves at large radii. We categorise these as either *rising* or *falling* RCs.

Rising	**Flat**	**Falling**
U94	U89	U232
U508	U1310	U624
U798	U4666	U2183
U2141	U5060	U4458
U2487	U5906	U4605
U5351	U6786	U7989
U5960	U6787	U8699
U6001	U12043	U8863
U6118	-	U11852
U7704	-	-
U9644	-	-
U11914	-	-
U12815	-	-

Table 5.1: Summary of the visual inspection performed on the approved (no red-border) PVDs in the Atlas of galaxies in Figure 2.3. We include only 30 of the 33 galaxies in our NSA-selected sample of galaxies from Noordermeer et al. (2005). The excluded galaxies are U6283, U6742, U12713 due to undetermined fits header inconsistencies.

One parametric rotation curve model which effectively maps the rising inner parts of galaxy rotation curves, as well as varying outer slopes, is the "Polyex" model (Giovanelli & Haynes, 2002). This model relies on a versatile mathematical form to fit an observed rotation curve with a minimum of free parameters, and takes the form:

$$v(r) = v_{pe}(r) + V_0 \left(1 - e^{\frac{-r}{r_{pe}}}\right)\left(1 + \frac{\alpha r}{r_{pe}}\right) \tag{5.1}$$

A more elaborate model which may be adopted, also formulated phenomenologically, is called the multi-parameter function (MPF) (Courteau, 1997). This model describes the features of non-flattened rotation curves, as is predominant in our WHISP-derived sample. Taking the functional form:

$$v(r) = v_0 + v_c \frac{(1+x)^{\beta}}{(1+x^{\gamma})^{\frac{1}{\gamma}}} \tag{5.2}$$

where $x = \frac{r_t}{(r-r_0)}$. The parameters v_0, r_0, v_c, and r_t denote the velocity centre of rotation, the spatial centre of the galaxy, an asymptotic velocity, and transition radius, respectively.

The parameters of note here, differing from the Tanh model outlined in Chapter 3.2.4 are r_t, β, and γ. The term r_t denotes the inflexion point between the rising and flat parts of the rotation curve; γ uncovers the degree of sharpness of turnover in the inner regions of the rotation; β models the drop-off or rise of the outer parts of the rotation-curve. Thus, by using this model, we are able to capture declining rotation curves and PVD distributions which are not incorporated in a simple Tanh model.

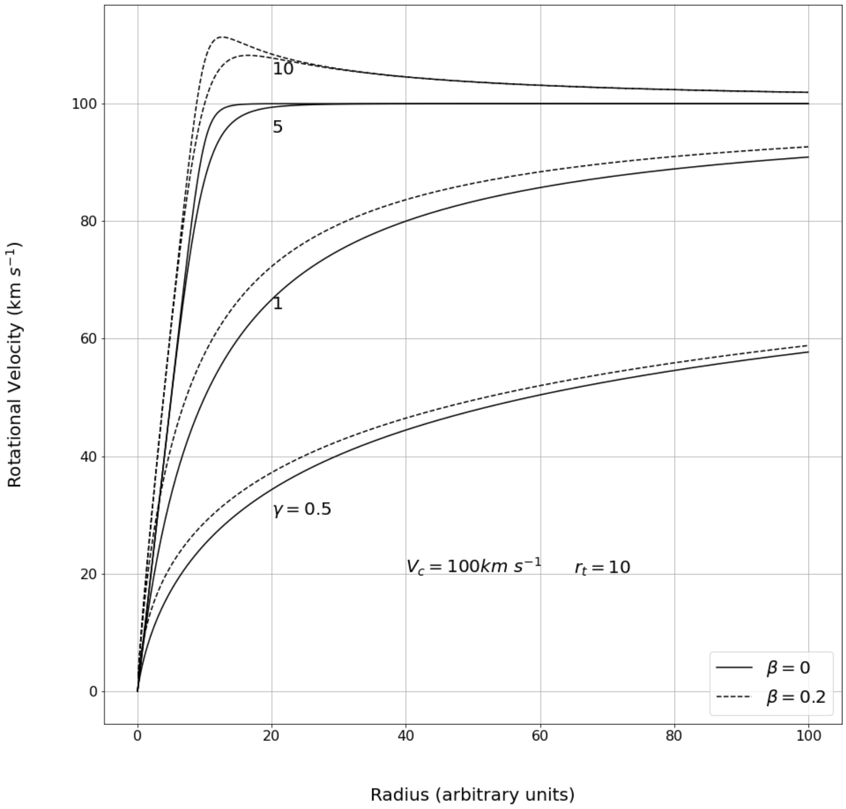


Figure 5.2: Example of MPF models for the fitting of rotation velocities. Varying the shape parameter, γ, across the five models (with a rotational velocity of $V_c = 100kms^{-1}$) demonstrates the increased ability to map decreasing rotation-curve shapes. This is presented in two variations of β.

In the case of blind survey stacking, we will not have access to many resolved sources with WSRT's 30-arcsecond beam - especially as we exceed $z = 0.1$. To this end, the coarse scaling relations in Chapter 1.4 serve as a reasonable proxy as galaxy sizes approach the beam size of the relevant telescopes.

In summary, independent of the exact model used to describe the HI distribution in PV space, careful consideration of the inner regions is required. As demonstrated in Figure 5.1, adopting a Tanh model, or MPF models for that matter, requires a secondary scaling to align the varying rising rates across galaxies - should an investigation of extraplanar gas be required. With an eye on developing an optically-based stacking tool, using scaling relations to predict HI distribution. the use of galaxy light concentration as a means to predict this rising rate is the motivation for the photometry work presented in Chapter 7.5.2.

Towards Optically-Based Scaling

During the course of this work, initial steps have been taken towards developing a stacking pipeline based purely on parameters of disk galaxies measured in the optical regime. Aided by existing scaling relations outlined in Chapter 1.4, Equation 1.1 from Broeils & Rhee (1997) (BR97, hereafter) provides a replacement of the existing parameters of R_{HI}.

That is, a future avenue of improvement is updating the 'Spatial Scaling' step in Figure 3.6, such that Equation 3.10 becomes:

$$a_S = \frac{D^{BR97}_{HI,max,sample}}{D^{BR97}_{HI,stackling}} \tag{5.3}$$

Furthermore, while we have explored the use of the amplitude of a Tanh model rotation-curve (V_{rot}) as an estimate for spectral-extent (see Chapter 3.2.4 and Equation 3.2.4), the line width at 20% peak intensity (W_{20}) provides a suitable alternative measure. In addition, by the Tully-Fisher relation (see Chapter 1.4), W_{20} correlates strongly with optical luminosity. Thus, a replacement of the 'Spectral Scaling' step in Figure 3.6 amends Equation 3.9 with:

$$v_S = \frac{V^{rot,TFR}_{max,sample}}{V^{rot,TFR}_{stackling}} \tag{5.4}$$

where $V^{rot,TFR}$ is obtained by rearranging Equation 1.4 to obtain the expected rotational velocity from the absolute B-band luminosity.

These two scaling relations provide sufficient proxies for a *coarse* signal alignment, ignoring issues relating to the specific rising rates and rotation-curve shapes dominating our approach.

Investigations of faint HI structure, however, require a more finely-tuned approach to signal alignment. In this work, we adopt a simple Tanh model to model the distribution of HI in position-velocity space. However, as discussed in Chapter 5.2.1, the behaviour of the disk at larger radii does not always lead to an asymptotic rotational-velocity (eg. V_{rot}) in Equation 3.25. Thus, the disc rising- or declining rates serve as important avenues for research - specifically finding optical proxies capable of determining rotation-curve shapes.

In Figure 5.1, we present a graphical summary of the HI alignment achieved by using a tanh function to model the HI distribution. Our omission of a non-uniform scaling of the inner regions of the PV plane is evident in the dispersion of rotation curves in this region - due to the difference in rising rates as a consequence of the varying mass distributions (both dark and luminous) within these regions.

In Chapter 1.4 we discuss such concepts such as 'Baryonic-scaling' and other mass-modelling approaches which may prove useful to predict rotation-curve shapes, and hence HI-distributions from, for instance, optical luminosity profiles in the case of maximal-disk galaxies. Using rotation curves from optical spectroscopy (tracing the $H\alpha$ kinematics) is a more direct avenue for tracing the PVD shape in HI before observation. Making use of the higher resolution at higher redshift ranges of the same galaxy will help to predict the HI rotation-curve shapes, particularly the inner rising rates where 21cm observations are prone to beam-smearing effects.

We present our initial forays toward optically determined rotation-curve correlations in Chapter 7.5. We present light-concentration ($C_{80,20}$) as well as specific radii (R_{20},R_{50},R_{80}) of the 33 WHISP galaxies with NSA-photometry.

5.2.2 Application of Cube Stacker to deep HI Surveys

The Square Kilometre Array (SKA)

In the present age of high-resolution, large-scale surveys being performed with telescope arrays such as the Square-Kilometer Array (SKA), the ability to resolve the HI in galaxies has improved significantly.

According to Braun et al. (2019), Phase 1 of the SKA consists of two telescopes, namely, SKA-Low and SKA1-Mid, covering 50 MHz to at least 15 GHz. Relevant to HI line studies, SKA1-Mid will consist of 133 15-metre SKA1 offset-Gregorian dishes, in addition to the 64 13.5-metre MeerKAT dishes, covering 350 MHz to at least 15 GHz, with baselines up to 150 kilometres. Similar to the MeerKAT L-band, SKA1-Mid2 will operate in the 0.95 – 1.76 GHz frequency range.

The *Cube Stacker* pipeline works under the assumption that the galaxies on which stacking might be performed (using the Cube Stacker algorithm) are well resolved, i.e. span several beam sizes. In the following discussion the term 'resolvable' is defined as a target with an apparent angular size θ_{gal} which spans at least three beam dimensions ($3\theta_{beam}$). Thus, by definition, galaxies remain resolvable when their angular size remains as $\theta_{gal} \geq 3 \times \theta_{beam}$. With the galaxy angular size of a galaxy of metric size R_{gal} given by:

$$\theta_{gal} = 206265 \times \frac{R_{gal}}{d}, \tag{5.5}$$

where d is the distance in the same units as R_{gal}, we can solve for the maximum distance for a galaxy to be resolved given the telescope beam size as:

$$d = 206265 \times \frac{R_{gal}}{3 \times \theta_{beam}}. \tag{5.6}$$

Braun et al. (2019) predict a maximum resolution of $\theta_{beam} = 0.4$ arcseconds for SKA1-Mid.

Using the smallest galaxy in our sample, U6742, with a major-axis length of $R_{HI} = 3.2kpc$, and applying Equation 5.6, we find that the *Cube Stacker* pipeline is applicable to galaxies out to a distance of $d = 550Mpc$. Using the Hubble law, and solving for redshift:

$$z = \frac{H_0 \times d}{c}, \tag{5.7}$$

where z is the redshift, $H_0 = 70km/s/Mpc$ as the Hubble constant, d is the distance to the object, and c is the speed of light. Thus, under maximal resolution, and using the smallest galaxy in our sample, the applicability of the *Cube Stacker* pipeline remains suitable out to at least a redshift of $z = 0.128$ for WHISP-NSA sized galaxies observed at maximal resolution on SKA Phase 1.

Several ongoing surveys performed using the SKA-precursor MeerKAT are well underway with data on hand. Three of the deepest are the LADUMA (Looking At the Distant Universe with MeerKAT Array, Baker et al. (2018)), CHILES (COSMOS HI Large Extragalactic Survey, Fernandez et al. (2015)), and MIGHTEE-HI (The HI emission project of the MeerKAT MIGHTEE survey, Maddox et al. (2021)). As we performed a cross-matching of WHISP radio targets with optical data in the NASA-Sloan Atlas, a similar cross-matching can be performed for targets in these listed surveys which are largely uncovered by the SDSS. In the case of CHILES, COSMOS data is on hand by definition of the sample selection. Thus, further development of the *Cube Stacker* pipeline is possible with data with greater spatial and spectral resolution.

Chapter 6

Summary and Conclusions

The present generation of high-resolution radio interferometers is already making groundbreaking discoveries. Advancements over the past few decades have allowed us to extend the limits of HI detections to higher redshifts and fainter emissions. Furthermore, statistical techniques have been developed to uncover fainter signals . One such technique is the co-addition of the HI global profile or one-dimensional (1D) stacking. In this work, we explore the benefits of stacking at higher dimensions. We present a novel approach to stacking imaged HI data cubes, or three-dimensional (3D) stacking, reporting valuable lessons and considerations which could be useful in extending the sensitivity limits in the age of the Square-Kilometre Array (SKA).

In general, stacking enables a deeper probing of HI in galaxies. Exploiting simple Poisson statistics, where co-added signal and noise rise by N and $\sqrt{N}$, respectively, where N is the number of spectra stacked (e.g. Delhaize et al. (2013)). Thus, an increased signal-to-noise ratio (SNR) is achieved - paving the way for uncovering faint features. Recent investigations by Chen et al. (2021) and Sinigaglia et al. (2022) into the merits of stacking in 3D focus on the data cubes *prior* to deconvolution, with the goal of extending UV coverage. These techniques address challenges in blind, unresolved HI surveys, facilitating the measurement of bulk properties such as HI-mass. In our approach, with sufficient galaxy samples, the focus is put on uncovering faint HI *structure* in a 3D sense, extending the spatial and spectral sensitivity.

Key Results

In developing the *Cube Stacker* pipeline, we have obtained rotation curve fits to 33 of the 68 galaxies in the Noordermeer et al. (2005) WHISP sample. Using these measures, and applying them in our 3D stacking algorithm, we have demonstrated:

1. Stacking unweighted 1D profiles performs identically to extracting a 1D profile from an unweighted 3D stack.
2. Application of a weighting regime results in the stack profiles (1D vs 3D) diverging.

However, they still fall within uncertainty - although we anticipate that this is due to low number statistics.

3. Applying an inverse-variance weighting has been shown to produce an additional increase in SNR over its equal-weighting counterpart, as demonstrated by the horizontal lines, denoting the noise level in the stackling cube, in Figure 4.9.

4. Figures 4.2 - 4.6 show that the HI radial profiles demonstrate an order of magnitude decrease in the range $R_{HI} \leq r \leq 2R_{HI}$, with the drop-off rate seemingly correlating with mass.

5. Across all stacking groups, the radial profiles begin to flatten beyond $r = 2R_{HI}$, suggesting a dampening of the noise from which, following the inclusion of a greater number of galaxies within a stack, we anticipate the emergence of any faint signal which may be present across the stack group.

Pending greater number statistics, we anticipate these preliminary results provide a promising avenue towards achieving a ‘Comprehensive Unearthing of Baryons in Extended-regions using a Stacker’ (*Cube Stacker*).

Further Work

Optimally aligning voxels within an *imaged* HI data cube is non-trivial. Careful consideration of galaxy distribution modelling must be considered, as this forms the basis of the definition of signal alignment in a 3D context. Our initial approach of adopting a Tanh model relied on the assumption that the galaxies had a steadily rising rate, culminating in a tangential velocity at a large radial extent. Depending on the galaxy sample, as later uncovered by utilising the WHISP sub-sample of galaxies, this assumption is often overly simplistic resulting in erroneous detection of, for instance, projected disc gas as extraplanar gas. Furthermore, the intrinsic disc thickness (q0) and galaxy axial-ratio (b/a) cannot be left unmatched, as the projection of HI due to misaligned galaxies, as demonstrated by Chapter 4.3, is a critical consideration for studies of faint HI emission in the circumgalactic medium. To this end, our initial iteration of the *Cube Stacker* algorithm has provided a fair coarse alignment of the HI signal in nearby spirals.

Upon concluding this study, we have identified several additional considerations that have yet to be addressed beyond voxel alignment. Of particular concern is the potential impact of the deconvolution algorithm on data cube systematics. Following our discussion in Chapter 5.1.2, we suggest the *Cube Stacker* algorithm be tailored toward a dataset which has undergone a deconvolution algorithm such as MSCLEAN which better handles extended sources.

Following the direct alignment of HI signal, and voxels specifically in the case of 3D stacking, there exists the potential for substituting these HI parameters with optical proxies. To this end, we present measures of central concentration (see Chapter 7.5.2) to be used for the further development of an optically based HI stacking suite.

Chapter 7

Appendices

The Appendices include six subsections that detail our development in the main Chapters of this thesis. For reference, these are connections between these subsections and the main Chapter of the thesis:

7.1 Describing a Transformed Beam: how we characterize a beam following the cube transformations as outlined in Chapter 3.1.1.

7.2 Signal alignment and results from envelope fitting that details the results outlined in Chapter 3.2.3 and 3.2.4.

7.3 Comparison of Masked SBPs which accompany the discussion demonstrating the improved SNR achieved by 3D stacking. This discussion can be found in Chapters 4.2-4.6.

7.4 Masked Stacklings are used to isolate physical counterparts, as done in Chapter 4.7, to features emerging in the SBPs.

7.5 Light concentrations from NSA Photometry, detailing initial steps towards an optically based stacking suite.

7.1 Describing a Transformed Beam

The radio beam produced by a radio interferometer can be described using three parameters:

1. The beam major-axis (BMAJ = a)

2. The beam minor-axis (BMIN = b)

3. The beam position-angle (BPA = θ)

In essence, these three parameters describe the size and orientation of a Gaussian ellipse on the sky, from which radio information in this instance is sampled.

In this work, we perform substantial transformations on the data cubes as outlined in Chapter 3, these transformations also affect the size and orientation of the beam. Below we discuss the relevant effects and how we track and update the beam parameters.

The motivation for tracking the effects of linear transformations on the beam stems from the effects of a beam on resolution. In broad terms a smaller beam leads to larger resolution and, with the aim of our project to stack data cubes with identical properties, we must account for the differences in beam shapes and orientations, convolving all cubes to the same spatial as well as spectral resolution, prior to stacking.

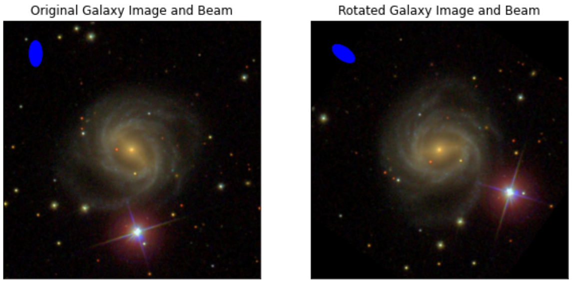


Figure 7.1: Demonstrating the effect of rotating an image on the beam (blue ellipse) associated with the image. We perform a rotation on each spectral channel of an HI data cube, treating these as images or 2D arrays. For illustrative purposes, we show the effect of a counterclockwise rotation of an image of NGC 2487 by 55°, and the corresponding rotation the beam undergoes. Credit: SDSS DR14.

7.1.1 The Transformed Beam

While the Westerbork Synthesis Radio Telescope (WSRT) beam will vary depending on the specific baseline configuration, the East-West geometry of the array means the default beam position-angle (BPA) will be 0°, lying North-South.

An ellipse aligned to the axes can be described in parametric form as:

$$\begin{aligned} x &= a\cos t \\ y &= b\sin t \end{aligned} \tag{7.1}$$

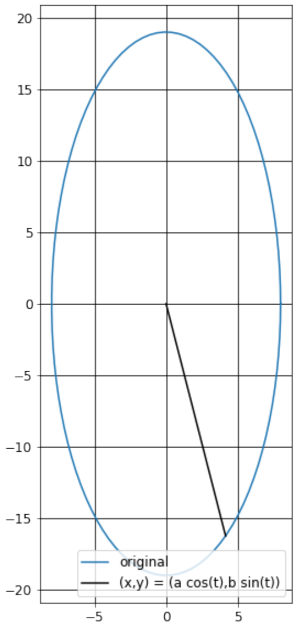


Figure 7.2: Demonstrating the parametric form of an ellipse as described in Equation 7.1. It is worth noting that the angle t is not the angle of (x(t),y(t)) with the x-axis, unlike θ referenced in the rest of this section.

By utilising matrix operations, we can derive the set of points which describe an ellipse which undergoes a rotation θ and scaling S_x and S_y along the x- and y-axis, respectively. We capture these transformations as:

$$P' = SR\begin{pmatrix} x \\ y \end{pmatrix}, \tag{7.2}$$

with scaling matrix S, and rotation matrix R, defined as:

$$S = \begin{pmatrix} S_x & 0 \\ 0 & S_y \end{pmatrix} \tag{7.3}$$

$$R = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{7.4}$$

Using Equation 7.2, we generate a set of (x,y) data points describing an ellipse under the necessary transformations. From this, the updated semi-major, semi-minor axes, and position angle of the transformed beam can be obtained by performing a least-squares fit.

7.1.2 Obtaining the Transformed-Beam Parameters

The easiest obtained parameter is the position-angle of the transformed beam. Modelling the beam as a smooth ellipse, we adopt the conic definition:

$$Ax^2 + Bxy + Cy^2 + F = 0 \tag{7.5}$$

where the coefficients combine the radio-beam parameters in (i-iii) as follows:

$$\begin{aligned} A &= (b\cos\theta)^2 + (a\sin\theta)^2 \\ B &= 2\sin\theta\cos\theta(b^2 - a^2) \\ C &= (a\cos\theta)^2 + (b\sin\theta)^2 \\ F &= -(ab)^2 \end{aligned} \tag{7.6}$$

Equation 7.5 describes an ellipse rotated counter-clockwise by an angle θ about its centre. To capture the effects of scaling, we apply the following mapping to the scaled reference frame:

$$\begin{aligned} x' &= S_x \cdot x \\ y' &= S_y \cdot y \end{aligned} \tag{7.7}$$

As demonstrated in Figure 7.3, this scaling affects both the ellipse position angle and axis lengths. To obtain the new radio-beam parameters (a',b',θ'), we adopt a least-squares approach to fitting an ellipse.

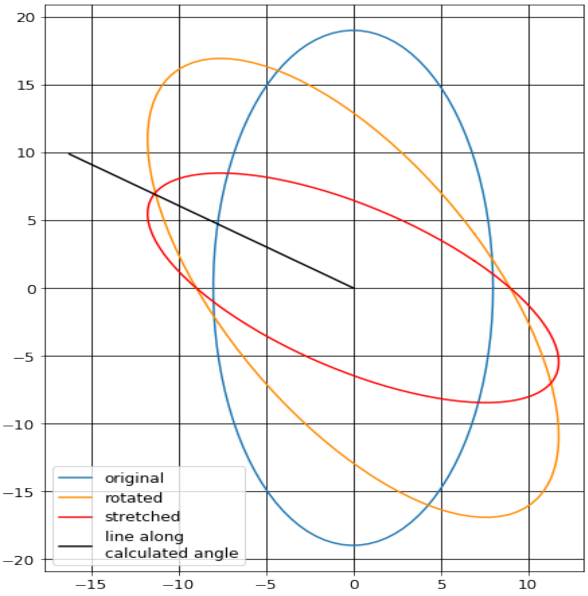


Figure 7.3: A sequenced demonstration of an ellipse undergoing a rotation by 30° counterclockwise (orange curve) and then a scaling of $S_x = 1$ and $S_y = 0.5$ along the x- and y-axis respectively (red curve). The black line shows the updated position angle of the scaled and rotated ellipse.

Applying the mapping in Equation 7.7 onto Equation 7.5, the coefficients in Equation 7.6 now become:

$$\begin{aligned} &\Rightarrow A' = S_x^2 A \\ &\Rightarrow B' = S_x S_y B \\ &\Rightarrow C' = S_y^2 C \end{aligned} \tag{7.8}$$

Using these updated coefficients, we obtain the new position-angle of the scaled and rotated ellipse as:

$$\theta' = \arctan\left(\frac{C' - A' - \sqrt{(A' - C')^2 + B'^2}}{B'}\right) \tag{7.9}$$

Obtaining the updated axial lengths is not as straightforward. Starting again with the conic definition of an ellipse:

$$f(a, P') = \mathbf{D} \cdot \mathbf{a} = 0 \tag{7.10}$$

where $\mathbf{D} = (x^2, xy, y^2, 1)$ and $\mathbf{a} = (A, B, C, F)$, the same coefficients from Equation 7.6.

We apply the geometrical arguments employed by Fitzgibbon et al. (1996), reducing Equation 7.10 to an eigenvalue problem, whereby **a** is minimised to obtain the major- and minor-axis of the transformed beam.

We perform these calculations using the Python linear algebra functions *numpy.linalg*.

7.2 Results from Envelope Fitting

In this chapter, we present the results from our novel envelope fitting routine as introduced in Chapter 3.2.3. Figures 7.4-7.8 present the position-velocity diagrams (PVDs) upon which the envelopes are plotted to demonstrate the tracing of the HI in PV space. Finally, we present the results of the modelling of the distribution of these envelopes performed using a Tanh model as described by Equation 3.25 in Chapter 3.2.4.

UGC	h_{rot} [arcsec]	V_{rot} [km/s]	dV_{rot} [km/s]
89	7.2	171.94	17.16
94	16.07	135.29	18.53
232	19.65	113.26	5.08
508	30.74	205.08	32.44
624	20.38	251.56	18.69
798	31.65	101.51	12.6
1310	4.12	61.64	4.69
2141	37.07	115.61	8.15
2183	30.08	129.08	7.61
2487	37.23	196.01	10.19
4458	21.24	110.77	0.68
4605	48.57	187.95	6.14
5060	29.57	34.89	7.69
5351	12.31	129.56	1.02
5906	9.84	46.50	1.35
5960	34.1	85.78	10.32
6001	8.06	57.35	3.7
6118	23.83	80.78	1.81
6283	66.44	97.17	9.22
6742	14.27	46.01	4.16
6786	25.54	202.29	1.17
6787	43.71	219.25	0.73
7704	21.82	37.28	4.75
7989	82.38	44.73	15.09
8271	5.83	59.93	10.16
8699	28.3	175.45	11.03
8863	50.21	176.49	9.59
9644	13.25	88.23	0.48
11852	18.97	133.45	5.16
11914	52.11	163.55	2.23
12043	29.48	167.29	3.0
12713	37.85	58.13	3.19
12815	16.76	265.35	62.81

Table 7.1: Table summarising the results from fitting a Tanh model (see Chapter 3.2.4) to envelopes fit obtained as outlined in Chapter 3.2.3.

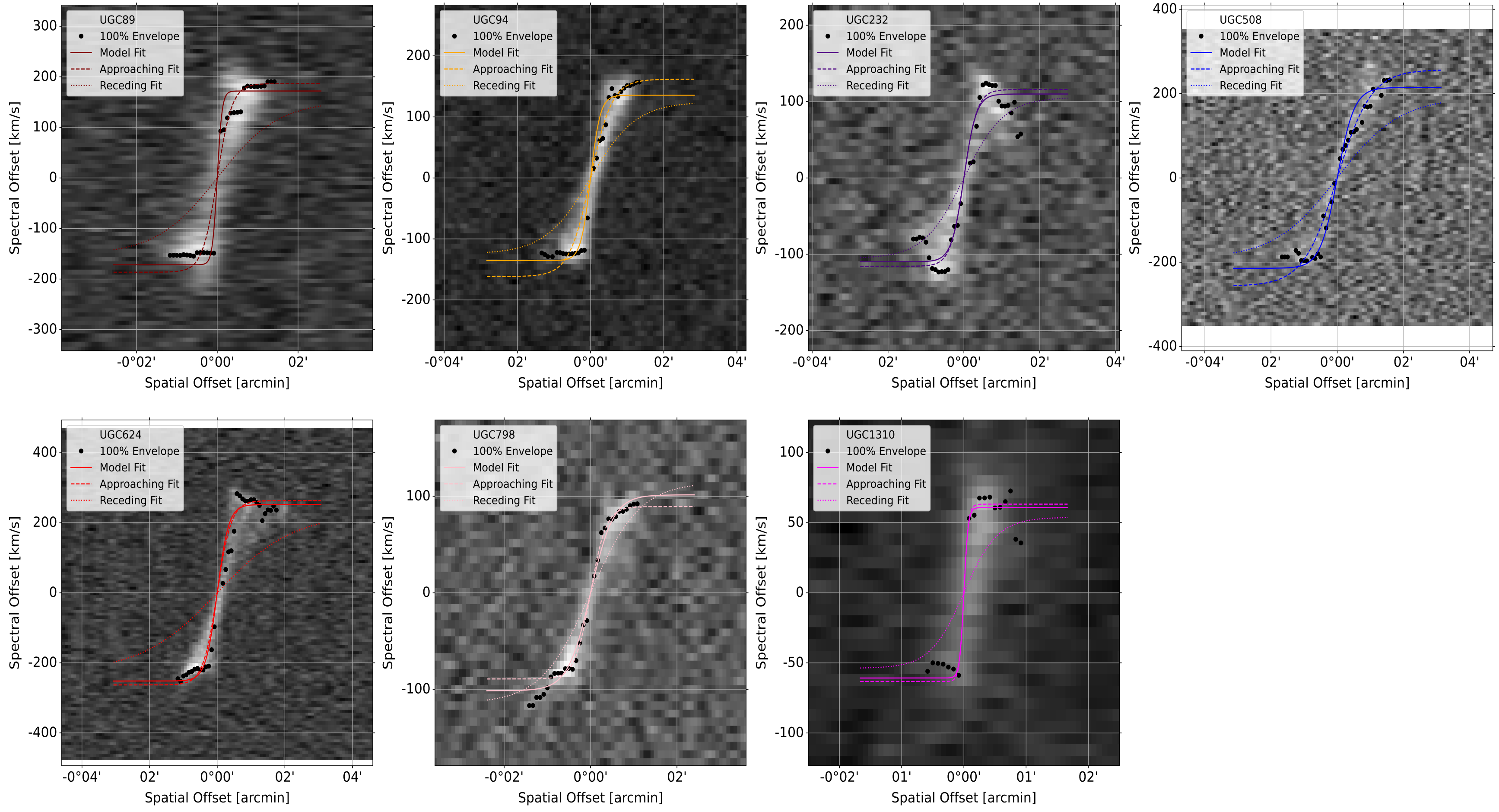


Figure 7.4: Presentation of PVDs, with envelope overlays, and Tanh models fit to the approaching (dashed), receding (dotted), and to both ends simultaneously (solid). The area between dashed and dotted curves estimates rotational amplitude error (dV_{rot}) as detailed in Table 7.1.

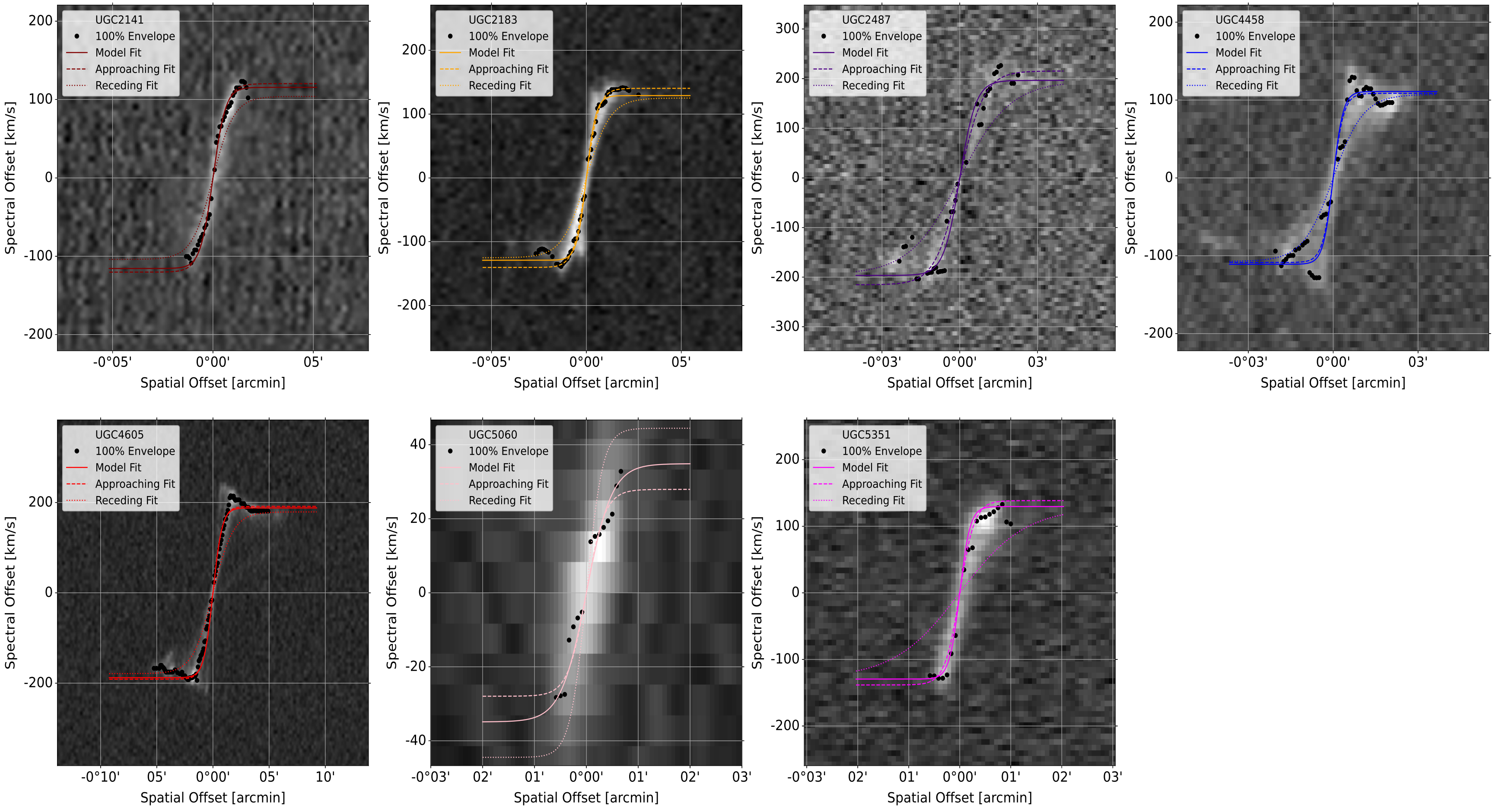


Figure 7.5: Presentation of PVDs, with envelope overlays, and Tanh models fit to the approaching (dashed), receding (dotted), and to both ends simultaneously (solid). The area between dashed and dotted curves estimates rotational amplitude error (dV_{rot}) as detailed in Table 7.1.

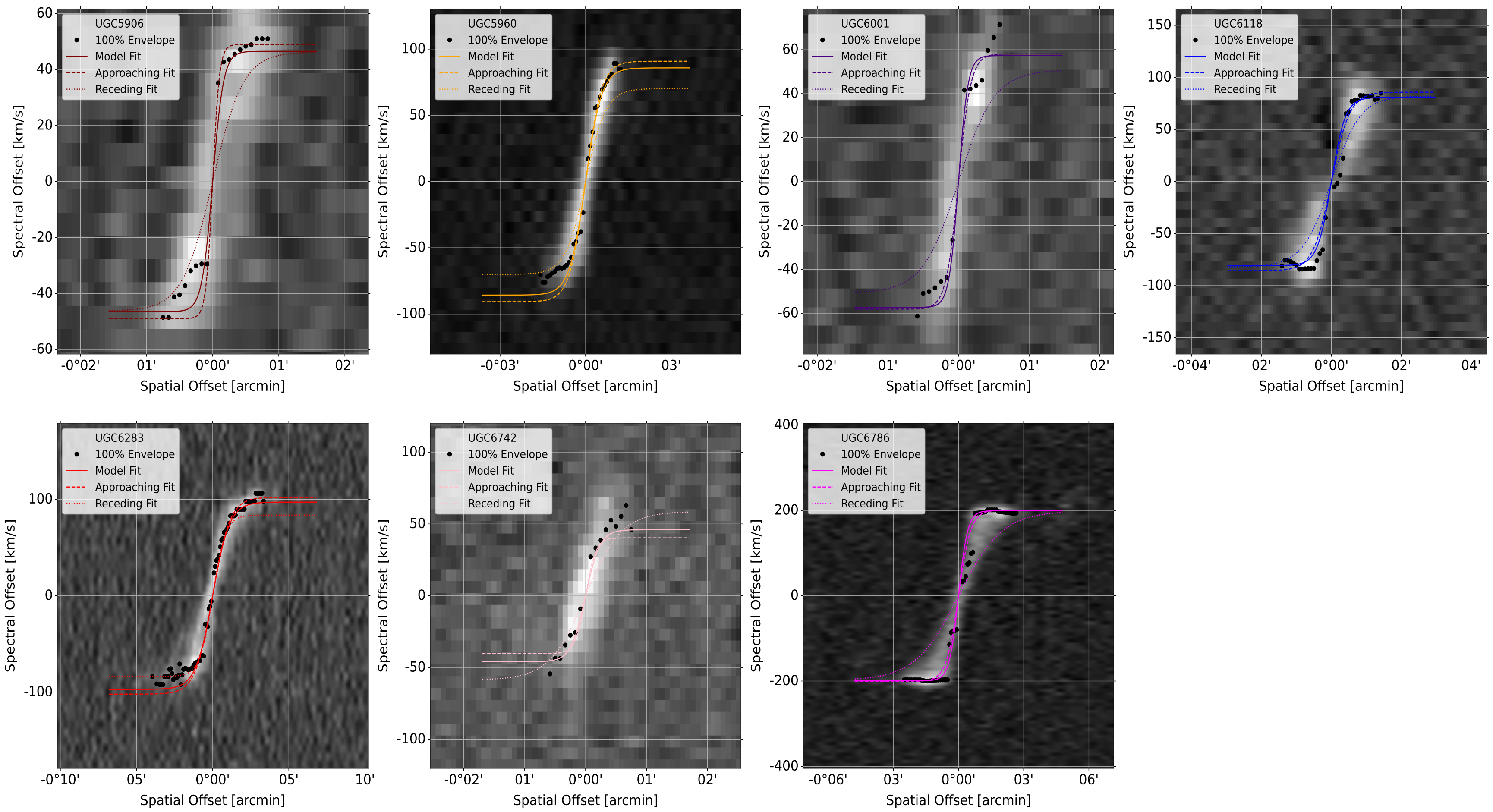


Figure 7.6: Presentation of PVDs, with envelope overlays, and Tanh models fit to the approaching (dashed), receding (dotted), and to both ends simultaneously (solid). The area between dashed and dotted curves estimates rotational amplitude error (dV_{rot}) as detailed in Table 7.1.

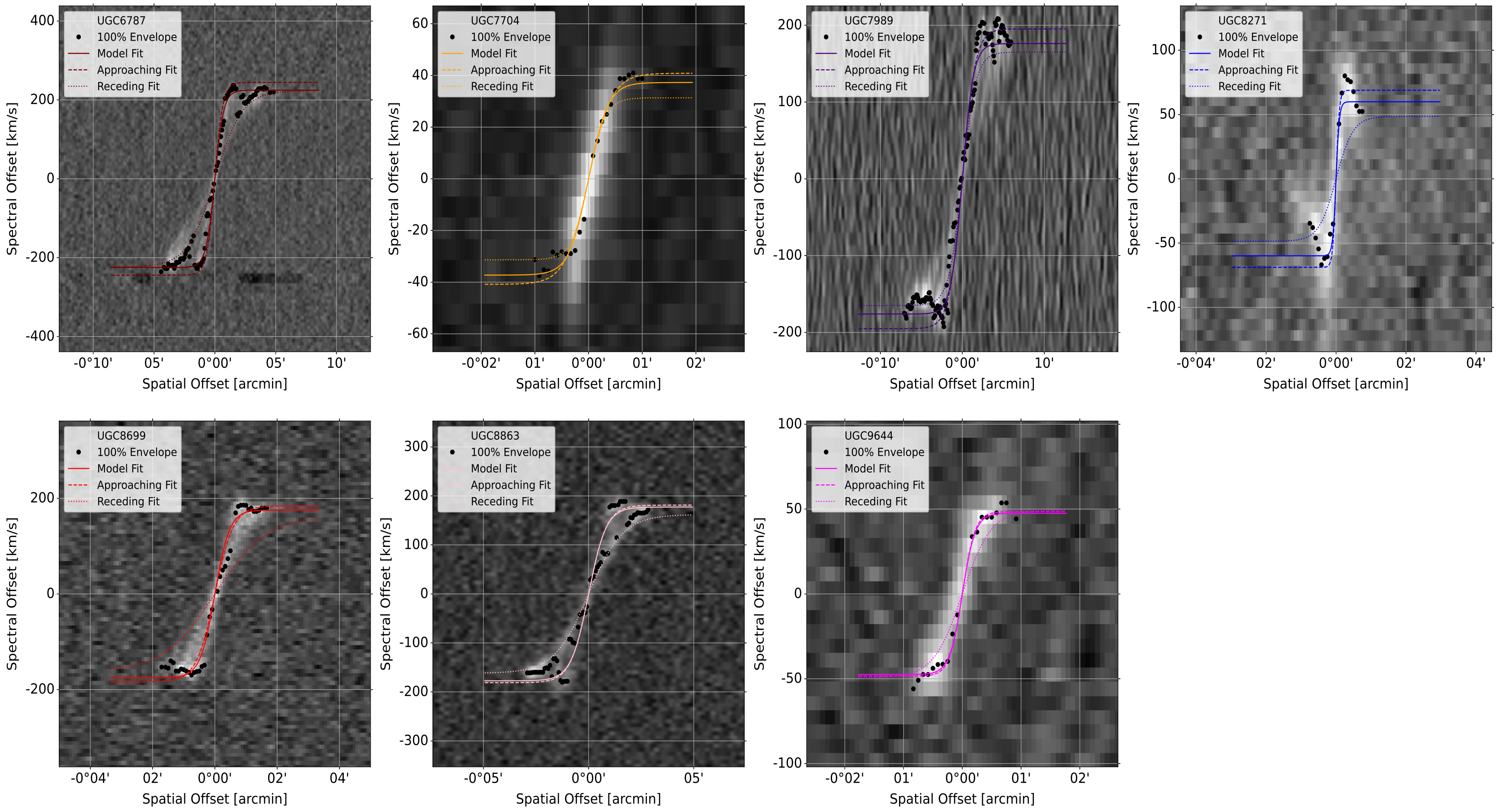


Figure 7.7: Presentation of PVDs, with envelope overlays, and Tanh models fit to the approaching (dashed), receding (dotted), and to both ends simultaneously (solid). The area between dashed and dotted curves estimates rotational amplitude error (dV_{rot}) as detailed in Table 7.1. Note the PVD of U7704 does not have a visible rollover as associated with major axis PVDs. The cause for this is discussed in Chapter 4.3.

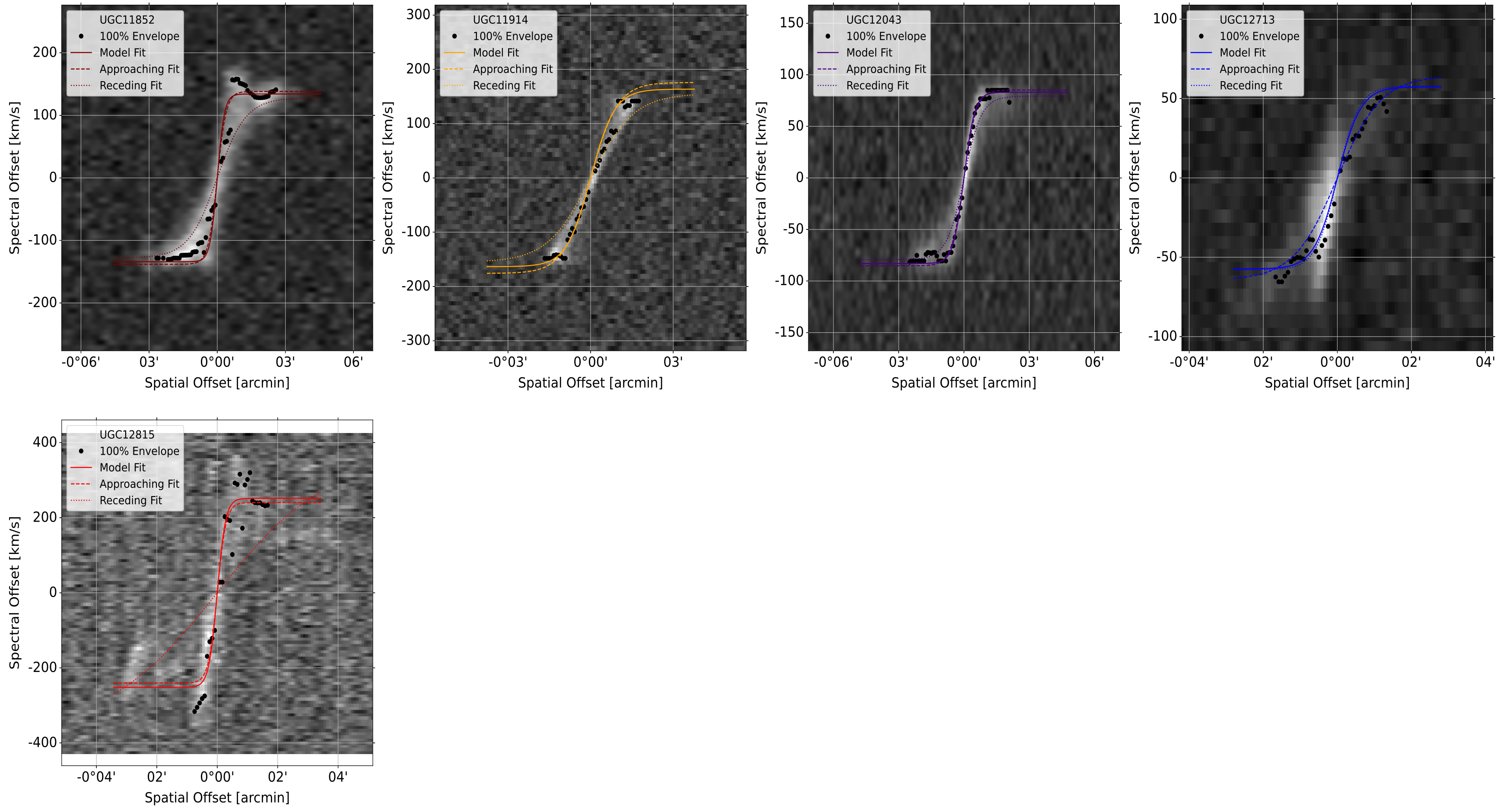


Figure 7.8: Presentation of PVDs, with envelope overlays, and Tanh models fit to the approaching (dashed), receding (dotted), and to both ends simultaneously (solid). The area between dashed and dotted curves estimates rotational amplitude error (dV_{rot}) as detailed in Table 7.1.

7.3 Comparison of Masked SBPs

In this chapter, we present the set of masked (sigma-clipped) surface brightness profiles as utilised to uncover extended features within the noise, while simultaneously demonstrating the merits of applying a weighting scheme dictated by the noise properties of the stacklings. Additionally, we use the surface brightness profiles of the masked cubes and one-dimensional stacks to demonstrate the improved signal-to-noise achieved by stacking in 3D as opposed to 1D.

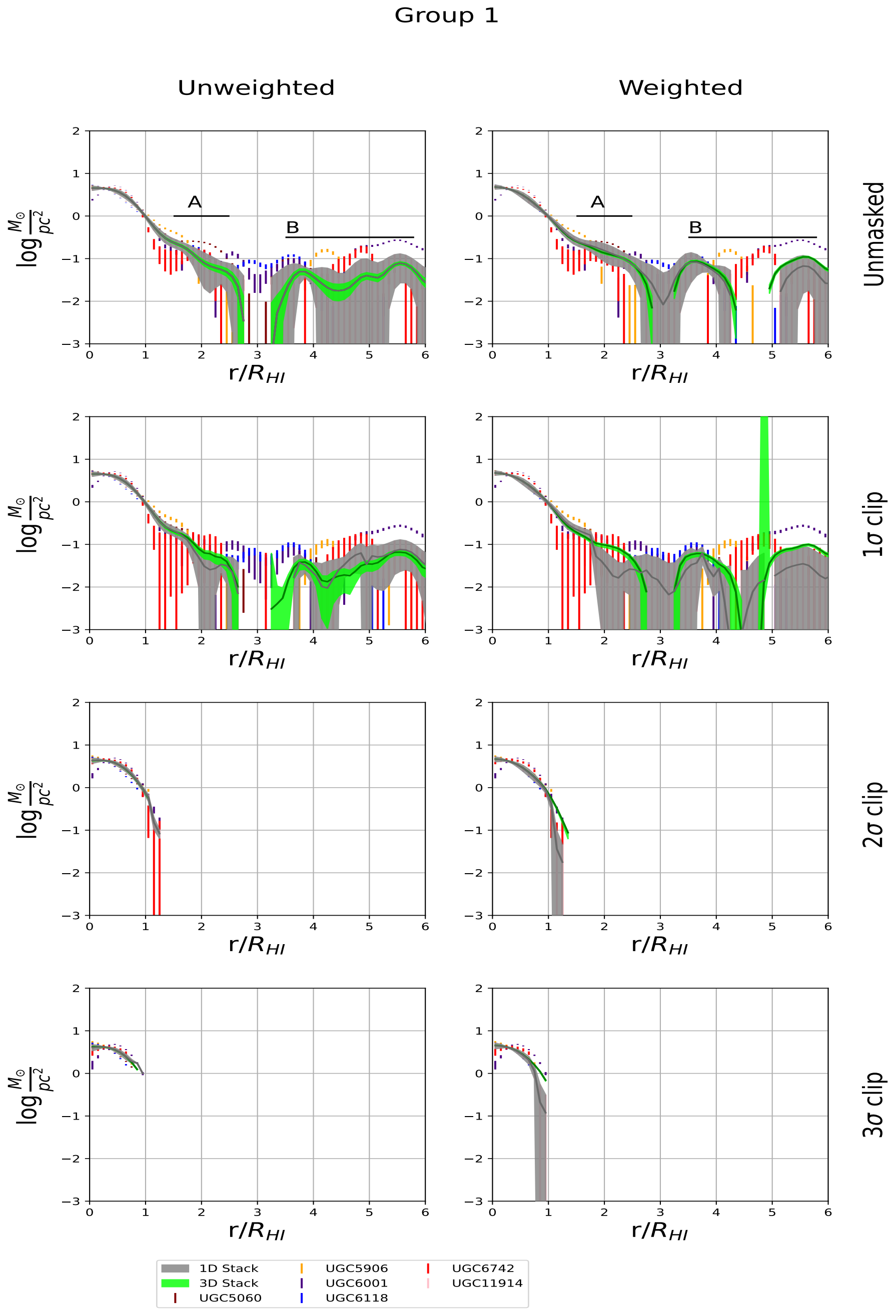


Figure 7.9: Stack Group 1: Low-mass with Low-inclination (i < 47.9°). As discussed in Chapter 4.2.2 the amplitude in the inner regions of the weighted 3D stack is consistently higher than that of the 1D stack.

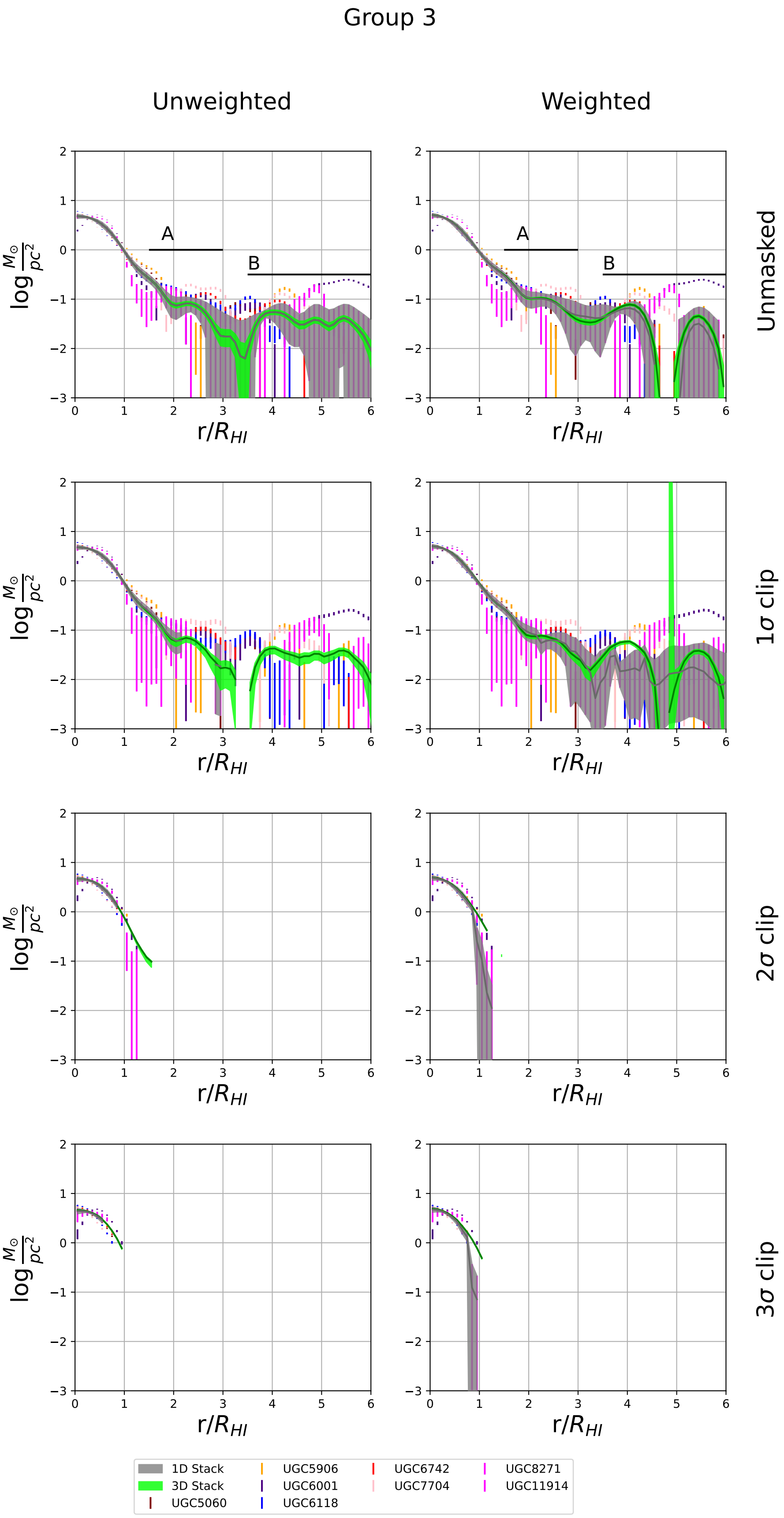


Figure 7.10: Stack Group 3: Low-mass with Low-to-moderate inclination (i <70). As discussed in Chapter 4.3.2 .

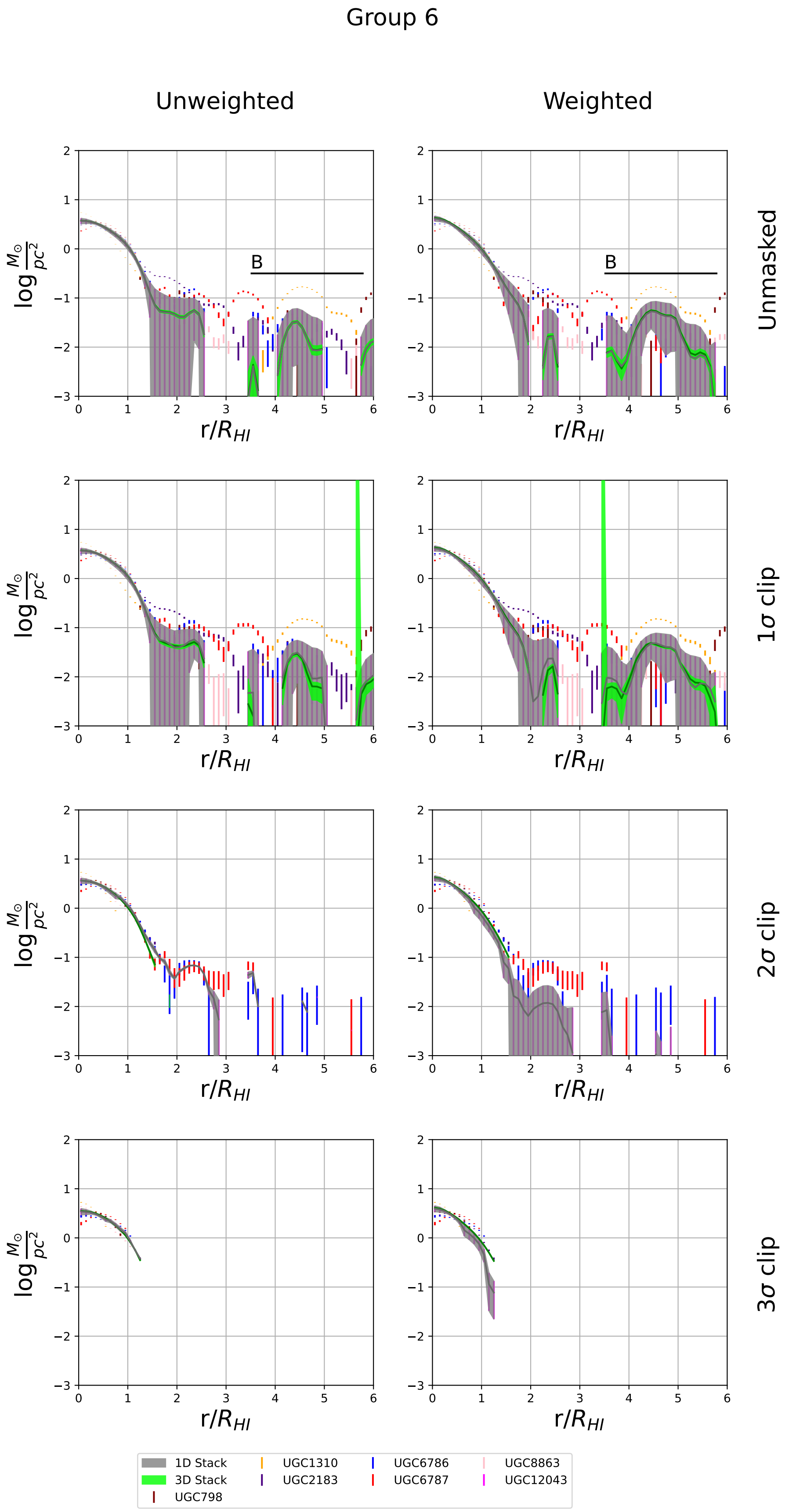


Figure 7.11: Stack Group 6: Moderate-mass with Low-to-moderate inclination (i <70).

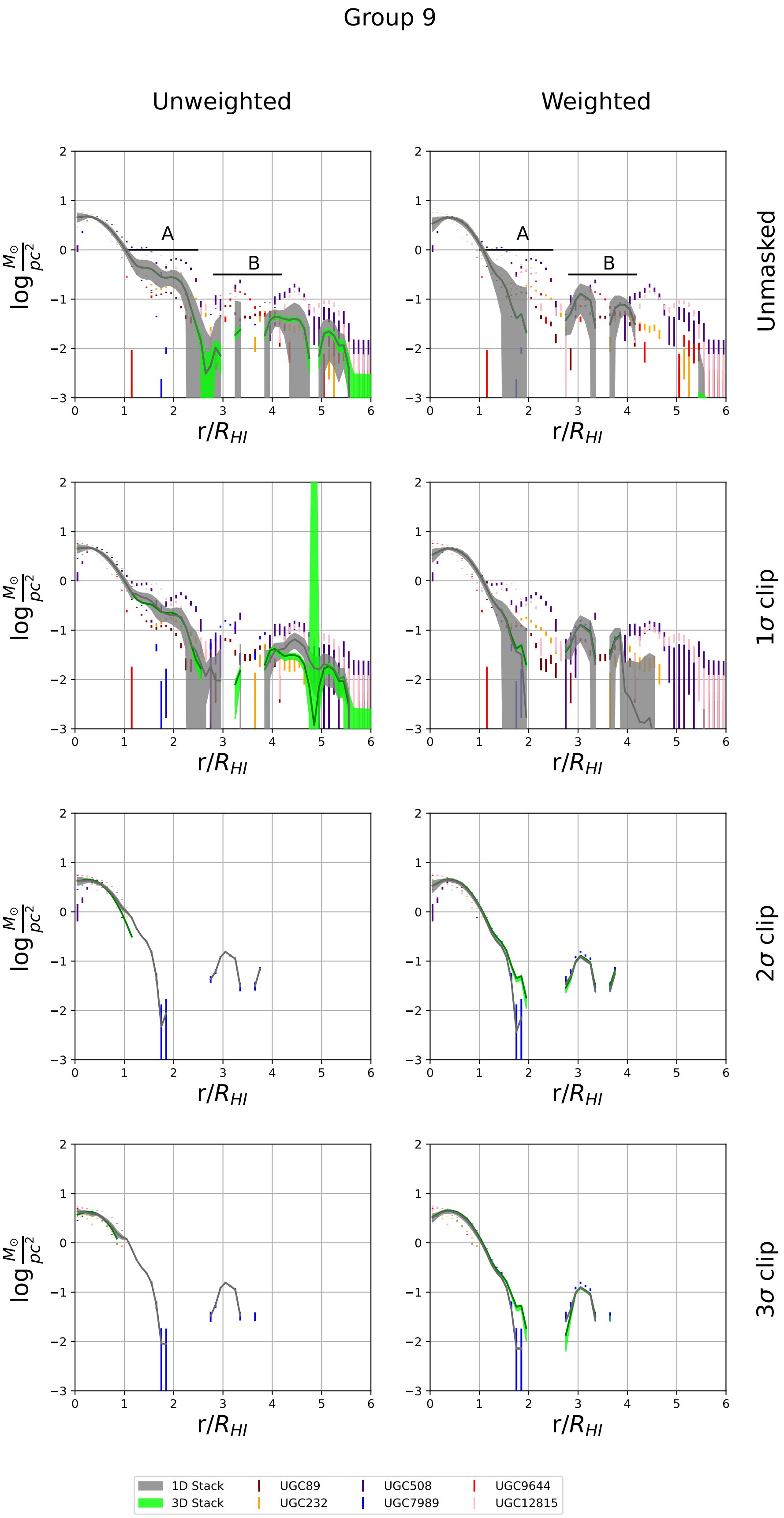


Figure 7.12: Stack Group 9: High-mass with Low-to-moderate inclination (i <70). We discuss the performance of 3D stacking over 1D stacking in Chapter 4.5.2.

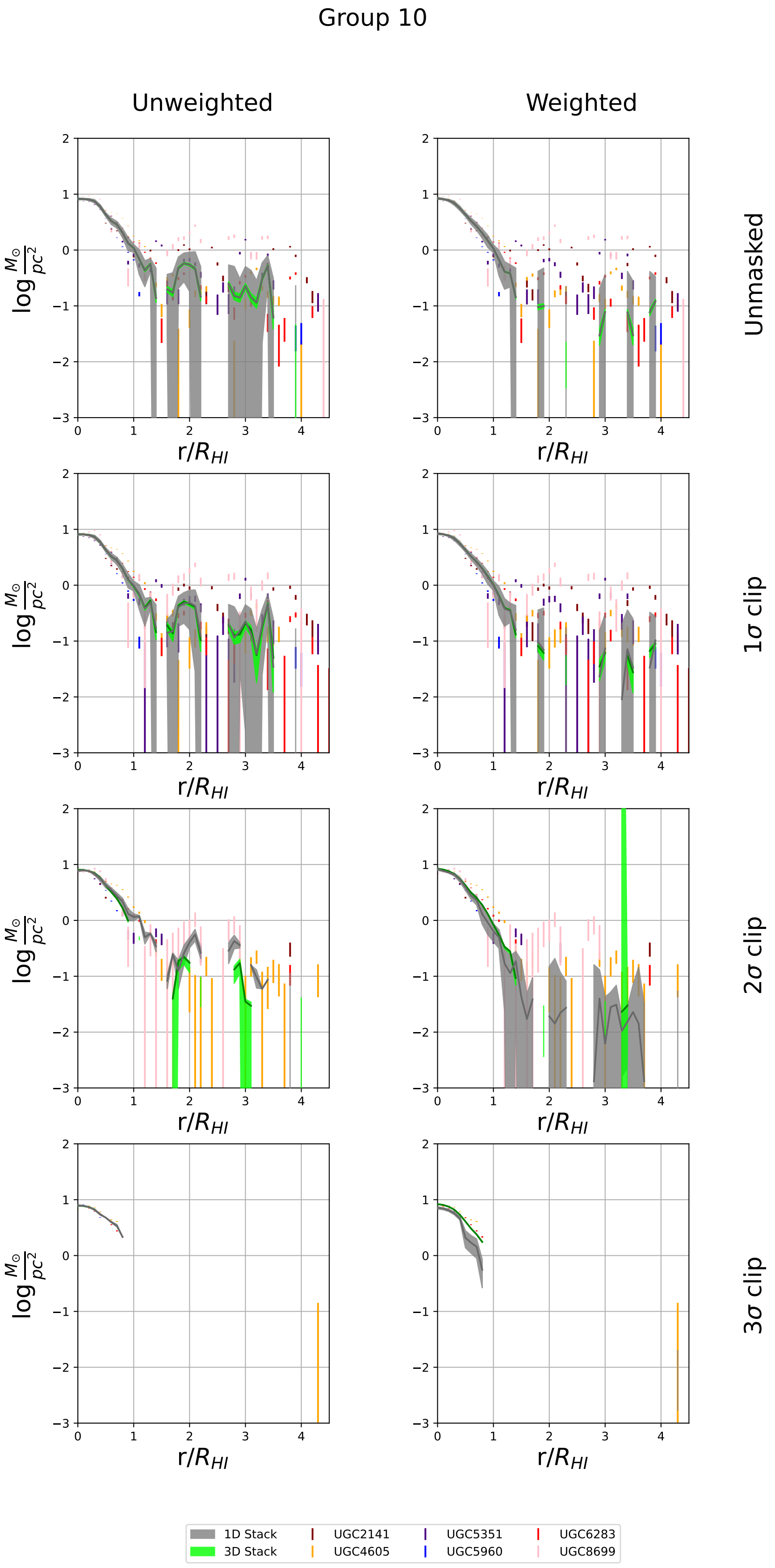


Figure 7.13: Stack Group 10: High-inclination ($i \geq 70°$).

7.4 Masked Stacklings

In this section, we present the moment 0 maps of the individual galaxies - grouped by stacking group. The analysis presented in Chapter 4.2.2 outlines significant features uncovered in the moment 0 maps by means of the application of a mask defined by a sigma-clipping relative to measured background noise. Construction of the moment 0 maps is described by Equation 4.1 in Chapter 4.1.1.

Figure 7.14: Below, Group 1 masked moment 0 maps. Each column visualises the moment 0 at the particular sigma-clipping level performed on the data cube. Each of the annuli at intervals of R_{HI} are overplotted, corresponding to the

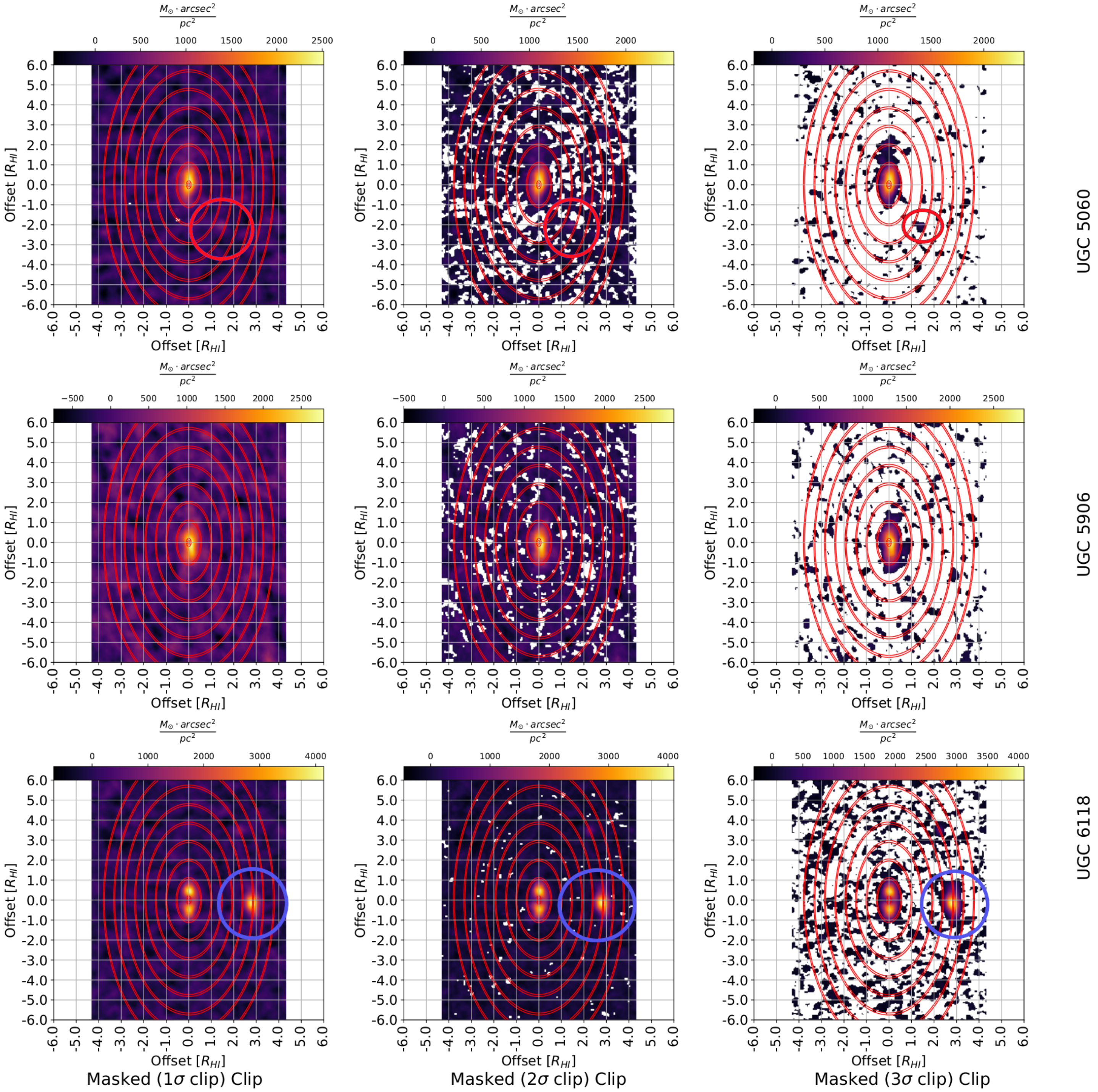


Figure 7.15: Group 1 masked moment 0 maps continued.

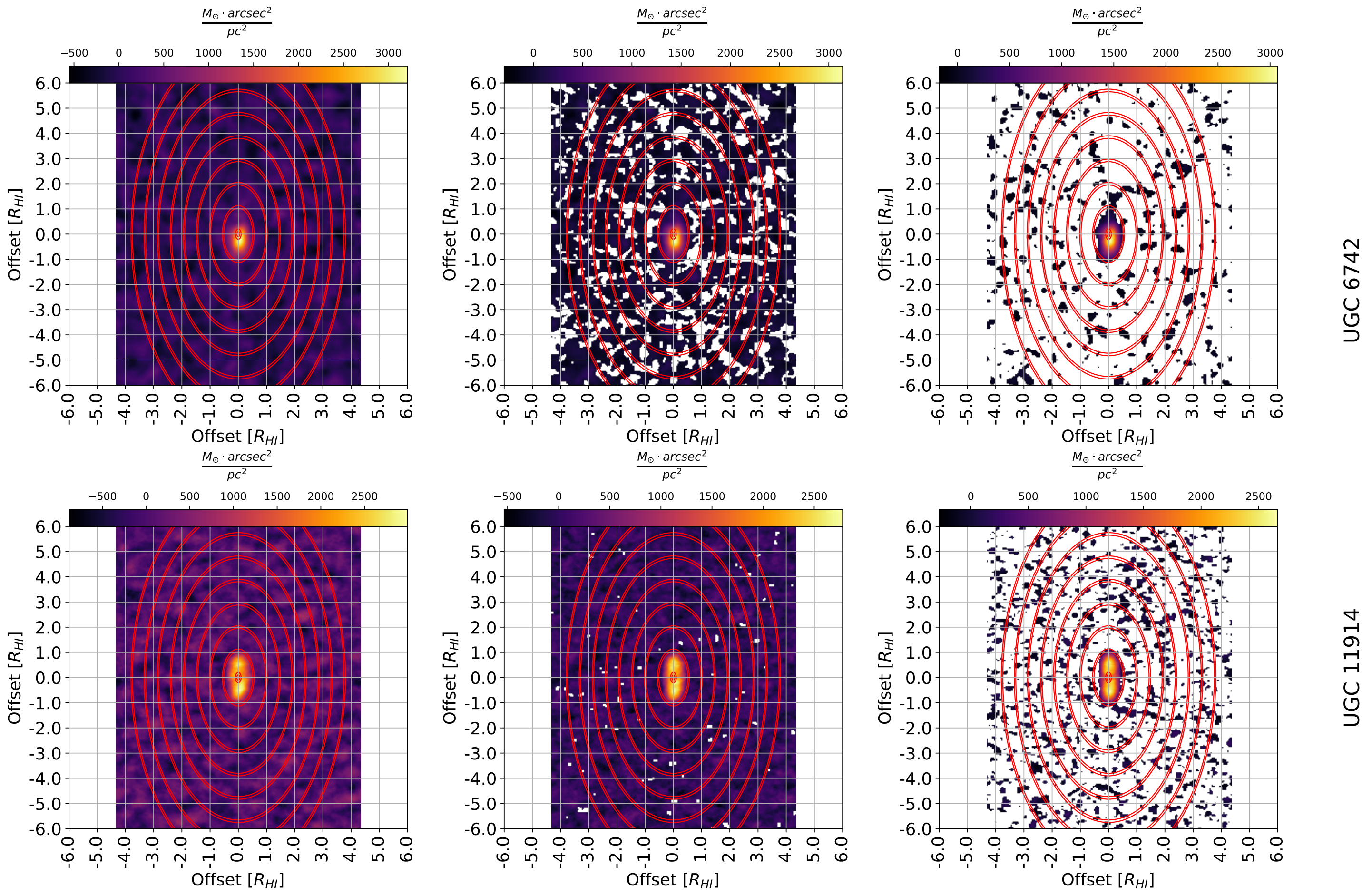
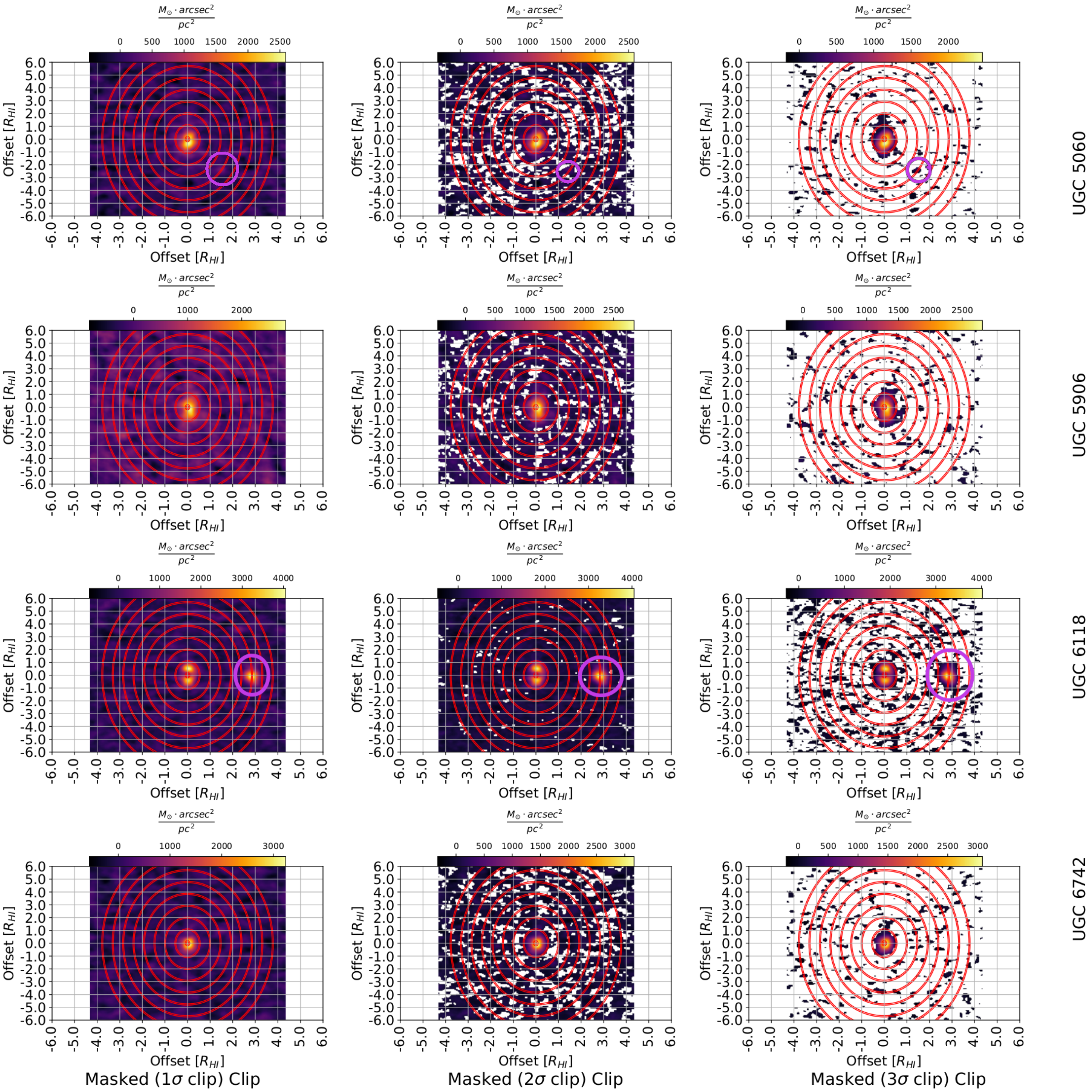


Figure 7.16: Below, Group 3 masked moment 0 maps. Each column visualises the moment 0 at the particular sigma-clipping level performed on the data cube. Each of the annuli at intervals of R_{HI} are overplotted, corresponding to the analysis in Chapter 4.3.2.



Masked (1σ clip) Clip
Masked (2σ clip) Clip
Masked (3σ clip) Clip

Figure 7.17: Group 3 masked moment 0 maps continued. Demonstrated by the purple line, we confirm an outdated measurement of the position angle for U7704 by Noordermeer et al. (2005). Despite the apparent east-west elongation, we attribute this to an error in inclination reprojection due to the contrasting reported inclinations for U8271. The addition of these galaxies serve as a demonstration of the sensitivity of the *Cube Stacker* to accurate measures of position-angle and inclinations.

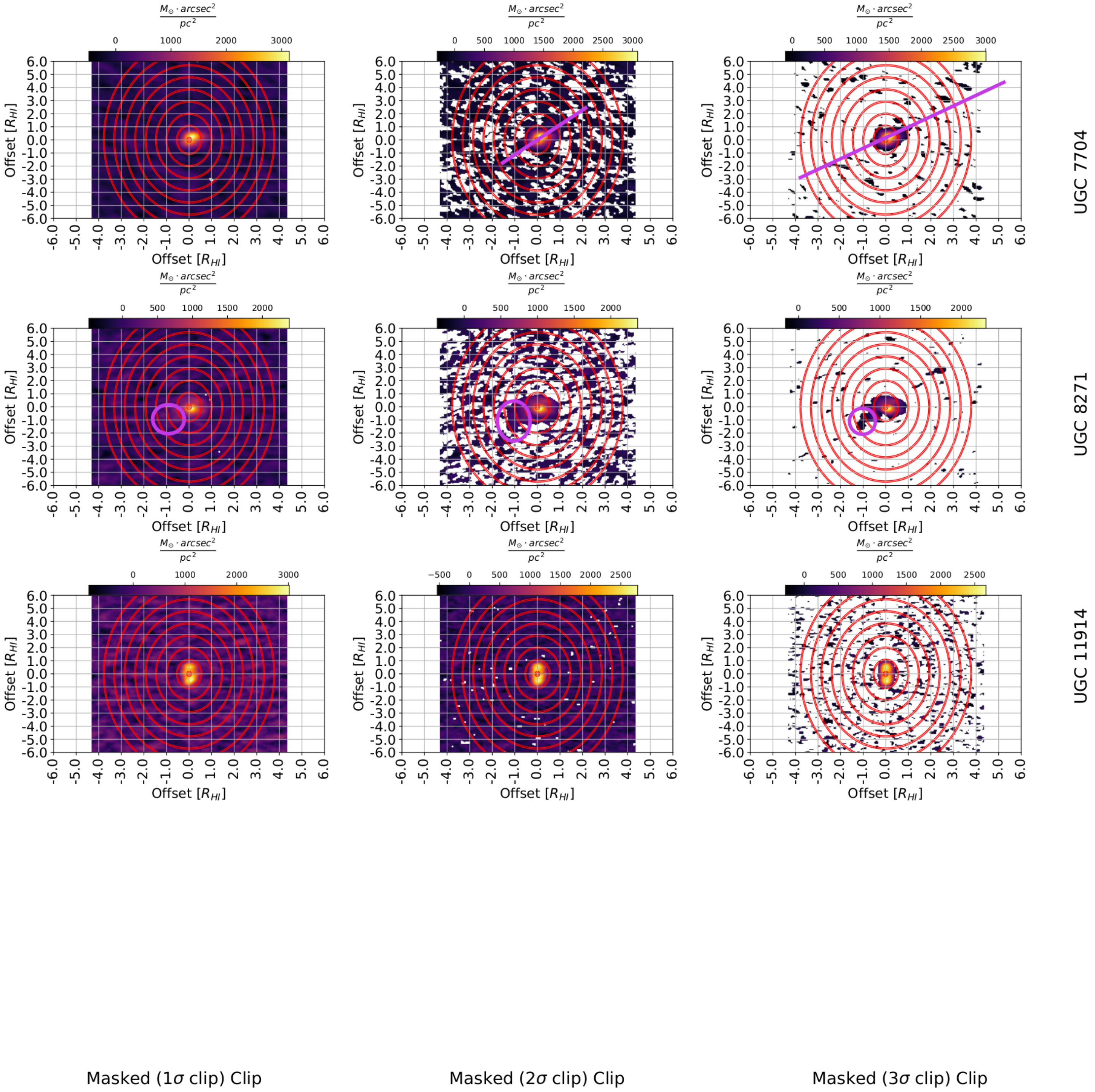


Figure 7.18: Below, Group 6 masked moment 0 maps. Each column visualises the moment 0 at the particular sigma-clipping level performed on the data cube. Each of the annuli at intervals of R_{HI} are overplotted, corresponding to the analysis in Chapter 4.4.2.

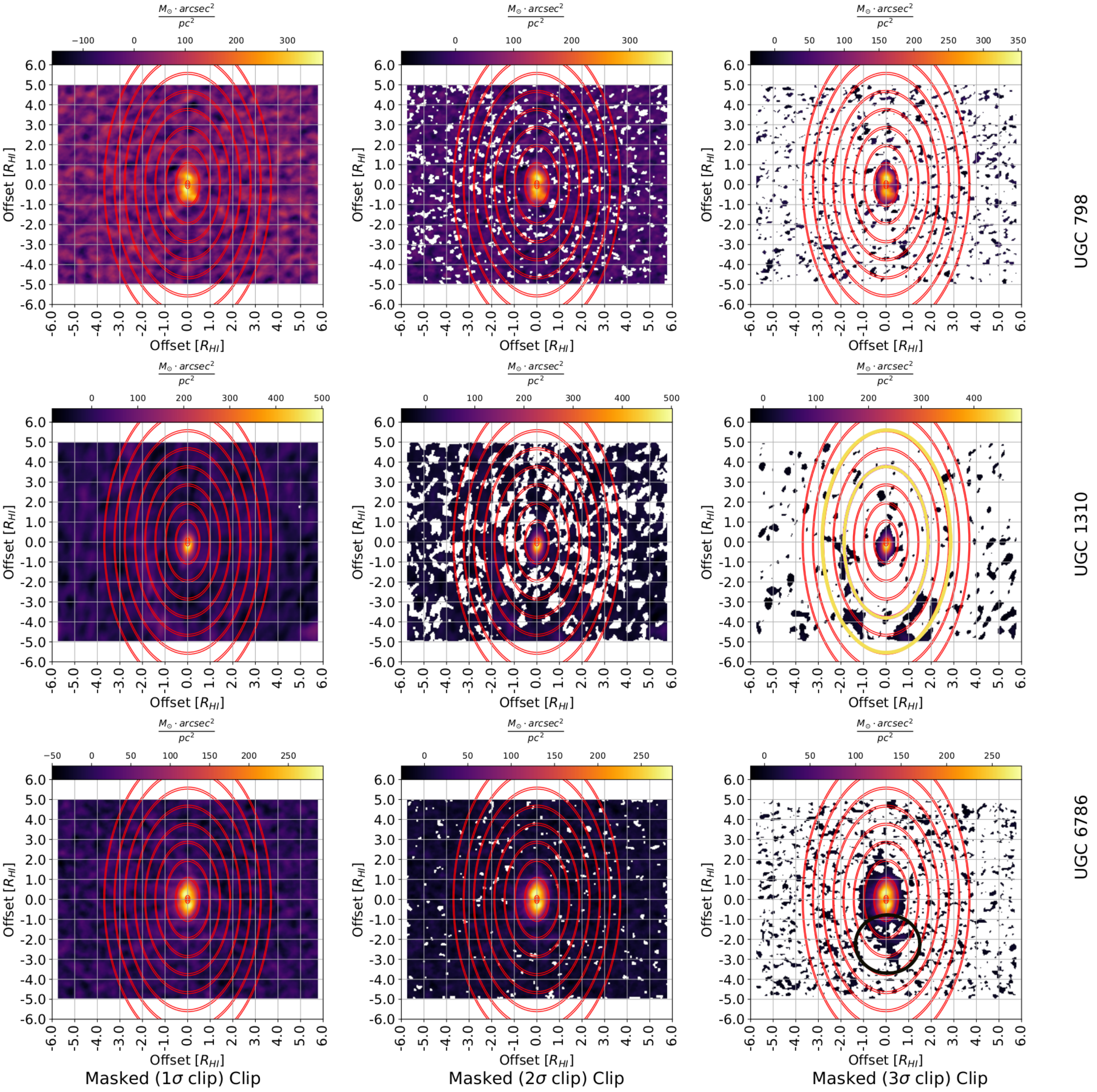


Figure 7.19: Group 6 masked moment 0 maps continued.

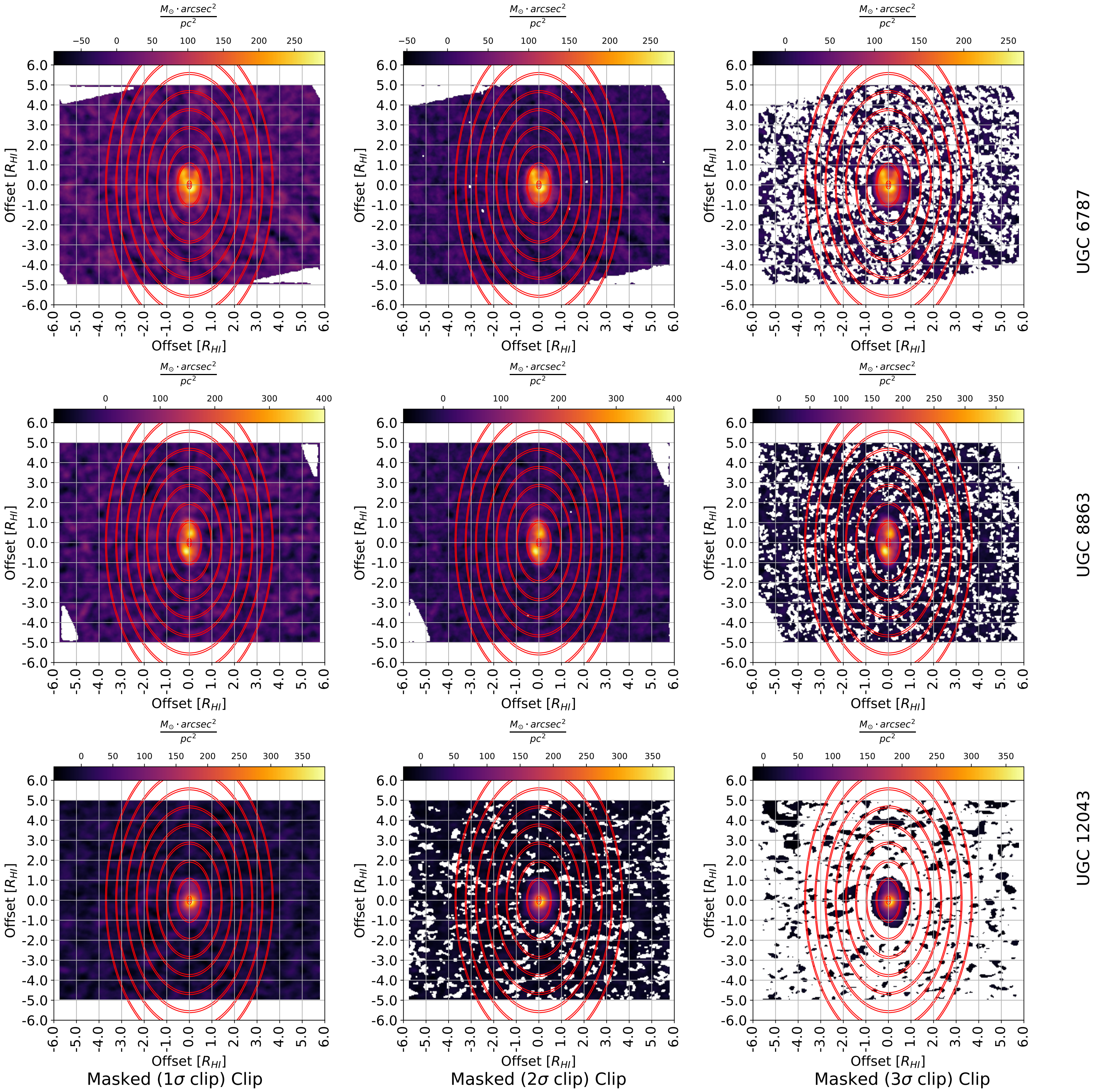


Figure 7.20: Below, Group 9 masked moment 0 maps. Each column visualises the moment 0 at the particular sigma-clipping level performed on the data cube. Each of the annuli at intervals of R_{HI} are overplotted, corresponding to the analysis in Chapter 4.5.2.

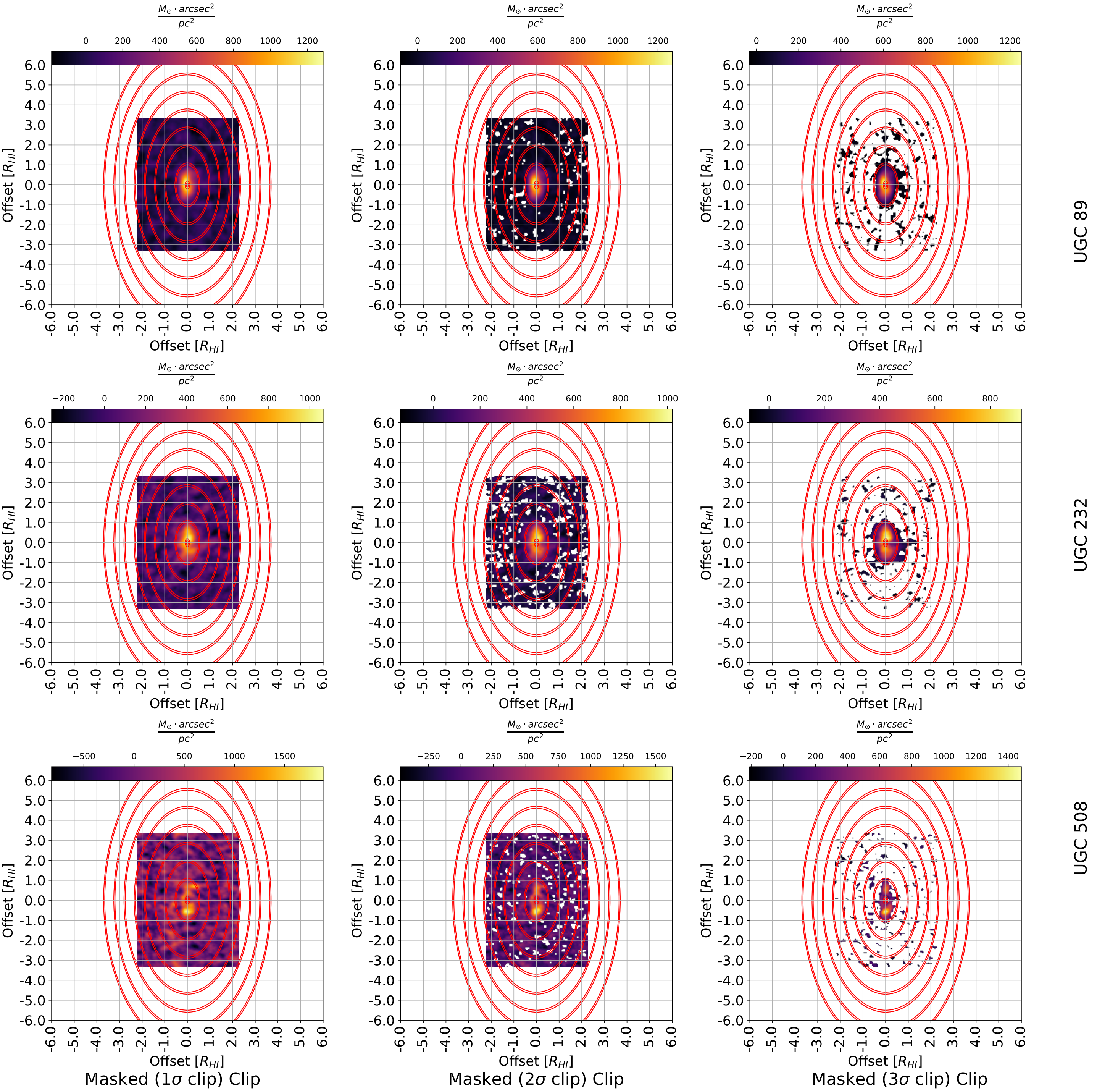


Figure 7.21: Group 9 masked moment 0 maps continued.

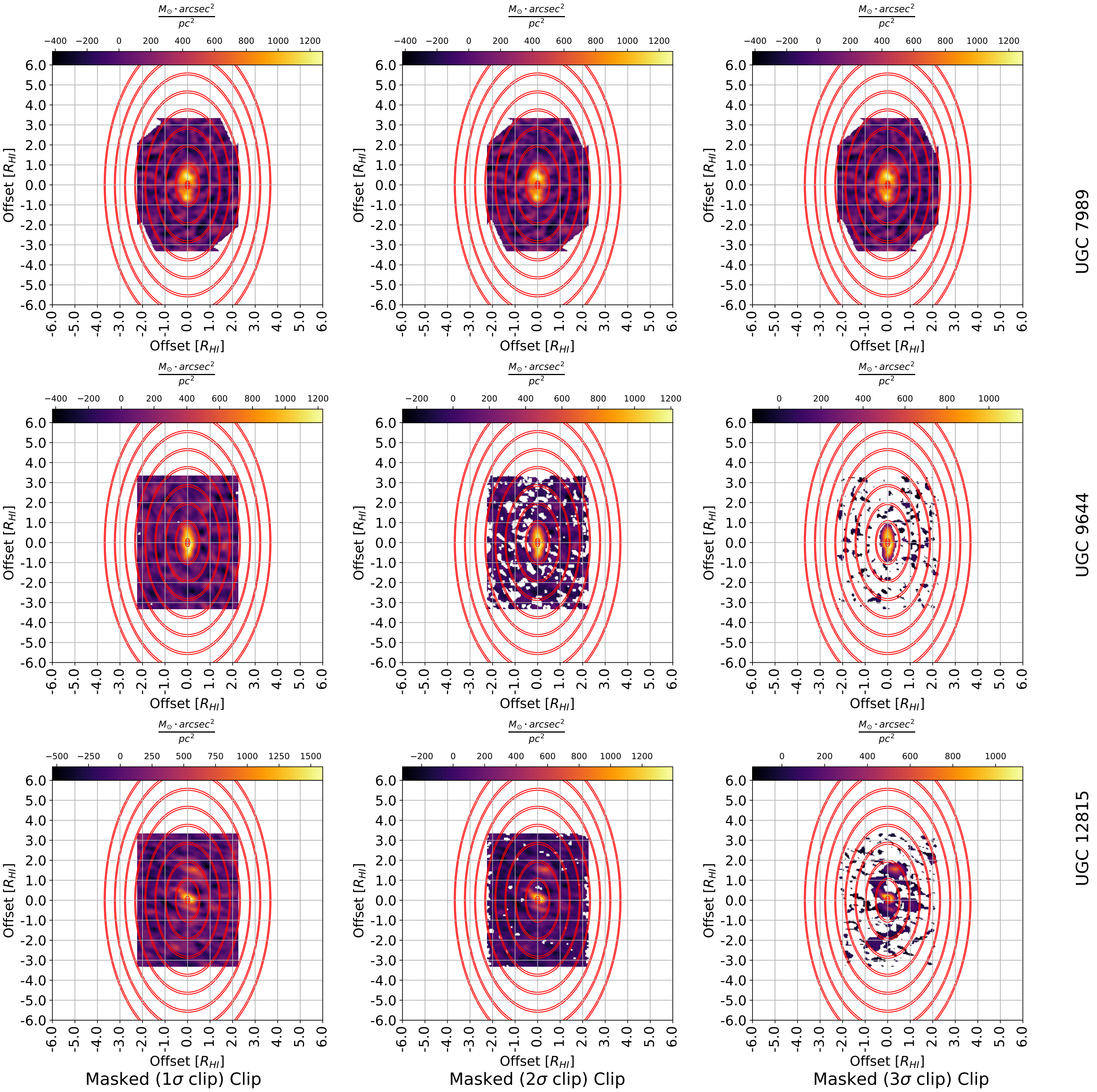


Figure 7.22: Below, Group 10 masked moment 0 maps. Each column visualises the moment 0 at the particular sigma-clipping level performed on the data cube. Each of the annuli at intervals of R_{HI} are overplotted, corresponding to the analysis in Chapter 4.6.2.

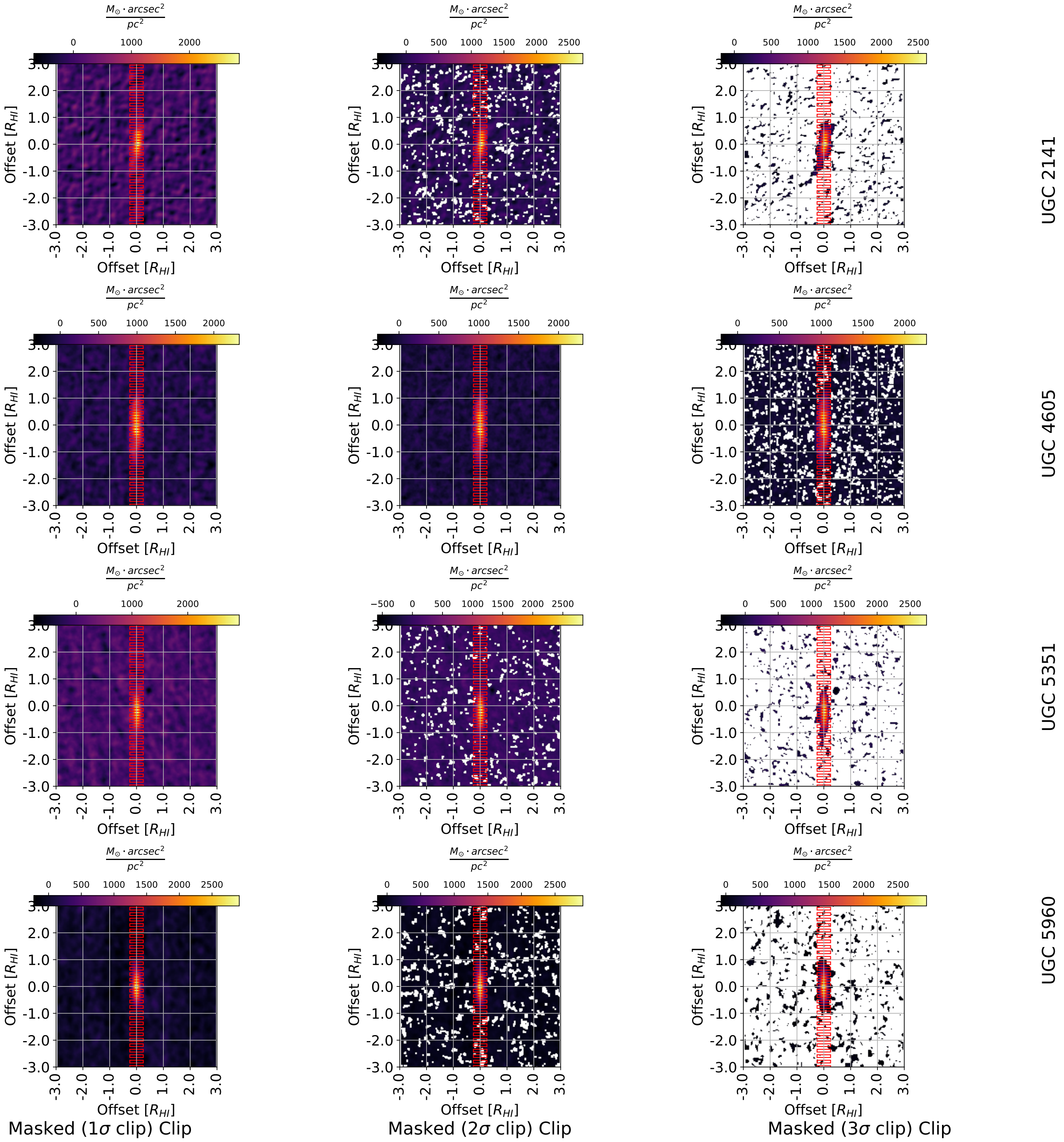
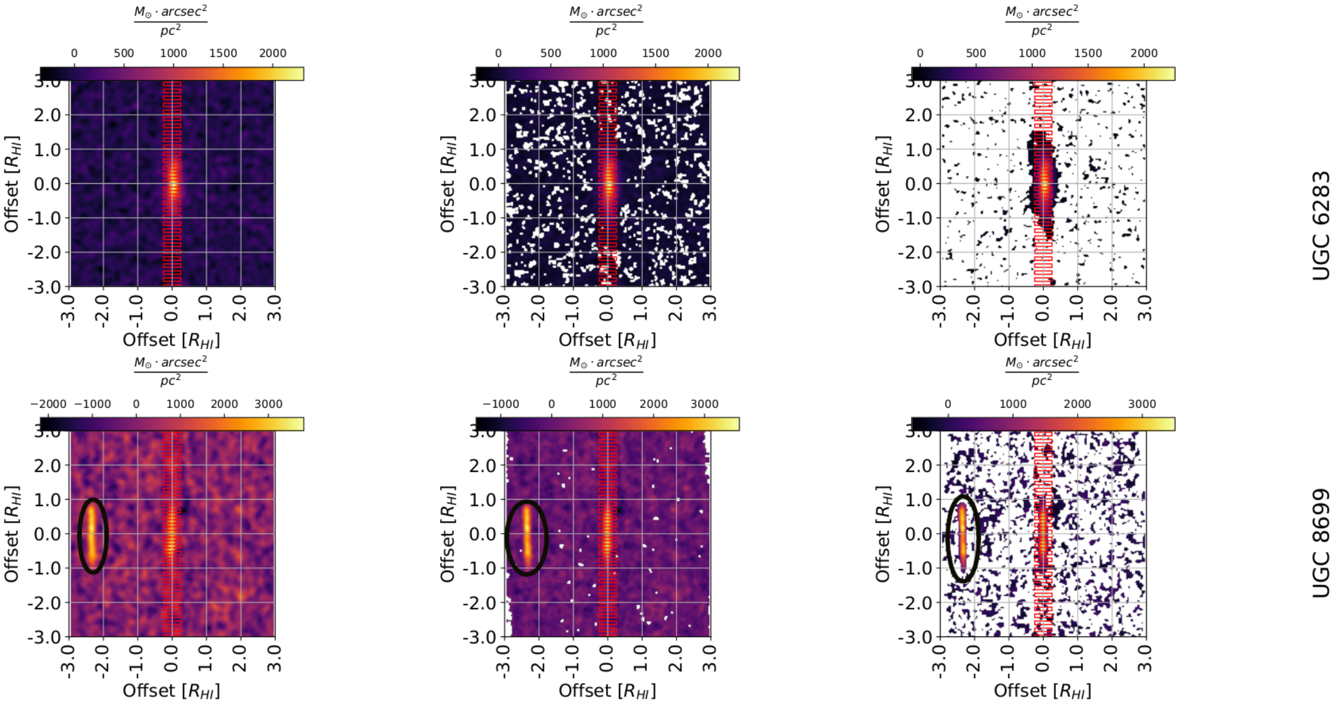


Figure 7.23: Group 10 masked moment 0 maps continued.



Masked (1σ clip) Clip

Masked (2σ clip) Clip

Masked (3σ clip) Clip

7.5 Light Concentrations from NSA Photometry

In this chapter, we present our initial explorations of the optical light distributions of the galaxies in our sample. We collect and present the bulge-disc decomposition present in the literature for galaxies in our WHISP-NSA sample. Furthermore, we calculate effective radii R_{20}, R_{50}, and R_{90} from the integrated combined model. We present the bulge-disc parameters from the literature.

7.5.1 NSA Photometry

Our initial explorations towards developing a pipeline for the stacking of radio data cubes based on optical measures of galaxy extent began with an investigation into the stellar luminosity profiles of our WHISP-NSA sample. Introduced in Chapter 2, the NASA-Sloan Atlas offers a comprehensive (and uniform) dataset comprising aperture photometry from which we derive surface-brightness profiles, as well as make use of existing measures of .

Restricting our attention to r-band photometry, to obtain the flux density profiles (I_{FD} in units of nanomaggies per square arcsecond), we divide the azimuthally-averaged flux by the corresponding aperture area:

$$I_{FD}(r) = \frac{f(r)}{A(r)}, \tag{7.11}$$

where f is the flux of an object, in nanomaggies, within an aperture A with an area in $arcseconds^2$. We then convert to standard magnitudes using the relation in Equation 7.12

$$\mu_{SB} = [22.5mag]2.5log_{10}\left\{I_{FD}(r)\right\}, \tag{7.12}$$

7.5.2 Light Concentrations

Included alongside the NSA's aperture photometry is the resulting Sérsic index (n) and effective-radius (r_{eff}) (Albareti et al., 2017). Obtained following the fitting of a two-dimensional Sérsic model of the form:

$$I_{sers}(x, y) = I_e \exp -b_n \left\{\left(\frac{r(x, y)}{r_{eff}}\right)^{\frac{1}{n}} - 1\right\} \tag{7.13}$$

Due to time constraints involved in performing our own lightcurve decompositions, we opted to use the Sérsic index from a two-dimensional fit as an estimate for the Sérsic index which would be obtained by fitting a one-dimensional Sérsic model (see Equation 7.16) used in conventional radial luminosity decomposition studies. The Sérsic indices of the one-dimensional fit and the two-dimensional fit are seldom identical. Factors such as the exact form of the Sérsic model used, noise in the data and instrumental effects can lead to slight discrepancies between the two indices, and we confirm this based on the discrepancy between the effective radii measurements in column (3) and our measured half-light radius (R50) in column (5) in Table 7.3.

However, in an ideal scenario with noise-free data and a perfect fit, the Sérsic index obtained from the two-dimensional fit should be consistent with the one derived from the one-dimensional profile.

Next, then, we integrate the galaxy's model Sérsic profile to obtain the curve-of growth (CoG). We obtain the integrated profile as:

$$F(r) = 2\pi \int_0^r I_{FD}(r')r'dr', \tag{7.14}$$

where $F(r)$ is the cumulative flux density (I_{FD}) within a radius r. By fitting a linear spline to this profile we derive the specific radii, namely r_{20}, r_{50}, r_{80}, and r_{90}. These radii denote the radial position enclosing 20%, 50%, 80%, and 90% of the galaxy's total luminosity, respectively.

Finally, we calculate and tabulate in Table 7.3, measures of luminosity concentration. For this we adopt the image concentration parameter C as defined by Kent (1985), based on the curve of growth:

$$C_{i/j} = 5\log_{10}\left(\frac{r_i}{r_j}\right) \tag{7.15}$$

Disc-Bulge Decompositions

Summarised in Table 7.2 are results from disc-bulge decomposition performed on the WHISP galaxies in our sample. The decomposition was performed in the R and Rc bands by Noordermeer & van der Hulst (2007) and Korsaga et al. (2018), respectively.

Among a myriad of decomposition techniques, galaxy photometric profiles can be modelled as the sum of an exponential disk and a chosen model for the bulge. Traditionally, this involved analyzing photometric profiles, typically employing a 1D approach by fitting an exponential or Sérsic model to these profiles. While this method allows rapid analysis of numerous galaxies, it is hindered by projection effects, where the observed intensity in a galaxy comprises contributions from both bulge and disk at different radii, resulting in systematic errors when deriving their properties.

To address these limitations, Noordermeer & van der Hulst (2007) have devised a more accurate 2D decomposition method. This approach utilizes the full 2D galaxy distribution to separate the galaxy components into bulge and disk components. Employing a flattened axisymmetric spheroidal Sérsic profile (Sersic, 1968) for the extended regions, the bulge parameters are derived from the residual of the raw data and a subtracted disk-estimate. This process involves creating a model bulge based on fitted parameters, subtracting these from the originals, and attributing the remaining light to the disks. In terms of surface density, in magnitude units, these models are described by:

$$\mu_b(r) = \mu_e + 1.0857 b_n \left\{ \left(\frac{r}{r_e}\right)^{\frac{1}{n}} - 1 \right\} \tag{7.16}$$

$$\mu_d(r) = \mu_{0,d} + 1.0857 \left(\frac{r}{h}\right) \tag{7.17}$$

where can be approximated as $bn1.9992n0.3271$ for 1<n<10 (Graham, 2001).

The authors begin by estimating bulge and disk parameters from a photometric profile. Due to projection effects, the central profile areas, dominated by bulges, face challenges, while the outer regions, mainly governed by flat disks, provide reliable initial disk estimations.

After generating an initial model of the disk, see Equation 7.17, subtracting it from the original image isolates bulge-related light. Bulge characteristics, like flattening and alignment with the disk's symmetry plane, are then determined through ellipse fitting.

Further steps involve obtaining bulge photometric profiles across colour bands and fitting these profiles to derive their contributions at larger radii. The method accounts for residuals remaining after bulge subtraction, attributing any remaining colour gradients to the disk. The authors model the disk as:

Finally, the radial distribution of the disk light is determined by measuring average intensities on the ellipses. This advanced approach offers a more accurate way to separate and analyze the bulge and disk components within spiral galaxies, mitigating the limitations of traditional 1D methods affected by projection effects and parameter degeneracy. In Table 7.2, we present the R-band results of a selection of galaxies in our WHISP-NSA sample performed by Noordermeer & van der Hulst (2007).

Galaxies U508, U2141, U2183, and U6118 are supplemented by Korsaga et al. (2018). Motivated by the same projection effects, Korsaga et al. (2018) make use of a 2D fitting Python routine to decompose the galaxy surface-brightness profiles. Unlike the simpler two-component decomposition employed by Noordermeer & van der Hulst (2007), their Python routine decomposes, where appropriate, the surface-brightness profile into multiple components (among which include the option of a disc, bulge, bar, spiral arm, ring, lens). With the exception of galaxies which display truncations or breaks in the surface-brightness profile, Korsaga et al. (2018) adopt identical models for the bulge and disc to Noordermeer & van der Hulst (2007), summarised in Equation 7.16 and Equation 7.17, respectively. In Equation 7.16, Korsaga et al. (2018) adopt a similar approximation of $bn2n0.33$.

UGC	Disk Parameters μ_0 $\frac{mag}{arcsec^2}$	 h kpc	Bulge Parameters μ_e $\frac{mag}{arcsec^2}$	 r_e arcsec	 n	Band	Source
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
89	20.19	6.65	17.53	2.90	1.20	R	Noordermeer & van der Hulst (2007)
94	20.04	3.61	21.43	3.70	2.50	R	Noordermeer & van der Hulst (2007)
508	19.40	7.50	17.75	1.95	2.16	Rc	Korsaga et al. (2018)
624	21.17	5.78	21.10	15.00	3.70	R	Noordermeer & van der Hulst (2007)
2141	19.30	1.20	20.32	8.37	0.61	Rc	Korsaga et al. (2018)
2183	20.70	2.30	19.05	6.14	2.35	Rc	Korsaga et al. (2018)
2487	20.35	7.97	19.89	6.70	1.70	R	Noordermeer & van der Hulst (2007)
4458	21.21	8.62	19.21	5.80	2.60	R	Noordermeer & van der Hulst (2007)
6118	18.10	1.50	16.77	1.80	0.64	Rc	Korsaga et al. (2018)
6786	18.27	1.54	22.43	43.50	5.50	R	Noordermeer & van der Hulst (2007)
6787	19.76	3.32	18.43	8.80	2.30	R	Noordermeer & van der Hulst (2007)
8699	20.65	3.70	19.58	6.90	2.70	R	Noordermeer & van der Hulst (2007)
11852	20.37	4.50	20.55	5.10	2.20	R	Noordermeer & van der Hulst (2007)
11914	20.03	2.66	20.90	26.20	3.10	R	Noordermeer & van der Hulst (2007)
12043	19.84	0.86	0.00	0.00	0.00	R	Noordermeer & van der Hulst (2007)

Table 7.2: Compilation of Disk-Bulge Decomposition for WHISP-NSA Galaxies. (1) UGC ID, (2) disk central surface brightness, (3) disk scale length, (4) bulge effective surface brightness, (5) bulge effective radius,(6) bulge Sérsic paramete, (7) colour band, (8) Literature Source.

In lieu of developing a data-cube stacking profile which relies solely on optical parameters, performing bulge-disc decomposition as performed by Noordermeer & van der Hulst (2007) and Korsaga et al. (2018) provides several more measures of mass concentration. An investigation of relations between HI distribution and optical measures of central concentration may prove to be an interesting avenue for future research. Specifically for predicting the rising rates of HI rotation curves, as outlined in Chapter 5.2.1, whereby the mass-concentration in the central regions is expected to dominate the kinematic profile of the galaxies at these radial offsets.

UGC (1)	n (2)	R_{eff} arcsec (3)	R_{20} arcsec (4)	R_{50} arcsec (5)	R_{80} arcsec (6)	R_{90} arcsec (7)	$C_{80/20}$ (8)	$C_{90/50}$ (9)
89	4.83	13.54	4.47	15.07	49.47	88.03	5.22	3.83
94	1.47	16.96	7.63	17.19	34.43	47.56	3.27	2.21
232	1.98	17.2	7.06	17.53	39.08	57.11	3.72	2.56
508	4.9	30.86	8.19	29.69	94.70	157.41	5.32	3.62
624	2.85	22.42	7.95	22.88	59.46	93.77	4.37	3.06
798	5.02	29.18	7.74	28.31	91.66	153.73	5.37	3.67
1310	1.14	7.27	3.72	7.51	13.73	18.13	2.84	1.91
2141	1.34	22.89	10.50	23.10	44.88	60.99	3.15	2.11
2183	3.99	19.32	6.17	20.13	60.73	102.73	4.96	3.54
2487	4.36	29.72	8.43	29.14	89.29	147.57	5.12	3.52
4458	5.89	19.83	5.48	20.75	73.58	130.85	5.64	4.00
4605	2.33	22.86	8.72	23.21	55.40	83.75	4.01	2.79
5060	4.49	10.56	3.82	12.12	38.14	67.23	5	3.72
5351	0.93	27.31	13.72	27.35	47.83	61.51	2.71	1.76
5906	2.07	20.22	8.10	20.55	46.70	68.90	3.8	2.63
5960	1.67	16.28	7.07	16.55	34.61	48.91	3.45	2.35
6118	3.64	40.94	12.10	38.95	108.28	169.43	4.76	3.19
6283	2.26	27.51	10.50	27.78	65.48	98.21	3.97	2.74
6786	3	31.96	10.72	31.86	83.98	131.68	4.47	3.08
6787	4.99	39.31	9.75	35.72	111.39	179.34	5.29	3.50
7989	5.36	48.77	10.94	41.12	126.94	198.72	5.32	3.42
8271	1.77	16.42	7.01	16.70	35.63	50.87	3.53	2.42
8699	6	20.53	5.57	21.31	75.90	134.68	5.67	4.00
8863	3.59	47.9	13.93	44.45	120.77	184.90	4.69	3.10
9644	6	88.51	15.16	57.59	164.09	237.26	5.17	3.07
11852	2.03	12.37	5.19	12.77	28.57	41.91	3.7	2.58
11914	2.56	32.52	11.70	32.58	80.48	122.90	4.19	2.88
12043	1.22	16.59	7.89	16.79	31.63	42.30	3.01	2.01
12713	0.86	16.37	8.43	16.43	28.15	35.82	2.62	1.69
12815	2.19	28.28	10.92	28.53	66.37	98.88	3.92	2.70

Table 7.3: Summary of Galaxy Concentration Measurements for WHISP-NSA Galaxies. (1) UGC ID, (2) Sérsic index from NSA modelling, (3) Effective radius from Sérsic modelling, (4)-(7) effective radii containing 20%, 50%, 80%, and 90% of the total light, (8) concentration index between r_{80} and r_{20} specific radii,(9)concentration index between r_{90} and r_{50} radii.

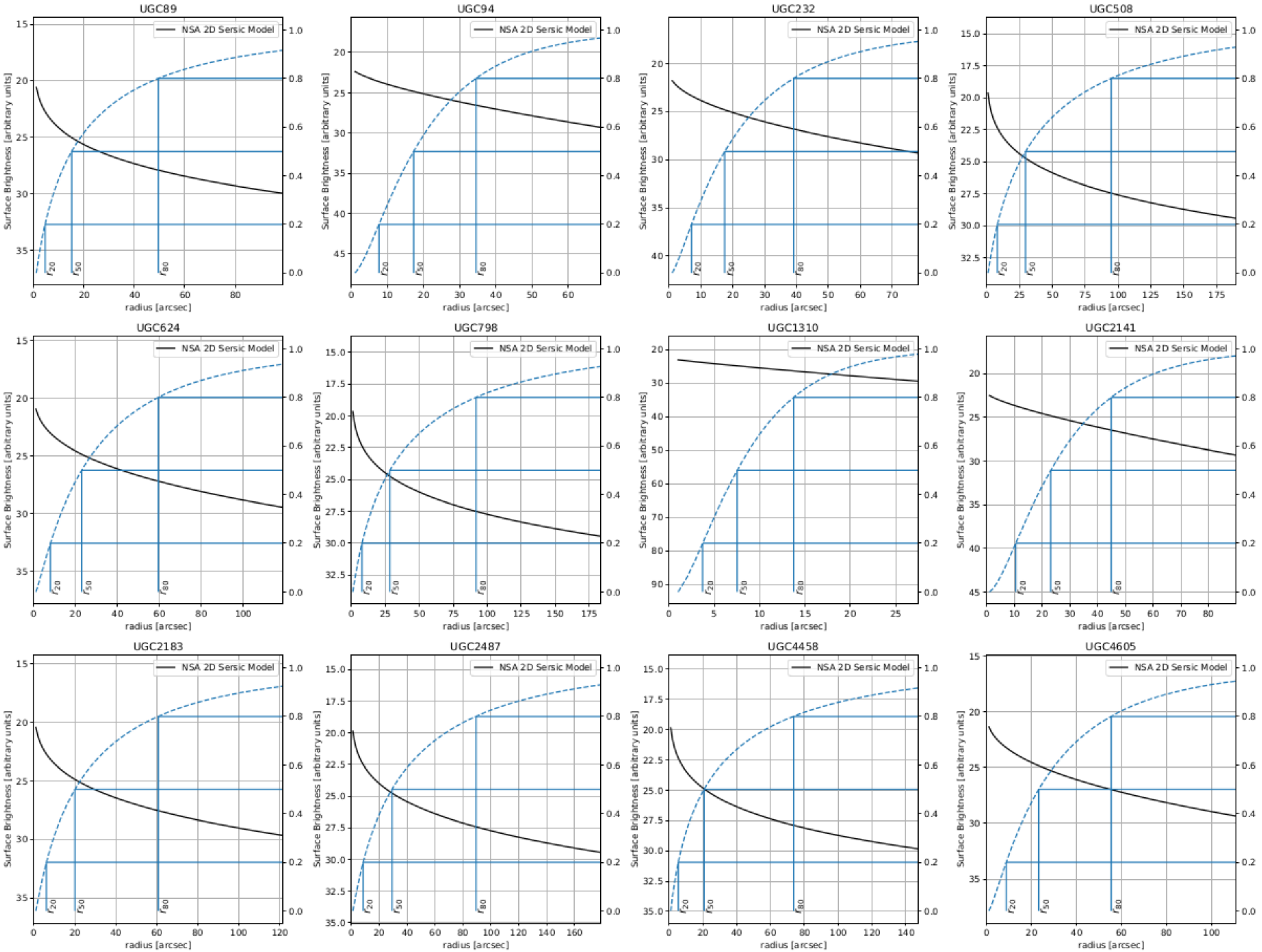


Figure 7.24: Compilation of WHISP galaxies with SDSS Photometry for which two-dimensional Sérsic models have been fit. We present the specific radii (R20, R50, R80, and R90) as measured from the one-dimensional Sérsic profile's Curve-of-Growth (dashed blue curve). We represent the measured specific radii as a the vertical and horizontal lines on each subplot.

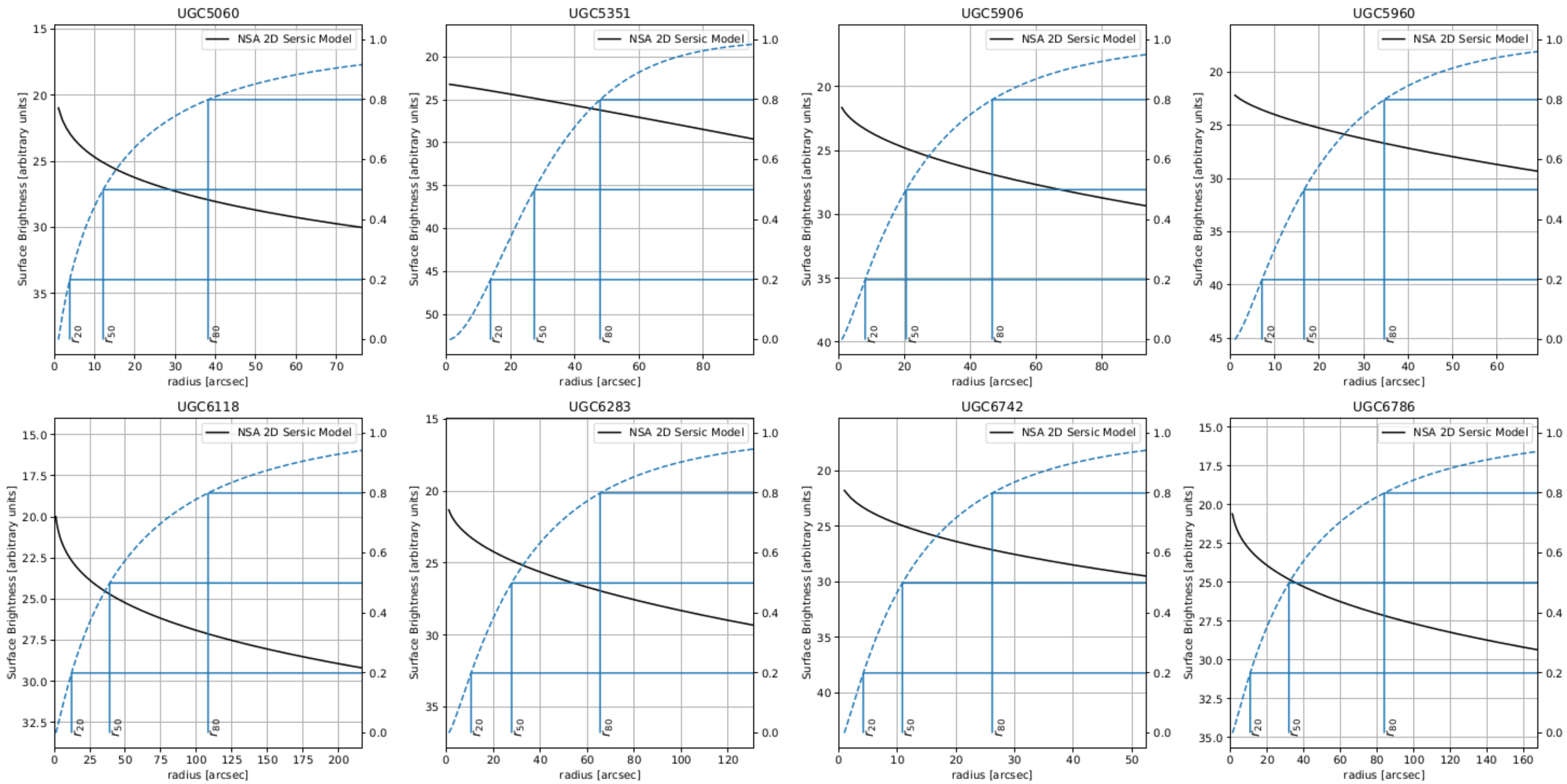


Figure 7.25: Compilation of WHISP galaxies with SDSS Photometry for which two-dimensional Sérsic models have been fit. We present the specific radii (R20, R50, R80, and R90) as measured from the one-dimensional Sérsic profile's Curve-of-Growth (dashed blue curve). We represent the measured specific radii as a the vertical and horizontal lines on each subplot

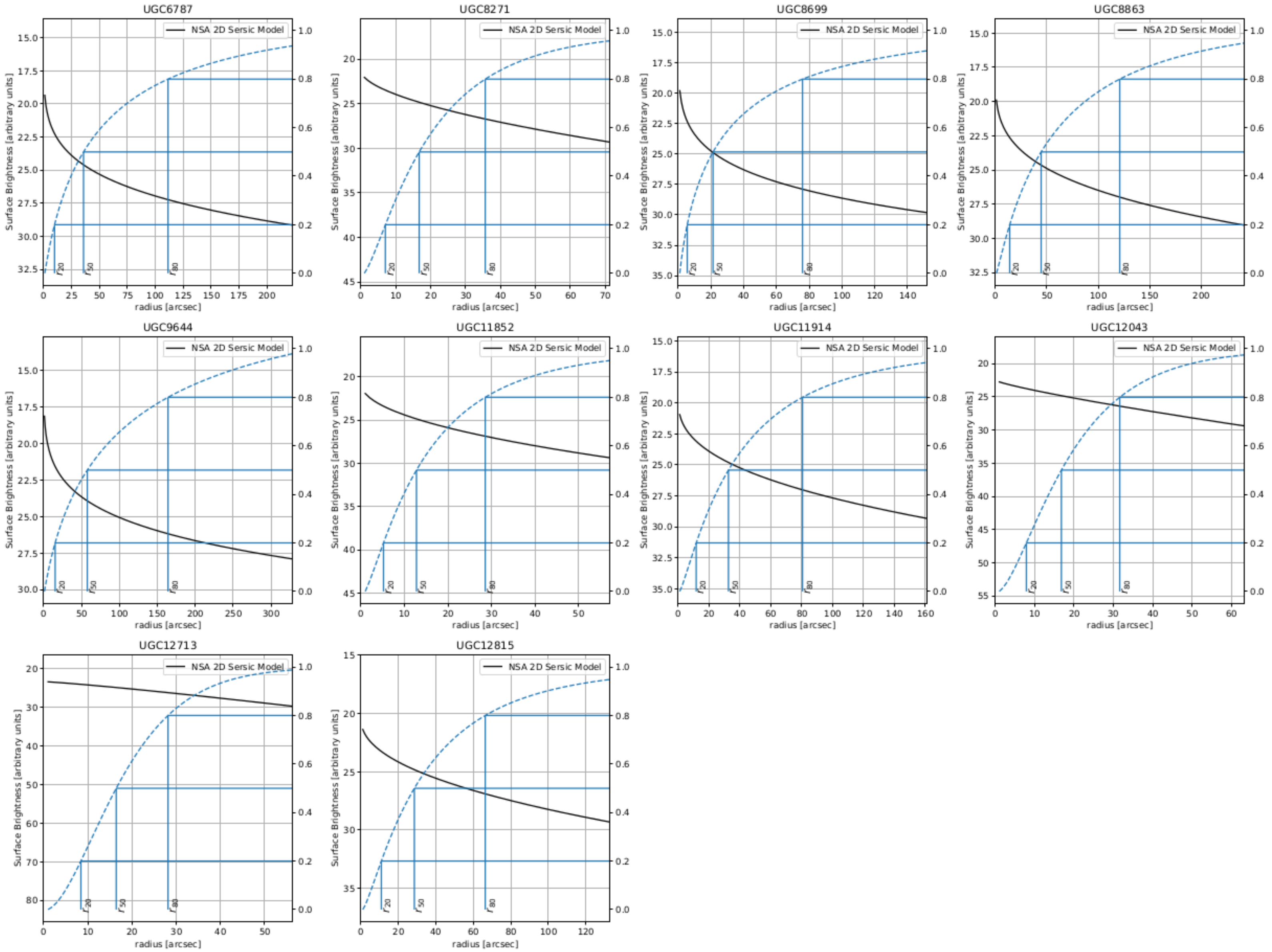


Figure 7.26: Compilation of WHISP galaxies with SDSS Photometry for which two-dimensional Sérsic models have been fit. We present the specific radii (R20, R50, R80, and R90) as measured from the one-dimensional Sérsic profile's Curve-of-Growth (dashed blue curve). We represent the measured specific radii as a the vertical and horizontal lines on each subplot

7.6 Auxiliary Tables

UGC	Channel Width (Hz)	Scaling Factor
89	-39062.5	1.0
94	-39062.5	1.0
232	-39062.5	1.0
508	-39062.5	1.0
624	-39062.5	1.0
798	-39062.5	1.0
1310	-39062.5	1.0
2141	-19531.25	0.5
2183	-39062.5	1.0
2487	-39062.5	1.0
4458	-39062.5	1.0
4605	-39062.5	1.0
5060	-19531.25	0.5
5351	-39062.5	1.0
5906	-19531.25	0.5
5960	-19531.25	0.5
6001	-19531.25	0.5
6118	-39062.5	1.0
6283	-19531.25	0.5
6742	-19531.25	0.5
6786	-39062.5	1.0
6787	-39062.5	1.0
7704	-19531.25	0.5
7989	-78125.0	2.0
8271	-19531.25	0.5
8699	-39062.5	1.0
8863	-39062.5	1.0
9644	-19531.25	0.5
11852	-39062.5	1.0
11914	-39062.5	1.0
12043	-19531.25	0.5
12713	-19531.25	0.5
12815	-39062.5	1.0

Table 7.4: Table summarising the spectral-axis scaling factors used to standardised the spectral sampling of the data-set. The fiducial spectral sampling was set at $CDELT3 = -39062.5$ Hz.

UGC	alternative name	Type	RA [J2000]	DEC [J2000]	V$_{sys}$	PA	i	R$_{HI}$	M$_{HI}$	W$_{20}$	σ_{map}
			°	°	km/s	°	°	kpc	$M_\odot$	km/s	$10^{19} atoms \cdot cm^{-2}$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
89	NGC23	SB(s)a	2.47	25.92	4555	170	50	24	8.63	440	3.25
94	NGC26	SA(rs)ab	2.61	25.83	4589	95	42	26	11.28	319	2.86
232	N/A	SB(r)a	6.16	33.26	4839	45	48	27	7.95	276	4.97
508	NGC266	SB(rs)ab	12.45	32.28	4647	120	25	29	7.40	478	2.82
624	NGC338	Sab	15.15	30.67	4772	109	59	29	14.42	560	11.3
798	IC1654	(R)SB(r)a	18.80	30.19	4896	45	40	23	4.08	220	3.13
1310	NGC694	S0? pec	27.74	22.00	2958	160	47	9.7	2.33	186	3.66
2141	NGC1012	S0/a?	39.81	30.15	987	24	90	10	2.30	232	71.8
2183	NGC1056	Sa:	40.70	28.57	1540	147	90	17	3.63	292	6.52
2487	NGC1167	SA0-	45.43	35.21	4950	252	36	39	17.09	466	1.29
4458	NGC2599	SAa	128.05	22.56	4757	106	27	34	11.82	284	3.14
4605	NGC2654	SBab: sp	132.30	60.22	1350	243	90	28	5.99	432	18.4
5060	NGC2893	(R)SB0/a	142.57	29.54	1699	170	24	7	0.70	185	2.04
5351	NGC3067	SAB(s)ab?	149.59	32.37	1486	107	70	6.3	0.99	276	4.11
5906	NGC3380	(Râ)SBa?	162.05	28.60	1601	36	29	5.3	0.31	123	2.04
5960	NGC3413	S0	162.84	32.77	643	178	77	5.7	0.75	180	6.45
6001	NGC3442	Sa?	163.28	33.91	1730	30	47	5.3	0.33	153	5.07
6118	NGC3504	(R)SAB(s)ab	165.80	27.97	1536	163	24	9.5	0.84	219	4.35
6283	NGC3600	Sa?	168.97	41.59	713	9	90	12	2.00	216	28.9
6742	NGC3870	S0?	176.49	50.20	752	58	36	3.2	0.17	127	1.84
6786	NGC3900	SA(r)0+	177.29	27.02	1799	182	65	18	3.93	445	1.48
6787	NGC3898	SA(s)ab	177.31	56.08	1171	110	67	23	3.96	484	2.14
7704	NGC4509	Sab pec?	188.28	32.09	937	37	60	4.3	0.34	93	4.69
7989	NGC4725	SAB(r)ab pec	192.61	25.50	1208	32	51	34	9.78	417	9.73
8271	NGC5014	Sa? sp	197.88	36.28	1128	49	60	7.7	0.78	166	0.98
8699	NGC5289	(R)SABab:	206.29	41.50	2521	280	72	17	3.26	387	12.4
8863	NGC5377	(R)SB(s)a	209.07	47.24	1791	210	51	19	2.52	391	2.11
9644	N/A	SB(r)a	224.89	27.12	6665	126	19	24	7.16	131	9.41
11852	N/A	SBa?	329.00	27.90	5846	-92	50	53	29.60	329	4.12
11914	NGC7217	(R)SA(r)ab	331.97	31.36	949	88	31	8.1	0.69	322	2.25
12043	NGC7286	S0/a	336.96	29.10	1007	160	67	10.8	1.18	190	9.50
12713	N/A	S0/a	354.56	30.71	295	68	72	2.3	0.08	147	1.78
12815	NGC7771	SB(s)a	357.85	20.11	4307	0	61	30	9.36	656	9.62

Table 7.5: The Sample: basic parameters of scale. (1) UGC number, (2) alternative name, (3) morphological type, (4) RA, (5) Dec, (6) systemic velocity,(7) position-angle, (8) inclination angle, (9) HI radius in kpc, (10) HI mass in solar masses, (11) line width of global profile at the 20% level, (12) 1 sigma column density in the (Noordermeer et al., 2005) HI maps.

Bibliography

Albareti, F. D., Allende Prieto, C., Almeida, A., et al. 2017, ApJS, 233, 25, doi: 10.3847/1538-4365/aa8992

Andersen, D. R., & Bershady, M. A. 2013, The Astrophysical Journal, 768, 41, doi: 10.1088/0004-637x/768/1/41

Arras, P., Frank, P., Leike, R., Westermann, R., & Enßlin, T. A. 2019, Astronomy amp; Astrophysics, 627, A134, doi: 10.1051/0004-6361/201935555

Baker, A. J., Blyth, S., Holwerda, B. W., & LADUMA Team. 2018, in American Astronomical Society Meeting Abstracts, Vol. 231, American Astronomical Society Meeting Abstracts #231, 231.07

Begeman, K. G. 1987, PhD thesis, University of Groningen, Kapteyn Astronomical Institute

Bianchi, L., & GALEX Team. 2000, Mem. Soc. Astron. Italiana, 71, 1123

Bigiel, F. 2010, in From Stars to Galaxies: Connecting our Understanding of Star and Galaxy Formation, 150

Binney, J., Dehnen, W., & Bertelli, G. 2000, Monthly Notices of the Royal Astronomical Society, 318, 658, doi: 10.1046/j.1365-8711.2000.03720.x

Blanton, M. R., Kazin, E., Muna, D., Weaver, B. A., & Price-Whelan, A. 2011, AJ, 142, 31, doi: 10.1088/0004-6256/142/1/31

Braun, R., Bonaldi, A., Bourke, T., Keane, E., & Wagg, J. 2019, Anticipated Performance of the Square Kilometre Array – Phase 1 (SKA1). https://arxiv.org/abs/1912.12699

Broeils, A. H. 1992, PhD thesis, University of Groningen. http://www.rug.nl/

Broeils, A. H., & Rhee, M. H. 1997, A&A, 324, 877

Caldú-Primo, A., Schruba, A., Walter, F., et al. 2013, The Astronomical Journal, 146, doi: 10.1088/0004-6256/146/6/150

Cannon, J. M., Giovanelli, R., Haynes, M. P., et al. 2011, ApJ, 739, L22, doi: 10.1088/2041-8205/739/1/L22

Catinella, B., Schiminovich, D., Kauffmann, G., et al. 2010, Monthly Notices of the Royal Astronomical Society, 403, 683, doi: 10.1111/j.1365-2966.2009.16180.x

Chen, Q., Meyer, M., Popping, A., & Staveley-Smith, L. 2021, Monthly Notices of the Royal Astronomical Society, 502, 2308, doi: 10.1093/mnras/stab168

Cornwell, T. J. 2008, IEEE Journal of Selected Topics in Signal Processing, 2, 793, doi: 10.1109/JSTSP.2008.2006388

Cornwell, T. J. 2009, A&A, 500, 65, doi: 10.1051/0004-6361/200912148

Courteau, S. 1997, The Astronomical Journal, 114, 2402, doi: 10.1086/118656

Dale, D. A., Smith, J. D. T., Schlawin, E. A., et al. 2009, The Astrophysical Journal, 693, 1821, doi: 10.1088/0004-637x/693/2/1821

de Blok, W. J. G., Walter, F., Brinks, E., et al. 2008, AJ, 136, 2648, doi: 10.1088/0004-6256/136/6/2648

de Blok, W. J. G., Walter, F., Brinks, E., et al. 2008, The Astronomical Journal, 136, 2648, doi: 10.1088/0004-6256/136/6/2648

de Blok, W. J. G., Adams, E. A. K., Amram, P., et al. 2016, in MeerKAT Science: On the Pathway to the SKA, 7, doi: 10.22323/1.277.0007

Delhaize, J., Meyer, M. J., Staveley-Smith, L., & Boyle, B. J. 2013, Monthly Notices of the Royal Astronomical Society, 433, 1398, doi: 10.1093/mnras/stt810

Dénes, H., Kilborn, V. A., & Koribalski, B. S. 2014, MNRAS, 444, 667, doi: 10.1093/mnras/stu1337

Donahue, M., & Voit, G. M. 2022, Physics Reports, 973, 1, doi: https://doi.org/10.1016/j.physrep.2022.04.005

Dressler, A. 1980, ApJ, 236, 351, doi: 10.1086/157753

Eke, V. R., Baugh, C. M., Cole, S., et al. 2005, Monthly Notices of the Royal Astronomical Society, 362, 1233, doi: 10.1111/j.1365-2966.2005.09384.x

Elson, E. C. 2013, Monthly Notices of the Royal Astronomical Society, 437, 3736, doi: 10.1093/mnras/stt2182

Fabello, S. 2012, PhD thesis, Ludwig-Maximilians University of Munich, Germany

Fernandez, X., van Gorkom, J. H., Momjian, E., & Chiles Team. 2015, in American Astronomical Society Meeting Abstracts, Vol. 225, American Astronomical Society Meeting Abstracts #225, 427.03

Fernández, X., Gim, H. B., van Gorkom, J. H., et al. 2016, ApJ, 824, L1, doi: 10.3847/2041-8205/824/1/L1

Fitzgibbon, A., Pilu, M., & Fisher, R. 1996, in Proceedings of 13th International Conference on Pattern Recognition, Vol. 1, 253–257 vol.1, doi: 10.1109/ICPR.1996.546029

Fraternali, F., van Moorsel, G., Sancisi, R., & Oosterloo, T. 2002, The Astronomical Journal, 123, 3124, doi: 10.1086/340358

Gereb, K., Morganti, R., Oosterloo, T., Hoppmann, L., & Staveley-Smith, L. 2015, Astronomy & Astrophysics, 580, A43, doi: 10.1051/0004-6361/201424810

Geréb, K., Morganti, R., Oosterloo, T. A., Guglielmino, G., & Prandoni, I. 2013, Astronomy & Astrophysics, 558, A54, doi: 10.1051/0004-6361/201322113

Giovanelli, R., & Haynes, M. P. 2002, ApJ, 571, L107, doi: 10.1086/341368

Giovanelli, R., Haynes, M. P., Kent, B. R., et al. 2005, The astronomical journal, 130, 2598

Graham, A. W. 2001, The Astronomical Journal, 121, 820, doi: 10.1086/318767

Heald, G. H., Rand, R. J., Benjamin, R. A., & Bershady, M. A. 2007, The Astrophysical Journal, 663, 933, doi: 10.1086/518087

Healy, J., Blyth, S.-L., Elson, E., et al. 2019, Monthly Notices of the Royal Astronomical Society, 487, 4901, doi: 10.1093/mnras/stz1555

Hoekstra, H., van Albada, T. S., & Sancisi, R. 2001, Monthly Notices of the Royal Astronomical Society, 323, 453, doi: 10.1046/j.1365-8711.2001.04214.x

Högbom, J. A. 1974, A&AS, 15, 417

Holwerda, B. W., Pirzkal, N., de Blok, W. J. G., et al. 2011, Monthly Notices of the Royal Astronomical Society, 416, 2401–2414, doi: 10.1111/j.1365-2966.2011.18938.x

J. N Chengalur, Braun, R., & Wieringa, M. 2001, Astronomy & Astrophysics, 372, 768, doi: 10.1051/0004-6361:20010547

Kalnajs, A. J. 1983, in Internal Kinematics and Dynamics of Galaxies, ed. E. Athanassoula, Vol. 100, 109–115

Kamphuis, P., Jütte, E., Heald, G. H., et al. 2022, Astronomy & Astrophysics, 668, A182, doi: 10.1051/0004-6361/202140704

Kang, J., Zhu, M., Ai, M., Yu, H., & Sun, C. 2022, Research in Astronomy and Astrophysics, 22, 065019, doi: 10.1088/1674-4527/ac6796

Kennicutt, Robert C., J. 1998, ApJ, 498, 541, doi: 10.1086/305588

Kent, S. M. 1985, ApJS, 59, 115, doi: 10.1086/191066

—. 1986, AJ, 91, 1301, doi: 10.1086/114106

Kereš, D., Katz, N., Weinberg, D. H., & Davé, R. 2005, Monthly Notices of the Royal Astronomical Society, 363, 2, doi: 10.1111/j.1365-2966.2005.09451.x

Kereš, D., & Hernquist, L. 2009, The Astrophysical Journal, 700, L1, doi: 10.1088/0004-637X/700/1/L1

Koribalski, B. S., Staveley-Smith, L., Westmeier, T., et al. 2020a, Ap&SS, 365, 118, doi: 10.1007/s10509-020-03831-4

—. 2020b, Ap&SS, 365, 118, doi: 10.1007/s10509-020-03831-4

Korsaga, M., Amram, P., Carignan, C., & Epinat, B. 2018, Monthly Notices of the Royal Astronomical Society, 482, 154–174, doi: 10.1093/mnras/sty2582

Kregel, M., & van der Kruit, P. C. 2004, MNRAS, 352, 787, doi: 10.1111/j.1365-2966.2004.07978.x

Lah, P., Chengalur, J. N., Briggs, F. H., et al. 2007, Monthly Notices of the Royal Astronomical Society, 376, 1357, doi: 10.1111/j.1365-2966.2007.11540.x

Leclercq, F., Verhamme, A., Epinat, B., et al. 2022, Astronomy & Astrophysics, 663, A11, doi: 10.1051/0004-6361/202142179

Lelli, F., Verheijen, M., & Fraternali, F. 2014, A&A, 566, A71, doi: 10.1051/0004-6361/201322657

Loomis, R. A., Öberg, K. I., Andrews, S. M., et al. 2018, AJ, 155, 182, doi: 10.3847/1538-3881/aab604

Maddox, N., Frank, B. S., Ponomareva, A. A., et al. 2021, A&A, 646, A35, doi: 10.1051/0004-6361/202039655

Makarov, D., Prugniel, P., Terekhova, N., Courtois, H., & Vauglin, I. 2014, A&A, 570, A13, doi: 10.1051/0004-6361/201423496

Marasco, A., & Fraternali, F. 2011, AA, 525, A134, doi: 10.1051/0004-6361/201015508

Mathewson, D. S., & Ford, V. L. 1996, ApJS, 107, 97, doi: 10.1086/192356

Mathewson, D. S., Ford, V. L., & Buchhorn, M. 1992, ApJS, 81, 413, doi: 10.1086/191700

Meyer, M. 2009. https://arxiv.org/abs/0912.2167

Meyer, M., Robotham, A., Obreschkow, D., et al. 2017, Publications of the Astronomical Society of Australia, 34, doi: 10.1017/pasa.2017.31

Meyer, S. A., Meyer, M., Obreschkow, D., & Staveley-Smith, L. 2016, Monthly Notices of the Royal Astronomical Society, 455, 3136, doi: 10.1093/mnras/stv2458

Naluminsa, E., Elson, E. C., & Jarrett, T. H. 2021, Monthly Notices of the Royal Astronomical Society, 502, 5711, doi: 10.1093/mnras/stab067

Nilson, P. 1973, Nova Acta Regiae Soc. Sci. Upsaliensis Ser. V, 0

Noordermeer, E., & van der Hulst, J. M. 2007, Monthly Notices of the Royal Astronomical Society, 376, 1480, doi: 10.1111/j.1365-2966.2007.11532.x

Noordermeer, E., van der Hulst, J. M., Sancisi, R., Swaters, R. A., & van Albada, T. S. 2005, Astronomy & Astrophysics, 442, 137, doi: 10.1051/0004-6361:20053172

Oosterloo, T., Fraternali, F., & Sancisi, R. 2007, AJ, 134, 1019, doi: 10.1086/520332

Oosterloo, T., Verheijen, M., van Cappellen, W., et al. 2009, Apertif - the focal-plane array system for the WSRT, arXiv, doi: 10.48550/ARXIV.0912.0093

Rhee, J., Zwaan, M. A., Briggs, F. H., et al. 2013, Monthly Notices of the Royal Astronomical Society, 435, 2693, doi: 10.1093/mnras/stt1481

Rhode, K. L., Salzer, J. J., Westpfahl, D. J., & Radice, L. A. 1999, The Astronomical Journal, 118, 323–336, doi: 10.1086/300926

Rich, J. W., de Blok, W. J. G., Cornwell, T. J., et al. 2008, The Astronomical Journal, 136, 2897, doi: 10.1088/0004-6256/136/6/2897

Sancisi, R., Fraternali, F., Oosterloo, T., & van der Hulst, T. 2008, A&A Rev., 15, 189, doi: 10.1007/s00159-008-0010-0

Sanders, R. H. 1996, ApJ, 473, 117, doi: 10.1086/178131

Schwarz, U. 1978, Astronomy and Astrophysics, 65, 345

Serra, P., de Blok, W. J. G., Bryan, G. L., et al. 2016, in MeerKAT Science: On the Pathway to the SKA, 8, doi: 10.22323/1.277.0008

Sersic, J. L. 1968, Atlas de Galaxias Australes

Sinigaglia, F., Elson, E., Rodighiero, G., & Vaccari, M. 2022, Monthly Notices of the Royal Astronomical Society, 514, 4205, doi: 10.1093/mnras/stac1584

Stern, J., Fielding, D., Faucher-Giguère, C.-A., & Quataert, E. 2019, Monthly Notices of the Royal Astronomical Society, 488, 2549, doi: 10.1093/mnras/stz1859

Swaters, R. A., & Balcells, M. 2002, Astronomy & Astrophysics, 390, 863, doi: 10.1051/0004-6361:20020449

Swaters, R. A., Sancisi, R., van der Hulst, J. M., & van Albada, T. S. 2012, Monthly Notices of the Royal Astronomical Society, 425, 2299, doi: 10.1111/j.1365-2966.2012.21599.x

Taylor, P., & Kobayashi, C. 2015, Monthly Notices of the Royal Astronomical Society, 448, 1835, doi: 10.1093/mnras/stv139

Teague, R. 2019, Research Notes of the AAS, 3, 74, doi: 10.3847/2515-5172/ab2125

Tempel, E., Tamm, A., Gramann, M., et al. 2014, Astronomy & Astrophysics, 566, A1, doi: 10.1051/0004-6361/201423585

Tully, R. B., & Fisher, J. R. 1977, A&A, 54, 661

van Albada, T. S., Bahcall, J. N., Begeman, K., & Sancisi, R. 1985, ApJ, 295, 305, doi: 10.1086/163375

van de Hulst, H. C., Raimond, E., & van Woerden, H. 1957, Bull. Astron. Inst. Netherlands, 14, 1

van der Hulst, J. M., & Sancisi, R. 2005, in Astronomical Society of the Pacific Conference Series, Vol. 331, Extra-Planar Gas, ed. R. Braun, 139, doi: 10.48550/arXiv.astro-ph/0411055

van der Hulst, J. M., van Albada, T. S., & Sancisi, R. 2001, in Astronomical Society of the Pacific Conference Series, Vol. 240, Gas and Galaxy Evolution, ed. J. E. Hibbard, M. Rupen, & J. H. van Gorkom, 451

van der Kruit, P. C., & Freeman, K. C. 2011, ARA&A, 49, 301, doi: 10.1146/annurev-astro-083109-153241

Verheijen, M. A. W., & Sancisi, R. 2001, Astronomy &amp Astrophysics, 370, 765, doi: 10.1051/0004-6361:20010090

Verheijen, M. A. W., & Sancisi, R. 2001, A&A, 370, 765, doi: 10.1051/0004-6361:20010090

Volders, L. M. J. S. 1959, Bull. Astron. Inst. Netherlands, 14, 323

Walter, F., Brinks, E., de Blok, W. J. G., et al. 2008a, AJ, 136, 2563, doi: 10.1088/0004-6256/136/6/2563

—. 2008b, AJ, 136, 2563, doi: 10.1088/0004-6256/136/6/2563

Wang, J., Koribalski, B. S., Serra, P., et al. 2016, Monthly Notices of the Royal Astronomical Society, 460, 2143, doi: 10.1093/mnras/stw1099

Wang, J., Fu, J., Aumer, M., et al. 2014, Monthly Notices of the Royal Astronomical Society, 441, 2159–2172, doi: 10.1093/mnras/stu649

Werner, M. W., Roellig, T. L., Low, F. J., et al. 2004, ApJS, 154, 1, doi: 10.1086/422992

York, D. G., Adelman, J., Anderson, John E., J., et al. 2000, AJ, 120, 1579, doi: 10.1086/301513

Zhang, S., Cai, Z., Xu, D., et al. 2023, Science, 380, 494, doi: 10.1126/science.abj9192