

Università degli Studi di Milano Bicocca

Dipartimento di Fisica

Corso di Laurea Triennale



A Catalogue of Starless Dark Matter Halo Candidates  
Detected with FAST

**Relatore:**

**Alejandro Benitez Llambay**

**Candidato:**

**Matteo Raschi**

**ANNO ACCADEMICO 2024/2025**

*This page is intentionally left blank.*

# Abstract

This thesis explores the possibility that there are invisible structures in the universe, made of dark matter, that contain gas but no stars. These objects are called Reionization-Limited HI Clouds (RELHICs). According to current theories, they may have formed in the early universe but never created stars because the gas inside them stayed too hot or too sparse. To search for these hidden structures in our universe, I used data from a large radio telescope survey called "FAST All Sky HI survey" (FASHI), which looks for hydrogen gas in emission. I compared these radio signals with a map of visible galaxies from the Sloan Digital Sky Survey (SDSS) project. When a radio source was found that didn't have a matching optical object, it became a potential candidate for one of these dark halos. After applying a set of selection criteria, namely, a narrow HI line width ( $W_{50} \leq 20\text{km/s}$ , including uncertainties), low neutral hydrogen mass ( $\log_{10}(M_{\text{HI}}/M_{\odot}) < 9$ ), nearly circular morphology ( $b/a > 0.8$ ), significant isolation (distance to nearest galaxy  $> 500\text{kpc}$ ), and moderate integrated HI flux ( $S_{\text{bf}} \leq 3000\text{mJy km/s}$ , within errors), I studied the filtered sample using statistical tools and graphical analysis. . The results show that many of these sources have properties that match what we expect for RELHICs. While more detailed observations are needed to elucidate the nature of this objects, this work helps take the first steps toward discovering and cataloging these hidden systems of the universe.

# **Dedication**

To my family.

# **Declaration**

I declare that this thesis has been composed solely by myself and that it has not been submitted, in whole or in part, in any previous application for a degree. Except where states otherwise by reference or acknowledgment, the work presented is entirely my own.

# Acknowledgements

I would like to express my sincere gratitude and appreciation to my thesis supervisor, Prof. Benitez, for his guidance, support and insightful feedback throughout the entire research process.

I am sincerely thankful to my parents, for their unwavering support, encouragement, and for never asking too often, "When will you finally graduate?"

A special thank you to my girlfriend, Bubu, for being the best possible companion throughout this journey, for her infinite patience, and for believing in me from the very beginning even when my first Calculus exam didn't go as planned. I look forward to sharing many more joyful moments together in the future.

A special thanks to my younger brother, whose presence reminded me that no matter how hard university gets, at least I don't have to do high school all over again. Thanks for the laughs and the lightness.

I would also like to thank my friends Stefano, Silvia, and Giovanni, who, even though they are not fellow physics students, have always been there for me throughout my studies. Their support and distractions made this journey much more bearable and sometimes even fun.

And finally, I would like to thank myself: for not giving up, for surviving all nighters, and for convincing myself that quantum mechanics was a good idea in the first place.

Grazie a tutti voi, di cuore!

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	The Dark Matter Problem and HI Surveys . . . . .	1
1.2	A Case Study: Cloud-9 . . . . .	1
1.3	Aim of the Thesis . . . . .	2
1.4	Thesis Structure . . . . .	2
<b>2</b>	<b>Cosmological Context</b>	<b>3</b>
2.1	$\Lambda$ CDM Model Fundamentals . . . . .	3
2.1.1	Cosmological Simulations . . . . .	3
2.1.2	Small Scale Challenges . . . . .	4
2.2	Structure Formation and Dark Matter Halos . . . . .	5
2.2.1	Formation and Growth of Dark Matter Halos . . . . .	5
2.2.2	Internal Structure of Dark Matter Halos . . . . .	5
2.2.3	Diffuse Gas in Dark Matter Halos . . . . .	6
2.2.4	Halo Substructure and Subhalos . . . . .	6
2.3	Galaxy Formation in $\Lambda$ CDM . . . . .	7
2.3.1	Gas Cooling and Reionization Effects . . . . .	8
2.3.2	Mergers and Feedback Processes . . . . .	8
2.3.3	Simulations and Observational Insights . . . . .	8
2.4	Reionization-Limited HI Clouds (RELHICs) . . . . .	9
2.4.1	Environmental Regulation of RELHICs . . . . .	9
2.4.2	Structure and Detectability . . . . .	9
2.4.3	Sensitivity to Cosmic Conditions and Detection Prospects . . . . .	10
2.4.4	Cosmological Relevance and Astrophysical Implications . . . . .	10
2.4.5	Future Prospects and Theoretical Significance . . . . .	11

<b>3 Data</b>	<b>12</b>
3.1 Instrument FAST: Five-hundred-meter Aperture Spherical Telescope . . . . .	12
3.1.1 FAST's Role in HI Surveys . . . . .	12
3.2 FASHI survey . . . . .	13
3.3 Sloan Digital Sky Survey (SDSS) . . . . .	14
<b>4 Methodology</b>	<b>16</b>
4.1 Sample selection . . . . .	16
4.1.1 FASHI . . . . .	16
4.1.2 SDSS . . . . .	16
4.2 Cross-matching FASHI and SDSS Sources . . . . .	17
4.2.1 Data Cleaning and Validation of Celestial Coordinates . . . . .	17
4.2.2 Initial Attempt: Brute-Force Cross-Matching and Its Limitations .	18
4.2.3 Optimizing Cross-Matching with KDTree . . . . .	19
4.2.4 Visual Inspection and KDTree-Based Cross-Matching Implementation . . . . .	19
4.2.5 Selection of Potential RELHIC Candidates . . . . .	22
<b>5 Results</b>	<b>25</b>
5.0.1 Statistical Characterization of RELHIC Candidates . . . . .	25
5.1 Statistical Analysis and Data Visualization . . . . .	26
5.1.1 Histograms . . . . .	26
5.1.2 Corner Plot . . . . .	32
5.1.3 Scatter plots . . . . .	35
<b>6 Conclusions</b>	<b>45</b>
6.0.1 Interpretation of Results and Cosmological Implications . . . . .	45
6.0.2 Future Prospects . . . . .	46
<b>A A Catalogue of Starless Dark Matter Halo Candidates Detected with FAST</b>	<b>47</b>

# List of Figures

2.1	Visual summary of recent cosmological simulations for structure and galaxy formation. . . . .	4
2.2	Halo occupation fraction as a function of halo mass. . . . .	7
3.1	Sky distribution of FASHI HI sources and comparison with ALFALFA and HIPASS. . . . .	13
4.1	SDSS sources (blue) and FASHI sources (red). SDSS sources were chosen to match the FASHI footprint . . . . .	18
4.2	Aitoff projection to visualize the spatial distribution of detected sources across the celestial sphere. . . . .	22
5.1	Distribution of HI Line Widths ( $W_{50}$ ). . . . .	26
5.2	Distribution of HI Line Widths ( $W_{50}$ ) and a bi-gaussian fit. . . . .	27
5.3	Distribution of HI Masses. . . . .	28
5.4	Distribution of heliocentric velocities ( $cz$ ). . . . .	29
5.5	Distribution of circularity ( $b/a$ ). . . . .	30
5.6	Distribution of a ( $ell_{maj}$ ) and b ( $ell_{min}$ ). . . . .	32
5.7	Corner plot of key parameters. . . . .	34
5.8	Relationships between $W_{50}$ and $M_{HI}$ . . . . .	35
5.9	Relationships between $W_{50}$ and $M_{HI}$ highlighted. . . . .	36
5.10	Relationships between $W_{50}$ and peak flux density $F_{peak}$ . . . . .	37
5.11	Relationships between $W_{50}$ and peak flux density $F_{peak}$ highlighted. . . . .	38
5.12	Relationships between $W_{50}$ and peak flux density $F_{peak}$ color-coded by $\log_{10}(M_{HI})$ . . . . .	39
5.13	Relationships between $W_{50}$ and peak flux density $F_{peak}$ color-coded by $S_{bf}$ . . . . .	41
5.14	Scatter plot of $W_{50}$ versus circularity ratio ( $b/a$ ). . . . .	42
5.15	Scatter plot of $W_{50}$ versus circularity ratio ( $b/a$ ) highlighted. . . . .	43



# 1. Introduction

## 1.1 The Dark Matter Problem and HI Surveys

One of the most fascinating and challenging problems that continues to grapple physicists in recent years concerns dark matter. Although it constitutes about 85% of the total matter in the universe, its nature remains unknown and its direct detection continues to be elusive. The study of the distribution of neutral hydrogen (HI) is a promising new method to detect dark matter halos. Its characteristic predicted emission at 21 cm allows us to explore the properties of gas trapped in these dark matter halos. In this context, the FAST All Sky HI Survey (FASHI), conducted with the five hundred meter aperture spherical telescope (FAST), represents one of the most advanced observational campaigns aimed at identifying and cataloguing extragalactic sources that emit HI. Due to its remarkable sensitivity, FASHI can detect faint, gas-rich objects with or without optical counterparts, typically missed by other astronomical surveys. This makes FAST an ideal instrument for indirectly searching for dark matter halos. By combining the FASHI catalog with data from the Sloan Digital Sky Survey (SDSS), it is possible to identify HI sources without clear optical counterparts. These sources can be classified as "Reionization-Limited HI Clouds" (RELHICs), dark matter halos devoid of stars containing gas in hydrostatic equilibrium within the halo's gravitational potential and in thermal equilibrium with the cosmic ultraviolet radiation background. Such objects are particularly relevant as they allow investigation of the small-scale distribution of dark matter and verification of predictions from the  $\Lambda$ CDM cosmological model.

## 1.2 A Case Study: Cloud-9

A recent example of a potential RELHIC candidate is "Cloud-9", an HI source detected near the galaxy M94, characterized by narrow-band HI emission and no detectable stellar counterpart in optical observations from the DESI Legacy Imaging Survey. These charac-

teristics make Cloud-9 a promising candidate for a starless dark matter halo, potentially representing the first observed example of a RELHIC. This is especially relevant in light of the recent discovery of others RELHICs candidate with FAST and the importance of these systems in testing the predictions of the  $\Lambda$ CDM model.

## 1.3 Aim of the Thesis

This thesis aims to produce a catalogue dark matter halo candidates through the study of FASHI data, in conjunction with the SDSS catalog, to deepen our understanding of the distribution and properties of dark halos in the local universe and provide a direct comparison with predictions from the  $\Lambda$ CDM cosmological model.

## 1.4 Thesis Structure

This thesis is organized as follows. We begin in Section 2 with a summary of the theoretical background, followed in Section 3 by a presentation of our data sets. In Section 4, we present the methodology with which the data were selected; we identify a possible population of starless dark matter halos, and in Section 5 we discuss their properties and our results. We conclude with a summary of our main conclusions in section 6.

## 2. Cosmological Context

### 2.1 $\Lambda$ CDM Model Fundamentals

Our modern theory of cosmology is based on an "educated guess" that the Universe is homogeneous and isotropic on large scales, and Einstein's theory of General Relativity (GR) that says that the structure of space-time is determined by the mass and energy content of the Universe. Observations have shown that the Universe started from a much denser, hotter, and nearly homogeneous state and has been expanding for approximately the past thirteen and a half billion years. The current standard model of cosmology is the Lambda Cold Dark Matter ( $\Lambda$ CDM) model. It describes a Universe that is flat in geometry and dominated by two mysterious components: cold dark matter (CDM) and dark energy, represented by the cosmological constant,  $\Lambda$ . The term "cold" refers to dark matter (DM) particles that move slowly (non-relativistically) and decoupled from radiation very early in the Universe's history. This feature allows DM to form structures hierarchically, from smaller to larger scales, over cosmic time [1] [2].

#### 2.1.1 Cosmological Simulations

Cosmological simulations based on the  $\Lambda$ CDM model have successfully matched many observed features of the Universe. These simulations model the growth of large-scale structures from small initial fluctuations [1] [2]. A summary of recent simulations, both dark matter-only and hydrodynamical, is shown in Figure 2.1. According to this model, small dark matter halos form prematurely and progressively coalesce to build larger and larger structures. However, many of these smaller halos remain dark, unable to retain baryonic matter, mostly due to cosmic reionization [3] [2].

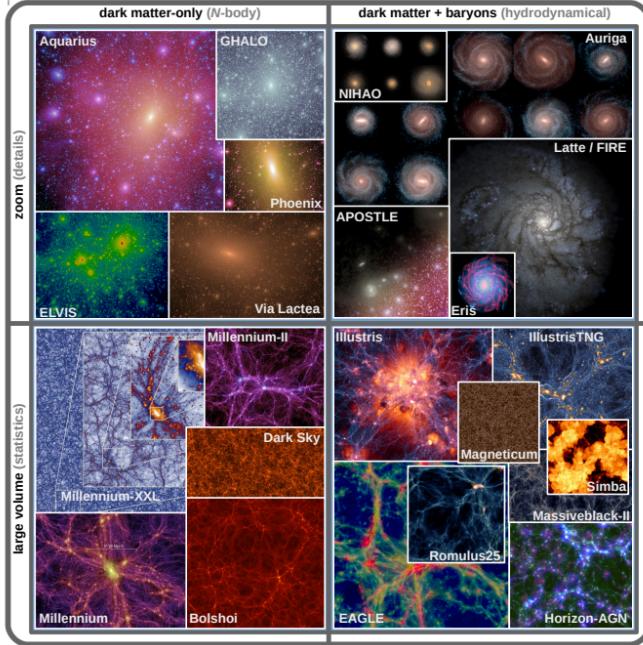


Figure 2.1: Visual representations of some selected recent structure and galaxy formation simulations. The simulations are divided in large volume simulations providing statistical samples of galaxies, and zoom-in simulations resolving smaller scales in more detail. Furthermore, they are also divided in dark matter-only, i.e. N-body, and dark matter plus baryons. Dark matter-only simulations have now converged on a wide range of predictions for the large-scale clustering of dark matter and the dark matter distribution within gravitationally bound dark matter halos. Recent hydrodynamical simulations reproduce galaxy populations that agree remarkably well with observational data. However, many detailed predictions of these simulations are still sensitive to the underlying implementation of baryonic physics. Picture and caption adapted from Paul Torrey Mark Vogelsberger Federico Marinacci and Ewald Puchwein “Cosmological Simulations of Galaxy Formation” (2019).

### 2.1.2 Small Scale Challenges

A very important prediction of  $\Lambda$ CDM is that dark matter halos have a universal density profile, commonly described by the Navarro-Frenk-White (NFW) profile, as discussed in detail by Benitez-Llambay and Frenk in their article “*The detailed structure and the onset*

*of galaxy formation in low-mass gaseous dark matter haloes”* (arXiv, 2020) . This profile effectively characterizes halos based on their mass and concentration, explaining various observations such as galaxy dynamics and gravitational lensing [4] [3]. Regardless of the success of the  $\Lambda$ CDM model, several unsolved problems remain on smaller scales, one example being the contradictions between the predicted and observed properties of DM halos. These discrepancies suggest that additional physics may be needed, either related to baryonic processes or to new properties of dark matter itself [5] [2]. As these issues are prevalent on small scales, further investigations into this regime are likely to provide valuable insights into the fundamental properties and nature of dark matter.

## 2.2 Structure Formation and Dark Matter Halos

### 2.2.1 Formation and Growth of Dark Matter Halos

Dark matter halos are basic elements of  $\Lambda$ CDM and are the main places where galaxies form. At first, they sprang from the density fluctuations in the Universe when it was still in the primordial stage, but later they succumbed to their own gravity, thus becoming self-bound and nearly spherical systems in which galaxies are born [5]. Through N-body simulations, researchers have investigated their characteristics and evolution. The main principle of growth in the halos is the merging of the smallest structures until they become continuously larger. These processes are driven by both gravitational interactions and baryonic effects mainly by radiative cooling and supernova-driven outflows, which change the inner distribution of matter within halos. Also, the interconnection of dark matter halos is engaged in the formation of the large-scale cosmic structure, which leads to the assembly of galaxy clusters and filaments that hold the cosmic web in place.

### 2.2.2 Internal Structure of Dark Matter Halos

The internal structure of dark matter halos is well described by universal density profiles, particularly the Navarro-Frenk-White (NFW) profile, which predicts a density distribution diverging towards the center as  $\rho \propto r^{-1}$ , steepening further out [5]. This profile emerges

naturally from N-body simulations and reflects the purely gravitational nature of dark matter interactions. Unlike galaxy-hosting halos, which are significantly affected by complex baryonic processes such as star formation and feedback mechanisms, starless dark matter haloes (RELBICs) remain simpler and retain this universal profile. Consequently, RELHICs offer a pristine environment for testing predictions of dark matter structure without the complications introduced by galaxy formation processes

### 2.2.3 Diffuse Gas in Dark Matter Halos

In addition to playing a fundamental role in the formation of galaxies, dark matter halos are also important reservoirs of diffuse gas, which can be detected in X-ray observations of galaxy clusters. These gas reservoirs provide valuable information on the thermodynamic state of halos and their interaction with the surrounding intergalactic medium.

### 2.2.4 Halo Substructure and Subhalos

On a small scale, recent studies have focused on the role of halo substructure, particularly in the form of subhalos, which can either host satellite galaxies or remain dark due to inefficient star formation. The abundance and spatial distribution of subhalos have profound implications for our understanding of the properties of dark matter and the nature of structure formation in the Universe. Figure 2.2 shows the predicted halo occupation fraction (HOF) as a function of halo mass at  $z = 0$ , under different assumptions about galaxy formation before reionization. The plot highlights how only halos above a certain mass threshold are likely to host a luminous galaxy, implying that low-mass halos may remain dark and undetectable in optical surveys.

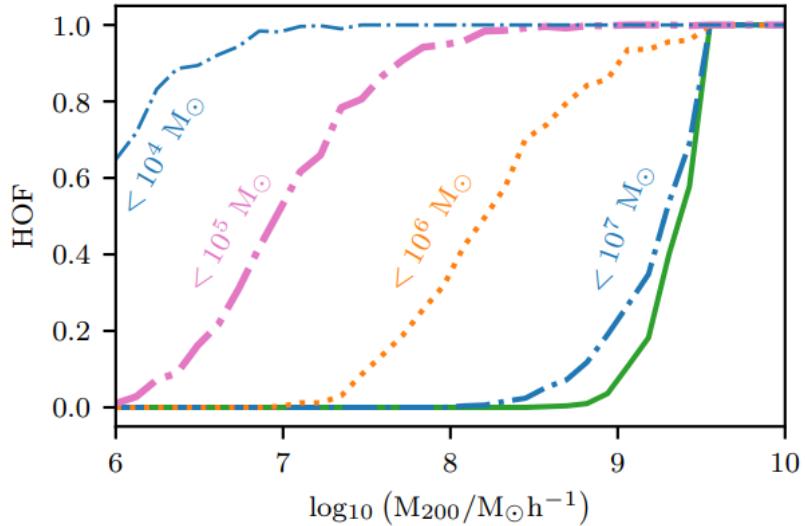


Figure 2.2: Halo Occupation Fraction (HOF) at the present day as a function of halo mass predicted by their model [1], assuming a redshift of reionization,  $z_{re} = 11.5$ . The green solid line shows the limiting case of assuming that all dark matter haloes reach  $z_{re}$  without hosting a luminous galaxy; the other curves assume that all haloes that exceeded the quoted mass prior to reionization host a galaxy by  $z_{re}$ . The  $z = 0$  HOF depends strongly on assumptions about galaxy formation prior to reionization, but only if galaxy formation proceeds in haloes well below the atomic hydrogen cooling limit at  $z_{re}$ . For  $z_{re} \sim 11.5$ , this corresponds to  $M_{200} \sim 10^7 h^{-1} M_\odot$ . Picture and caption taken from Alejandro Benitez-Llambay and Carlos Frenk "The detailed structure and the onset of galaxy formation in low-mass gaseous dark matter haloes" (2020).

## 2.3 Galaxy Formation in $\Lambda$ CDM

Galaxy formation is a fundamental process within the cosmological framework, describing how galaxies originate and evolve over cosmic time. Galaxies form within dark matter halos, which are gravitationally bound regions dominated by dark matter [5]. According to the  $\Lambda$ -Cold Dark Matter ( $\Lambda$ CDM) cosmological model, dark matter halos form hierarchically, with smaller structures merging progressively to create larger ones, providing the gravitational wells necessary for galaxy formation [6].

### **2.3.1 Gas Cooling and Reionization Effects**

Initially, the Universe was filled with nearly uniform gas composed primarily of hydrogen and helium. Over time, this gas began to collapse gravitationally into the potential wells of dark matter halos. When gas falls into a dark matter halo, it heats up and can eventually cool and condense, initiating the formation of stars and thus leading to the emergence of galaxies [1]. The critical mass scale for galaxy formation is closely linked to the ability of gas to cool efficiently within a halo, which occurs primarily through radiative cooling mechanisms. Before cosmic reionization, galaxy formation was possible only in halos massive enough for atomic hydrogen cooling to be effective. After reionization, the situation changed significantly due to the increased heating of the gas, requiring halos to have deeper gravitational wells to retain gas and continue forming stars [1]. The assembly history of galaxies depends strongly on their host halos' growth through smooth accretion and mergers with other halos [5].

### **2.3.2 Mergers and Feedback Processes**

Major mergers, events in which two halos of similar mass collide and merge, can dramatically alter galaxy structure, potentially transforming disk-like galaxies into elliptical ones. Conversely, minor mergers, involving a large galaxy absorbing a significantly smaller one, contribute to the gradual buildup of galaxy mass without drastically changing their morphology [2]. Galaxy evolution is also influenced by various feedback mechanisms, processes that regulate star formation and gas dynamics within galaxies. These include stellar feedback, where massive stars explode as supernovae, injecting energy and momentum into the surrounding interstellar medium, and feedback from active galactic nuclei (AGN), where supermassive black holes at galaxy centers release enormous amounts of energy, significantly affecting galaxy growth and evolution [6].

### **2.3.3 Simulations and Observational Insights**

Observational and computational advancements, including large-scale cosmological simulations, have greatly improved our understanding of galaxy formation [2]. These simula-

tions model the complex interplay of gravity, gas dynamics, star formation, and feedback processes, accurately reproducing many observed galaxy properties such as size, structure, mass distribution, and chemical composition. Overall, galaxy formation is a complex, multiscale, and multifaceted phenomenon governed by gravitational and baryonic physics. Studying this process provides crucial insights into the history and evolution of the Universe and helps refine our cosmological models [6].

## 2.4 Reionization-Limited HI Clouds (RELHICs)

Of particular interest are low-mass dark matter halos, specifically Reionization Limited HI Clouds (RELHICs). RELHICs are dark matter-dominated objects that have experienced significant effects from cosmic reionization, a process that regulates their gas content and stifles star formation. They have an equilibrium between gravitational confinement and external UV background heating from the cosmic ultraviolet background, which puts their gas in hydrostatic and thermal equilibrium [3]. In comparison to normal galaxies, these halos have no stellar populations but have neutral hydrogen (HI) [4].

### 2.4.1 Environmental Regulation of RELHICs

RELHICs exist under the regulations of two important processes. On one hand, reionization, ionized a significant fraction of the intergalactic medium, inhibiting gas and subsequent star formation. Further depletion of their gas content by interaction with the intergalactic medium, specifically via ram pressure stripping, is also possible in denser environments [3]. Consequently, RELHICs are largely located in field areas, where they experience less disturbance.

### 2.4.2 Structure and Detectability

Hydrodynamical simulations predict that RELHICs possess gas distributions that are concordant with Navarro-Frenk-White (NFW) dark matter halos, which follow a universal density profile characterized by a central cusp ( $\rho \propto 1/r$ ) and a steep outer fall-off ( $\rho \propto$

$1/r^3$ ). Their neutral hydrogen content is small and centrally concentrated. They are expected to have small velocity dispersions because their gas is, to a large extent, in thermal equilibrium at  $T \sim 10^4$  K and devoid of large-scale turbulent motion. The best observational method to find RELHICs is through 21 cm emission line surveys, which can trace their HI content even if they are starless [4].

### 2.4.3 Sensitivity to Cosmic Conditions and Detection Prospects

It has been suggested in recent studies that HI content of RELHICs is very sensitive to cosmic UV background intensity and to the subtle thermal evolution of the intergalactic medium. While some RELHICs are able to preserve substantial HI reservoirs, others can be nearly completely ionized and hence difficult to detect. The real proportion of dark matter halos that are still RELHICs relies on several astrophysical factors, including the timing and extent of cosmic reionization and the local environment of the halos. Ongoing attempts at detecting RELHICs have intensified, especially by a series of broad HI surveys such as the FAST All Sky HI Survey (FASHI), which has begun detecting possible candidates in agreement with theoretical expectations [7]. The observations here provide critical tests for the predictions on the small-scale structure in the context of  $\Lambda$ CDM and give valuable insight into the dynamics of low-mass dark matter halos. In addition, numerical simulations that are currently underway are increasingly refining predictions on the mass function and spatial distribution of RELHICs, which helps in the interpretation of observational data.

### 2.4.4 Cosmological Relevance and Astrophysical Implications

The study of RELHICs is of significant importance in understanding the era of reionization and the underlying physics that govern galaxy formation. As these objects carry a record of the environment that existed during the reionization epoch, they serve as valuable probes for inferring the thermal and ionization history of the Universe and provide hints to the nature of dark matter . Future observational campaigns, especially those specifically targeting isolated HI clouds with no stellar companions, have great potential for advanc-

ing our understanding of dark matter and baryonic interaction in low-mass halos [1]. The characterization of RELHICs will also impose constraints on star formation thresholds and the intricate interactions between baryonic matter and dark matter.

#### **2.4.5 Future Prospects and Theoretical Significance**

Next-generation radio telescopes, along with future deep HI surveys, will make detecting of RELHICs increasingly feasible, returning unprecedented insights into the nature of dark matter. The study of RELHICs not only functions to test crucial elements of the  $\Lambda$ CDM model but forms a key link between cosmic Reionization and subsequent formation of luminous galaxies. Through the merging of empirical observations with improved theoretical tools, future work on RELHICs has the potential to shed light on the intricate interactions between dark matter, baryonic physics, and cosmic evolution.

## 3. Data

### 3.1 Instrument FAST: Five-hundred-meter Aperture Spherical Telescope

The Five-hundred-meter Aperture Spherical Telescope (FAST) is the world's largest and most sensitive single-dish radio telescope, built in a naturally formed dip in Guizhou Province, China. The telescope is a huge five hundred meter spherical disc, with the purpose of collecting and concentrating radio waves coming from the Universe upon its receiver instruments. While the 500-meter diameter is spanned by the reflector, only a part (around 300 meters) is actually employed at a time to point at different positions in the sky. FAST has a highly selective 19-beam receiver array that boosts its observation efficiency by allowing them to make simultaneous observations in different directions in the sky. The telescope works in a frequency range of roughly 1050 to 1450 MHz, optimized to detect neutral hydrogen (HI) by its distinctive 21-centimeter wavelength emission line. FAST's sensitivity of around 0.76 mJy in each of its beams at a spectral resolution of some 6.4 km/s renders it extremely adept at detecting faint astrophysical phenomena that are imperceptible by smaller radio telescopes. This unprecedented sensitivity and survey rate have allowed FAST to take on a grand program like the FAST All Sky HI Survey (FASHI).

#### 3.1.1 FAST's Role in HI Surveys

FASHI is sampling a vast area of the sky accessible by FAST, roughly 22000 square degrees, to survey tens of thousands of HI sources like galaxies and clouds of gas. In comparison with earlier surveys such as the Arecibo Legacy Fast ALFA (ALFALFA) survey, FAST has much better sensitivity, spatial, and spectral resolution. As a result, FAST has played a very significant role in the detection of the low-luminosity and isolated astrophysical objects like dim dwarf galaxies and dark matter halo potential gas clouds. The sky coverage and source density achieved by FASHI, as illustrated in Figure 3.1, highlight the

contrast with previous surveys like ALFALFA and HIPASS, both in depth and sky footprint. Such abilities make FAST a versatile instrument of modern astrophysical research that contributes extensively to understanding the cosmic evolution and basic architecture of the Universe.

## 3.2 FASHI survey

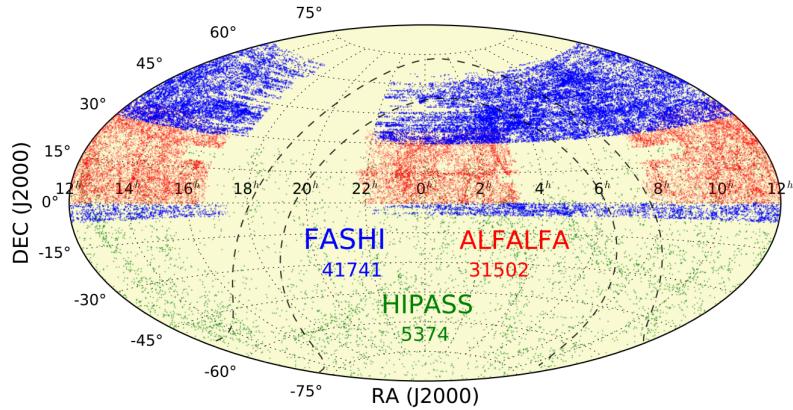


Figure 3.1: FASHI sky distribution of currently released 41741 HI sources (in blue points) included in the Galactic hemispheres, showing the roughness of boundaries imposed by practical and scheduling constraints. For comparison, ALFALFA (Haynes et al., 2018) and HIPASS galaxies (Koribalski et al., 2004; Meyer et al., 2004; Wong et al., 2006) are also indicated with red and green points, respectively. Image and caption taken from Chuan-Peng Zhang and FAST Collaboration "The FAST All Sky H I Survey (FASHI): the first release of catalog" (2023)

The FAST All Sky HI Survey (FASHI) is a wide-area astronomy survey initiative aimed at surveying neutral hydrogen (HI) over the full sky accessible to the FAST telescope, covering declinations between  $-14^{\circ}$  and  $+66^{\circ}$ . Operating within a range of 1050 to 1450 MHz, FASHI will discover in excess of 100, 000 extragalactic HI objects and deliver the largest HI catalog to date upon completion. Over the August 2020 to June 2023 time frame, FASHI successfully imaged some 7, 600 square degrees, approximately 35% of its planned observation area. In this time, the survey detected 41, 741 extragalactic HI

objects. The survey sensitivity is estimated at a median of 0.76 mJy/beam, and a velocity resolution of around 6.4 km/s at frequencies of approximately 1.4 GHz. FASHI's sensitivity and resolution are outstanding compared with earlier wide-area surveys like ALFALFA, and FASHI is especially proficient at detecting weak, low-mass objects that are isolated. FASHI's data product offers essential information on the HI content of the nearby universe, useful to galaxy evolution, star-forming potential, and dark matter halo distribution and characteristics. FASHI data are analyzed with sophisticated techniques of data analysis, such as stringent cross-matching with optical surveys like the Sloan Digital Sky Survey (SDSS) to identify and flag the presence of source-optical counterparts, enabling studies of dark galaxies and isolated HI clouds. As a result, FASHI greatly enhances insight into both the universe's baryonic and dark matter components and contributes a wide-ranging observation framework to cosmology and astrophysics research.

### 3.3 Sloan Digital Sky Survey (SDSS)

The Sloan Digital Sky Survey (SDSS) is one of the largest and most significant astronomical surveys to be made so far. Started in 2000, SDSS has made systematic use of a special 2.5-meter wide-field optical telescope at the Apache Point Observatory in New Mexico, USA, to map a large area of the sky. Using multi-spectral images and redshift measurements made by spectroscopy, SDSS has gathered detailed information on millions of objects in the sky, such as stars, galaxies, and quasars. Using a five-filter photometric system ( $u$ ,  $g$ ,  $r$ ,  $i$ , and  $z$ ) to take images at various wavelengths, SDSS measures the brightness and color of objects with a very good degree of accuracy. Its spectroscopy capability also allows the measurement of redshifts, giving information about the distance and velocity of objects. Imaging and spectroscopy combined have created a vast dataset that can be used to aid a wealth of different areas of astronomy research. In the case of the FAST All Sky HI Survey (FASHI) research, SDSS is an essential tool to cross-refer and validate detections of HI. Cross-referencing FASHI's radio observations of neutral hydrogen (HI) with the observations made in the optical by SDSS, researchers can have a way to identify the HI source with no optical counterpart. These sources are of interest because they could

be dark galaxies or Reionization-Limited HI Clouds (RELHICs) dark matter halos with gas but without any star. Combining the data of SDSS with FASHI enhances the capability to classify and tell what kind of source of HI was detected. For example, if a source of HI in FASHI coincides with an object in the optical in SDSS, the source can be linked to a known galaxy, helping to study its characteristics. In turn, a lack of an optical counterpart to an FASHI-detected HI source could be a dark galaxy or a RELHIC candidate, and therefore a source of investigation. In addition, the very large catalog and easy data-access interfaces offered by SDSS, like the SkyServer, make for quick cross-matching and analysis. Images, spectra, and object parameters can be retrieved by researchers to accompany and contextualize FASHI radio observations. Such a combination of the radio and optical surveys highlights the value in multi-wavelength strategies in contemporary astrophysics, allowing for a better insight into the universe in both structure and content.

## **4. Methodology**

### **4.1 Sample selection**

#### **4.1.1 FASHI**

The datasets used in this research were sourced from the FAST All Sky HI Survey (FASHI) and the Sloan Digital Sky Survey (SDSS). The FASHI dataset was acquired directly from tables publicly available on the FASHI project's official online platform. These tables include comprehensive information on detected extragalactic HI sources, covering aspects such as positional coordinates, flux densities, and velocity widths.

#### **4.1.2 SDSS**

For the SDSS dataset, data were retrieved using the CASJOB platform, a specialized tool designed for efficient data retrieval from SDSS. On CASJOB, SQL (Structured Query Language) code must be used to perform specific queries and data extraction, the code used is reported in code 4.1. The sample was carefully selected by extracting data points corresponding to galaxies and quasars within specific regions of the sky, ensuring effective and precise cross-matching between the optical observations from SDSS and radio observations from FASHI. This combined dataset forms the foundation for subsequent analyses aimed at identifying potential starless dark matter halo candidates and examining their astrophysical characteristics.

```

1  SELECT objID, class, ra, dec, z, zErr
2  INTO mydb.MyTable
3  FROM SpecPhotoAll
4  WHERE ra BETWEEN 0. AND 360.
5    AND ((dec BETWEEN -8 AND 1) OR (dec BETWEEN 29 AND 67))
6    AND photoRa IS NOT NULL
7    AND photoDec IS NOT NULL
8    AND (class = 'GALAXY' OR class = 'QSO')

```

Code 4.1: SQL query used to download data from SDSS.

## 4.2 Cross-matching FASHI and SDSS Sources

The analysis began with the careful preparation and validation of data sets to ensure accuracy and coherence in preparation for subsequent analysis. Two master catalogs served as the basis for this cross-matching process: the FAST All Sky HI Survey (FASHI) and its observations of extragalactic neutral hydrogen (HI) sources and the Sloan Digital Sky Survey (SDSS) that includes detailed optical data for galaxies and quasars. The programming language Python, a popular language within the scientific community, was used to manage and manipulate the large data sets, relying largely upon the pandas library and its immense data manipulation power.

### 4.2.1 Data Cleaning and Validation of Celestial Coordinates

In the data preparation stage, special care was paid to the format and numerical integrity of essential parameters like right ascension (RA) and declination (DEC) since the values are fundamental to cross-match the data sets. Any entries that lacked, were corrupted, or contained wrong values were identified and either cleaned or excluded to preserve the reliability of the output. To ensure the reliability of the subsequent analysis, we performed checks for duplicated coordinate values (same RA and DEC) and validated the numerical accuracy of the data. We verified that the right ascension and declination values were within

physically meaningful ranges. Specifically, right ascension (RA) values were checked to fall within the range  $[0^\circ, 360^\circ]$ , and declination (DEC) values within  $[-90^\circ, +90^\circ]$ , as required by celestial coordinate conventions. This step helped eliminate inconsistencies that could otherwise compromise the robustness of the final results, as visually confirmed in Figure 4.1.

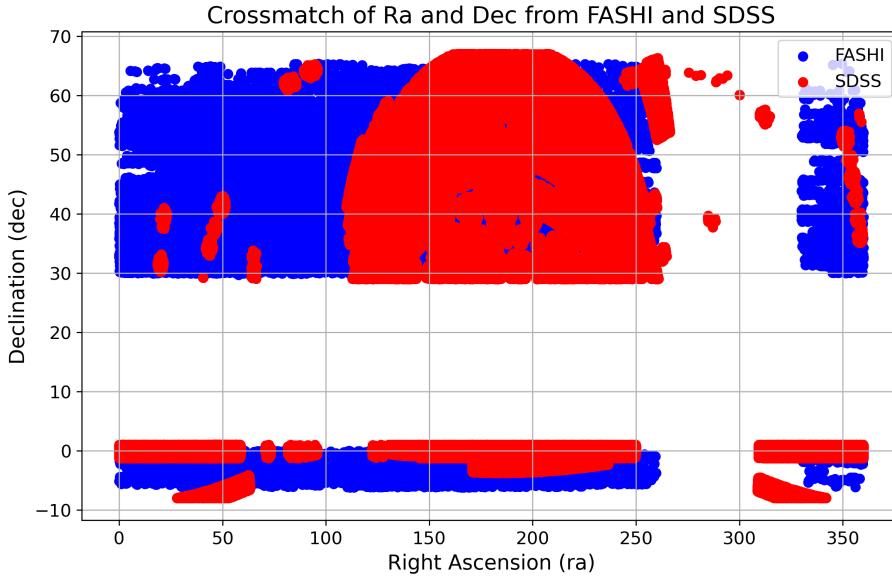


Figure 4.1: SDSS sources (blue) and FASHI sources (red). SDSS sources were chosen to match the FASHI footprint

#### 4.2.2 Initial Attempt: Brute-Force Cross-Matching and Its Limitations

Initially, the cross-matching analysis between the FASHI and SDSS data sets was created with a brute-force, direct comparison approach. This course of action would compare every object within the FASHI catalog to all objects within the SDSS catalog within a desired angular separation. This was, however, quickly discarded since it became immediately obvious that this was a computationally prohibitive task. The data volume within each of the catalogs, and with the additional inclusion of redshift tolerance and error margins within the positions, overwhelmed the processor of a standard personal computer. Execution times were excessively high, and sometimes even aborting entirely due to a

lack of memory and processor power.

### 4.2.3 Optimizing Cross-Matching with KDTree

In order to overcome this issue, a more effective strategy was employed, using a data structure called a KDTree. A KDTree (K-Dimensional Tree) is a spatial index data structure that permits the efficient organization and querying of points in N-dimensional space. Think of a KDTree like a nested hierarchy that partitions space to store and efficiently retrieve point close to a given location. Suppose that you are trying to locate a book within a large library where all the books are randomly stored. Searching each one until finding what you were searching for would take a lot of time. A KDTree, however, stores the books within the category to and within the location where they are situated, allowing for much faster and more effective location. For astronomy, this means finding rapidly objects within a catalog that are near objects within another.

### 4.2.4 Visual Inspection and KDTree-Based Cross-Matching Implementation

Once the FASHI and SDSS data were cleaned and sorted, a preliminary visual examination was performed. Scatter plots were created of RA vs DEC to visually inspect the spatial coverage and possible overlap within the two datasets (shown in figure 4.1). This preliminary analysis step helped identify regions of interest and ensure that the data were aligned as expected. After this qualitative analysis, the cross-matching was performed by applying the cKDTree library of the SciPy library. The SDSS dataset was indexed into cKDTree by its celestial coordinates to ensure that nearby neighbors were rapidly looked-up for each object in FASHI catalog. The cross-matching algorithm was formulated to compare position and redshift. A match was considered to be a duo of objects with an angular separation of 3 arcminutes and a redshift difference within the combined errors of both observations. The choice of a 3 arcminutes threshold for the positional cross-matching is consistent with the methodology adopted in the official FASHI data release, where counterparts in both the Siena Galaxy Atlas (SGA) and the Sloan Digital Sky Sur-

vey (SDSS) were identified using a maximum angular separation of  $3'$  in right ascension and declination [7]. This value accounts for the FAST beam size ( $\sim 2.9'$ ), as well as alignment differences. Angular separation was determined by the haversine formula, a standard technique in astronomy to calculate distances on the celestial sphere, and the errors associated with this were included within the match criteria. Using this methodology led to the compilation of two separate datasets:

- **Matched Sources:** HI objects identified in both FASHI and SDSS catalogs, indicating HI sources with detectable stellar counterparts (2837 sources).
- **Non-Matched Sources:** Objects found exclusively in the FASHI dataset, representing candidate isolated HI clouds potentially lacking stellar populations (39064 sources).

In total, the implementation of the cKDTree algorithm made what would have been an intractable computational challenge a reasonable and efficient process. This not only made the cross-matching step possible, but also made possible subsequent analyses and selection of candidates based upon the matched and unmatched resulting population. The implementation details of the cross-matching routine using KDTree, including angular separation filtering and redshift consistency checks, are illustrated in Code 4.2. The full code and supplementary material supporting this thesis are publicly available on GitHub at: <https://github.com/Matteo-Raschi/relhic-bachelor-thesis>.

```

1  from scipy.spatial import cKDTree
2  import numpy as np
3
4  # Build KDTree from SDSS coordinates in radians
5  sdss_coords = np.radians(sdss_data[['ra', 'dec']].dropna().
6      to_numpy())
7  sdss_tree = cKDTree(sdss_coords)
8
9  # Define max separation in radians
10 max_distance_rad = np.radians(0.05)
11
12 # Iterate over FASHI entries
13 for i, row in fashi_data.iterrows():
14     ra_f, dec_f = row['ra'], row['dec']
15     if pd.notna(ra_f) and pd.notna(dec_f):
16         idx_list = sdss_tree.query_ball_point([np.radians(ra_f),
17                                             np.radians(dec_f)], max_distance_rad)
18         for idx in idx_list:
19             # Additional redshift and error checks
20             ...

```

Code 4.2: Core matching routine using cKDTree to crossmatch FASHI and SDSS catalogs.

#### 4.2.5 Selection of Potential RELHIC Candidates

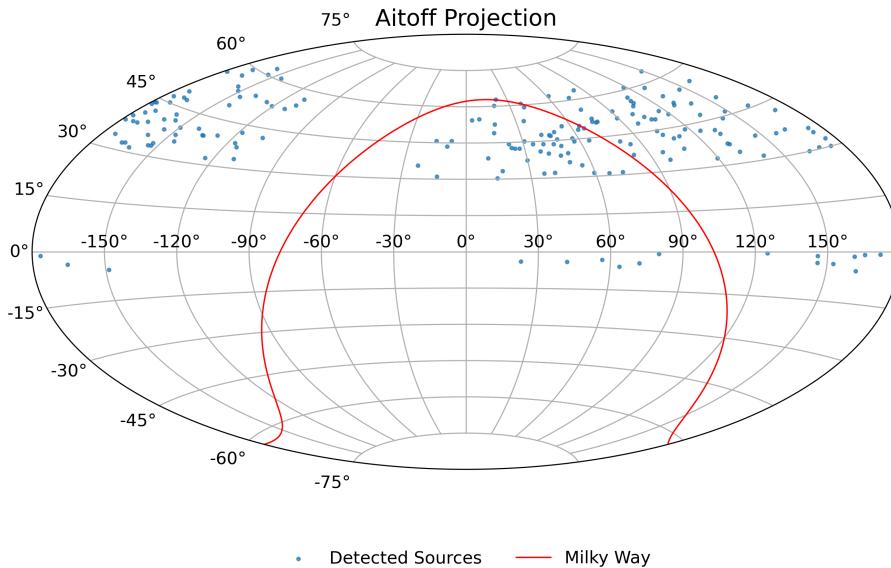


Figure 4.2: Aitoff projection to visualize the spatial distribution of detected sources across the celestial sphere.

After cross-matching FASHI and SDSS sources, a targeted analysis was performed of non-matched sources to search for Reionization-Limited HI Cloud (RELHIC) candidates. RELHICs are predicted starless dark matter halos by the  $\Lambda$ CDM theory. These systems have neutral hydrogen and are in hydrostatic and thermal equilibrium with the cosmic UV background (UVB) radiation. Importantly, they never formed stars, either because their mass was too small to permit the gas to cool efficiently prior to the Universe reionization, or due to photoheating by the UVB radiation of the reionization era that halted gas accretion. The selection of RELHIC candidates was based on strict requirements motivated by expectations from hydrodynamical simulations and theory. Every condition was selected for a specific physical motivation, so that the resulting selected objects corresponded to the expected characteristics of true RELHICs:

1. HI line width ( $W_{50}$ ) equal to or below 20 km/s, considering measurement uncertainties. This narrow line width is a sign of minimal internal motion within the gas cloud. In RELHICs, without the presence of stars and other kinetic feedback drivers,

there is little or no turbulence or rotational velocity. Thermal motions dominate their HI profiles, and a small line width is therefore expected under the predicted conditions of thermal equilibrium in such low-mass halos ( $T \sim 10^4$  K).

2. HI mass  $M_{HI} < 10^9 M_\odot$  – RELHICs are expected to reside in dark matter halos with masses lower than the critical threshold for star formation, typically under  $10^{10} M_\odot$ . A low HI mass ensures consistency with simulations showing that halos below this limit retain small amounts of gas, mostly in ionized form, with a neutral hydrogen component [3]. RELHICs are expected to be found in dark matter halos with mass less than the star-forming critical mass, usually less than  $10^{10} M_\odot$ . A low mass of HI guarantees conformity to simulations that demonstrate that halos of this mass or less hold minimal quantities of gas, predominantly ions, with a neutral hydrogen component, moderate in mass. [3].
3. Circularity ratio  $> 0.8$  – This condition guarantees that objects chosen are approximately spherical in projection. RELHICs have been found to be nearly spherical in shape by simulations. A large circularity cuts off irregular or filaments that are highly unlikely to be in equilibrium. [2].
4. Distance  $> 500$  kpc from known objects – This will guarantee that the source stands by itself, reducing the chances that it is a known galaxy associated with a gas cloud (i.e., tidal debris or a satellite galaxy). RELHICs are predicted to be found at low-density regions where interactions are minimal, like the regions at the edge of galaxy groups.
5. Integrated HI flux  $S_{bf} < 3$  Jy km/s – This flux cut-off preferentially selects weak, low-luminosity objects. Such objects match the anticipated observation signature of RELHICs: radio-detected but radio-faint in surveys, with no counterparts in the optical.

These requirements were not random but are directly inherited from cosmological hydrodynamical simulations, e.g., those in the APOSTLE and EAGLE projects, and investigations like [1][3], which simulated the competition between photoionization heating and

gravitational trapping. RELHICs reside in a confined part of parameter space in which gas exists in a neutral state without giving rise to star formation, a very rare but very fundamental result of structure buildup within the  $\Lambda$ CDM framework. Applying all of these filters returned a cleaned population of RELHIC candidates (188), stored in a specialized dataset awaiting further analysis. Figure 4.2 shows an Aitoff projection of the end sample of candidates, plotting their sky distribution and tracing the discovery of isolated, compact neutral hydrogen sources tentatively representing dark matter halo populations at their faint end. The complete catalogue is listed at the end of this thesis in appendix A and can be downloaded on the public GitHub of this work.

## 5. Results

In this chapter we present the overall results of our analysis, namely a catalogue of potential Reionization-Limited HI Clouds (RELHICs) in the FASHI dataset. The findings are based on a cleaned sample of HI sources with no optical counterpart in the Sloan Digital Sky Survey (SDSS) and are, therefore, strong candidates for isolated, dark matter halos without stars. The organization of this section is to gradually lead the reader to an even deeper understanding of the data.

### 5.0.1 Statistical Characterization of RELHIC Candidates

We start with a statistical summary of some of the key physical parameters, such as HI line width, total mass of HI, and recession velocity. One-dimensional distributions of these parameters give insight into the general characteristics of the population of candidates in our catalogue. Then, the analysis turns to multivariate distributions and relationships between parameters, via scatter plots and pairplots, to examine the relationships between the fundamental parameters. Particular emphasis is placed upon parameters that index the gravitational stability and isolation of the sources, in accordance with the requirements of the standard theory of RELHIC. Introduction and interpretation of each figure, in the context in which it should be viewed, are presented, together with an assessment of interesting trends and discussion of their physical meaning. The concluding part of the chapter brings together the results and relates them to expectations within theory and earlier research and considers to what extent the chosen candidates conform to the RELHIC hypothesis.

## 5.1 Statistical Analysis and Data Visualization

### 5.1.1 Histograms

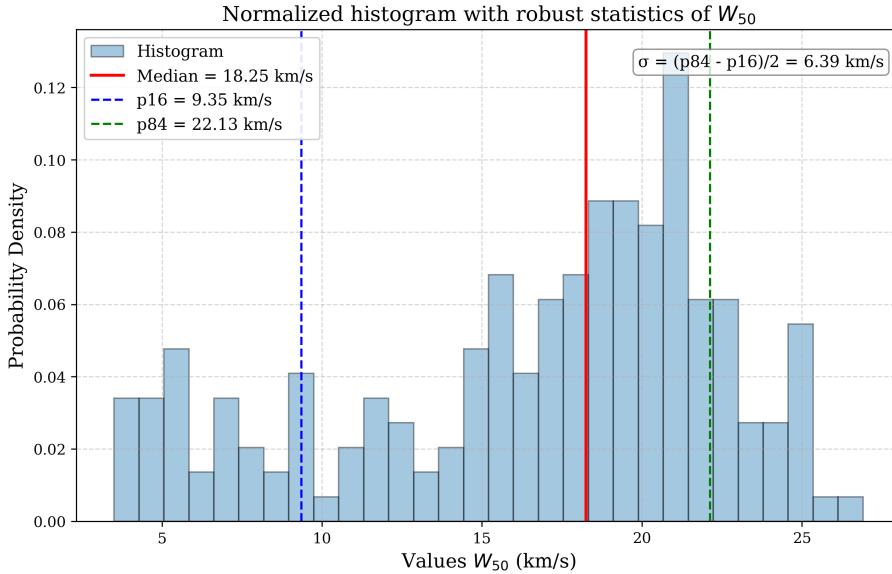


Figure 5.1: Distribution of HI Line Widths ( $W_{50}$ ).

The normalized histogram shown in Figure 5.1 displays the distribution of the HI line width,  $W_{50}$ , of the selected candidate RELHICs. The  $W_{50}$  parameter is the full width at half maximum of the HI line and gives direct insight into the internal kinematics of the detected objects. The histogram is evidently skewed towards lower values, with a dense concentration at 15 to 21 km/s. The statistical measures, a median of 18.25 km/s, a velocity dispersion of 6.39 km/s defined as  $\sigma = (p_{84} - p_{16})/2$ . These values confirm that the majority of objects have narrow line widths, a characteristic feature to be expected for Reionization-Limited HI Clouds. Low values of  $W_{50}$  indicate little internal velocity dispersion, consistent with the theoretical prediction of RELHICs to be gas-rich but dynamically cold halos. In the absence of a star population, feedback processes like stellar winds or supernovae do not operate to disrupt the gas, and the HI is in near hydrostatic equilibrium. The average gas temperature in such systems - heated by the cosmic ultraviolet field - accommodates thermal velocities of this order of magnitude (10-20 km/s), consistent

with the peak of the distribution. In addition, the absence of a wide, high-velocity tail supports the notion that these are not rotation-dominated galaxies, but pressure-supported, spherically symmetric HI clouds in shallow dark matter potentials. In conclusion, this distribution offers compelling early proof that the screened sample agrees with the anticipated kinematic profile of RELHICs by  $\Lambda$ CDM cosmological simulations [3], demonstrating that our selection of potential RELICH candidates is sensible and correctly applied.

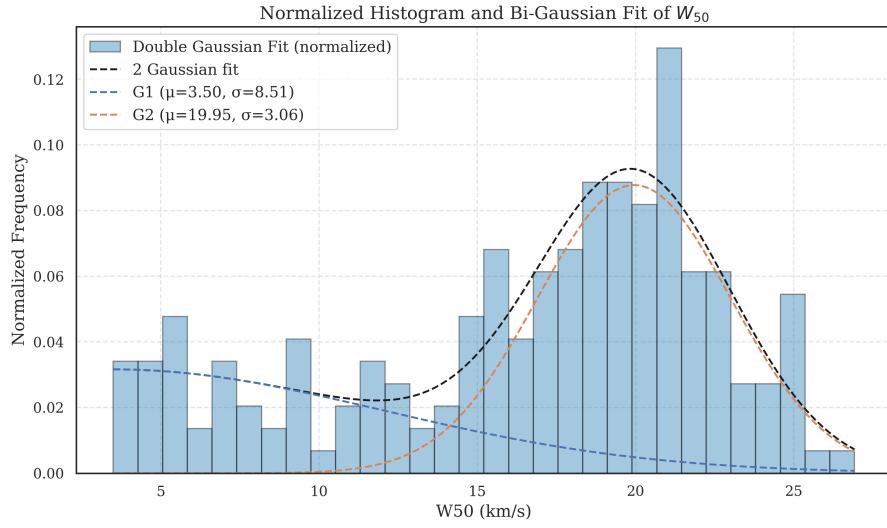


Figure 5.2: Distribution of HI Line Widths ( $W_{50}$ ) and a bi-gaussian fit.

In order to study further the suggestion that the measured distribution might hide the presence of more than a single dynamical population, a bi-Gaussian fit was applied. This fit method represents the histogram by the sum of two components with a Gaussian distribution, enabling one to test whether the asymmetry could be due to the superposition of discrete subgroups within the sample. Figure 5.2 displays the result of this fit, where the overall distribution is broken down to two curves: the first with center at roughly 3.5 km/s and wider dispersion, and the second with center at roughly 20 km/s and smaller dispersion. The second component falls in line with the expectations formulated by the theory of RELHICs, namely, compact, thermally in equilibrium gas clouds with minimal turbulence as measured in hydrodynamical simulations [3]. The model predicts that low-mass dark matter halos at the star formation limit edge will have neutral hydrogen trapped in a quasi-static regime, heated and confined by the cosmological ultraviolet background. The

lower- $W_{50}$  component could be due to observational artifacts at or just above the detection limit or to a population of extremely cold and quiet systems predicted in isolated low-mass halos, especially those below the filtering mass ceiling. Two peaks don't necessarily suggest two different physical causes; instead, the result could be the intrinsic scatter in properties seen even in similar halos, especially with the help of environmental factors and the uncertainties involved in  $W_{50}$  measures of very faint objects. This decomposition validates the premise that the majority of candidates show dynamical characteristics consistent with the RELHIC model, but a proper statistical analysis discovers subtle structure in the data that can be attributed to the halo mass or thermal state diversity offered by  $\Lambda$ CDM-based simulations. Specifically, the distinct separation of the two Gaussian peaks further supports the validity of the initial selection criterion by an upper limit to  $W_{50}$  of 20 km/s.

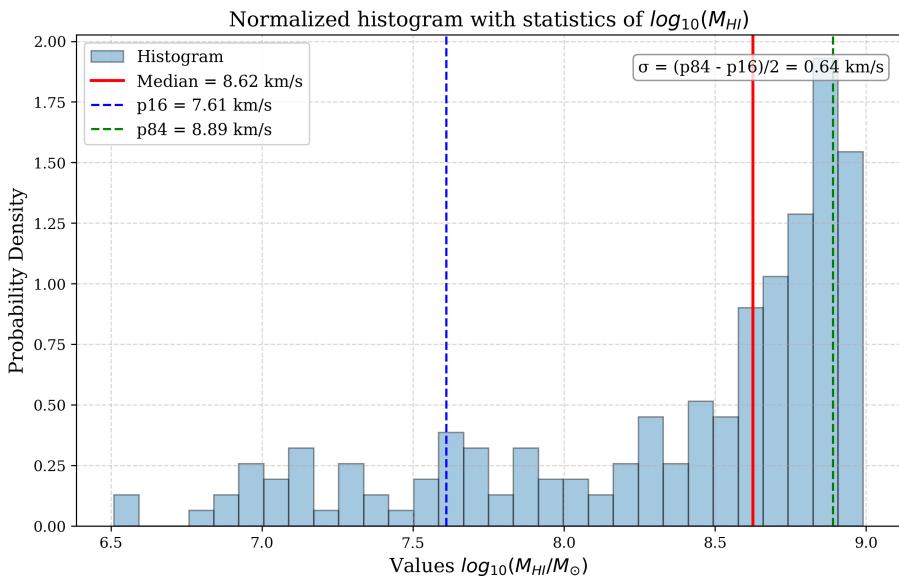


Figure 5.3: Distribution of HI Masses.

Figure 5.3 presents the normalized histogram of the total neutral hydrogen mass for the candidate RELHICs, expressed as  $\log_{10}(M_{HI})$ . The distribution is notably asymmetric, with a clear accumulation of sources toward higher masses and a sharp decline below  $\log_{10}(M_{HI}/M_\odot) \sim 8.5$ . The calculated median value is 8.62 km/s, with a  $\sigma$  of 0.64 km/s. This mass range is broadly consistent with theoretical expectations for low-mass halos just

below the threshold for star formation. In particular, cosmological hydrodynamical simulations [3] predict that halos hosting RELHICs typically retain HI masses between  $10^6 M_\odot$  and  $10^9 M_\odot$ , depending on the strength of the ionizing background, their formation history and their dark matter mass. The presence of candidates with  $\log_{10}(M_{HI}/M_\odot) < 7.5$  is especially interesting, as these could represent extremely faint systems at the edge of detectability, potentially corresponding to the lowest-mass halos capable of retaining any neutral gas. The sharp rise near  $\log_{10}(M_{HI}/M_\odot) \approx 8.8$  likely reflects the intrinsic mass distribution of the selected RELHIC candidates. While selection effects such as the flux sensitivity threshold (e.g.,  $S_{bf} < 3 \text{ Jy km s}^{-1}$ ) may limit the detectability of lower-mass systems, they would not account for a sharp increase at high HI masses. Thus, the observed steep rise is more plausibly associated with the underlying population rather than observational bias. Nevertheless, the absence of candidates with  $\log_{10}(M_{HI}/M_\odot) > 9$  is consistent with the upper mass cutoff imposed by the RELHIC selection criteria, which aim to isolate starless halos rather than normal gas-rich galaxies. This mass distribution reinforces the interpretation that the selected candidates are not part of the typical galaxy population, but instead lie in the transition regime between baryonic failure and star-forming systems a hallmark domain for RELHICs in  $\Lambda$ CDM cosmology.

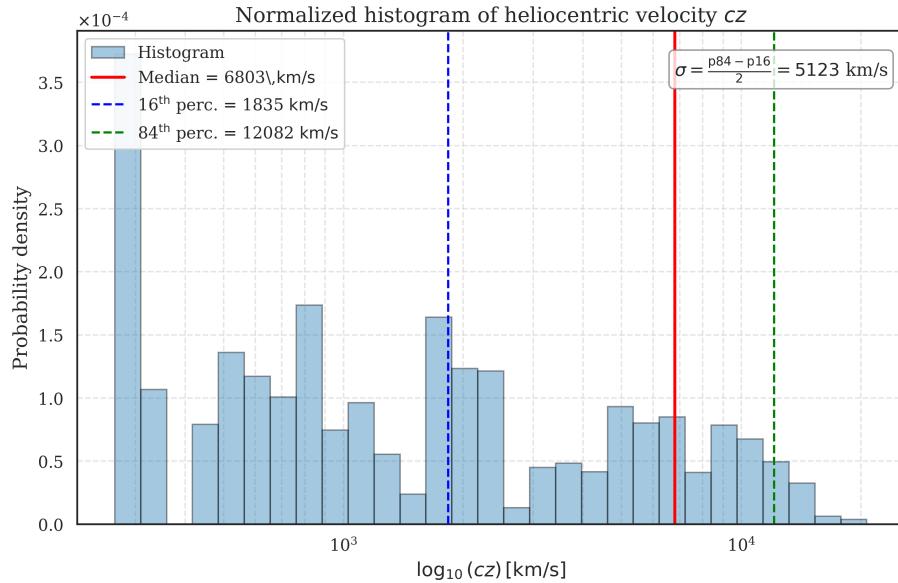


Figure 5.4: Distribution of heliocentric velocities ( $cz$ ).

Figure 5.4 displays the normalized histogram of heliocentric velocities ( $cz$ ) for the candidate RELHICs. The  $cz$  value, expressed in km/s, corresponds to the radial velocity of each source with respect to the Sun and provides a direct proxy for its cosmological distance under the assumption of Hubble flow. The distribution reveals a broad spread of velocities, with a median of 6803.07 km/s and a  $\sigma$  of 5123.43 km/s calculated as  $\sigma = (p_{84} - p_{16})/2$ . A mild peak appears in the first few bins (below 2500 km/s), followed by a relatively flat tail that extends to around 20,000 km/s. The flatness and extension of the distribution suggest that the sample includes candidates spread across a wide range of distances, from a few Mpc to over 250 Mpc. This is consistent with the capabilities of the FASHI survey, whose sensitivity allows the detection of low-HI-flux objects even at intermediate redshifts. However, the sharp decline beyond  $cz \sim 15,000$  km/s may reflect the practical detection limit imposed by both the spectral resolution of the instrument and the lower limit of HI flux FAST can detect. From a theoretical perspective, RELHICs are expected to inhabit underdense environments and be relatively isolated. The distribution shown here supports this hypothesis, as it lacks any pronounced clustering at velocities corresponding to large-scale structures like galaxy groups or clusters. Moreover, the inclusion of distant objects reinforces the idea that the search for RELHICs is not limited to the local Universe but spans a significant cosmological volume.

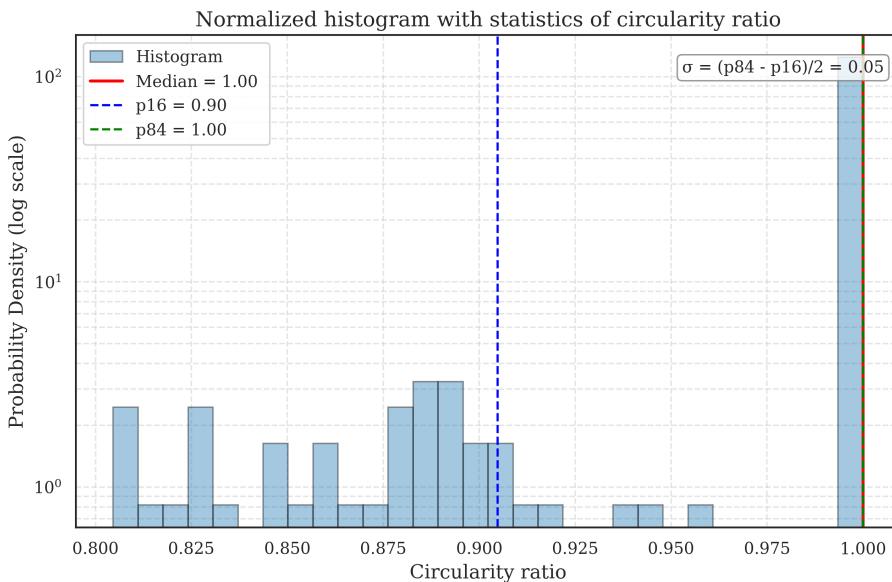


Figure 5.5: Distribution of circularity (b/a).

The normalized histogram in Figure 5.5 displays the distribution of the circularity ratio for the selected RELHIC candidates. The circularity ratio, defined as the ratio of the minor to the major axis of the HI emission, tells us how spherical the gas distribution is. The data show an overwhelming peak at a ratio of 1.0, with the 84th percentile at  $p_{84} = 1.00$  and the 16th at  $p_{16} = 0.90$ , yielding a robust dispersion of  $\sigma = (p_{84} - p_{16})/2 = 0.05$ . This tight distribution, concentrated toward nearly perfect circular shapes, strongly supports the theoretical prediction that RELHICs are nearly spherical gas clouds in hydrostatic equilibrium with their host dark matter halos. As shown in simulations [3], RELHICs form in low-density environments where the absence of forces, such as ram-pressure stripping, allows the gas to settle into stable configurations. The sharp peak and small dispersion are consistent with the expected morphology of pressure-supported systems, as opposed to flattened, rotation-dominated galaxies. This is in agreement with the predicted structural properties of RELHICs from  $\Lambda$ CDM simulations, further reinforcing the interpretation of these systems as the observable counterparts of dark minihalos. The figures presented thus validate the sources of our catalogue as potential RELHICs candidates. However, it is important to consider an alternative explanation for this high circularity. Given the resolution limit of the FAST telescope (beam FWHM  $\approx 3$  arcmin), many of these sources may be unresolved. In such cases, the observed morphology would reflect the projected shape of the instrument beam rather than the intrinsic structure of the HI source. To verify this, one can examine the absolute values of the major and minor axes ( $a$  and  $b$ ). If both are close to 3 arcmin, this suggests that the source is spatially unresolved, and the high circularity may not be a reliable indicator of a spherical gas distribution. Here we report the distribution of  $a$  ( $ell_{maj}$ ) and  $b$  ( $ell_{min}$ ) after, in the scatter section, we analyze in more detail this problem.

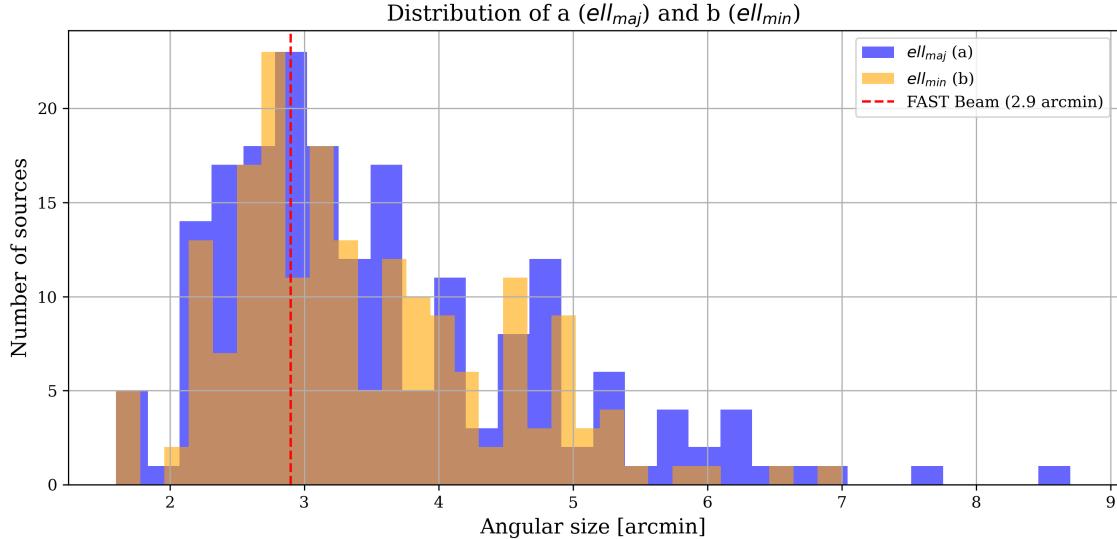


Figure 5.6: Distribution of a ( $ell_{maj}$ ) and b ( $ell_{min}$ ).

Figure 5.6 shows the distribution of angular sizes (major and minor axes,  $ell_{maj}$  and  $ell_{min}$ ) for all RELHIC candidates. The vertical dashed line marks the beam size of FAST (2.9 arcmin), which represents the resolution limit. A significant number of sources are clustered near or below this value, suggesting that unresolved detections are common.

### 5.1.2 Corner Plot

Figure 5.7 presents a corner plot of six key observational parameters for the selected RELHIC candidates: right ascension (RA), declination (Dec), heliocentric velocity ( $cz$ ), HI line width at 50% and 20% of peak intensity ( $W_{50}$  and  $W_{20}$ ), and peak flux density ( $F_{peak}$ ). Each diagonal panel shows the kernel density estimate of the distribution for a single variable, while the off-diagonal panels display scatter plots for every pairwise combination. The corner plot provides a comprehensive visual overview of potential correlations, clustering trends, and outliers across the dataset. Some notable observations include:

- Lack of strong correlation between RA and Dec: The sky distribution appears roughly uniform within the observational window of the FASHI survey, indicating no obvious clustering in celestial coordinates. This spatial uniformity supports the idea that the selection criteria did not bias the sample toward any sky region.

- $c_z$  vs.  $W_{50}$  and  $W_{20}$ : There is no clear trend between redshift and HI line width, suggesting that the internal dynamics of the clouds are independent of their distance within the range probed. This is consistent with the interpretation that  $W_{50}$  is primarily governed by internal thermal and gravitational equilibrium rather than large-scale environmental factors.
- $F_{peak}$  shows a weak inverse correlation with both  $W_{50}$  and  $W_{20}$ : While subtle, this may reflect the fact that more compact, narrow-linewidth sources tend to have higher peak fluxes, since their emission is more concentrated in frequency space. However, this effect could also be influenced by instrumental sensitivity and noise, and should be interpreted cautiously.
- Distribution asymmetries: The marginal distributions confirm the findings of the previous sections:  $W_{50}$  and  $W_{20}$  are strongly peaked toward low values, and  $c_z$  is widely dispersed with a slight accumulation at lower redshifts.

Overall, the corner plot does not reveal any dramatic substructures or correlations that would suggest contamination by known galactic systems or systematic observational bias. Instead, it supports the idea that the candidate RELHICs form a diverse but physically coherent population, largely consistent with expectations from cosmological simulations of starless dark matter halos.

Corner Plot of Key Parameters

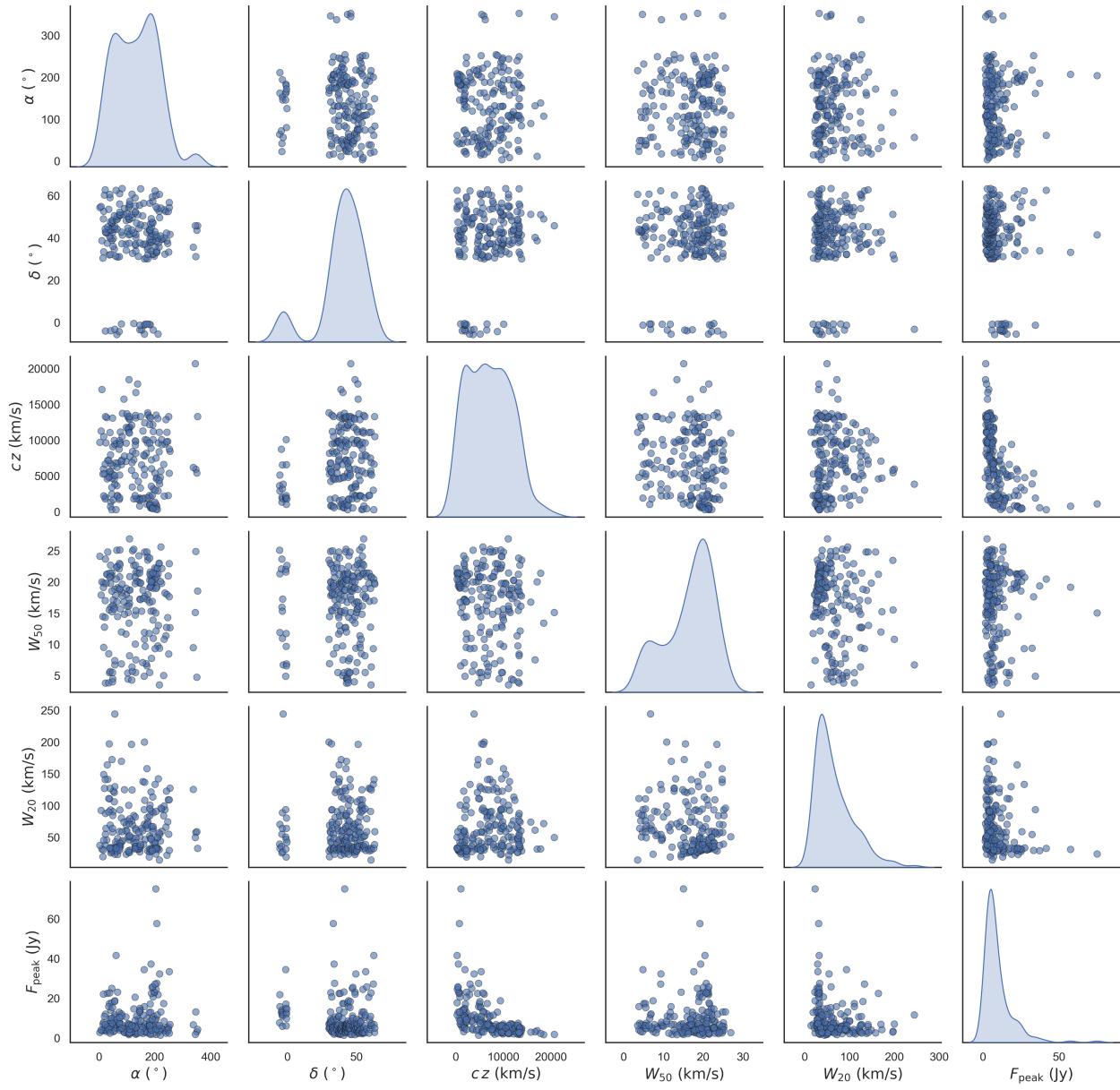


Figure 5.7: Corner plot of key parameters.

### 5.1.3 Scatter plots

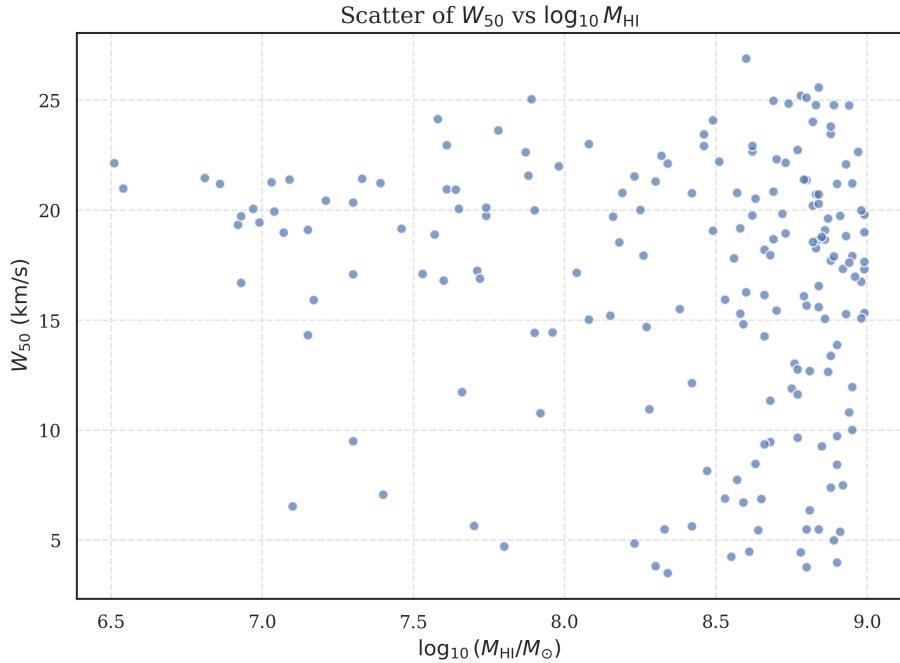


Figure 5.8: Relationships between  $W_{50}$  and  $M_{\text{HI}}$ .

Figure 5.8 shows the scatter plot of  $W_{50}$  versus  $\log_{10}(M_{\text{HI}})$  for the candidate RELHICs. This figure is particularly important for investigating whether there exists a correlation between the amount of neutral hydrogen present in the systems and their internal kinematic state. The distribution is notably dispersed, without a clear trend. However, some patterns emerge upon closer inspection:

- Most sources are clustered in the upper-right region (48.9%), with  $\log_{10}(M_{\text{HI}}/M_{\odot}) > 8$  and  $W_{50} \gtrsim 15$  km/s. This group likely represents the most massive and kinematically “warm” clouds in the sample and could possibly be galaxies missed in the SDSS catalogue.
- A secondary population occupies the left quadrant (26.1%), with  $\log_{10}(M_{\text{HI}}/M_{\odot}) < 8$  and line widths spanning a broader range, including values below 10 km/s. These may be more consistent with thermally supported RELHICs in the lowest-mass halos.

- An additional group (16%) is composed of systems with relatively high HI masses ( $\log_{10}(M_{\text{HI}}/M_{\odot}) > 8$ ) but narrow line widths ( $W_{50} < 10 \text{ km/s}$ ). These may represent either nearby RELHICs with higher gas content or particularly cold, pressure-supported systems. Their inclusion as extended RELHIC-like candidates is motivated by their low dynamical activity, despite their higher mass.
- A small number of sources with intermediate properties ( $8 < \log_{10}(M_{\text{HI}}) < 9$ ,  $10 < W_{50} < 15 \text{ km/s}$ ) remain unclassified. These systems fall outside the conservative thresholds used for the three main categories and may represent a mixture of unresolved, transitional, or ambiguous cases.
- The absence of objects with both very low mass and high line width is consistent with physical expectations: less massive halos do not have deep enough gravitational potentials to sustain fast internal motions.

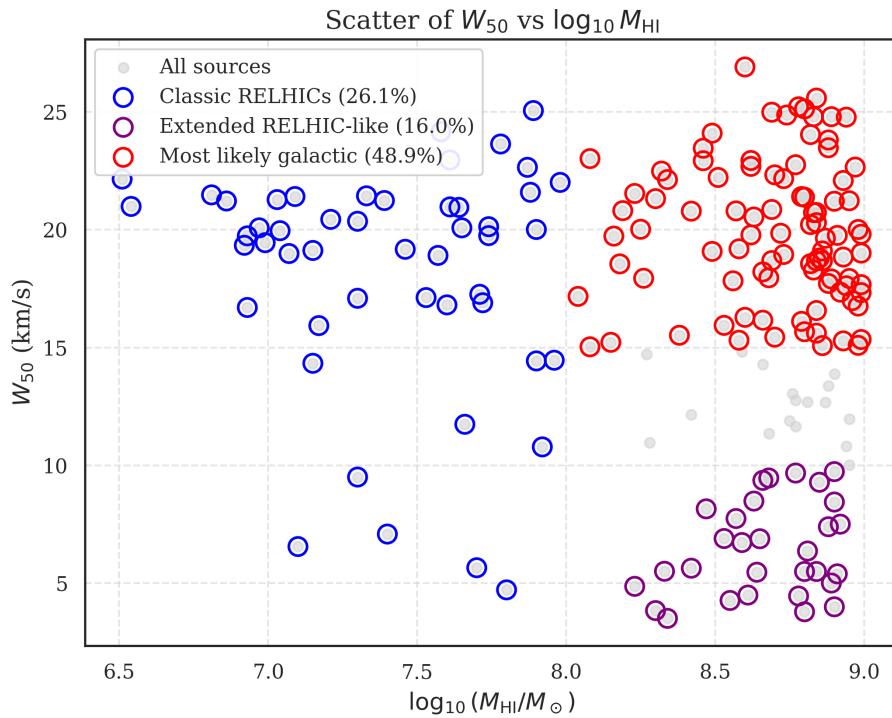


Figure 5.9: Relationships between  $W_{50}$  and  $M_{\text{HI}}$  highlighted.

From a theoretical standpoint, the relationship between mass and velocity dispersion is expected in virialized systems, where  $W_{50}$  should increase with increasing mass. However,

for RELHICs — which may not be fully virialized and are regulated by photoionization heating — the relation is not necessarily monotonic and may show a large intrinsic scatter [1]. The lack of a tight correlation may also reflect observational uncertainties, selection effects, and the intrinsic complexity of these systems. Nevertheless, figure 5.8 helps illustrate the range of physical regimes spanned by the sample and reinforces the interpretation of RELHICs as diverse objects inhabiting a narrow window between gravitational collapse and thermal pressure support.

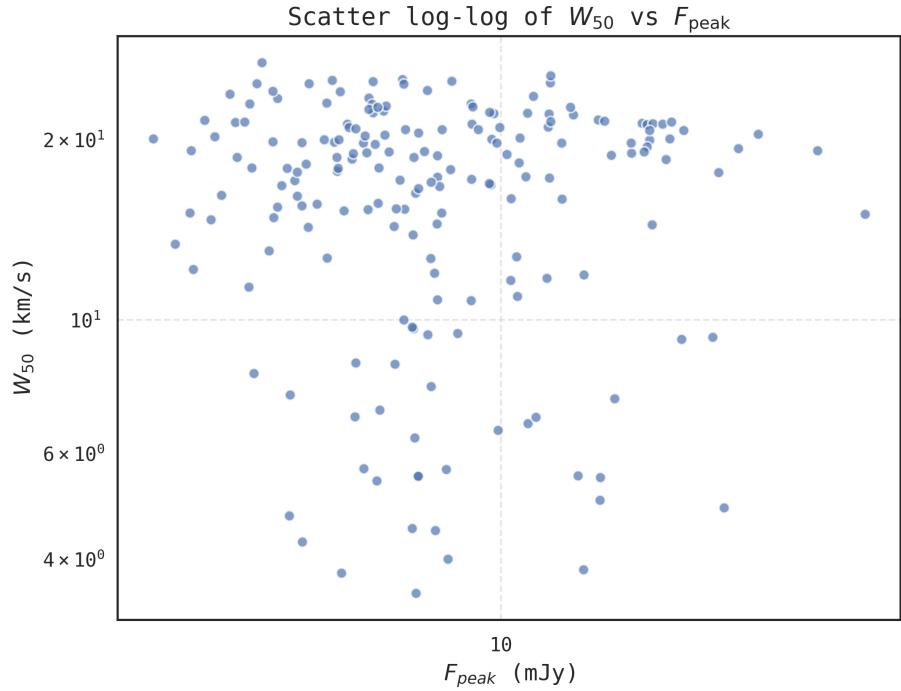


Figure 5.10: Relationships between  $W_{50}$  and peak flux density  $F_{peak}$ .

Figure 5.10 shows the scatter plot of HI line width ( $W_{50}$ ) versus peak flux density ( $F_{peak}$ ) with both axes plotted on a logarithmic scale to better visualize the broad dynamic range of the sample. This representation is particularly insightful for understanding the relationship between the brightness of the source and its kinematic state. The majority of sources cluster in the upper part of the plot, where  $F_{peak}$  typically ranges between 5 and 20 mJy and  $W_{50}$  lies around 15–25 km/s. These objects are moderately bright and dynamically active HI clouds, but their relatively large line widths indicate that they are not the most dynamically cold systems. According to theoretical predictions [3], the most

promising RELHIC candidates are expected to occupy the lower-left region of the diagram, characterized by both low  $F_{\text{peak}}$  and narrow linewidths ( $W_{50} \lesssim 10$  km/s), reflecting their faintness and dynamically cold nature. In our sample, only a minority of sources falls in this region, which is consistent with the predicted rarity and intrinsic faintness of RELHICs. At lower  $F_{\text{peak}}$  values (below 6 mJy), we observe a wider spread in  $W_{50}$ , including several sources with narrow line widths (<10 km/s). These may represent low-signal, thermally supported clouds, possibly near the sensitivity limit of the FASHI survey. Conversely, objects with both high  $F_{\text{peak}}$  and narrow  $W_{50}$  are rare, likely due to the inherent challenge in detecting strong, unresolved lines in extended, starless structures. The absence of a strong positive or negative correlation suggests that peak flux alone is not a reliable predictor of line width in this regime. However, there is a mild trend indicating that narrower lines tend to appear in fainter sources, a feature that could be explained by the thermal origin of their emission and the absence of turbulence or rotation.

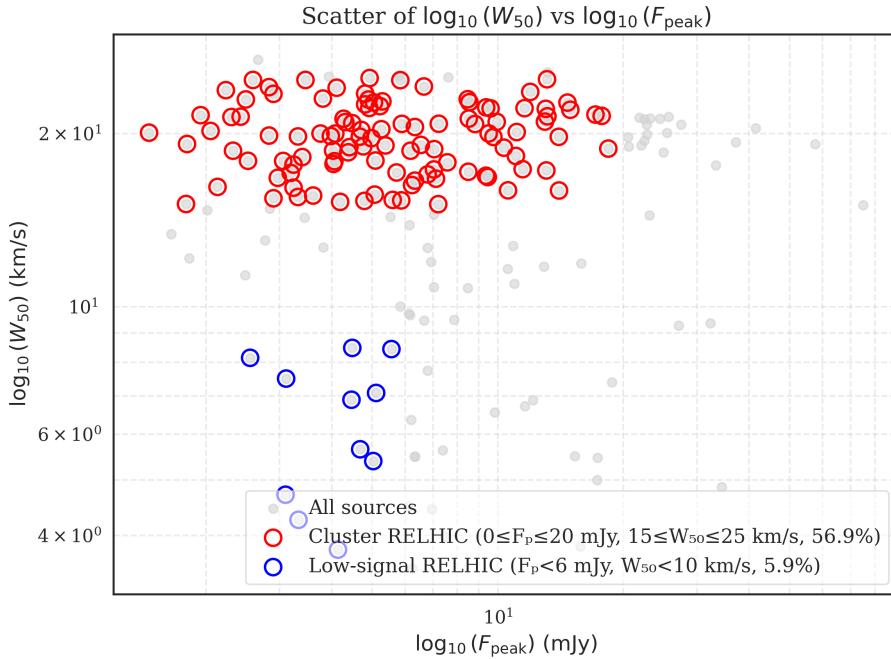


Figure 5.11: Relationships between  $W_{50}$  and peak flux density  $F_{\text{peak}}$  highlighted.

Figure 5.11 further illustrates the distribution of sources in the  $W_{50}(F_{\text{peak}})$  plane, highlighting two categories: *Cluster RELHIC* (red circles) and *Low-signal RELHIC* (blue circles). Most sources fall within the *Cluster RELHIC* region, which is characterized by

moderate flux ( $0 \leq F_{\text{peak}} \leq 20$  mJy) and relatively broad lines ( $15 \leq W_{50} \leq 25$  km/s). These criteria are designed to capture compact HI sources that may exhibit the morphological and kinematic signatures of RELHICs while still remaining detectable with sufficient signal-to-noise ratio. In contrast, the *Low-signal RELHIC* group includes only a small fraction (5.9%) of sources with  $F_{\text{peak}} < 6$  mJy and  $W_{50} < 10$  km/s values expected for thermally broadened, dynamically cold gas at  $T \sim 10^4$  K, consistent with  $\Lambda$ CDM predictions for genuine RELHICs residing in low-mass halos [4]. These selection boundaries are not arbitrary: the thermal line width at  $T \approx 2 \times 10^4$  K implies  $W_{50} \sim 10$  km/s, and fluxes below 6 mJy are characteristic of HI clouds with very low HI mass or located at larger distances. By isolating these two regimes, the visualization highlights both the typical observational footprint of potential RELHICs (*Cluster*) and the rare, but theoretically most pristine candidates (*Low-signal*). This dual classification underscores the difficulty in detecting RELHICs in surveys and the need for robust criteria to separate them from other HI structures or noise. Thus, the visualization not only organizes the data but also reinforces the theoretical expectations derived from hydrodynamic simulations [4].

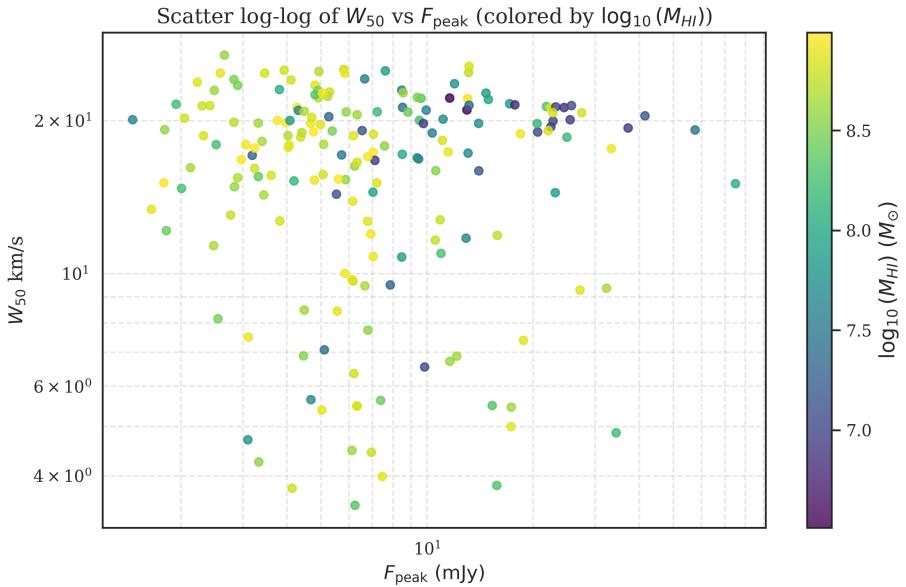


Figure 5.12: Relationships between  $W_{50}$  and peak flux density  $F_{\text{peak}}$  color-coded by  $\log_{10}(M_{\text{HI}})$ .

Figure 5.12 presents the same scatter plot as in Figure 5.10, but now the data points

are color-coded according to the value of  $\log_{10}(M_{\text{HI}})$ . This coloring allows for a deeper understanding of how mass correlates with the dynamics (represented by  $W_{50}$ ) and the observed brightness ( $F_{\text{peak}}$ ). Several key trends emerge from this plot:

- High-mass systems (with  $\log_{10}(M_{\text{HI}}) \gtrsim 8.5$ , shown in yellow and green) are distributed across a wide range of  $F_{\text{peak}}$  and  $W_{50}$ , but are often found among sources with relatively high peak flux and/or broader line widths. This reflects the tendency for more massive HI clouds to be more luminous or dynamically hotter, possibly due to their larger gravitational potential wells, or possible presence of undetected galaxies.
- Contrary to theoretical expectation, low-mass systems (with  $\log_{10}(M_{\text{HI}}/M_{\odot}) \lesssim 7.0$ , shown in blue and purple) are not exclusively found among the faintest and dynamically coldest sources. Instead, many low-mass clouds are also present in the region with moderate  $F_{\text{peak}}$  and intermediate  $W_{50}$  (15–25 km/s), which corresponds to the so-called "Cluster RELHIC" area in Figure 4.9. This suggests that even clouds with relatively low HI mass can display a variety of kinematic properties, and are not strictly limited to the lower-left corner of the diagram. The physical diversity of these low-mass systems may reflect a combination of observational limits, intrinsic differences in halo properties, or environmental effects that broaden their linewidths.

Overall, the plot highlights that, although the distribution does not show a tight correlation between  $F_{\text{peak}}$  and  $W_{50}$ , the properties of each system are strongly influenced by its HI mass. The higher-mass HI clouds span a broad range of dynamics and luminosities, while the lower-mass systems are primarily found among the faint and dynamically cold population. This analysis supports the physical diversity of the observed sample, which encompasses a range of mass, dynamics, and luminosity, in agreement with the theoretical expectations for RELHICs within the  $\Lambda$ CDM framework.

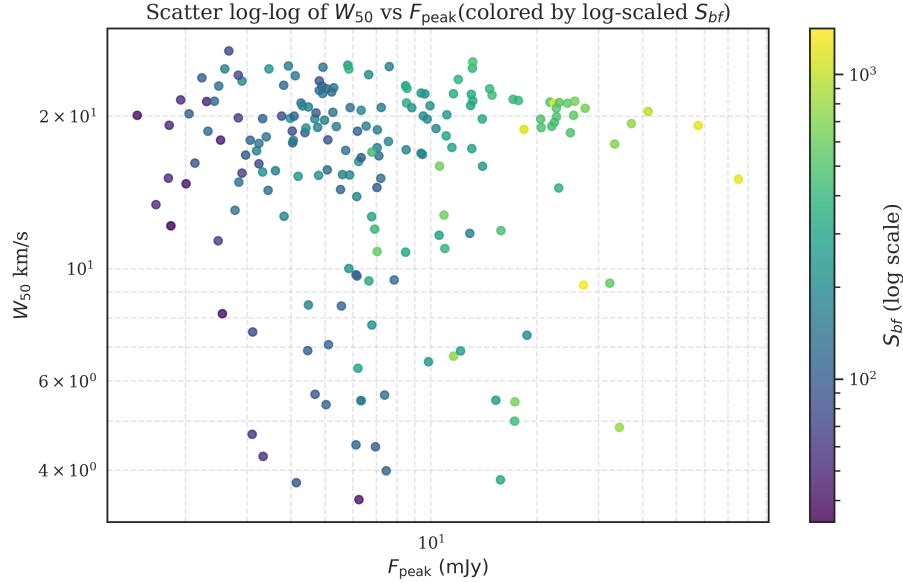


Figure 5.13: Relationships between  $W_{50}$  and peak flux density  $F_{\text{peak}}$  color-coded by  $S_{bf}$ .

Figure 5.13 shows the scatter plot of HI line width ( $W_{50}$ ) versus peak flux density ( $F_{\text{peak}}$ ), with each data point color-coded by the integrated HI flux ( $S_{bf}$ ). Both axes use logarithmic scales, and the color bar to the right indicates the  $S_{bf}$  value, from purple (lowest integrated flux) to yellow (highest).

The main features emerging from the plot are:

- The majority of sources are clustered in the upper left and central upper regions of the diagram, corresponding to high  $W_{50}$  (typically  $>15$  km/s) and low to intermediate  $F_{\text{peak}}$  (ranging from about 4 to 20 mJy).
- Most points in these areas are colored in dark purple, indicating that they have low integrated flux ( $S_{bf}$ ), generally below 1000 mJy km/s.
- Only a small fraction of sources, mainly at higher  $F_{\text{peak}}$  and high  $W_{50}$ , show colors shifting toward blue or cyan, which correspond to higher  $S_{bf}$ . The brightest integrated flux sources (yellow) are rare in the sample.
- The lower left region of the plot (low  $F_{\text{peak}}$  and low  $W_{50}$ ), theoretically favored for the faintest and most dynamically cold RELHIC candidates, is sparsely populated, suggesting both intrinsic rarity and observational limits in detecting such objects.

In summary, Figure 5.13 indicates that the sample is dominated by HI clouds with moderate to low peak flux, relatively large line widths, and low integrated flux, while only a few sources are both bright and kinematically broad. This distribution highlights the physical diversity of the detected systems and is broadly consistent with theoretical expectations for a population of HI-rich halos within the  $\Lambda$ CDM model, where only a minority represent the most pristine, faint RELHICs.

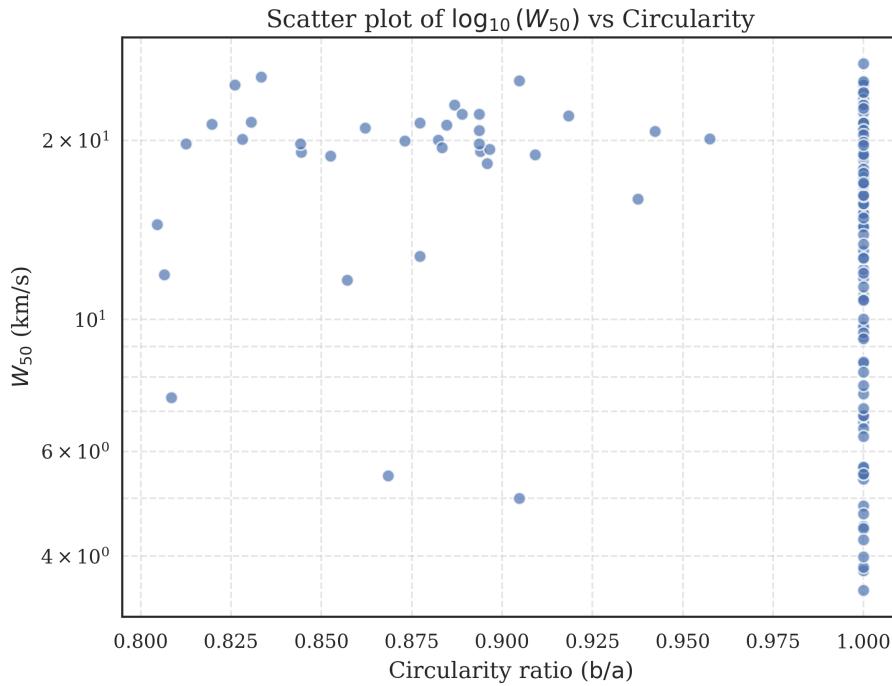


Figure 5.14: Scatter plot of  $W_{50}$  versus circularity ratio ( $b/a$ ).

Figure 5.14 shows the scatter plot of HI line width ( $W_{50}$ ) versus circularity of the gas content. Most of the objects cluster at very high circularity values (near  $b/a = 1$ ), consistent with spherically symmetric gas distributions. The vertical spread in  $W_{50}$  spans from  $\sim 4$  km/s up to  $\sim 25$  km/s, suggesting a mix of thermally supported systems and more massive, dynamically warmer ones. Notably, the plot indicates a sharp accumulation of objects at the high circularity end, while more flattened sources are comparatively rare. While this may suggest intrinsically spherical gas distributions, it is important to consider that unresolved sources can appear artificially round due to the beam convolution, potentially biasing the circularity estimate, as already mentioned at page 31.

### RELHIC-like candidates: Cold vs Warm classification

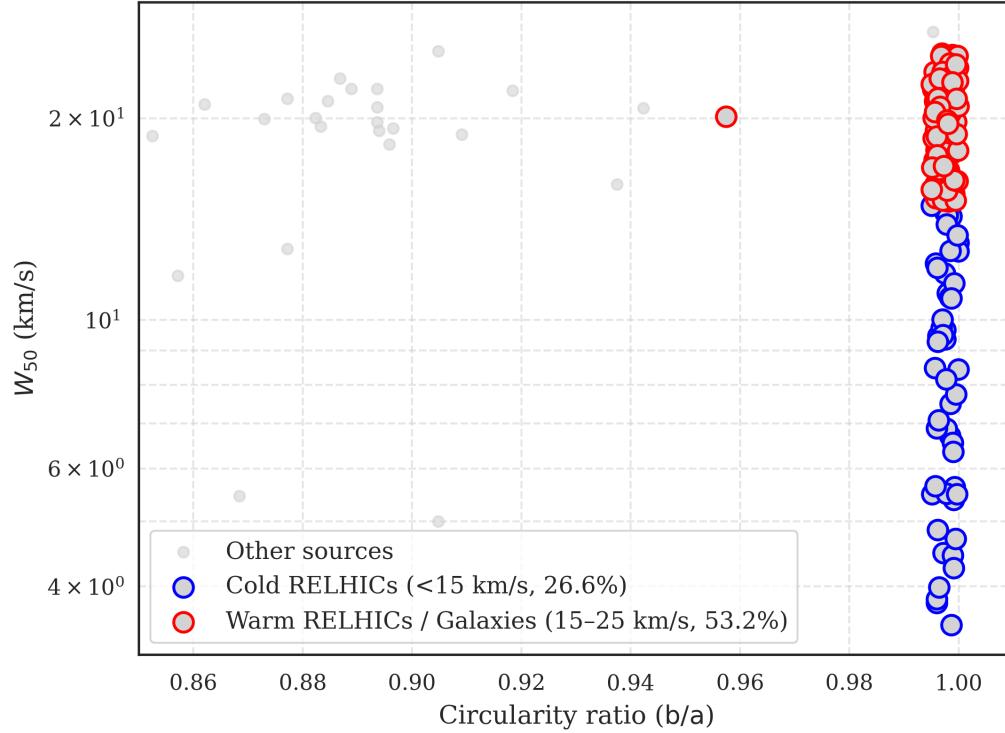


Figure 5.15: Scatter plot of  $W_{50}$  versus circularity ratio ( $b/a$ ) highlighted.

In Figure 5.15, we isolate the subset of sources that match theoretical expectations for RELHICs. We define two categories: “Cold RELHICs”, with circularity  $> 0.95$  and  $W_{50} < 15$  km/s, and “Warm RELHICs/potential Galaxies” with the same geometric constraint but broader linewidths ( $15 \leq W_{50} < 25$  km/s). The former likely trace low-mass halos in hydrostatic equilibrium, where thermal motion dominates the internal dynamics, and account for 26.6% of the total sample. The latter group may correspond to higher-mass halos or other systems like galaxies, and comprise 52.7% of the sample. This classification reinforces the notion that RELHICs occupy a well defined region of the circularity  $W_{50}$  plane, characterized by high morphological symmetry and modest internal kinematics. However, it is important to note that some of the Warm RELHICs, despite lacking optical counterparts, may in fact be faint or unresolved galaxies. Their larger linewidths ( $W_{50} > 15$  km/s) are consistent with either more massive pressure-supported halos or systems exhibiting rotational broadening.

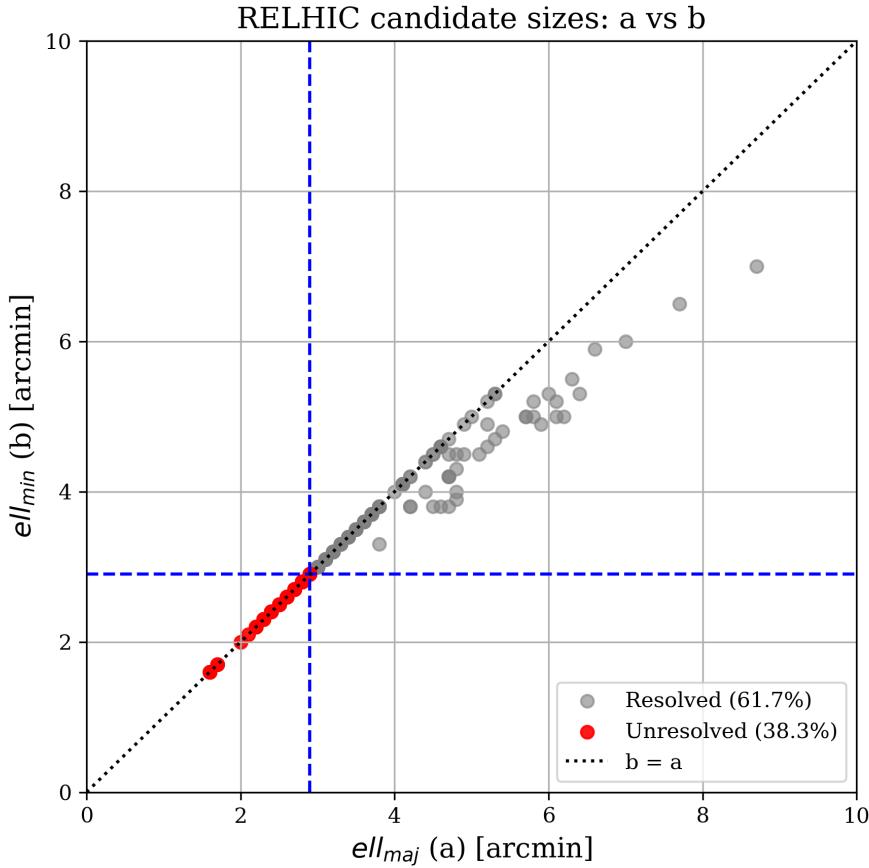


Figure 5.16: Scatter plot of a ( $ell_{maj}$ ) versus b ( $ell_{min}$ ).

Figure 5.16 shows the scatter plot of major versus minor axes for our sample. The dashed lines indicate the FAST resolution limit (2.9 arcmin). Sources falling within the lower left quadrant are considered unresolved. We find that 38.3% of the RELHIC candidates are unresolved, suggesting that their observed circularity may reflect the beam profile rather than intrinsic morphology. Follow-up HI mapping or deep optical imaging would be needed to confirm their true nature.

## 6. Conclusions

This thesis has aimed to identify, produce the first catalogue and analyze a population of candidate Reionization-Limited HI Clouds (RELHICs), theoretical objects predicted by the  $\Lambda$ CDM cosmological model in the low-mass regime. Through a systematic analysis of radio data from the FAST All Sky HI Survey (FASHI), and subsequent cross-matching with the optical catalog of the Sloan Digital Sky Survey (SDSS), a catalog of HI sources without optical counterparts has been constructed. The methodology involved the use of stringent selection criteria, informed by cosmological hydrodynamical simulations, to isolate only those sources that met the expected physical conditions for RELHICs. This approach yielded a significant subsample of plausible candidates.

### 6.0.1 Interpretation of Results and Cosmological Implications

Statistical analysis revealed that the kinematic and morphological properties of these sources are consistent with theoretical predictions: the HI lines are narrow and symmetric, suggesting dynamic stability dominated by thermal pressure rather than turbulent or rotational motions. The mass values are distributed within a range compatible with that predicted for subcritical halos, incapable of initiating star formation but still able to retain neutral gas due to gravitational confinement and thermal equilibrium imposed by the cosmic UV background. Visualizations through scatter plots and pair plots further confirmed the lack of strong correlations with distance or luminosity, reinforcing the hypothesis that these objects constitute a heterogeneous yet coherent population. The apparent spatial distribution of the sources, not concentrated in any particular region of the sky, supports the idea that they are isolated objects, not associated with known galaxies or local structures. Overall, the results obtained support the existence of a population of dark matter halos without stars, consistent with the RELHICs predicted by the  $\Lambda$ CDM model. While further observational confirmations are necessary, such as high-resolution follow-up radio observations and deeper photometric observation in optical wavelengths, our work constitutes a concrete first step toward the compilation of a systematic catalog of these objects.

## 6.0.2 Future Prospects

This catalogue open promising avenues for future observational and theoretical analysis. With the advent of the next generation astronomical surveys, significant progress can be expected in the coming years and decades. Upcoming deeper and more sensitive optical surveys, such as the Dark Energy Spectroscopic Instrument (DESI) and the Vera C. Rubin Observatory's Legacy Survey of Space and Time (LSST), will provide essential data for cross-matching. These surveys will achieve unprecedented depths and sky coverage, giving us the confirmation or exclusion of stellar populations in the currently identified RELHIC candidates. On the radio side, the upcoming Square Kilometre Array (SKA) will bring revolutionary progress in detecting and characterizing neutral hydrogen structures. Given its sensitivity and resolution, SKA would best be suited to add to our present catalogue of RELHIC candidates. The catalogue created in this thesis would make an excellent foundation and calibration set, for future surveys with the SKA. In addition, hydrodynamical simulations will continue to parametrize our theoretical framework of the  $\Lambda$ CDM model. In the end, the synergy of future observational surveys with cosmological models will further our understanding of starless dark matter halos, their contribution to galaxy formation, and their importance in the cosmological framework.

# A. A Catalogue of Starless Dark Matter Halo Candidates Detected with FAST

Table A.1: Catalogue of candidate starless dark matter halos.

ID	RA [deg]	DEC [deg]	$\log_{10}(M_{\text{HI}}/M_{\odot})$	$W_{50}$ [km/s]	$F_{\text{peak}}$ [mJy]	$b/a$
20230042101	63.7487	-5.8178	7.8900	25.0600	7.6100	1.0000
20230001032	162.2959	-5.5824	8.7500	11.8900	15.8200	0.8065
20230051923	211.2052	-5.5264	8.3000	21.3200	21.8100	1.0000
20230060231	71.9649	-4.3679	8.9300	22.0900	13.0300	1.0000
20230001920	22.8130	-4.0396	7.7800	23.6400	11.9600	1.0000
20230060238	41.8749	-4.0140	7.1700	15.9300	14.0200	0.9375
20230002250	194.3416	-3.6777	8.2800	10.9600	10.9500	1.0000
20230060253	152.5343	-3.6640	8.5800	15.3100	5.8700	1.0000
20230002479	146.0171	-3.4474	7.7100	17.2600	13.0700	1.0000
20230060269	56.5118	-3.2042	8.5900	6.7200	11.6000	1.0000
20230004171	161.2406	-1.3552	8.2300	4.8600	34.3900	1.0000
20230004180	145.9646	-1.3250	7.9800	22.0100	14.9200	0.9184
20230004300	183.4720	-1.0897	7.8800	21.5900	17.1400	1.0000
20230004437	165.4888	-0.9132	8.7700	9.6700	6.1600	1.0000
20230004467	179.8896	-0.8736	7.6600	11.7400	12.9000	1.0000
20230004560	79.9688	-0.7669	8.6500	6.8800	12.1400	1.0000
20230004629	171.8309	-0.7301	7.1000	6.5500	9.8400	1.0000
20230004813	124.9601	-0.4454	7.8700	22.6500	14.6900	1.0000
20230058023	162.7119	30.0593	8.9400	10.8100	7.0300	1.0000
20230058029	174.2158	30.0897	8.8800	23.4800	4.8100	1.0000

Continued on next page

Table A.1: Catalogue of candidate starless dark matter halos.

ID	RA [deg]	DEC [deg]	$\log_{10}(M_{\text{HI}}/M_{\odot})$	$W_{50}$ [km/s]	$F_{\text{peak}}$ [mJy]	$b/a$
20230060788	14.4494	30.3543	7.9600	14.4600	7.0200	1.0000
20230006431	168.6336	30.4979	8.5300	6.8900	4.4600	1.0000
20230006836	66.2202	30.9361	8.4900	19.0800	5.3800	1.0000
20230042219	72.6484	31.0720	7.6100	22.9600	8.4700	0.8868
20230007060	346.5075	31.1281	8.7400	24.8700	13.1200	1.0000
20230065109	141.0249	31.2857	8.2500	20.0200	9.5100	0.8824
20230007268	59.3309	31.3573	8.5600	17.8300	7.5700	1.0000
20230007323	198.3234	31.4079	7.0400	19.9600	22.7700	0.8730
20230008133	36.4147	31.9650	8.3800	15.5200	3.3200	1.0000
20230008317	41.4255	32.1237	8.6200	22.6900	8.5500	1.0000
20230058517	192.9119	32.1816	7.0700	18.9900	20.5900	1.0000
20230008574	175.4561	32.2458	8.9500	17.9300	3.0600	1.0000
20230008700	235.9656	32.3444	8.6600	18.2100	3.4000	1.0000
20230008787	194.3485	32.4083	8.9900	15.3400	5.6000	1.0000
20230053742	119.6368	32.6981	8.6900	24.9900	4.9300	1.0000
20230061145	205.3185	33.1330	8.3400	22.1300	9.6100	0.8889
20230061168	206.8948	33.2024	7.4600	19.1700	57.6000	0.8939
20230010235	249.0831	33.3212	8.9000	8.4400	5.5600	1.0000
20230061208	17.2210	33.3791	8.9900	19.8100	3.9800	1.0000
20230010719	132.1558	33.7000	8.2300	21.5400	1.9400	1.0000
20230010881	185.7150	33.8268	6.9200	19.3400	37.1900	0.8966
20230011160	206.5286	34.0412	8.8300	20.7300	27.4700	1.0000
20230011202	97.6083	34.0990	8.2600	17.9400	2.5200	1.0000
20230052081	173.7489	34.5540	7.5300	17.1200	5.7200	1.0000
20230011688	55.2591	34.6583	8.5300	15.9400	10.5800	1.0000

Continued on next page

Table A.1: Catalogue of candidate starless dark matter halos.

ID	RA [deg]	DEC [deg]	$\log_{10}(M_{\text{HI}}/M_{\odot})$	$W_{50}$ [km/s]	$F_{\text{peak}}$ [mJy]	$b/a$
20230061663	337.0120	35.5500	8.6800	9.4600	6.6600	1.0000
20230058895	215.0038	35.7062	8.4600	22.9300	4.9000	1.0000
20230053964	173.8639	35.8005	8.8200	24.0400	4.1100	1.0000
20230013053	200.1429	35.8833	8.5100	22.2200	9.3800	1.0000
20230052219	49.9433	36.0400	8.6000	16.2800	6.2300	1.0000
20230054537	130.9198	36.5720	8.8000	5.4900	6.3300	1.0000
20230055005	157.2901	36.8948	8.8400	18.6800	2.3200	1.0000
20230014939	130.3780	36.9431	8.2700	14.7100	2.0100	1.0000
20230061728	96.6853	37.0822	8.7700	11.6400	10.5500	0.8571
20230015282	163.1264	37.1643	8.8300	24.8000	3.4600	1.0000
20230054777	82.2617	37.2332	8.8300	18.2900	11.0500	0.8958
20230015606	245.9890	37.3178	8.6600	14.2800	3.4400	1.0000
20230015793	108.2297	37.4375	8.9900	17.3300	7.0300	1.0000
20230015820	29.8185	37.4584	8.9000	21.2100	22.4400	1.0000
20230016103	146.7828	37.6178	7.7400	19.7500	14.0000	0.8125
20230054509	212.2800	37.6637	8.7600	13.0400	2.7700	1.0000
20230054754	181.4496	37.8250	8.8800	23.8100	2.2300	1.0000
20230054763	185.6696	38.0861	8.9300	15.2800	4.7900	1.0000
20230016828	188.2739	38.1163	6.9900	19.4500	22.4500	0.8833
20230017251	204.0725	38.3758	8.4200	12.1500	1.8200	1.0000
20230050960	10.8050	38.6316	8.9800	20.0100	3.7600	1.0000
20230054588	243.4606	38.6576	8.6800	17.9600	5.0900	1.0000
20230042416	46.8392	38.9615	8.1500	15.2200	4.1900	1.0000
20230018173	192.4863	39.0115	8.8200	20.2200	2.0500	1.0000
20230061878	224.9860	39.2561	8.7300	22.1700	4.9200	1.0000

Continued on next page

Table A.1: Catalogue of candidate starless dark matter halos.

ID	RA [deg]	DEC [deg]	$\log_{10}(M_{\text{HI}}/M_{\odot})$	$W_{50}$ [km/s]	$F_{\text{peak}}$ [mJy]	$b/a$
20230061879	224.4421	39.2642	8.9100	19.7600	4.6700	1.0000
20230142538	181.8789	39.2916	7.1500	19.1200	6.5400	1.0000
20230019574	50.8640	39.7420	7.3000	17.0900	3.1900	1.0000
20230019820	41.3846	39.9261	6.9300	19.7400	9.7600	0.8936
20230019891	210.1121	39.9816	8.9000	9.7300	6.1200	1.0000
20230019963	231.5699	40.0417	8.8900	5.0000	17.3000	0.9048
20230020202	131.8818	40.1814	8.9200	7.5000	3.1100	1.0000
20230021424	197.9224	40.8621	7.0300	21.2800	23.1800	0.8197
20230021454	185.0453	40.8773	6.5400	20.9900	12.9700	0.8621
20230021456	37.6114	40.8798	8.6400	5.4600	17.3300	0.8684
20230021521	219.2547	40.9188	7.6500	20.0800	1.4600	1.0000
20230021993	84.2581	41.2556	8.8800	17.7200	4.0400	1.0000
20230055325	107.7326	41.3994	8.8600	18.6800	6.1800	1.0000
20230057538	203.5468	41.4645	8.0800	15.0300	74.9800	1.0000
20230022468	131.4512	41.6698	8.7000	15.4400	2.9000	1.0000
20230062160	254.0326	41.7218	8.8700	12.6700	6.7800	1.0000
20230062178	111.3706	41.7791	8.6900	18.6900	4.0300	1.0000
20230023111	188.1473	42.1180	8.6300	20.5400	6.3300	1.0000
20230055657	241.7548	42.2803	8.8100	12.6900	3.8200	1.0000
20230023570	27.4370	42.3620	8.9500	21.2300	4.2700	1.0000
20230023631	25.1002	42.4081	8.6200	19.7800	3.3200	1.0000
20230023665	124.2846	42.4306	8.7800	25.2200	5.8000	0.9048
20230055733	171.7249	42.4495	8.8900	24.8000	2.5900	1.0000
20230024193	80.1378	42.7209	8.9500	11.9700	6.9200	1.0000
20230024196	23.3365	42.7239	8.8000	3.7800	4.1400	1.0000

Continued on next page

Table A.1: Catalogue of candidate starless dark matter halos.

ID	RA [deg]	DEC [deg]	$\log_{10}(M_{\text{HI}}/M_{\odot})$	$W_{50}$ [km/s]	$F_{\text{peak}}$ [mJy]	$b/a$
20230024245	49.0750	42.7527	8.6100	4.4900	6.1200	1.0000
20230024607	200.1025	42.9814	7.3000	20.3600	5.2600	1.0000
20230024800	104.1525	43.1570	8.7800	4.4500	6.9500	1.0000
20230024885	181.2554	43.2367	6.9300	16.7000	7.1100	1.0000
20230025255	350.1352	43.5568	7.8000	4.7100	3.1000	1.0000
20230025382	45.5806	43.6630	6.8100	21.4800	17.7400	0.8305
20230025553	117.7013	43.7923	8.5500	4.2600	3.3300	1.0000
20230025566	40.2894	43.8015	8.4200	5.6300	7.3800	1.0000
20230055819	181.4133	43.8048	8.8400	20.3000	4.7100	1.0000
20230025721	182.3359	43.9523	7.3900	21.2400	8.5200	1.0000
20230025798	34.9173	44.0249	8.3000	3.8300	15.7700	1.0000
20230025973	56.8552	44.1489	8.8800	7.3900	18.7600	0.8085
20230025978	64.9047	44.1550	8.3200	22.4900	4.8100	1.0000
20230026148	22.4363	44.3173	7.1500	14.3300	5.5300	1.0000
20230026367	142.3560	44.5144	7.5800	24.1600	6.6500	1.0000
20230026495	103.1265	44.6699	8.9100	5.3900	5.0300	1.0000
20230026886	188.0430	45.0944	8.5700	7.7400	6.7900	1.0000
20230062494	30.9936	45.2374	8.9400	24.7800	5.8400	0.8261
20230027158	48.5480	45.3830	8.3300	5.5000	15.3000	1.0000
20230027183	56.6983	45.4106	8.9300	18.8400	18.4000	0.8525
20230027598	163.3059	45.7294	8.9900	17.6700	3.2400	1.0000
20230027669	352.2484	45.7755	8.8200	18.5600	4.3800	1.0000
20230027678	344.1973	45.7862	8.9800	15.0900	1.7900	1.0000
20230028271	148.3387	46.2378	8.0400	17.1700	8.5000	1.0000
20230028755	184.4659	46.6776	8.5800	19.1900	1.8000	1.0000

Continued on next page

Table A.1: Catalogue of candidate starless dark matter halos.

ID	RA [deg]	DEC [deg]	$\log_{10}(M_{\text{HI}}/M_{\odot})$	$W_{50}$ [km/s]	$F_{\text{peak}}$ [mJy]	$b/a$
20230028995	45.9986	46.9427	8.1800	18.5500	24.9300	1.0000
20230029142	134.1079	47.1047	7.7400	20.1300	11.1000	0.9574
20230029191	193.2892	47.1787	8.9000	13.8800	6.1500	1.0000
20230029291	45.3534	47.2863	7.6100	20.9700	9.9400	1.0000
20230029371	190.1250	47.3654	6.9700	20.0800	25.4600	0.8281
20230029624	41.2180	47.6190	8.5700	20.8000	5.8900	1.0000
20230030154	97.6179	48.2739	8.9900	19.0100	4.7600	1.0000
20230030315	48.8554	48.4183	8.9000	3.9900	7.4600	1.0000
20230030514	62.8702	48.6293	8.7000	22.3300	5.2200	1.0000
20230030583	109.0069	48.7087	8.8400	16.5700	6.3300	1.0000
20230030771	66.5101	48.9194	7.9000	14.4300	23.1000	0.8046
20230030778	106.9859	48.9240	8.8800	13.3800	1.6500	1.0000
20230030827	15.6531	49.0017	8.8600	19.1000	22.0600	0.8444
20230030894	186.7549	49.0712	7.9000	20.0100	4.0800	1.0000
20230052741	118.9058	49.5124	7.7000	5.6500	4.6800	1.0000
20230031322	216.8291	49.5429	8.6600	9.3600	32.2900	1.0000
20230031744	64.8857	50.0712	7.6000	16.8100	9.4600	1.0000
20230056245	104.7198	50.0946	8.8600	15.0800	7.2000	1.0000
20230052692	155.9231	50.4139	7.4000	7.0800	5.1100	1.0000
20230032104	54.9115	50.5038	8.8400	15.6100	3.6100	1.0000
20230032225	75.6704	50.6599	8.7700	22.7600	5.2900	1.0000
20230032351	76.6681	50.8221	8.8000	15.6700	5.0700	1.0000
20230051574	138.5849	50.9502	8.8000	21.3800	2.3000	1.0000
20230032533	73.8977	51.0147	8.8500	18.8000	7.0400	1.0000
20230032608	89.3599	51.1299	8.9800	16.7600	2.9700	1.0000

Continued on next page

Table A.1: Catalogue of candidate starless dark matter halos.

ID	RA [deg]	DEC [deg]	$\log_{10}(M_{\text{HI}}/M_{\odot})$	$W_{50}$ [km/s]	$F_{\text{peak}}$ [mJy]	$b/a$
20230032623	116.3409	51.1419	8.4600	23.4600	2.9000	1.0000
20230033177	100.4002	51.7522	8.7900	16.1100	3.2400	1.0000
20230033225	128.9050	51.7908	8.4200	20.7900	8.8200	0.8936
20230033603	75.0729	52.2194	8.8000	25.1300	3.9300	1.0000
20230033829	239.1335	52.4376	8.6800	11.3500	2.4800	1.0000
20230033861	130.2455	52.4644	8.9500	10.0100	5.8400	1.0000
20230033894	248.0596	52.4922	7.6400	20.9500	4.3100	1.0000
20230063227	215.0129	52.5120	8.0800	23.0200	3.8100	1.0000
20230033982	16.8604	52.5792	7.3000	9.5000	7.8600	1.0000
20230034861	232.8778	53.3137	7.9200	10.7800	8.4700	1.0000
20230035152	34.6572	53.5590	8.6300	8.4800	4.4800	1.0000
20230063286	208.5268	53.7813	6.5100	22.1500	11.5900	0.8936
20230035576	90.9894	53.9004	8.5900	14.8200	2.8400	1.0000
20230035788	221.6758	54.1170	8.8400	25.5900	13.1500	0.8333
20230036013	251.4592	54.3078	8.6200	22.9400	2.4900	1.0000
20230056767	106.2364	54.3236	8.6600	16.1600	2.1300	1.0000
20230056672	3.4857	54.3531	8.4900	24.1000	2.8300	1.0000
20230036458	203.9813	54.7409	6.8600	21.2200	24.4700	0.8846
20230036518	8.0845	54.7925	8.8900	17.9100	4.0600	1.0000
20230036896	130.7697	55.1540	8.9700	22.6600	5.0500	1.0000
20230037009	108.8599	55.2496	8.6000	26.9100	2.6600	1.0000
20230037149	145.7536	55.4253	7.3300	21.4400	13.1500	1.0000
20230037757	109.1264	56.0702	8.4700	8.1500	2.5500	1.0000
20230037771	73.7700	56.0921	7.5700	18.9100	10.3400	0.9091
20230037977	181.6337	56.3976	8.8400	5.4900	6.3200	1.0000

Continued on next page

Table A.1: Catalogue of candidate starless dark matter halos.

ID	RA [deg]	DEC [deg]	$\log_{10}(M_{\text{HI}}/M_{\odot})$	$W_{50}$ [km/s]	$F_{\text{peak}}$ [mJy]	$b/a$
20230038003	160.2244	56.4392	8.9200	17.3500	11.4700	1.0000
20230063559	251.2237	56.7313	8.9400	17.6200	33.3300	1.0000
20230038542	192.7749	57.3717	7.0900	21.4100	25.6800	0.8772
20230038574	110.1877	57.4062	8.6900	20.8600	4.4800	1.0000
20230039657	129.5101	58.8854	8.1600	19.7300	20.5200	0.8442
20230057104	39.0619	59.0936	8.7900	21.4100	2.4200	1.0000
20230063718	190.5155	59.4473	7.7200	16.9000	9.3700	1.0000
20230040464	204.1978	60.1713	8.7300	18.9500	4.4100	1.0000
20230040474	19.3800	60.1869	8.7700	12.7600	10.8900	0.8772
20230040692	38.1854	60.5939	8.8400	20.7300	22.7400	0.9423
20230040710	215.8008	60.6375	8.3400	3.5000	6.2400	1.0000
20230040774	118.0977	60.7453	8.8100	6.3600	6.2000	1.0000
20230040892	74.0915	60.9881	8.1900	20.8100	7.2200	1.0000
20230063937	200.3831	62.2472	8.8500	9.2800	27.1600	1.0000
20230063962	61.4593	62.5611	7.2100	20.4400	41.5100	1.0000
20230064006	21.9845	62.7867	8.9600	16.9900	6.7900	1.0000
20230064090	149.0296	63.2041	8.7200	19.8500	2.8300	1.0000
20230064113	83.2880	63.3466	8.8700	19.6400	4.9800	1.0000

# Bibliography

- [1] Alejandro Benitez-Llambay and Carlos Frenk. “The detailed structure and the onset of galaxy formation in low-mass gaseous dark matter haloes”. In: *arXiv* (2020).
- [2] Paul Torrey Mark Vogelsberger Federico Marinacci and Ewald Puchwein. “Cosmological Simulations of Galaxy Formation”. In: *arXiv* (2019).
- [3] Alejandro Benitez-Llambay and Julio F. Navarro. “The properties of “dark” CDM halos in the Local Group”. In: *arXiv* (2016).
- [4] Alejandro Benitez-Llambay and Julio F. Navarro. “Is a recently discovered HI cloud near M94 a starless dark matter halo?” In: *arXiv* (2023).
- [5] Carlos S. Frenk Jesús Zavala. “Dark matter haloes and subhaloes”. In: *arXiv* (2019).
- [6] Romeel Dave † Rachel S. Somerville. “Physical Models of Galaxy Formation in a Cosmological Framework”. In: *arXiv* (2014).
- [7] Chuan-Peng Zhang and FAST Collaboration. “The FAST All Sky H I Survey (FASHI): the first release of catalog”. In: *arXiv* (2023).