



Imaging the 21-cm Signal from the Cosmic Dawn & Epoch of Reionization and the Connection with the Global Signal

Satadru Bag^{1,2}, Michele Bianco^{3,4}, Samir Choudhuri⁵, Saswata Dasgupta⁶,
Abhirup Datta⁷, Kanan K. Datta⁸, Ivelin Georgiev^{9,10}, Sambit K. Giri^{9,11}, Ian
Hothi¹², Akanksha Kapahtia¹³, Suman Majumdar⁷, Garrelt Mellema¹⁰, Arnab
Mishra⁸, Samit Kumar Pal⁷ and Feiyu Zhao¹⁴

¹Max-Planck-Institut für Astrophysik, Garching D-85748, Germany

²Technical University of Munich, TUM School of Natural Sciences, Physics Department,
James-Franck-Straße 1, 85748 Garching, Germany

³Laboratoire d'Astrophysique, École Polytechnique Fédérale de Lausanne (EPFL), Observatoire de
Sauverny, Chemin Pegasi 51, CH-1290 Versoix, Switzerland

⁴Institute for Particle Physics and Astrophysics, ETH Zurich, Wolfgang-Pauli-Str 27, CH-8093 Zurich,
Switzerland

⁵Centre for Strings, Gravitation and Cosmology, Department of Physics, Indian Institute of Technology
Madras, Chennai 600036, India

⁶Institute of Astronomy & Kavli Institute for Cosmology, University of Cambridge, Cambridge, UK

⁷Department of Astronomy, Astrophysics and Space Engineering, Indian Institute of Technology Indore,
Indore 453552, India

⁸Department of Physics, Jadavpur University, Kolkata, India

⁹Department of Astronomy & Oskar Klein Centre, Stockholm University, AlbaNova University Centre,
Stockholm, Sweden

¹⁰ARCO (Astrophysics Research Center), Department of Natural Sciences, The Open University of Israel, 1
University Road, PO Box 808, Ra'anana 4353701, Israel

¹¹Van Swinderen Institute for Particle Physics and Gravity, University of Groningen, Nijenborgh 3, 9747 AG
Groningen, The Netherlands

¹²Laboratoire de Physique de l'ENS, ENS, Université PSL, CNRS, Sorbonne Université, Université Paris
Cité, 75005, Paris, France

¹³Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany

¹⁴Shanghai Astronomical Observatory, Chinese Academy of Sciences, 80 Nandan Road, Shanghai 200030,
China

E-mail: garrelt.mellema@astro.su.se

The original baseline design for SKA-Low was motivated by the ability to produce tomographic images of the redshifted 21-cm signal, thus allowing the research field to move beyond the simple statistic of the power spectrum. In this chapter we review the imaging capabilities of SKA-Low, the wide variety of methods proposed for quantitatively analysing image data, as well as the connection with the global 21-cm signal.

Contents

1	Introduction	2
2	Imaging capabilities of AA* and AA4	3
3	Identifying structures in images	6
3.1	Structures from image data	6
3.2	Structures from non-image data	7
4	Analysing structures identified in images	9
4.1	Size distributions	9
4.2	Topological quantities	10
5	Analysing images without identifying structures	12
5.1	Scattering transform	12
5.2	Neural networks	13
5.3	Other analysis methods	13
6	Combining image data with other observables	13
6.1	Cross-correlation of images with other observables	13
6.2	Ly- α forest - IGM tomography	14
6.3	Connecting individual ionized bubbles to the ionizing sources: a multiwavelength synergy with SKA	14
7	The connection with the global 21-cm signal	15
7.1	Measuring the global 21-cm signal with short baselines	15
7.2	Measuring the global 21-cm signal from the images	17
8	Conclusions	17
9	Author contributions	18

1 Introduction

The original baseline design for SKA-Low (Dewdney et al., 2013) envisaged an interferometer which would completely transform the study of the 21-cm signal from the Cosmic Dawn (CD) and Epoch of Reionization (EoR) by providing sufficient signal-to-noise ratio to enable imaging the signal at scales of a few arcminutes. This was revolutionary as none of the interferometers operational or planned at that time had images as its science goal; they rather focused on a detection of the statistical 21-cm power spectrum. Producing images of an intergalactic medium consisting of ionized regions embedded in still neutral areas will not only make the reionization process more visible and real for a wider science and lay audience, but also enable answering a range of questions which the power spectrum is unable to address.

The system currently under construction, with a first goal defined as AA* and a final goal as AA4, retains this imaging capability. The first results from AA* will focus on a statistical characterization of the signal in the form of power spectra. Given that none of the precursors has yet achieved a detection, this may be the first detection of the redshifted 21-cm signal. However, subsequent studies with SKA-Low will undoubtedly focus on extracting images, or rather image cubes, since the broad frequency coverage of SKA-Low allows tomographic imaging of the cosmological 21-cm signal.

The SKA Science Book from 2015 contained two chapters which specifically focused on tomographically imaging the 21-cm CD / EoR signal (Mellema et al., 2015; Wyithe et al., 2015). Those chapters still provide a good description of the basic concepts. The current chapter provides an update, considering the specific imaging capabilities of SKA-Low AA* and AA4 (Section 2). Furthermore, the past decade has seen major progress in the development of techniques to identify structures in noisy images, which we will summarize in Section 3. Section 4 will provide an overview of the large body of work describing different methods to quantitatively analyze tomographic image data. These quantitative analysis methods can be used in inference studies, and the chapter on inference (Acharya et al., 2026a) describes some examples of this. However, it is not necessary to identify structures in order to quantitatively analyze tomographic image data. Section 5 describes methods that can produce a quantitative characterization of the data without relying on the identification of structures. A large advantage of imaging data is that it is associated with a specific part of the sky and thus allows synergistic studies with other types of observables for that patch of sky. Section 6 provides an overview of what synergetic research is enabled by imaging data, which is described in more detail in the chapter on synergetic observations (Chakraborty et al., 2026). A special case is the connection with the global 21-cm signal, sometimes called the sky-averaged signal or the monopole. Section 7 describes how imaging can be used to derive the 21-cm global signal, as well as how short SKA baselines could be used to measure the global signal.

2 Imaging capabilities of AA* and AA4

Figure 1 shows simulated images of the 21-cm signal with the AA* roll-out phase of SKA-Low. The three columns show the noiseless image, an image for an integration time of 1000 hours, and one with an integration time of 250 hours. The top row uses baselines up to 1 km (the full core), whereas the bottom row has half the resolution by only retaining baselines up to 500 m. The chosen redshift is $z = 7$ ($\nu = 177.6$ MHz). The bandwidth for these images is matched to their angular resolution and thus 0.92 and 1.85 MHz respectively. All images include the effect of the primary (station) beam which is assumed to have a Gaussian shape. The extend of its FWHM is indicated by a black dashed circle. The lower integration time of 250 hours can on the one hand be considered to give an indication of the possibilities of early imaging. However, as most interferometric images struggle to reach their theoretical noise level, they could also be seen as imperfectly calibrated images of 1000 hours of integration. These images do not contain any residual foreground effects and were produced with the `TOOLS21CM` package (Giri et al., 2020).

We can conclude that the high resolution ($B_{\max} = 1$ km or $\sim 6'$) images are heavily noise-dominated and cannot be used to extract any features. Reducing the resolution to $12'$ yields a much better

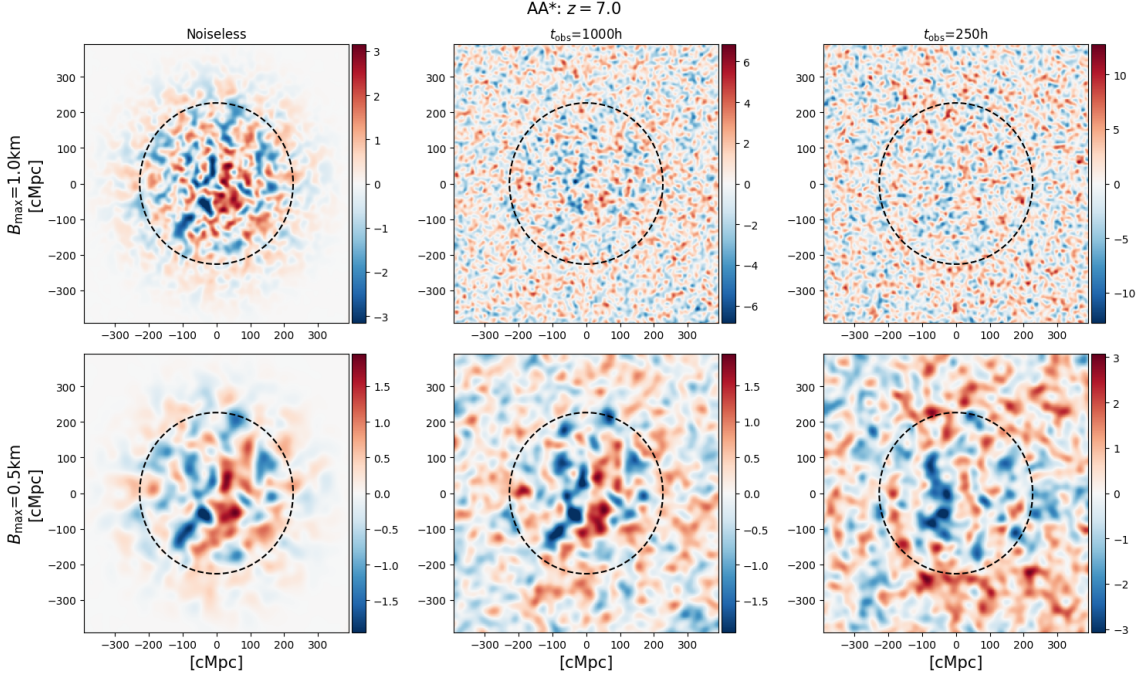


Figure 1: Sample 21-cm images for the AA* configuration of SKA-Low for a redshift $z = 7$ ($\sim 60\%$ ionized universe in the chosen model). The top row shows images using all baselines up to 1 km, the bottom row only uses baselines up to 500 m. The left column shows the noiseless image (including the primary beam), the middle column the image for a nominal integration time of 1000 hours, and right column the image for an integration time of 250 hours.

representation of the underlying structures, at least at the lower noise level. At the higher noise level, the features are clearly affected by noise and the analysis should be performed keeping this in mind.

Figure 2 shows the same images but now using the full Baseline Design, AA4. The additional sensitivity produces improvements for all cases. For example, the higher resolution, low noise case (middle panel top row) shows better correlation with the noiseless image. We can therefore conclude that the AA4 version of SKA-Low will have superior imaging to the AA* one.

The above results are somewhat idealized as they did not consider the impact of residual foregrounds or calibration errors. Some insights on the impact of those observational effects can be obtained from the results of Dasgupta et al. (2023). Those authors produced realistic SKA-Low 21-cm maps simulated via a synthetic radio observation as per the characteristics of the upcoming SKA-Low AA4 configuration. The synthetic maps were produced using the 21cmE2E pipeline (Mazumder et al., 2022; Pal et al., 2025b), based on OSKAR¹ software for SKA-Low, and CASA². They used different deconvolution algorithms (such as Högbom and Multiscale) and various weighting schemes (natural and Briggs weighting) to investigate their effects on the morphological features of ionized regions. Pal et al. (2025a) extended this analysis by introducing the effects of antenna-based

¹<https://ska-telescope.gitlab.io/sim/oskar/>

²<https://casa.nrao.edu/>

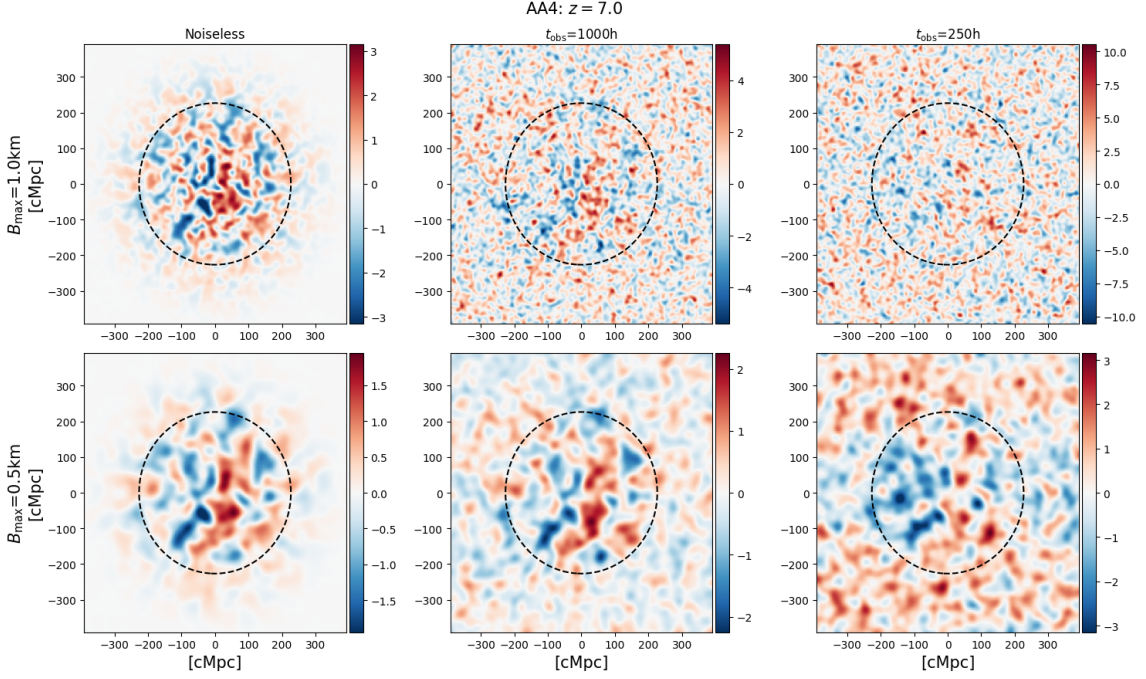


Figure 2: As Fig. 1 but for the Baseline Design AA4.

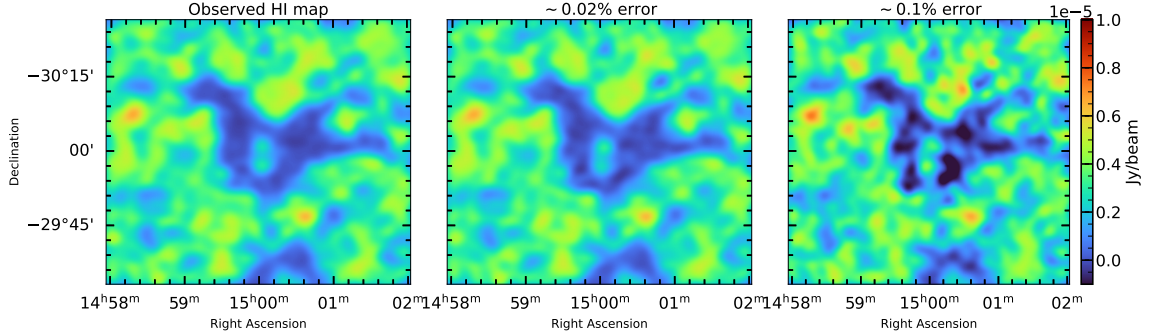


Figure 3: Impact of residual post calibration errors on the observed 21-cm maps via synthetic SKA-Low (AA4) observation at $z = 7.76$ ($\sim 40\%$ ionized universe in the chosen model) with angular resolution of $2.7'$ (7.1cMpc). Results from Pal et al. (2025a).

gain calibration errors into the 21-cm observation maps. Figure 3 illustrates these effects, showing how varying levels of residual post-calibration errors affect the reconstruction of the 21-cm maps. This work highlights how calibration errors and residual foreground contamination artifacts can introduce artificial structures which can affect the identification of structures (see Section 3).

These results highlight that although SKA-Low in principle can produce tomographic images of the 21-cm signal, it will not be trivial to produce data in which structures can be robustly identified. Both low noise levels and excellent calibration and foreground removal will be critical to achieve high image fidelity.

3 Identifying structures in images

As is evident from the results in the previous section, image cubes from SKA-Low can be expected to contain a fair level of noise and artifacts such as residual foreground structures or instrumental effects. Most of the techniques for analyzing tomographic image data are based on analyzing structures seen in the data, e.g. ionized regions or neutral islands. The first step is therefore to identify reliably such structures in noisy data. This section provides an overview of the different methods that have been proposed to achieve this. In the first subsection, we consider techniques that work on the image data, whereas the second subsection describes how certain types of structure can be identified directly in the visibility data.

3.1 Structures from image data

The 21-cm signal image data contains patterns that can be exploited to identify structures. For example, the 21-cm signal is produced by the neutral areas in the IGM, which in absence of spin temperature fluctuations leads to a bimodal distribution of ionized areas with no signal and neutral areas with signal. In the neutral areas the signal will have a range of values due to the density fluctuations. The response of the interferometer may still assign some specific value to the areas without signal but that can be considered as a constant offset. It is this bimodal distribution of signal values which allows the separation of ionized and neutral regions.

The actual separation will be complicated by the presence of noise, residual foregrounds, and finite resolution. Several methods have been proposed to choose a threshold value to segment image data into ionized and neutral regions. [Kakiichi et al. \(2017\)](#) simply used the mean value of the image data as the threshold. [Giri et al. \(2018a\)](#) proposed using the K-means clustering algorithm (see [Kodinariya et al., 2013](#), for a review) to separate ionized and neutral resolution elements. [Giri et al. \(2018b\)](#) extended this approach adding the maximum deviation method to deal with the cases when the distribution is no longer clearly bimodal (see also [Gazagnes et al., 2021](#)). [Dasgupta et al. \(2023\)](#) instead used a gradient-descent method to find the local minima between the two peaks of the bimodal distribution. Yet another method was proposed by [Pal et al. \(2025a\)](#) in which unsharp masking is used to enhance edges between neutral and ionized regions before applying a global thresholding method.

A more advanced method was proposed by [Giri et al. \(2018b\)](#). It not only considers the value of the resolution elements, but also their environment by implementing the Simple Linear Iterative Clustering (SLIC) method. This method calculates distances in the four dimensional space of position and 21-cm signal. The grouping of resolution elements according to the SLIC method leads to the Superpixel method for image segmentation. This method can reconstruct the bimodal distribution even in the presence of noise. It was further extended to employ cross-entropy in [Giri et al. \(2025\)](#) to classify these superpixels into ionized and neutral regions.

Another approach to segment images into ionized and neutral regions is to train machine learning methods to recognize meaningful patterns by training it on simulated noisy data. For example, neural networks with U-shaped architecture can learn a mapping between two image datasets ([Ronneberger et al., 2015](#)). These networks have been employed to learn how to construct binary maps of ionized/neutral regions from noisy 21-cm images (e.g., [Bianco et al., 2021](#)). This approach

has been quite successful in identifying ionized/neutral regions in these images even in the presence of foreground residuals (e.g., [Gagnon-Hartman et al., 2021](#); [Bianco et al., 2024](#); [Kennedy et al., 2024](#)). We refer the interested readers to the chapter on machine learning methods ([Acharya et al., 2026b](#)) for more discussion.

The above studies focus on segmenting image data into neutral and ionized regions, assuming the bimodal distribution of signal values which can be expected when there are no substantial spin temperature fluctuations. During the Cosmic Dawn when the ionized regions are still very small, we expect such spin temperature fluctuations due to local differences in Lyman- α coupling and X-ray heating. Under these circumstances it may be much harder to segment the image data as spin temperature differences are expected to produce a much more continuous distribution of signal values. Given this it could of course be argued that segmentation of images into two different regimes is not really useful. Still, the problem of how to quantitatively analyse 21-cm image data in case of substantial spin temperature fluctuations has to date not been systematically studied, even though the quite large scales on which differences are likely to exist may allow the construction of very low resolution images of the Cosmic Dawn from SKA-Low observations ([Ross et al., 2021](#)).

3.2 Structures from non-image data

If it is known what kind of structure you are looking for, it is possible to look for its signature in the visibility data, without the need to construct images. This is the idea behind the so-called matched-filter method ([Datta et al., 2007, 2008](#); [Datta et al., 2012](#); [Majumdar et al., 2011, 2012](#); [Mishra et al., 2025](#)). This method searches for isolated spherical ionized regions which are expected to exist around high-redshift luminous sources such as quasars and galaxy clusters ([Mortlock et al., 2011](#); [Bañados et al., 2017](#); [Wang et al., 2018](#); [Matsuoka et al., 2019a,b](#); [Wang et al., 2021](#); [Witstok et al., 2025](#)). As it is working directly on the visibilities it is very efficient and this method may be particularly useful if it is hard to produce reliable images. This approach has already been applied to existing data from the MWA ([Trott et al., 2021](#)) and HERA ([Chen et al., 2026](#)).

As an example, the left panel of figure 4 shows a simulated HII bubble around a bright quasar at redshift $z = 7.1$, similar to the one reported in [Mortlock et al. \(2011\)](#). The map is simulated using parameters consistent with a SKA-Low AA4 observation. The input bubble parameters used are: radius $R_b \sim 24$ Mpc, centered in the field of view, and mean neutral hydrogen fraction $x_{\text{HI}} = 0.88$. A 100-hour observation with the AA * and AA4 configurations yields matched-filter detections with $\text{SNR} \approx 10.3$ and $\text{SNR} \approx 11.8$ (right panel of Figure 4), respectively for this HII bubble. The number of short baselines in the SKA-Low AA * and AA4 configurations is very similar; however, AA4 includes significantly more long baselines. Consequently, the difference in the SNR between the two configurations is small, since the HI 21-cm signal contributes very little at large baselines. However, the long baselines play an important role in foreground mitigation, particularly for modelling and subtracting point sources. Studies further show that the SNR scales with bubble size approximately as $\text{SNR} \propto R_b^{1.7}$ for SKA-Low AA * and $\text{SNR} \propto R_b^{1.95}$ for SKA-Low AA4. This highlights that larger ionized regions can be detected more readily, and the scaling relation provides a useful guide for predicting detectability across a range of bubble sizes and observational setups.

A Bayesian parameter estimation framework ([Ghara and Choudhury, 2020](#)) has also been applied

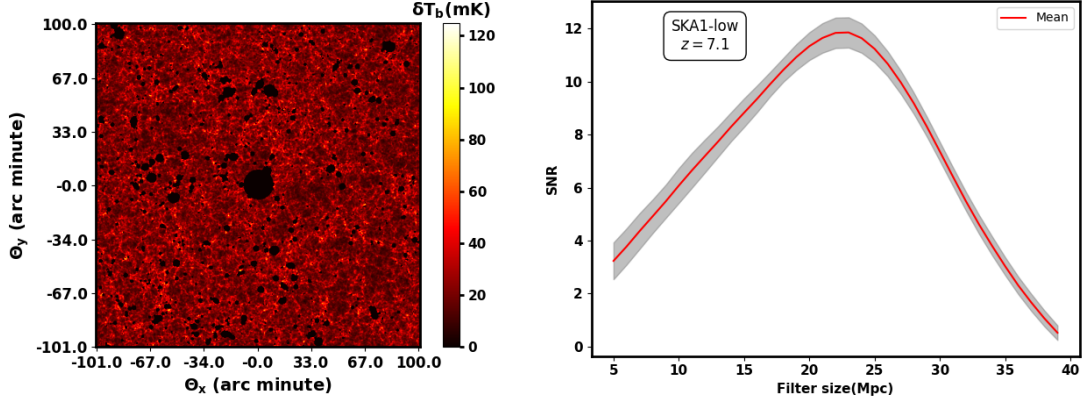


Figure 4: Left panel: Differential brightness temperature map of the HI 21-cm signal around an ionized bubble of comoving size ~ 24 Mpc at $z = 7.1$ corresponding to an observing frequency of 175 MHz. The angular size of simulated maps is $\sim 3.3^\circ \times 3.3^\circ$ with $\sim 23''$ resolution. Right panel: SNR of 20 independent noise realizations as a function of filter size for 100 hours of the SKA-Low AA4 observations at redshift 7.1. The shaded region shows 1σ uncertainty estimated from independent realizations. Both panels are taken from Mishra et al. (2025) for reference.

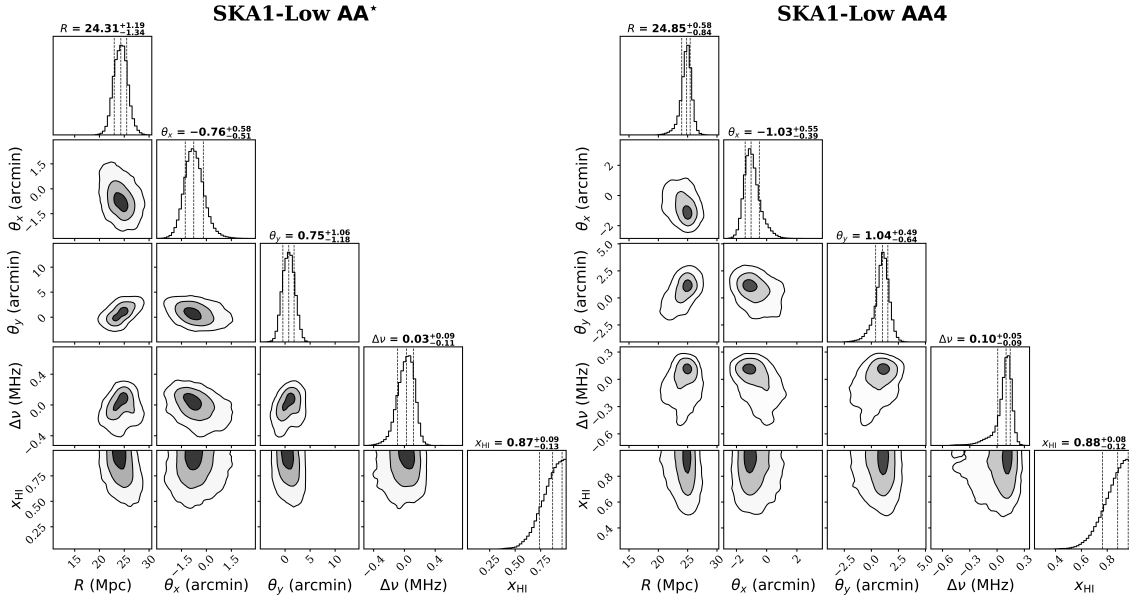


Figure 5: Posterior probability distributions of ionized bubble parameters derived from Bayesian analysis of the matched-filter outputs. The left panel corresponds to SKA-Low AA* and the right panel to SKA-Low AA4. The contours represent the 1σ , 2σ , and 3σ credible intervals, constraining bubble size, position, frequency offset, and neutral hydrogen fraction.

to constrain bubble properties such as size and location, as well as IGM parameters such as the neutral hydrogen fraction. The resulting posterior distributions are shown in Figure 5. These results demonstrate that a visibility-based matched filtering, combined with Bayesian inference, provides a robust framework for detecting and characterizing isolated ionized bubbles using SKA-Low AA4 and AA* observations.

4 Analysing structures identified in images

Once structures have been identified, the goal is to quantitatively analyse them. As the structures characteristic for CD/EoR are expected to be complex in shape, and when considering tomographic data, are actually three-dimensional, their analysis is far from trivial. In fact, there exist many different ways to analyse them. Over the past 10 years a considerable amount of effort has been dedicated to exploring different kinds of quantitative methods for describing the morphology of regions. This section provides an overview. The most obvious quantitative description is one which captures the sizes of identified structures and the first subsection is dedicated to such size distributions. The mathematical theory for describing shapes in multiple dimensions is known as topology, and the second subsection describes how topological quantities can be used to characterize structures in image data.

4.1 Size distributions

The size distribution and evolution of ionized bubbles during reionization are closely tied to the astrophysical processes driving the growth of ionization fronts and to the statistical properties of the ionization field (Furlanetto et al., 2006). Bubble size distributions (BSDs) quantify the characteristic scales of these regions and provide valuable information about the timing, morphology, and topology of reionization. A range of methods has been developed to model BSDs (Furlanetto et al., 2004; Doussot and Semelin, 2022) and to estimate them from simulated or observed 21 cm signal data (Giri et al., 2018a; Lin et al., 2016).

One of the earliest approaches is the friends-of-friends (FoF) algorithm (Iliev et al., 2006), which identifies connected ionized pixels and groups them into contiguous regions, from which bubble sizes are derived. As this method uses connected regions, it is closely connected to the topological quantities discussed below in Section 4.2. The typical size distribution produced by the FOF method does not resemble the other size distributions described next. As described in Furlanetto and Oh (2016), after percolation the FOF size distribution is dominated by one large connected region, the percolation cluster. The volume fraction of this dominant cluster relative to the total ionized volume is called the largest cluster statistic (LCS) and will be discussed in more detail below.

The spherical average method (SPA) (Zahn et al., 2007) derives a BSD by finding, for each ionized pixel, the largest sphere centered on it for which the average ionized fraction exceeds a chosen threshold. This method has been found to quite diffusive and has therefore become less popular. A more popular algorithm is the mean free path (MFP) method (Mesinger and Furlanetto, 2007; Mesinger et al., 2011), implemented for example in 21CMFAST, which employs a Monte Carlo procedure: random ionized pixels are selected, and the distance to the nearest neutral pixel along a random direction is recorded; the resulting histogram of distances traces the bubble size distribution. Another estimator for size distribution uses granulometry (Kakiichi et al., 2017). In this method the ionization field is repeatedly sieved with spherical holes of increasing radius R and the fraction of ionized pixels removed at each step is used to build a probability distribution of bubble sizes. Lin et al. (2016) introduced a three-dimensional extension of the watershed transform, which segments the ionization field into regions by identifying isodensity surfaces that trace bubble boundaries.

Each of these four methods produce somewhat different size distributions for the same input data.

However, the distributions are comparable in that they typically display a peak value and a roughly log-normal distribution of sizes. Comparisons of these methods applied to simulated maps without noise or foreground contamination have been presented in [Lin et al. \(2016\)](#), while [Giri et al. \(2018a\)](#) extended the analysis to include observational effects.

More recently, machine learning approaches have been explored; for example, [Shimabukuro et al. \(2022\)](#) trained artificial neural networks to recover MFP size distributions directly from the 21 cm power spectrum, demonstrating that this direct recovery yields higher accuracy than inferring size distributions from model parameter fits. However, [Choudhury et al. \(2025\)](#) showed artificial neural networks can only recover size distributions from power spectra produced by the reionization code they were trained on.

4.2 Topological quantities

In a 2D scalar field such as the 21-cm brightness temperature map, the excursion set is defined as regions above a certain chosen threshold. Its shape and connectivity captures both the morphology and topology of the underlying field. These excursion sets evolve during the EoR: ionized regions grow from isolated bubbles into a percolating, connected network, while neutral regions fragment over time ([Furlanetto and Oh, 2016](#)). The Friends-of-Friends size distribution introduced in Section 4.1 is in fact the size distribution of the region defined by the excursion set. Real-space statistics from images retain both amplitude and phase information, making them ideal for probing the highly non-Gaussian 21-cm field.

Perhaps the simplest topological quantity which can be defined is the Largest Cluster Statistic (LCS), already briefly mentioned above in Section 4.1. This single number, defined as

$$\text{LCS} = \frac{\text{volume of the largest ionized region}}{\text{total volume of all the ionized regions}}, \quad (1)$$

has been shown to be able to distinguish extreme reionization models ([Pathak et al., 2022](#)) and thus could be a simple but useful quantity to consider. A series of papers has further explored the potential of the LCS, also considering the practical challenges of measuring it in noisy and imperfectly calibrated image data ([Dasgupta et al., 2023](#); [Pal et al., 2025a](#)). For this and the other topology-related quantities described in this section, both noise, residual foreground and artifacts introduced by calibration errors can be especially damaging as they affect the connectivity. [Pal et al. \(2025a\)](#) concludes that for a minimum of 2000 hours of SKA-Low (AA4) integration and for a maximum antenna-based calibration error tolerance of $\sim 0.02\%$ after calibration, the reionization history can be recovered in a robust and relatively unbiased manner using LCS.

Morphological and topological descriptors have been widely used to study neutral and ionized regions in simulations and mock 21-cm images during reionization. One of the earliest morphological descriptors used for EoR fields is the *Scalar Minkowski Functionals (SMFs)*. In n dimensions, there are $n + 1$ SMFs. In 2D, they are contour length, area fraction, and genus (connected regions minus holes) of the excursion set. The genus has been used to study EoR phases ([Lee et al., 2008](#)) and distinguish reionization scenarios in noisy mock 21-cm images ([Hong et al., 2014](#); [Wang et al., 2015](#)). In 3D—relevant for SKA-low observations—SMFs are volume, surface area, integrated mean curvature, and Euler characteristic. The latter is a single number that expresses the topology

of a field by counting connected components, tunnels, and cavities. These have been used to study reionization morphology (Gleser et al., 2006; Friedrich et al., 2011; Yoshiura et al., 2017; Chen et al., 2019). It has been shown that when used in conjunction with the 21-cm power spectrum, the 3D SMFs give a 30% improvement in constraining reionization model parameters than when using the power spectrum alone (Diao et al., 2024).

Another interesting approach focuses on the geometry of individual ionized or neutral regions, including their morphological and topological properties. In n dimensions, n *Shapefinders* (SFs) are defined as ratios of $(n + 1)$ SMFs, with each SF having the dimension of length and characterizing the spatial extension of a region bounded by a closed surface (Sahni et al., 1998). SFs can be further combined to quantify the morphology of individual regions, e.g., their *planarity* and *filamentarity* in 3D. Using the shape diagnostic tool SURFGEN2 (Sheth et al., 2003; Sheth and Sahni, 2005), which implements a Marching Cubes 33 triangulation scheme to model the surfaces in 3D, Bag et al. (2018) showed that large ionized regions maintain a nearly constant cross-section at the percolation transition, while their lengths grow sharply, rendering them highly filamentary. The morphological properties of these regions, such as their characteristic cross-sections and the distributions of planarity and filamentarity at the onset of percolation, can be used to distinguish drastically contrasting reionization scenarios (Pathak et al., 2022). From the excursion set approach, the morphology of HI overdense and underdense regions similarly shows that, as reionization progresses, the cross-sections of large overdense regions become increasingly uniform (Bag et al., 2019).

Minkowski tensors (MTs) are generalizations of the SMFs, of which the *Contour Minkowski Tensor* (CMT) is the only linearly independent, translation-invariant 2D MT that provides anisotropy information beyond SMFs (Chingangbam et al., 2017). Since the eigenvalues of CMT encode both scale and shape information, it was shown that it could be used to define a characteristic ionized bubble size, without assuming spherical symmetry (Kapahtia et al., 2018). By mapping the morphology of the ionization, density, and spin temperature fields to that of the observable 21-cm brightness temperature Kapahtia et al. (2019) qualitatively showed that the CMT could serve as a promising tool for discerning various sources of non-Gaussianity in the 21-cm field.

Betti numbers (β_k) are topological invariants that count k -dimensional holes in the excursion set. In 3D, β_0 counts isolated regions, β_1 tunnels, and β_2 cavities (Pranav et al., 2019). They are related to the Euler characteristic χ through

$$\chi = \beta_0 - \beta_1 + \beta_2, \quad (2)$$

and thus carry more information than this SMF. They have been used to trace reionization stages and percolation in mock 21-cm cubes (Giri and Mellema, 2021), showing that they evolve in a characteristic and parameterizable way, from which meaningful topological insights can be obtained. It has also been shown that when combined with the CMT, can constrain EoR model parameters from a smoothed noisy image at a single frequency slice (Kapahtia et al., 2021).

Tracking Betti numbers as a function of a parameter (filtration) such as the field threshold, allows for *persistent homology*, which measures the birth and death of features as the filtration is varied. The richer topological information reveals distinct post-overlap reionization stages (Elbers and van

de Weygaert, 2019) and shows prospects for distinguishing astrophysical models of EoR with 1000h of HERA data (Elbers and van de Weygaert, 2023).

5 Analysing images without identifying structures

Instead of focusing on structures or shapes identified in image data, it is of course also possible to analyse the image data directly, that is, to subject it to some mathematical operation which should capture some of its essential properties. Actually, deriving the power spectrum is in essence such a method. However, as described elsewhere, the power spectrum is not sensitive to the non-Gaussian nature of the 21-cm fields. In this section we describe some methods proposed to process image data in order to capture its non-Gaussian character.

5.1 Scattering transform

Scattering transforms have emerged as powerful summary statistics for characterizing non-Gaussian fields (Mallat, 2012; Bruna and Mallat, 2013). Inspired by convolutional neural networks (CNNs), the scattering transforms rely on successive wavelet transforms, that separate a given field on its different oriented scales, as well as non-linear operations such as modulus, which allow to efficiently characterize the interactions between these different scales. Several generations of Scattering Transform statistics have been introduced : the initial Wavelet Scattering Transform (WST) (Allys et al., 2019; Cheng et al., 2020), Wavelet Phase Harmonics (WPH) (Allys et al., 2020), and the Scattering Covariance (Cheng et al., 2024). While the WST is highly compact set of summary statistics for characterization and parameter inference, the WPH and Scattering Covariances allow for additional application as generative modeling and component separation (see, for instance Auclair et al., 2024)

To get a sense of scattering transform statistics, let us consider the WST. Starting from an input field $I(x)$, the first layer of WST is built from the convolution with a set of wavelets parametrized by scale j and orientation θ :

$$S_{j,\theta}^1 = \langle |I \star \psi_{j,\theta}| \rangle, \quad (3)$$

where convolution is denoted by \star and spatial average by $\langle \cdot \rangle$. These first-order coefficients $S_{j,\theta}^1$ measure the amplitude of fluctuations at various oriented scales, resembling a scale-resolved power spectrum, but allowing to characterize sparsity and local structures (Cheng et al., 2020). To characterize interactions between scales, a second convolution with the set of wavelet if performed:

$$S_{j_1,\theta_1,j_2,\theta_2}^2 = \langle | |I \star \psi_{j_1,\theta_1}| \star \psi_{j_2,\theta_2} | \rangle. \quad (4)$$

These second-order scattering coefficients directly characterize the interaction between two scales. They capture non-Gaussian, higher-order correlations in the data, thereby circumventing the computational issues linked to conventional N-point statistics (Allys et al., 2019; Cheng et al., 2024).

Unlike the bispectrum, which characterises the interaction between individual Fourier modes, the scattering transforms rely on wavelets. These are bandpass filters that span an ensemble of Fourier modes. This broader mode coverage results in significantly lower variance for the estimated statistics, thereby improving robustness and stability (Cheng et al., 2024).

Scattering Transforms have already been applied to EoR simulations for parameter constraints (Greig et al., 2022; Hothi et al., 2024; Prelogović and Mesinger, 2024; Zhao et al., 2024; Shimabukuro et al., 2025), as well as to build generative model of EoR lightcones (Hothi et al., 2026).

5.2 Neural networks

Beyond summary statistics, machine learning techniques, particularly Convolutional Neural Networks (CNNs), offer a powerful paradigm for "end-to-end" parameter inference. This approach aims to directly map 21-cm images (or image cubes) to the underlying astrophysical parameters, bypassing human-defined summary statistics such as the power spectrum or size distributions. The primary motivation is that NNs are exceptionally adept at identifying the complex, non-linear, and non-Gaussian patterns in the 21-cm data that other statistics may miss. Several studies (e.g., Gillet et al., 2019; Hassan et al., 2019) have demonstrated the feasibility of this approach, showing that CNNs can recover key reionization physics from 21-cm images. Subsequent research has focused on improving robustness against realistic observational challenges such as instrumental noise and foreground contamination (Prelogović et al., 2022; Zhao et al., 2022). For detailed discussion, we refer the interested readers to the chapter on machine learning techniques (Acharya et al., 2026b).

5.3 Other analysis methods

In addition to the methods already discussed, several other statistics have been proposed to extract non-Gaussian information from 21-cm images. The local variance, for example, is a one-point statistic that leverages sample variance along the frequency direction of an image cube to trace the reionization history and ionization morphology (Gorce et al., 2021). Marked statistics, such as the marked power spectrum, re-weight the field based on local density, allowing for enhanced sensitivity to information in low- or high-density regions that the standard power spectrum may not capture well (Kamran et al., 2025). The evolving morphology of the neutral hydrogen distribution can also be characterized using generalized fractal dimensions, which quantify the complex, multifractal nature of the ionized regions and can be used to constrain the reionization history (Bandyopadhyay et al., 2017).

6 Combining image data with other observables

Image data contains positional information and so allows the 21-cm data to be combined with other observables. This section addresses some combinations of image data with data from different telescopes.

6.1 Cross-correlation of images with other observables

The most obvious path to combine the tomographic image data with other observables is to cross-correlate them. This has the added benefit that as calibration errors, foreground residuals and general systematics will be uncorrelated between the two data sets, the correlation can be studied even if the 21-cm image data is severely affected by such effects. As the synergy chapter (Chakraborty et al., 2026) contains an extensive description of the different kinds of cross-correlations that have been explored, we will not repeat them here but refer to that chapter. However, the next section does discuss one particular cross-correlation study which is not contained in that chapter.

6.2 Ly- α forest - IGM tomography

The Lyman- α forest, observed as absorption features in the spectra of distant quasars, is a cornerstone probe of the IGM (e.g., Fan et al., 2023, for a review). It provides 1D tomography of the (mostly ionized) IGM along individual lines-of-sight. This technique is most powerful at $z \lesssim 6$; at higher redshifts, the rapidly increasing IGM neutral fraction renders the forest completely saturated (the "Gunn-Peterson trough") (e.g., Becker et al., 2015), limiting its utility deep in the EoR. A complementary technique, photometric IGM tomography, aims to move beyond 1D skewers by using deep narrow-band (NB) imaging of numerous faint background galaxies to reconstruct a 2D or 3D map of the Ly- α forest transmission (e.g., Kakiichi et al., 2022). This mapping approach is particularly promising for identifying the final, large-scale neutral islands at the very end of reionization ($z \sim 5.7$) (Giri et al., 2025).

This is precisely where 21-cm tomography with SKA-Low becomes the dominant probe, as it is designed to directly image the 3D distribution of neutral gas within the EoR. A key synergy is the cross-correlation of SKA-Low 21-cm maps with Lyman- α forest data, which can be done using the reconstructed 2D/3D maps from Lyman- α forest transmissions. As demonstrated by Giri et al. (2025), these reconstructed Lyman- α transmission maps (which have high values in ionized regions) and the SKA-Low 21-cm maps (which have high values in neutral regions) are expected to be strongly anti-correlated. A detection of this anti-correlation would provide a "smoking gun" confirmation of the cosmological 21-cm signal, as the foregrounds and systematics for radio and optical/NB surveys are entirely independent.

6.3 Connecting individual ionized bubbles to the ionizing sources: a multiwavelength synergy with SKA

Apart from cross-correlation another approach is multi-frequency studies of particular areas. If the 21-cm image data from SKA-Low identifies certain ionized areas, follow-up observations (or archival observations) in optical and near-infrared, of the same patch of the sky, will reveal the galaxy population there. As it is these galaxies which produced the photons that made the ionized region, such a joint study can reveal how galaxies reionized the Universe, and for example give constraints on the escape fraction of ionizing photons.

Zackrisson et al. (2020); Mishra et al. (2025); Mellema et al. (2015); Lu et al. (2024); Giri et al. (2018a); Chen et al. (2026); Mason and Gronke (2020); Gazagnes et al. (2021); Ghara et al. (2018); Ghara and Choudhury (2020) and others suggest that SKA-Low will be able to detect ionized regions of size $\geq 5 - 10$ cMpc, corresponding to angular scales exceeding 5 arcminutes at redshifts $z = 7 - 10$. This resolution threshold enables the detection of ionized bubbles with volumes ≥ 1000 (cMpc)³. Several of the above studies independently established, via various reionization simulations, that there is a strong correlation between the volume of the ionized regions and the number of escaped ionizing photons (mostly UV) from the sources of ionization residing inside those bubbles.

Further Zackrisson et al. (2020) suggest that using JWST, Euclid, ELT, Nancy Grace Roman Space Telescope, etc., one can, in principle, detect the rest-frame UV emission from star-forming galaxies within ionized bubbles already detected via SKA-Low. However, the number of detectable bubble

galaxies, either via photometry or spectroscopy, will depend on the time-integrated escape fraction of ionizing photons and ionizing photon production efficiency. In the case of a galaxy-dominated reionization model, an escape fraction in the range of 5 – 20% will ensure even the smallest ionized region detectable by SKA-Low at $z \leq 10$ will contain tens of galaxies detectable above JWST and ELT spectroscopic thresholds. The JWST photometry will increase this number by more than a dozen; Roman and Euclid photometry will detect fewer galaxies than that. However, photometry comes with its own challenges, as the line-of-sight extent of the bubble ($\Delta z \approx 0.03 - 0.06$) is far smaller than the photometric redshift uncertainties ($\Delta z \sim 1$). Thus, photometrically detected galaxies within the bubbles suffer from confusion from interloper galaxies, whose number decreases with increasing redshift. Thus, a follow-up spectroscopic confirmation of the photometrically detected bubble galaxies via ELT-MOSAIC or JWST-NIRSpec would be essential for confirmation of their host bubble membership (Chen et al., 2026; Mason and Gronke, 2020; Zackrisson et al., 2020).

7 The connection with the global 21-cm signal

The global or sky-averaged 21-cm signal is another powerful observable. Although it does not carry any positional information, it captures the global evolution of the 21-cm signal and also measures its absolute strength. This section describes different ways in which SKA-Low measurements connect to the global signal. The first subsection describes a proposed method for measuring the global signal with the shortest baselines of SKA-Low and the second subsection outlines how high quality 21-cm image can be used to constrain the value of the global 21-cm signal.

7.1 Measuring the global 21-cm signal with short baselines

The global 21 cm signal is a key observable for the thermal and ionization history during the Cosmic Dawn/EoR, yet it is obscured by bright diffuse foregrounds. Besides total-power experiments, very short baselines inside an SKA-Low station provide an interferometric response to a uniform sky, so that one station can be treated as a global-signal interferometer.

For a two-element interferometer with baseline vector \mathbf{b} at frequency ν (wavelength $\lambda = c/\nu$), the visibility is

$$V(\mathbf{b}, \nu) = \frac{1}{4\pi} \int A(\hat{\mathbf{s}}, \nu) T(\hat{\mathbf{s}}, \nu) e^{-2\pi i(\mathbf{b} \cdot \hat{\mathbf{s}})/\lambda} d\Omega, \quad (5)$$

where A is the primary beam pattern and T is the sky temperature. Writing $T(\hat{\mathbf{s}}, \nu) = T_0(\nu) + \delta T(\hat{\mathbf{s}}, \nu)$ gives

$$V(\mathbf{b}, \nu) = T_0(\nu) G(\mathbf{b}, \nu) + V_{\text{ang}}(\mathbf{b}, \nu), \quad G \equiv \frac{1}{4\pi} \int A e^{-2\pi i(\mathbf{b} \cdot \hat{\mathbf{s}})/\lambda} d\Omega, \quad (6)$$

where G describes the interferometric response to a uniform sky. For short baselines with $\rho \equiv |\mathbf{b}|/\lambda \lesssim 0.5$, the amplitude $|G|$ decreases smoothly as the baseline grows, reflecting the reduced sensitivity to the uniform sky component. This means that the coupling to the monopole remains strong on the shortest spacings, but rapidly weakens as the baseline length increases.

Prototype work has established the feasibility and practical requirements for these measurements. Using the Engineering Development Array-2 (EDA-2), the All-Sky Signal Short-Spacing Interferometer (ASSASSIN) recovered a smooth Galactic spectrum from short spacings and emphasized

the need for wide, smooth bandpasses, embedded-element beam/coupling models, and conservative baseline masks (McKinley et al., 2020). The two-element Short-spacing Interferometer Telescope probing cosmic dawn and epoch of Reionization (SITARA) quantified correlated receiver noise and coupling on very short spacings and proposed calibration strategies (Thekkepattu et al., 2022; Thekkepattu et al., 2023). The Short-spacing Interferometer for Global-signal Measurement and Analysis (SIGMA) concept is a purpose-built 1D short-spacing array that maintains $b < \lambda/2$ over 65–90 MHz with explicit coupling treatment (Zhao et al., 2025). Design and sensitivity considerations for interferometric monopole recovery are discussed in (Presley et al., 2015; Singh et al., 2015).

This approach can be evaluated with a one-day single-station SKA Low simulation with $N = 256$ elements. For this simulation drift scan visibilities are generated with OSKAR³ over 50–100 MHz in 1 MHz channels and 240 s integrations. The antenna response is simulated in FEKO⁴ using the SKALA 4.1 design (Bolli et al., 2020) to obtain the element beam patterns. Only cross correlations between distinct antenna elements are analyzed, applying a strict short baseline mask of $\rho \leq 0.5$. Since the shortest physical spacing in the station is about 1.76 m, no baseline satisfies the $\lambda/2$ condition above approximately 85.4 MHz.

To extract the global sky component, a visibility-weighted estimator is used for each frequency channel using the short baseline data. Averaging these estimates over one sidereal day yields a smooth Galactic spectrum that follows the expected synchrotron power law between 50 and 80 MHz. To test signal recovery, a flattened Gaussian profile similar to the EDGES measurement ($A = 0.5$ K) is injected at the visibility level using the unit sky kernel for Stokes I . A two component fit at each (ν, t) , $V \simeq a(\nu, t) G + b(\nu, t) M + n$, where M is an angular foreground control derived from background only visibilities, followed by averaging of $a(\nu, t)$ over local sidereal time, successfully recovers the injected signal with an amplitude of $A_{\text{rec}} = -0.505 \pm 0.002$ K from the single day simulation.

In extending this approach to real observations, the dominant limitations are expected to arise from instrumental chromaticity of the primary beam and bandpass, mutual coupling and correlated noise on the shortest spacings, polarization leakage that mixes foregrounds into Stokes I , in-band reflections and slow gain variations, incomplete or imperfect sky models (including weak sidelobe sources), and environmental effects such as the ionosphere, radio interference, and near-field scattering. Future developments may explore strategies such as monitoring the spectral flatness of $G(\nu)$, performing visibility-level injection and recovery tests, applying conservative short-baseline selections, and combining data from multiple stations to reduce thermal noise and assess station-dependent systematics.

For band averaged amplitudes, the thermal uncertainty scales as

$$\sigma \propto (\Delta\nu \tau N_{\text{bl,eff}} N_{\text{day}})^{-1/2}. \quad (7)$$

In an idealized single-day simulation for SKA-Low, the thermal noise is already well below the level that would limit the analysis. Extending the integration to multiple days or combining measurements

³<https://ska-telescope.gitlab.io/sim/oskar>

⁴<https://altairhyperworks.com/feko>

from several stations is therefore most valuable for improving internal consistency and allowing stricter short baseline selections, rather than for further reducing statistical noise. At the later stages of the SKA Array Assembly, such as AA4 and AA*, thermal sensitivity will no longer be the dominant constraint. The focus will instead shift toward controlling instrumental chromaticity and beam stability so that the effective coupling kernel $K(\nu)$ remains well characterized. Cross checks between stations, observing days, and baseline masks will then provide the basis for a credible detection of the global 21 cm signal.

In summary, short baselines within an SKA-Low station enable direct interferometric estimation of the sky-averaged spectrum from visibilities, provided that the station kernel and bandpass remain smooth and stable. Our single-day simulation recovers both the smooth Galactic spectrum and the injected EDGES-like profile with high fidelity, in line with the experience from EDA2/ASSASSIN, SITARA, and SIGMA. These results point toward a clear, technically grounded path for achieving a credible detection of the global 21 cm signal with SKA-Low.

7.2 Measuring the global 21-cm signal from the images

[Giri et al. \(2018b\)](#) proposed a method to estimate the global signal by utilizing ionized regions identified in 21-cm signal images (see Section 3.1 for methods to identify these ionized regions). Due to the absence of a zero-spacing baseline in interferometric observations, the 21-cm signal at each frequency channel has a mean value of zero. However, pixels within the ionized regions contain values that correspond to the negative of the global signal value at that frequency, and by inverting these pixel values, the global signal can be recovered.

This estimation method has important caveats, as it is sensitive to both foreground residuals and noise levels in the images. [Giri et al. \(2018b\)](#) demonstrated that the method can be successfully applied to 21-cm signal images obtained with 1000-hour observations using the final SKA-Low configuration (AA4), provided that foreground residuals are subdominant and very large ionized regions are used for the measurement. Under these conditions, the technique can reliably recover the global signal from the 21-cm images.

8 Conclusions

SKA-Low will have the capability to produce tomographic image data of the redshifted 21-cm signal from the Cosmic Dawn and Epoch of Reionization. Producing high-fidelity image data will however be quite challenging and likely require substantial amounts of observing time (> 1000 hours) on selected fields and superb calibration and foreground removal techniques. As noise levels are lower at lower redshifts, the prospects are best for the reionization epoch, where the focus will most likely be on identifying ionized and neutral regions, effectively segmenting the image data into those two types of regions. Producing tomographic images data will be easier in AA4 than in AA*.

A wide range of analysis methods have been developed to quantitatively describe different aspects of such segmented image data, ranging from different algorithms extracting size distributions to many different types of topological quantities. Advances in machine learning are now also enabling

analysis methods which do not rely on identifying specific features in the data. A different use of image data is to combine it in different ways with data from other telescopes for the same field.

Although interferometers in principle cannot measure the sky-averaged signal, techniques are being developed that allow extracting the global signal by using very short, intra-station, baselines, so essentially from very low resolution images.

9 Author contributions

The authors of this chapter contributed to its text in the following way. SB, SD, AD, SM and SKP contributed to Sects. 2, 3.1, 4.1, 4.2. This includes the production of Fig. 3. SM also wrote the first version of Sect. 6.3. MB contributed to the machine learning part of Section 3.1 SC, together with KKD and AM mainly contributed to Sect. 3.2 on the matched filtering technique, including Figs. 4 and 5. IG wrote a large part of Sect. 4.1 on size distributions. SKG primarily contributed to Sects. 2, 6.2 and 7.2 including producing Figure 1 and 2 and also helped with the text in Sect. 3.1. IH wrote Section 5.1 on scattering transforms and their applications. AK contributed to writing of Sect. 4.2 on topological quantities. GM coordinated the writing of this chapter, wrote the abstract, Sects. 1, 8, the preambles of Sects. 3, 4, 5, 6 and 7, and was responsible for the final editing of the text. FZ contributed with the text of Sect. 7.1 on measuring the global signal with short baselines.

References

- A. Acharya et al. In *Advancing Astrophysics with the SKA – II (AASKAII)*. 2026a. arXiv search: Report number AASKAII/Acharya01.
- A. Acharya et al. In *Advancing Astrophysics with the SKA – II (AASKAII)*. 2026b. arXiv search: Report number AASKAII/Acharya02.
- E. Allys et al. *A&A*, 629:A115, Sept. 2019. doi: 10.1051/0004-6361/201834975.
- E. Allys et al. *Phys. Rev. D*, 102(10):103506, Nov. 2020. doi: 10.1103/PhysRevD.102.103506.
- C. Auclair et al. *A&A*, 681:A1, Jan. 2024. doi: 10.1051/0004-6361/202346814.
- S. Bag et al. *MNRAS*, 477(2):1984–1992, June 2018. doi: 10.1093/mnras/sty714.
- S. Bag et al. *MNRAS*, 485(2):2235–2251, May 2019. doi: 10.1093/mnras/stz532.
- B. Bandyopadhyay, T. R. Choudhury, and T. R. Seshadri. *MNRAS*, 466(2):2302–2311, Apr. 2017. doi: 10.1093/mnras/stw3347.
- E. Bañados et al. *Nature*, 553(7689):473–476, Dec. 2017. doi: 10.1038/nature25180.
- G. D. Becker et al. *MNRAS*, 447(4):3402–3419, Mar. 2015. doi: 10.1093/mnras/stu2646.
- M. Bianco, S. K. Giri, I. T. Iliev, and G. Mellema. *MNRAS*, 505(3):3982–3997, May 2021. doi: 10.1093/mnras/stab1518.
- M. Bianco et al. *MNRAS*, 528(3):5212–5230, Jan. 2024. doi: 10.1093/mnras/stae257.
- P. Bolli et al. *IEEE Open Journal of Antennas and Propagation*, 1:253–263, June 2020. doi: 10.1109/OJAP.2020.2999109.
- J. Bruna and S. Mallat. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8): 1872–1886, 2013. doi: 10.1109/TPAMI.2012.230.
- A. Chakraborty et al. In *Advancing Astrophysics with the SKA – II (AASKAII)*. 2026. arXiv search: Report number AASKAII/Chakraborty01.

- K.-F. Chen et al. *ApJ*, 1001(2):186, Apr. 2026. doi: 10.3847/1538-4357/ae57a8.
- Z. Chen, Y. Xu, Y. Wang, and X. Chen. *ApJ*, 885(1):23, Nov. 2019. doi: 10.3847/1538-4357/ab43e6.
- S. Cheng, Y.-S. Ting, B. Ménard, and J. Bruna. *MNRAS*, 499(4):5902–5914, Dec. 2020. doi: 10.1093/mnras/staa3165.
- S. Cheng et al. *PNAS Nexus*, 3(4):pgae103, 03 2024. doi: 10.1093/pnasnexus/pgae103.
- P. Chingangbam et al. *JCAP*, 2017(12):023, Dec. 2017. doi: 10.1088/1475-7516/2017/12/023.
- M. Choudhury et al. *JCAP*, 2025(6):003, June 2025. doi: 10.1088/1475-7516/2025/06/003.
- S. Dasgupta et al. *JCAP*, 2023(5):014, May 2023. doi: 10.1088/1475-7516/2023/05/014.
- K. K. Datta, S. Bharadwaj, and T. R. Choudhury. *MNRAS*, 382(2):809–818, Dec. 2007. doi: 10.1111/j.1365-2966.2007.12421.x.
- K. K. Datta, S. Majumdar, S. Bharadwaj, and T. R. Choudhury. *MNRAS*, 391(4):1900–1912, Dec. 2008. doi: 10.1111/j.1365-2966.2008.14008.x.
- K. K. Datta et al. *MNRAS*, 424(1):762–778, July 2012. doi: 10.1111/j.1365-2966.2012.21268.x.
- P. Dewdney et al. *SKA1 SYSTEM BASELINE DESIGN, SKA-TEL-SKO-DD-001, Square Kilometre Array Organisation*. 2013.
- K. Diao, Z. Chen, X. Chen, and Y. Mao. *ApJ*, 974(1):141, Oct. 2024. doi: 10.3847/1538-4357/ad6c40.
- A. Doussot and B. Semelin. *A&A*, 667:A118, Nov. 2022. doi: 10.1051/0004-6361/202244108.
- W. Elbers and R. van de Weygaert. *MNRAS*, 486(2):1523–1538, June 2019. doi: 10.1093/mnras/stz908.
- W. Elbers and R. van de Weygaert. *MNRAS*, 520(2):2709–2726, Apr. 2023. doi: 10.1093/mnras/stad120.
- X. Fan, E. Bañados, and R. A. Simcoe. *ARA&A*, 61:373–426, Aug. 2023. doi: 10.1146/annurev-astro-052920-102455.
- M. M. Friedrich et al. *MNRAS*, 413:1353–1372, May 2011. doi: 10.1111/j.1365-2966.2011.18219.x.
- S. R. Furlanetto and S. P. Oh. *MNRAS*, 457(2):1813–1827, Apr. 2016. doi: 10.1093/mnras/stw104.
- S. R. Furlanetto, M. Zaldarriaga, and L. Hernquist. *ApJ*, 613(1):1–15, Sept. 2004. doi: 10.1086/423025.
- S. R. Furlanetto, M. McQuinn, and L. Hernquist. *MNRAS*, 365(1):115–126, Jan. 2006. doi: 10.1111/j.1365-2966.2005.09687.x.
- S. Gagnon-Hartman, Y. Cui, A. Liu, and S. Ravanbakhsh. *MNRAS*, 504(4):4716–4729, July 2021. doi: 10.1093/mnras/stab1158.
- S. Gazagnes, L. V. E. Koopmans, and M. H. F. Wilkinson. *MNRAS*, 502(2):1816–1842, Apr. 2021. doi: 10.1093/mnras/stab107.
- R. Ghara and T. R. Choudhury. *MNRAS*, 496(1):739–753, July 2020. doi: 10.1093/mnras/staa1599.
- R. Ghara et al. In V. Jelić and T. van der Hulst, editors, *Peering towards Cosmic Dawn*, volume 333 of *IAU Symposium*, pages 122–125, May 2018. doi: 10.1017/S1743921318000728.
- N. Gillet et al. *MNRAS*, 484(1):282–293, Mar. 2019. doi: 10.1093/mnras/stz010.
- S. Giri, G. Mellema, and H. Jensen. *The Journal of Open Source Software*, 5(52):2363, Aug. 2020. doi: 10.21105/joss.02363.
- S. K. Giri and G. Mellema. *MNRAS*, 505(2):1863–1877, Aug. 2021. doi: 10.1093/mnras/stab1320.
- S. K. Giri, G. Mellema, K. L. Dixon, and I. T. Iliev. *MNRAS*, 473:2949–2964, Jan. 2018a. doi:

- 10.1093/mnras/stx2539.
- S. K. Giri, G. Mellema, and R. Ghara. *MNRAS*, 479(4):5596–5611, Oct. 2018b. doi: 10.1093/mnras/sty1786.
- S. K. Giri, K. Kakiichi, M. Bianco, and P. D. Meerburg. *MNRAS*, 544(4):3076–3093, Dec. 2025. doi: 10.1093/mnras/staf1947.
- L. Gleser, A. Nusser, B. Ciardi, and V. Desjacques. *MNRAS*, 370(3):1329–1338, Aug. 2006. doi: 10.1111/j.1365-2966.2006.10556.x.
- A. Gorce, A. Hutter, and J. R. Pritchard. *A&A*, 653:A58, Sept. 2021. doi: 10.1051/0004-6361/202140515.
- B. Greig, Y.-S. Ting, and A. A. Kaurov. *MNRAS*, 513(2):1719–1741, June 2022. doi: 10.1093/mnras/stac977.
- S. Hassan, A. Liu, S. Kohn, and P. La Plante. *MNRAS*, 483(2):2524–2537, Feb. 2019. doi: 10.1093/mnras/sty3282.
- S. E. Hong et al. *Journal of Korean Astronomical Society*, 47(2):49–67, Apr. 2014. doi: 10.5303/JKAS.2014.47.2.49.
- I. Hothi, E. Allys, B. Semelin, and F. Boulanger. *A&A*, 686:A212, June 2024. doi: 10.1051/0004-6361/202348444.
- I. Hothi, E. Allys, B. Semelin, and R. Meriot. *A&A*, 706:A15, Jan. 2026. doi: 10.1051/0004-6361/202555896.
- I. T. Iliev et al. *MNRAS*, 369:1625–1638, July 2006. doi: 10.1111/j.1365-2966.2006.10502.x.
- K. Kakiichi et al. *MNRAS*, 471:1936–1954, Oct. 2017. doi: 10.1093/mnras/stx1568.
- K. Kakiichi, T. Schmidt, and J. Hennawi. *MNRAS*, 516(1):582–601, Oct. 2022. doi: 10.1093/mnras/stac2026.
- M. Kamran, M. Sahlén, D. Sarkar, and S. Majumdar. *JCAP*, 2025(7):054, July 2025. doi: 10.1088/1475-7516/2025/07/054.
- A. Kapahtia, P. Chingangbam, S. Appleby, and C. Park. *JCAP*, 2018(10):011, Oct. 2018. doi: 10.1088/1475-7516/2018/10/011.
- A. Kapahtia, P. Chingangbam, and S. Appleby. *JCAP*, 2019(9):053, Sept. 2019. doi: 10.1088/1475-7516/2019/09/053.
- A. Kapahtia et al. *JCAP*, 2021(5):026, May 2021. doi: 10.1088/1475-7516/2021/05/026.
- J. Kennedy et al. *MNRAS*, 529(4):3684–3698, Apr. 2024. doi: 10.1093/mnras/stae760.
- T. M. Kodinariya, P. R. Makwana, et al. *International Journal*, 1(6):90–95, 2013.
- K.-G. Lee, R. Cen, J. R. Gott, III, and H. Trac. *ApJ*, 675(1):8–15, Mar. 2008. doi: 10.1086/525520.
- Y. Lin, S. P. Oh, S. R. Furlanetto, and P. M. Sutter. *MNRAS*, 461:3361–3374, Sept. 2016. doi: 10.1093/mnras/stw1542.
- T.-Y. Lu et al. *MNRAS*, 528(3):4872–4890, Mar. 2024. doi: 10.1093/mnras/stae266.
- S. Majumdar, S. Bharadwaj, K. K. Datta, and T. R. Choudhury. *MNRAS*, 413(2):1409–1418, May 2011. doi: 10.1111/j.1365-2966.2011.18223.x.
- S. Majumdar, S. Bharadwaj, and T. R. Choudhury. *MNRAS*, 426(4):3178–3194, Nov. 2012. doi: 10.1111/j.1365-2966.2012.21914.x.
- S. Mallat. *Communications on Pure and Applied Mathematics*, 65(10):1331–1398, 2012. doi: <https://doi.org/10.1002/cpa.21413>.
- C. A. Mason and M. Gronke. *MNRAS*, 499(1):1395–1405, Nov. 2020. doi: 10.1093/mnras/staa2910.

- Y. Matsuoka et al. *ApJ*, 883(2):183, Oct. 2019a. doi: 10.3847/1538-4357/ab3c60.
- Y. Matsuoka et al. *ApJL*, 872(1):L2, Feb. 2019b. doi: 10.3847/2041-8213/ab0216.
- A. Mazumder, A. Datta, A. Chakraborty, and S. Majumdar. *MNRAS*, 515(3):4020–4037, Sept. 2022. doi: 10.1093/mnras/stac1994.
- B. McKinley et al. *MNRAS*, 499(1):52–67, Nov. 2020. doi: 10.1093/mnras/staa2804.
- G. Mellema et al. In *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, page 10, Apr. 2015. doi: 10.22323/1.215.0010.
- A. Mesinger and S. Furlanetto. *ApJ*, 669:663–675, Nov. 2007. doi: 10.1086/521806.
- A. Mesinger, S. Furlanetto, and R. Cen. *MNRAS*, 411:955–972, Feb. 2011. doi: 10.1111/j.1365-2966.2010.17731.x.
- A. Mishra et al. *JCAP*, 02:055, 2025. doi: 10.1088/1475-7516/2025/02/055.
- D. J. Mortlock et al. *Nature*, 474(7353):616–619, June 2011. doi: 10.1038/nature10159.
- S. K. Pal et al. *JCAP*, 2025(10):096, Oct. 2025a. doi: 10.1088/1475-7516/2025/10/096.
- S. K. Pal, A. Datta, and A. Mazumder. *JCAP*, 2025(2):058, Feb. 2025b. doi: 10.1088/1475-7516/2025/02/058.
- A. Pathak et al. *JCAP*, 2022(11):027, Nov. 2022. doi: 10.1088/1475-7516/2022/11/027.
- P. Pranav et al. *MNRAS*, 485(3):4167–4208, May 2019. doi: 10.1093/mnras/stz541.
- D. Prelogović and A. Mesinger. *A&A*, 688:A199, Aug. 2024. doi: 10.1051/0004-6361/202449309.
- D. Prelogović et al. *MNRAS*, 509(3):3852–3867, Jan. 2022. doi: 10.1093/mnras/stab3215.
- M. E. Presley, A. Liu, and A. R. Parsons. *ApJ*, 809(1):18, Aug. 2015. doi: 10.1088/0004-637X/809/1/18.
- O. Ronneberger, P. Fischer, and T. Brox. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015. doi: 10.1007/978-3-319-24574-4_28.
- H. E. Ross et al. *MNRAS*, 506(3):3717–3733, Sept. 2021. doi: 10.1093/mnras/stab1822.
- V. Sahni, B. Sathyaprakash, and S. F. Shandarin. *ApJL*, 495:L5–L8, 1998. doi: 10.1086/311214.
- J. V. Sheth and V. Sahni. *arXiv e-prints*, art. astro-ph/0502105, Feb. 2005. doi: 10.48550/arXiv.astro-ph/0502105.
- J. V. Sheth, V. Sahni, S. F. Shandarin, and B. Sathyaprakash. *MNRAS*, 343:22, 2003. doi: 10.1046/j.1365-8711.2003.06642.x.
- H. Shimabukuro, Y. Mao, and J. Tan. *Research in Astronomy and Astrophysics*, 22(3):035027, Mar. 2022. doi: 10.1088/1674-4527/ac4ca3.
- H. Shimabukuro, Y. Xu, and Y. Shao. *Phys. Rev. D*, 112(6):063557, Sept. 2025. doi: 10.1103/nxr4-14gb.
- S. Singh, R. Subrahmanyam, N. Udaya Shankar, and A. Raghunathan. *ApJ*, 815(2):88, Dec. 2015. doi: 10.1088/0004-637X/815/2/88.
- J. N. Thekkepattu et al. *PASA*, 39:e018, Apr. 2022. doi: 10.1017/pasa.2022.13.
- J. N. Thekkepattu, C. M. Trott, and B. McKinley. *MNRAS*, 520(4):6040–6052, 2023.
- C. M. Trott et al. *MNRAS*, 507(1):772–780, Oct. 2021. doi: 10.1093/mnras/stab2235.
- F. Wang et al. *ApJL*, 869(1):L9, Dec. 2018. doi: 10.3847/2041-8213/aaf1d2.
- F. Wang et al. *ApJL*, 907(1):L1, Jan. 2021. doi: 10.3847/2041-8213/abd8c6.
- Y. Wang et al. *ApJ*, 814(1):6, Nov. 2015. doi: 10.1088/0004-637X/814/1/6.
- J. Witstok et al. *MNRAS*, 536(1):27–50, Jan. 2025. doi: 10.1093/mnras/stae2535.

- S. Wyithe, P. Geil, and H. Kim. In *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, page 15, Apr. 2015. doi: 10.22323/1.215.0015.
- S. Yoshiura, H. Shimabukuro, K. Takahashi, and T. Matsubara. *MNRAS*, 465(1):394–402, Feb. 2017. doi: 10.1093/mnras/stw2701.
- E. Zackrisson et al. *MNRAS*, 493(1):855–870, Jan. 2020. doi: 10.1093/mnras/staa098.
- O. Zahn et al. *ApJ*, 654:12–26, Jan. 2007. doi: 10.1086/509597.
- F. Zhao et al. *Research in Astronomy and Astrophysics*, 25(5):055011, May 2025. doi: 10.1088/1674-4527/adcb8d.
- X. Zhao, Y. Mao, C. Cheng, and B. D. Wandelt. *ApJ*, 926(2):151, Feb. 2022. doi: 10.3847/1538-4357/ac457d.
- X. Zhao, Y. Mao, S. Zuo, and B. D. Wandelt. *ApJ*, 973(1):41, Sept. 2024. doi: 10.3847/1538-4357/ad5ff0.