

Constraining Astrophysics using Line Intensity Mapping in the Post-reionization regime

M.Sc Thesis

By

Parth Hitesh Kothari

Supervisor: Dr. Suman Majumdar



DEPARTMENT OF ASTRONOMY, ASTROPHYSICS & SPACE ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY INDORE

May 20, 2025

Constraining Astrophysics using Line Intensity Mapping in the Post-reionization regime

A THESIS

*Submitted in partial fulfillment of the
requirements for the awards of the degree
of*
Master of Science

by

Parth Hitesh Kothari



**DEPARTMENT OF ASTRONOMY, ASTROPHYSICS & SPACE ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY INDORE**

May 20, 2025



INDIAN INSTITUTE OF TECHNOLOGY INDORE

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the thesis entitled **Constraining Astrophysics using Line Intensity Mapping in the Post-reionization regime** in the partial fulfillment of the requirements for the award of the degree of **MASTER OF SCIENCE** and submitted in the **DEPARTMENT OF Astronomy, Astrophysics and Space Engineering, Indian Institute of Technology Indore**, is an authentic record of my own work carried out during the time period from **June 2024 to May 2025** under the supervision of **Dr. Suman Majumdar, Associate Professor, Department of Astronomy, Astrophysics and Space Engineering**.

The matter presented in this thesis has not been submitted by me for the award of any other degree of this or any other institute.

Parth B
12/05/2025

Signature of the student with date
(Parth Hitesh Kothari)

This is to certify that the above statement made by the candidate is correct to the best of my/our knowledge.

Suman Majumdar
12/05/2025

Signature of the Supervisor of
M.Sc. thesis (with date)
Dr. SUMAN MAJUMDAR

Parth Hitesh Kothari has successfully given his/her M.Sc. Oral Examination held on **06/05/2025**.

Suman Majumdar

Signature(s) of Supervisor(s) of MSc thesis
Date: 12/05/2025

Manoueta Chakraborty

Convenor, DPGC
Date: 19/05/2025

Manoueta Chakraborty
Programme Coordinator, M.Sc.
Date: 19/05/2025

S. Das
HoD, DAASE
Date:

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my heartfelt gratitude to all those who have played a significant role—academically and morally—in the completion of this thesis.

First and foremost, I am deeply grateful to my supervisor, Dr. Suman Majumdar, for his invaluable guidance, unwavering support, and insightful feedback throughout the course of my research. His mentorship has been instrumental in shaping my academic journey and refining my approach to scientific problem-solving.

Research is never a solitary pursuit; it thrives on collective effort and diverse perspectives. I am fortunate to have been surrounded by a group of inspiring PhD seniors—Dr. Chandrashekar, Leon, Mohit, Yashraj, and Shiriny—whose constructive advice and timely support have guided me throughout this journey. Their encouragement and shared knowledge played a vital role in bringing this thesis to fruition.

I would also like to thank my labmates and friends, Prasad and Vishrut, with whom I shared both laughter and rigorous academic discussions. These exchanges often sparked insights that helped me overcome difficult problems. I extend my gratitude to my batchmates—Aryan, Anushka, Ashutosh, Daisy, Hari, Gitaj, Vijay, Devesh, Amardeep, Navnit, and Annie—for making my academic life memorable and filled with shenanigans and lighthearted moments.

I am deeply indebted to my parents and family, whose unwavering support and strength gave me the resolve to persevere through challenges. Finally, I would like to acknowledge my own resilience in pushing through difficult times and staying committed to completing this thesis.

— Parth Hitesh Kothari

ABSTRACT

Line Intensity Mapping (LIM) is an emerging technique in cosmology for quantifying clustering and strength of astrophysical sources at large length scales. In this method, instead of resolving each individual source of emission, one generates a coarse map of the total emission being observed in a broad resolution element. This helps us to detect emissions from very faint and distant sources, which are otherwise difficult to detect with the resolution and sensitivity of the presently operating telescopes. This technique thus presents a unique window for studying the large-scale structures by generating tomographic maps of the emission lines, tracking the evolution of the structures through different redshifts. Many surveys, using LIM, are ongoing in their initial phase, and many future surveys are planned to extend upon the instrumental techniques derived from them. Carbon monOxide Mapping Array Project (COMAP) and MeerKAT radio telescope are two such experiments targeting the CO(1-0) and HI 21-cm line transition, respectively, reporting upper limits at $z \approx 3$ and detections at $z \approx 0.32$ and 0.44 , respectively.

In this study, we analyze data from the COMAP and MeerKAT surveys to constrain the model parameters α , β , and neutral hydrogen density parameter, Ω_{HI} , which govern an empirical CO(1-0) emission model and a semi-numerical framework for HI 21-cm line emission. We employ a Bayesian framework with Markov Chain Monte Carlo (MCMC) sampling to explore the model parameter space. Our analysis shows that the current data is insufficient to constrain the parameters α and β , whereas we can obtain stringent constraints on Ω_{HI} , given a specific HI injection model for halos. At $z = 0.32$, we find $\Omega_{\text{HI}} = 6.7^{+0.1}_{-0.3} \times 10^{-4}$ when priors drawn from galaxy surveys are applied, and $\Omega_{\text{HI}} = 7.2^{+0.3}_{-0.3} \times 10^{-4}$ when no priors are imposed. Similarly, at $z = 0.44$, the parameter is constrained to $\Omega_{\text{HI}} = 10.2^{+0.1}_{-0.3} \times 10^{-4}$ with priors from radio observations, and $\Omega_{\text{HI}} = 10.9^{+0.3}_{-0.3} \times 10^{-4}$ without priors. We additionally perform a joint analysis combining data from both surveys; however, this does not yield significant constraints on the model parameters, underscoring the need for higher signal-to-noise ratio (SNR) observations and improved data quality for robust joint inference.

CONTENTS

List of Figures	iii
List of Tables	vi
1 Introduction	1
1.1 The Cosmic Microwave Background	1
1.2 Dark Ages and Cosmic Dawn	2
1.3 Epoch of Reionization	2
1.4 Post Epoch of Reionization	3
1.5 Motivation for the project	3
2 Line intensity Mapping	5
2.1 Introduction	5
2.2 Formalism	6
2.3 Emission lines	8
2.3.1 Singly Ionized Carbon [CII]	8
2.3.2 CO	8
2.3.3 Lyman α	9
2.3.4 HI 21 cm	9
2.4 LIM Surveys	10
2.4.1 CONCERTO	10
2.4.2 FYST	10
2.4.3 COMAP	11
2.4.4 SphereX	11
2.4.5 MeerKAT	12
2.5 Major obstacles for LIM	12

3	Methodology	14
3.1	Observations	14
3.1.1	Power Spectrum	14
3.1.2	COMAP Data	16
3.1.3	MeerKAT Data	17
3.2	Model	19
3.2.1	IllustrisTNG Simulations	19
3.2.2	Friends of Friends (FoF) Halo Finder	20
3.2.3	Line Intensity Model	20
3.2.3.1	CO model	21
3.2.3.2	HI 21-cm model	22
3.3	Inference	23
3.3.1	Parameter estimation via Bayesian Inference	23
3.3.2	Bayesian Framework	25
3.3.3	Outline of the Analysis	27
3.3.3.1	Independent Study	27
3.3.3.2	Joint Analysis	27
4	Results	28
4.1	Line Intensity Maps and Power Spectrum	28
4.1.1	CO Maps and COMAP Power spectrum	28
4.1.2	HI 21-cm Maps and MeerKAT Power Spectrum	29
4.2	Exploration of CO and HI Model Parameter Space	32
4.2.1	Constraints on α and β	32
4.2.2	Constraints on Ω_{HI}	33
4.3	Parameter Estimation using Bayesian Inference	34
4.3.1	COMAP Analysis	34
4.3.2	MeerKAT Analysis	34
4.3.2.1	$z = 0.33$	34
4.3.2.2	$z = 0.44$	36
4.4	Joint Analysis of COMAP and MeerKAT	38
5	Summary and Future Plan	40

LIST OF FIGURES

1.1	Universe's evolution along Redshift (Credit: NAOJ).	1
3.1	<i>Top row:</i> Power spectrum measurements for the eight k -bins targeted by COMAP across the three fields. <i>Bottom row:</i> Signal-to-noise ratio for respective bins. Image credit: Stutzer et al. 2024 [15].	16
3.2	MeerKAT power spectra at redshifts $z = 0.32$ and $z = 0.44$ of the observed field. Image credit: Paul et al. 2023[2].	17
3.3	Flowchart of Markov Chain Monte Carlo Framework for parameter estimation using emcee (Foreman Mackey et al. 2012) [37].	26
4.1	<i>Left:</i> Halo overdensity map of a slice of the TNG300-1 box. <i>Right:</i> The corresponding CO intensity map created with the LIM simulator using the empirical model based on Li et al. 2016 [3]. Both fields have a grid resolution of 0.56 Mpc.	28
4.2	COMAP data points having high SNR considered for the Bayesian analysis. As noted in Section 3.1.2, the uncertainty associated with the second k -bin is smaller than that of the other three data points, with its lower bound extending into negative values.	29
4.3	<i>Left:</i> Same realization of the CO intensity map as figure 4.1. <i>Right:</i> The power spectrum is estimated for the k -bins being considered for the analysis and is computed as in Murmu et al. 2021[19].	29
4.4	<i>Left Column:</i> Halo overdensity maps derived from the TNG300-1 simulation at $z = 0.33$ and $z = 0.44$; <i>Right Column:</i> 21-cm maps generated using scaling relations of Bagla et al. 2010[4], followed by conversion to 21-cm brightness temperature map using the Villaescusa-Navarro et al. 2018 scheme[36]. Both are coarse-gridded to a resolution of 0.42 Mpc, corresponding to the MeerKAT detections at these redshifts.	30

4.5	HI 21-cm power spectrum detected by the MeerKAT telescope in interferometric mode. The left and right power spectrum plots show the data points with $\text{SNR} > 2.0$ chosen for the analysis, at $z = 0.32$ and $z = 0.44$, respectively.	31
4.6	<i>Left:</i> HI 21-cm intensity maps (same as those shown in Figure 4.4). <i>Right:</i> Power spectrum computed for the chosen k -bins from the corresponding intensity maps.	31
4.7	<i>Left:</i> Values of α and β derived from empirical fits to galaxy surveys, with 1σ uncertainties, for the line emission model. The Tabulated values are shown in Table 4.1. <i>Right:</i> Theoretical Power spectrum plotted with the COMAP observation; the power spectrum is calculated from the simulated map, which takes the mean values of α and β from the three surveys as input. .	32
4.8	Values of Ω_{HI} reported by Rao et al. 2017 [41] and Rhee et al. 2018 [42].	33
4.9	<i>Left and Right</i> figures show the model power spectra corresponding the the Ω_{HI} estimates of Rao et al. 2017 [41] and Rhee et al. 2018 [42] plotted in comparison to the MeerKAT data for $z = 0.33$ and $z = 0.44$, respectively. The power spectra are computed for the mean and edges of the parameter uncertainty to visualize how they vary with Ω_{HI}	34
4.10	<i>Left:</i> MCMC posteriors for 3000 steps with four walkers after discarding 1000 points and a thinning factor 25. No constraints are visible over the two parameters. <i>Right:</i> MCMC posteriors for 6000 steps with four walkers after discarding 840 points and a thinning factor 35. Results remain unchanged despite the increase in walker lengths.	35
4.11	<i>Left:</i> Probability distribution of Ω_{HI} for 6000 steps with three walkers initialized at the edge of parameter space. Convergence of Ω_{HI} is visible at the edge of the parameter space described in table 4.2, the values being $\Omega_{\text{HI}} = 6.7^{+0.1}_{-0.3} \times 10^{-4}$ and the distribution peaking at 6.8×10^{-4} . <i>Right:</i> The power spectrum for $\Omega_{\text{HI}} = 6.8 \times 10^{-4}$	35
4.12	<i>Left:</i> Figure shows the results of the second MCMC run in which the prior is updated as $0 \leq \Omega_{\text{HI}}$ with no upper bounds. The probability distribution shows the convergence at $\Omega_{\text{HI}} = 7.2^{+0.3}_{-0.3} \times 10^{-4}$. <i>Right:</i> The power spectrum for the $\Omega_{\text{HI}} = 7.2 \times 10^{-4}$	36
4.13	<i>Left:</i> Marginalized probability distribution of Ω_{HI} at $z = 0.44$ from an MCMC run with 6000 steps and three walkers initialized near the boundary of the parameter space. The distribution peaks at 10.2×10^{-4} , with a best-fit value of $\Omega_{\text{HI}} = 10.2^{+0.1}_{-0.2} \times 10^{-4}$, consistent with the bounds given in Table 4.2. <i>Right:</i> Power spectrum corresponding to $\Omega_{\text{HI}} = 10.2 \times 10^{-4}$	37
4.14	<i>Left:</i> Posterior distribution of Ω_{HI} at $z = 0.44$ from a second MCMC run using a modified prior: $0 \leq \Omega_{\text{HI}}$, without an upper bound. The parameter converges to $\Omega_{\text{HI}} = 10.9^{+0.3}_{-0.3} \times 10^{-4}$. <i>Right:</i> Corresponding power spectrum for $\Omega_{\text{HI}} = 10.9 \times 10^{-4}$	37

4.15	The results are validated by comparing the constraints on Ω_{HI} with those obtained from galaxy surveys and radio observations. The results show tighter constraints on the neutral hydrogen density parameter, which are consistent with the findings of Rao et al. 2017 [41] and Rhee et al. 2018 [42]. This work highlights the potential of combining Line Intensity Mapping with other observational probes to substantially enhance our understanding of both astrophysical processes and cosmological parameters in the future.	38
4.16	Joint Analysis of COMAP and MeerKAT Surveys: We performed an MCMC run with 6000 steps and 4 walkers, exploring the parameter space by varying α and Ω_{HI} , while fixing $\beta = 2.00$ as suggested by Greve et al. 2014 [40]. The chain did not converge even after a runtime of approximately 168 hours. We discarded the first 1000 steps as burn-in and applied a thinning factor of 50. The results yielded nonphysical parameter values, which is expected due to the lack of convergence. These findings suggest that a joint analysis is not feasible with the current data from the COMAP and MeerKAT surveys.	39

LIST OF TABLES

3.1	Observation for the specific central k bin, measured spherical power spectrum and their respective uncertainties. The bins are equispaced in log scale with the individual bin width, $\Delta(\log k \text{ Mpc}^{-1}) = 0.155$	17
3.2	Observed power spectrum measurements at redshifts $z = 0.32$ and $z = 0.44$	18
3.3	Specifications of the TNG 300-1 box used for the simulation.	19
4.1	Best-fit values of α and β from different studies.	32
4.2	Reported estimates of neutral hydrogen density parameter by galaxy and radio observations in the redshift range.	33

CHAPTER 1

INTRODUCTION

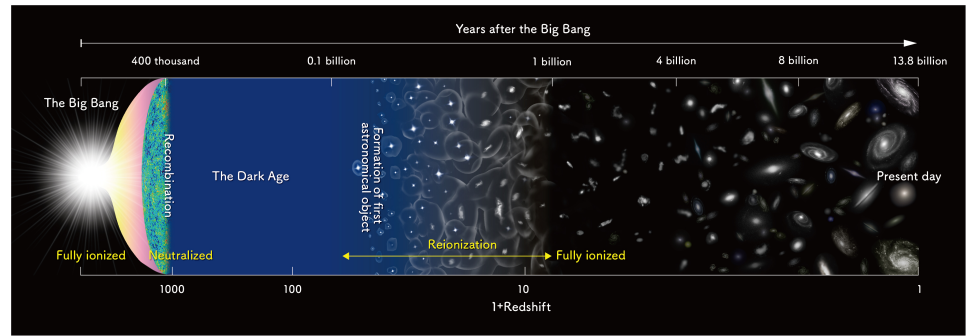


Figure 1.1: Universe's evolution along Redshift (Credit: NAOJ).

The main objective of the modern cosmology era has been to uncover and understand the history of the Universe, followed by the evolution in the various epochs observed in the present era. The Hot Big bang or more precisely the Λ -Cold Dark Matter (CDM), is the current accepted model explaining the formation of structures, expansion of the universe and the detection of the Cosmic Microwave Background. Understanding the present-day Universe necessitates an understanding of its origin and the various stages it has traversed to get to its current state.

1.1 The Cosmic Microwave Background

Matter and radiation decoupled as a result of the recombination of free electrons and protons. After decoupling from matter, radiation became free streaming, achieving a mean free path greater than the Hubble Horizon. This is called the epoch of the last scattering, where the Universe is in a neutral

state composed of neutral hydrogen and free-streaming radiation. The wavelength of the radiation increased due to the expansion of the Universe as the photons lost energy to travel the expanding space. Due to this phenomenon, the radiation is shifted to the microwave end of the electromagnetic spectrum. The radiation was termed as Cosmic Microwave Background as the Intensity was uniform coming from every direction when it was first detected in the late 1960s.

1.2 Dark Ages and Cosmic Dawn

The formation of the first atoms leads to the decoupling of matter-radiation. Hydrogen and Helium were the first atoms formed, with hydrogen dominating the helium population. Diffuse hydrogen permeating the space can be traced via the 21-cm transition even though no stars are present in this age. Luminous sources were yet to be produced; hence, this era was devoid of light, called the Dark Ages. The evolution of perturbations in the dark matter field led to slightly under dense and over dense regions, which became eventual seeds for the dark matter potential well. These regions accumulated more particles over time, creating a distinct separation between over-dense and under-dense regions. Over-dense regions continued to grow, leading to the creation of dark matter halos, which caused baryonic matter to start to fall into the halo. Once a certain threshold for mass was reached and the temperature was sufficiently low, the first stars formed. These were the first-generation stars, also known as Population III (Pop III) stars, illuminating the dark Universe after a long period. The formation of the stars and release of the first light is termed the cosmic dawn.

1.3 Epoch of Reionization

Cosmic Dawn spawned a cascade of structure formation ranging from star formation to the galaxies residing in the halos. Initially, the Pop III stars went supernova, enriching the Inter Stellar Medium (ISM) inside the halos, producing different atoms other than neutral hydrogen(HI) and Helium, which would act as coolants to induce and increase the star formation rate. The

radiation from stars started to ionize the HI residing in the halos; however, recombination of the ionized hydrogen (HII) due to a high density kept the ionizing radiation inside the halos. The increase in star formation led to a high number of ionizing photons, eventually escaping the HI present in the Inter Galactic Medium (IGM). The density of HI in IGM was not large enough to ensure the recombination rate remains at par with the reionization rate. Increased rate of reionization led to ionization of the IGM. This period is called the Epoch of Reionization (EoR).

1.4 Post Epoch of Reionization

The entire IGM was ionized by $z \sim 5$, reionization being the last major phase transition for the Universe. In this era, the matter is concentrated in the halos surrounding a galaxy or cluster of galaxies. The astrophysics inside these cupolas plays a major role in the evolution of the Universe in its present state. Different emissions are present in this epoch as the later generation stars, also known as Population I stars (Pop I), combined with different atoms and molecules, permeate the medium, paving a way to probe the properties of the Universe through complementary surveys.

1.5 Motivation for the project

Line Intensity Mapping (LIM) is a technique that focuses on observing specific line emissions in coarse resolution to map spatial fluctuations of the signal. LIM is able to probe the Universe from the present time to the cosmic dawn and further in the past. Multiple LIM surveys are planned to target different eras of the Universe, and many are underway. Carbon monOxide Mapping Array Project (COMAP) is a LIM survey targeting rotational CO(1-0) lines from Post-EoR at $z = 2.4-3.4$, CO(2-1) from the EoR observing of the signal to put a constraint on the power spectrum (Cleary et al. 2022[1]). MeerKAT telescope has demonstrated the use of LIM at near redshift for targeting HI 21 cm emission and has confirmed detection of power spectrum for HI at redshifts $z \sim 0.32$ and 0.44 (Paul et al. 2023 [2]). In this project, we explore both the CO emission model (Li et al. 2016 [3]) and the HI 21-

cm model (Bagla et al. 2010[4]) to investigate whether tighter constraints could be placed on the CO model parameters, α and β , using the CO power spectrum upper limits from COMAP, and on the cosmic neutral hydrogen content Ω_{HI} using MeerKAT observations. Building on this, we attempt a joint analysis of the surveys—each targeting different line emissions—to assess whether such a multi-tracer approach can yield improved constraints on the astrophysical and cosmological parameters.

CHAPTER 2

LINE INTENSITY MAPPING

2.1 Introduction

Studying the Universe on the largest scale has posed many difficulties in the past, including but not limited to the need for bigger telescopes and larger landmark areas for setting up the telescopes. Efforts have been made to launch multiple space-based telescopes that would avoid atmospheric distortions of signals and make a comprehensive data set to be studied. A multifold increase in difficulties is posed for studying the evolution of the Universe as the sources of light and the structures change with redshift as we probe deeper into the Universe. Resolving individual galaxies at high redshift poses huge problems, as massive apertures are needed. Astrophysical obstacles are one of the factors in observing the galaxies. The emissions from sources are absorbed by the intervening galactic medium, making the telescopes blind to those particular emissions. Different emissions require particular instruments, increasing the cost of the mission, combined with the difficulty of studying the sources as the look back time increases.

Probing the Universe at high redshift requires a telescope with sophisticated instruments, such as the James Webb Space Telescope (JWST), as the expansion of the universe stretches the wavelength of the photon, resulting in the detection in the low-frequency range of the spectrum. Resolving each individual source leads to an increase in observation time, yet the survey will be limited by the magnitude it can observe. The brightest sources in the field will be visible, but they may not capture the overall matter distribution, which is required to study the underlying cosmology. Increased time

and expensive structures are some of the essential drawbacks of traditional telescopes.

Line Intensity Mapping (LIM) is a unique technique to study large-scale structures by targeting specific line emissions (Kovetz et al. 2017[5]). LIM detects spectral line emission without isolating individual sources, instead aggregating photons within a broad-resolution element. This method allows for mapping the distribution of sources with lower spatial resolution combined with the high-frequency resolution needed along the line of sight, compared to the traditional surveys over large fields of view. The technique involves creating a tomographic cosmological cube where the fluctuation of line intensity on the sky plane is mapped in coarse resolution elements and the fluctuations along the line of sight are contained in different frequency channels, effectively generating a three-dimensional map of a specific line emission. The major advantage of this technique is the reduction of time for observation, leading to bigger area coverage. The idea that individual sources do not have to be resolved contributes to decreased telescope costs. LIM can essentially follow the evolution of the Universe by tracking emissions resulting from varied astrophysical processes traced to galaxy formation and the evolution in subsequent redshift. This dynamical dependence of cosmology and astrophysics embedded in the technique effectively allows us to leverage it in this era of precision cosmology.

2.2 Formalism

Line intensity Mapping, in essence, measures the fluctuation in intensity or the mean intensity of the specific emission targeted. A multitude of luminous sources are present in a voxel with different luminosities. A voxel is a 3-D resolution element in the sky where the area of the face is the sky plane, and the width of the frequency channel defines the line of sight information. Thus, we can write the specific luminosity density within the voxel as

$$\rho_L = \frac{\sum_i L_i}{\Delta V}$$

where L_i = Luminosity of the i th source in the voxel and ΔV = Volume of the voxel. The quantity will have a flux volume density

$$f_L = \frac{\rho_L}{4\pi D_L^2}$$

associated with it, D_L being the luminosity distance.

Decomposing the voxel volume element into sky coordinates and the resolution of the frequency composed by the survey instrument, we get the specific intensity

$$I = f_L \frac{dA}{d\Omega} \frac{d\chi}{dv_{obs}}$$

where

- dA = area of the voxel on the sky
- $d\Omega$ = solid angle subtended by the voxel to the observer
- $d\chi$ = comoving distance to the voxel
- dv = Observed frequency interval

$dA/d\Omega$ can be further decomposed as D_A^2 , angular diameter distance and $d\chi/dv$ as $c(1+z)/H(z)v_{obs}$. The specific intensity can be rewritten as

$$I(z) = \frac{c}{4\pi\nu H(z)} \rho_L(z)$$

In the regime of Rayleigh-Jeans, we can relate the intensity to the brightness temperature as

$$T(z) = \frac{c^3(1+z)^2 \rho_L(z)}{8\pi k_B \nu^3 H(z)}$$

where ν is defined as the rest-frame frequency of the line emission. Mean luminosity density can be estimated as

$$\begin{aligned} \langle I(z) \rangle &= K \langle \rho_L(z) \rangle \\ &= K \int L \frac{dn}{dL}(L, z) dL \\ &= K \int dM L(M, z) \frac{dn}{dM}(M, z) \end{aligned}$$

where $K = c/4\pi\nu H(z)$. The expression $K \int L \frac{dn}{dL}(L, z) dL$ describes the mean intensity $\langle I(z) \rangle$ as an integral over the line luminosity function, while $K \int dM L(M, z) \frac{dn}{dM}(M, z)$ represents $\langle I(z) \rangle$ as an integral over the halo mass function.

2.3 Emission lines

2.3.1 Singly Ionized Carbon [CII]

[CII] emission is one of the brightest tracers ($158\mu\text{m}$) to the cold ISM where star formation occurs. [CII] origins are varied in nature ranging from Photo Disassociation regions (PDRs), HII regions, cold atomic gas and dark CO clouds (Silva et al. 2015 [6]). The main source of [CII] is PDR, which is found at the boundaries of HII regions. PDR are warm and dense regions of the ISM located between HII regions and molecular clouds. PDRs are mostly composed of neutral gas, yet, being close to the star-forming region, they are mostly governed by the UV radiation emitted by the stars. Due to UV penetrating the PDR, the dust grains absorb this radiation and re-emit it, warming the gas. [CII] line transition acts as a coolant for the region by emitting radiation and thereby decreasing the temperature of its surroundings, hence aiding star-formation. Other places where [CII] can be found are the HII regions, as [CII] has a lower ionizing potential than HII, resulting in high emission of the [CII] lines.

2.3.2 CO

CO molecule is the most abundant molecule found in the galaxies observed at low redshifts after H_2 (Carilli et al. 2013 [7]). H_2 is the preferred rotational line for tracing the cool molecular content of the region, but due to the lack of a permanent dipole, the lowest vibrational state is forbidden, requiring temperatures higher than 500 K, which are not available in the cold medium of the ISM. Contrary to this, CO's lowest vibrational state requires a temperature of 5K, achievable in the cold medium. The rotational transition follows a ladder of frequencies: $\nu_{J \rightarrow J-1} = J \times 115.27 \text{ GHz}$. The transitions are among the brightest that can be observed in galaxy spectra and

can easily be observed from a ground-based telescope. CO is estimated to be an effective tracer for cold molecular content in the galaxy, the preceding stage after which the gas collapses and star formation happens. This means that CO can probe the galaxy's stellar content and star formation rate.

2.3.3 Lyman α

The most prominent emission to be targeted in the UV spectrum would be the Ly- α . It is the result of the UV photons produced by stars traversing through the ISM, ionizing the HI in the vicinity (Bernall et al. 2022 [8]). The recombination of the ionized hydrogen results in the emission of Ly- α . The wavelength of this emission is 1216 Å, falling in the absorption range of HI situated around the star-forming region, increasing the HI 21 cm emission from those neutral regions. Thus, Ly- α emission is a probe for active star-forming regions, and its absence shows a sign of HI in the surroundings.

2.3.4 HI 21 cm

21 cm emission results from the transition in the ground state of the hydrogen atom. The magnetic moment of electrons and protons leads to the formation of a hyperfine structure in the ground state of atomic hydrogen. The flipping of spin results in the emission of the 21 cm line. This emission is an ideal probe for studying the IGM as it permeates the whole region. 21 cm emission is driven by processes that vary across cosmic time. During the early stages, it is driven by the coupling of Cosmic Microwave Background (CMB) photons to the HI field, dominating the collision excitation. After the cosmic dawn with the first stars being formed, Ly- α radiation couples the neutral hydrogen to the kinetic temperature of the gas permeating. This mechanism is termed the Wouthuysen-Field effect (Wouthuysen 1952[9], Field 1958[10]). In the Post-EoR regime, all the HI is concentrated in the galaxies as the HI permeating the IGM is ionized. The HI survives in the ISM as a closely packed gas cloud. The UV from the stars constantly ionizes the gas, but due to proximity, they recombine quickly, making sure they are present in the region.

2.4 LIM Surveys

2.4.1 CONCERTO

The CONCERTO (CarbON [CII] line in post-rEionization and ReionizaTiOn epoch) instrument, installed in the APEX telescope located at 5105 m altitude on the Llano de Chajnantor in Northern Chile (CONCERTO Collaboration et al. 2020 [11]). CONCERTO's science goal is to find out the role of star-forming regions in the early galaxy evolution, trying to probe the history of metal enrichment and how these star-forming galaxies have contributed to cosmic reionization. The mission will measure the [CII] line emission at redshifts $4.5 < z < 8.5$ by making 3D tomographic cubes of Intensity fluctuations, effectively probing the reionization and Post-EoR era of the cosmic evolution. CONCERTO will be sensitive to CO intensity fluctuations at the redshift range $0.3 < z < 2$, tracing the volume of molecular gas and its progression across this time frame.

2.4.2 FYST

Fred Young Submillimeter Telescope (FYST)(CCAT-Prime Collaboration et al. 2023[12]), a 6-meter aperture telescope working in the submillimeter range designed for a wide range of views, is located at a height of 5600 meters on Cerro Chajnantor in northern Chile. The operational frequency range of FYST is 100 GHz to 900 GHz, translating to wavelengths of 3 to 0.33 millimeters. The goal of FYST lies in the broader range of astrophysics and cosmology, which includes line intensity mapping of [CII] from Post-EoR to EoR. EoR Spec DSS survey on the FYST is designed to observe and capture tomographic maps of [CII] line from the faint star-forming regions in the cold regions of galaxy in the range of $z \sim 3.5 - 8.05$. The proposed plan for the IM observation is scanning more than 8 deg^2 sky plane in at least two target fields where the deep multi-wavelength survey will be performed. It is expected that full survey will provide enough data to obtain a claim for the detection of a clustering signal of [CII] in the $z \sim 5-8$. The broad frequency range has the ability to observe rotational CO lines falling in the survey range for [CII]. CO(2-1) transition and above are the emission that

can be targeted by FYST, the brightest in the $z \sim 0 - 2$, enabling FYST to map the galaxies from cosmic noon to the present day.

2.4.3 COMAP

The Carbon monOxide Mapping Array Project is a survey aimed at studying the CO emissions from the nascent stage of the universe when the first stars were being formed. The first stage of the survey, COMAP Pathfinder, is a spectrometer consisting of a single polarization 19 feed-array working in the 26-36 GHz range. The 19 feeds are placed on a single 10-meter dish at the Owens Valley Radio Observatory (Lamb et al. 2022[13]). Redshift for CO (1-0) transition will be $z = 2.4 - 3.4$ which will be tracing the epoch of galaxy assembly. In the early science phase of the pathfinder for the 13 months of observations (Ihle et al. 2022[14]), they reported an upper limit for CO power spectrum over the range $k = 0.051 - 0.62$ (Mpc) giving the $P(k) = -2.7 \pm 1.7 \times 10^4 \mu\text{K}^2 \text{Mpc}^3$. In the 5-year-long survey time of the Pathfinder, they have reported an improved upper limit on the power spectrum (Stutzer et al. 2024 [15]).

2.4.4 SphereX

The Spectro-Photometer for the History of the Universe, Epoch of Reionization (EoR), and Ices Explorer (SPHEREx) is a space-based telescope under NASA that is designed to operate for 2 years conducting a 4 full-sky survey in the optical and near-infrared spectrum (Doré et al. 2014[16]). The scientific goal of the telescope is to probe the origin and the history of how galaxies are formed by doing a full spectroscopic deep survey, paving the way to generate intensity maps of the large-scale structure at near-infrared wavelength. It will investigate the inflation history of the universe and progression of the galaxies from the EoR period. SPHEREx will also make a catalogue of spectral imaging over the entire scan. SPHEREx at low redshifts will be able to detect multiple line emissions such as $\text{H}\alpha$ between redshift range $0.1 < z < 5$, $\text{H}\beta$ between $0.5 < z < 2$ and $[\text{OIII}]$ between $0.5 < z < 3$. In the redshift range $5.2 < z < 8$, it will be able to detect the Lyman- α line, an important tracer of star-forming regions in the EoR era.

2.4.5 MeerKAT

MeerKAT is a 64-dish telescope array having a single dish of 13.5 m situated in South Africa and is a precursor to the SKA mission. MeerKAT consists of two frequency bands that cover the redshift range from $0 < z < 1.4$. It will use the intensity mapping technique to probe the local Universe at large scales and observe the HI 21-cm signal. The first results from employing LIM are presented in Paul et al. 2023[2], confirming the detection of HI 21-cm power spectrum in the local Universe.

2.5 Major obstacles for LIM

Multi-line analysis of LIM provides a ground for flexible study of the target properties. LIM surveys can be utilized to cross-correlate with galaxy surveys, enabling vigorous study of the astrophysical process coupled with cosmological scales. The drawback of the immense potential is the contamination of the observations via Interlopers and foregrounds present in the line of sight.

Continuum foregrounds constitute synchrotron radiation or thermal emissions from the galaxy dust, which is a continuous spectrum. The radiation dominates in the infrared region, making a Cosmic Infrared Background (CIB) for the LIM target (Bernal et al. 2022[8]). Removal of the CIB does not eliminate the continuum emissions from galaxies present in the foreground, but these are mostly contained in the shortest line of sight Fourier modes. Foregrounds can easily be eliminated by cross-correlating LIM with galaxy surveys, resulting in the cancellation of the uncorrelated noise, foregrounds and other systematics, and the remainder of the emissions are left over from the targets that LIM traces.

The bigger concern for the LIM survey is the contamination of the signal by interlopers. Interlopers are the spectral lines residing in a different redshift, which are redshifted into the observation frequency channel, resulting in the addition of the line into the target emission (Visbal and Loeb 2010[17]). Considering an example of [CII] emission, we are targeting [CII] from a redshift, $z = 7$, having an observed frequency of $\nu = 238$ GHz. The interloper for this redshift emission would be from rotational CO transition.

The CO (3-2) from $z = 0.45$, CO(4-3) from $z = 0.88$, CO(5-4) from $z = 1.4$, and CO(6-5) from $z = 1.8$ are the emissions overlapping with the frequency range of [CII]. The overlap will have an effect on the integrated emission as the CO signal can dominate the [CII] signal, concluding false evidence for [CII] clustering (Lidz and Taylor 2016[18], Silva et al. 2015[6]). Removing interloper signal strength would involve cross-correlating the map with another map or galaxy survey to remove the uncorrelated signal, confirming the detection of the target emission.

CHAPTER 3

METHODOLOGY

3.1 Observations

3.1.1 Power Spectrum

The Power spectrum is a statistic that calculates the fluctuations in a field at different length scales, quantifying the variance of the field with respect to the mean at that scale in the Fourier space. The power spectrum is defined as the Fourier transform of the two-point correlation function in real space. Starting with the basic formalism, density contrast at a point in a field is defined as:

$$\delta(\mathbf{x}, t) = \frac{\rho(\mathbf{x}, t) - \bar{\rho}(t)}{\bar{\rho}(t)}$$

where

- $\rho(\mathbf{x}, t)$ is density at the point in time t , $\mathbf{x}$ being a vector
- $\bar{\rho}(t)$ mean of the field at time t

The two-point correlation function is then defined as

$$\xi(\mathbf{x}, \mathbf{y}, t) = \langle \delta(\mathbf{x}, t) \delta(\mathbf{y}, t) \rangle$$

$\langle \delta(\mathbf{x}, t) \delta(\mathbf{y}, t) \rangle$ is the average of the density contrast at two different points in space at a given time quantifying the correlation between all the points considered in the ensemble. In the context of cosmology, the function will change as the cosmological principle assumes statistical homogeneity and isotropy. The function will only depend on the absolute distance between

the two points considered in the iteration, effectively changing the functional form.

$$\xi(|\mathbf{x} - \mathbf{y}|, t) = \langle \delta(\mathbf{x}, t) \delta(\mathbf{y}, t) \rangle$$

Taking the Fourier transform of the density contrast yields the form as

$$\Delta(k, t) = \int d^3x e^{i\mathbf{k} \cdot \mathbf{x}} \delta(\mathbf{x}, t)$$

Applying this to our correlation function yields

$$\begin{aligned} \langle \Delta(\mathbf{k}, t) \Delta(\mathbf{k}', t) \rangle &= \int d^3x d^3y e^{i\mathbf{k} \cdot \mathbf{x} + i\mathbf{k}' \cdot \mathbf{y}} \langle \delta(\mathbf{x}, t) \delta(\mathbf{y}, t) \rangle \\ &= \int d^3x d^3y e^{i\mathbf{k} \cdot \mathbf{x} + i\mathbf{k}' \cdot \mathbf{y}} \xi(r, t) \\ &= \int d^3r d^3y e^{i\mathbf{k} \cdot \mathbf{r} + i(\mathbf{k} + \mathbf{k}') \cdot \mathbf{y}} \xi(r, t) \\ &= (2\pi)^3 \delta_D^3(\mathbf{k} + \mathbf{k}') \int d^3r e^{i\mathbf{k} \cdot \mathbf{r}} \xi(r, t). \end{aligned}$$

where we have used $\mathbf{r} = \mathbf{x} - \mathbf{y}$, $\delta_D^3(\mathbf{k} + \mathbf{k}') = \text{Dirac delta function}$. The power spectrum is then defined as

$$P(k) = \int d^3r e^{i\mathbf{k} \cdot \mathbf{r}} \xi(r, t)$$

OR

$$\xi(r) = \int \frac{d^3k}{(2\pi)^3} P(k) e^{-i\mathbf{k} \cdot \mathbf{r}}$$

Assuming the field is Gaussian, the power spectrum captures all the information present by quantifying the fluctuations. Power spectrum statistic is widely used in the field of astronomy, and observations are provided for this particular statistic.

Following the methodology of Murmu et al. 2021 [19], we compute the power spectrum by first taking the Fourier transform of the simulated field and then calculating the power spectrum as the squared amplitude of the resulting Fourier modes.

3.1.2 COMAP Data

COMAP recently finished the second season survey lasting for 5 years (Lunde et al. 2024[20], Stutzer et al. 2024 [15], Chung et al. 2024[21]). The outcome of the survey gave insights into new and improved upper limits on the CO (1-0) spherically averaged power spectrum in the redshift $z \approx 3$ (Stutzer et al. 2024 [15]). Three fields were chosen for the scans, each having a field of view around 2 - 3 deg² and the combined observations lasting for 17500 hours. The Fiducial cosmology incorporated for translating frequencies and sky plane separations from the maps observed into cosmological distances is $\Omega_m = 0.286$, $\Omega_\Lambda = 0.714$, $\Omega_b = 0.047$, $H_0 = 100 h \text{ km s}^{-1} \text{ Mpc}^{-1}$ with $h = 0.7$, $\sigma_8 = 0.82$, and $n_s = 0.96$, consistent with WMAP (Hinshaw 2013 [22]).

The voxel size is estimated to be $2 \times 2 \text{ arcmin}^2$ which translates to a comoving cosmological volume of around $3.7 \times 3.7 \times 4.1 \text{ Mpc}^3$. Comoving volume for the three fields approximated at $150 \times 150 \times 1000 \text{ Mpc}^3$.

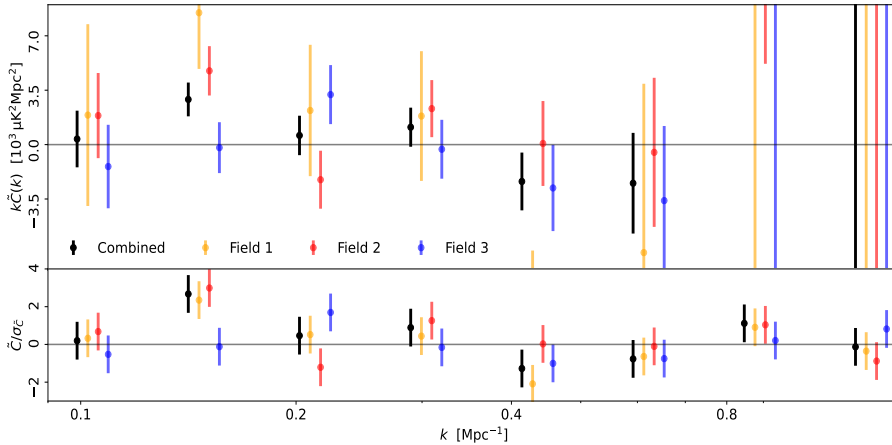


Figure 3.1: *Top row:* Power spectrum measurements for the eight k -bins targeted by COMAP across the three fields. *Bottom row:* Signal-to-noise ratio for respective bins. Image credit: Stutzer et al. 2024 [15].

Figure 3.1 shows the power spectrum measurement of a recently completed survey and the errors associated with each data point. The paper emphasized that the first six bins are the most sensitive bins of the survey, with the second bin giving an optimistic result of around 2.7σ above zero. The power spectrum limits and respective uncertainties are listed below in the Table 3.1.

k [Mpc^{-1}] (center)	$k\tilde{C}(k)$ [$10^3 \mu\text{K}^2 \text{Mpc}^2$]	$k\sigma_{\tilde{C}(k)}$ [$10^3 \mu\text{K}^2 \text{Mpc}^2$]
0.1	0.36	1.82
0.15	2.9	1.09
0.21	0.59	1.27
0.3	1.19	1.26
0.44	-2.37	1.86
0.62	-2.48	3.24
0.89	101.5	90.9
1.27	-5.05×10^5	3.9×10^6

Table 3.1: Observation for the specific central k bin, measured spherical power spectrum and their respective uncertainties. The bins are equispaced in log scale with the individual bin width, $\Delta(\log k \text{ Mpc}^{-1}) = 0.155$.

Apart from the continuum foregrounds as discussed in 2.5, interlopers are another contributor to the LIM observation. The current operating frequency of COMAP has the contribution from the targeted signal CO(1-0) at $z = 2.4\text{--}3.4$ and the interloper contribution from CO(2-1) transition at $z = 6\text{--}8$ (Foss et al. 2022)[23]. Hence, it is important to note that techniques are employed to mitigate the foreground and systematic noise; still, the data can possess the clustering signal from CO(2-1) as well.

3.1.3 MeerKAT Data

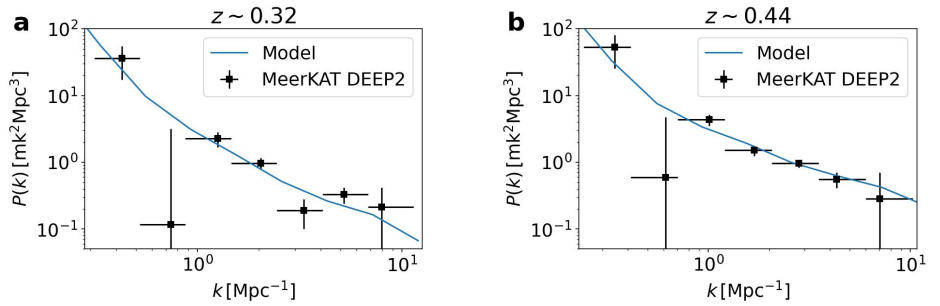


Figure 3.2: MeerKAT power spectra at redshifts $z = 0.32$ and $z = 0.44$ of the observed field. Image credit: Paul et al. 2023[2].

MeerKAT, the precursor to the SKA observatory, using the Line Intensity Mapping technique, has detected the HI intensity power spectrum at two different redshifts $z \approx 0.44$ and $z \approx 0.32$ corresponding to the central frequency channel 986 MHz and 1077.5 MHz with a width of 46MHz used

in the survey (Paul et al. 2023[2]). The study targeted the J2000 field, observing it for 96 hours over a 2 deg^2 field of view. Taking into account the cosmology used to calculate the sky and frequency plane size, $\Omega_m = 0.311$, $\Omega_b = 0.049$, $H_0 = 100 h \text{ km s}^{-1} \text{ Mpc}^{-1}$ with $h = 0.677$, $\sigma_8 = 0.8102$, and $n_s = 0.967$ (Planck Collaboration et al. 2018 [24]), the average comoving cosmological volume is $40 \times 40 \times 300 \text{ Mpc}^3$ for both redshifts.

The power spectrum is shown in Figure 3.2. There is a significant drop in the second k -mode of each power spectrum, which is also reflected in their signal-to-noise ratio. This drop has been credited for systematic errors in the shorter baseline configurations.

$z = 0.32$			
$k \text{ [Mpc}^{-1}]$	$P(k) \text{ [mK}^2 \text{ Mpc}^3]$	$\sigma_P \text{ [mK}^2 \text{ Mpc}^3]$	P/σ_P
0.43	36.48	19.03	1.92
0.74	0.12	3.04	0.04
1.25	2.26	0.57	3.98
2.04	0.96	0.19	4.96
3.30	0.19	0.09	2.12
5.19	0.33	0.09	3.80
7.96	0.21	0.20	1.07
$z = 0.44$			
$k \text{ [Mpc}^{-1}]$	$P(k) \text{ [mK}^2 \text{ Mpc}^3]$	$\sigma_P \text{ [mK}^2 \text{ Mpc}^3]$	P/σ_P
0.34	53.43	27.80	1.92
0.61	0.60	4.13	0.14
1.01	4.34	0.85	5.12
1.68	1.51	0.27	5.62
2.81	0.96	0.13	7.54
4.31	0.56	0.15	3.80
7.04	0.28	0.41	0.69

Table 3.2: Observed power spectrum measurements at redshifts $z = 0.32$ and $z = 0.44$.

Table 3.2 shows the central k -value, power spectrum measurement, uncertainties and the signal-to-noise ratio for the two detections.

3.2 Model

3.2.1 IllustrisTNG Simulations

The comoving cosmological box of the observation is asymmetric, with the frequency (redshift) axis for COMAP elongating up to 1000 Mpc and for MeerKAT up to 300 Mpc compared to sky plane sizes of around 150 Mpc and 40 Mpc, respectively.

The simulation box we opt for should contain the k -modes with sufficient samples in each bin to match the observed data. For this particular approach, the IllustrisTNG simulation presents a comparable simulation box. IllustrisTNG ([25], [26], [27], [28], [29]) is a collaborative project aimed at performing cosmological gravo-magnetohydrodynamical simulations and providing coeval boxes of approximate volume 50, 100, and 300 Mpc³. Throughout the run, 100 snapshots are taken from the different epochs of the universe for each simulation box. Two types of N-body simulations are initialized, one with Dark matter (DM) particles only and one with DM plus baryonic particles, constituting six simulation data frames, each having 100 unique snapshots at different redshifts.

For this work, we incorporate the TNG 300-1 hydrodynamical box corresponding to a 300 Mpc cube which offers the highest mass resolution at this scale. Table 3.3 lists the relevant simulation parameters. We extract from snapshots at redshifts $z = 3.01, 0.44$ and 0.33 , similar to the redshifts at which power spectrum observations are provided. We will specifically use the group catalog data, which consists of the halo catalog, at that redshift. Fiducial cosmological parameters incorporated for the simulations are $\Omega_{\Lambda,0} = 0.6911, \Omega_{m,0} = 0.3089, \Omega_{b,0} = 0.0486, \sigma^8 = 0.8159, n_{s_s} = 0.9667$ and $h = 0.6774$ based on Planck 2015([30]).

Number of DM particles	2500 ³
mass of dark matter particle	$5.9 \times 10^7 M_{\odot}$
baryonic matter	$1.1 \times 10^7 M_{\odot}$
Length of the box	302.6 Mpc

Table 3.3: Specifications of the TNG 300-1 box used for the simulation.

3.2.2 Friends of Friends (FoF) Halo Finder

A general N-body hydrodynamical simulation initializes with a specific number of DM particles and gas particles. Perturbation is provided to the particles by calculating the power spectrum based on linear theory, and the perturbation is passed on to the next time, eventually evolving into large-scale structures.

The outputs of the simulations are then processed via the FoF algorithm to generate a halo catalog. A halo consists of DM particles creating a site for baryonic matter infall, forming galaxies in the structure. FoF algorithm identifies halos based on the nearest particle method (Davis et al. 1985[31]). The procedure is as follows: The algorithm goes to a grid and identifies a dark matter particle. It searches for companions around the particle within some length scale. This length scale is termed the linking length. The formula for the linking length is

$$l = b \left(\frac{V}{N} \right)^{\frac{1}{3}}$$

where V is the Simulation Volume, and N is the number of dark matter particles. b is the dimensionless parameter, which is free to be changed. The fraction $(V/N)^{1/3}$ becomes the mean interparticle distance. In general, $b = 0.2$ is used as it captures the over-dense regions, and the computed halo mass function matches the analytical one. So, if the algorithm finds the particle in the length scale, it bundles it together. When it cannot find any more particles, it moves on to the next particle. This is how halos are identified using FoF. The linking length used in the IllustrisTNG is $b = 0.2$. In the simulation context, this means that 32 Dark matter particles will make up a halo. The halo catalog contains information on the number of halos identified in the simulation, including their position, velocities, and mass.

3.2.3 Line Intensity Model

In this work, we will be using the LIM simulator (Murmu et al. 2021) [19] to simulate the brightness temperature maps of CO and HI 21-cm. The simulator requires a halo catalog with the mass and position of each halo.

The simulator based on the model will generate the line luminosity maps, which will then be converted to a luminosity density field using a cloud in cell (CIC) algorithm, after which the brightness temperature map is simulated.

3.2.3.1 CO model

We will simulate the CO maps based on an empirical model proposed by Li et al. 2016 [3]. We start by relating the halo mass function to the Star Formation Rate (SFR). The SFR model is borrowed from Silva et al. 2013 [32] based on the observations in the local universe and models explaining the observed data.

$$\text{SFR}(M, z) \left[\frac{M_\odot}{\text{yr}} \right] = 2.25 \times 10^{-26} (1 + 0.075 \times (z - 7)) M^a \times \left(1 + \frac{M}{c_1} \right)^b \left(1 + \frac{M}{c_2} \right)^d \left(1 + \frac{M}{c_3} \right)^e$$

where , M = halo mass, $a = 2.59$, $b = -0.62$, $d = 0.4$, $e = -2.25$, $c_1 = 8 \times 10^8 M_\odot$, $c_2 = 7 \times 10^9 M_\odot$, and $c_3 = 1 \times 10^{11} M_\odot$ are parameters for the SFR-Halo relation.

SFRs cannot be directly measured from any observations, making it essential to relate to a quantity that is observable. SFRs can be related to CO luminosity via intermediate observables based on empirical relations. Infrared continuum luminosity L_{IR} is the ideal probe for SFR as the energetic UV photons are absorbed by the dust around the ISM and re-radiated in the Infrared. The emission in this range is impervious to dust, and as such, it escapes the region where we can observe the light. Range of L_{IR} spans from 8-1000 μm wavelength. Assuming a correlation between SFR and L_{IR} based on Kennicutt 1998 [33]:

$$\text{SFR} = \delta_{\text{MF}} \times 10^{-10} L_{\text{IR}}$$

where SFR is in units of $M_\odot \text{yr}^{-1}$ and L_{IR} in units of $L_\odot$. The normalization δ_{MF} depends on the assumption about the initial mass function, star formation duration and dust. $\delta_{\text{MF}} = 1.0$ is adopted for the model (Carilli et al. 2013 [7], Behroozi et al. 2013a [34]). L_{IR} is converted to CO luminosity based

on an empirical relation in the form of a power law.

$$\log L_{\text{IR}} = \alpha \log L'_{\text{CO}} + \beta$$

where L'_{CO} is in units of $\text{K km s}^{-1} \text{ pc}^2$ (areal integrated source brightness temperature), α and β are free parameters. L'_{CO} is converted to CO luminosity as

$$L_{\text{CO}} = 4.9 \times 10^{-5} L_{\odot} \left(\frac{\nu_{\text{CO,rest}}}{115.27 \text{ GHz}} \right)^3 \left(\frac{L'_{\text{CO}}}{\text{K km s}^{-1} \text{ pc}^2} \right)$$

where $\nu_{\text{CO,rest}} = 115.27 \text{ GHz}$ (rest frame frequency). CO emission lies in the radio range of the spectrum, making it possible to invoke the Rayleigh-Jeans approximation. The brightness temperature of CO relates to the intensity density as

$$T = \frac{\lambda^3 (1+z)^2 \rho_L}{8\pi k_B H(z)}$$

where k_B = Boltzmann constant, λ = rest frame wavelength, ρ_L = luminosity density and $H(z)$ = Hubble parameter.

3.2.3.2 HI 21-cm model

As previously discussed, HI in the IGM is fully ionized in the post-reionization regime, which leads us to select schemes that inject HI into the dark matter haloes. We will use a semi-empirical prescription to populate the halos with neutral hydrogen based on equation (6) from Bagla et al. 2010 [4]. The scheme starts by discussing the implications of ionization on HI at very low redshifts, where the amount of neutral hydrogen is found to be quite low. The gas inside halos with circular velocity in excess of 60 per second can cool down and produce stars. The relation between the circular velocity of the gas and virialized halo mass is as follows:

$$M_{\text{vir}} \simeq 10^{10} \left(\frac{v_{\text{circ}}}{60 \text{ km s}^{-1}} \right)^3 \left(\frac{1+z}{4} \right)^{-3/2} M_{\odot}$$

Simulations show that neutral hydrogen can effectively shield itself from ionizing radiation only if the host halo's circular velocity is at least $v_{\text{circ}} \sim 30 \text{ km s}^{-1}$. This limit will constitute the minimum mass a halo should have to sustain hydrogen. Similarly, neutral hydrogen content decreases in higher

mass halos, setting the limit of Maximum HI mass a massive halo can have which leads to the circular velocity cut off at $v_{\text{circ}} \sim 200 \text{ km s}^{-1}$ (Pontzen et al. 2008 [35]). These velocities set the lower and upper mass limits of halos hosting HI. The scheme, according to these limits, assigns HI mass to halos in the form:

$$M_{\text{HI}}(M_{\text{h}}) = \begin{cases} f_3 \frac{M_{\text{h}}}{1 + \frac{M_{\text{h}}}{M_{\text{max}}}} & \text{if } M_{\text{min}} \leq M_{\text{h}} \\ 0 & \text{otherwise} \end{cases}$$

where f_3 is a free parameter proportional to HI density parameter, Ω_{HI} . The dependence is as follows:

$$\begin{aligned} \frac{\bar{\rho}_{\text{HI}}}{\rho_{\text{critical},0}} &= \Omega_{\text{HI}} \\ \bar{\rho}_{\text{HI}} &= \Omega_{\text{HI}} \times \rho_{\text{critical},0} \\ \frac{M_{\text{HI}}}{V_{\text{sim}}} &= \Omega_{\text{HI}} \times \rho_{\text{critical},0} \\ M_{\text{HI}} &= \Omega_{\text{HI}} \times \rho_{\text{critical},0} \times V_{\text{sim}} \\ \sum_i f_3 \frac{M_{\text{h},i}}{1 + (\frac{M_{\text{h},i}}{M_{\text{max}}})} &= \Omega_{\text{HI}} \times \rho_{\text{critical},0} \times V_{\text{sim}} \\ f_3 &= \frac{\Omega_{\text{HI}} \times \rho_{\text{critical},0} \times V_{\text{sim}}}{\sum_i \frac{M_{\text{h},i}}{1 + (\frac{M_{\text{h},i}}{M_{\text{max}}})}} \end{aligned}$$

where $\rho_{\text{critical},0}$ is the critical density of the universe at the present epoch.

After the mass assignment, the density field, $\rho_{\text{HI}}(\vec{x})$, is created using the cloud in cell algorithm. The 21-cm brightness temperature field is then simulated following the Villaescusa-Navarro et al. 2018 scheme[36].

$$T_b(\vec{x}) = 189h \left(\frac{H_0(1.0+z)^2}{H(z)} \right) \left(\frac{\rho_{\text{HI}}(\vec{x})}{\rho_c} \right) \text{mK}$$

3.3 Inference

3.3.1 Parameter estimation via Bayesian Inference

Estimating the free parameters α , β and Ω_{HI} , discussed in the model section will be estimated given the new data by COMAP and MeerKAT, which will

help determine whether to accept or reject the model. A Bayesian framework is used for further analysis and constraining the parameters. Bayes' theorem is a fundamental theorem in probability and statistics that relates conditional probabilities. It is expressed as:

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)} \quad (3.1)$$

where:

- $P(\theta|D)$ is the posterior probability of the model parameters θ given data D .
- $P(D|\theta)$ is the likelihood of observing the data D given the parameters θ .
- $P(\theta)$ is the prior probability of the parameters θ before observing any data.
- $P(D)$ is the marginal likelihood or evidence, a normalizing constant that ensures the posterior is a proper probability distribution.

In Bayesian inference, we use Bayes' theorem to update our belief about the parameters θ of a model after observing new data D . The end goal is to find the underlying probability distribution of the parameters θ , given the observed data. In practice, the posterior distribution is proportional to the product of the likelihood and the prior:

$$P(\theta|D) \propto P(D|\theta)P(\theta) \quad (3.2)$$

This forms the basis for the parameter estimation and hypothesis testing in Bayesian Inference.

Markov Chain Monte Carlo (MCMC) method is a framework in Bayesian inference used to approximate the posterior distribution parameter space by making use of the likelihood for the data and prior knowledge of the space to sample values from the posterior distribution. The general approach is:

- Define the posterior distribution $P(\theta|D)$ using Bayes' theorem.
- Apply the MCMC algorithm

- Analyze the generated samples to make inference on the underlying distribution of the samples.

3.3.2 Bayesian Framework

We are using Bayesian-based MCMC to estimate the free parameters of the model. Figure 3.3 concisely depicts each step taken by the algorithm for which can be briefly explained as follows:

1. Initializing values of the free parameters in the parameter space.
2. Forward Modeling via LIM simulator and power spectrum calculation for the simulated map.
3. Proposing new parameter values and using the LIM-PS process.
4. Calculating probability of Likelihood using power spectrum values.
5. Computing Acceptance ratio based on derived Likelihood values.
6. Acceptance and rejection of points lead to storing or retaining the values, respectively, and proposing new points.
7. The algorithm runs for a specific number of iterations before breaking out of the loop

The main point for consideration in the above process is the perpetual use of simulating intensity maps based on the parameters at every iteration, followed by calculating the power spectrum, resulting in increased computation time for the algorithm. `emcee` (Foreman Mackey et al. 2012) [37] is an open-source MCMC library based on an affine invariant sampler proposed by Goodman-Weary(2010)[38]. We will use `emcee` for further analysis as it can explore the parameter space efficiently and faster than a home-based MCMC sampler.

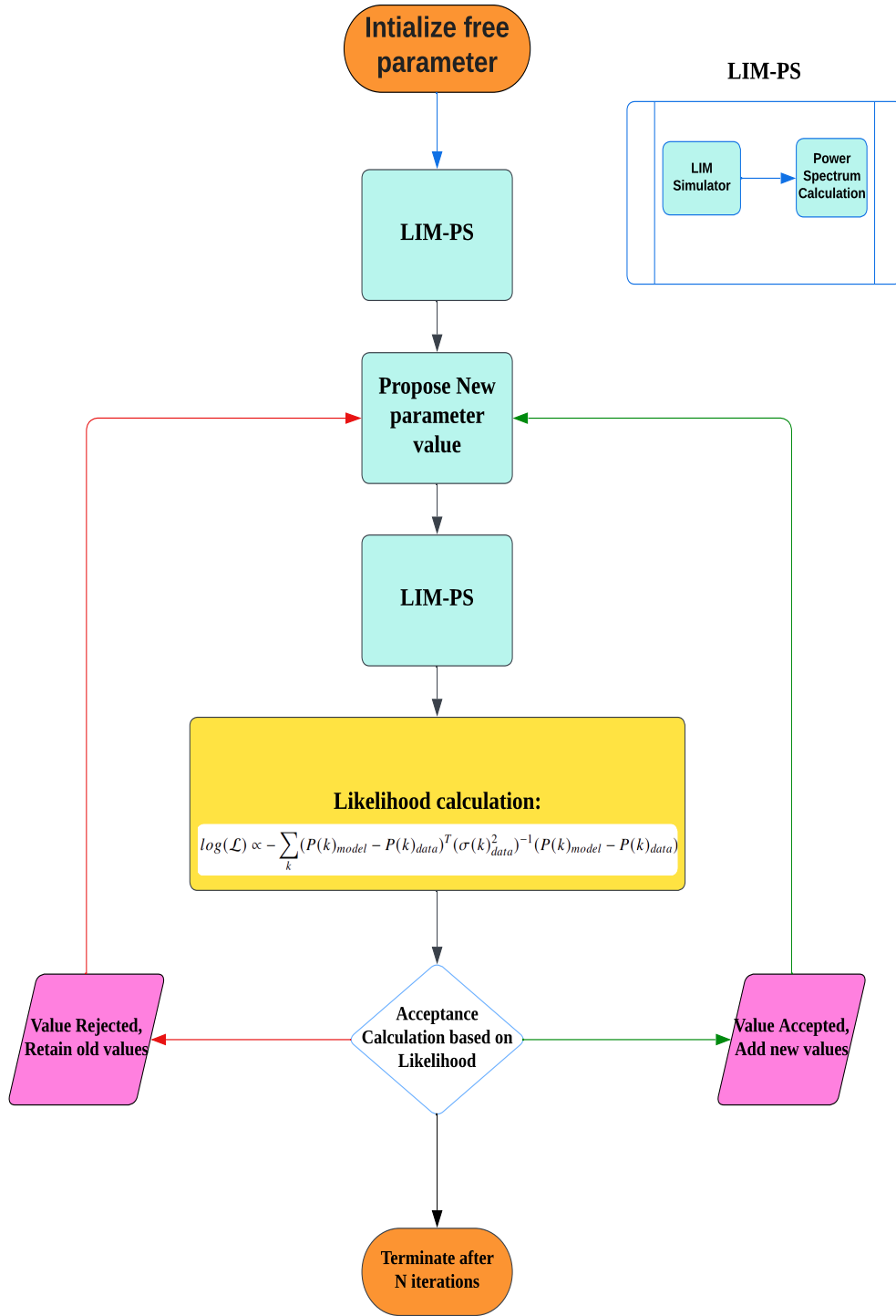


Figure 3.3: Flowchart of Markov Chain Monte Carlo Framework for parameter estimation using emcee (Foreman Mackey et al. 2012) [37].

3.3.3 Outline of the Analysis

3.3.3.1 Independent Study

From the previous section, it is evident that there is a need to define a likelihood that enables sampling correctly from the posterior distribution. For our work, we use multivariate Gaussian log-likelihood, which is defined as follows:

$$\log(\mathcal{L}) \propto - \sum_k (P(k)_{\text{model}} - P(k)_{\text{data}})^T (\sigma(k)_{\text{data}}^2)^{-1} (P(k)_{\text{model}} - P(k)_{\text{data}})$$

where $P(k)_{\text{model}}$ = Model Power spectrum, $P(k)_{\text{data}}$ = Data Power spectrum and $\sigma(k)_{\text{data}}^2$ = Data error.

3.3.3.2 Joint Analysis

Our aim is to start by analyzing the observations independently to draw inferences on the parameter space to be probed. Following the individual study, we are proposing to do a joint parameter estimation by defining a likelihood that takes into account both the CO and HI 21-cm emissions models. The joint likelihood can be written as:

$$\mathcal{L}_{\text{joint}} = \mathcal{L}_{\text{CO}} \times \mathcal{L}_{\text{HI}}$$

$$\log(\mathcal{L}_{\text{joint}}) = \log(\mathcal{L}_{\text{CO}}) + \log(\mathcal{L}_{\text{HI}})$$

CHAPTER 4

RESULTS

4.1 Line Intensity Maps and Power Spectrum

4.1.1 CO Maps and COMAP Power spectrum

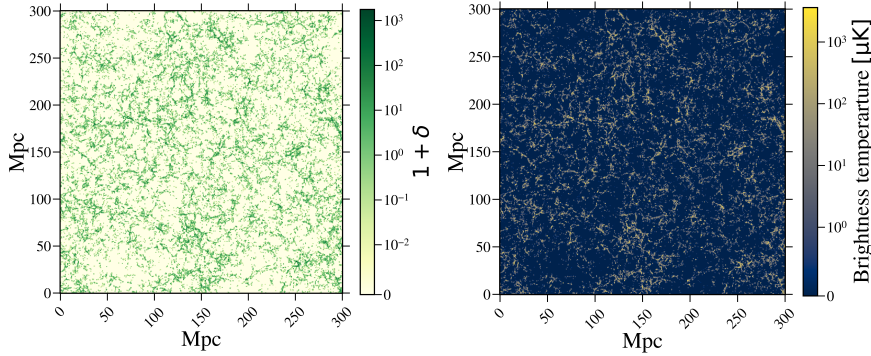


Figure 4.1: *Left:* Halo overdensity map of a slice of the TNG300-1 box. *Right:* The corresponding CO intensity map created with the LIM simulator using the empirical model based on Li et al. 2016 [3]. Both fields have a grid resolution of 0.56 Mpc.

A slice of the halo catalog from the IllustrisTNG coeval box of size 300 Mpc^3 converted to density fluctuations is shown on the left of figure 4.1. The right side of the figure 4.1 shows the CO intensity map generated using the LIM simulator.

The COMAP data consists of eight points for an eight bin averaged spherical power spectrum. In the analysis, we are considering the first four data points as shown in figure 4.2 since the last two bins enter the regime where noise from the system and resolution limit endowing for less sampling in the bins results in much larger uncertainties. The fifth and sixth bin have measured power spectrum in having negative values combined with the

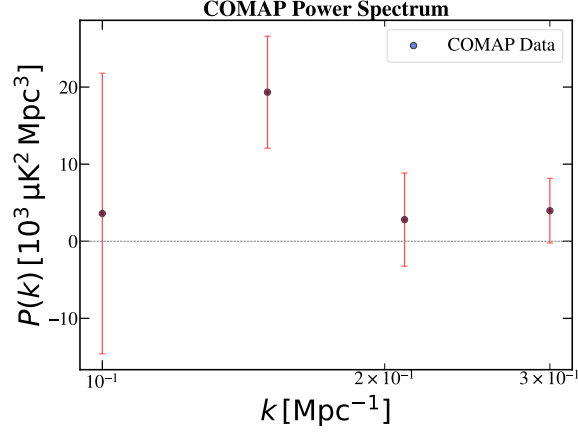


Figure 4.2: COMAP data points having high SNR considered for the Bayesian analysis. As noted in Section 3.1.2, the uncertainty associated with the second k -bin is smaller than that of the other three data points, with its lower bound extending into negative values.

error not crossing the zero value, hence a stricter measure has been taken by discarding those data points.

Figure 4.3 provides a visualization of how, from simulated CO maps, we are calculating the summary statistic, power spectrum, for the field. The power spectrum is calculated as in Murmu et al. 2021[19]. The power spectrum is estimated for the k -bins mentioned above.

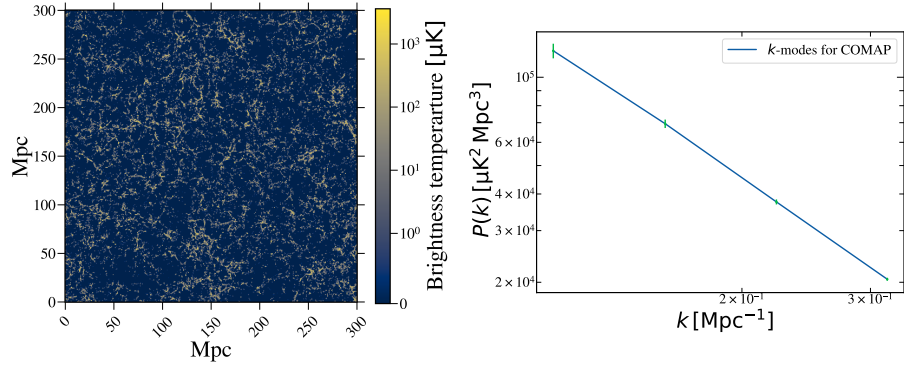


Figure 4.3: *Left:* Same realization of the CO intensity map as figure 4.1. *Right:* The power spectrum is estimated for the k -bins being considered for the analysis and is computed as in Murmu et al. 2021[19].

4.1.2 HI 21-cm Maps and MeerKAT Power Spectrum

Following the same flow of the explanation, Figure 4.4 shows the simulated intensity maps at two redshifts, $z = 0.32$ and $z = 0.44$, corresponding to the detection of MeerKAT.

MeerKAT has seven data points for seven logarithmically spaced bin-averaged spherical power spectrum at each redshift from which we will consider only those values having an SNR factor of > 2.0 . The selection criteria leave us to work with four data points, similar to COMAP, which are shown in the figure 4.5. The Power spectrum for the HI maps is thus calculated for those specific k -bins in the same way as the CO maps and is shown in Figure 4.6.

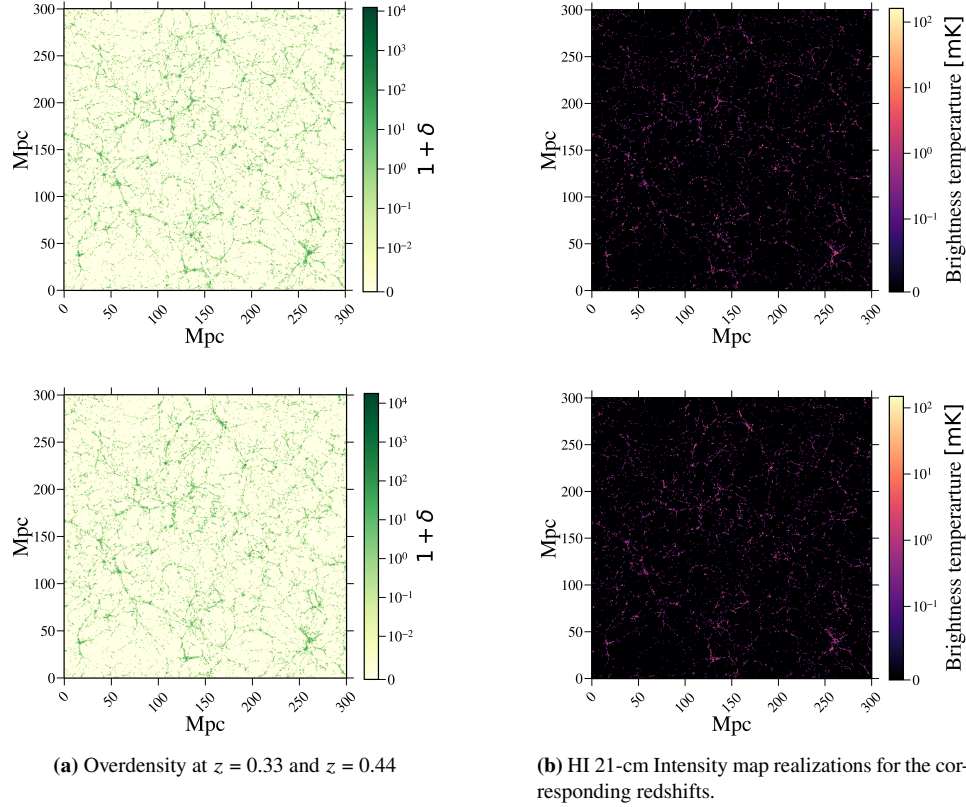


Figure 4.4: *Left Column:* Halo overdensity maps derived from the TNG300-1 simulation at $z = 0.33$ and $z = 0.44$; *Right Column:* 21-cm maps generated using scaling relations of Bagla et al. 2010[4], followed by conversion to 21-cm brightness temperature map using the Villaescusa-Navarro et al. 2018 scheme[36]. Both are coarse-gridded to a resolution of 0.42 Mpc, corresponding to the MeerKAT detections at these redshifts.

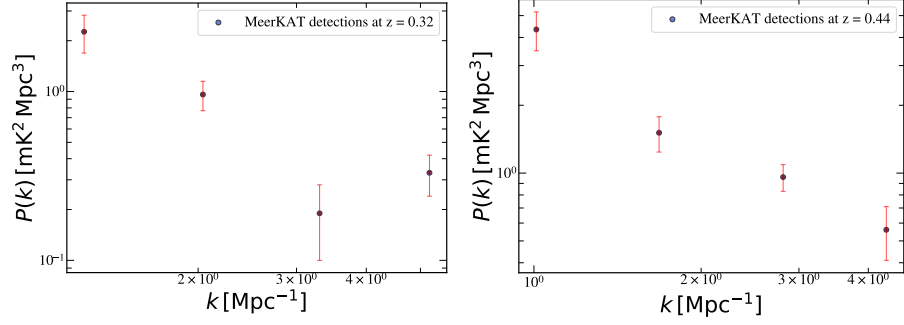
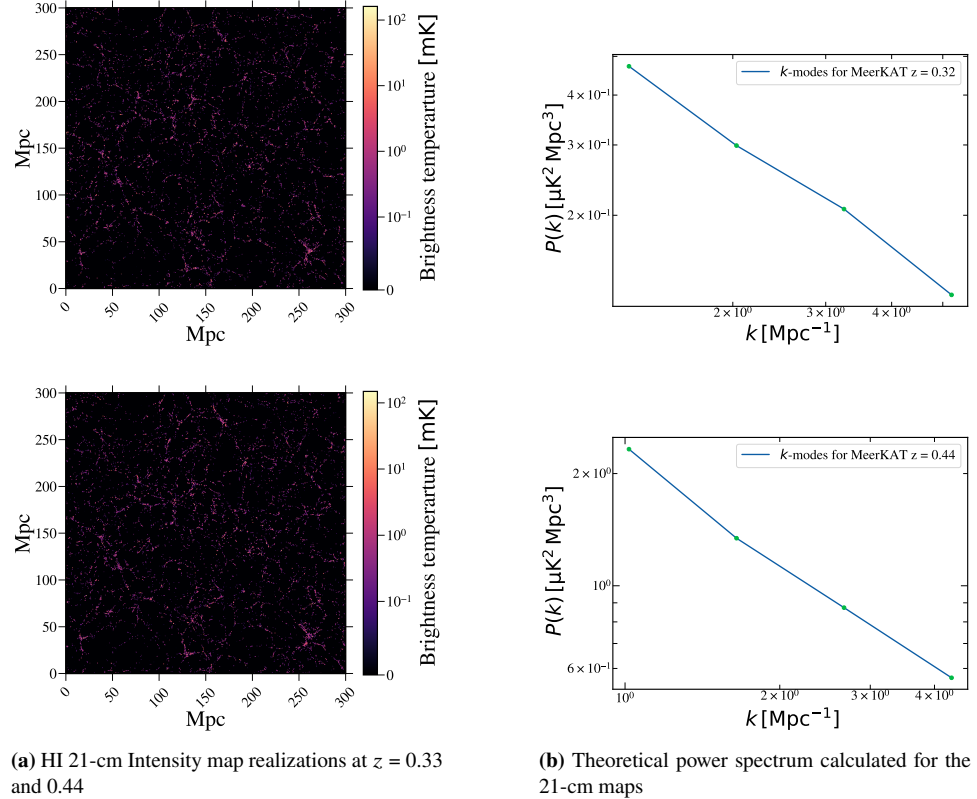


Figure 4.5: HI 21-cm power spectrum detected by the MeerKAT telescope in interferometric mode. The left and right power spectrum plots show the data points with $\text{SNR} > 2.0$ chosen for the analysis, at $z = 0.32$ and $z = 0.44$, respectively.



(a) HI 21-cm Intensity map realizations at $z = 0.33$ and 0.44

(b) Theoretical power spectrum calculated for the 21-cm maps

Figure 4.6: *Left:* HI 21-cm intensity maps (same as those shown in Figure 4.4). *Right:* Power spectrum computed for the chosen k -bins from the corresponding intensity maps.

4.2 Exploration of CO and HI Model Parameter Space

4.2.1 Constraints on α and β

Independent observations at low redshifts have enabled the formulation of an empirical CO line emission model, one of which is the model we have used for our analysis. Kamenetzky et al. 2016 [39] analysis of Herschel Space telescope data combined with data from ground-based telescopes put the value of $\alpha = 1.27 \pm 0.04$ and $\beta = -1.0 \pm 0.4$. Carilli et al. 2013[7] and Greve et al. 2014 [40] did the same process of deriving an empirical relation from the data. Figure 4.7 shows the parameter space spanned by the observations and the power spectrum, respectively, to the mean values the studies gave. It also displays the COMAP power spectrum plotted against the power spectrum of simulated maps. All the values are summarised in Table 4.1.

Observation	α	β
Kamenetzky et al. 2016 [39]	1.27 ± 0.04	-1.0 ± 0.4
Carilli et al. 2013 [7]	1.37 ± 0.04	-1.74 ± 0.40
Greve et al. 2014 [40]	1.00 ± 0.05	2.0 ± 0.5

Table 4.1: Best-fit values of α and β from different studies.

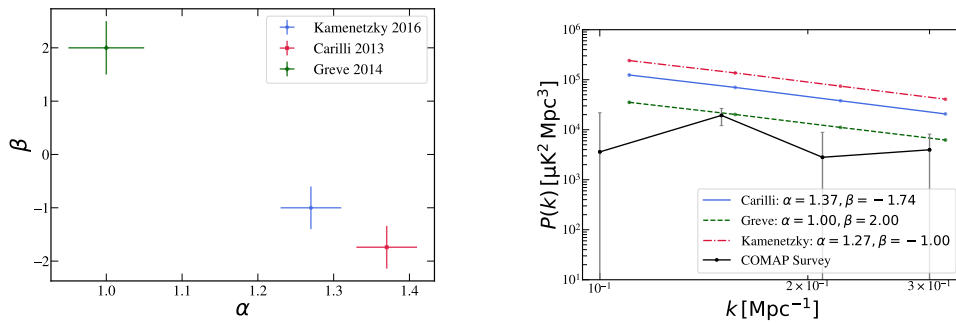


Figure 4.7: *Left:* Values of α and β derived from empirical fits to galaxy surveys, with 1σ uncertainties, for the line emission model. The Tabulated values are shown in Table 4.1. *Right:* Theoretical Power spectrum plotted with the COMAP observation; the power spectrum is calculated from the simulated map, which takes the mean values of α and β from the three surveys as input.

4.2.2 Constraints on Ω_{HI}

Detecting Ω_{HI} has been one of the prominent efforts of research through various means of approach such as proxy tracers for Ω_{HI} used in galaxy surveys where line emissions of metals and molecules are used as tracer to put stringent constraints on the density parameter. Rao et al. 2017 [41] used the archival data of the telescopes such as GALEX and SDSS to trace the Ω_{HI} . They targeted the redshift range 0.11 - 0.61 and fitted a power law to the data to infer the Ω_{HI} . Radio observations from interferometric data are also used to detect the HI signal as 21-cm falls into the radio range without any interlopers. Rhee et al. 2018 [42] used GMRT data to study the Ω_{HI} distribution in the redshift range 0.30 - 0.34 with the central z at 0.32. They used HI spectral stacking to increase the SNR and detect the HI content in the galaxies. Figure 4.8 presents the range of Ω_{HI} values from the two surveys, which encompasses the redshift range of our interest. Observations are summarised in table 4.2.

Figure 4.9 shows the nature of the theoretical power spectrum compared to the data power spectrum bins we are considering for the analysis.

Observation	$\Omega_{\text{HI}} (\times 10^{-4})$	z
Rao et al. 2017 [41]	7.7 ± 2.6	0.11 - 0.61
Rhee et al. 2018 [42]	5.0 ± 1.8	0.30 - 0.34

Table 4.2: Reported estimates of neutral hydrogen density parameter by galaxy and radio observations in the redshift range.

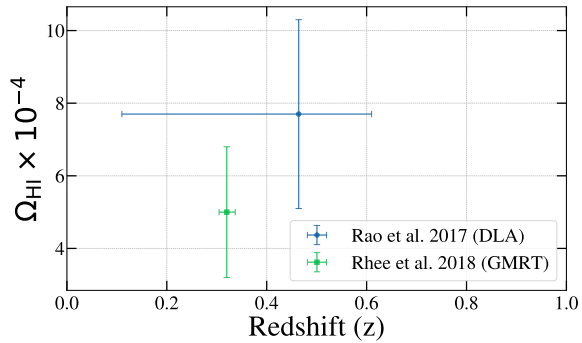


Figure 4.8: Values of Ω_{HI} reported by Rao et al. 2017 [41] and Rhee et al. 2018 [42].

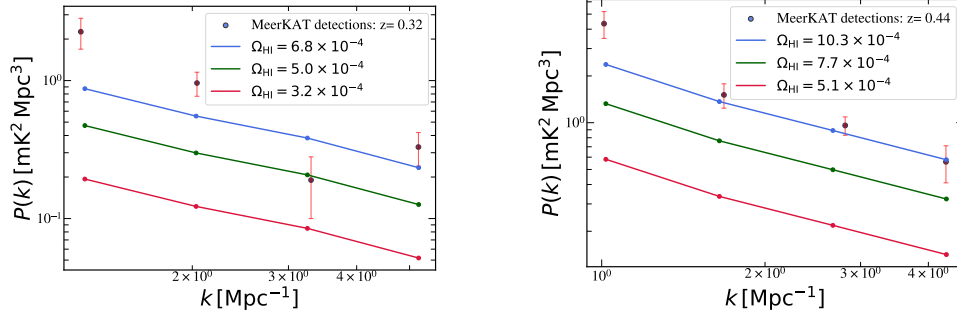


Figure 4.9: *Left* and *Right* figures show the model power spectra corresponding the the Ω_{HI} estimates of Rao et al. 2017 [41] and Rhee et al. 2018 [42] plotted in comparison to the MeerKAT data for $z = 0.33$ and $z = 0.44$, respectively. The power spectra are computed for the mean and edges of the parameter uncertainty to visualize how they vary with Ω_{HI} .

4.3 Parameter Estimation using Bayesian Inference

4.3.1 COMAP Analysis

Initially, in the testing phase, we did a run for 3000 steps with four walkers in the parameter space. The walkers are initialized at the corners of the space to explore thoroughly, 1000 points are discarded as burn-in and the thinning value is kept at 25. The results of the chain are shown in 4.10. α has a less spread in the parameter space; however, there are multiple peaks in the values, suggesting no stringent constraint can be introduced. β is spread over the entire region with multiple peaks, inferring no constraints.

A longer chain of 6000 steps with the same conditions is initialized, discarding 840 steps for burn-in and picking up the 35th point from the chain. The results, however, remains the same as shown in 4.10.

4.3.2 MeerKAT Analysis

We perform 4 MCMC runs, 2 for each $z = 0.33$ and $z = 0.44$. The results are discussed in the subsections below.

4.3.2.1 $z = 0.33$

The first run consists of 6000 steps with three walkers initialized in the parameter space at three points, the mean and the edges of the parameter

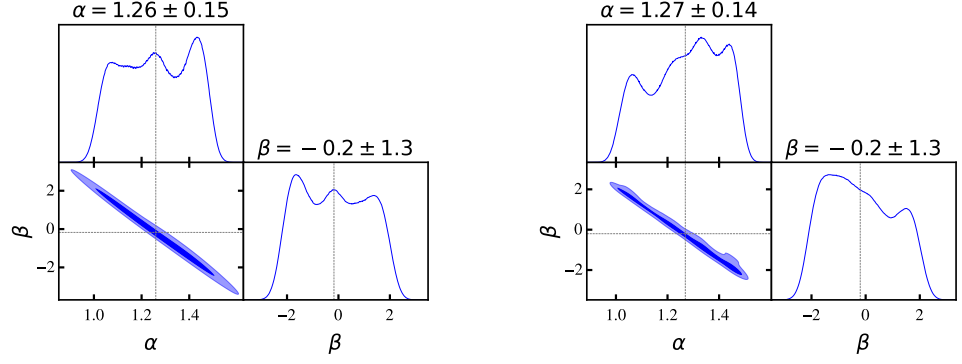


Figure 4.10: *Left:* MCMC posteriors for 3000 steps with four walkers after discarding 1000 points and a thinning factor 25. No constraints are visible over the two parameters. *Right:* MCMC posteriors for 6000 steps with four walkers after discarding 840 points and a thinning factor 35. Results remain unchanged despite the increase in walker lengths.

space defined for the redshift in table 4.2. 300 points are discarded as the burn-in phase, and the thinning factor is set to be every 20th point to break the correlation of the chains. The results are plotted in the figure 4.11, the distribution of Ω_{HI} on the left and the power spectrum for the peak value on the right, plotted against the data. We can infer from the plots that values are significantly clustered at the edge of the priors. Due to the skewed nature of the distribution, we use the median as the central line to draw the confidence intervals, and in that range, we can see the peak at $\Omega_{\text{HI}} = 6.8 \times 10^{-4}$. Yet we clearly see due to skewed distribution, a possible chance of constraints beyond the set priors. Therefore, we do a second run with the same initial

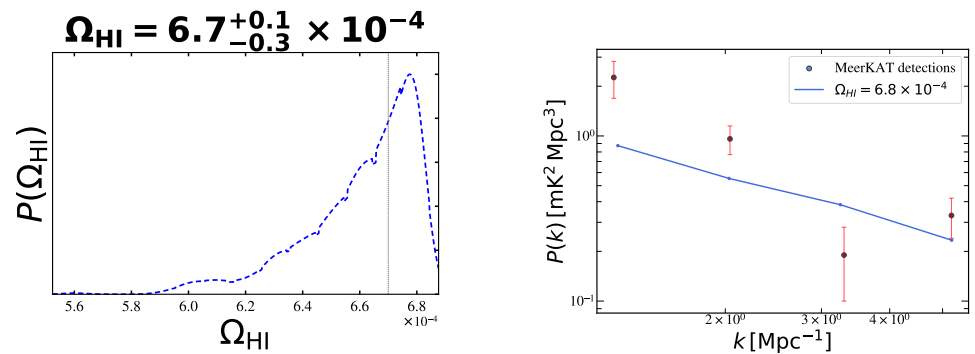


Figure 4.11: *Left:* Probability distribution of Ω_{HI} for 6000 steps with three walkers initialized at the edge of parameter space. Convergence of Ω_{HI} is visible at the edge of the parameter space described in table 4.2, the values being $\Omega_{\text{HI}} = 6.7^{+0.1}_{-0.3} \times 10^{-4}$ and the distribution peaking at 6.8×10^{-4} . *Right:* The power spectrum for $\Omega_{\text{HI}} = 6.8 \times 10^{-4}$.

conditions for the walker and update the priors by setting the lower limit to 0 and the upper limit unbounded. We discard the first 380 points from each

walker as burn-in, followed by thinning the chain by a factor of 20. This set of changes yields tighter constraints on the parameter as shown in Figure 4.12, peaking at $\Omega_{\text{HI}} = 7.2 \times 10^{-4}$ and the power spectrum for calculated for the peak.

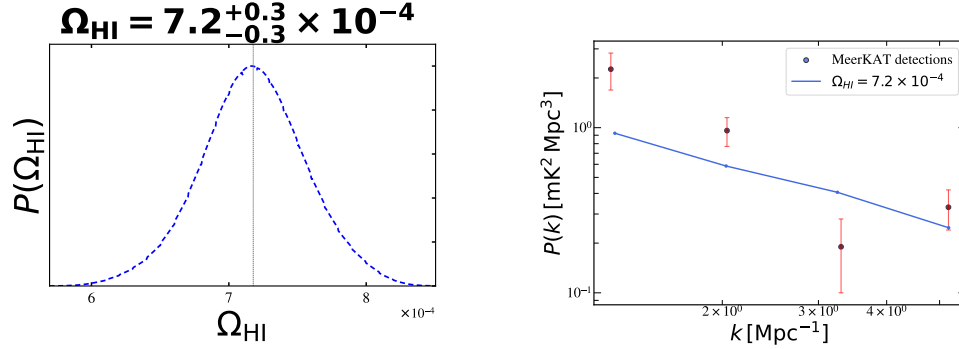


Figure 4.12: *Left:* Figure shows the results of the second MCMC run in which the prior is updated as $0 \leq \Omega_{\text{HI}}$ with no upper bounds. The probability distribution shows the convergence at $\Omega_{\text{HI}} = 7.2^{+0.3}_{-0.3} \times 10^{-4}$. *Right:* The power spectrum for the $\Omega_{\text{HI}} = 7.2 \times 10^{-4}$.

4.3.2.2 $z = 0.44$

The same analysis procedure is followed for $z = 0.44$. We perform an initial run with 6000 steps and three walkers, initialized at the mean and edges of the parameter space defined in Table 4.2. A burn-in of the first 300 steps is discarded, and the chains are thinned by a factor of 10 to reduce autocorrelation. Results are shown in Figure 4.13, where we can see the constraints with the mean value peaking at the median, $\Omega_{\text{HI}} = 10.2 \times 10^{-4}$. Skewed distribution emphasizes repeating the same exercise as before by updating the priors. The second run follows the same MCMC initializing and discarding procedure, for which results are shown in Figure 4.14. Similar to the results achieved in the previous section, we are getting significant constraints with the value peaking at $\Omega_{\text{HI}} = 10.9 \times 10^{-4}$.

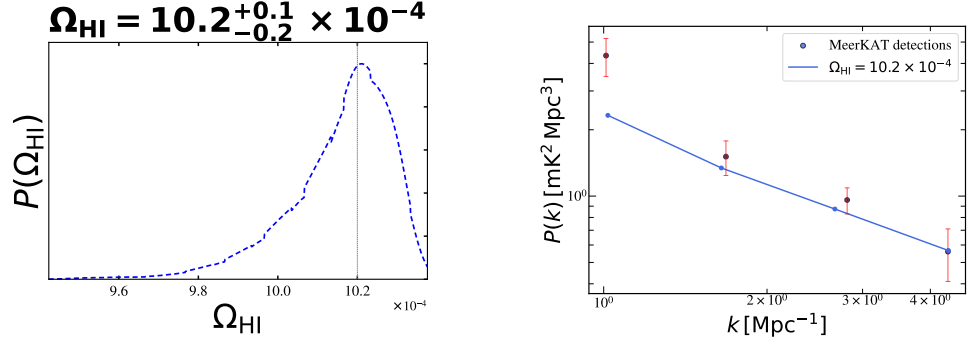


Figure 4.13: *Left:* Marginalized probability distribution of Ω_{HI} at $z = 0.44$ from an MCMC run with 6000 steps and three walkers initialized near the boundary of the parameter space. The distribution peaks at 10.2×10^{-4} , with a best-fit value of $\Omega_{\text{HI}} = 10.2^{+0.1}_{-0.2} \times 10^{-4}$, consistent with the bounds given in Table 4.2. *Right:* Power spectrum corresponding to $\Omega_{\text{HI}} = 10.2 \times 10^{-4}$.

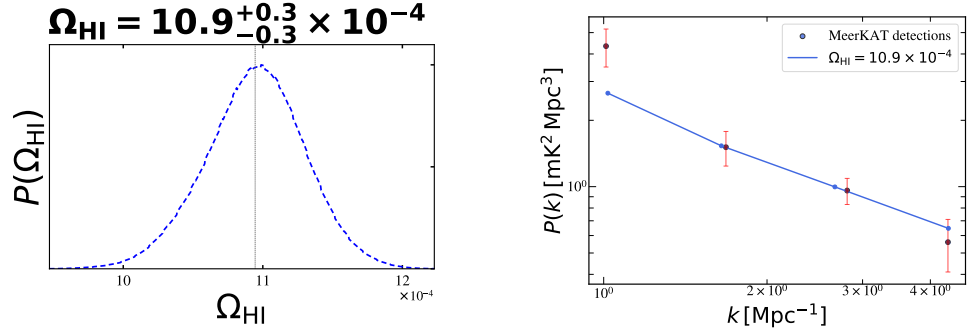


Figure 4.14: *Left:* Posterior distribution of Ω_{HI} at $z = 0.44$ from a second MCMC run using a modified prior: $0 \leq \Omega_{\text{HI}}$, without an upper bound. The parameter converges to $\Omega_{\text{HI}} = 10.9^{+0.3}_{-0.3} \times 10^{-4}$. *Right:* Corresponding power spectrum for $\Omega_{\text{HI}} = 10.9 \times 10^{-4}$.

Given that we have achieved tighter constraints on the Ω_{HI} independently using LIM, we now compare our results to the priors we had taken from literature and show how our study validates the observations and demonstrates the feasibility of the LIM technique for the future in the Figure 4.15.

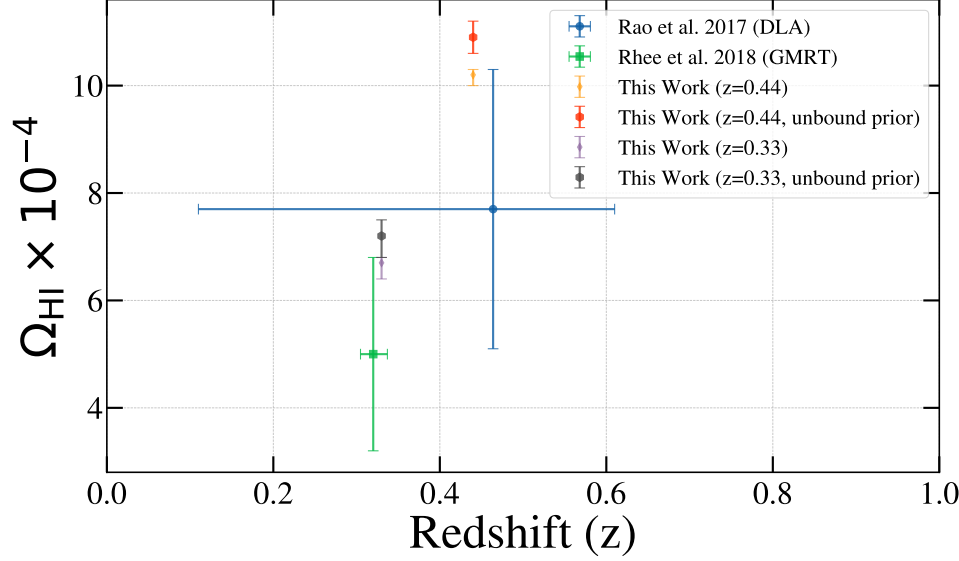


Figure 4.15: The results are validated by comparing the constraints on Ω_{HI} with those obtained from galaxy surveys and radio observations. The results show tighter constraints on the neutral hydrogen density parameter, which are consistent with the findings of Rao et al. 2017 [41] and Rhee et al. 2018 [42]. This work highlights the potential of combining Line Intensity Mapping with other observational probes to substantially enhance our understanding of both astrophysical processes and cosmological parameters in the future.

4.4 Joint Analysis of COMAP and MeerKAT

This work aims to investigate whether a joint analysis of two independent surveys, each targeting different emissions, can provide tighter constraints on various astrophysical parameters tracing distinct components of the interstellar medium (ISM), thereby validating existing empirical models.

We used the same procedure described in the methodology section to perform the MCMC run with four walkers and 6000 steps. The parameters α and Ω_{HI} are treated as free, while β is fixed to 2, following the value from Table 4.1, as the estimate from Greve et al. 2014[40] shows better agreement with the data and because the COMAP analysis indicated a degeneracy between α and β . Figure 4.16 displays the outcome of the joint analysis. The main problem we are facing here is that the autocorrelation is significant for α , resulting in the chain being shorter than needed for convergence, and the run time for this analysis took ~ 168 hours, as for each step, simulation of CO and HI was computed. Even if we bypass this and use the results by discarding 1000 points and thinning by a factor of 50, we can see no relation between the parameters, and we are getting back the same distribution for

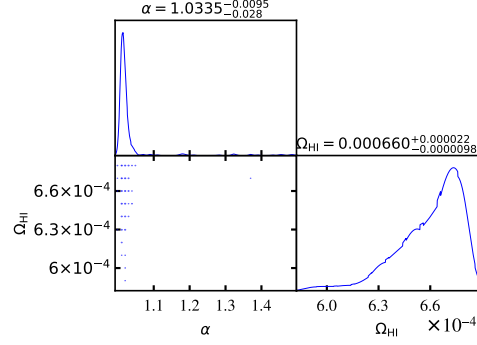


Figure 4.16: Joint Analysis of COMAP and MeerKAT Surveys: We performed an MCMC run with 6000 steps and 4 walkers, exploring the parameter space by varying α and Ω_{HI} , while fixing $\beta = 2.00$ as suggested by Greve et al. 2014 [40]. The chain did not converge even after a runtime of approximately 168 hours. We discarded the first 1000 steps as burn-in and applied a thinning factor of 50. The results yielded nonphysical parameter values, which is expected due to the lack of convergence. These findings suggest that a joint analysis is not feasible with the current data from the COMAP and MeerKAT surveys.

Ω_{HI} at $z = 0.33$ as previously viewed in the individual analysis. The main factor driving this behaviour of the likelihood space can be attributed to the large error bars in the COMAP observation, altering the likelihood valley, leading to nonphysical results.

CHAPTER 5

SUMMARY AND FUTURE PLAN

This research aimed to perform a Bayesian analysis using recent observational constraints from Line Intensity Mapping surveys such as COMAP and MeerKAT. The study focuses on constraining the astrophysical parameters of a specific line emission model for CO and HI 21-cm, as proposed in the literature, and assesses the model’s consistency with current observational data.

We summarise the work accomplished during the process here.

- COMAP targets the rotational CO line emission from the post-reionization regime in its pathfinder phase, which recently put deeper upper limits on the CO(1-0) power spectrum. Using these upper limits as detections, we aimed to evaluate how well the current models used in simulations of intensity maps can predict the observations and, if so, how stringent constraints can be imposed on the model parameters of the emission model. For this work, we used the widely used empirical model by Li et al. 2016 [3] to simulate our CO Intensity maps. The details of the CO model are discussed in section 3, where we choose α and β as free parameters. Simulation of CO and HI 21-cm maps is post-processed on halo catalogs, derived from numerical simulations, using the LIM simulator (Murmu et al. 2021[19]).
- Power spectrum of CO is observed at large length scales while the HI 21-cm is detected in smaller scales. Appropriately sampling the theoretical power spectrum at the relevant scales requires selecting a simulation box that encompasses these scales. To achieve this, we utilised the TNG 300-1 simulation, extracting the dark matter halo

catalogs at three redshifts — 3.01, 0.33, and 0.44 — corresponding to the target redshifts of the COMAP and MeerKAT surveys.

- Only the high-SNR data points from the COMAP observations, as discussed in Section 3.1.2, were selected for analysis. We used these points to perform a Bayesian inference study employing the `emcee` library for MCMC sampling. The initial results suggest that no meaningful joint constraints can be placed on the free CO model parameters. Achieving tighter constraints would require access to larger Fourier modes of the power spectrum with reduced uncertainties.
- MeerKAT, the precursor to the SKA, targets the HI 21-cm emission from low redshifts. Paul et al. 2023 [2] reported the detection of the HI 21-cm power spectrum from the DEEP2 field using the intensity mapping technique at two redshifts, 0.32 and 0.44. Bagla et al. 2010 [4] scheme is used to paint the halos with neutral hydrogen, where Ω_{HI} is the free parameter, followed by Villaescusa-Navarro et al. 2018 scheme[36] to convert it to 21-cm brightness scheme. Following the protocol, we discarded the low SNR data points and performed the Bayesian analysis. The results show tighter constraints on the parameter in both the redshifts, which we validated against previous detections via galaxy surveys and are shown in Figure 4.15.
- We performed a joint analysis of the two surveys, incorporating the likelihood described in Section 3.3.3.2. As anticipated, the results exhibited several inconsistencies due to the large uncertainties in the CO observations. Although the aim was to achieve tighter constraints on the astrophysical parameters, our findings demonstrate that this is not feasible with the current data.

In this work, we systematically explored the parameter spaces of CO and HI intensity mapping models using independent and joint analyses. Our results demonstrated that individual constraints on model parameters of CO remain weak due to observational limitations and the associated uncertainties in detection; we have displayed tighter restrictions on the HI model parameter tracing the underlying astrophysical processes. Finally, we also performed

a joint analysis, which did not reveal many insights into the astrophysical processes due to significantly large error bars from the COMAP observation.

Given the limitations identified in this study, we propose several avenues for future work:

- **Bayesian Model Selection:** Exploring different halo–SFR relations across multiple models, which can be compared using the Bayes factor and Bayesian model selection techniques, would open new possibilities for integrating CO and HI models. These models could be informed by empirical relations, hydrodynamical simulations, as well as semi-numerical and analytical approaches, thereby enabling a more comprehensive exploration of the underlying astrophysical processes.
- **Redshift Space Distortions:** Inclusion of redshift space distortions (RSD) in the simulations to accurately compare the observations and simulation outputs. The telescope will observe the RSD due to the peculiar velocities of galaxies, altering the signal statistics. In further studies, we plan to include RSD into the Bayesian inference pipeline to study its effect on parameter constraints.
- **Cosmological parameter estimation:** Estimating cosmological parameters using approximate N-body simulations, followed by astrophysical parameters, is one idea we can branch into the future.
- **Emulation:** Running LIM simulator for each run becomes computationally expensive as the run would take days to complete. We can employ techniques such as Deep learning to emulate the power spectra of various models after making the training data by varying the free parameters. This way we can reduce the inference time, try out multiple models, and validate or reject them by comparing them to observational outcomes.

BIBLIOGRAPHY

- [1] Kieran A. Cleary et al. “COMAP Early Science. I. Overview”. In: *The Astrophysical Journal* 933.2 (July 2022), p. 182. ISSN: 1538-4357. DOI: [10.3847/1538-4357/ac63cc](https://doi.org/10.3847/1538-4357/ac63cc). URL: <http://dx.doi.org/10.3847/1538-4357/ac63cc>.
- [2] Sourabh Paul et al. *A first detection of neutral hydrogen intensity mapping on Mpc scales at $z \approx 0.32$ and $z \approx 0.44$* . 2023. arXiv: [2301.11943](https://arxiv.org/abs/2301.11943) [astro-ph.CO]. URL: <https://arxiv.org/abs/2301.11943>.
- [3] Tony Y. Li et al. “CONNECTING CO INTENSITY MAPPING TO MOLECULAR GAS AND STAR FORMATION IN THE EPOCH OF GALAXY ASSEMBLY”. In: *The Astrophysical Journal* 817.2 (Jan. 2016), p. 169. ISSN: 1538-4357. DOI: [10.3847/0004-637x/817/2/169](https://doi.org/10.3847/0004-637x/817/2/169). URL: <http://dx.doi.org/10.3847/0004-637X/817/2/169>.
- [4] J. S. Bagla, Nishikanta Khandai, and Kanan K. Datta. “HI as a probe of the large-scale structure in the post-reionization universe”. In: *Monthly Notices of the Royal Astronomical Society* 407.1 (Aug. 2010), pp. 567–580. ISSN: 0035-8711. DOI: [10.1111/j.1365-2966.2010.16933.x](https://doi.org/10.1111/j.1365-2966.2010.16933.x). eprint: <https://academic.oup.com/mnras/article-pdf/407/1/567/3087523/mnras0407-0567.pdf>. URL: <https://doi.org/10.1111/j.1365-2966.2010.16933.x>.
- [5] Ely D. Kovetz et al. *Line-Intensity Mapping: 2017 Status Report*. 2017. arXiv: [1709.09066](https://arxiv.org/abs/1709.09066) [astro-ph.CO]. URL: <https://arxiv.org/abs/1709.09066>.
- [6] Marta B. Silva et al. *Prospects for detecting CII emission during the Epoch of Reionization*. 2015. arXiv: [1410.4808](https://arxiv.org/abs/1410.4808) [astro-ph.GA]. URL: <https://arxiv.org/abs/1410.4808>.
- [7] C.L. Carilli and F. Walter. “Cool Gas in High-Redshift Galaxies”. In: *Annual Review of Astronomy and Astrophysics* 51.1 (Aug. 2013), pp. 105–161. ISSN: 1545-4282. DOI: [10.1146/annurev-astro-082812-140953](https://doi.org/10.1146/annurev-astro-082812-140953). URL: <http://dx.doi.org/10.1146/annurev-astro-082812-140953>.

- [8] José Luis Bernal and Ely D. Kovetz. “Line-intensity mapping: theory review with a focus on star-formation lines”. In: *The Astronomy and Astrophysics Review* 30.1 (Sept. 2022). ISSN: 1432-0754. DOI: [10.1007/s00159-022-00143-0](https://doi.org/10.1007/s00159-022-00143-0). URL: <http://dx.doi.org/10.1007/s00159-022-00143-0>.
- [9] S. A. Wouthuysen. “On the excitation mechanism of the 21-cm (radio-frequency) interstellar hydrogen emission line.” In: *AJ* 57 (Jan. 1952), pp. 31–32. DOI: [10.1086/106661](https://doi.org/10.1086/106661).
- [10] George B. Field. “Excitation of the Hydrogen 21-CM Line”. In: *Proceedings of the IRE* 46.1 (1958), pp. 240–250. DOI: [10.1109/JRPROC.1958.286741](https://doi.org/10.1109/JRPROC.1958.286741).
- [11] P. Ade et al. “A wide field-of-view low-resolution spectrometer at APEX: Instrument design and scientific forecast”. In: *Astronomy & Astrophysics* 642 (Oct. 2020), A60. ISSN: 1432-0746. DOI: [10.1051/0004-6361/202038456](https://doi.org/10.1051/0004-6361/202038456). URL: <http://dx.doi.org/10.1051/0004-6361/202038456>.
- [12] CCAT-Prime Collaboration et al. “CCAT-prime Collaboration: Science Goals and Forecasts with Prime-Cam on the Fred Young Submillimeter Telescope”. In: *The Astrophysical Journal Supplement Series* 264.1 (Dec. 2022), p. 7. ISSN: 1538-4365. DOI: [10.3847/1538-4365/ac9838](https://doi.org/10.3847/1538-4365/ac9838). URL: <http://dx.doi.org/10.3847/1538-4365/ac9838>.
- [13] James W. Lamb et al. “COMAP Early Science. II. Pathfinder Instrument”. In: *The Astrophysical Journal* 933.2 (July 2022), p. 183. ISSN: 1538-4357. DOI: [10.3847/1538-4357/ac63c6](https://doi.org/10.3847/1538-4357/ac63c6). URL: <http://dx.doi.org/10.3847/1538-4357/ac63c6>.
- [14] Håvard T. Ihle et al. “COMAP Early Science. IV. Power Spectrum Methodology and Results”. In: *The Astrophysical Journal* 933.2 (July 2022), p. 185. ISSN: 1538-4357. DOI: [10.3847/1538-4357/ac63c5](https://doi.org/10.3847/1538-4357/ac63c5). URL: <http://dx.doi.org/10.3847/1538-4357/ac63c5>.
- [15] N. -O. Stutzer et al. *COMAP Pathfinder – Season 2 results II. Updated constraints on the CO(1-0) power spectrum*. 2024. arXiv: [2406.07511](https://arxiv.org/abs/2406.07511) [[astro-ph.CO](https://arxiv.org/abs/2406.07511)]. URL: <https://arxiv.org/abs/2406.07511>.
- [16] Olivier Doré et al. *Cosmology with the SPHEREX All-Sky Spectral Survey*. 2015. arXiv: [1412.4872](https://arxiv.org/abs/1412.4872) [[astro-ph.CO](https://arxiv.org/abs/1412.4872)]. URL: <https://arxiv.org/abs/1412.4872>.
- [17] Eli Visbal and Abraham Loeb. “Measuring the 3D clustering of undetected galaxies through cross correlation of their cumulative flux fluctuations from multiple spectral lines”. In: *Journal of Cosmology and Astroparticle Physics* 2010.11 (Nov. 2010), pp. 016–016. ISSN: 1475-7516. DOI: [10.1088/1475-7516/2010/11/016](https://doi.org/10.1088/1475-7516/2010/11/016). URL: <http://dx.doi.org/10.1088/1475-7516/2010/11/016>.
- [18] Adam Lidz and Jessie Taylor. “ON REMOVING INTERLOPER CONTAMINATION FROM INTENSITY MAPPING POWER SPECTRUM MEASUREMENTS”. In: *The Astrophysical Journal* 825.2 (July 2016), p. 143. ISSN: 1538-4357. DOI: [10.3847/0004-637X/825/2/143](https://doi.org/10.3847/0004-637X/825/2/143). URL: <http://dx.doi.org/10.3847/0004-637X/825/2/143>.

- [19] Chandra Shekhar Murmu, Suman Majumdar, and Kanan K Datta. “C ii and H i 21-cm line intensity mapping from the EoR: impact of the light-cone effect on auto and cross-power spectra”. In: *Monthly Notices of the Royal Astronomical Society* 507.2 (Aug. 2021), pp. 2500–2509. ISSN: 1365-2966. DOI: [10.1093/mnras/stab2347](https://doi.org/10.1093/mnras/stab2347). URL: <http://dx.doi.org/10.1093/mnras/stab2347>.
- [20] J. G. S. Lunde et al. *COMAP Pathfinder – Season 2 results I. Improved data selection and processing*. 2024. arXiv: [2406.07510](https://arxiv.org/abs/2406.07510) [[astro-ph.CO](https://arxiv.org/abs/2406.07510)]. URL: <https://arxiv.org/abs/2406.07510>.
- [21] D. T. Chung et al. *COMAP Pathfinder – Season 2 results III. Implications for cosmic molecular gas content at "Cosmic Half-past Eleven"*. 2024. arXiv: [2406.07512](https://arxiv.org/abs/2406.07512) [[astro-ph.CO](https://arxiv.org/abs/2406.07512)]. URL: <https://arxiv.org/abs/2406.07512>.
- [22] G. Hinshaw et al. “NINE-YEAR WILKINSON MICROWAVE ANISOTROPY PROBE (WMAP) OBSERVATIONS: COSMOLOGICAL PARAMETER RESULTS”. In: *The Astrophysical Journal Supplement Series* 208.2 (Sept. 2013), p. 19. ISSN: 1538-4365. DOI: [10.1088/0067-0049/208/2/19](https://doi.org/10.1088/0067-0049/208/2/19). URL: <http://dx.doi.org/10.1088/0067-0049/208/2/19>.
- [23] Marie K. Foss et al. “COMAP Early Science. III. CO Data Processing”. In: *The Astrophysical Journal* 933.2 (July 2022), p. 184. ISSN: 1538-4357. DOI: [10.3847/1538-4357/ac63ca](https://doi.org/10.3847/1538-4357/ac63ca). URL: <http://dx.doi.org/10.3847/1538-4357/ac63ca>.
- [24] N. Aghanim et al. “Planck2018 results: VI. Cosmological parameters”. In: *Astronomy and Astrophysics* 641 (Sept. 2020), A6. ISSN: 1432-0746. DOI: [10.1051/0004-6361/201833910](https://doi.org/10.1051/0004-6361/201833910). URL: <http://dx.doi.org/10.1051/0004-6361/201833910>.
- [25] Volker Springel et al. “First results from the IllustrisTNG simulations: matter and galaxy clustering”. In: *Monthly Notices of the Royal Astronomical Society* 475.1 (Dec. 2017), pp. 676–698. ISSN: 1365-2966. DOI: [10.1093/mnras/stx3304](https://doi.org/10.1093/mnras/stx3304). URL: <http://dx.doi.org/10.1093/mnras/stx3304>.
- [26] Dylan Nelson et al. “First results from the IllustrisTNG simulations: the galaxy colour bimodality”. In: *Monthly Notices of the Royal Astronomical Society* 475.1 (Nov. 2017), pp. 624–647. ISSN: 1365-2966. DOI: [10.1093/mnras/stx3040](https://doi.org/10.1093/mnras/stx3040). URL: <http://dx.doi.org/10.1093/mnras/stx3040>.
- [27] Annalisa Pillepich et al. “First results from the IllustrisTNG simulations: the stellar mass content of groups and clusters of galaxies”. In: *Monthly Notices of the Royal Astronomical Society* 475.1 (Dec. 2017), pp. 648–675. ISSN: 1365-2966. DOI: [10.1093/mnras/stx3112](https://doi.org/10.1093/mnras/stx3112). URL: <http://dx.doi.org/10.1093/mnras/stx3112>.
- [28] Jill P Naiman et al. “First results from the IllustrisTNG simulations: a tale of two elements – chemical evolution of magnesium and europium”. In: *Monthly Notices of the Royal Astronomical Society* 477.1 (Mar. 2018), pp. 1206–1224. ISSN: 1365-2966. DOI: [10.1093/mnras/sty618](https://doi.org/10.1093/mnras/sty618). URL: <http://dx.doi.org/10.1093/mnras/sty618>.

- [29] Federico Marinacci et al. “First results from the IllustrisTNG simulations: radio haloes and magnetic fields”. In: *Monthly Notices of the Royal Astronomical Society* (Aug. 2018). ISSN: 1365-2966. DOI: [10.1093/mnras/sty2206](https://doi.org/10.1093/mnras/sty2206). URL: <http://dx.doi.org/10.1093/mnras/sty2206>.
- [30] P. A. R. Ade et al. “Planck2015 results: XIII. Cosmological parameters”. In: *Astronomy & Astrophysics* 594 (Sept. 2016), A13. ISSN: 1432-0746. DOI: [10.1051/0004-6361/201525830](https://doi.org/10.1051/0004-6361/201525830). URL: <http://dx.doi.org/10.1051/0004-6361/201525830>.
- [31] Marc Davis et al. “The evolution of large-scale structure in a universe dominated by cold dark matter”. In: *The Astrophysical Journal* 292 (1985), pp. 371–394.
- [32] Marta B. Silva et al. “INTENSITY MAPPING OF Ly alpha EMISSION DURING THE EPOCH OF REIONIZATION”. In: *The Astrophysical Journal* 763.2 (Jan. 2013), p. 132. ISSN: 1538-4357. DOI: [10.1088/0004-637x/763/2/132](https://doi.org/10.1088/0004-637x/763/2/132). URL: <http://dx.doi.org/10.1088/0004-637X/763/2/132>.
- [33] Robert C. Kennicutt Jr. “The Global Schmidt Law in Star-forming Galaxies”. In: *The Astrophysical Journal* 498.2 (May 1998), pp. 541–552. ISSN: 1538-4357. DOI: [10.1086/305588](https://doi.org/10.1086/305588). URL: <http://dx.doi.org/10.1086/305588>.
- [34] Peter S. Behroozi, Risa H. Wechsler, and Charlie Conroy. “THE AVERAGE STAR FORMATION HISTORIES OF GALAXIES IN DARK MATTER HALOS FROM $z=0-8$ ”. In: *The Astrophysical Journal* 770.1 (May 2013), p. 57. ISSN: 1538-4357. DOI: [10.1088/0004-637x/770/1/57](https://doi.org/10.1088/0004-637x/770/1/57). URL: <http://dx.doi.org/10.1088/0004-637X/770/1/57>.
- [35] Andrew Pontzen et al. “Damped Lyman alpha systems in galaxy formation simulations”. In: *Monthly Notices of the Royal Astronomical Society* 390.4 (Oct. 2008), pp. 1349–1371. ISSN: 0035-8711. DOI: [10.1111/j.1365-2966.2008.13782.x](https://doi.org/10.1111/j.1365-2966.2008.13782.x). eprint: <https://academic.oup.com/mnras/article-pdf/390/4/1349/3018831/mnras0390-1349.pdf>. URL: <https://doi.org/10.1111/j.1365-2966.2008.13782.x>.
- [36] Francisco Villaescusa-Navarro et al. “Ingredients for 21 cm Intensity Mapping”. In: *The Astrophysical Journal* 866.2 (Oct. 2018), p. 135. DOI: [10.3847/1538-4357/aadba0](https://doi.org/10.3847/1538-4357/aadba0). URL: <https://dx.doi.org/10.3847/1538-4357/aadba0>.
- [37] Daniel Foreman-Mackey et al. “emcee: The MCMC Hammer”. In: *Publications of the Astronomical Society of the Pacific* 125.925 (Mar. 2013), pp. 306–312. ISSN: 1538-3873. DOI: [10.1086/670067](https://doi.org/10.1086/670067). URL: <http://dx.doi.org/10.1086/670067>.
- [38] Jonathan Goodman and Jonathan Weare. “Ensemble samplers with affine invariance”. In: *Communications in Applied Mathematics and Computational Science* 5.1 (2010), pp. 65–80. DOI: [10.2140/camcos.2010.5.65](https://doi.org/10.2140/camcos.2010.5.65).
- [39] J. Kamenetzky et al. “RELATIONS WITH CO ROTATIONAL LADDERS OF GALAXIES ACROSS THE HERSCHEL SPIRE ARCHIVE”. In: *The Astrophysical Journal* 829.2 (Sept. 2016), p. 93. ISSN: 1538-4357. DOI: [10.3847/0004-637x/829/2/93](https://doi.org/10.3847/0004-637x/829/2/93). URL: <http://dx.doi.org/10.3847/0004-637X/829/2/93>.

- [40] T. R. Greve et al. “STAR FORMATION RELATIONS AND CO SPECTRAL LINE ENERGY DISTRIBUTIONS ACROSS THE J-LADDER AND REDSHIFT”. In: *The Astrophysical Journal* 794.2 (Oct. 2014), p. 142. ISSN: 1538-4357. DOI: 10.1088/0004-637x/794/2/142. URL: <http://dx.doi.org/10.1088/0004-637X/794/2/142>.
- [41] Sandhya M. Rao et al. “The statistical properties of neutral gas at $z < 1.65$ from UV measurements of Damped Lyman Alpha systems”. In: *Monthly Notices of the Royal Astronomical Society* 471.3 (July 2017), pp. 3428–3442. ISSN: 0035-8711. DOI: 10.1093/mnras/stx1787. eprint: <https://academic.oup.com/mnras/article-pdf/471/3/3428/19518380/stx1787.pdf>. URL: <https://doi.org/10.1093/mnras/stx1787>.
- [42] Jonghwan Rhee et al. “Neutral hydrogen (H I) gas content of galaxies at $z \approx 0.32$ ”. In: *Monthly Notices of the Royal Astronomical Society* 473.2 (Sept. 2017), pp. 1879–1894. ISSN: 0035-8711. DOI: 10.1093/mnras/stx2461. eprint: <https://academic.oup.com/mnras/article-pdf/473/2/1879/21515410/stx2461.pdf>. URL: <https://doi.org/10.1093/mnras/stx2461>.