

**Study of Inter-Galactic Medium in the Epoch of Reionization era from
Simulated HI 21cm images using Wavelet Scattering Transform**

M.Sc Thesis (ongoing)

*To be submitted in partial fulfillment of the
requirements for the degree of*

Master of Science in Astronomy

by

Praneeth Avasarala

Under the guidance of Dr. Suman Majumdar



**Department of Astronomy, Astrophysics and Space Engineering
INDIAN INSTITUTE OF TECHNOLOGY INDORE**

ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisor Dr. Suman Majumdar for granting me the opportunity to work on this project and providing me with freedom and guidance whenever required. I would like to thank my colleagues from my lab – Leon, Chandrasekhar & others ; and my batch-mates – Sanjay, Hemanth, Asad, Sparsh & others for their willingness to have discussions with me, unwavering support and belief in me particularly during challenging times. I am immensely grateful to the Department of Astronomy, Astrophysics and Space Engineering for allowing me to make a contribution to this vibrant field. Finally, I am proud of my past self for making the difficult decision to embark on a new journey into Astronomy even though I am from a mechanical engineering background & for giving my present and future self the confidence and belief in myself to keep pursuing my dreams without doubting my capabilities.

ABSTRACT

Epoch of Reionization (EoR) is a crucial phase **when the Intergalactic Medium (IGM), primarily composed of neutral hydrogen (HI), began to undergo ionization due to the influence of high-energy photons emitted by various sources.** This epoch **holds critical significance because it sheds light on the timing and nature of the first luminous sources, provides context for interpreting high-redshift observations affected by intervening IGM, and serves as a test for dark matter and galaxy formation models, which must produce sources consistent with reionization constraints.** Observations so far haven't been able to probe the exact duration, timing of reionization & the characteristics of the primary ionizing sources.

The HI 21-cm signal holds significant promise as a direct means of investigating these critical aspects of this epoch. **By directly measuring HI 21-cm signal at different redshifts along with its spatial distribution in the IGM, we can trace the history of reionization as a three dimensional movie of the early universe. The upcoming Square Kilometer Array (SKA) telescope is expected to be capable of such 21-cm observations.**

SKA is expected to make tomographic images of the neutral hydrogen distribution during the Epoch of Reionization via the redshifted 21-cm signal. Non-Gaussian information from these observations can reveal, in principle, the nature of ionizing sources and the dominant astrophysical processes in the Intergalactic medium. **We apply Wavelet Scattering Transform (WST) to analyze simulated HI 21-cm images & extract information from 21-cm images to help answer these questions about the IGM.** We consider different reionization sources in our work which emit UV, soft X-ray & hard X-ray photons. We have also considered the effects of halo mass dependent power-law efficiency and non-uniform recombination. We compare the WST statistic with power spectrum. We demonstrate that 2D WST outperforms 3D spherically averaged power spectrum when it comes to its ability to probe non-Gaussian information. Our work also shows that 2D WST performs better than 3D Bispectrum in distinguishing some of the source models. We further quantify the impact of light cone effect on WST Analysis by generating lightcones from **our simulated coeval cubes** and found that lightcone results follow the same trends as those of the simulated coeval cubes, which is expected from the 2D nature of WST.

CONTENTS

List of Figures	iii
List of Tables	iii
1 Epoch of Reionization & HI 21-cm Signal	1
1.1 History of the Universe :	1
1.2 Observations & Challenges :	3
2 21-cm Brightness Temperature maps of various Reionization Source Models :	7
2.1 Simulation Parameters :	7
2.2 Reionization Source models considered in my thesis work :	8
2.2.1 UV Photons from Galaxies :	8
2.2.2 Uniform Ionizing Background :	8
2.2.3 Soft X-Ray photons :	9
2.2.4 Power Law mass dependent efficiency :	9
2.2.5 Inhomogenous Clumping Scenario :	9
3 Wavelet Scattering Transform (WST)	11
3.1 Wavelets:	12
3.2 Scattering Transform :	14
4 Results	17
4.1 WST Analysis of different reionization source models :	17
4.1.1 Choice of Wavelet filter – Morlet Wavelet :	17
4.1.2 Fiducial Model :	18
4.1.3 Fiducial vs PL 2.0 vs PL 3.0 Models :	25
4.1.4 Fiducial vs Inhomogenous Clumping Models :	28
4.1.5 Fiducial vs UIB Models :	33
4.1.6 Fiducial vs SXR Dom vs UV+SXR+UIB Models :	37
4.2 WST Analysis of Lightcones:	44

4.2.1	Creating Lightcones from Simulated Volumes:	44
4.2.2	Calculating neutral fractions for Lightcone slices:	45
4.2.3	Binning Lightcone slices – Reducing Sample Variance:	45
4.2.4	WST Results comparison – Lightcones Vs Simulated Volumes:	45
4.3	Comparing WST with other Statistics :	50
4.4	Summary of the results so far :	52
5	Future Scope	53
5.1	Impact of Noise :	53
5.2	Impact of Telescope effects:	54
5.3	Reionization Model Selection with WST :	54

LIST OF FIGURES

1.1	Image Credits : European Space Agency, Planck Collaboration [7]	1
1.2	“The Evolution of Lyman limit absorption systems to $z \approx 6$ ” – Songaila and Cowie [1]	2
1.3	Constraining the Evolution of the Ionizing Background and the Epoch of Reionization with $z \approx 6$ Quasars. II. A Sample of 19 Quasars – Fan and Strauss et al [2]	2
1.4	Expression for Lyman- α optical depth of the IGM [20]. RHS of this equation makes optical depth highly sensitive to the presence of neutral hydrogen. Even a 0.001 \bar{x}_{HI} is enough to saturate the Lyman- α optical depth which limits our understanding to lower neutral fractions of EoR.	4
1.5	How an evolving 21-cm signal tomographic image is expected to look like. Y dimension is along the sky plane & X dimension is along the radially outward direction from us. Image from “Light cone effect on the reionization 21-cm signal II: Evolution, anisotropies and observational implications” - Datta et al. [3].	4
2.1	Redshift space brightness temperature maps for six of our reionization scenarios. The line of sight is along the y-axis. The mass averaged neutral fraction for all the panels is $\bar{x}_{HI} = 0.5$. Image from “Effects of the sources of reionization on 21-cm redshift-space distortions” - Majumdar et al. [4].	8
2.2	The relative contribution from different source types in our reionization scenarios. UV=Ultra-violet; UIB=Uniform ionizing background; SXR=Soft X-ray background; PL=Power-law. Table from “Effects of the sources of reionization on 21-cm redshift-space distortions” - Majumdar et al. [4].	10
3.1	Two fields with visibly different textures but same power spectrum. The left panel shows a simulated 21-cm brightness temperature image during EoR which is highly non-Gaussian. The right panel is a Gaussian Random Field. Image from "HI tomographic imaging of the Cosmic Dawn and Epoch of Reionization with SKA" - Mellema et al. [5]	11
3.2	Different types of Wavelets. Image from "How Wavelets allow researchers to transform and understand data" - Samuel Velasco [44]	13

3.3	Comparing fourier filters & wavelet filters [37]	13
3.4	Scattering Transform and its operations visualised. Image from “Astrophysics and cosmology with the scattering transform” - [37].	15
4.1	Wavelet Filters used in our work	17
4.2	Evolution of S_1 coefficients vs. J with \bar{x}_{HI} for Fiducial Model	18
4.3	Evolution of $S_1(j)$ coefficients with \bar{x}_{HI} for Fiducial Model	18
4.4	Evolution of $S_{21}(j_1, j_2)$ coefficients with \bar{x}_{HI} for Fiducial model	21
4.5	Evolution of S_1 coefficients vs. J with \bar{x}_{HI} for Fiducial, PL2.0 & PL 3.0 models	25
4.6	Evolution of $S_1(j)$ coefficients with \bar{x}_{HI} for Fiducial, PL 2.0 & PL 3.0 models	25
4.7	Evolution of $S_{21}(j_1, j_2)$ coefficients with \bar{x}_{HI} for Fiducial, PL 2.0 & PL 3.0 models	26
4.8	Evolution of S_1 coefficients vs. J with \bar{x}_{HI} for Fiducial & Clumping models	28
4.9	Evolution of S_{21} coefficients with \bar{x}_{HI} for Fiducial & Clumping models	30
4.10	Evolution of $S_1(j)$ coefficients with varying \bar{x}_{HI} for Fiducial & UIB models	33
4.11	Evolution of $S_{21}(j_1, j_2)$ coefficients with varying \bar{x}_{HI} for Fiducial & UIB models	34
4.12	Evolution of S_1 coefficients vs. J with \bar{x}_{HI} for Fiducial, SXR Dom & UV+SXR+UIB Dom models	37
4.13	Evolution of $S_1(j)$ coefficients with varying \bar{x}_{HI} for Fiducial, SXR Dom & UV+SXR+UIB Dom models	37
4.14	Evolution of Fiducial, SXR Dom & UV+SXR+UIB Dom models	39
4.15	Evolution of Fiducial, SXR Dom & UV+SXR+UIB Dom models	41
4.16	Evolution of $S_{21}(j_1, j_2)$ coefficients with varying \bar{x}_{HI} for Fiducial, SXR Dominated and UV+SXR+UIB Models	42
4.17	Comparing WST results of lightcone with WST results of simulated volumes	45
4.18	The figure on the top shows 3D power spectrum results on 21-cm line brightness temperature maps for all our models considered from Majumdar et al. [4]. The figure on the bottom shows 2D WST S_1 coefficient results for the same models. We can clearly see the similar trends in Power spectrum and WST S_1 results.	50
4.19	The figure on the top shows 3D Bispectrum results on 21-cm line brightness temperature maps for Fiducial & Clumping models. The figure on the bottom shows 2D WST S_2 coefficient results for the same models. We can clearly see that WST performs better at distinguishing the two models.	51
5.1	Interpreting the HI 21-cm cosmology maps through Largest Cluster Statistics - I: Impact of the synthetic SKA1-Low observations - Dasgupta et al. (2023) [46]	53
5.2	Telescope model considered in Ghara et al. [6] to match the SKA1-low capabilities as given in SKA documentation.	54

LIST OF TABLES

4.1 Sizes of the Wavelet filters used in our WST analysis	18
---	----

CHAPTER 1

EPOCH OF REIONIZATION & HI 21-CM SIGNAL

1.1 History of the Universe :

The cosmic microwave background radiation observed by Planck spacecraft has provided compelling evidence indicating a transformative phase in the universe's history. This period witnessed a pivotal shift from a state of hot, dense plasma to a universe filled predominantly with neutral atoms, with hydrogen emerging as the most abundant element. However, insights gleaned from studying quasars in our local universe present a different narrative.

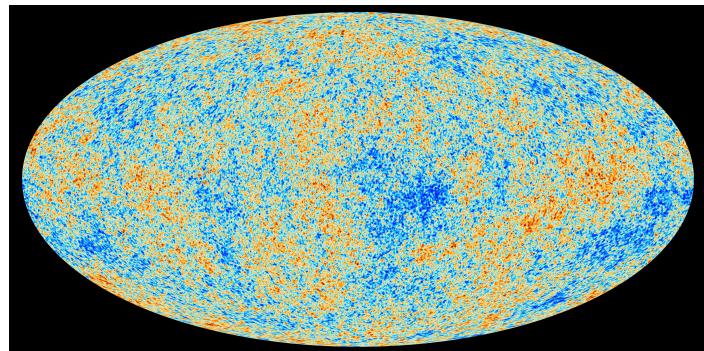


Figure 1.1: Image Credits : European Space Agency, Planck Collaboration [7]

As we scrutinize radiation emitted by quasars from earlier epochs, a gradual increase in absorption in Lyman alpha radiation and lower wavelengths becomes apparent. This intriguing observation finds its explanation in the rising prevalence of neutral hydrogen within the Intergalactic Medium (IGM) during the earlier stages of the universe's existence. The evolving nature of the IGM becomes evident as observations of quasars closer to our current cosmic epoch reveal a contrasting trend. In these later stages of the universe, the observation of radiation from quasars highlights an increasing presence of ionized hydrogen, which exhibits transparency to Lyman alpha radiation. This transition from an era characterized by neutral hydrogen dominating the IGM to a phase where ionized hydrogen becomes prevalent defines a significant cosmic epoch known as the **Epoch of Reionization (EoR)**. This transformation

from a state of neutral to ionized hydrogen marks a pivotal phase in the cosmic timeline, shaping the fundamental composition and characteristics of the universe as we understand it today.

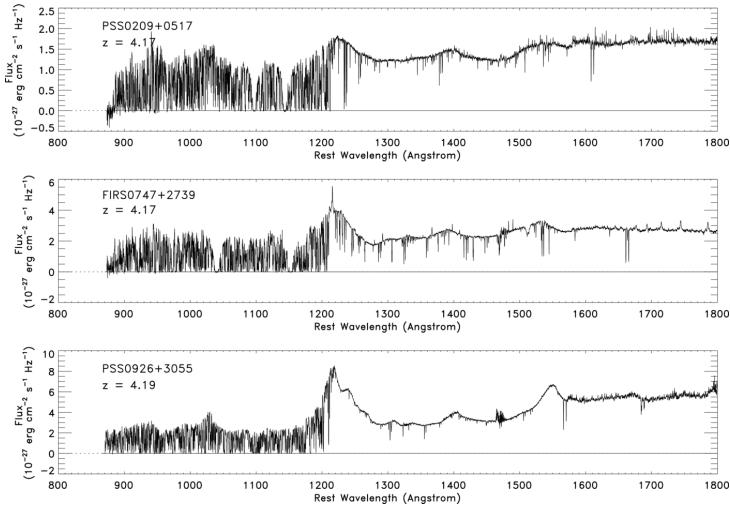


Figure 1.2: “The Evolution of Lyman limit absorption systems to $z \approx 6$ ” – Songaila and Cowie [1]

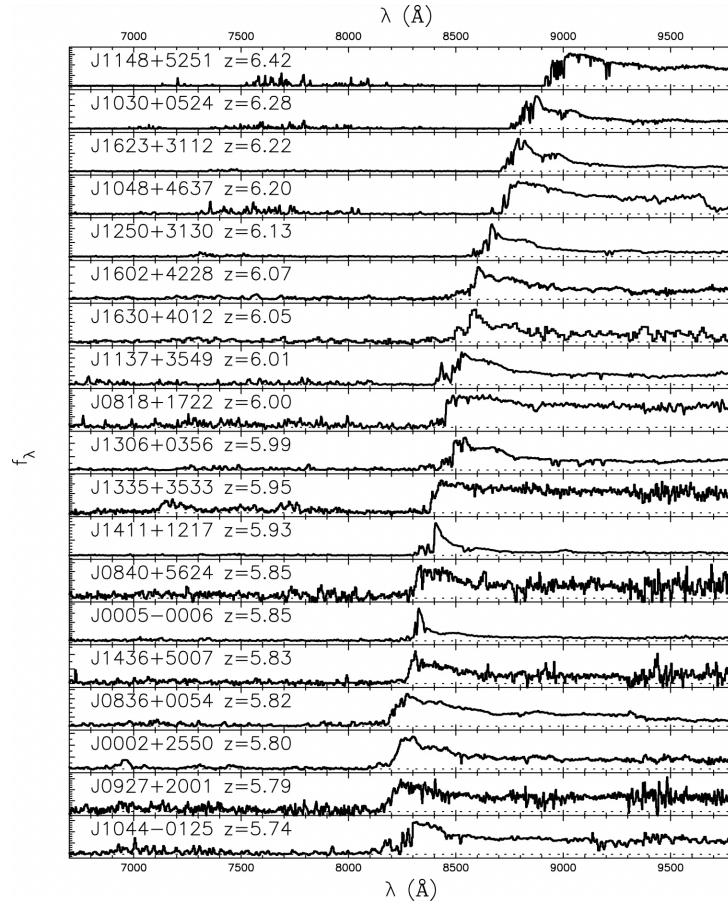
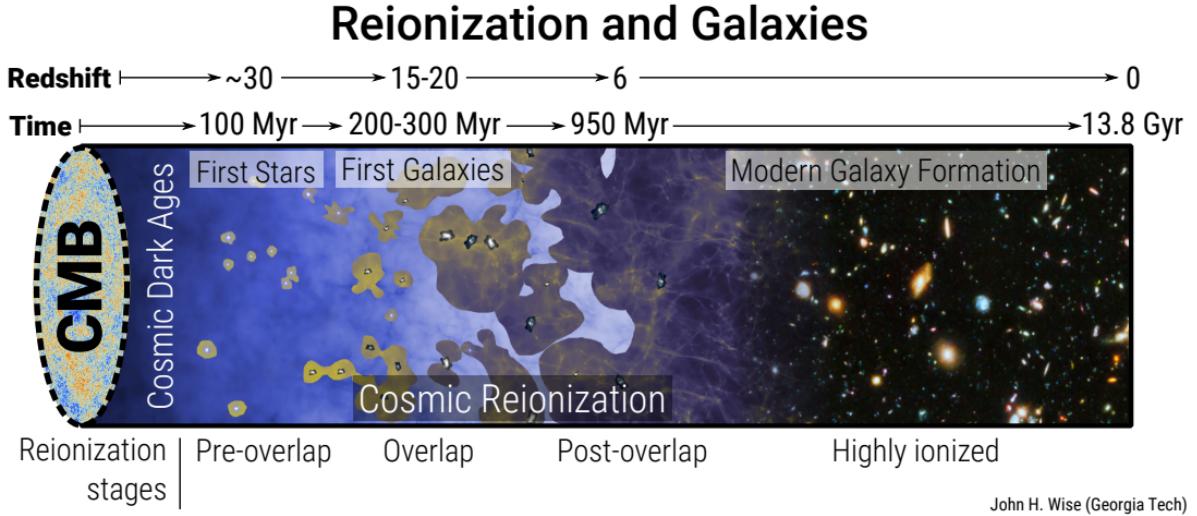


Figure 1.3: Constraining the Evolution of the Ionizing Background and the Epoch of Reionization with $z \approx 6$ Quasars. II. A Sample of 19 Quasars – Fan and Strauss et al [2]

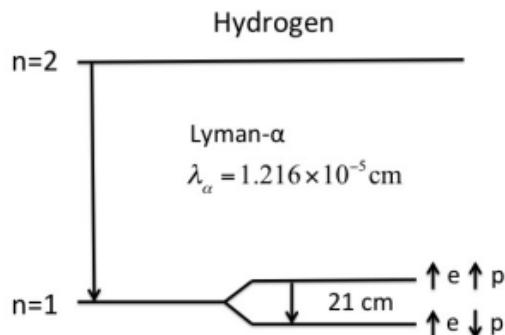
The Epoch of Reionization (EoR) is a crucial phase in cosmic history, marking a significant transition in the evolution of the universe. During this period, **the Intergalactic Medium (IGM)**, primarily composed of neutral hydrogen (HI), began to undergo ionization due to the influence of high-energy photons emitted by various sources. The EoR represents a pivotal moment when the universe’s first stars, galaxies, and quasars started to emit ionizing

radiation, dispelling the cosmic darkness that prevailed earlier. Comprehending the intricate details of the EoR is essential for understanding the formation of the cosmic structure and the emergence of luminous objects within it. This epoch holds critical significance because it sheds light on the timing and nature of the first luminous sources, provides context for interpreting high-redshift observations affected by intervening IGM, and serves as a test for dark matter and galaxy formation models, which must produce sources consistent with reionization constraints.



1.2 Observations & Challenges :

Several observations suggest that a substantial portion of reionization occurred between redshifts (z) approximately 6 and 8. Damping wings in quasar spectra at $z \geq 7$ indicate that the IGM was still largely neutral at those redshifts (Mortlock et al. [8]; Greig et al. [9]; Banados et al. [10]; Greig, Mesinger, and Banados [11]; Davies et al. [12]; Wang et al. [13]). Galaxy surveys infer the persistence of significant neutral regions in the universe at $z \approx 7$ - 8 based on the fraction of UV-selected galaxies exhibiting Lyman-alpha emission (Morales et al. [14]; Jung et al. [15]; Mason et al. [16]). Measurements of the Cosmic Microwave Background (CMB) also suggest a midpoint around $z \approx 8$ (Planck Collaboration 2018 – Aghanim et al.[17]). The thermal history of the IGM down to $z \approx 5$ further supports the notion that a substantial portion of reionization took place at $z \approx 7$ -8. (Gaikwad et al. [18]; Boera et al. [19])



Indirect observations with Lyman- α line have provided valuable insights into the EoR, but they have certain limitations when it comes to addressing crucial questions about this epoch. These outstanding issues encompass the exact duration and timing of reionization, the characteristics of the primary ionizing sources, the relative contributions of different source types to the ionizing photon budget, and the typical sizes and distribution patterns of ionized bubbles. While these indirect observations have been informative, they do not provide complete answers to these fundamental questions, highlighting the need for more comprehensive investigations and data to further our understanding of the EoR.

$$\tau_{\text{GP}} = \frac{\chi_\alpha n_{\text{HI}}(z) c}{H(z)\nu_\alpha} \approx 3 \times 10^5 \bar{x}_{\text{HI}} \left(\frac{1+z}{7} \right)^{3/2}.$$

Figure 1.4: Expression for Lyman- α optical depth of the IGM [20]. RHS of this equation makes optical depth highly sensitive to the presence of neutral hydrogen. Even a 0.001 \bar{x}_{HI} is enough to saturate the Lyman- α optical depth which limits our understanding to lower neutral fractions of EoR.

The HI 21-cm signal, predicted by Dutch astronomer H. van de Hulst in 1944, then directly observed by E. M. Purcell and his student H. E. Ewen in 1951 [21], holds significant promise as a direct means of investigating these critical aspects of this epoch. This signal emerges when the spin states of the electron and proton in the ground state of HI transition from parallel to anti-parallel configurations. As a result, it serves as a direct probe of the distribution of HI within the IGM, offering the potential to address key questions related to the EoR. This signal, measured relative to a background radiation source, such as the CMB, can be seen in either emission or absorption depending on both the thermal and ionisation states of the IGM. By directly measuring this signal at different cosmic epochs along with its spatial distribution in the IGM, the HI 21-cm signal, in principle, can trace the history of reionization as a three dimensional movie of the early universe, shedding light on the timing and progression of this pivotal cosmic event. For detailed review on the potential of 21cm line in Cosmology, see Furlanetto, Oh, and Briggs [20]; Pritchard and Loeb [22].

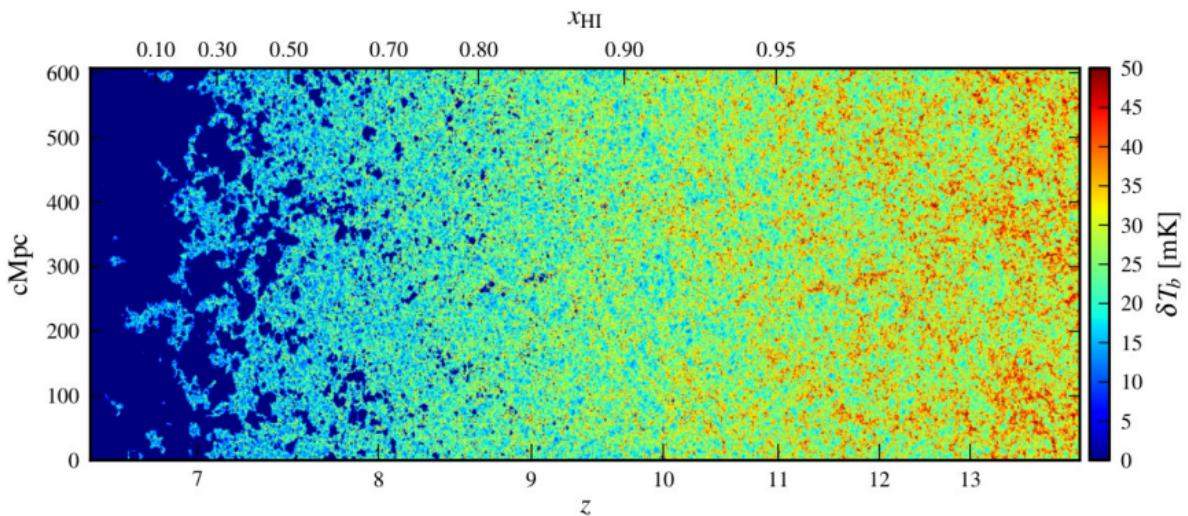


Figure 1.5: How an evolving 21-cm signal tomographic image is expected to look like. Y dimension is along the sky plane & X dimension is along the radially outward direction from us. Image from “Light cone effect on the reionization 21-cm signal II: Evolution, anisotropies and observational implications” - Datta et al. [3].

A number of radio interferometer experiments such as the uGMRT (Gupta et al. [23]), LOFAR (Haarlem et al. [24]), MWA (Tingay et al. [25]; Wayth et al. [26]), PAPER (Parsons et al. [27]), HERA (DeBoer et al. [28]) are specifically designed to detect the statistical fluctuations in the signal in the Fourier domain and constrain its power spectrum. Some of them have dedicated a considerable amount of their observing time to achieve a statistically significant detection of 21cm signal from the EoR. However, presence of strong foreground emissions ([29],[30], [31], [32]) in the same frequency range that are several orders of magnitude stronger than the expected signal and the signal distortions due to the system noise have made such detections very unlikely so far.

$$\delta T_b \approx 27x_{HI}(1 + \delta_b) \left(\frac{\Omega_b h^2}{0.023} \right) \left(\frac{0.15}{\Omega_m h^2} \frac{1+z}{10} \right) \left(\frac{T_S - T_R}{T_S} \right) \left[\frac{\delta_r v_r}{(1-z)H(z)} \right] mK \quad (1.1)$$

The power spectrum fully characterizes the statistical properties of a field when dealing with a Gaussian random field. However, in the case of the 21cm signal from 1.1, various factors like structures arising from gravity-induced effects in the underlying matter density field, as well as the evolving sizes of ionized regions and their interactions with each other, contribute to a **substantial degree of non-Gaussian behavior in the 21-cm signal**. For this reason, the spherically averaged power spectrum does not fully describe its statistical properties. This non-Gaussianity can be studied using higher order statistics such as onepoint skewness and kurtosis (Watkinson and Pritchard [33]) or the full bispectrum and trispectrum (Watkinson et al. [34], Mondal et al. [35]). However, such statistics are known to be difficult to interpret physically. **With the forthcoming Square Kilometer Array (SKA) and with its expected heightened sensitivities for capturing tomographic images of the EoR 21cm signal (Mellema et al. [5]; Koopmans et al. [36]) , there arises a strong incentive for the development of novel analysis techniques.** These techniques must aim to adeptly extract both Gaussian and non-Gaussian morphological features from the tomographic images, ensuring that the wealth of data expected from SKA can be effectively harnessed for a deeper understanding of the Epoch of Reionization (EoR).

Wavelet Scattering Transform (WST) is a technique that is relatively new in the literature of EoR and cosmology. Sihao Cheng ([37],[38],[39]) conducted an extensive analysis of the utility of the scattering transform in the domains of astrophysics and cosmology. His research illuminated the fact that the **scattering transform occupies a favorable intermediate position, bridging the analytical gap between the power spectrum and convolutional neural networks (CNNs). This positioning makes it particularly advantageous for the extraction of information from physical fields characterized by non-Gaussian properties, all while preserving the interpretability of its results.** This attribute of interpretability within the scattering transform proves invaluable in establishing meaningful connections between the WST coefficients and the distinctive features embedded within theoretical models.

The application of Wavelet Scattering Transform (WST) on simulated 21cm images represents a novel approach that has been explored in previous research by Greig, Ting, and Kaurov ([40], [41]). In their work, WST analysis was performed on various EoR models – Fiducial model, Cold reionization model, a model driven by large haloes and an extended reionization model. Their research demonstrated the remarkable sensitivity of WST to non-

Gaussian information within the 21-cm signal 2D images. A comparison between the 2D WST coefficients and 3D 21-cm Power Spectrum (PS) through Fischer matrix analysis revealed that 2D WST outperforms the PS in terms of constraining model parameters and extracting astrophysical insights. Even when instrumental effects are introduced onto the simulated images, WST remains capable of distinguishing between different EoR models. Greig and his team's work has opened up exciting possibilities for leveraging WST in the study of the EoR.

In my thesis work, we extend the application of WST to various EoR source models in an effort to determine the primary sources responsible for driving reionization. To enhance our understanding, we begin by applying WST to the fiducial model to examine how the coefficients respond to the evolving morphological features of ionized regions. Subsequently, we employ WST on other source models and compare them to the Fiducial case, scrutinizing the similarities and differences between the fiducial model and the alternative models.

CHAPTER 2

21-CM BRIGHTNESS TEMPERATURE MAPS OF VARIOUS REIONIZATION SOURCE MODELS :

My supervisor Dr. Majumdar had worked on simulating reionization scenarios consisting of different combination of various sources of ionizing photons. Here I briefly mention the important simulation parameters considered, describe these different source types and the different reionization scenarios. For more details on the simulation techniques used and reionization models, please see Majumdar et al. [4]

2.1 Simulation Parameters :

- Cosmological parameters used were taken from WMAP five year data release $h = 0.7, \Omega_m = 0.27, \Omega_\lambda = 0.73, \Omega_b h^2 = 0.0226$
- Size of simulation volume = $500h^{-1} = 714$ Mpc (comoving) along each side.
- 6912^3 particles of mass $4.0 \times 10^7 M_\odot$ on a 13824^3 mesh, which was then down-sampled to a 600^3 grid for modelling the reionization.
- The minimum halo mass considered for reionization simulations is $2.02 \times 10^9 M_\odot$. For each redshift output of the N-body simulation, haloes were identified using a spherical overdensity scheme.
- The resolution of this grid is 1.19 Mpc.

2.2 Reionization Source models considered in my thesis work :

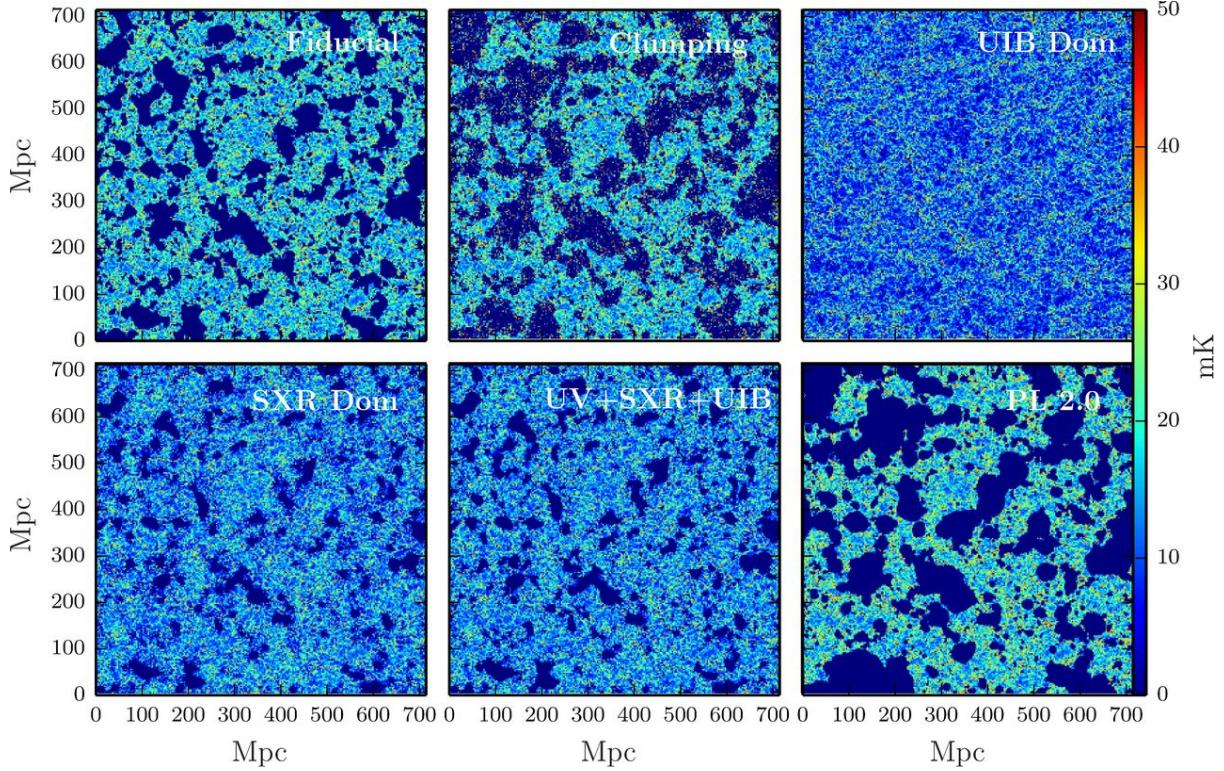


Figure 2.1: Redshift space brightness temperature maps for six of our reionization scenarios. The line of sight is along the y-axis. The mass averaged neutral fraction for all the panels is $\bar{x}_{HI} = 0.5$. Image from “Effects of the sources of reionization on 21-cm redshift-space distortions” - Majumdar et al. [4].

What are the different source models considered in this project? What are the differences (sources) and similarities (ionization history) in the incorporated physics of these different models ?

2.2.1 UV Photons from Galaxies :

- Galaxies residing in the collapsed dark matter halos are assumed to be the major sources of ionizing photons. Total number of ionizing photons contributed by a halo of mass M_h which is hosting such galaxies is assumed to be:

$$N_\gamma(M_h) = N_{ion} \frac{M_h \Omega_b}{m_p \Omega_m}$$

- Here, N_{ion} is a dimensionless constant that represents the number of photons entering the IGM per baryon in collapsed objects & m_p is the mass of proton or hydrogen atom. We assume that all ionizing photons produced by such sources are in the ultraviolet part of the spectrum. Thus, they only effect the IGM locally, up to a distance limited by their mean free path.

2.2.2 Uniform Ionizing Background :

- If sources of hard X-ray photons such as Active Galactic Nuclei (AGNs) or X-ray binaries were common in the early universe, they could give rise to a more or less uniform ionizing background. Hard X-ray photons would easily escape their host galaxies and travel

long distances before getting redshifted into the UV regime due to cosmic expansion and start ionizing hydrogen.

- These types of sources are modelled as a completely uniform ionizing background that provides the same number of ionizing photons at every location. The extreme case of a 100% contribution of the ionizing photons from this kind of background would lead to a global "outside-in" reionization.

2.2.3 Soft X-Ray photons :

- For simplicity, we assume that all the halos that we have identified as sources will produce soft X-ray photons and all of these photons will have the same energy (200 eV). the case when major portion of photons are in ultraviolet, a significant contribution by these soft X-ray photons may lead to a more homogenous reionization scenario.

2.2.4 Power Law mass dependent efficiency :

- For UV source type, we have assumed the number of UV photons generated by a galaxy is proportional to its host halo mass. But it can also be assumed that the number of photons contributed by a halo follows a power law such as

$$N_\gamma(M_h) \propto M_h^n$$

- We consider two cases where the power law index is 2 & 3. Higher mass halos produce relatively more ionizing photons giving rise to fewer but larger ionized regions.

2.2.5 Inhomogenous Clumping Scenario :

- In all of our reionization scenarios except Clumping model, it was assumed that the rate of recombination is uniform everywhere in the IGM. But in reality, the recombination rate is expected to depend on the density of the ionized medium. Dense structures with sizes on the order of a few kpc, mostly unresolvable in this type of simulations, are expected to boost the recombination rate significantly and this increases the number of photons required to complete reionization. They will also give rise to self-shielded regions like Lyman limit systems. Effects of these self-shielded regions in the IGM were included by using eq 15 in Choudhury, Haehnelt, and Regan [42].
- Although this approach somewhat overestimates the impact of non-uniform recombinations (due to the coarse resolution of the density fields that were used in semi-numerical models) it can still serve to illustrate their effect on the reionization topology.

Reionization scenario	UV	UIB	SXR	PL <i>n</i>	Non-uniform recombination
Fiducial	100%	–	–	1.0	No
Clumping	100%	–	–	1.0	Yes
UIB dominated	20%	80%	–	1.0	No
SXR dominated	20%	–	80%	1.0	No
UV+SXR+UIB	50%	10%	40%	1.0	No
PL 2.0	–	–	–	2.0	No
PL 3.0	–	–	–	3.0	No

Table 1. The relative contribution from different source types in our reionization scenarios. UV=Ultra-violet; UIB=Uniform ionizing background; SXR=Soft X-ray background; PL=Power-law.

Figure 2.2: The relative contribution from different source types in our reionization scenarios. UV=Ultra-violet; UIB=Uniform ionizing background; SXR=Soft X-ray background; PL=Power-law. Table from “Effects of the sources of reionization on 21-cm redshift-space distortions” - Majumdar et al. [4].

CHAPTER 3

WAVELET SCATTERING TRANSFORM (WST)

In the pursuit of comprehending the fundamental laws governing the natural world, physicists are consistently confronted with the formidable task of deriving suitable accurate and precise insights from data, whether it originates from sensor measurements or computer-generated simulations. Diverse sets of data analysis methodologies are employed across various scientific disciplines, their selection often driven by specific characteristics of the subject of inquiry, such as inherent symmetries and invariants, as well as intrinsic features of the data, including noise levels. At times, however, the choice may be influenced by mere convenience or established convention. Among these analytical tools, a handful find universal applications, notably the power spectrum and correlation functions. Unfortunately, as second-order moment analyses, their applicability is constrained by their inability to effectively capture the nuances of non-Gaussian structures. Such complex structures, of considerable significance, show up across a broad spectrum of domains within the realm of scientific investigation. Fig 3.1 shows a comparison of non-Gaussian and Gaussian field textures with same power spectrum. To analyse such fields, one needs to go beyond Power Spectrum.

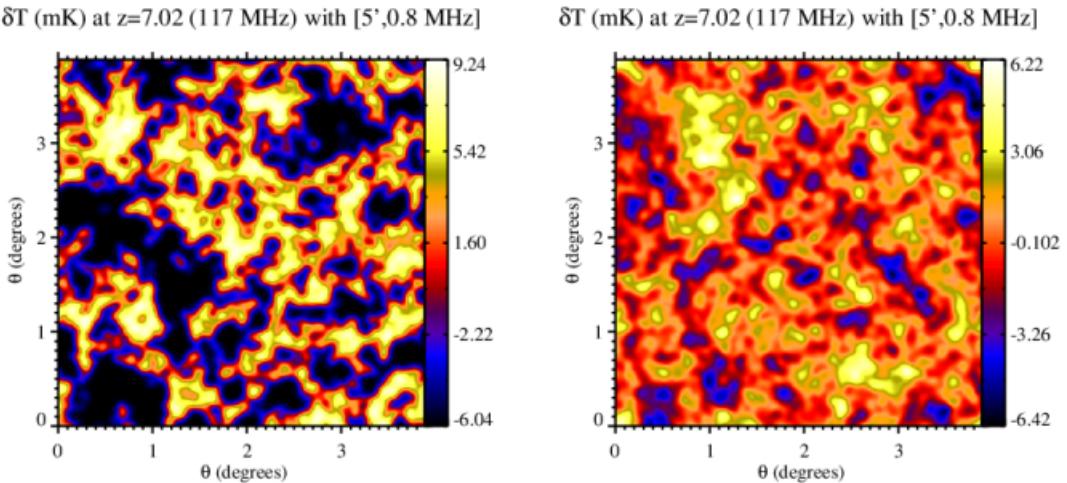


Figure 3.1: Two fields with visibly different textures but same power spectrum. The left panel shows a simulated 21-cm brightness temperature image during EoR which is highly non-Gaussian. The right panel is a Gaussian Random Field. Image from "HI tomographic imaging of the Cosmic Dawn and Epoch of Reionization with SKA" - Mellema et al. [5]

Over the past decade, a novel type of estimator has garnered increasing prominence across diverse scientific disciplines—namely, convolutional neural networks (CNNs). This innova-

tive paradigm has exhibited remarkable efficiency in extracting specific forms of information from pixelated data. However, it remains characterized by properties that are not yet comprehensively elucidated. Unlike traditional mathematical tools, neural networks suffer from a lack of transparency, stable mathematical properties, and interpretability, all of which hold significant importance in the context of scientific research. With increasing depth, neural networks achieve a spectacular degree of expressivity, capable of capturing the incredibly complex sets of variable characteristics created by the biological world, ranging from cat look to human voice. However, when we study a wide range of physical areas, we find ourselves in a different regime. We work with data that has a far lower level of complexity than the standard images or sounds addressed in ‘deep learning’ applications.

In his comprehensive doctoral dissertation and following work, Sihao Cheng [37] conducted an extensive analysis of the utility of the scattering transform in the domains of astrophysics and cosmology. His work [[38],[39]] illuminated the fact that the scattering transform occupies a favorable intermediate position, bridging the analytical gap between the power spectrum and convolutional neural networks (CNNs). This positioning makes it particularly advantageous for the extraction of information from physical fields characterized by non-Gaussian properties, all while preserving the interpretability of its outcomes. Notably, this attribute of interpretability within the scattering transform proves invaluable in establishing meaningful connections between the obtained results and the distinctive features embedded within theoretical models.

The scattering transform was first introduced as a signal processing tool by Mallat [43]. Since then, many people working across different fields including audio signal processing, image classification, texture classification, intermittency in turbulence and finance data, material science, graph-structured data, and astrophysics & cosmology have realised this technique’s effectiveness. In several of these applications, the scattering transform reached state-of-the-art performance compared to the CNNs in use at the time.

3.1 Wavelets:

Wavelets are mathematical constructs that represent short, wavelike oscillations characterized by diverse frequency ranges and shapes. Their exceptional flexibility allows them to embody a wide array of forms, enabling researchers to discern and align specific wave patterns within virtually any continuous signal. This adaptability has led to a transformative impact on various fields, including image processing, communication, and the analysis of scientific data streams.

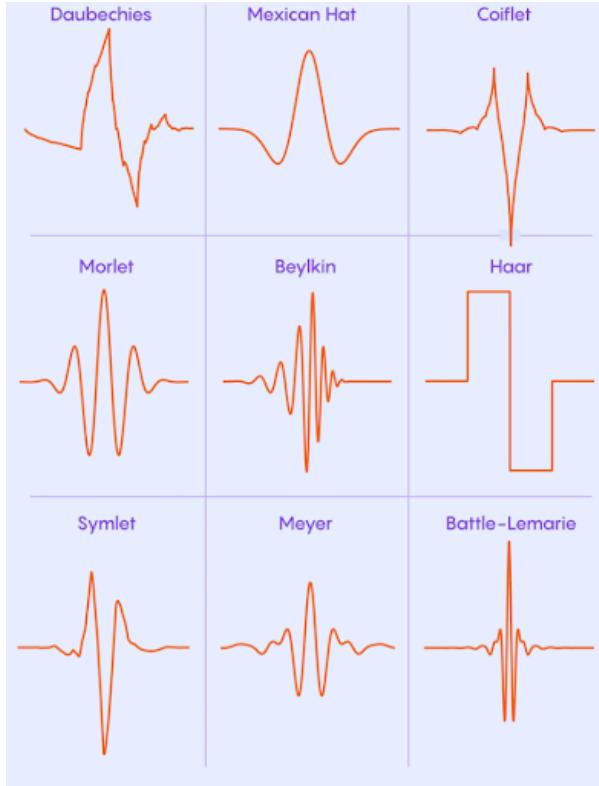


Figure 3.2: Different types of Wavelets. Image from "How Wavelets allow researchers to transform and understand data" - Samuel Velasco [44]

The key strength of wavelets lies in their ability to capture intricate details in signals by representing them at different scales and resolutions. Unlike traditional Fourier analysis, which decomposes a signal into sinusoidal components with fixed frequencies, wavelet analysis adapts to the varying frequency content of a signal. This adaptability makes wavelets particularly well-suited for identifying and characterizing complex wave phenomena in diverse applications. The versatility of wavelets has ushered in a new era in signal processing and data analysis, empowering researchers to gain deeper insights into the underlying structures of signals across different domains.

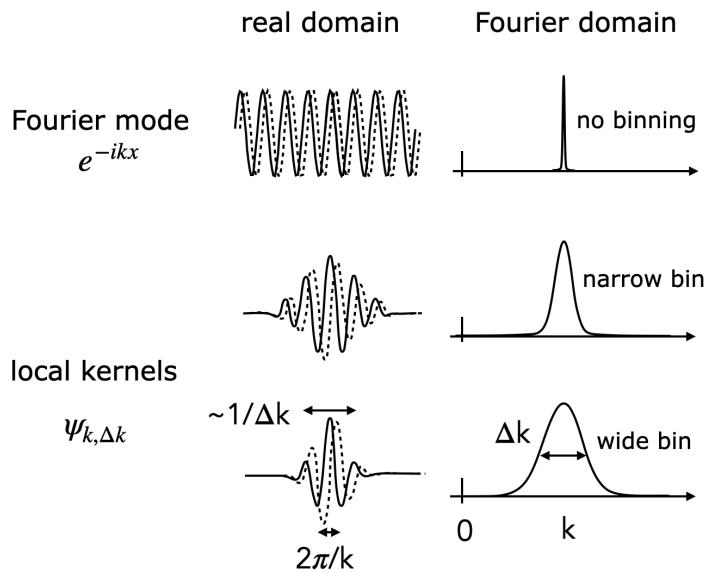


Figure 3.3: Comparing fourier filters & wavelet filters [37]

The wavelet transform, functioning as a convolution operation, embodies characteristics that lie between Fourier and pixel-based decompositions. **Wavelet filters demonstrate partial locality in both real space and Fourier space**, responding to translations smaller than the wavelet scale akin to Fourier coefficients (pure phase shift), while larger displacements are akin to pixel behavior. These localized, multi-scale kernels aren't arbitrary choices but are necessitated by the properties of fields in the physical world. Similar wavelet-like kernels have been observed in animal vision systems and the initial layers of Convolutional Neural Networks (CNNs), irrespective of the training set. This implies a fundamental role played by wavelets in processing various types of signals.

The physical world often involves signals and fields comprising sparse superpositions of local structures like peaks, edges, filaments, and transient features in time series signals. Wavelet representations, notably the outcomes of wavelet convolution, reveal these structures in a clearer and simpler manner compared to Fourier or pixel representations. Since the 1990s, it has been acknowledged that wavelets offer a set of bases that enable a sparse representation, contributing to a more streamlined representation of signals.

A frequently employed option is to utilize wavelets with dyadic scales, where the width in Fourier space is approximately equal to its central frequency, and these wavelets can be sampled at factor-of-two scales. This choice offers a favorable balance between the requirement to distinguish scales and the ability to investigate scale interactions within the scattering framework. **When the field extends beyond one dimension, a uniform tiling of orientations is also applied to capture the anisotropic features in the data.**

3.2 Scattering Transform :

Now I outline the formulation of the scattering transform as presented by Mallat [43] in 2012. Although I specifically focus on the 2D case, the approach can be readily extended to any other dimensionality. The scattering transform extracts information from an input field by iteratively applying two operations: a wavelet convolution and a modulus. This process generates a group of new fields, and the expected values of these fields serve as the scattering coefficients, characterizing the statistical properties of the original field. The hierarchical structure, utilization of localized convolution kernels, and incorporation of non-expansive nonlinear operators are key elements shared with the architecture of Convolutional Neural Networks (CNNs).

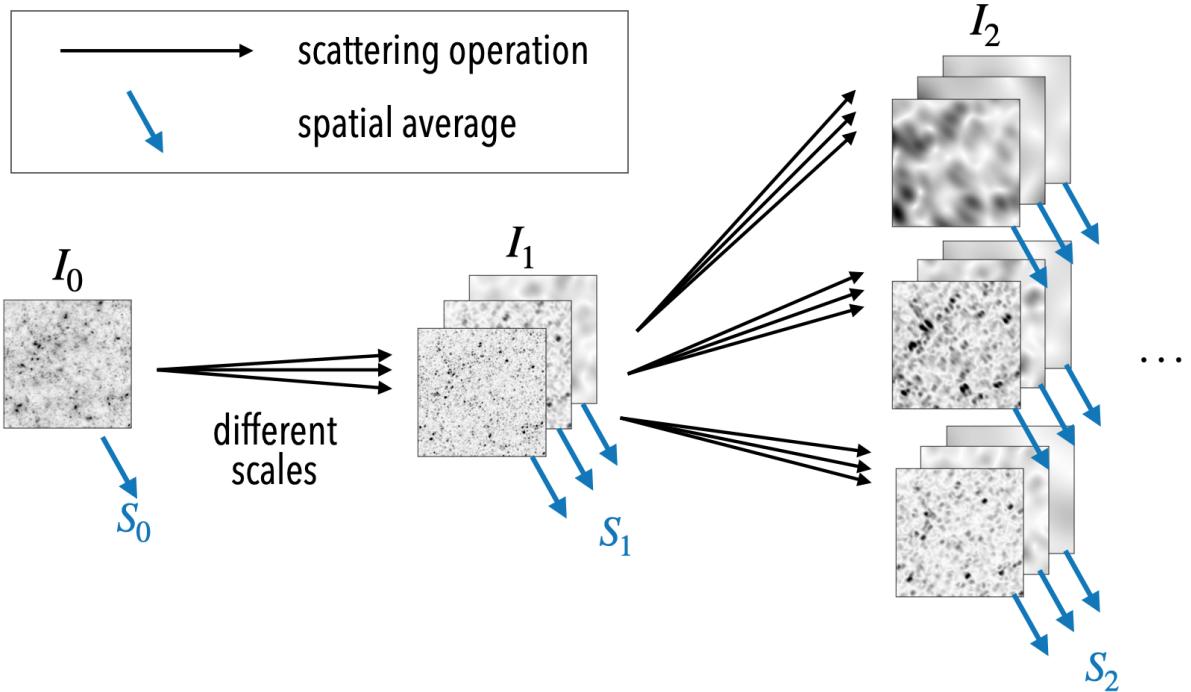


Figure 3.4: Scattering Transform and its operations visualised. Image from “Astrophysics and cosmology with the scattering transform” - [37].

Given an input field $I_0(x, y)$, the scattering transform produces a set of 1^{st} -order fields $I_1(x, y)$ by convolving it with a family of wavelets Ψ^{j_1, l_1} and subsequently taking the modulus. In general, a family of wavelets covers the whole Fourier space. They all have the same shape but different sizes and orientations, labelled by j and l respectively for a total of J sizes and L orientations.

$$I_1 = |I_0 \circledast \Psi^{j_1, l_1}|$$

where I_1 represents a group of fields labelled by the wavelet index j_1, l_1

Once the 1^{st} -order fields are generated, the same process can be iterated to create 2^{nd} -order fields $I_2(x, y)$

$$I_2 = |I_1 \circledast \Psi^{j_2, l_2}| = ||I_0 \circledast \Psi^{j_1, l_1}| \circledast \Psi^{j_2, l_2}|$$

where I_2 represents a group of fields labelled by two sets of wavelet index j_1, l_1 & j_2, l_2 .

In the case where the input field I_0 is homogeneous, the generated fields I_n will also remain homogeneous. As a result, the expected values of their intensity can be utilized as translation-invariant descriptors of the input field.

$$\begin{aligned} S_0 &= \langle I_0(x, y) \rangle \\ S_1^{j_1, l_1} &= \langle I_1(x, y) \rangle = \langle |I_0 \circledast \Psi^{j_1, l_1}| \rangle \\ S_2^{j_1, l_1, j_2, l_2} &= \langle I_2(x, y) \rangle = \langle ||I_0 \circledast \Psi^{j_1, l_1}| \circledast \Psi^{j_2, l_2}| \rangle \end{aligned}$$

When dealing with an isotropic field, particularly relevant in cosmology, the scattering coefficients S_n can be further condensed. To build isotropic statistics, we average over all orientation indices, leading to a reduction in the number of coefficients by a factor of L^n . This process results in a more concise and robust set of statistical descriptors. Therefore, we define our reduced scattering coefficients as:

$$s_0 = S_0$$

$$s_1^{j_1} = \langle S_1^{j_1, l_1} \rangle_{l_1}$$

$$s_2^{j_1, j_2} = \langle S_2^{j_1, l_1, j_2, l_2} \rangle_{l_1, l_2}$$

where S_n represent the standard scattering coefficients, s_n represent our reduced coefficients, and $\langle \rangle_l$ denotes an average over orientation indices.

CHAPTER 4

RESULTS

4.1 WST Analysis of different reionization source models :

- Simulation volume = 600^3 grids. Simulation box size = $500h^{-1}$ Mpc ≈ 714 Mpc

4.1.1 Choice of Wavelet filter – Morlet Wavelet :

Within the framework of the scattering transform, there is flexibility in choosing the family of wavelets. But real wavelets have drawbacks such as –

1. **Lack of Shift Invariance** - small shifts in the input signal can cause major variations in the distribution of energy between DWT coefficients at different scales.
2. **Poor directional selectivity for diagonal features**, because the wavelet filters are separable and real.

Complex-valued wavelets are shown by Kingsbury [45] to satisfy both these properties which are important for robustness & completeness of the statistic. Among the various options in complex-valued wavelets, Morlet wavelets offer a balanced approach, providing locality in both real space and Fourier space. Morlet wavelets essentially have Gaussian profiles. For more details on the mathematical form of the Morlet wavelet and how to generate them, see Appendix A in Cheng et al. [38].

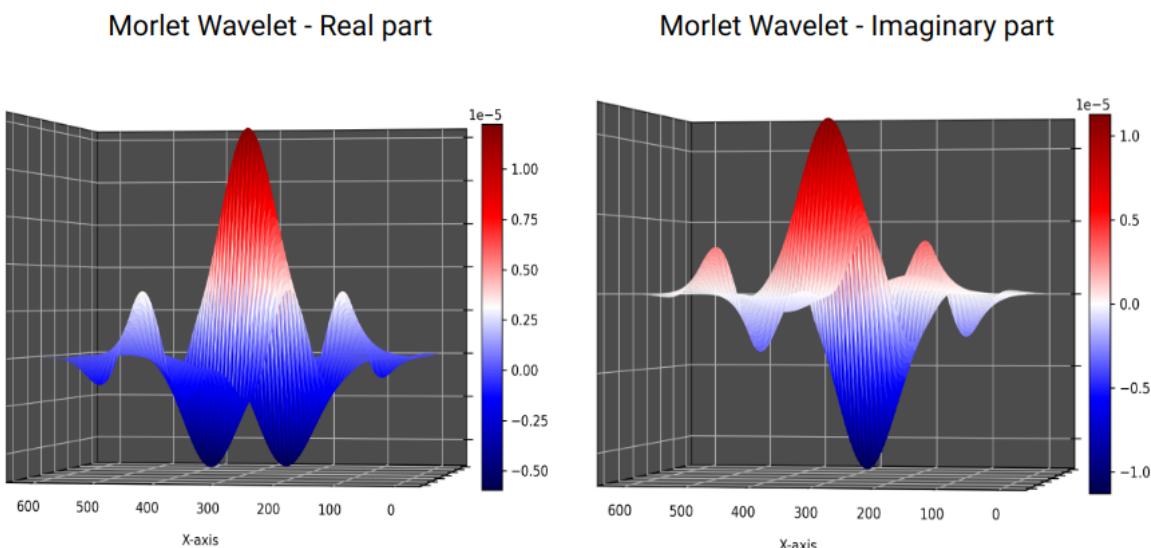


Figure 4.1: Wavelet Filters used in our work

Wavelet Sizes in Mpc	
Wavelet Scale Parameter (j)	Wavelet Size in Mpc (approx)
$j=0$	3
$j=1$	6
$j=2$	12
$j=3$	25
$j=4$	50
$j=5$	101
$j=6$	203
$j=7$	406

Table 4.1: Sizes of the Wavelet filters used in our WST analysis

4.1.2 Fiducial Model :

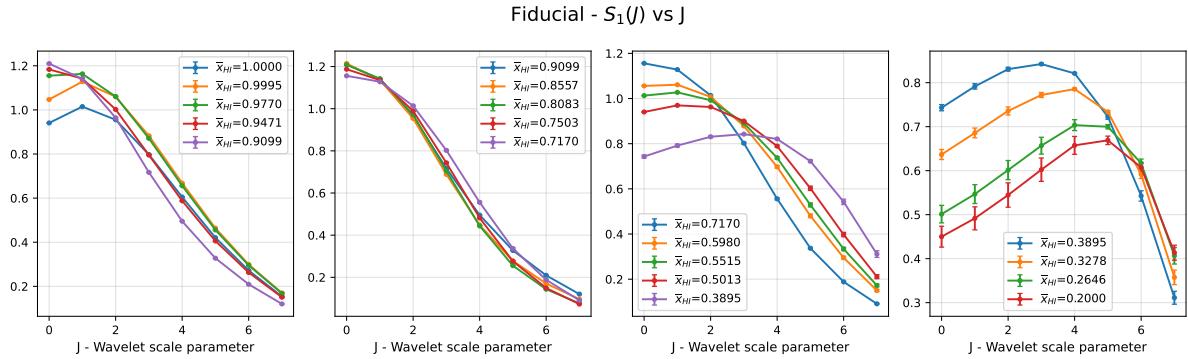


Figure 4.2: Evolution of $S_1(j)$ coefficients vs. J with \bar{x}_{HI} for Fiducial Model

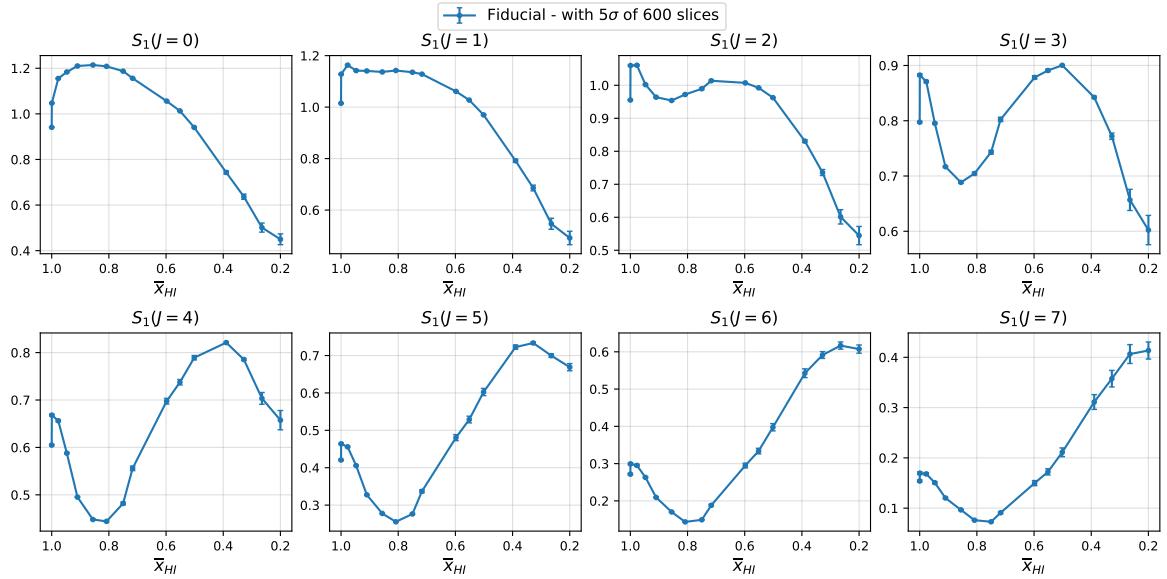


Figure 4.3: Evolution of $S_1(j)$ coefficients with \bar{x}_{HI} for Fiducial Model

In the early stages of the Epoch of Reionization (EoR), fluctuations in the 21cm images are mainly influenced by the interplay between the underlying neutral hydrogen density field

and peculiar velocity-induced redshift space distortions. However, as ionization progresses, particularly in the Fiducial case where local inside-out UV photo-ionization is dominant, a significant shift occurs:

1. **Ionized Region Growth:** Initially, ionized regions start forming and expanding around massive halos. These regions result from a substantial decrease in neutral hydrogen density due to photo-ionization, with regions around more massive halos ionizing first.
2. **Dominance of Ionization Fronts:** Over time, ionization fronts become increasingly influential in shaping the fluctuations of the 21cm signal. Eventually, there comes a crucial stage in the EoR where ionization fronts become the dominant source of fluctuations, marking a significant transition in the primary contributor to signal fluctuations.
3. **Correlation with Ionization Maps:** This shift from fluctuations driven by neutral hydrogen density to those dominated by ionization can be verified by comparing trends in scattering coefficients over time for both 21cm maps and ionization maps generated using excursion set methods. During the period when ionized regions take precedence, a distinct pattern emerges: fluctuations decrease on smaller scales while increasing on larger scales.
4. **Peak Correspondence:** Once ionization becomes the primary driver of fluctuations, the peak in Fig 4.2 aligns with the size of the dominant ionized regions that exert the most influence on the fluctuations in the 21cm signal.

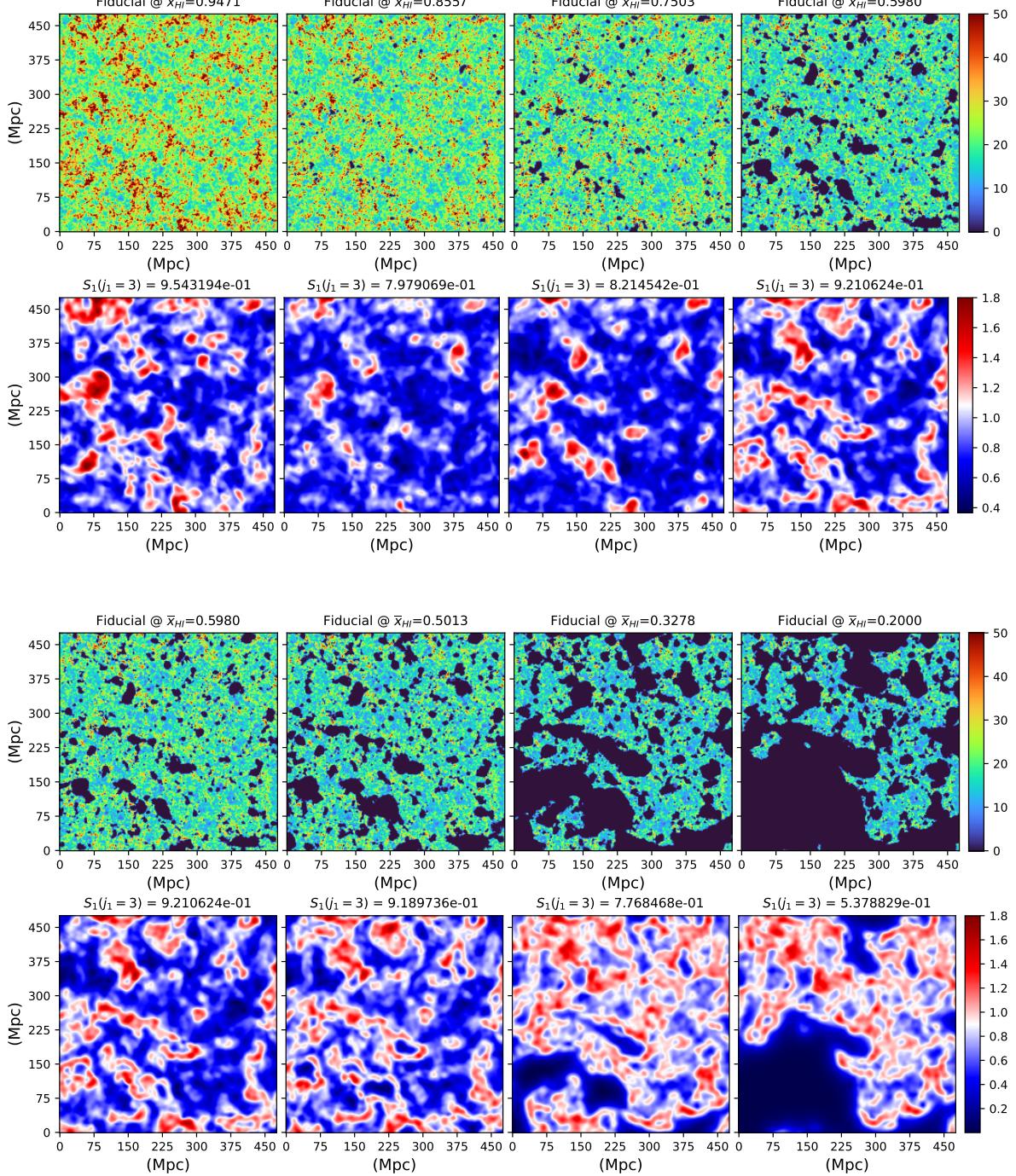
This transition from hydrogen density fluctuations to ionization-dominated fluctuations can also be observed in Fig 4.3. The increasing trend following the dip can be attributed to the transition of dominant sources of fluctuations in the 21cm signal from hydrogen density fluctuations to ionization fronts. As ionization fronts become the primary contributors, their growth over time results in increased fluctuations, explaining the rising trend after the dip in Fig 4.3.

The subsequent decreasing trend following the peak can be interpreted as a transition in the dominant sources of 21cm signal fluctuations at the scale of interest. This transition involves a shift from ionized regions at the scale of interest to larger scales. As ionized regions continue to grow, they eventually surpass the detection range of the wavelet filter at scale j_1 . This triggers the wavelet property of zero filter response, effectively rendering the convolution output nearly zero within these larger-than-scale regions that lack fluctuations, essentially producing a constant zero signal. As more and more of these ionized regions, which were once sources of fluctuations, transform into near-zero response regions, it naturally leads to a decrease in the magnitude of fluctuations. This decrease in signal magnitude is the underlying reason for the declining trend observed in the later stages of the 21cm signal fluctuations.

In Fig 4.3, it is also evident that the initiation of the decreasing trend in the $S_1(j)$ coefficients during the later stages of the Epoch of Reionization (EoR) occurs earlier at lower scales, while it shifts to later times as we examine larger scales. This scale-dependent variation can also be attributed to the dynamics of ionized region growth.

The dominance of the zero filter response of the wavelet filter becomes prevalent when the majority of ionized regions surpass the scale of interest. When ionized regions become

larger than the filter scale at one level, these regions still fall within the range of higher scales. Consequently, there will be ionized regions that contribute to a zero filter response at lower scales but contribute as high fluctuations at higher scales. This temporal transition in the dominance of ionized regions relevant to the scale is detected by the multi-scale wavelet analysis of $S_1(j)$ coefficients, resulting in earlier and later peaks in small-scale and large-scale coefficients, respectively.



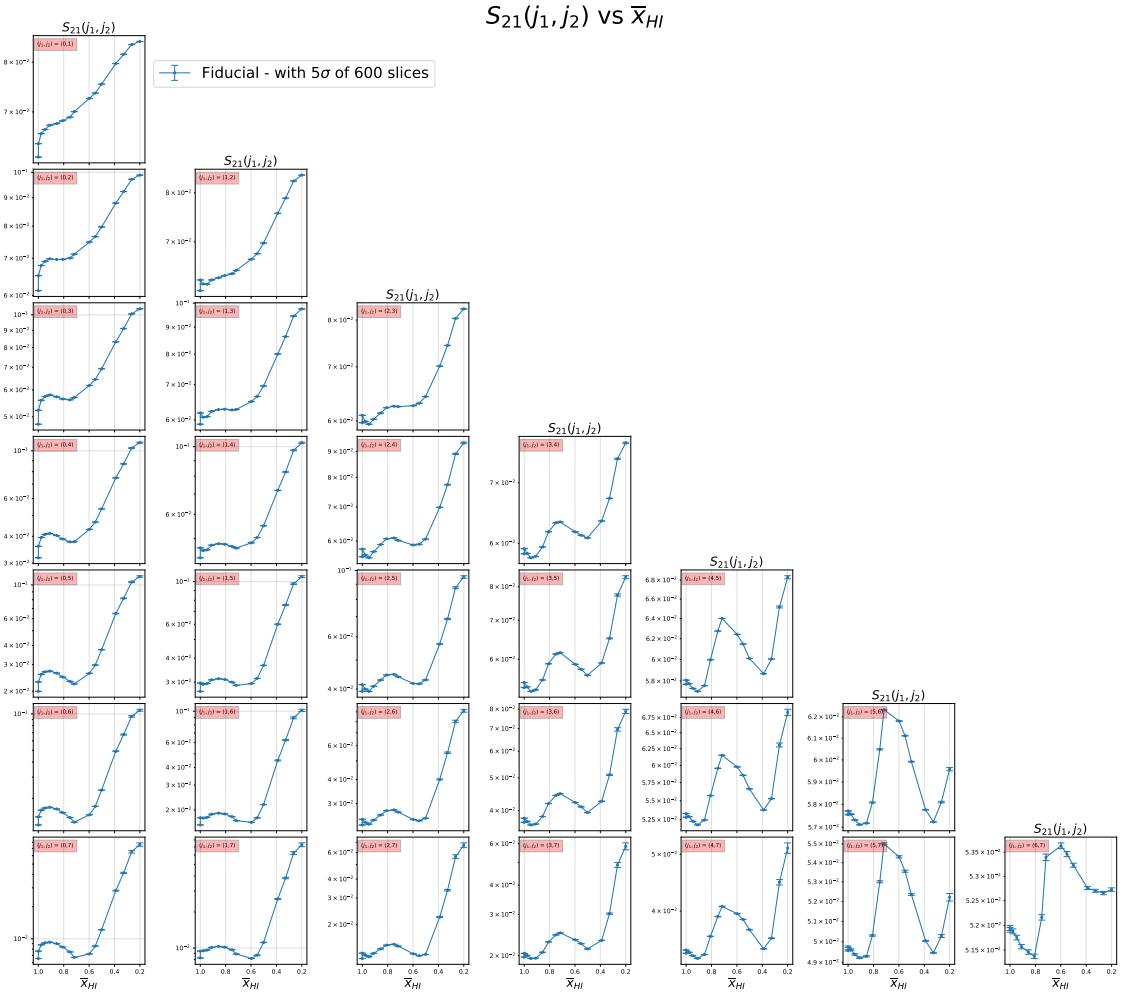


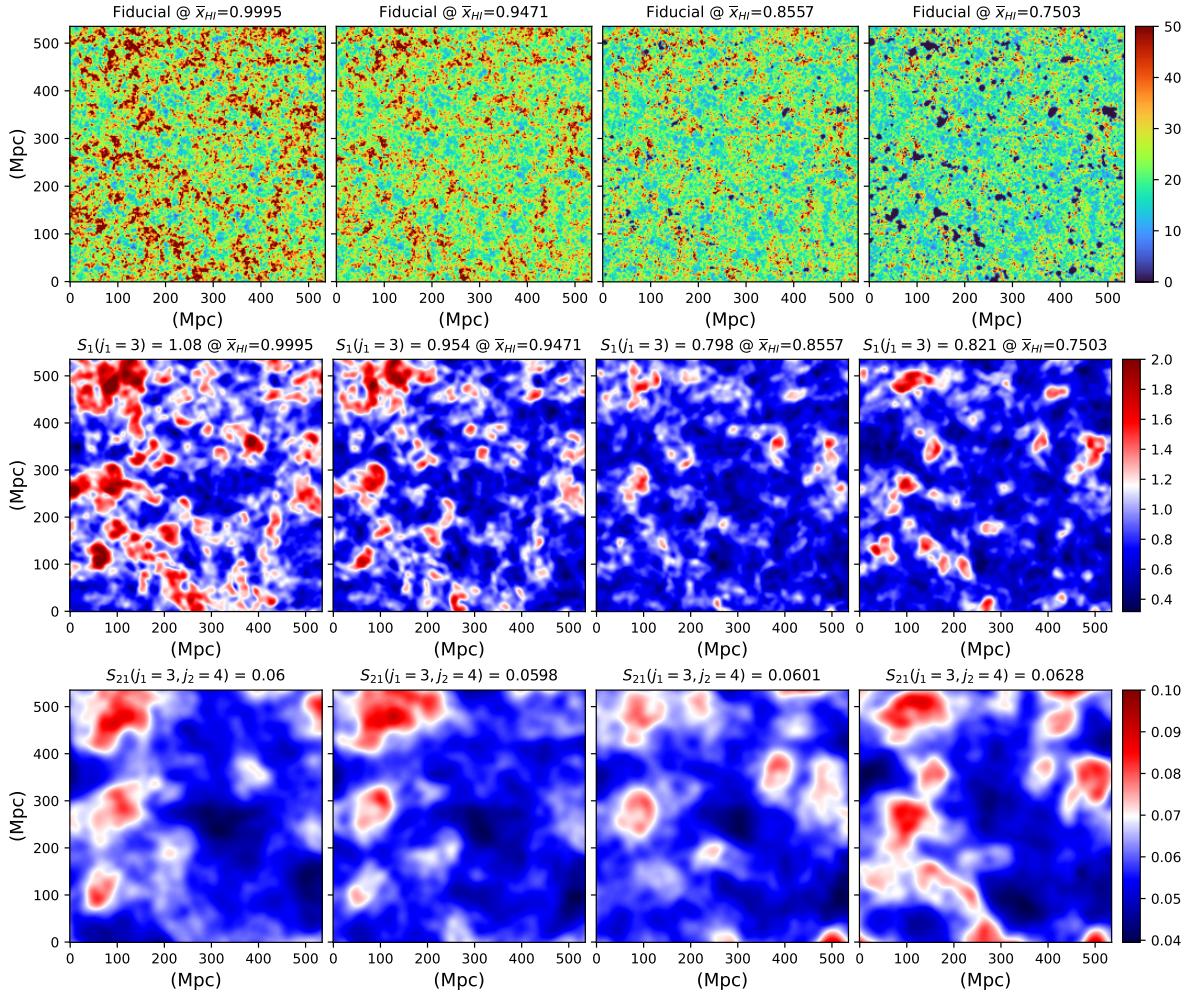
Figure 4.4: Evolution of $S_{21}(j_1, j_2)$ coefficients with \bar{x}_{HI} for Fiducial model

To get 2nd layer coefficients, we do a convolution once more with $j_2 (> j_1)$ wavelet filters on top of the layer 1 convolution output images with j_1 wavelet filters. It measures clustering strength of fluctuations from j_1 scale on j_2 scale. Since the layer 2 input is the convolution output of scattering layer 1, S_2 coefficient are highly correlated with the S_1 coefficient of the j_1 scale. To reduce this correlation, we divide S_2 with S_1 and study this new coefficient – S_{21}

To gain an intuitive understanding of the trends in the coefficient $S_{21}(j_1 = 2, j_2 = 4)$ from Fig 4.4, let's consider a region with a scale of j_2 filter size on the convolved image at scale j_1 .

As the universe evolves, the behavior of this coefficient can be conceptualized in the following way: If the clustering effect of j_1 fluctuations statistically dominates over their divergence effect, then the coefficient increases. In simpler terms, when the ionization regions become the primary contributors to fluctuations at the j_1 scale, there comes a phase where the growth in the sizes of these ionized regions contributes to an increase in the response of the j_2 filter.

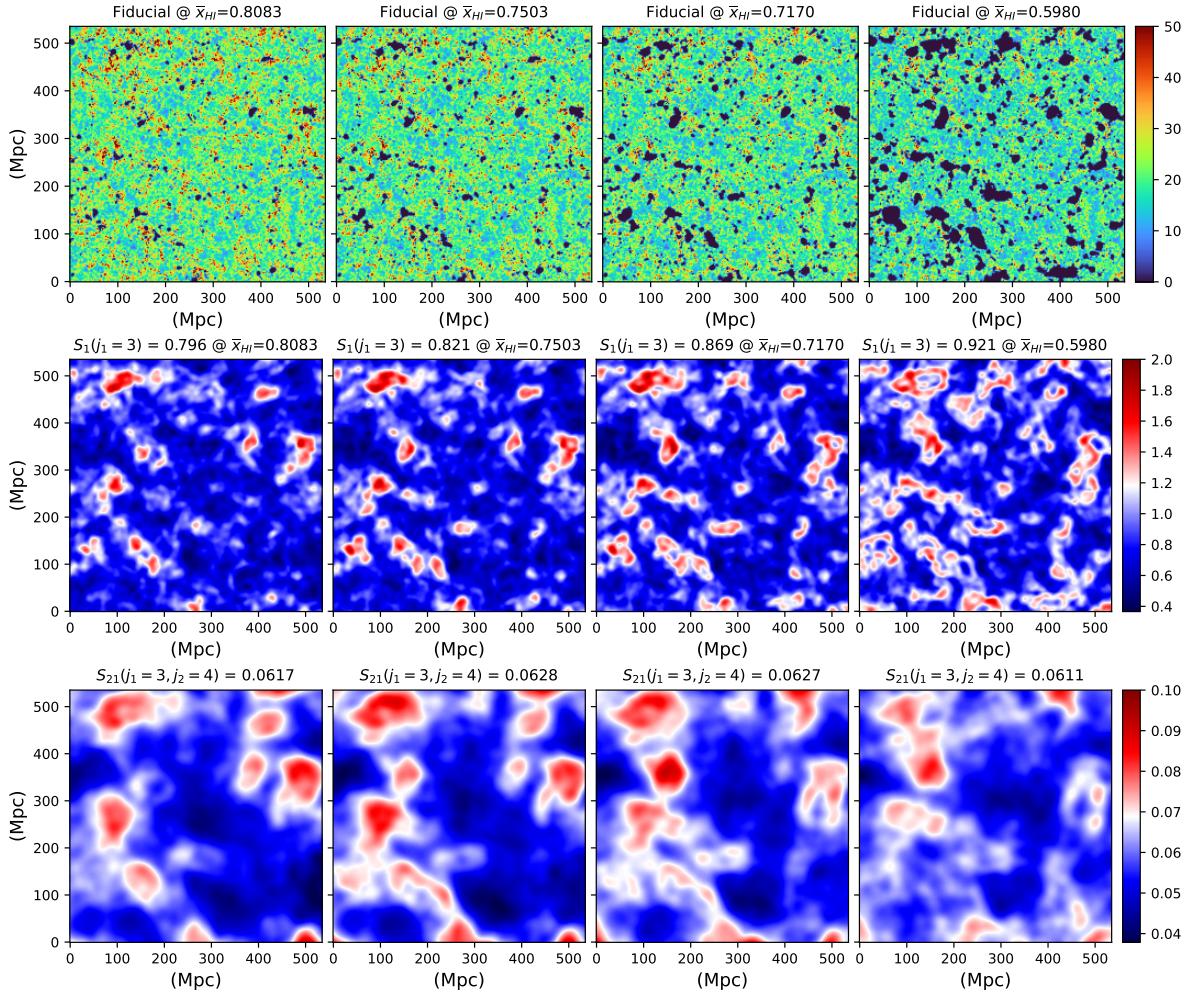
In essence, during this phase, as the ionized regions grow in size and become more abundant, they exert a statistically significant clustering effect on the fluctuations at the j_1 scale. This clustering effect, in turn, enhances the response of the j_2 filter, leading to the observed increasing trend in the coefficient $S_{21}(j_1, j_2)$.



For $S_{21}(j_1 = 3, j_2 = 4)_{Fidu}$, Initially, there is a rapid decline in fluctuations at scale $j_2 = 4$, primarily attributable to the prevalence of dominant fluctuations arising from the neutral hydrogen density field. During this phase, the high-density regions within this field experience a rapid reduction due to preferential UV photo-ionization concentrated in massive halos.

This initial decline is subsequently followed by a period of increasing fluctuations. This upward trend can be comprehended by acknowledging that the dominant fluctuations are currently undergoing a growth phase at the $j_1 = 3$ scale. This behavior exhibited by dominant ionized region fluctuations at the $j_1 = 3$ scale contributes to an escalation in the $S_{21}(j_1 = 3, j_2 = 4)_{Fidu}$ coefficient at the $j_2 = 4$ scale through two distinct mechanisms:

- Within the detection zone of the j_2 filter centered around any high fluctuation location of $j_1 = 3$ scale, the enlargement in the size of fluctuations concurrently augments the response of the j_2 filter.
- Inside the detection zone of j_2 filter centered around any low fluctuation location of $j_1 = 3$ scale, ionization fronts originating from various ionized regions begin to converge due to their expanding dimensions. This convergence leads to heightened fluctuations within the range of the j_2 filter, thus contributing to the observed upward trend.

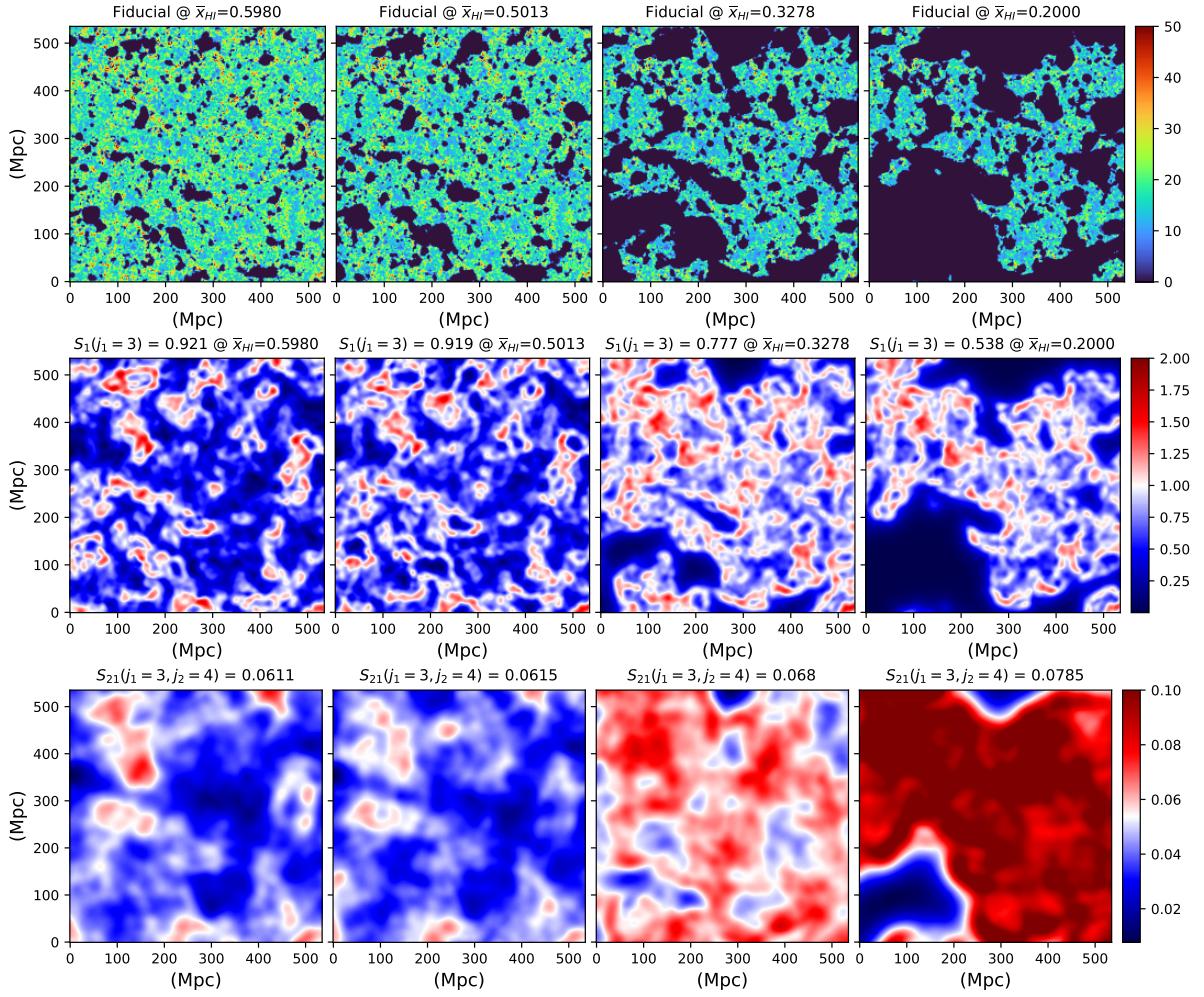


At the local level, specifically at the ($j_1 = 3$) scale, ionization fronts remain the primary contributors to the 21cm signal. However, a notable development occurs as the ionized regions begin to surpass the detection range of the $j_1 = 3$ filter in terms of their size. Consequently, in the convolution output at the $j_1 = 3$ scale, we begin to observe a zero filter response corresponding to these larger-than-scale ionized regions that lack fluctuations, situated in the spaces between the fluctuations associated with ionization fronts.

This novel behavior within the $j_1 = 3$ convolved field gives rise to two key outcomes:

- The stronger fluctuations represented by the ionization fronts at the $j_1 = 3$ scale start to move apart from each other.
- The newly formed regions eliciting zero responses generate fluctuations with magnitudes lower than those observed previously in the $j_1 = 3$ convolution field.

Upon subjecting this field to convolution once more, this time using the $j_2 = 4$ filter, it becomes evident that these two factors collectively contribute to a downward trend in the $S_{21}(j_1 = 3, j_2 = 4)_{Fidu}$ coefficient.



As the ionized regions continue to expand, yet another noteworthy phenomenon unfolds: ionization fronts from distinct ionized regions begin to converge. The sharp ascending trend observed in $S_{21}(j_1 = 3, j_2 = 4)_{Fidu}$ towards the conclusion of the Epoch of Reionization reflects this clustering effect of ionization front fluctuations at the $j_1 = 3$ scale, now impacting the $j_2 = 4$ scale.

4.1.3 Fiducial vs PL 2.0 vs PL 3.0 Models :

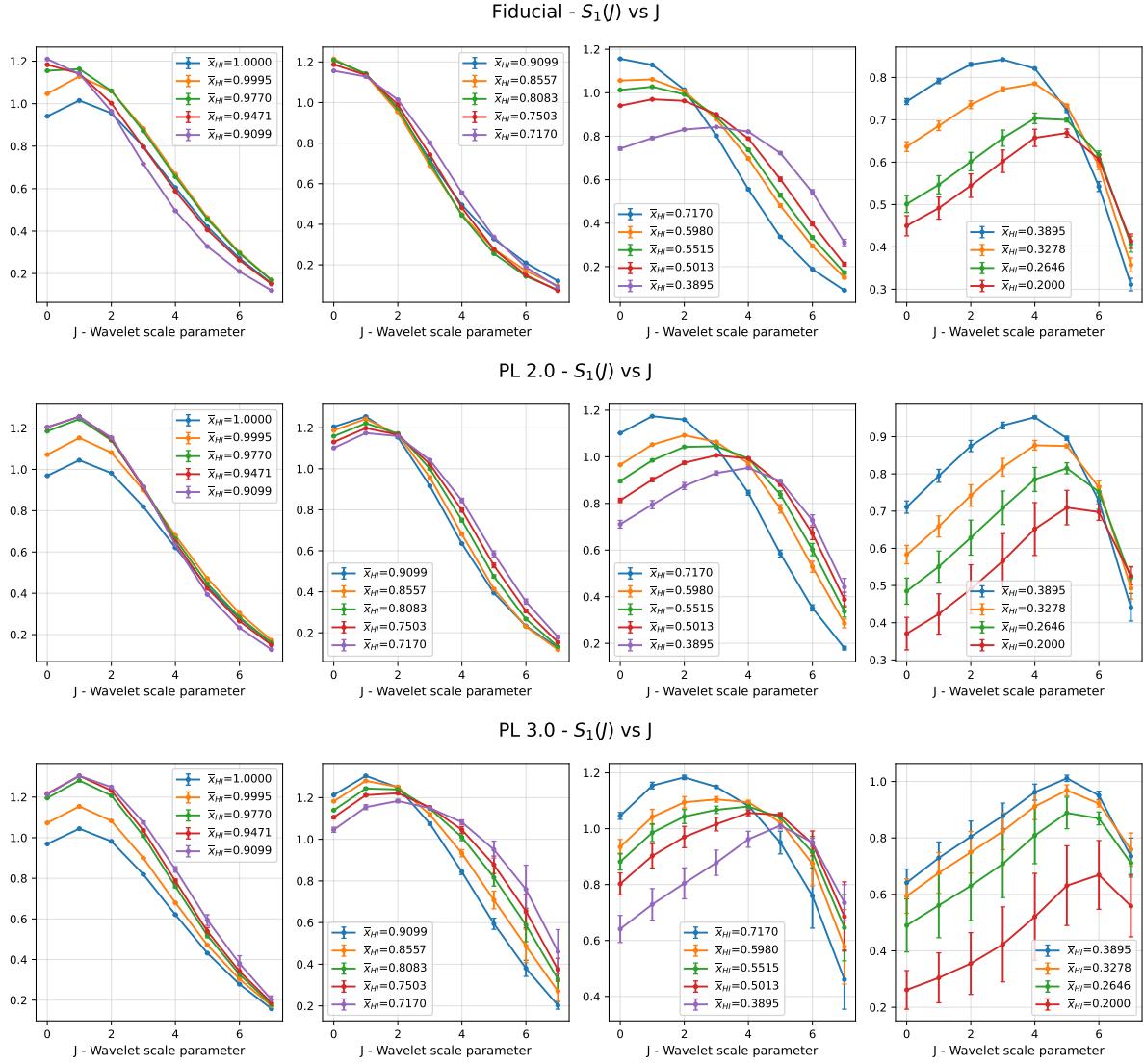


Figure 4.5: Evolution of S_1 coefficients vs. J with \bar{x}_{HI} for Fiducial, PL2.0 & PL 3.0 models

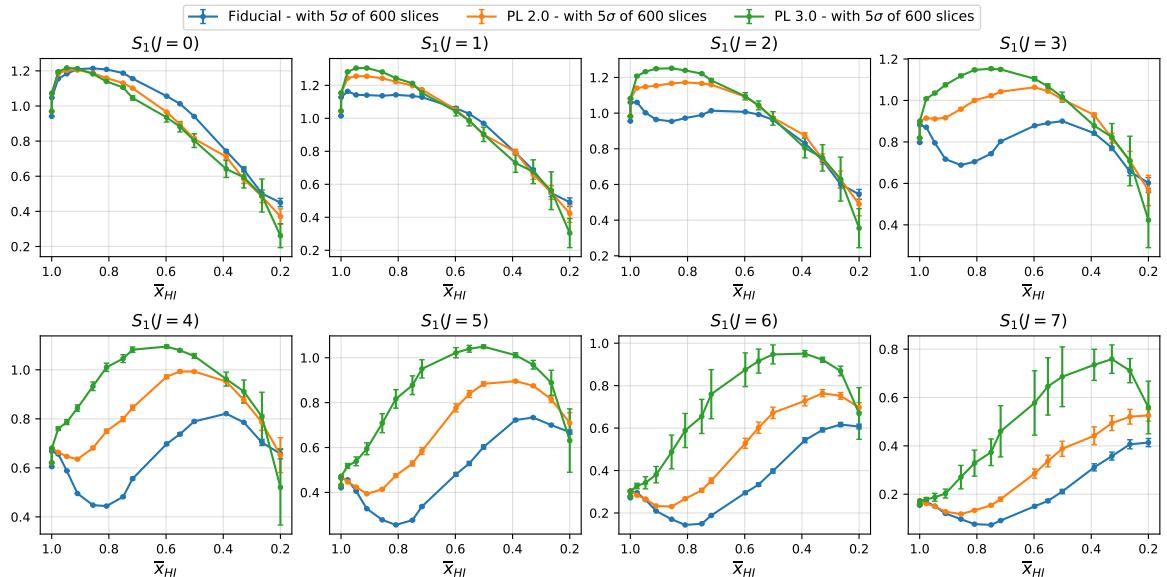


Figure 4.6: Evolution of $S_1(j)$ coefficients with \bar{x}_{HI} for Fiducial, PL 2.0 & PL 3.0 models

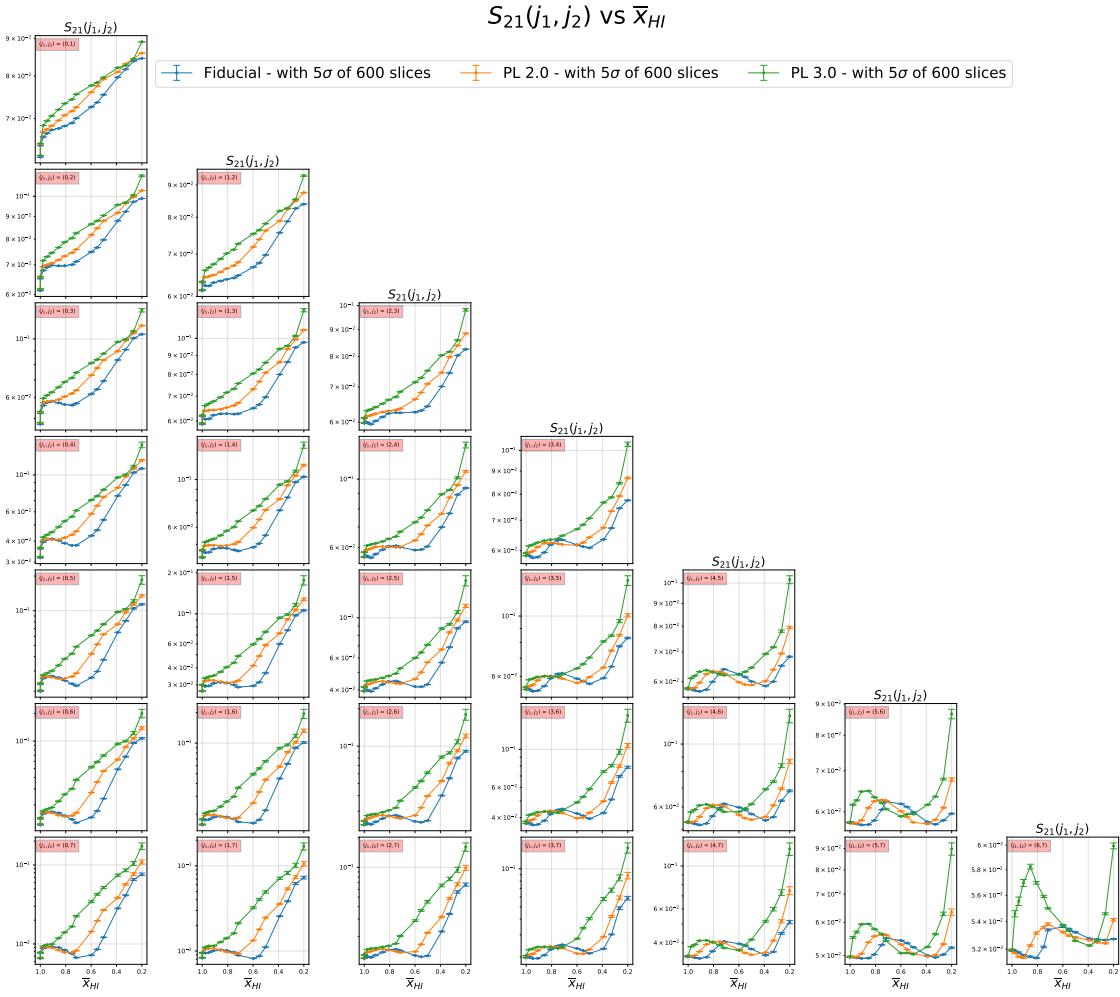
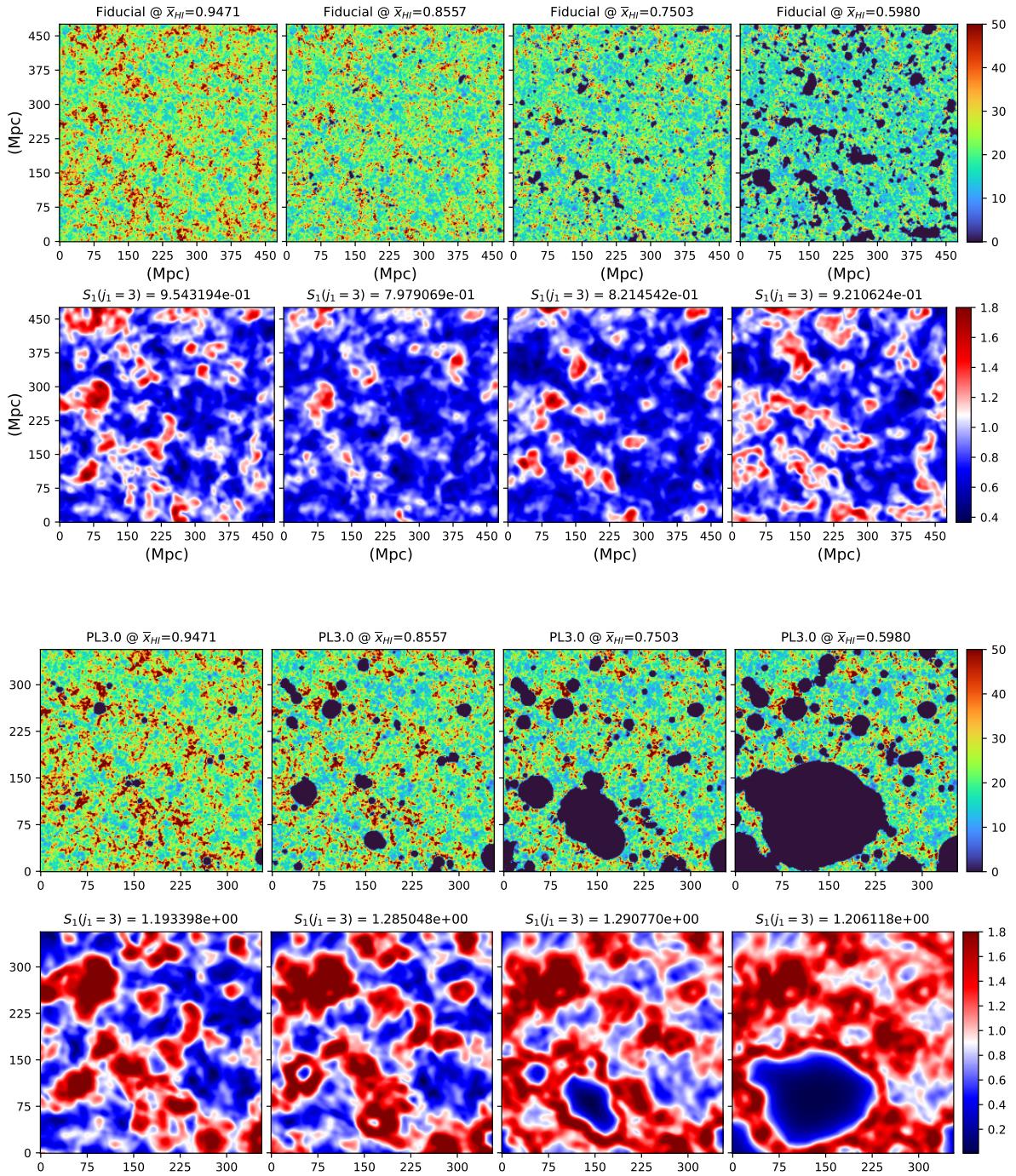


Figure 4.7: Evolution of $S_{21}(j_1, j_2)$ coefficients with \bar{x}_{HI} for Fiducial, PL 2.0 & PL 3.0 models

In the case of the PL 2.0 and PL 3.0 models, where there is a significant increase in the number of ionizing photons, the growth rate of ionized regions is notably faster compared to the Fiducial case. This accelerated growth becomes evident when observing the shift in the onset of the increasing trend in S_1 and S_{21} coefficients, as shown in Fig 4.6 and Fig 4.7 respectively.

In contrast to the Fiducial case, where fluctuations in neutral hydrogen remained dominant for an extended period, in the PL 2.0 and PL 3.0 models, the rapid expansion of ionized regions leads to an earlier transition to their dominance in 21cm signal fluctuations. As a result, there is very little or no observable decreasing trend during the early stages of the Epoch of Reionization (EoR) in these models.

The effect of rapid increase in ionized region sizes can also be readily observed in Fig 4.5. As previously explained, the peak in these curves corresponds to the size of the dominant ionized regions that have the greatest impact on the fluctuations in the 21cm signal. In the PL 2.0 and PL 3.0 models, the larger ionized regions manifest much earlier in comparison to the Fiducial model.



4.1.4 Fiducial vs Inhomogenous Clumping Models :

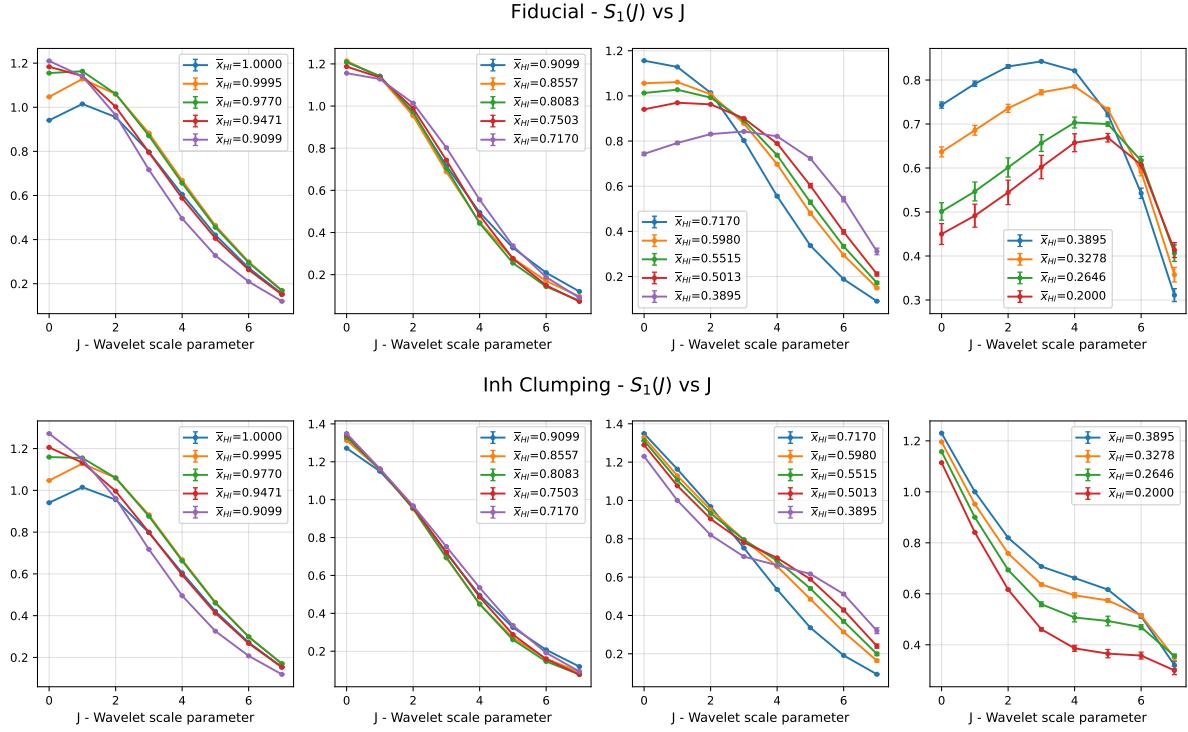
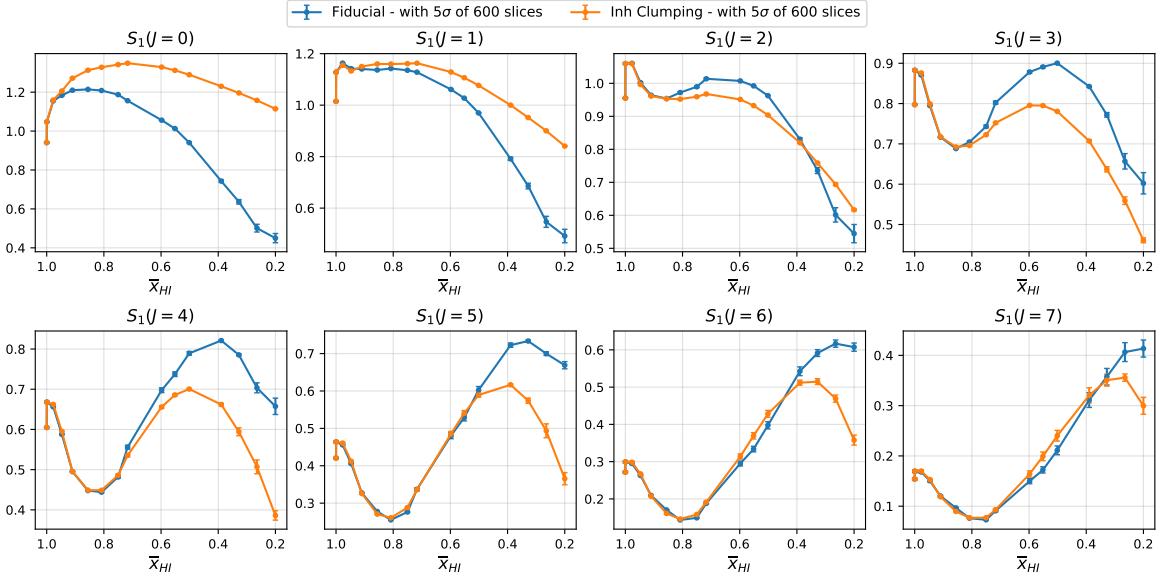


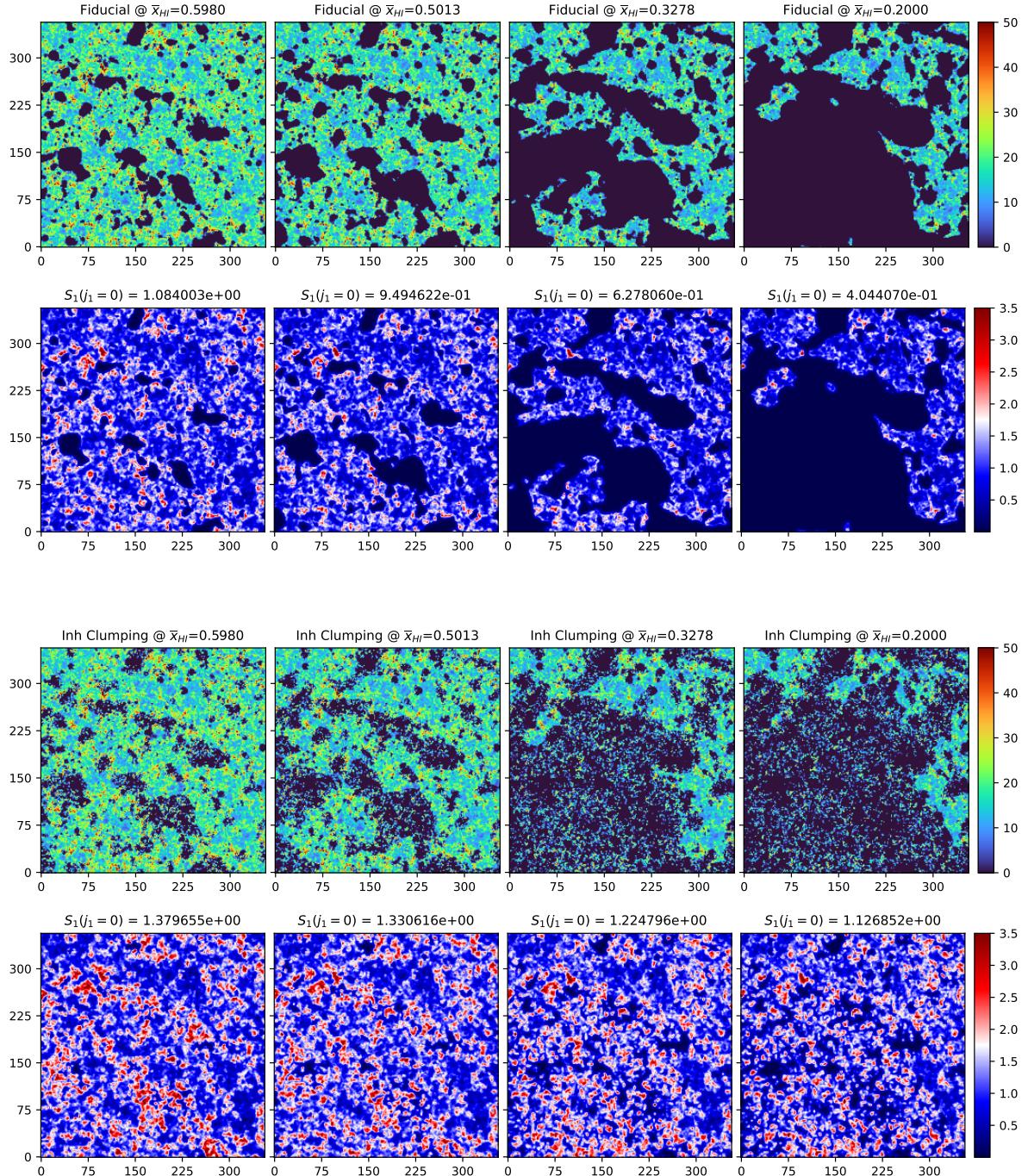
Figure 4.8: Evolution of S_1 coefficients vs. J with \bar{x}_{HI} for Fiducial & Clumping models



As one might anticipate, owing to the underlying physics encapsulated in simulating these models, lower neutral fractions exhibit a notable divergence in the distribution of fluctuations between the Inhomogeneous Clumping model and the Fiducial Model. This distinction is also evident in the evolution of small-scale S_1 coefficients for both of these models. In Fig 4.8, initially, both models commence with identical values, but the Clumping model diverges by not displaying the pronounced decrease in the magnitude of small-scale S_1 coefficients observed in the Fiducial case.

Toward the culmination of the Epoch of Reionization (EoR), when examining $S_1(j)$ vs. j plots, it becomes apparent that, in the Clumping model, there is a considerably greater contribution from small-scale fluctuations compared to large-scale fluctuations. This represents a

marked departure from the Fiducial case, where fluctuations of larger scales dominate during these later stages.



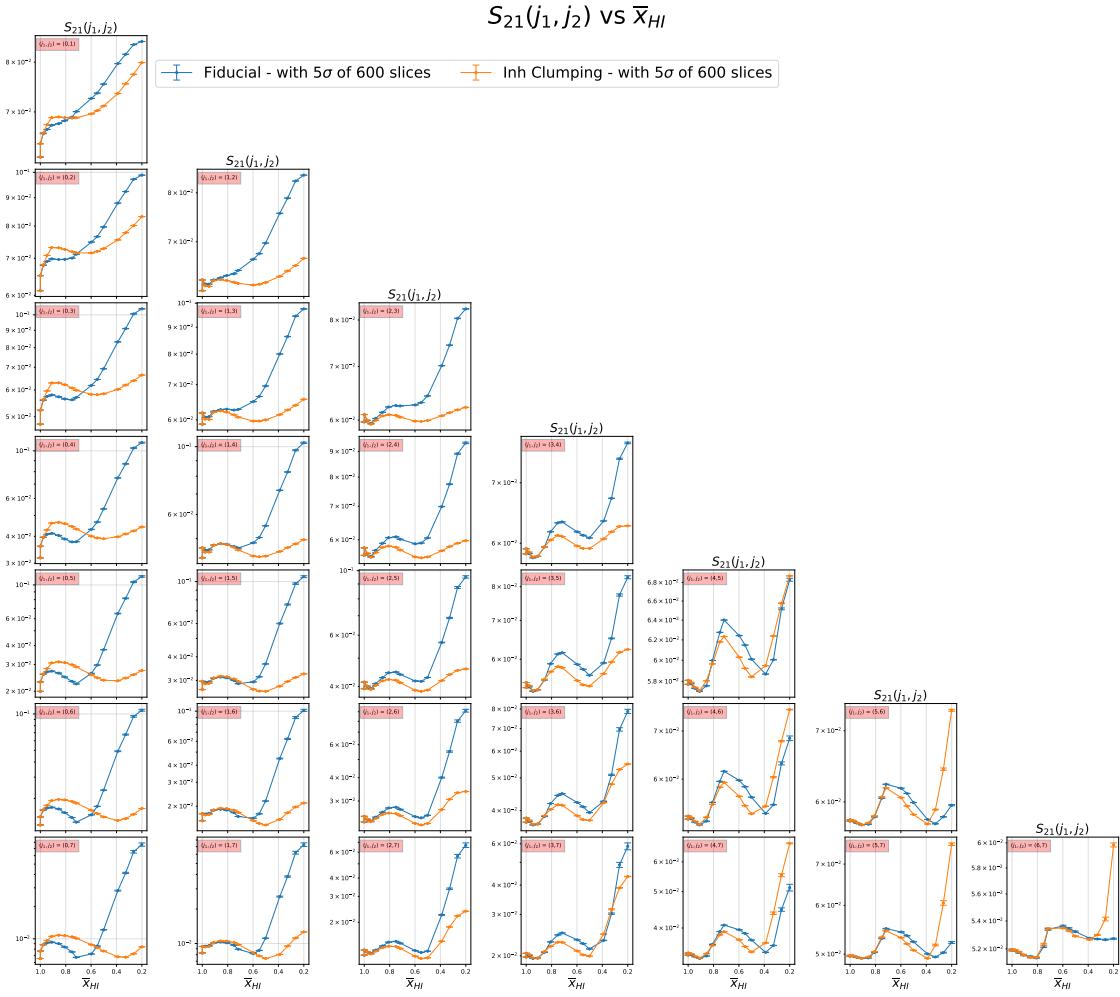
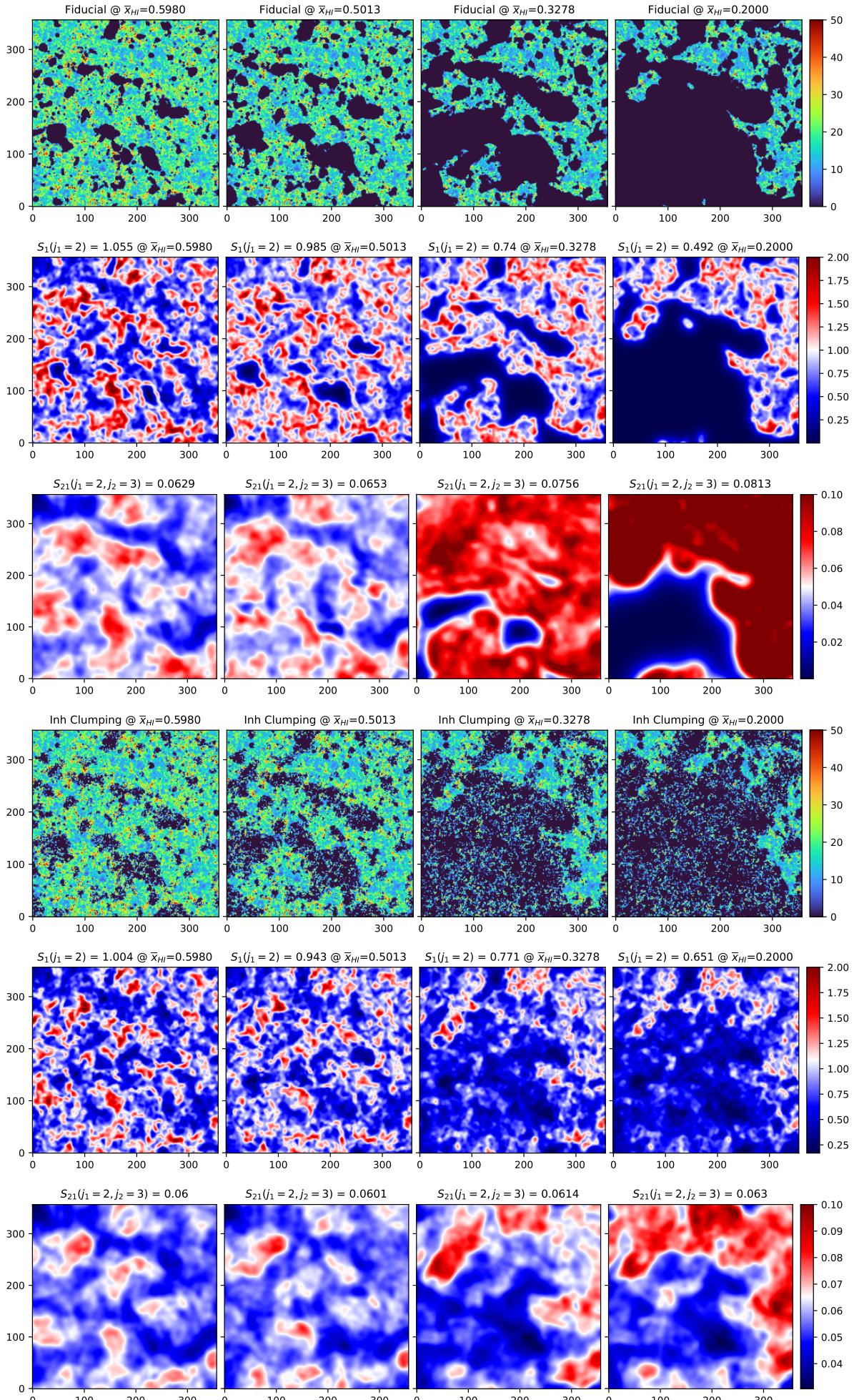


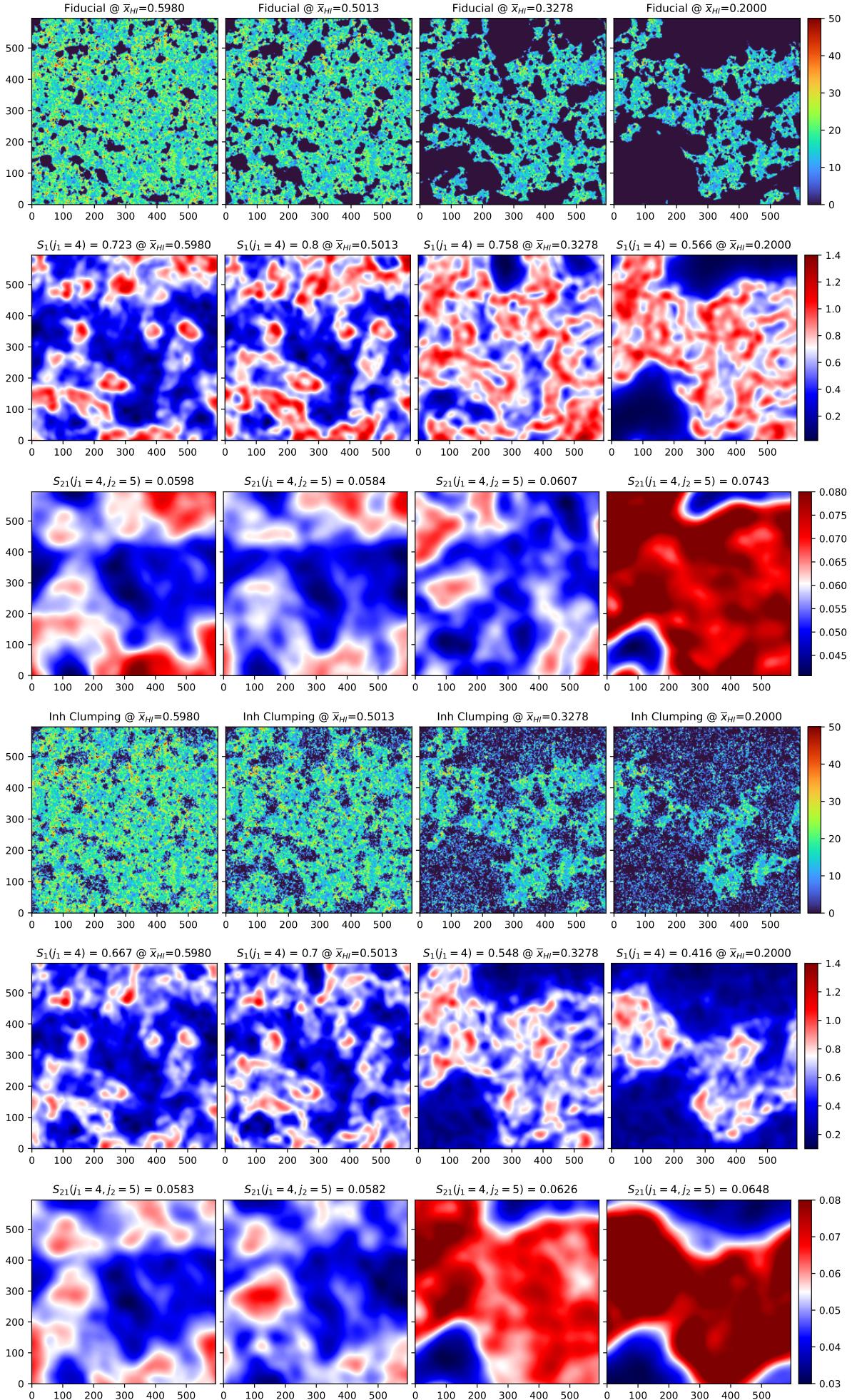
Figure 4.9: Evolution of S_{21} coefficients with \bar{x}_{HI} for Fiducial & Clumping models

The S_{21} plots from Fig 4.9 reveal that, in the early stages of reionization, the trends in the Clumping model and the Fiducial model align. However, as we delve into lower scales, a clear divergence emerges during the later phases of reionization. In the Fiducial case, a sharp increasing trend is observed, while the Clumping case exhibits a much slower increase. The tail end of the Clumping model begins to resemble the trend observed in the Fiducial model at higher j_1 scales.

At smaller scales, in the Fiducial model, when the j_1 filter detects larger-than-scale ionized regions and registers them as zero filter responses, these regions do not appear in the j_1 convolved output of the Clumping model due to the presence of recombinations within these ionized regions. In the Fiducial case, where ionization fronts dominate the fluctuations at the j_1 scale, their growth in size inevitably leads to their convergence. In contrast, in the Clumping model, both ionized regions and recombined regions contribute to fluctuations at the j_1 scale. While clustering of ionized front fluctuations still occurs in the Clumping model, the overall dominance of such clustering is noticeably diminished. This plays a pivotal role in the divergence observed in the S_{21} coefficients for small scales during the later stages of reionization.

Nevertheless, as we progress to higher j_1 scales, the contribution of recombination regions to the overall fluctuations also diminishes, causing the trend to align more closely with the dominant clustering effect of ionized regions, akin to what is observed in the Fiducial model.





4.1.5 Fiducial vs UIB Models :

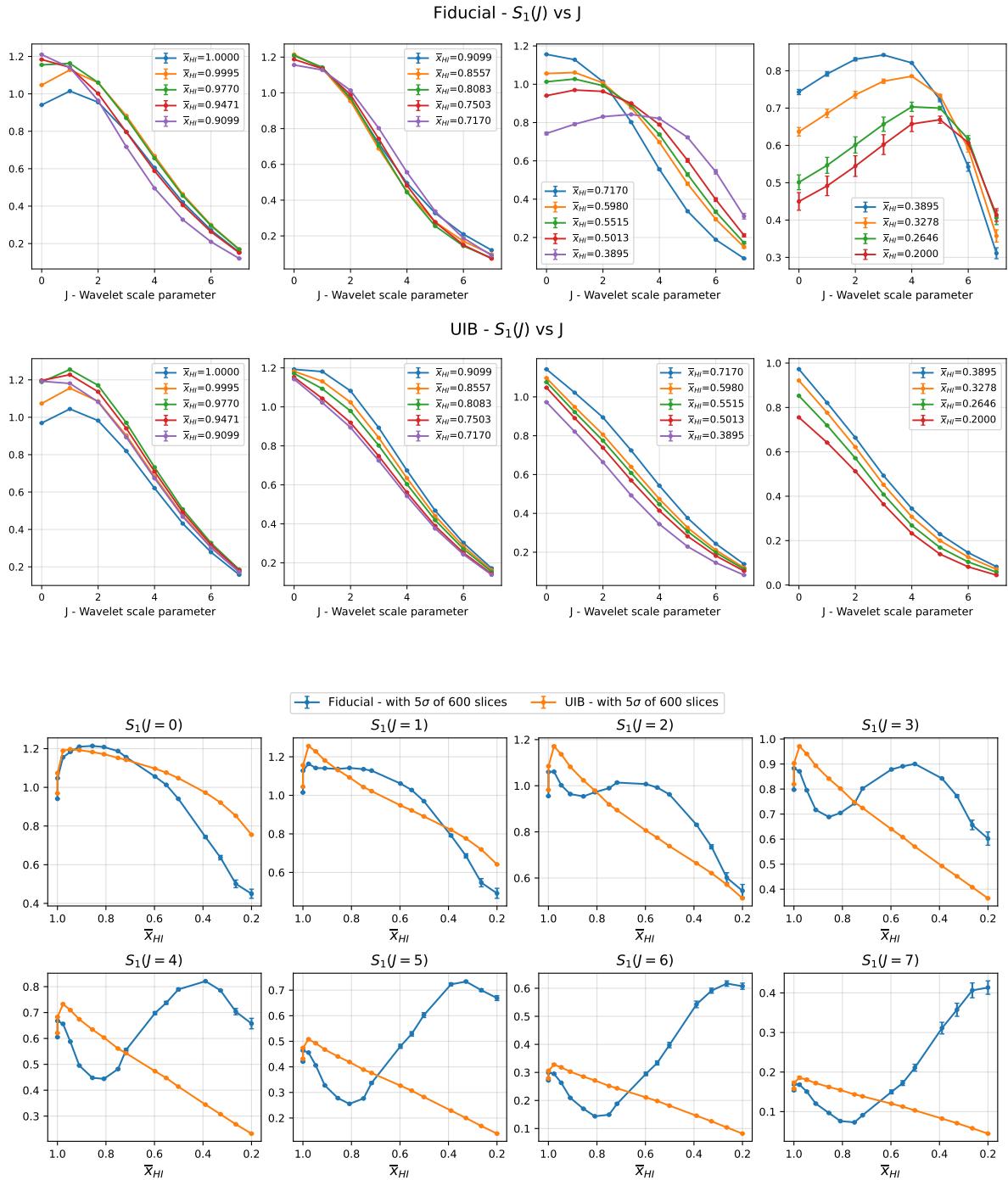


Figure 4.10: Evolution of $S_1(j)$ coefficients with varying \bar{x}_{HI} for Fiducial & UIB models

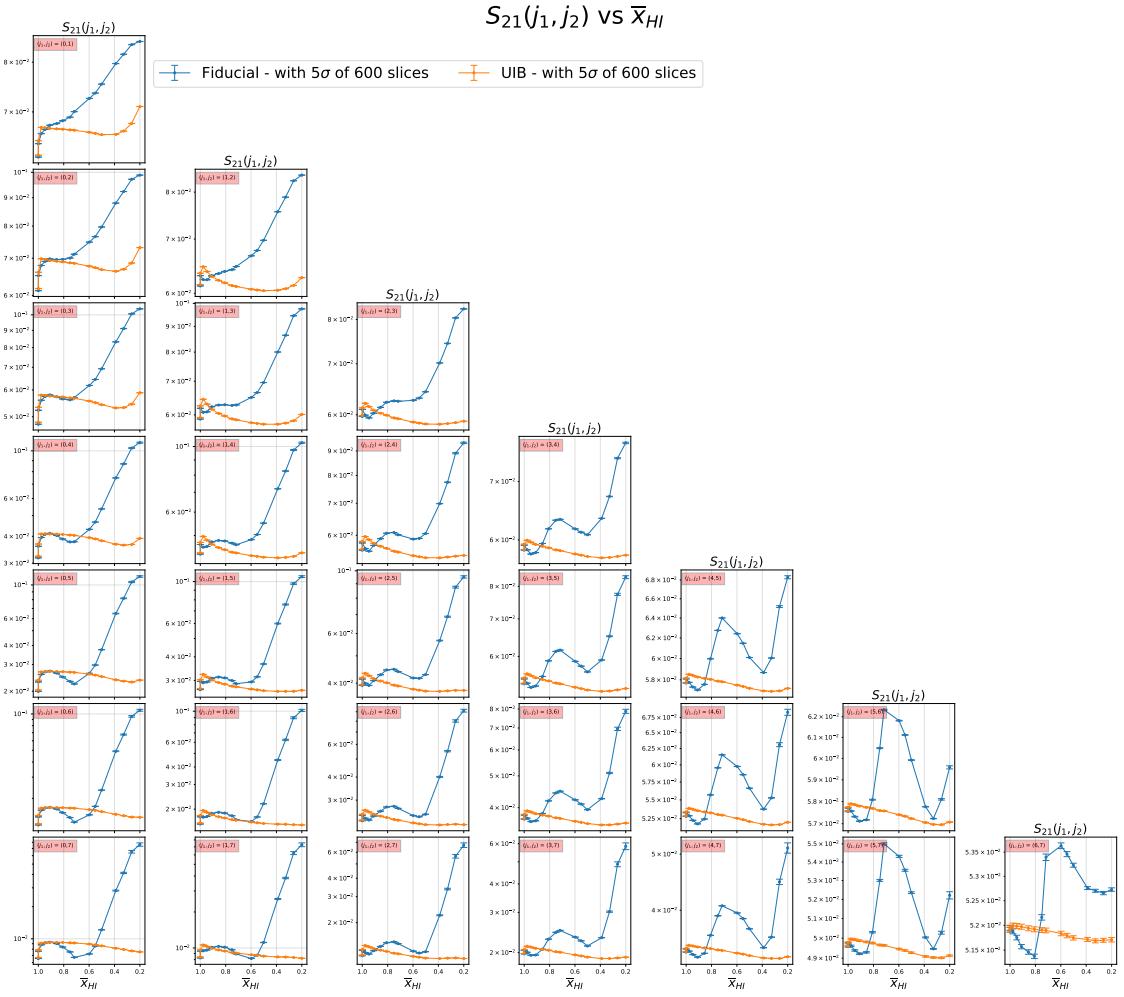
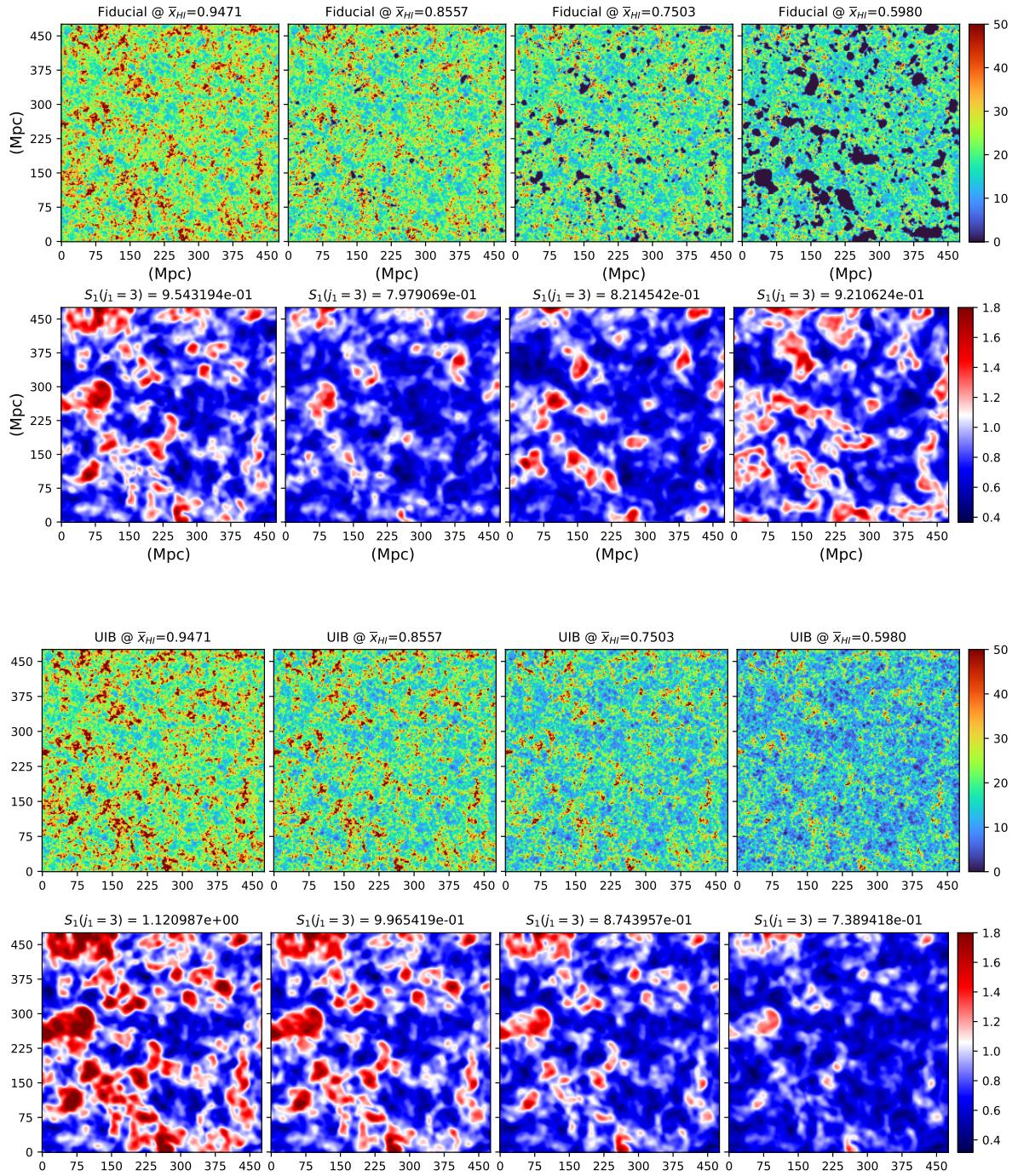
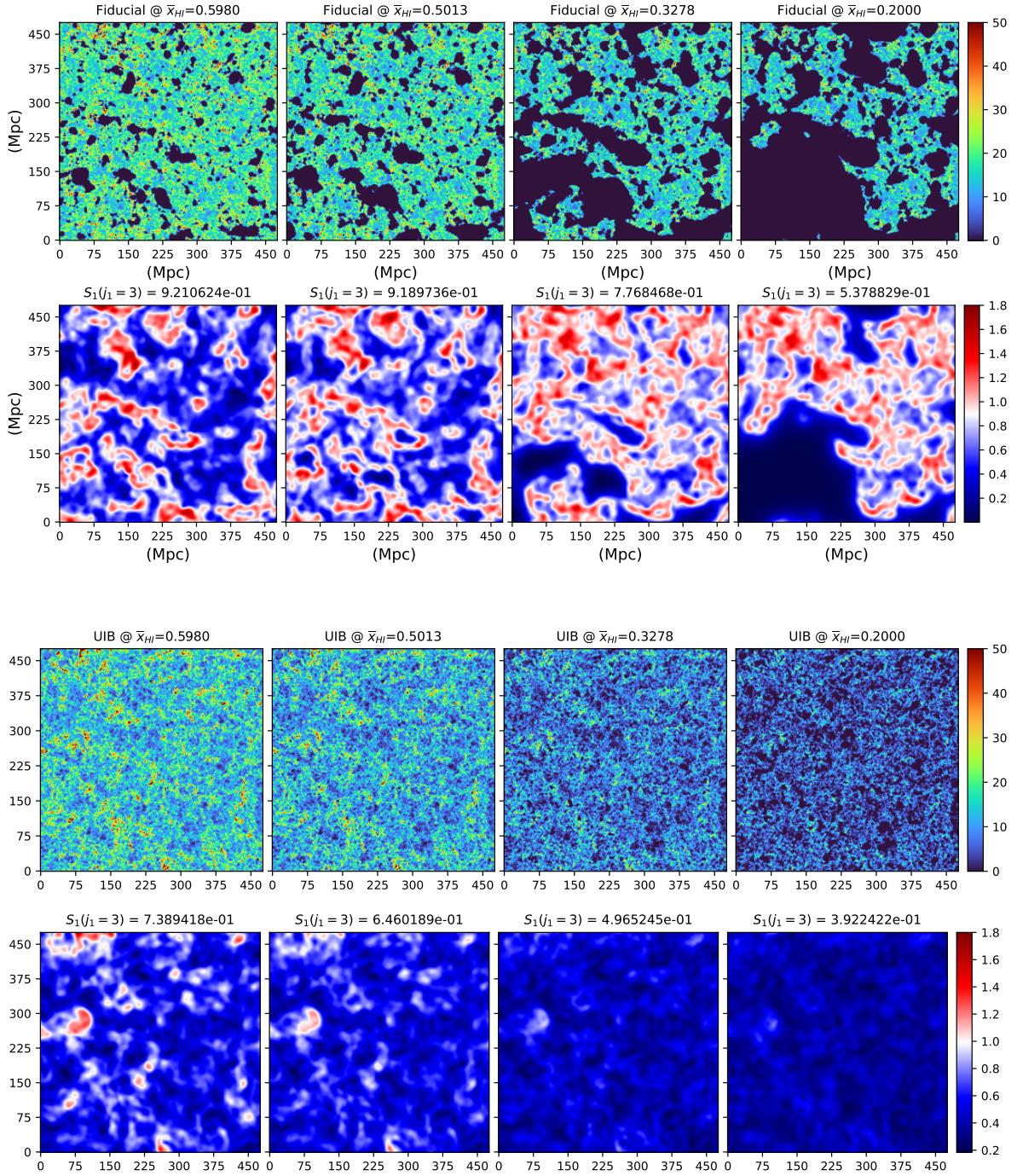


Figure 4.11: Evolution of $S_{21}(j_1, j_2)$ coefficients with varying \bar{x}_{HI} for Fiducial & UIB models

In the case of the UIB Model, where the UV contribution is 20% and the hard X-ray uniform ionization contribution is 80%, a notable difference in S_1 and S_{21} coefficients is observed when compared to the Fiducial case, as depicted in Fig 4.10 and Fig 4.11. For $S_1(j \geq 2)$ coefficients in Fiducial case, there is an initial decreasing trend followed by a subsequent increasing trend during the Epoch of Reionization (EoR). However, in the UIB model, the decreasing trend dominates throughout the entire EoR stage.





The absence of the increasing trend in the UIB model can be understood by a closer examination of 21cm maps. Unlike the Fiducial model where ionized regions grow, prompting the transition to an increasing trend, the UIB model lacks this growth of ionized regions. Consequently, the 21cm signal fluctuations persistently respond to the fluctuations in neutral hydrogen density. Moreover, these density fluctuations gradually decrease due to the influence of hard X-ray uniform ionization, which is the primary factor contributing to the observed persistent decreasing trend.

4.1.6 Fiducial vs SXR Dom vs UV+SXR+UIB Models :

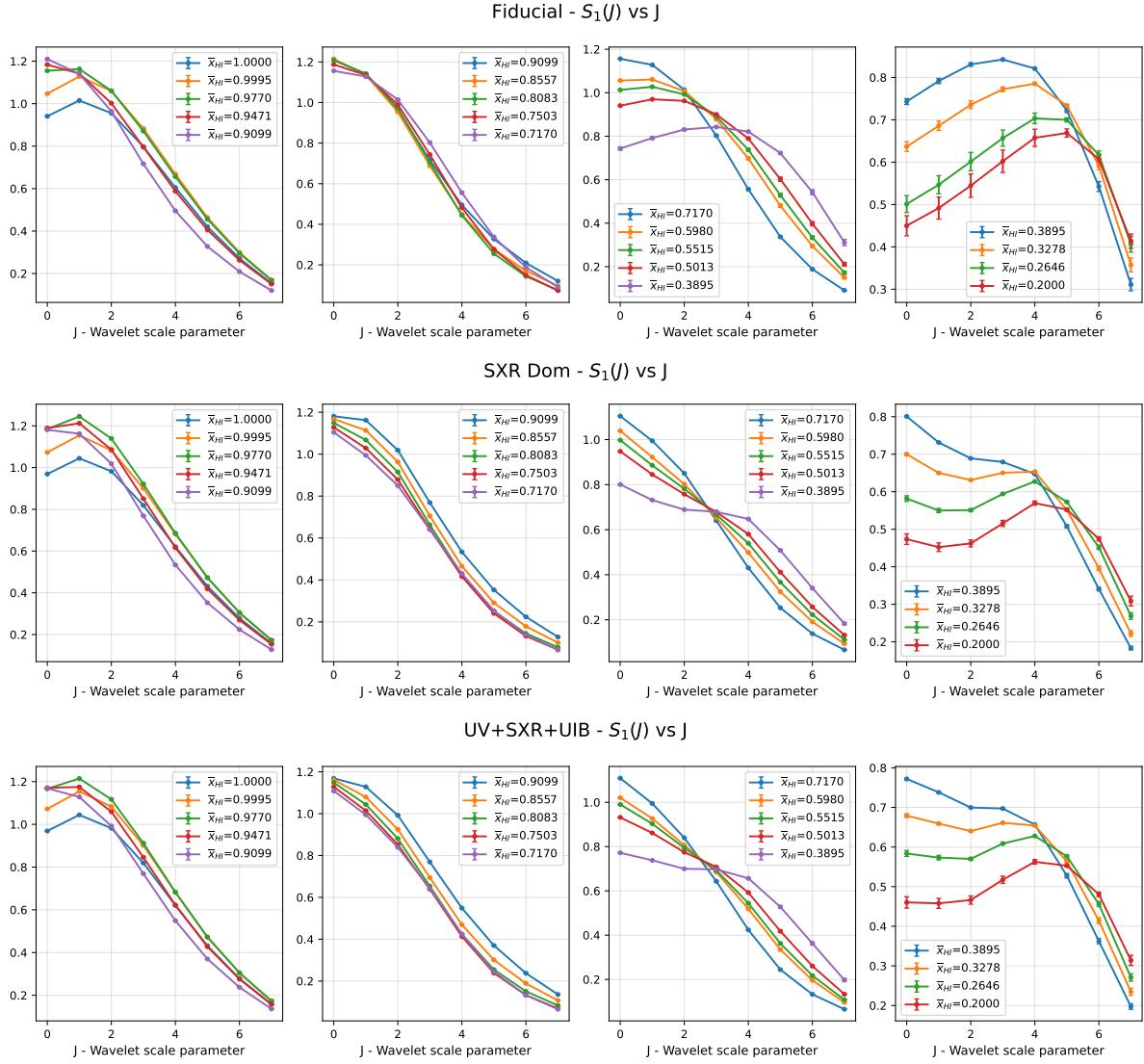


Figure 4.12: Evolution of S_1 coefficients vs. J with \bar{x}_{HI} for Fiducial, SXR Dom & UV+SXR+UIB Dom models

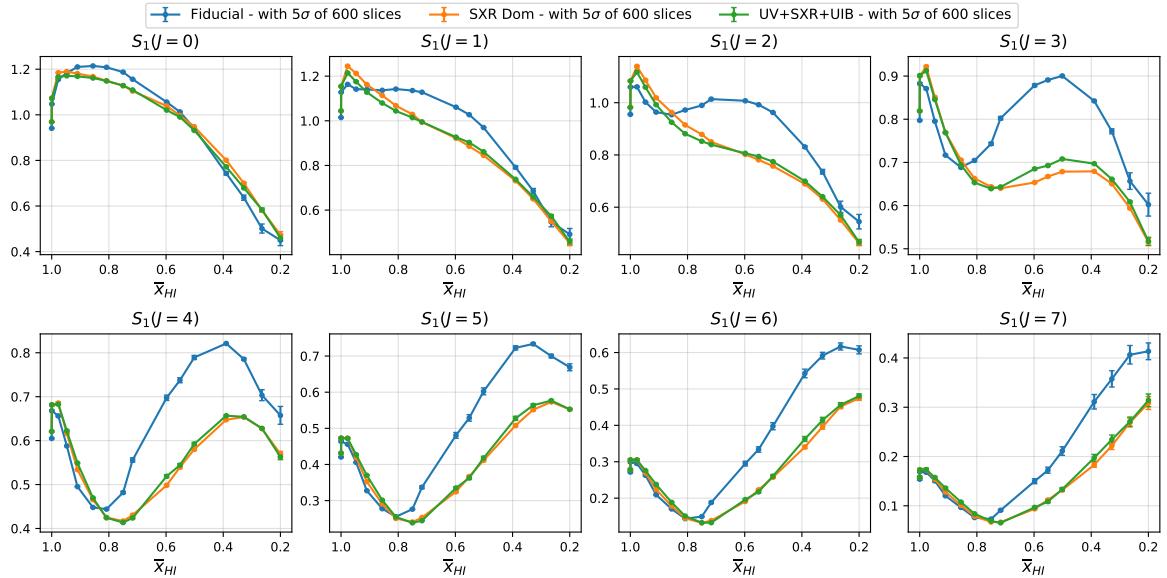
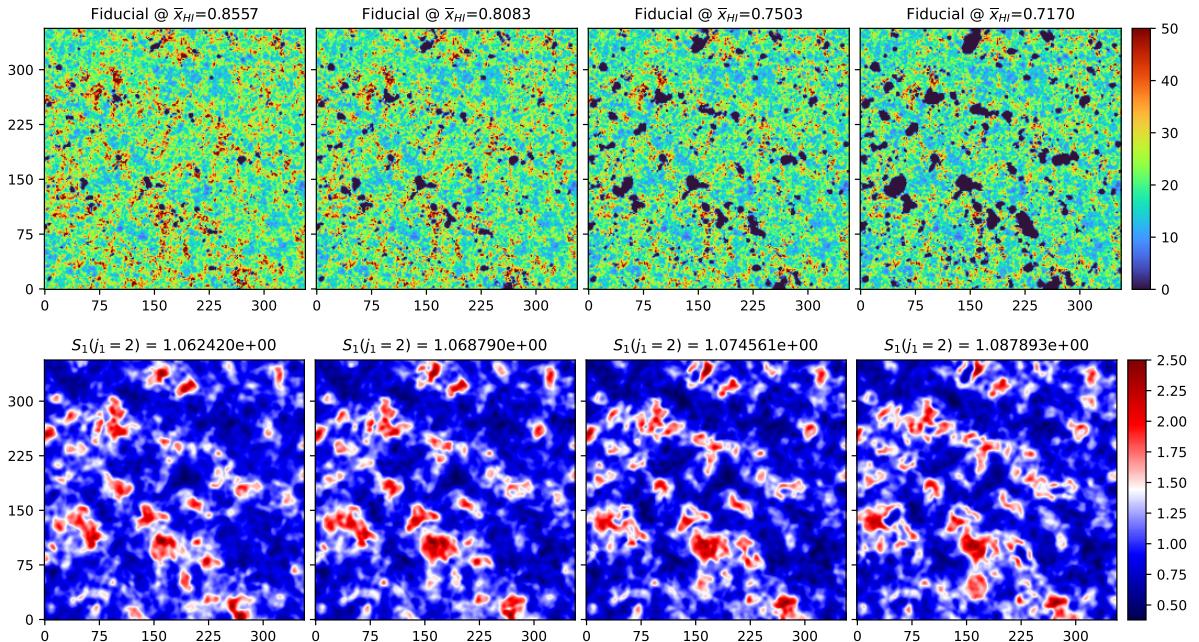


Figure 4.13: Evolution of $S_1(j)$ coefficients with varying \bar{x}_{HI} for Fiducial, SXR Dom & UV+SXR+UIB Dom models

From the evolution of S_1 coefficient in Fig 4.12 & Fig 4.13, these are the main differences that can be observed.

- The distinct timing of the onset of the increasing trend in the Fiducial case, occurring earlier in reionization, in contrast to the SXR and UV+SXR+UIB models, where this manifests a bit later in Fig 4.13
- The distinct timing and the amplitude of the onset of the decreasing trend in the late stages of the EoR in the Fiducial case, occurring earlier in contrast to the SXR and UV+SXR+UIB models, where this peak manifests a little later in Fig 4.13
- Higher contributions of small scale fluctuations in SXR Dom and UV+SXR+UIB models unlike the Fiducial model in Fig 4.12

An examination of Fig 4.14 & Fig 4.15, to understand how the source models considered influence the primary contributors to 21cm signal fluctuations at different scales, is necessary to explain these differences in WST coefficient trends. In the Fiducial model, the predominant ionization sources are UV photons, constituting 100% of the photon budget. As a result, regions with high-density neutral hydrogen experience rapid ionization, leading to the emergence of ionization fronts as the primary contributors to fluctuations in the 21cm signal. In the SXR Dominated model, the contribution of UV photons is significantly diminished, accounting for only 20% of the photon budget. Instead, this model heavily relies on Soft X-ray (SXR) photons, constituting the other 80% of the photon budget. In the UV+SXR+UIB model, multiple ionization sources are integrated, including UV photons, Soft X-ray (SXR) photons, and hard X-ray photons. The combined effects of these diverse sources introduce complex dynamics into the ionization process.



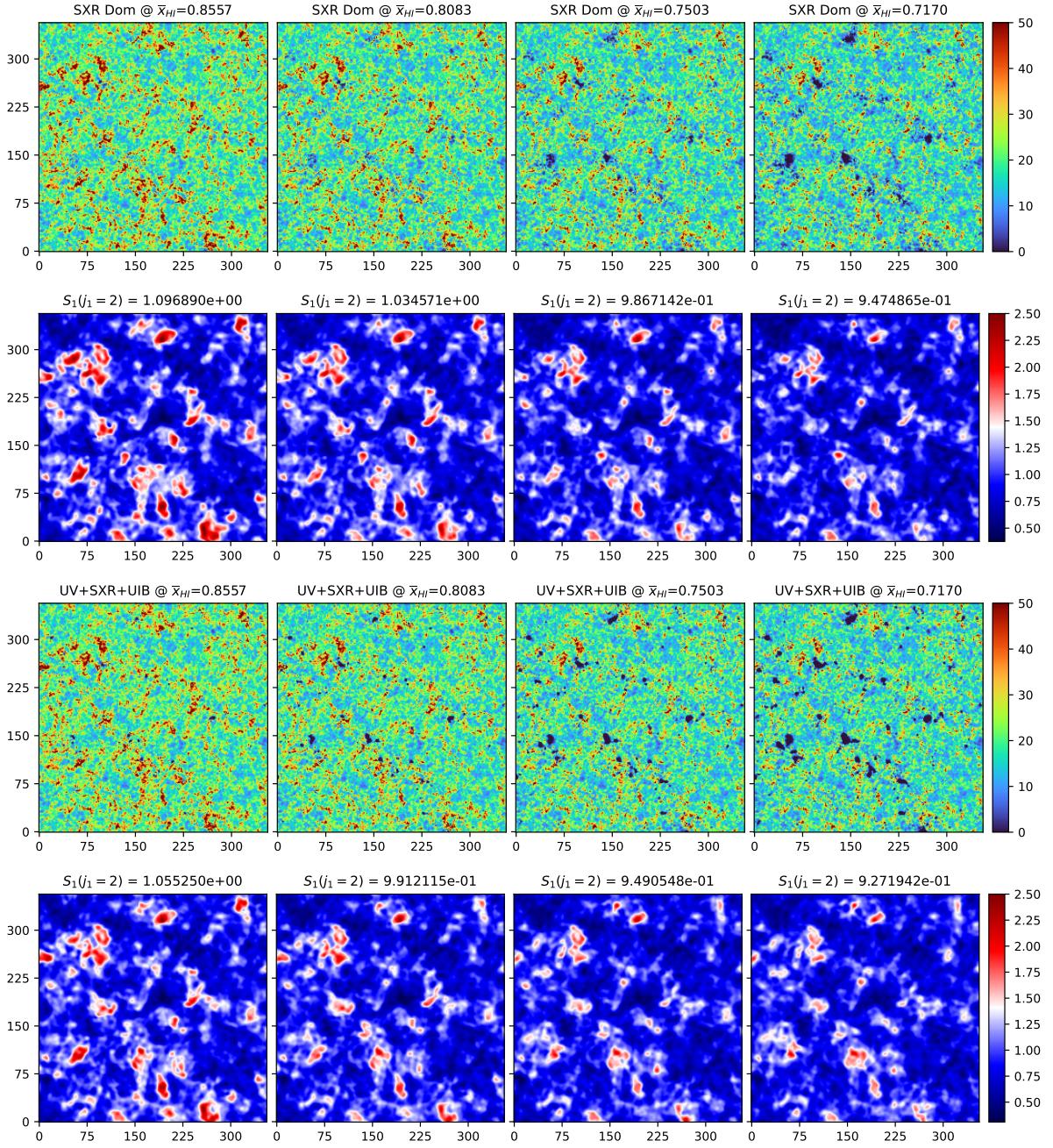


Figure 4.14: Evolution of Fiducial, SXR Dom & UV+SXR+UIB Dom models

SXR and hard X-ray photons inherently possess lower ionization power compared to UV photons due to their photoionization cross-sections. X-ray photons need to travel a certain distance from their source to get redshifted into the UV regime, where their ionizing power becomes equivalent to that of UV photons. This means that these SXR photons do not ionize the immediate surroundings of their emission regions effectively. Consequently, massive halos with high-density regions of neutral hydrogen retain their neutral state for an extended duration, delaying the emergence of ionization fronts as the primary contributors to fluctuations in the 21cm signal. This delay accounts for the late appearance of the dip in these models. The swift ionization of high-density regions because of the higher UV contribution in the photon budget accounts for the early appearance of the dip in the Fiducial model.

Additionally, due to the significant contribution of UV photons, ionized regions also expand in size at a faster rate in the Fiducial case compared to the other two models. The timing of the peak in the 21cm signal corresponds to the moment when the dominant ionized

regions attain a size that matches the wavelet filter under consideration. Since the Fiducial model benefits from a high contribution of UV photons, the ionized regions reach this critical size earlier in the reionization process. Therefore, the peak in the WST coefficients of 21cm signal appears earlier in the Fiducial model compared to the other models. It is also interesting to note that the dip effect is seen at $j = 2$ and higher scales in the Fiducial model, but is only seen at $j = 3$ and higher scales in SXR Dom and UV+SXR+UIB models. The absence of the dip effect at $j = 2$ scale in the latter models suggests that, there is no phase in EoR at this scale in these models where the evolution of fluctuations is dominated by the growth of ionized regions. The depletion effect of neutral hydrogen fluctuations is strong enough to dictate the evolution trends in SXR Dom and UV+SXR+UIB models. This causes the difference in the trends in $S_1(j = 2)$ coefficient in Fig 4.13

In the Fiducial case, ionization proceeds primarily through inside-out ionization, allowing the inside-out ionized regions to grow faster than in the other models. The transition from smaller to larger sizes occurs relatively smoothly in the Fiducial case, except when two ionized bubbles merge together. Conversely, in the SXR Dom and UV+SXR+UIB models, due to the combined influence of partial UV and X-ray ionization, both inside-out and outside-in ionization processes occur concurrently around the sources. Sizes of ionized regions change abruptly when UV-ionized regions and X-ray-ionized regions around the same source connect with each other.

There are instances in the Fiducial case where certain ionized regions start out as separate but closely situated bubbles, and later they merge with each other after growing to larger sizes. However, in the SXR Dom and UV+SXR+UIB models, due to the previously mentioned combined effect of inside-out and outside-in ionization, such closely located ionized regions merge when ionized regions are smaller than fiducial case, reducing the overall number of ionized regions and introducing an more abrupt transition in sizes of ionized regions. This effect can be seen in trends of Fig 4.12 – the transition of major fluctuations from smaller scales to larger scales is smoother for Fiducial model but relatively quicker for SXR Dom and UV+SXR+UIB models. Also, due to the partial x-ray ionization which mainly effects small scale neutral hydrogen density fluctuations without directly contributing to growth of ionized regions, small scale fluctuations still dominate in later stages for SXR Dom and UV+SXR+UIB models unlike Fiducial model.

At $j = 4$ scale and higher scales, trends in the SXR Dom and UV+SXR+UIB models align with those of the Fiducial case, as the ionized regions become the dominant fluctuations and their growth dictates the evolution trends similar to the Fiducial Model.

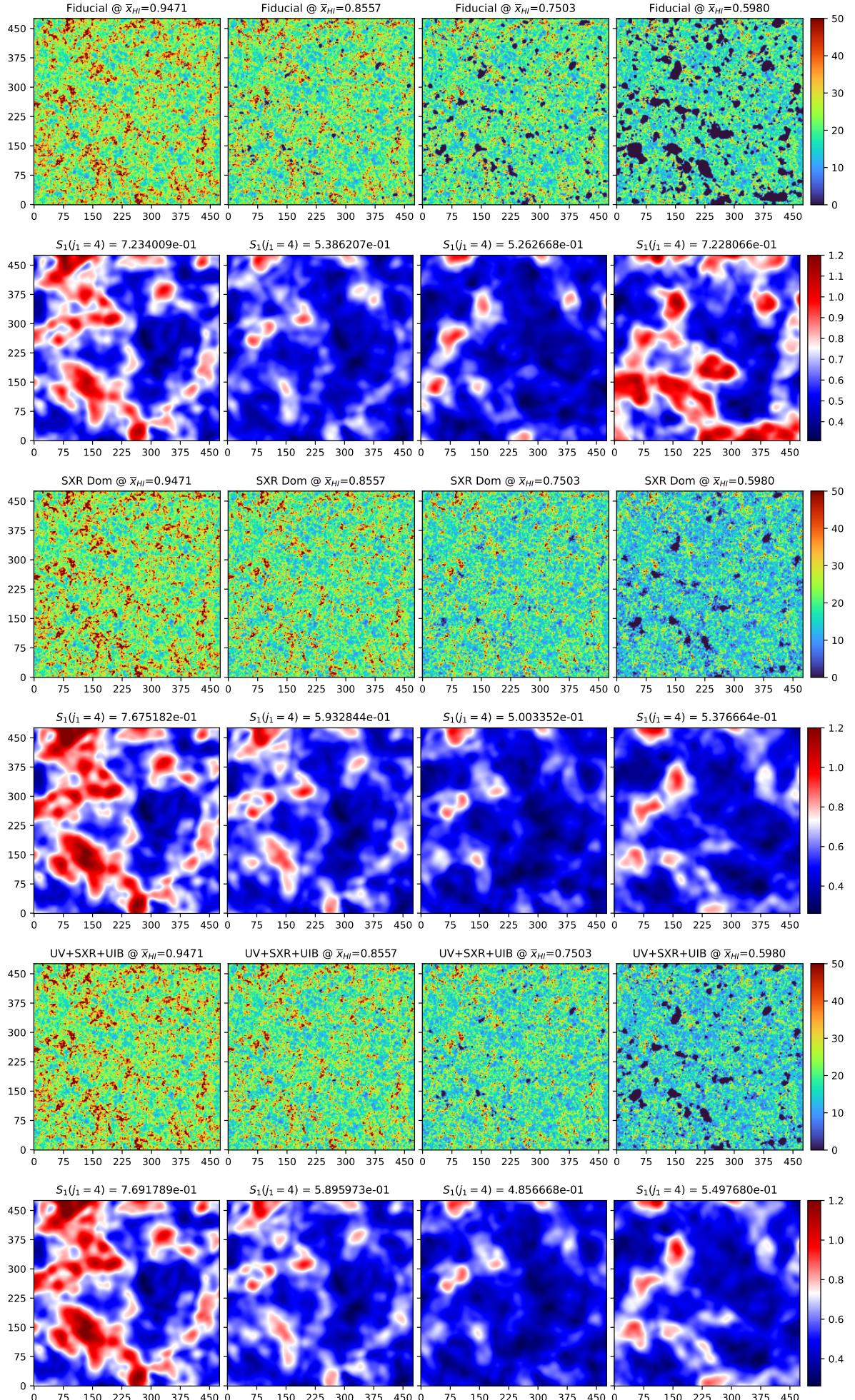


Figure 4.15: Evolution of Fiducial, SXR Dom & UV+SXR+UIB Dom models

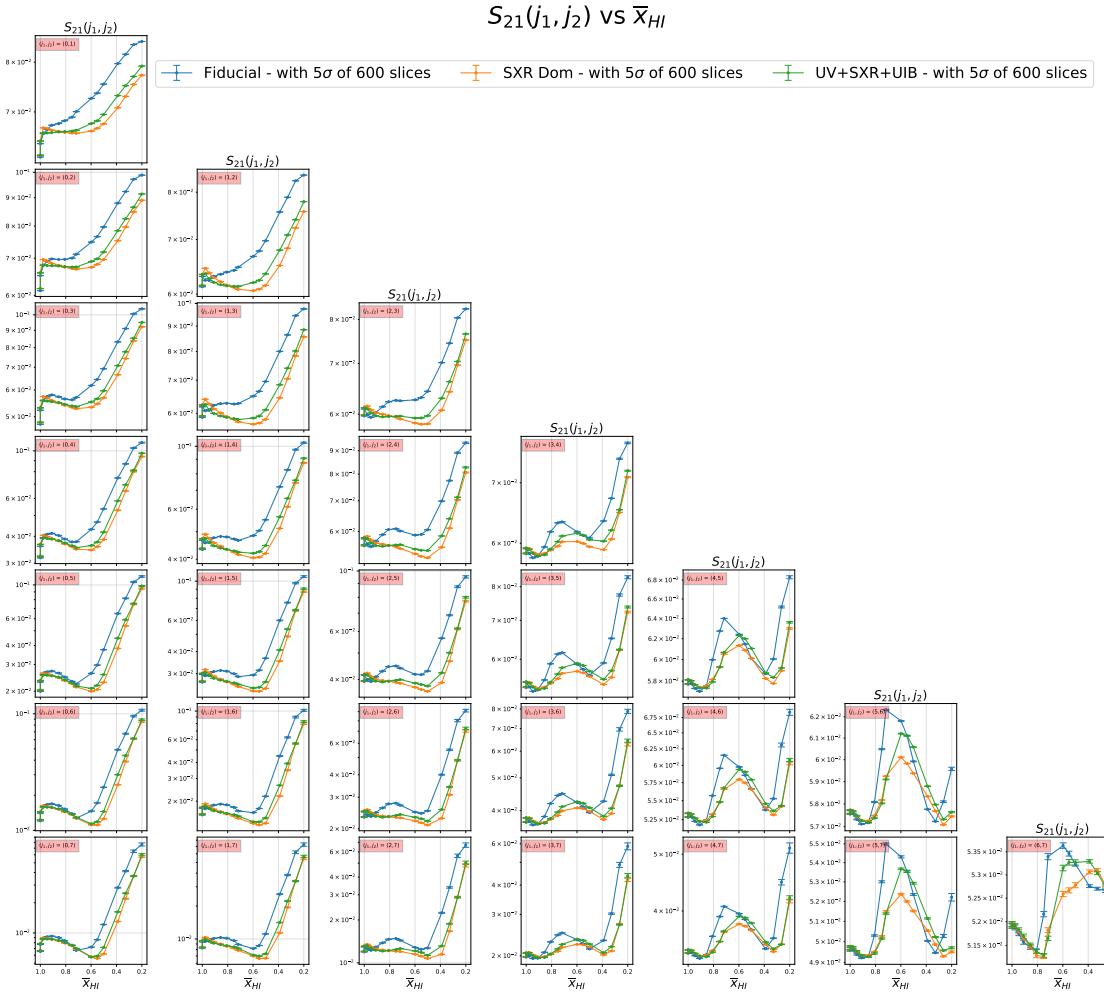
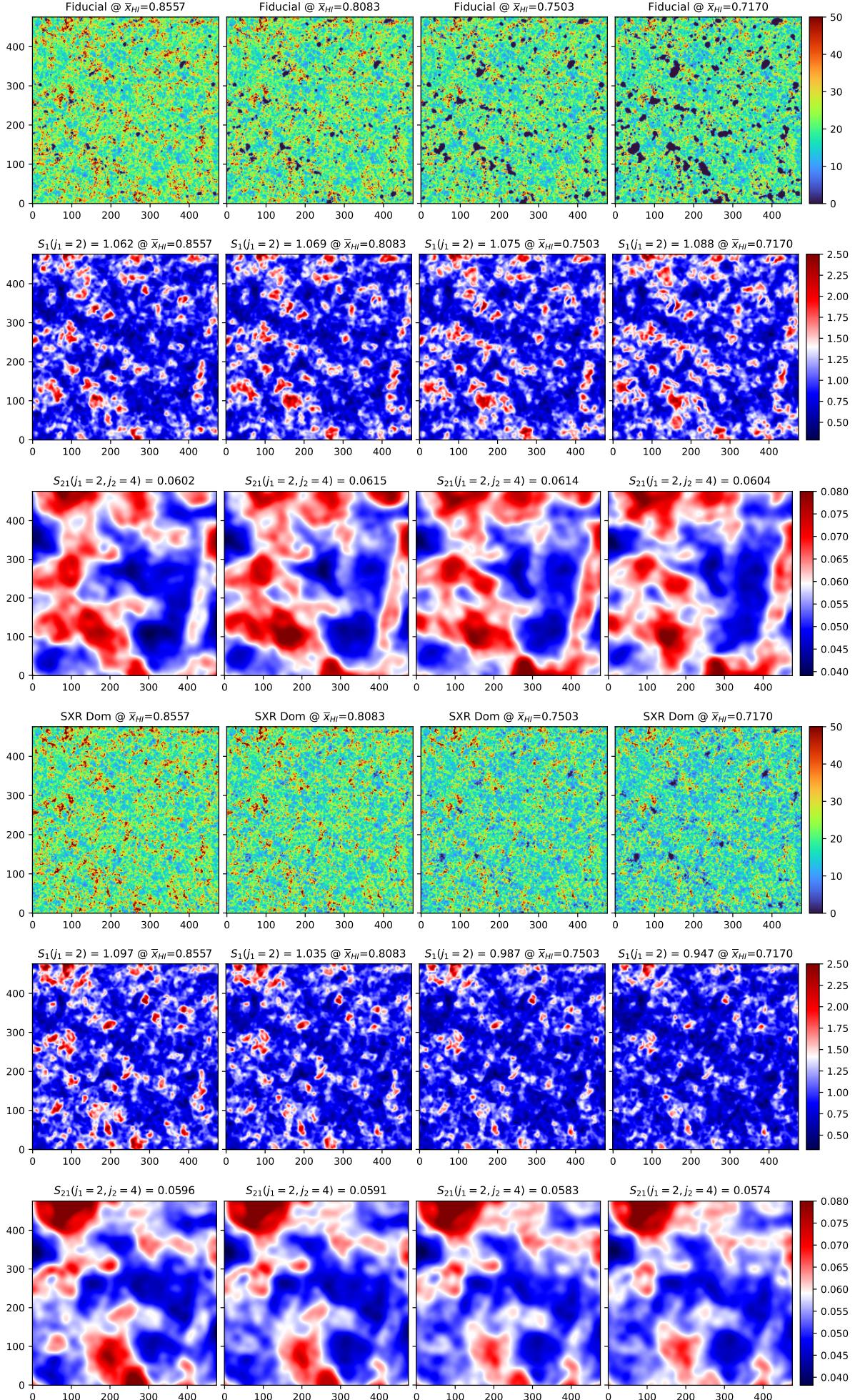
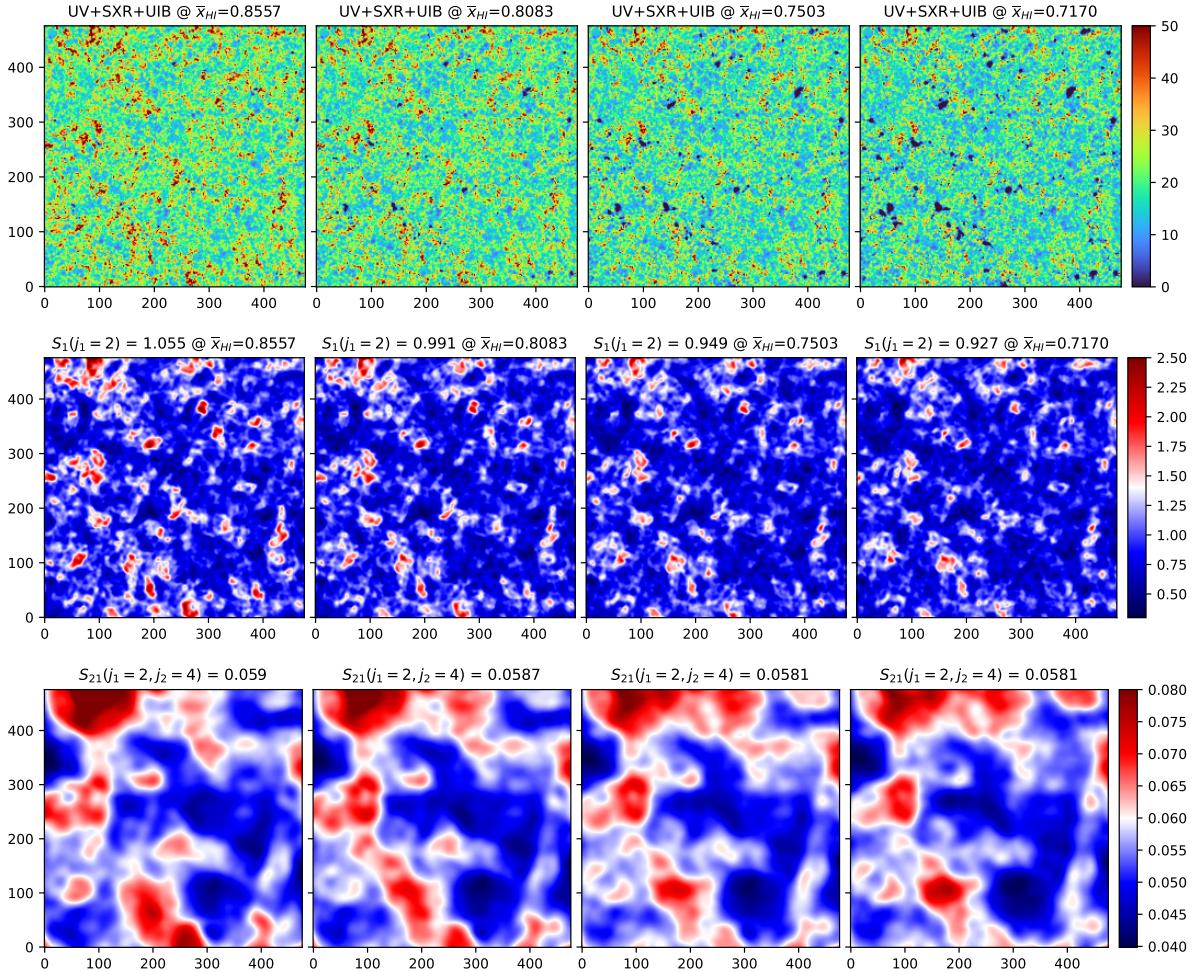


Figure 4.16: Evolution of $S_{21}(j_1, j_2)$ coefficients with varying \bar{x}_{HI} for Fiducial, SXR Dominated and UV+SXR+UIB Models

Fig 4.16 indeed shows that, at most lower and higher scales, the trends in all three models closely align, demonstrating a high degree of similarity in their evolution. However, a notable distinction emerges at a particular scale, $j_1 = 2$, where differences become apparent. This scale corresponds to the same scale where dissimilar trends were observed in the $S_1(j_1 = 2)$ coefficient across these models. The deviations at this scale likely play a crucial role in shaping the distinct behaviors and patterns observed in the 21cm signal fluctuations within each model.

In Fig 4.16, let us focus on $S_{21}(j_1 = 2, j_2 = 4)$ in particular and analyse the differences.





The reduced contribution of UV photons in the source models has significant consequences for the growth rate and abundance of ionized regions. Furthermore, the non-local partial X-ray ionization further diminishes the fluctuations in neutral hydrogen density. Collectively, these factors combined can explain the divergence observed in the $S_{21}(j_1 = 2, j_2 = 4)$ coefficient from the Fiducial model. When compared to the increasing trend observed in the Fiducial model during the phase when ionization regions become the dominant contributors at scale $j_1 = 2$, there could be a diminished clustering effect of the ionized regions in SXR Dom and UV+SXR+UIB models owing to the combined effect of fewer ionized regions, slower growth of these regions and the x-ray ionization of neutral hydrogen regions.

In summary, the diminished presence of UV photons, slower growth of ionized regions, and the influence of X-ray radiation collectively could explain the distinct behavior of the S_{21} coefficients in comparison to the Fiducial model. This divergence is particularly evident during the phase when ionization regions assume dominance in the Fiducial model case at the $j_1 = 2$ scale.

4.2 WST Analysis of Lightcones:

4.2.1 Creating Lightcones from Simulated Volumes:

I have taken inspiration from Tools21cm package and used their ideas for generating lightcones from simulated volumes. I consider two methods of interpolation in lightcone generation – Linear & Sigmoid. To see if the choice of interpolation affects the WST trends, I have

generated lightcones using both and plotted the WST lightcone results below alongside the WST coeval cube results.

4.2.2 Calculating neutral fractions for Lightcone slices:

Lightcone output directly only returns redshift array of all its slices. But to get the neutral fractions at these redshifts, I generated lightcones for hydrogen gas density fields and ionization fields at the same redshifts as the brightness temperature lightcone. Then with straightforward calculation for each slice, neutral fraction values are computed.

4.2.3 Binning Lightcone slices – Reducing Sample Variance:

As there were 600 samples for each neutral fraction in simulated volume, we were able to take the mean and variance, plot the means at different redshifts to see the mean trend. But for lightcone, we only have one sample at each neutral fraction. If we plot WST coefficients directly as a function of neutral fractions for lightcones, sample variance may significantly disturb the mean trends that we're interested in. So, 25 neutral fractions are binned together and averaged to reduce this sample variance. WST lightcone results before and after binning are plotted below and we can see there is a clear impact on the smoothness of WST coefficients.

4.2.4 WST Results comparison – Lightcones Vs Simulated Volumes:

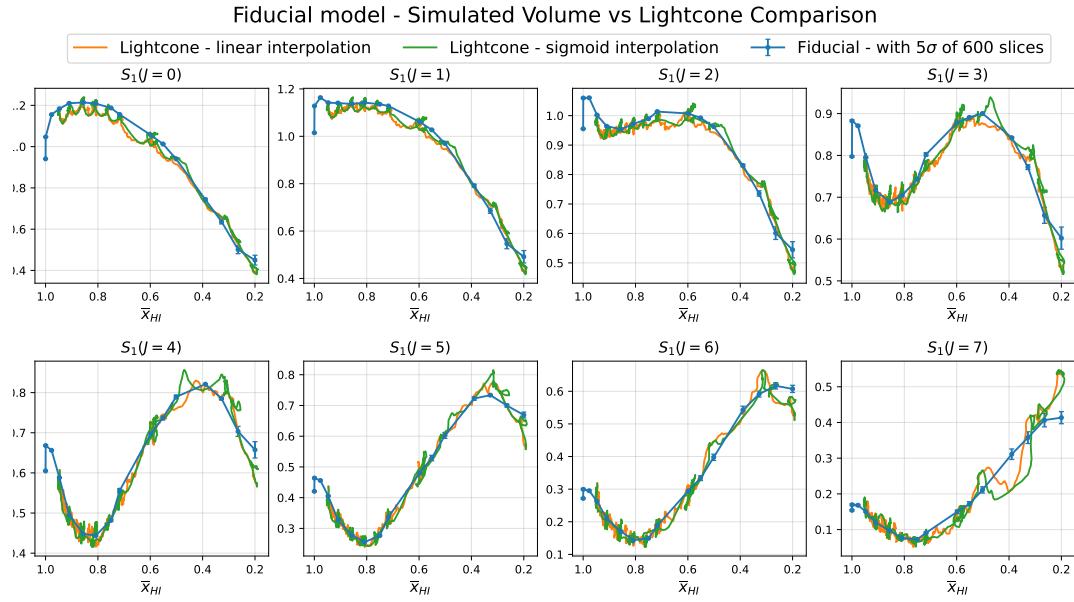
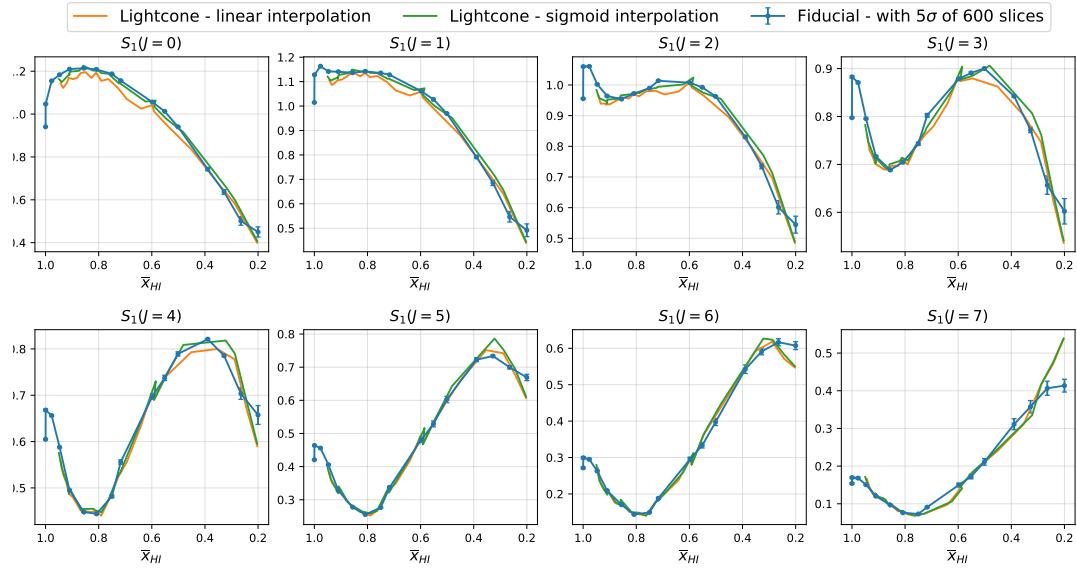
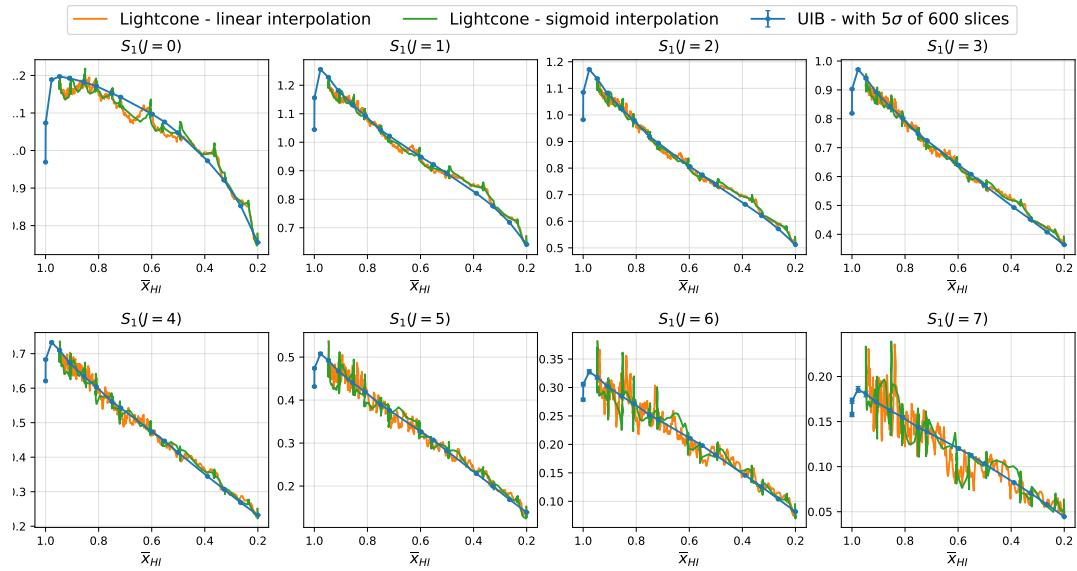


Figure 4.17: Comparing WST results of lightcone with WST results of simulated volumes

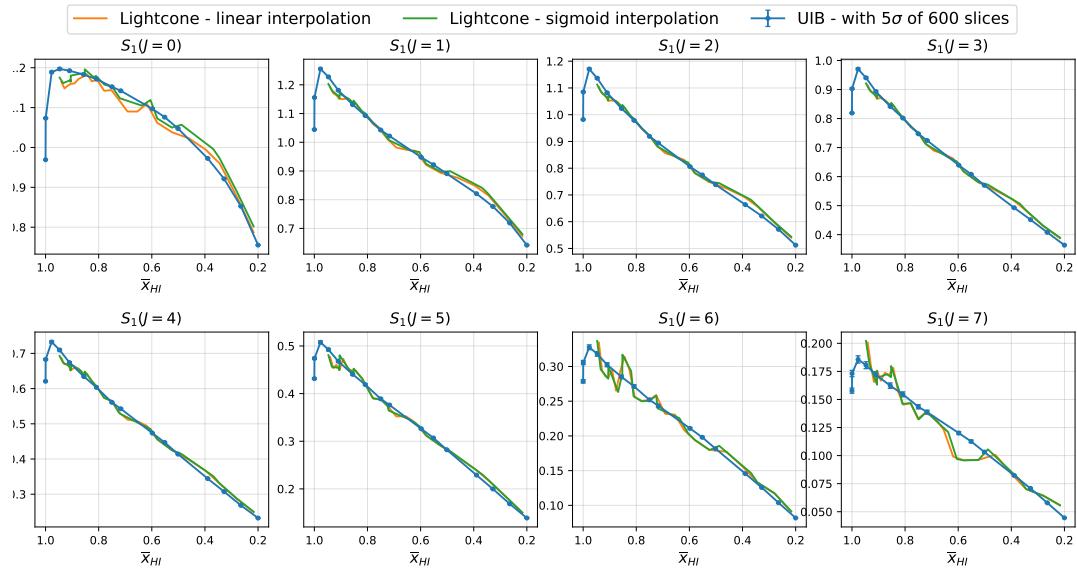
Lightcone : After Binnig



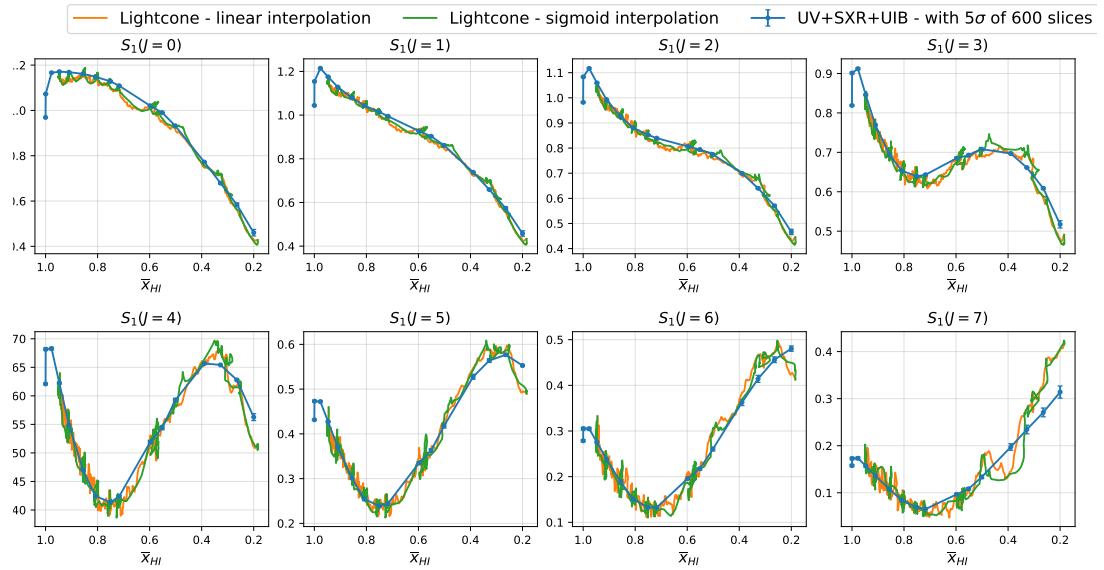
UIB model - Simulated Volume vs Lightcone Comparison



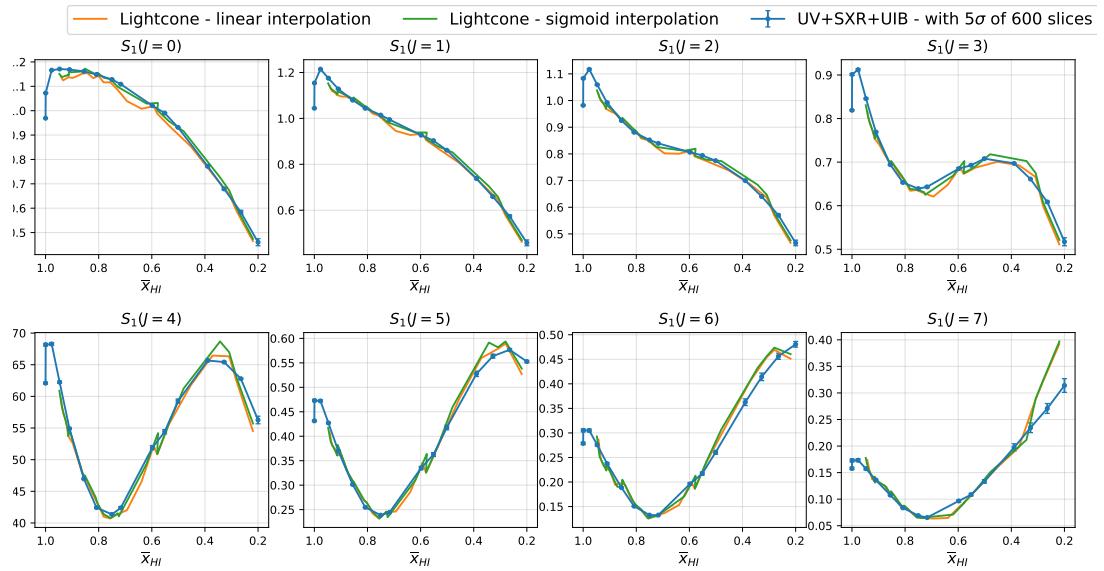
Lightcone : After Binnig



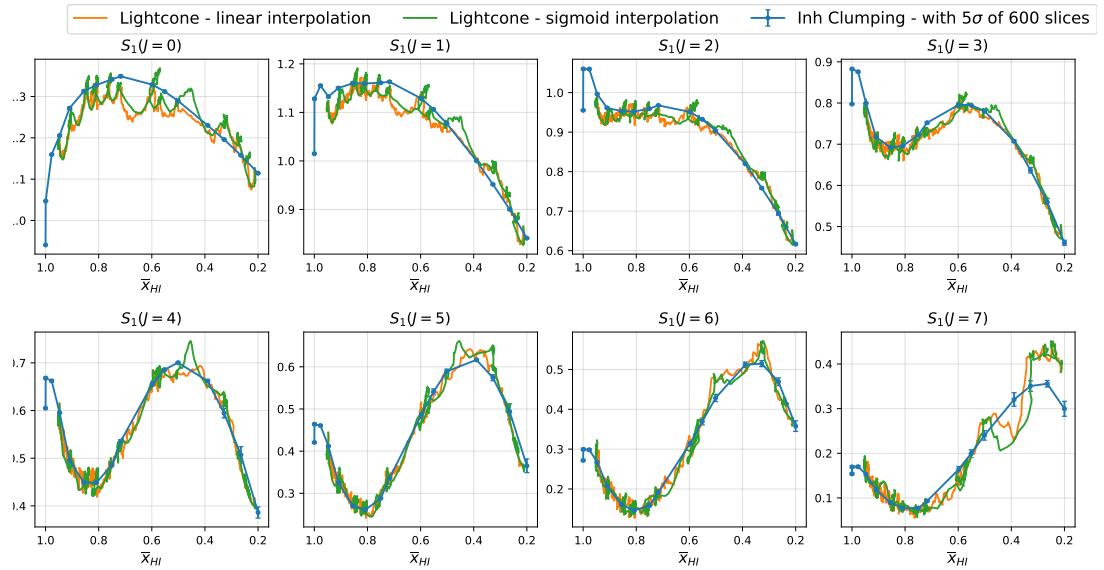
UV+SXR+UIB model - Simulated Volume vs Lightcone Comparison



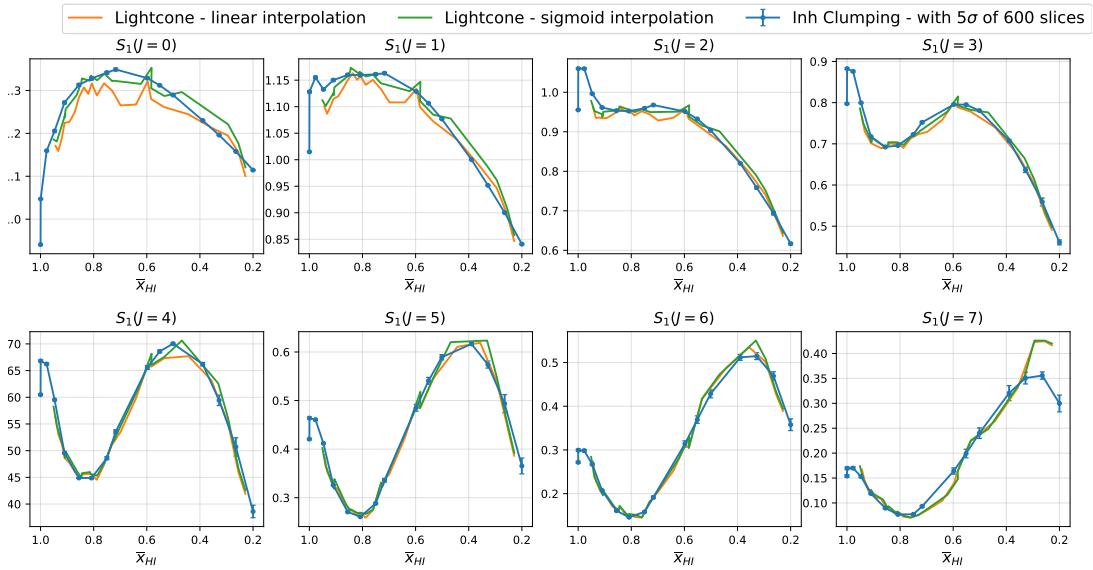
Lightcone : After Binnig



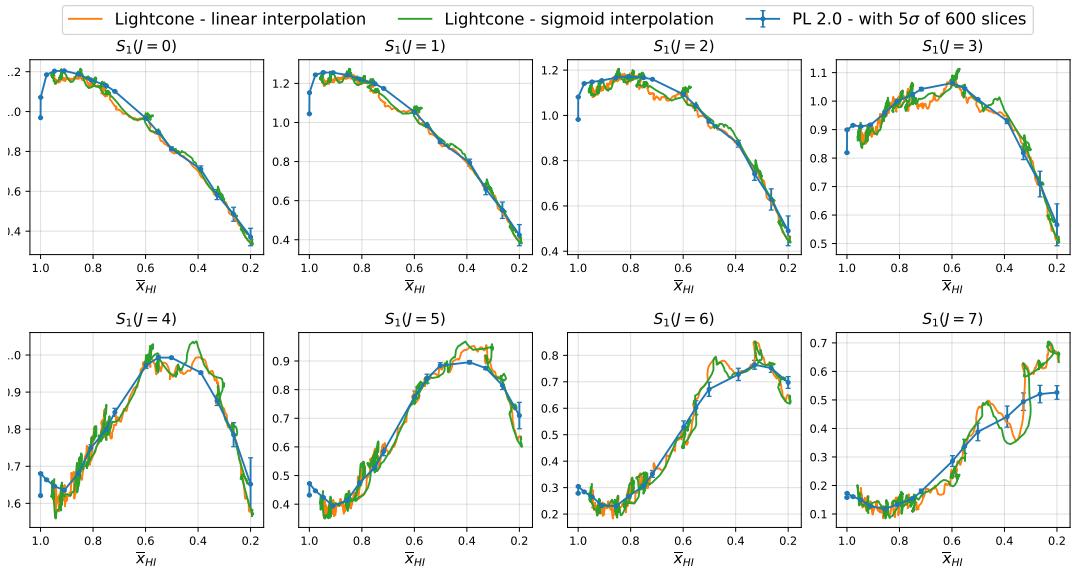
Inh Clumping model - Simulated Volume vs Lightcone Comparison



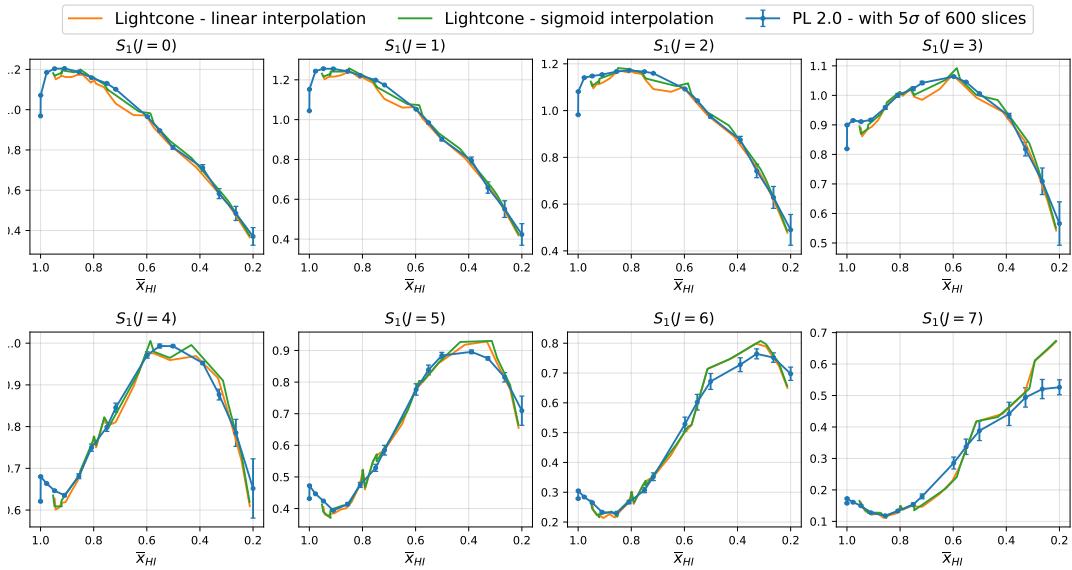
Lightcone : After Binnig



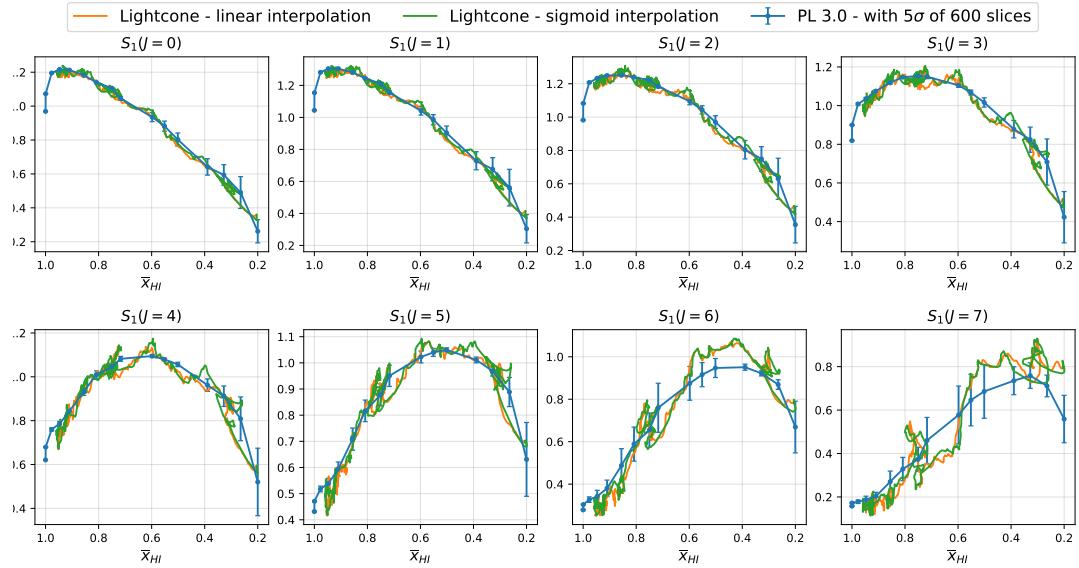
PL 2.0 model - Simulated Volume vs Lightcone Comparison



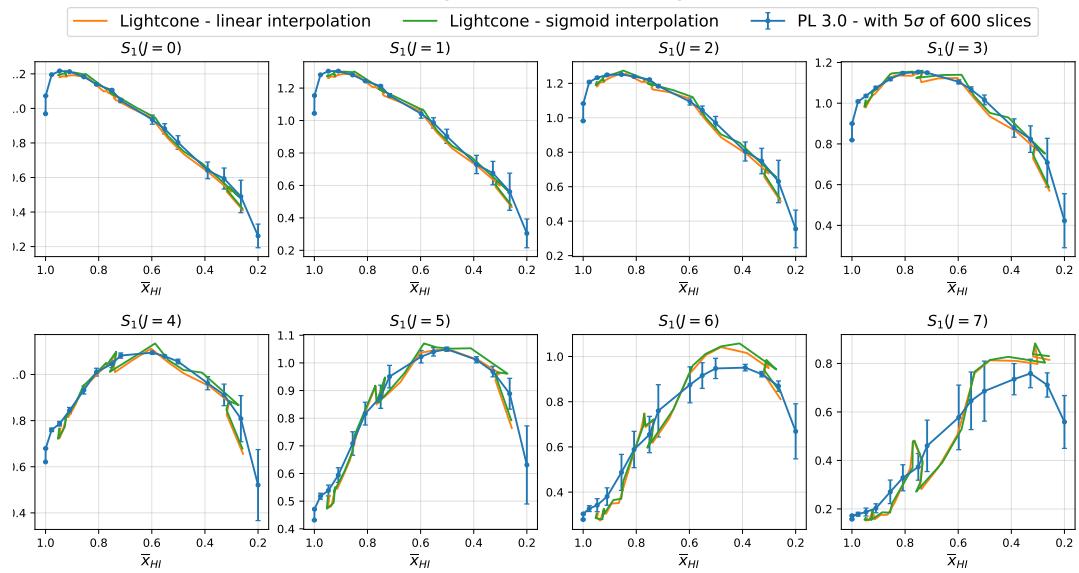
Lightcone : After Binnig



PL 3.0 model - Simulated Volume vs Lightcone Comparison



Lightcone : After Binnig



4.3 Comparing WST with other Statistics :

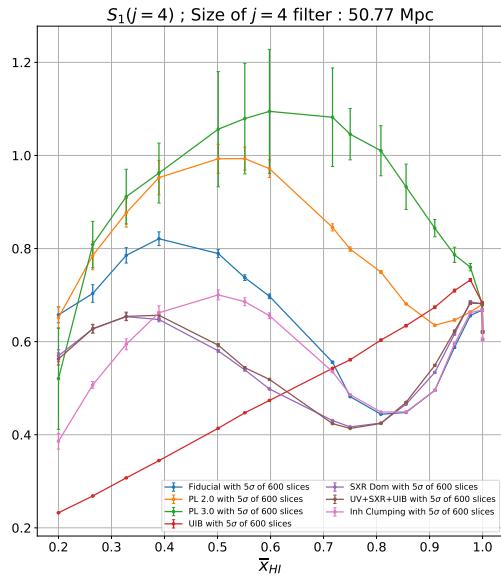
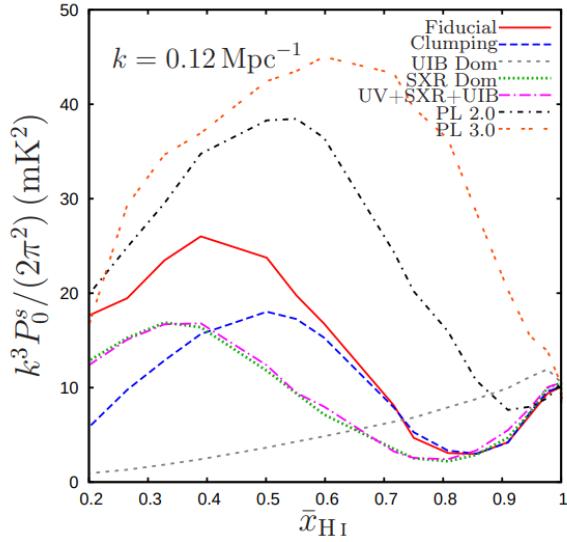


Figure 4.18: The figure on the top shows 3D power spectrum results on 21-cm line brightness temperature maps for all our models considered from Majumdar et al. [4]. The figure on the bottom shows 2D WST S_1 coefficient results for the same models. We can clearly see the similar trends in Power spectrum and WST S_1 results.

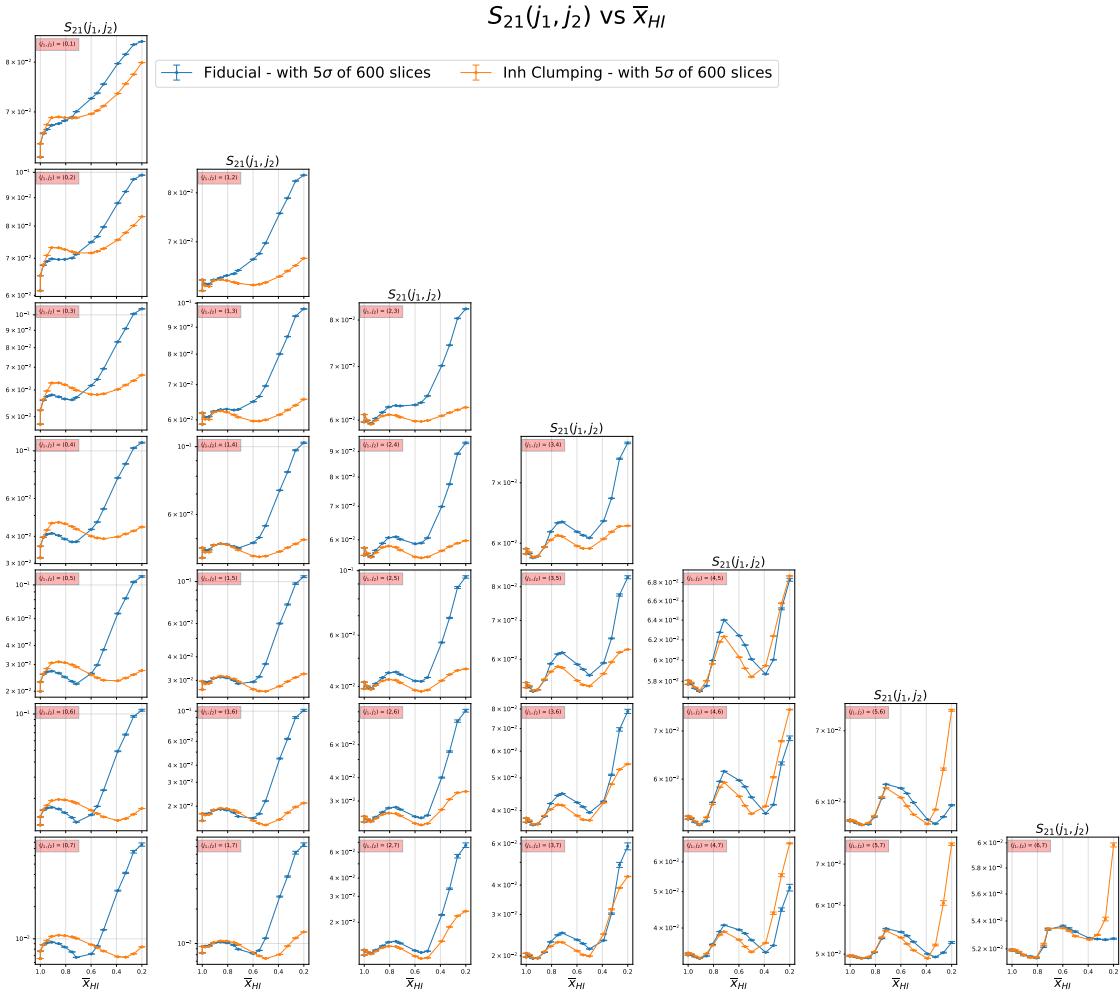
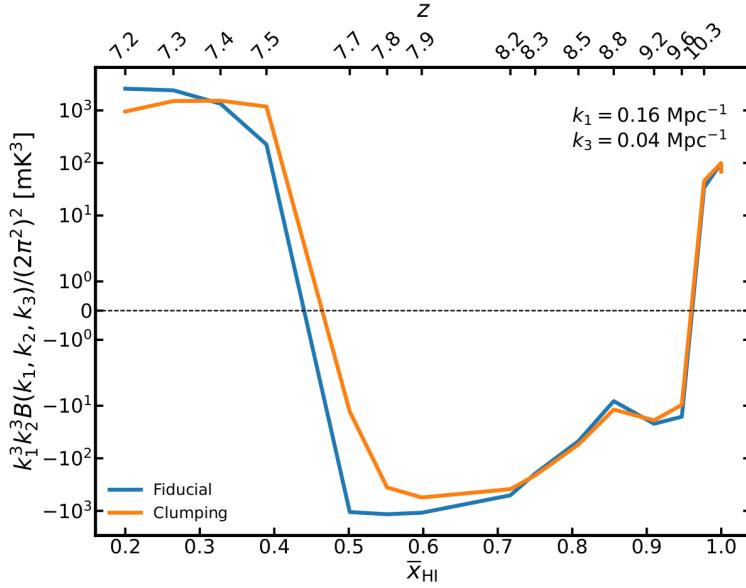


Figure 4.19: The figure on the top shows 3D Bispectrum results on 21-cm line brightness temperature maps for Fiducial & Clumping models. The figure on the bottom shows 2D WST S_2 coefficient results for the same models. We can clearly see that WST performs better at distinguishing the two models.

4.4 Summary of the results so far :

1. WST S_1 coefficient measures strength of fluctuations as a function of scale – **just like Power Spectrum from Fig 4.18**.
2. WST S_2 coefficient measures clustering strength of fluctuations of one scale at higher scales – **Scale-to-scale correlation : non-Gaussian information beyond Power Spectrum**.
3. With WST S_2 coefficient, 2D WST is as good as 3D Bispectrum, even better for some models considered in our work as shown in Fig 4.19.
4. Compared to 3D statistics, 2D WST has the advantage of being able calculate sample variance of images from just one simulation.
5. Lightcone WST S_1 results are consistent with WST S_1 results of evolving co-eval cubes as shown in Fig 4.17. This is in sharp contrast to power spectrum as its results are significantly effected by introducing lightcone effect to the simulations.

CHAPTER 5

FUTURE SCOPE

My work so far has been dedicated to assess WST's ability to extract information from 21-cm signal and distinguish various reionization source models from simulated data. An important follow-up to this is to include telescope beam effects & system noise and then study if the reionization models can still be distinguished using WST. With this step, we move one step closer to the observations we expect SKA to make, which would also come with foreground contamination from galactic synchotron and extragalactic point radio sources of the local universe. **Aim for my future work is to simulate mock observations constructed from reionization simulations and a telescope model to examine whether WST can still be used for extracting the properties of underlying sources on 21-cm brightness temperature images with realistic observation-like effects.**

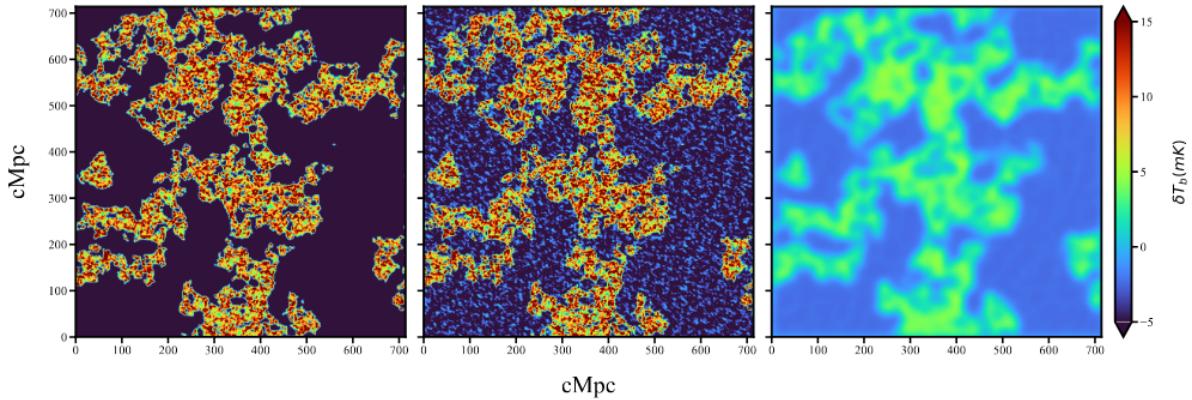


Figure 5.1: Interpreting the HI 21-cm cosmology maps through Largest Cluster Statistics - I: Impact of the synthetic SKA1-Low observations - Dasgupta et al. (2023) [46]

5.1 Impact of Noise :

As discussed in Ghara et al. [6], the system noise at different baselines and frequency channels are uncorrelated and expected to be Gaussian random variables with zero mean. The rms noise for each baseline and frequency channel of width $\Delta\nu_c$ and correlator integration time Δt_c is given by

$$\sigma = \frac{\sqrt{2}k_B T_{sys}}{A_{eff}\sqrt{\Delta\nu_c\Delta t_c}}$$

5.2 Impact of Telescope effects:

As the configuration and the antennas for SKA1-low are still being developed, for simulating mock observations we need to consider telescope parameters like number of antenna stations, diameter of each antenna, distribution of these stations, *uv* coverage and the observation time etc., as indicative for what SKA1-low is planned for and is able to achieve.

Parameters	Values
Observation time (t_{int})	1000 h
System temperature (T_{sys})	$60 \left(\frac{\nu}{300\text{MHz}} \right)^{-2.55} \text{ K}$
Effective collecting area (A_D)	962 m ²
Critical frequency (ν_c)	110 MHz

Figure 5.2: Telescope model considered in Ghara et al. [6] to match the SKA1-low capabilities as given in SKA documentation.

5.3 Reionization Model Selection with WST :

Previous studies [40] have shown that, for different reionization scenarios achieved by varying parameters of the same source model (fiducial), 2D WST can be used for constraining parameters. Our work clearly shows that WST excels at distinguishing various source models. So, our work can be extended to test WST's ability for model selection in Epoch of Reionization by considering different reionization source models and varying their parameters. To achieve this, WST needs to be applied on many simulated outputs by varying parameters from each model.

References :

- [1] Antoinette Songaila and Lennox L Cowie. "The evolution of Lyman limit absorption systems to redshift six". In: *The Astrophysical Journal* 721.2 (2010), p. 1448. URL: <https://iopscience.iop.org/article/10.1088/0004-637X/721/2/1448/meta>.
- [2] Xiaohui Fan and Michael A. Strauss. "Constraining the Evolution of the Ionizing Background and the Epoch of Reionization with $z \approx 6$ Quasars. II. A Sample of 19 Quasars". In: *The Astrophysical Journal* 132.1 (2006), p. 117. URL: <https://iopscience.iop.org/article/10.1086/504836/meta>.
- [3] Kanan K Datta et al. "Light cone effect on the reionization 21-cm signal-II. Evolution, anisotropies and observational implications". In: *Monthly Notices of the Royal Astronomical Society* 442.2 (2014), pp. 1491–1506.
- [4] Suman Majumdar et al. "Effects of the sources of reionization on 21-cm redshift-space distortions". In: *Monthly Notices of the Royal Astronomical Society* 456.2 (2016), pp. 2080–2094.
- [5] Garrelt Mellema et al. "HI tomographic imaging of the Cosmic Dawn and Epoch of Reionization with SKA". In: *arXiv preprint* (2015). URL: <https://arxiv.org/abs/1501.04203>.
- [6] Raghunath Ghara et al. "Imaging the redshifted 21 cm pattern around the first sources during the cosmic dawn using the SKA". In: *Monthly Notices of the Royal Astronomical Society* 464.2 (Sept. 2016), pp. 2234–2248. ISSN: 0035-8711. doi: 10.1093/mnras/stw2494. eprint: <https://academic.oup.com/mnras/article-pdf/464/2/2234/18518470/stw2494.pdf>. URL: <https://doi.org/10.1093/mnras/stw2494>.
- [7] N. Aghanim and Y. Akrami. "Planck 2018 results I. Overview and the cosmological legacy of Planck". In: *Astronomy & Astrophysics* 641.A1 (2018). URL: https://www.aanda.org/articles/aa/full_html/2020/09/aa33880-18/aa33880-18.html.

- [8] Daniel J Mortlock et al. "A luminous quasar at a redshift of z= 7.085". In: *Nature* 474.7353 (2011), pp. 616–619. URL: <https://www.nature.com/articles/nature10159>.
- [9] Bradley Greig et al. "Are we witnessing the epoch of reionization at z= 7.1 from the spectrum of J1120+0641?" In: *Monthly Notices of the Royal Astronomical Society* 466.4 (2017), pp. 4239–4249. URL: <https://academic.oup.com/mnras/article/466/4/4239/2738738>.
- [10] Eduardo Banados et al. "An 800-million-solar-mass black hole in a significantly neutral Universe at a redshift of 7.5". In: *Nature* 553.7689 (2018), pp. 473–476. URL: <https://www.nature.com/articles/nature25180>.
- [11] Bradley Greig, Andrei Mesinger, and Eduardo Banados. "Constraints on reionization from the z= 7.5 QSO ULASJ1342+ 0928". In: *Monthly Notices of the Royal Astronomical Society* 484.4 (2019), pp. 5094–5101.
- [12] Frederick B Davies et al. "Quantitative constraints on the reionization history from the IGM damping wing signature in two quasars at z> 7". In: *The Astrophysical Journal* 864.2 (2018), p. 142. URL: <https://iopscience.iop.org/article/10.3847/1538-4357/aad6dc/meta>.
- [13] Feige Wang et al. "A Significantly Neutral Intergalactic Medium Around the Luminous z= 7 Quasar J0252–0503". In: *The Astrophysical Journal* 896.1 (2020), p. 23. URL: <https://iopscience.iop.org/article/10.3847/1538-4357/ab8c45/meta>.
- [14] Alexa M Morales et al. "The evolution of the lyman-alpha luminosity function during reionization". In: *The Astrophysical Journal* 919.2 (2021), p. 120. URL: <https://iopscience.iop.org/article/10.3847/1538-4357/ac1104/meta>.
- [15] Intae Jung et al. "Texas spectroscopic search for Ly α emission at the end of reionization. III. The Ly α equivalent-width distribution and ionized structures at z> 7". In: *The Astrophysical Journal* 904.2 (2020), p. 144. URL: <https://iopscience.iop.org/article/10.3847/1538-4357/abbd44/meta>.
- [16] Charlotte A Mason et al. "The universe is reionizing at z > 7: bayesian inference of the IGM neutral fraction using Ly α emission from galaxies". In: *The Astrophysical Journal* 856.1 (2018), p. 2. URL: <https://iopscience.iop.org/article/10.3847/1538-4357/aab0a7/meta>.
- [17] N Aghanim et al. "Planck 2018 results". In: *A&A* 641 (2020), A12. URL: <https://www.aanda.org/articles/aa/abs/2020/09/aa33910-18/aa33910-18.html>.
- [18] Prakash Gaikwad et al. "Probing the thermal state of the intergalactic medium at z> 5 with the transmission spikes in high-resolution Ly α forest spectra". In: *Monthly Notices of the Royal Astronomical Society* 494.4 (2020), pp. 5091–5109. URL: <https://academic.oup.com/mnras/article-abstract/494/4/5091/5818338>.
- [19] Elisa Boera et al. "Revealing reionization with the thermal history of the intergalactic medium: new constraints from the Ly α flux power spectrum". In: *The Astrophysical Journal* 872.1 (2019), p. 101. URL: <https://iopscience.iop.org/article/10.3847/1538-4357/aafcc4/meta>.
- [20] Steven R Furlanetto, S Peng Oh, and Frank H Briggs. "Cosmology at low frequencies: The 21 cm transition and the high-redshift Universe". In: *Physics reports* 433.4-6 (2006), pp. 181–301. URL: <https://www.sciencedirect.com/science/article/pii/S0370157306002730>.
- [21] Harold Irving Ewen and Edward Mills Purcell. "Observation of a Line in the Galactic Radio Spectrum: Radiation from Galactic Hydrogen at 1,420 Mc./sec." In: *Nature* 168.4270 (1951), pp. 356–356. URL: <https://www.nature.com/articles/168356a0>.
- [22] Jonathan R Pritchard and Abraham Loeb. "21 cm cosmology in the 21st century". In: *Reports on Progress in Physics* 75.8 (2012), p. 086901. URL: <https://iopscience.iop.org/article/10.1088/0034-4885/75/8/086901/meta>.
- [23] Yashwant Gupta et al. "The upgraded GMRT: opening new windows on the radio Universe". In: *Current Science* (2017), pp. 707–714. URL: <https://www.jstor.org/stable/26293915>.
- [24] Michael P van Haarlem et al. "LOFAR: The low-frequency array". In: *Astronomy & astrophysics* 556 (2013), A2. URL: <https://www.aanda.org/articles/aa/abs/2013/08/aa20873-12/aa20873-12.html>.
- [25] Steven John Tingay et al. "The Murchison widefield array: The square kilometre array precursor at low radio frequencies". In: *Publications of the Astronomical Society of Australia* 30 (2013), e007. URL: <https://www.cambridge.org/core/journals/publications-of-the-astronomical-society-of-australia/article/murchison-widefield-array-the-square-kilometre-array-precursor-at-low-radio-frequencies/ED20FE56B17C253DAB94836785D887F0>.

- [26] Randall B Wayth et al. "The phase II Murchison widefield array: design overview". In: *Publications of the Astronomical Society of Australia* 35 (2018), e033. URL: <https://www.cambridge.org/core/journals/publications-of-the-astronomical-society-of-australia/article/phase-ii-murchison-widefield-array-design-overview/A416B810071E43E759C528E3888A83A6>.
- [27] Aaron R Parsons et al. "The precision array for probing the epoch of re-ionization: eight station results". In: *The Astronomical Journal* 139.4 (2010), p. 1468. URL: <https://iopscience.iop.org/article/10.1088/0004-6256/139/4/1468/meta>.
- [28] David R DeBoer et al. "Hydrogen epoch of reionization array (HERA)". In: *Publications of the Astronomical Society of the Pacific* 129.974 (2017), p. 045001. URL: <https://iopscience.iop.org/article/10.1088/1538-3873/129/974/045001>.
- [29] G Bernardi et al. "Foregrounds for observations of the cosmological 21 cm line-I. First Westerbork measurements of Galactic emission at 150 MHz in a low latitude field". In: *Astronomy & Astrophysics* 500.3 (2009), pp. 965–979. URL: <https://www.aanda.org/articles/aa/abs/2009/24/aa11627-09/aa11627-09.html>.
- [30] Tiziana Di Matteo et al. "Radio foregrounds for the 21 centimeter tomography of the neutral intergalactic medium at high redshifts". In: *The Astrophysical Journal* 564.2 (2002), p. 576. URL: <https://iopscience.iop.org/article/10.1086/324293/meta>.
- [31] Sk Saiyad Ali, Somnath Bharadwaj, and Jayaram N Chengalur. "Foregrounds for redshifted 21-cm studies of reionization: Giant Meter Wave Radio Telescope 153-MHz observations". In: *Monthly Notices of the Royal Astronomical Society* 385.4 (2008), pp. 2166–2174. URL: <https://academic.oup.com/mnras/article/385/4/2166/1035667>.
- [32] V Jelić et al. "Foreground simulations for the LOFAR–epoch of reionization experiment". In: *Monthly Notices of the Royal Astronomical Society* 389.3 (2008), pp. 1319–1335. URL: <https://academic.oup.com/mnras/article/389/3/1319/1019026>.
- [33] Catherine A Watkinson and Jonathan R Pritchard. "Distinguishing models of reionization using future radio observations of 21-cm 1-point statistics". In: *Monthly Notices of the Royal Astronomical Society* 443.4 (2014), pp. 3090–3106. URL: <https://academic.oup.com/mnras/article-abSTRACT/443/4/3090/1016826>.
- [34] A Watkinson et al. "A fast estimator for the bispectrum and beyond—a practical method for measuring non-Gaussianity in 21-cm maps". In: *Monthly Notices of the Royal Astronomical Society* 472.2 (2017), pp. 2436–2446. URL: <https://academic.oup.com/mnras/article/472/2/2436/4091435>.
- [35] Rajesh Mondal et al. "The effect of non-Gaussianity on error predictions for the Epoch of Reionization (EoR) 21-cm power spectrum". In: *Monthly Notices of the Royal Astronomical Society: Letters* 449.1 (2015), pp. L41–L45. URL: <https://academic.oup.com/mnrasl/article/449/1/L41/1028566>.
- [36] LVE Koopmans et al. "The cosmic dawn and epoch of reionization with the square kilometre array". In: *arXiv preprint* (2015). URL: <https://arxiv.org/abs/1505.07568>.
- [37] S Cheng. "Astrophysics and cosmology with the scattering transform". PhD dissertation. John Hopkins University, 2021. URL: https://pages.jh.edu/scheng40/Dissertation_SihaoCheng.pdf.
- [38] Sihao Cheng et al. "A new approach to observational cosmology using the scattering transform". In: *Monthly Notices of the Royal Astronomical Society* 499.4 (2020), pp. 5902–5914. URL: <https://academic.oup.com/mnras/article/499/4/5902/5924461>.
- [39] Sihao Cheng and Brice Ménard. "How to quantify fields or textures? A guide to the scattering transform". In: *arXiv preprint* (2021). URL: <https://arxiv.org/abs/2112.01288>.
- [40] Bradley Greig, Yuan-Sen Ting, and Alexander A Kaurov. "Exploring the cosmic 21-cm signal from the epoch of reionization using the wavelet scattering transform". In: *Monthly Notices of the Royal Astronomical Society* 513.2 (2022), pp. 1719–1741. URL: <https://academic.oup.com/mnras/article/513/2/1719/6565815>.
- [41] Bradley Greig, Yuan-Sen Ting, and Alexander A Kaurov. "Detecting the non-Gaussianity of the 21-cm signal during reionization with the wavelet scattering transform". In: *Monthly Notices of the Royal Astronomical Society* 519.4 (2023), pp. 5288–5303. URL: <https://academic.oup.com/mnras/article/519/4/5288/6969426>.
- [42] Tirthankar Roy Choudhury, Martin G Haehnelt, and John Regan. "Inside-out or outside-in: the topology of reionization in the photon-starved regime suggested by Ly α forest data". In: *Monthly Notices of the Royal Astronomical Society* 394.2 (2009), pp. 960–977.

- [43] Stéphane Mallat. "Group invariant scattering". In: *Communications on Pure and Applied Mathematics* 65.10 (2012), pp. 1331–1398. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/cpa.21413>.
- [44] Velasco Sameul. "How Wavelets allow researchers to transform and understand data". In: *Quanta Magazine* (2021). URL: <https://www.quantamagazine.org/how-wavelets-allow-researchers-to-transform-and-understand-data-20211013/>.
- [45] Nick Kingsbury. "Complex Wavelets for Shift Invariant Analysis and Filtering of Signals". In: *Applied and Computational Harmonic Analysis* 10.3 (2001), pp. 234–253. ISSN: 1063-5203. URL: <https://doi.org/10.1006/acha.2000.0343>.
- [46] Saswata Dasgupta et al. "Interpreting the HI 21-cm cosmology maps through Largest Cluster Statistics. Part I. Impact of the synthetic SKA1-Low observations". In: *Journal of Cosmology and Astroparticle Physics* 2023.05 (May 2023), p. 014. DOI: [10.1088/1475-7516/2023/05/014](https://dx.doi.org/10.1088/1475-7516/2023/05/014). URL: <https://dx.doi.org/10.1088/1475-7516/2023/05/014>.