











Interferometric H_I Intensity Mapping of the Late Time Universe with SKA-Mid

Aishrila Mazumder* ¹, Zhaoting Chen[†] ², Junaid Townsend ³, Suman Chatterjee ³,
Zhixing Li,⁴ Sourabh Paul ¹, Reza Ansari ⁵, Laura Wolz ¹ and Mario G. Santos ^{3,6}

¹Jodrell Bank Centre for Astrophysics, Department of Physics and Astronomy, The University of Manchester, Manchester M13 9PL, UK

²Institute for Astronomy, The University of Edinburgh, Royal Observatory, Edinburgh EH9 3HJ, UK

³Department of Physics and Astronomy, University of the Western Cape, Robert Sobukwe Road, Bellville, Cape Town 7535, South Africa

⁴Department of Astronomy, Tsinghua University, Beijing 100084, China

⁵Université Paris-Saclay, Université Paris Cité, CEA, CNRS, AIM, 91191, Gif-sur-Yvette, France

⁶South African Radio Astronomy Observatory (SARAO), Liesbeek House, River Park, Gloucester Road, Mowbray, Cape Town, 7700, South Africa

E-mail: *aishrila.mazumder@manchester.ac.uk, [†]zhaoting.chen@roe.ac.uk

We discuss the progress towards using the SKA-Mid for interferometric neutral hydrogen (H_I) intensity mapping surveys. By mapping the distribution of cosmic H_I distribution through the 21 cm line, SKA-Mid will be able to measure the H_I power spectrum at small angular separations in interferometric mode. We review the measurements made from the precursor MeerKAT telescope, using the MeerKAT DEEP2 as well as the MIGHTEE survey data, yielding tentative detection as well as upper limits on H_I clustering. The methodology for MeerKAT can be naturally extended to SKA-Mid. Forecasts suggest that SKA-Mid AA4 will be able to measure the H_I power spectrum with high statistical significance across a wide range of redshifts from $z \sim 1.0$ to $z \sim 3.0$, around nonlinear scales $k \sim 1.0 \text{ Mpc}^{-1}$. The precise measurements can be used to constrain the properties of H_I galaxies, providing a novel window into probing galaxy evolution at $1.0 \lesssim z \lesssim 3.0$.

AM and ZC contribute equally to the chapter.

1 Introduction

In the post-reionisation Universe at $z \lesssim 5.5$ (Bosman et al., 2022), 21 cm emission line from neutral hydrogen (HI) can be used to map the three-dimensional distribution of the underlying dark matter. These maps can then be used to constrain the cosmological parameters, particularly the equation of state for dark energy, through galaxy clustering and the Baryon Acoustic Oscillation (BAO) using analysis techniques similar to those employed in optical surveys. HI galaxy surveys with SKA-Mid will provide a large sample of HI selected galaxies up to redshift $z < 0.4$, with precise measurements of redshifts (Nasirudin et al., 2026; Mayor et al., 2026). Meanwhile, single dish intensity mapping with SKA-Mid will significantly extend the mappable redshift range for HI to $z \sim 3$ (Cunnington & Wang et al., 2026).

However, the single dish (SD) HI intensity maps are limited to large angular scales. The angular resolution is limited to the size of the primary beam of the telescope, given by $\sim \lambda/D$, where D is the dish diameter and λ is the observed wavelength. For arrays such as the SKA-Mid, with $D = 13.5\text{m}$ and $D = 15\text{m}$ for MeerKAT and SKA-Mid dishes respectively, the primary beam gives an angular resolution of $\gtrsim 1$ deg at $z = 1$. Although this is sufficient to detect the BAO features in the power spectrum up to $z \sim 3$, there are valuable cosmological and astrophysical information at smaller angular scales which will be missed by SD surveys. The SKA-Mid telescope in the interferometric mode will provide information on complementary scales and extending the redshift reach of HI galaxy surveys.

In recent years, significant progress has been made in HI observations with interferometers in the post-reionization Universe. The Canadian Hydrogen Intensity Mapping Experiment (CHIME) has detected large-scale structure signal with HI stacking using eBOSS galaxies and Ly α systems (Amiri et al., 2023; CHIME Collaboration et al., 2023), and preliminary detection on the auto-power spectrum has been reported (CHIME Collaboration et al., 2025). The upgraded Giant Metrewave Radio Telescope (uGMRT) has placed an upper limit on the post-reionization HI power spectrum (Chakraborty et al., 2021; Elahi et al., 2026). Experiments such as the Hydrogen Intensity and Real-Time Analysis Experiment (Crichton et al., 2022) and the Canadian Hydrogen Observatory and Radio-transient Detector (Vanderlinde et al., 2019) will further explore measurements of HI clustering at BAO scales using interferometry. Interferometry intensity mapping with SKAO has also been studied, in the context of its precursor MeerKAT telescope. Using datasets of various survey areas and depth, measurements of the HI power spectrum have been made, including a tentative measurement (Paul et al., 2023) and robust upper limits at $z < 0.5$ (Mazumder et al., 2025), as well as measurement at $z < 0.1$ that matches the expected signal from HI selected galaxies (Townsend et al., 2026). The precursor measurements are discussed in detail in Section 2.

The progress in data analysis reveals challenges in mitigating systematics, both in observation and in signal modelling. In particular, for SKA-Mid, there are significant challenges in dealing with low-level radio frequency interference (RFI) contamination, foreground leakage in high delay modes near the horizon, calibration errors due to field sources around calibrators, and modelling the signal and its covariance at nonlinear scales. To address these problems, robust simulation tools are needed to understand the observational effects and theoretical framework needs to be established to understand the modelling.

The SKA-Mid will be transformative for H_I intensity mapping, providing unprecedented sensitivity and dense coverage of relatively large angular scales. Interferometric H_I intensity mapping benefits from the instrumental power of the SKAO as well as being commensal to continuum surveys. The existing MeerKAT measurements, for example, utilise datasets that are meant for the MeerKAT DEEP2 observation (Mauch et al., 2020) and the MeerKAT International GHz Tiered Extragalactic Exploration (MIGHTEE) Survey (Jarvis et al., 2016). It can be foreseen that, without relying on dedicated observation time, SKA-Mid will be able to measure the small-scale H_I power spectrum with high statistical significance at a wide range of redshifts $1 \lesssim z \lesssim 3$, as we present the forecast in Section 3. The precise measurements of H_I clustering can be used to probe H_I science, shedding light on galaxy evolution, as we discuss in Section 4. Throughout this chapter, we assume a fixed cosmology corresponding to the model reported in Planck Collaboration et al. (2020).

2 Results from the SKAO precursor MeerKAT telescope

We describe the results of H_I intensity mapping power spectrum measurements using MeerKAT surveys in this section. The MeerKAT array, with its dense core of short baselines and large collecting area, provides excellent sensitivity to large angular scales (low k_{\perp}), which are crucial for 21-cm cosmology. The results so far make measurements at k values $\gtrsim 0.3 \text{ Mpc}^{-1}$, using observations in the L-band, corresponding to H_I emission in the redshift range of $0 \lesssim z \lesssim 0.5$.

The studies described here using MeerKAT observations follow a similar power spectrum reconstruction approach. No imaging is performed in this approach, rather visibilities are used directly. They are gridded into uv grids, with Δu & Δv values that to make the primary beam width in k_{\perp} space negligible, thereby minimising the effect of mode mixing. “Delay-space” visibilities are obtained by Fourier transforming the gridded visibilities along the frequency axis. The Fourier space visibilities are used to measure the 3D power spectrum. Since the visibilities contain an additive noise bias, it removed either by calculating and removing the expected noise bias or by dividing visibilities into sets of equal but independent time samples, called “odd” and “even” samples, and correlating them removes the noise bias that is uncorrelated between the samples. These methods give power spectrum measurements via visibility auto-correlation and cross-correlation respectively ¹. The 3D reconstructed power spectrum in Fourier space is given by (Morales, 2005; Parsons et al., 2012; Paul et al., 2016):

$$P_D(\mathbf{k}_{\perp}, k_{\parallel}) = \frac{A_e}{\lambda^2 B} \frac{D^2 \Delta D}{B} \left(\frac{\lambda^2}{2k_B} \right)^2 \Re\{V_1(\mathbf{u}, \tau) V_2^*(\mathbf{u}, \tau)\}, \quad (1)$$

where A_e is the effective area of the dish, λ is the central wavelength of the corresponding chosen frequency range, B is the total bandwidth, D is the comoving distance calculated at the central redshift, z , ΔD is the comoving width corresponding with the effective bandwidth in frequency, k_B is the Boltzmann constant, and $V_1(\mathbf{u}, \tau) = V_2(\mathbf{u}, \tau)$ for auto-correlation. All interferometric

¹It is to be noted that this is different from auto- and cross-correlations mentioned in context of single dish observations, where they mean self-correlations of intensity maps or between intensity maps and optical galaxy number counts respectively. Instead, here the terms denote either squaring the same visibilities, i.e. $|V_i|^2$ or internal cross-correlation between different sets, $|V_i| * |V_j|$

intensity mapping measurements made with MeerKAT thus far have used the foreground avoidance technique, using only modes above the horizon limit is given by (see, for example [Thyagarajan et al., 2015](#)):

$$k_{\parallel} = \frac{D(z)E(z)H_0}{c(1+z)} \sin(\theta) k_{\perp}, \quad (2)$$

where $E(z) = (H(z)/H_0)^2$, H_0 is the Hubble constant, and θ is the field-of-view (FoV) of the observation. The 1D power spectrum is estimated using an inverse noise variance weighting, taken in logarithmic k bins, Δk_i , which can be expressed as:

$$\hat{P}_D^i(k) = \sum_j P_D^j w_j / \sum_j w_j, \quad (3)$$

where $w_j = 1/\sigma_{j,\text{TN}}$ is the inverse thermal noise variance weighting. The error on the power spectrum is then estimated using:

$$(\Delta P_D^i)^2 = \sum_j (P_D^j - \hat{P}_D^i)^2 w_j^2 / (\sum_j w_j)^2. \quad (4)$$

To obtain the weights, w_j , multiple thermal noise realisations are simulated, occupying the visibility grid using random values generated from a complex circular Gaussian distribution with a mean of zero, $\mu_{\text{TN}} = 0$, and a standard deviation derived from the radiometer equation ([Morales, 2005](#)):

$$\sigma_{\text{TN}} = \frac{2k_{\text{B}}T_{\text{sys}}}{A_{\text{eff}}\sqrt{\delta f \delta t}}, \quad (5)$$

where k_{B} is the Boltzmann constant, T_{sys} is the system temperature, A_{eff} is the effective collecting area for one dish, δf is the frequency resolution and δt is the time resolution.

The standard deviation of the power spectra of these realisations corresponds to $\sigma_{P_{\text{TN}}}$ used in $w = 1/\sigma_{P_{\text{TN}}}^2$. We note that, the thermal noise simulation is only used to obtain the weights, and the error bar is estimated using the variance of the measured 3D power as described in [Equation 4](#).

Using the aforementioned formalism, the visibility data taken from the tracking observations of the MeerKAT telescope can be used to measure the H_I power spectrum, as we review for the rest of this section.

2.1 Detection of the H_I power spectrum from the DEEP2 field

[Paul et al. \(2023\)](#) presents the first statistically significant interferometric detection of the H_I intensity mapping power spectrum at low redshifts using data from the MeerKAT radio telescope in the L-band. This measurement was performed in the DEEP2 field, selected for its lack of bright continuum sources and long integration time, making it ideal for deep intensity mapping studies. The analysis used approximately 96 hours of MeerKAT observations collected across nine nights during the telescope's commissioning phase in 2018 ([Mauch et al., 2020](#)). This data has been used

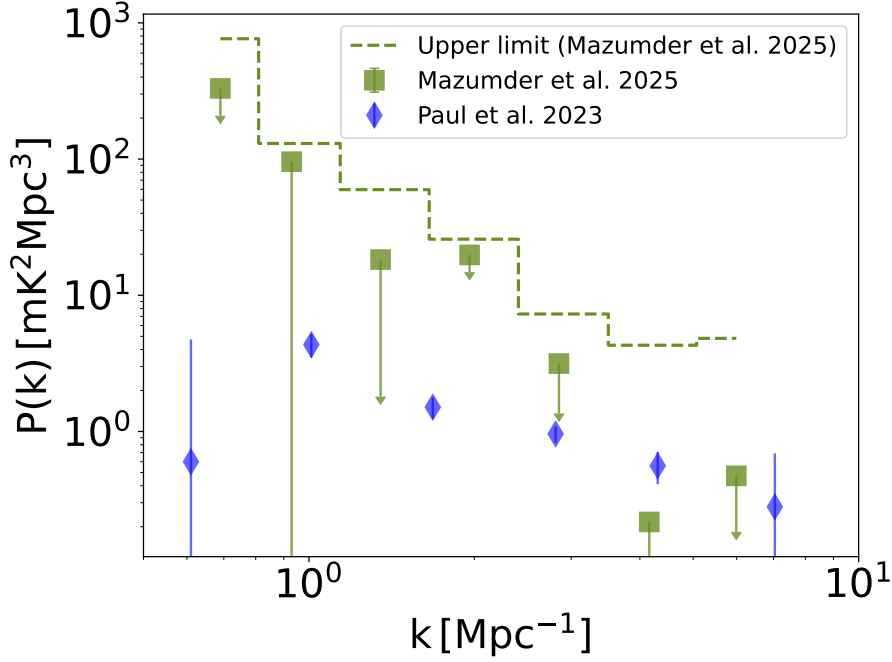


Figure 1: HI power spectrum estimates from MeerKAT observations at $z \sim 0.44$. The purple squares represent the upper limits from HI power spectrum obtained using cross correlation, and purple dashed lines represent the absolute upper limits using MIGHTEE data (Mazumder et al., 2025). Blue diamonds show the tentative detection using the MeerKAT DEEP2 data reported in Paul et al. (2023).

for investigating a wide variety of science cases, for example, sub-mJy source counts (Matthews et al., 2021a), measuring the local luminosity function of radio sources (Matthews et al., 2021b), indicating the data quality and reliability.

The data span two RFI-free frequency windows centred at 1077.5 MHz ($z \approx 0.32$) and 986 MHz ($z \approx 0.44$), each with a bandwidth of ~ 46 MHz. Our calibration and power spectrum pipeline was specifically developed to enable robust delay-spectrum estimation directly from visibilities, avoiding the need for image cube construction. Even after rigorous RFI flagging, there was evidence of low-level RFI present in the data. Hence, using a combination of automated CASA-based routines and novel per-baseline delay filtering, we applied two complementary mitigation techniques:

- Baseline-Level Flagging (BLF): identifies and removes antenna pairs exhibiting anomalous delay-space power outside the foreground wedge;
- uv-delay Flagging (UVDF): excludes individual contaminated modes in the gridded delay-space power spectrum using a 5σ threshold over thermal noise.

The BLF method identifies and flags contaminated antenna pairs early, on a per-scan basis, before any power-spectrum calculation. By contrast, UVDF applies a threshold after gridding the visibilities, comparing the auto-power spectrum to the expected thermal noise. As a result, BLF, being closer to traditional radio-data flagging, tends to be more aggressive, while UVDF, applied nearer to the final power-spectrum stage, typically flags less. Both approaches yield consistent results.

We employed a conservative foreground avoidance strategy, applying a cut at $k_{\parallel} > 0.3k_{\perp}$ to ensure robustness against residual contamination. Following these mitigations, we obtained clear detections of the 3-d H_I auto-power spectrum:

- $\sim 5.9\sigma$ significance at $z \approx 0.32$, and
- $\sim 9.2\sigma$ significance at $z \approx 0.44$.

Figure 1 shows the 1D H_I power spectrum at $z \approx 0.44$ in blue diamonds, overplotted with MIGHTEE results (described in Section 2.2 below). These measurements represent the first detection of interferometric auto-power spectra at these redshifts, achieved without relying on cross-correlation with galaxy surveys. The amplitude and scale dependence of the measured power spectra are consistent with predictions from halo-based models, and the signal is robust under multiple calibration and systematic error tests.

2.2 Upper limits of the H_I power spectrum with multiple fields from the MIGHTEE survey

Mazumder et al. (2025) presents a follow-up to the DEEP2 power spectrum measurement using the MeerKAT International GigaHertz Tiered Extragalactic Exploration, i.e. the MIGHTEE survey (Jarvis et al., 2016). It covers ~ 20 square degrees with a total of 1000 hours of observation time distributed over four well-known extragalactic deep fields. A subset of the data covering ~ 100 hours in the COSMOS field was used for this study (Heywood et al., 2024). The observations span ~ 4 square degrees, consisting of 15 pointings with 94.2 hours on-source. Owing to the different phase centres of each pointing, they sample slightly different parts of the sky. Hence, the gridded, delay-transformed visibilities were incoherently averaged in the power spectrum domain for the final measurements.

The frequency range chosen in this work covers the same RFI free window centred at ~ 986 MHz (or H_I emission at $z \approx 0.44$) as that in Paul et al. (2023). Since incoherent averaging is the so-called “square-then-average” method, contrary to coherent averaging or the “average-then-square” method, the noise scales down as $1/\sqrt{N}$, since N independent measurements (pointings) are averaged after squaring, as opposed to averaging N independent measurements (visibilities) and then squaring, which reduces the noise by $1/N$. Hence, we get a root-mean-squared thermal noise fluctuation ~ 10 mK using an observation spread over 15 pointings (for a single deep observation of the same time, it would be ~ 5 mK).

The calibrated visibilities for each observing night were gridded in the uv -plane with a grid size 60λ . The gridded cubes were Fourier transformed along the frequency axis to produce a (u, v, τ) cube, where τ is the Fourier conjugate of frequency, or the delay. The 3D delay power per pointing P_D given by Equation 1 is weighted by the respective baseline density for each pointing, to account for the differences in baseline coverage among the pointings. This gives the final averaged 3D power:

$$P_D(\mathbf{k}_\perp, k_\parallel) = \frac{\sum_{i=1}^{15} w_i(\mathbf{k}_\perp, k_\parallel) P_D^i(\mathbf{k}_\perp, k_\parallel)}{\sum_{i=1}^{15} w_i(\mathbf{k}_\perp, k_\parallel)}, \quad (6)$$

where the i is over all 15 pointings and $P_D^i(\mathbf{k}_\perp, k_\parallel)$ is the delay power for the i th pointing, weighted by the factor w_i which is the baseline density per 3D k pixel. Finally, the incoherently averaged 3D power spectrum is binned into spherical and cylindrical bins. The data usability and feasibility of using incoherent averaging were checked through jackknife tests, checks for power distribution per 1D k -bin and varying the foreground avoidance window. These tests showed that while residual systematics may be present in some pointings, on combination and cross-correlation, most systematic contributions reduce considerably.

Employing the same foreground avoidance window described in Section 2.1, we set the first upper limits on the H_I power spectrum. It should be noted that, unlike Paul et al. (2023), we do not use any baseline level or uv-delay space flagging in this work. Using both auto-correlation and cross-correlation visibilities (see Figure 1), we find consistent upper limits of 29.8 mK²Mpc³ from the 2 σ cross-correlation measurements and 25.82 mK²Mpc³ from auto-correlation at $k \sim 2$ Mpc⁻¹. The data used here constitutes a small fraction of the MIGHTEE survey and demonstrates that combined analysis of the full MIGHTEE survey can potentially detect the H_I power spectrum at $z \lesssim 0.5$, in the range $0.1 \text{ Mpc}^{-1} \lesssim k \lesssim 10 \text{ Mpc}^{-1}$ or quasi-linear scales.

Figure 1 shows the H_I clustering power spectrum upper limits from auto- (dashed purple lines) and the cross- (purple squares) correlation of MIGHTEE-COSMOS observations. The MIGHTEE COSMOS data presented in this subsection has a similar amount of observational time with the DEEP2 field. However, the observation consists of 15 pointings whereas the DEEP2 data is one single pointing, and the final measurement is an incoherent averaging of the power spectrum measurements from each pointing. Therefore, the signal-to-noise ratio of the MIGHTEE measurements is lower, with only upper limits reported. Nevertheless, these consistent results demonstrate the viability of MeerKAT as a precision interferometric instrument for cosmological H_I intensity mapping across multiple pointings over a large survey area, providing a strong foundation for future observations with the SKA-Mid array.

2.3 A comparative analysis of H_I clustering between intensity maps and direct detections

MeerKAT observations have also been used to test the robustness of using the interferometric H_I intensity mapping in the low-redshift Universe ($z < 0.1$). In this redshift regime, large-scale H_I galaxy surveys are feasible, since the sensitivity requirements are less strenuous on radio telescopes. As one goes to higher redshifts, this becomes harder to do, making intensity mapping ideal for probing the H_I content in the post-EoR Universe beyond $z = 0.1$.

With the MIGHTEE Early-Science COSMOS field data (Jarvis et al., 2016), there are known H_I galaxy detections. Townsend et al. (2026) explores an interesting way to test the performance of intensity mapping, using known H_I sources. Here, both interferometric observations and H_I selected galaxies in the same field of view are used to estimate and compare H_I power spectra from each of

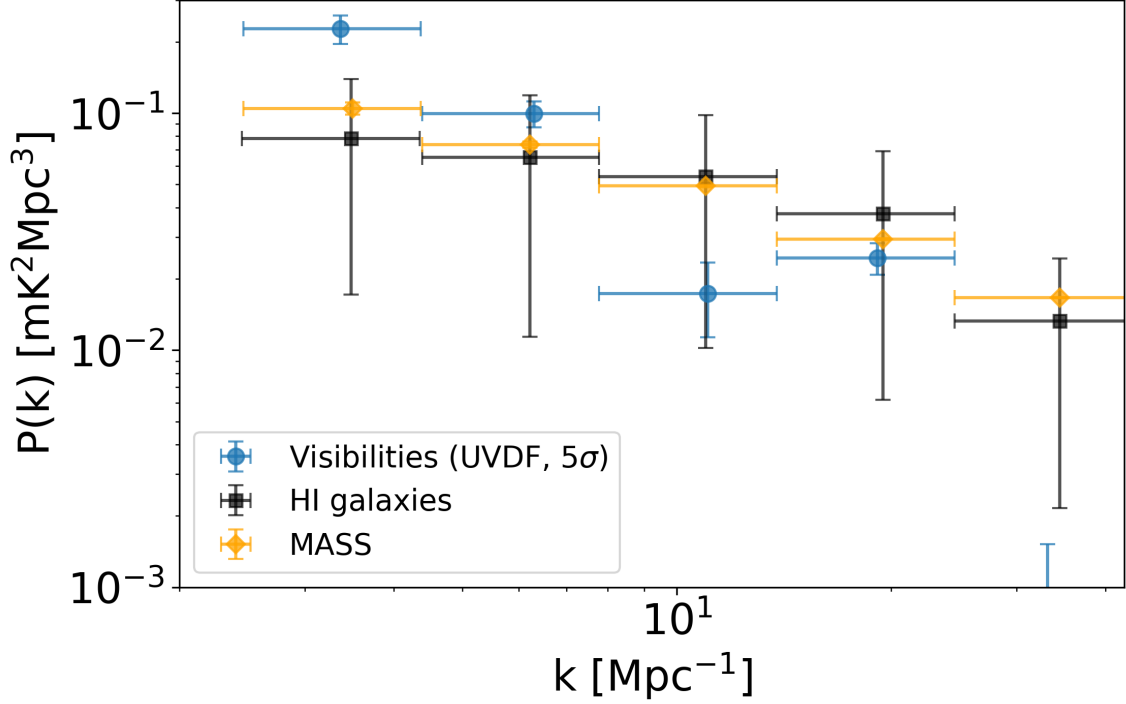


Figure 2: H I power spectrum estimates from MeerKAT observations at $z \lesssim 0.1$ using the MIGHTEE Early-Science COSMOS data (Townsend et al., 2026). The blue circles denote the measurement from visibility data, with a 5σ UVDF cut. The black squares denote the measurement from the H I galaxy catalogue. Orange diamonds denote measurements from simulated visibilities (MASS) using the H I galaxy catalogue.

these data sets. Such tests can be used for (a) testing how much neutral gas each of the methods (intensity mapping and galaxy detections) misses, and (b) testing the robustness of the methodology developed for processing visibility data for H I power spectrum estimation (used in Paul et al. 2023; Mazumder et al. 2025). Further, such analyses allow us to probe the H I clustering present in both the galaxy samples and the interferometric data sets on ultra-small scales ($3 \lesssim k \lesssim 40 \text{ Mpc}^{-1}$).

In this analysis, the data used consists of 3 nights on the same central region of the COSMOS field, combining observations from both Early-Science data (Heywood et al., 2022) and MIGHTEE spectral line data (Heywood et al., 2024), with a total observation time of ~ 17.45 hours and a redshift range of $0.02 < z < 0.066$. Conversely, (Mazumder et al., 2025) use all 15 pointings on the COSMOS field described in (Heywood et al., 2024). The calibrated visibility data were used to perform the cross-correlation, as discussed above. The UVDF approach described in Section 2.1 is adopted, flagging delay modes with a 5σ threshold. The results are shown in Figure 2 as the blue circle points for the 5σ cut.

Since this analysis aims to use the H I galaxy detections of the same data to compare and validate the power spectrum measurements, we also require the H I selected galaxy catalogue. This catalogue is generated following Maddox et al. (2021), with a measurement of the H I mass and the velocity width of the H I profile for each galaxy. We then use two methods to estimate the H I power spectrum from the galaxy catalogue. First, we grid the measured H I emission into a rectangular grid, and

perform power spectrum estimation directly on this grid. The results are shown in Figure 2 as the black square points (“HI galaxies”). Second, we use the source information to simulate HI-only visibilities using the MeerKAT All-Sky Simulator (MASS; see Section 5.1 of Townsend et al. 2026). The HI-only visibilities are then used to perform the power spectrum estimation and compare to the results from the data, as shown in the orange diamond points (“MASS”) in Figure 2.

As one can see, for the presented range of k scales, the measurements of the power spectrum amplitude from the intensity maps agree with the overlapping HI-selected galaxy power spectrum. This analysis indicates the robustness of the approach adopted in the MeerKAT data analysis for interferometric intensity mapping.

In conclusion, interferometric intensity mapping with MeerKAT has successfully demonstrated that both deep single-pointing and multi-pointing wide-area surveys are capable of measuring the clustering signal from HI. The precursor MeerKAT data analysis indicates that interferometric intensity mapping with SKA-Mid holds great potential for measuring the HI power spectrum at small scales across a wide range of redshifts. It is worth noting that these measurements are made using data designed for continuum and emission line galaxy science. In principle, any tracking observations with SKA-Mid can be used to perform power spectrum analysis, yielding great synergistic potential with other science goals of the SKAO.

2.4 MeerKAT All-Sky Simulator (MASS)

Besides observations, there have also been developments in simulating and modelling systematics in interferometric data. While several such examples are present for EoR, one example in the post-reionization regime is the MeerKAT All-Sky Simulator (MASS).

MASS is designed to simulate interferometric visibilities for a given astrophysical sky model. This package is developed for large-scale HI intensity mapping with interferometric observations using SKA-Mid precursor MeerKAT. MASS can be used for simulation based forward modelling approaches where realistic mock data is generated to study the impact of systematics and other complex processes applied to the observed data. In its current format, the MASS sky model can be a point-like source catalogues (Dirac delta functions) or a HEALPix (Hierarchical Equal Area isoLatitude Pixelization of a sphere; Gorski et al., 2005) map. MASS can also incorporate analytical models as input and construct source catalogues or HEALPix maps. With a given antenna configuration or baseline distribution, the package generates measurement sets that can be directly used in the data processing pipeline.

The visibilities are simulated using:

$$\mathcal{V}(\theta, U, \nu) = Q_\nu \sum_{q=0}^{N-1} \Delta\Omega_{\hat{n}_q} T(\hat{n}_q, \nu) A(\Delta n_q, \nu) e^{2\pi i U \cdot \Delta n_q}, \quad (7)$$

where the antenna primary beam (PB) pattern $A(\Delta n_q, \nu)$ quantifies how the individual antenna responds to the signal from different directions \hat{n} on the sky, \hat{m} refers to the direction in which the antennas are pointing with $\Delta\hat{n}_q = \hat{n}_q - \hat{m}$, $Q_\nu = 2k_B/\lambda^2$ is the conversion factor from brightness

temperature to specific intensity in the Rayleigh - Jeans limit, $T(\hat{n}_q, \nu)$ is the brightness temperature distribution on the sky and $\Delta\Omega_{\hat{n}_q}$ is the elemental solid angle in the direction \hat{n}_q . The summation can be over HEALPix pixel or the source catalogue. The primary advantage of this simplistic approach is that one can include different components at scales that they dominate. For instance, it is seen that bright point sources far away from the central lobe of the telescope PB will impact the H_I signal that is in the primary lobe of the beam. Current simulation capabilities of MASS include diffused galactic synchrotron emission, free-free emission, extra-galactic point sources, the Sun, Radio frequency interference (RFI, narrow and broad band), noise, gain error, and primary beam with side-lobes that extend to the horizon. However, the current version of MASS is not suitable for simulating small scale diffused emission.

3 Forecasts of power spectrum measurements for SKA-Mid

This section describes the forecasts from SKA-Mid for interferometric intensity mapping. H_I intensity mapping experiments require good coverage in spatial frequencies, i.e. the uv plane to ensure sufficient sensitivity to the faint cosmological signal, and to limit the impact of mode mixing. Hence, most of the dedicated interferometers for intensity mapping are configured as packed arrays. Although SKA-Mid is a general-purpose interferometer designed to reach high angular resolution, the configuration of the central part of the array, with baseline lengths $\lesssim 1$ km, is suitable for intensity mapping and provides good coverage of the uv plane. As discussed in Section 2, intensity mapping can be performed using continuum survey data of SKA-Mid, with multiple pointings of relatively deep integration time. Here, a demonstration is given of how these features translate to signal-to-noise expectations and the detectability of the H_I clustering power spectrum at redshifts beyond 0.5.

3.1 Signal-to-Noise Estimates

Using the AA4 array configuration specified in Braun et al. (2019)², Figure 3 shows the array layout for the 197 SKA-Mid dishes, with a zoom-in on the dense core within 2 km of the array centre. It shows that 118 out of 197 dishes are within the 1 km of the array centre, corresponding to $\sim 36\%$ of the baselines. Thus, it will be possible to map out the H_I distribution with ~ 1 arcmin resolution at redshift $z = 1$ with high sensitivity using the SKA-Mid array.

To calculate the expected uv coverage and the corresponding signal-to-noise level for the survey, a 4hr tracking observation pointing at the DEEP2 field is simulated. It is assumed that the visibility data will have a time resolution of 10 s. The resulting uv coverage is shown in Figure 4.

The radiometer equation, as described in Equation 5, is used to simulate thermal noise realisations, and forecasts are presented for three redshift bins at $0.5 < z < 1.5$ ($z \sim 1$), $1.5 < z < 2.5$ ($z \sim 2$) and $2.5 < z < 3.0$ ($z \sim 3$), with a frequency resolution of 0.38 MHz for each bin³. The total observation datasets are assumed to consist of 100 trackings of different pointings, each with 10 hr of observation time.

²<https://gitlab.com/ska-telescope/ost/ska-ost-array-config>

³Note that the frequency resolution of the visibility data can be higher, and the choice of frequency resolution in our forecast corresponds to the visibility data being averaged down to a lower resolution.

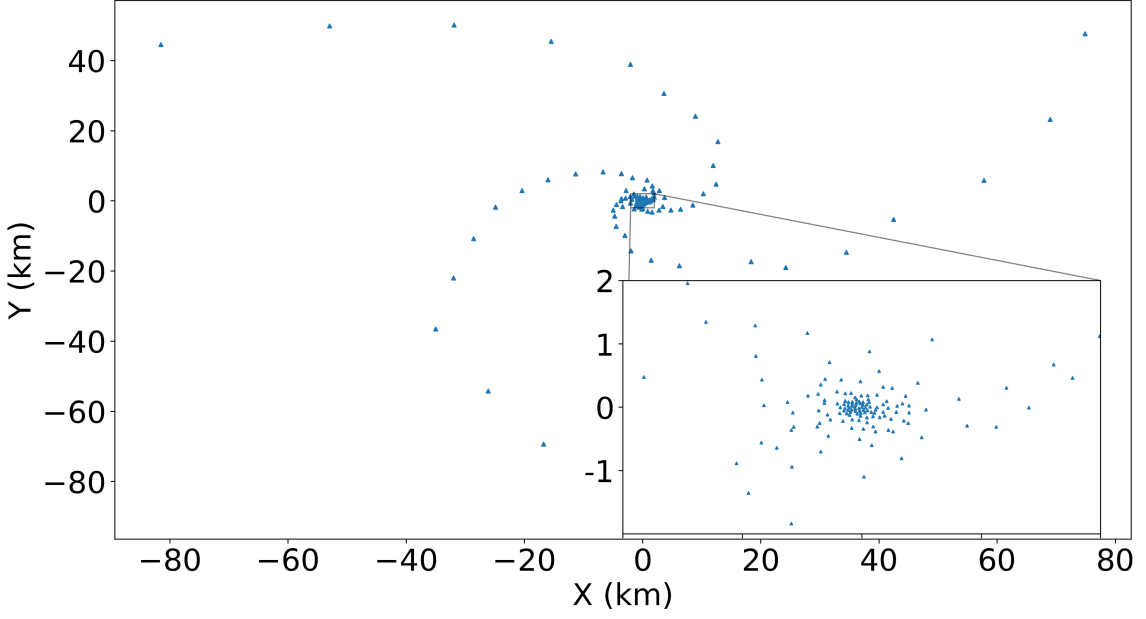


Figure 3: Left: (u, v) plane values of all dish pairs for the SKA-Mid array with 197 dishes at 700MHz. Right: (u, v) values for baselines with $\sqrt{u^2 + v^2} < 1500$

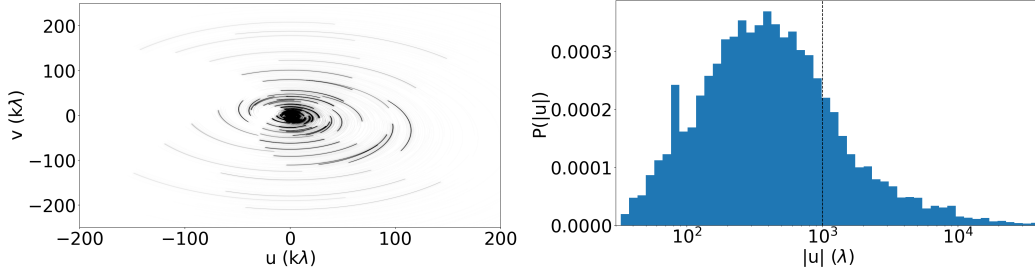


Figure 4: Left: The redshift distribution of the baselines for SKA-Mid AA4 for a tracking observation. Right: The probability distribution of the uv distance $|u| = \sqrt{u^2 + v^2}$ for SKA-Mid AA4. The dashed line denotes $|u| = 1000 \lambda$, which is a crude threshold for baselines that probe scales useful for clustering.

Then the visibilities are gridded to a uv grid following Paul et al. (2023). A grid size of $\Delta u = \Delta v = 20 \lambda$ is used here. The thermal noise generated per grid follows

$$\sigma_{\text{TN}}^{\text{grid}} = \sigma_{\text{TN}} \sqrt{\frac{t_{\text{track}}}{t_{\text{field}} N_{\text{baseline}}^{\text{grid}}}}, \quad (8)$$

where σ_{TN} is the noise level for a single baseline following Equation 5, $t_{\text{track}} = 4$ hr is the observation time of the simulated tracking, $t_{\text{field}} = 10$ hr is the total observation time of one field.

Then the thermal noise realisations are generated and the brightness temperature power spectrum is calculated based on Equation 1 for 128 realisations. The 3D power spectrum is averaged into the 1D power spectrum. The 1D $|k|$ -bins are logarithmically spaced between 0.075 Mpc^{-1} to 10 Mpc^{-1} for $z \sim 1$, 0.05 Mpc^{-1} to 8 Mpc^{-1} for $z \sim 2$ and 0.04 Mpc^{-1} to 6 Mpc^{-1} for $z \sim 3$. Finally, following equation Equation 2, a horizon limit corresponding to $\theta = 5$ deg is chosen. This is a conservative

choice as the FoV of the SKA-Mid dishes will be ~ 1 deg, which effectively considers potential foreground leakage at high delay. The 3D k -modes above the limit is then used for averaging into 1D. The variance of the 1D power spectrum among the realisations is used for the thermal noise variance so that

$$\sigma_{P_{\text{TN}}}^2 = \text{var}[P_{\text{TN}}^i] / N_{\text{field}}, \quad (9)$$

where $\text{var}[P_{\text{TN}}^i]$ is the variance among the realisations for a single field, and the final thermal noise variance is the result of incoherent averaging of all the fields.

Additionally, the fiducial signal $P_{\text{H I}}$ as well as its variance are both calculated. The chosen fiducial model corresponds to the halo model described later in Section 4.1. The signal variance is then $P_{\text{H I}} / \sqrt{N_k}$, where N_k is the number of k -modes in each 1D k -bin. The correlations between different k -bins are ignored for simplicity.

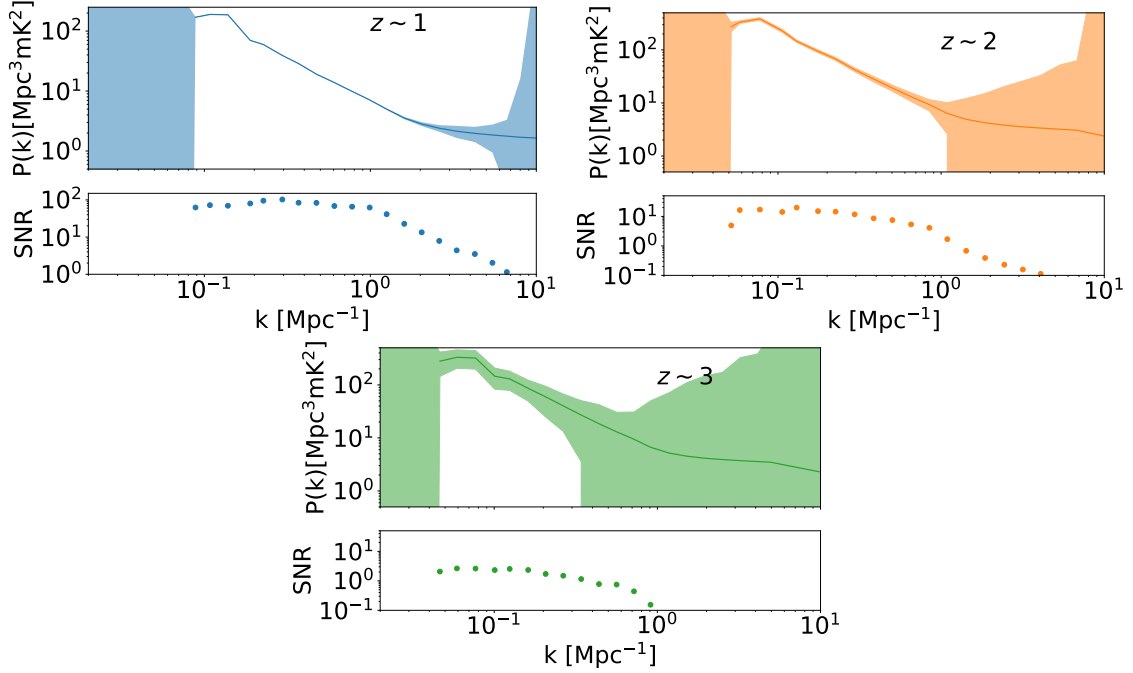


Figure 5: Forecasts for the interferometric H I intensity mapping measurements for SKA-Mid AA4 configuration. The top half of each panel shows the fiducial H I power spectrum and the expected 1σ measurement error. The bottom half of each panel shows the expected signal-to-noise ratio of each 1D k -bin. The redshift bins shown are $0.5 < z < 1.5$ ($z \sim 1$), $1.5 < z < 2.5$ ($z \sim 2$) and $2.5 < z < 3.0$ ($z \sim 3$).

The results are shown in Figure 5. For SKA-Mid AA4, the H I clustering with interferometric observations across a wide range of scales can be probed. High accuracy measurements, with a signal-to-noise ratio larger than 10, can be achieved around $0.1 \lesssim k \lesssim 1 \text{ Mpc}^{-1}$ for $z \sim 1$ and $z \sim 2$. For $z \sim 1$, scales up to $k \sim 10 \text{ Mpc}^{-1}$ can be probed. For $z \sim 3$, detection of the H I power spectrum can be made at $k \sim 0.1 \text{ Mpc}^{-1}$.

4 Anticipated science outputs from interferometric intensity mapping using SKA-Mid

In Section 3, it is demonstrated that SKA-Mid can be used to probe the H_I clustering with high accuracy. The measurements can then be used to probe H_I science, and the potential science outputs are discussed in this section.

4.1 Halo Model

The measured H_I power spectrum at quasi-linear and non-linear scales can be used to constrain the halo model of cosmic neutral hydrogen (Chen et al., 2021). In halo model formalism, the power spectrum is modelled as a combination of a 2-halo, a 1-halo term and a shot noise term (Cooray and Sheth, 2002). Following Chen et al. (2021), the H_I power spectrum can be modelled as:

$$P_{\text{H}_I}(k_{\perp}, k_{\parallel}) = P_{2\text{h}}(k_{\perp}, k_{\parallel}) + P_{1\text{h}}(k_{\perp}, k_{\parallel}) + P_{\text{SN}}(k_{\parallel}). \quad (10)$$

The 2-halo term is

$$P_{2\text{h}}(k, k_{\parallel}) = C_{\text{H}_I}^2 \left[\int dm n(m) b(m) \langle M_{\text{H}_I}(m) \rangle u_{\text{H}_I}(k|m) \right]^2 \left(1 + \frac{f k_{\parallel}^2}{b_{\text{H}_I}^0 k^2} \right)^2 \frac{P_m(k)}{1 + (k_{\parallel} \sigma_p)^2 / 2}, \quad (11)$$

where m integrates over the halo mass range, C_{H_I} is the conversion factor from H_I mass density to temperature, $n(m)$ is the halo mass function (Tinker et al., 2008), $b(m)$ is the linear halo bias (Tinker et al., 2010), $\langle M_{\text{H}_I}(m) \rangle$ is the H_I-halo mass relation, $u_{\text{H}_I}(k|m) = \mathcal{F}(\rho_{\text{H}_I}(r|m))$ is the normalised Fourier-transform of the density profile $\rho_{\text{H}_I}(r|m)$, $k = \sqrt{k_{\perp}^2 + k_{\parallel}^2}$ and $b_{\text{H}_I}^0$ is the linear H_I bias. The $(1 + (f k_{\parallel}^2)/(b_{\text{H}_I}^0 k^2))^2$ term denotes the Kaiser effect (Kaiser, 1987), and $1 + (k_{\parallel} \sigma_p)^2 / 2$ denotes the finger-of-god (FoG) effect (Jackson, 1972). f is the growth rate and $\sigma_p = \sigma_v(1+z)/(Hf)$ where σ_v is the velocity dispersion and H is the Hubble parameter. $P_m(k)$ is the matter power spectrum.

The linear H_I bias is calculated as

$$b_{\text{H}_I}^0 = \frac{\int dm n(m) b(m) \langle M_{\text{H}_I}(m) \rangle}{\int dn n(m) \langle M_{\text{H}_I}(m) \rangle}, \quad (12)$$

while the 1-halo term can be modelled as

$$P_{1\text{h}}(k, k_{\parallel}) = C_{\text{H}_I}^2 \int dm n(m) \langle M_{\text{H}_I}(m) \rangle^2 u_{\text{H}_I}(k|m) \frac{1}{1 + (k_{\parallel} \sigma_p)^2 / 2}. \quad (13)$$

The shot noise term $P_{\text{SN}}(k_{\parallel})$ can be written as

$$P_{\text{SN}}(k_{\parallel}) = \frac{P_{\text{SN}}^0}{1 + (k_{\parallel} \sigma_p)^2 / 2}. \quad (14)$$

The shot noise in the H_I power spectrum is also attenuated by the FoG effect. This is because the 21 cm line is broadened by the internal velocity dispersion of H_I galaxies. The broadening breaks the point-source assumption along the line of sight, smearing the overall power spectrum along the k_{\parallel} direction, including the shot noise (e.g. Zhang et al. 2020).

	$\log_{10}[A_{\text{HI}}/M_{\odot}]$	β	γ	σ_v [km/s]	P_{SN}^0 [Mpc ³]	$\Omega_{\text{HI}} \times 10^4$
Prior	(7,10)	(0,4)	(1,3)	(0,250)	(1,500)	(1,100)
Fiducial	8.600	0.700	1.50	50.00	120.00	4.95
Forecast	$8.599^{+0.004}_{-0.007}$	$0.697^{+0.033}_{-0.030}$	$1.48^{+0.28}_{-0.26}$	$49.93^{+5.85}_{-5.77}$	$118.10^{+15.22}_{-13.58}$	$4.96^{+0.12}_{-0.14}$

Table 1: The forecasts for halo model constraints using SKA-Mid interferometric HI intensity mapping at $z \sim 1$. The first row lists the model parameters and the derived quantity Ω_{HI} . The second row lists the boundary of the flat priors for the parameters. The third row lists the fiducial values of the parameters. The fourth row lists the forecasted constraints with the [16%, 50%, 84%] confidence interval.

Using halo model, we can also calculate the average HI density as

$$\Omega_{\text{HI}} = \int dn n(M) \langle M_{\text{HI}}(m) \rangle / \rho_c, \quad (15)$$

where ρ_c is the critical density of the Universe at $z = 0$.

We follow the parameterisation in Paul et al. (2023) so that

$$\begin{aligned} \langle M_{\text{HI}}(m) \rangle &= A_{\text{HI}} \left(\frac{m}{M_0} \right)^\beta, \\ \rho_{\text{HI}}(r|m) &= \rho_0 \left(\frac{r}{r_0} \right)^\gamma \exp[-ar/r_0], \end{aligned} \quad (16)$$

where r is the radius within a dark matter halo and $M_0 = 10^{10} M_{\odot} h^{-1}$. A_{HI} is a free parameter that controls the amplitude of the HI-halo mass relation, β determines the slope of the relation, γ determines the slope of the HI mass profile inside the halos. Together with the velocity dispersion σ_v and the shot noise P_{SN}^0 , they form the 5-parameter model which we use for forecasts. For simplicity, we present the forecasts only for the $z \sim 1$ bin. The fiducial values adopted are listed in Table 1.

Using the fiducial model and the measurement error presented in Section 3, we perform Bayesian inference on the model parameters, assuming a diagonal Gaussian likelihood. We use NAUTILUS (Lange, 2023) to perform importance nested sampling. We adopt wide, flat priors for all model parameters listed in Table 1. The γ parameter has a prior of (1,3) as values outside this region gives unphysical divergence. We choose the number of live points to be 2000 and consider the sampling to have converged if the evidence in the live set is less than 1% of the total evidence (see Eq. 8 of Lange 2023). The resulting 1σ confidence interval is listed in Table 1, and the posterior presented in Figure 6. As shown, the halo model parameters can be precisely constrained, resulting in unprecedented measurements on the relation between the dark matter halos and their HI content. As we demonstrate in Figure 6, despite parameter degeneracies, the high signal-to-noise measurements of SKA-Mid will be able to break the parameter degeneracy and provide unbiased estimation of the model parameters. In particular, extremely accurate measurements of the HI-halo mass relation will be probed, providing important insights into hydrodynamical simulations of cosmic structure. Moreover, the halo model constraints can in turn produce an accurate measurement of the HI density $\Omega_{\text{HI}} = 4.96^{+0.12}_{-0.14} \times 10^{-4}$, extending HI density measurements using emission line to $z \sim 1.0$ and beyond.

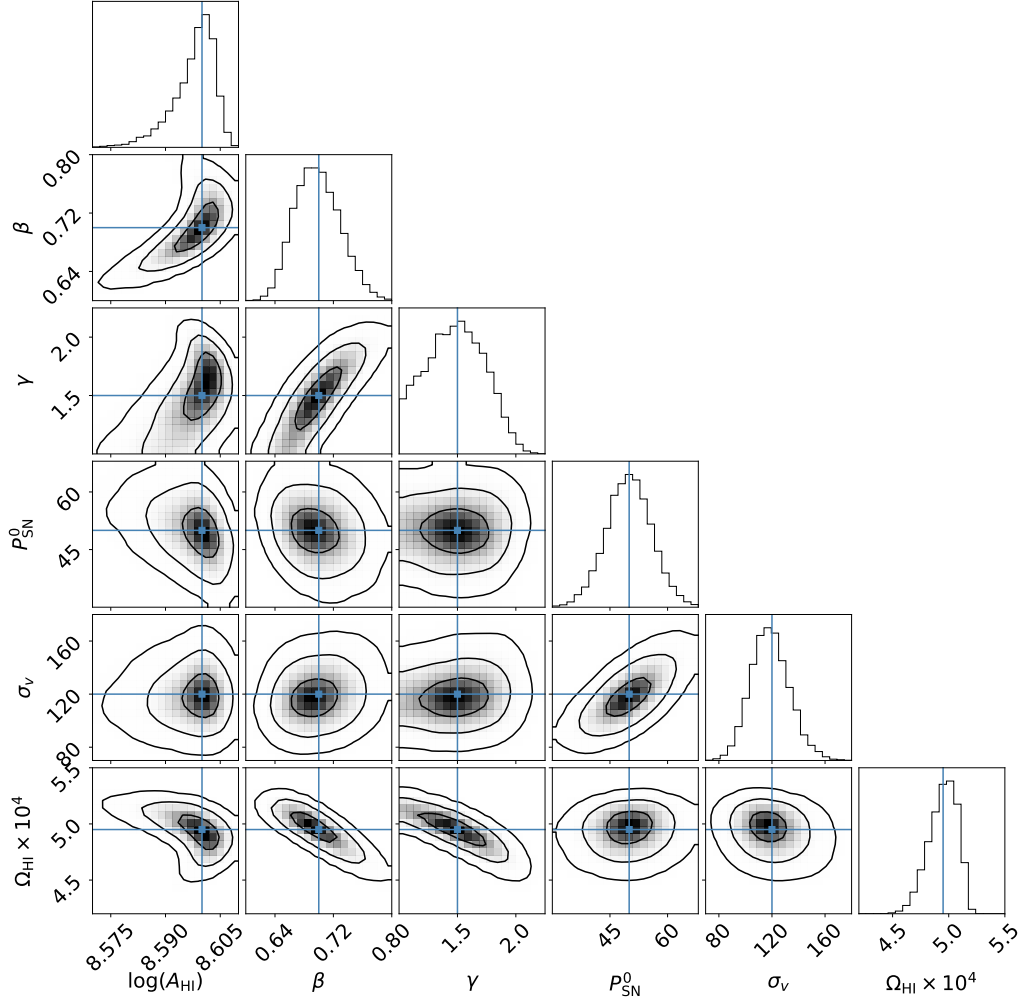


Figure 6: Forecasts on the posterior distribution of the halo model parameters for the interferometric HI intensity mapping measurements for SKA-Mid AA4 configuration at $z \sim 1$. The contours show the 1σ , 2σ and 3σ confidence level, respectively. The solid lines denote the fiducial values listed in [Table 1](#).

4.2 HIMF

The halo model constraints from the measured HI power spectrum can be turned into constraints on the HI mass function (HIMF), as suggested in [Chen et al. \(2021\)](#). The HIMF is typically parameterised as ([Schechter, 1976](#)):

$$\phi_{\text{HI}}(M_{\text{HI}}) \equiv \frac{dn_{\text{HI}}}{d\log M_{\text{HI}}} = \ln 10 \phi_* \left(\frac{M_{\text{HI}}}{M_*}\right)^{\alpha+1} e^{-\frac{M_{\text{HI}}}{M_*}}, \quad (17)$$

where $[\phi_*, M_*, \alpha]$ are the parameters that control the amplitude, knee mass and the slope of the HIMF respectively.

The HIMF can then be used to calculate the HI density and the real-space shot noise ([Chen et al.](#),

2021):

$$\Omega_{\text{H}\text{I}} = \int d\log M_{\text{H}\text{I}} \phi_{\text{H}\text{I}}(M_{\text{H}\text{I}}) M_{\text{H}\text{I}} / \rho_c \quad (18)$$

$$P_{\text{SN}}^0 = (C_{\text{H}\text{I}} \rho_c \Omega_{\text{H}\text{I}})^2 \frac{\int d\log M_{\text{H}\text{I}} \phi_{\text{H}\text{I}}(M_{\text{H}\text{I}}) M_{\text{H}\text{I}}^2}{\left(\int d\log M_{\text{H}\text{I}} \phi_{\text{H}\text{I}}(M_{\text{H}\text{I}}) M_{\text{H}\text{I}}\right)^2} \quad (19)$$

Equation 17 contains three HIMF parameters. In principle, the H I power spectrum provides the measurements of the overall H I density as well as the shot noise. Combining H I power spectrum with complimentary measurements, such as stacking, can be used to fully constrain the HIMF. In this chapter, we focus on using only the H I power spectrum. In this case, the HIMF parameters will not be fully constrained, due to the degeneracy of HIMF parameters in $\Omega_{\text{H}\text{I}}$ and P_{SN}^0 . However, the parameter posterior still gives informative constraints on the shape of the HIMF, especially on the “knee mass” of the HIMF.

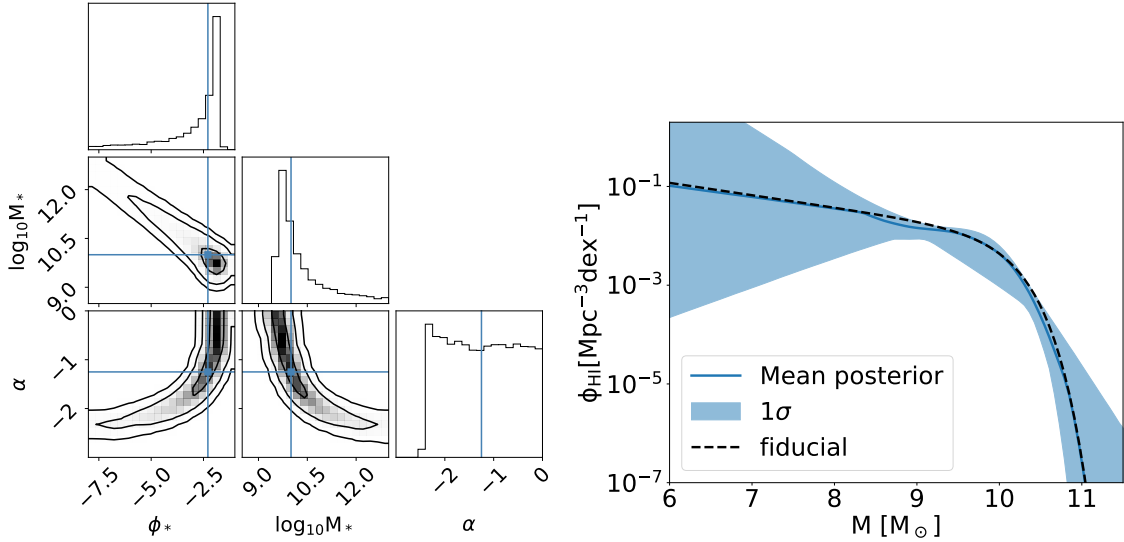


Figure 7: Left panel: Forecasts on the posterior distribution of the HIMF parameters for SKA-Mid AA4 configuration at $z \sim 1$. The contours show the 1 σ , 2 σ and 3 σ confidence level, respectively. Right panel: Forecast on the posterior distribution of the measured HIMF. The black dashed line denotes the fiducial HIMF. The blue solid line denotes the mean of the posterior. The shaded region denotes the 1 σ confidence interval of the measurement.

We repeat the same Bayesian inference routine in Section 4.1, and substitute the shot noise free parameter with Equation 19. M_* and α parameters are sampled with wide flat priors of $\log[M_*/M_{\odot}] \in (6, 13)$ and $\alpha \in (-5, 0)$, whereas the ϕ_* is calculated in each sample point based on $\Omega_{\text{H}\text{I}}$ according to Equation 15 and Equation 18. The results are shown in Figure 7. The HIMF parameter constraints are significantly degenerate due to the lack of information on the number density of H I sources from intensity mapping. The slope of the HIMF is not sensitive to intensity mapping measurements and is largely unconstrained, except for a lower limit of $\alpha \gtrsim -2.5$. The amplitude and the knee mass of the HIMF, on the other hand, can be measured. The reconstructed HIMF function is shown in the right panel of Figure 7. As shown, the posterior can reconstruct

the overall shape of the HIMF, especially in the mass range around the knee mass. The results demonstrate that H I intensity mapping with SKA-Mid can be used to constrain HIMF, pushing the measurements beyond previous limits $z \lesssim 0.5$.

4.3 Progress in modelling non-linear power spectrum

The anticipated science outputs shown above demonstrate the power of H I intensity mapping, and the importance of modelling the H I power spectrum at non-linear scales. In particular, the modelling of the H I shot noise is important, as the shot noise dominates the clustering amplitude in small scales, and enables constraints on the HIMF.

In a low-redshift universe ($z \leq 1.0$), most of the atomic hydrogen (H I) gas resides in the cold dense region of galaxies, where it can be self-shielded from ultraviolet photons (see e.g., [Krumholz et al., 2009](#); [Diemer et al., 2018](#); [Villaescusa-Navarro et al., 2018](#)). Ideally, particle information is needed to obtain the non-linear ($k \gtrsim 1 \text{ Mpc}^{-1}$) power spectrum given the scales. However, there might not be much information contained in the power spectrum with very small scales (say, $k \gtrsim 10 \text{ Mpc}^{-1}$). For the scale of our concern ($1 \text{ Mpc}^{-1} \lesssim k \lesssim 10 \text{ Mpc}^{-1}$), several factors might play a role, including the diameter of H I discs, stretching length of H I velocity dispersion in redshift space and the distances between galaxies. When modelling the H I field for some specific surveys, it is strongly recommended to compare the scale of different factors with the resolution or the scale of your concern along and perpendicular to the line-of-sight direction, and then to determine which factors to be modelled.

Take the detection by [Paul et al. \(2023\)](#) as an example, [Li et al. \(2024\)](#) compares different factors qualitatively by assuming that the H I statistics would not change dramatically as z changes from 0.0 to 1.0. It is found that the sizes of H I discs have subtle effects, while other factors are more important. Although the H I IM surveys performed by the MeerKAT interferometers do not identify individual sources as the extragalactic surveys do, these sources can still be resolved by the surveys and leave some imprints on the measured nonlinear H I power spectrum. Specifically, sub-structures of H I insides halos alter the shot noise term of the power spectrum, and galaxy peculiar velocities exacerbate the Finger of God effect. Hence, halo models without substructure modelling are not ideal for predictions on nonlinear scales. Velocity dispersions due to the H I gas random motions and circular rotations suppress the non-linear power spectrum, especially the shot noise term. [Li et al. \(2024\)](#) further finds that the H I power spectrum is relatively insensitive to the profile shape of the H I emission line at these scales, while showing a strong correlation with the profile width. By assuming a Gaussian shape for each emission line profile, the shot noise term can be modelled as Equation 20.

$$P_{\text{SN}}(k, \mu) = \frac{V}{M_t^2} \cdot \sum_i M_i^2 \cdot e^{-\sigma_i^2 \mu^2 k^2} \quad , \quad \mu = k_z/k \quad (20)$$

where V is the volume of the simulated box; M_t is the total mass of all H I sources; M_i is the H I mass of each galactic H I source; σ_i is the σ parameter in the Gaussian profiles. σ_i is used to describe the profile width and is thus determined by the Full Width Half Maximum (FWHM) w_{50} obtained in extragalactic H I surveys as Equation 21.

$$\sigma_i = \frac{l_{w50}}{2\sqrt{2 \ln 2}}, \quad (21)$$

where the l_{w50} is the stretched length due to the profile width in the redshift space.

If we have a 2-dimensional distribution $\phi(M, w_{50})$, Equation 20 can also be written as Equation 22. It is apparent that shot noise term have strong correlation with the 2D HI distribution $\phi(M, w_{50})$ obtained in extragalactic surveys, and shot noise term in the power spectrum can thus provide constraint on the HI velocity function and also HI mass function.

$$P_{\text{SN}}(k, \mu) = V \cdot \frac{\int \int \phi(M, w_{50}) M^2 \cdot e^{-\sigma(w_{50})^2 \mu^2 k^2} dM dw_{50}}{\int \int \phi(M, w_{50}) M dM dw_{50}}, \quad \mu = k_z/k. \quad (22)$$

Additionally, although it is easily overlooked, the observational settings, like horizon limit and frequency resolution, can have a relatively large impact on the results (see Figure 11 in Li et al., 2024). Different surveys might utilise different observational settings, so forecasting based on specific settings of the different surveys is always essential.

4.4 Cross-correlation with galaxy surveys

The 21-cm signal traces the neutral gas distribution over a wide redshift range, probing cosmological volumes of the order of a few tenths of Gpc^3 around $z \sim 1$. The inherently higher resolution of radio data in both spatial and radial directions (compared to optical spectroscopic surveys) would enable efficient mitigation of instrumental and astrophysical effects originating from separate sources in optical and radio data through cross-correlation measurements. While such investigations are routinely conducted with single dish maps (see Cunnington&Wang et al. (2026) for a discussion of progress with MeerKAT), direct cross-correlation between interferometric intensity maps and other large-scale structure tracers is yet to be achieved, though there have been positive results using stacking out to redshift ~ 1 (for example Amiri et al. (2023)). In this section, we discuss the prospects of cross-correlation without using stacking.

As discussed earlier, the prospects of intensity mapping are very promising with the SKA-Mid owing to its compact baselines design and excellent noise characteristics. Using these datasets, we can realistically forecast the prospects of cross-correlation with galaxy redshift surveys. Several of the surveys with SKA-Mid will overlap in sky coverage with Stage IV optical surveys such as Euclid and LSST. Thus, a realistic cross-correlation measurement can be achieved with a survey covering only a few hundred square degrees. Figure 8 shows the forecasts on the uncertainties of the cross-power measurements for such a scenario. It assumes an SKAO ‘‘medium-deep’’ type radio survey covering 100 square degrees with 1000 hours observing time and an overlapping survey with a volume density 10^{-4} for galaxy number counts. The forecast assumes a redshift range $0.5 \lesssim z \lesssim 1.5$, accessible with the Band-1 of SKA-Mid. As seen in Figure 8, the uncertainties achieved under these realistic survey specifications are quite low between scales ~ 0.5 to 10 hMpc^{-1} . Beyond $\sim 10 \text{ hMpc}^{-1}$, the cross-power starts being shot noise dominated, hence, enhanced observing times are of little help. However, the additive shot noise in the cross-power spectrum is also a useful probe of gas astrophysics. It is to be noted that the optical galaxy number density is representative of the Euclid

survey (Euclid Collaboration et al., 2020), and the expected signal-to-noise would vary depending on the specifications of the galaxy survey used. A lower sensitivity or shallower survey may reduce the galaxy number counts or enhance the shot noise levels or both, which affects the SNR. Thus, such a survey with higher uncertainties on the cross-power spectrum would lead to looser constraints on the subsequent parameters estimated. However, these forecasts are still valid under the realistic condition that the patch of the sky with radio data is also covered by Euclid.

Cross-correlation measurement can be used as a complementary tool to auto-correlation for not only parameter estimation but also diagnostics. While cross-correlation alone would not independently measure the full clustering amplitude, different tracers affected by uncorrelated systematic biases can help us quantify them (for example, any bias due to foreground residuals in H I auto-power spectrum would be largely cancelled in the cross-power). Sections 4.1 to 4.3 discuss the constraints on global H I properties using auto-correlation data. However, using scales accessible with interferometers, cross-correlations with optical galaxy catalogues can also be used to determine the average H I content in galaxies. Galaxy catalogues binned using different selection criteria, such as stellar mass and colour, can be used to probe H I scaling relations. Further, Wolz et al. (2017) showed that the additive shot noise term in the cross-power spectrum scales with the average H I brightness temperature of the optically selected galaxies. Thus, the power spectrum obtained from cross-correlation between H I intensity maps and optical galaxies on small scales presents a unique opportunity for studying galaxy evolution independently.

Further, cross-correlation of H I intensity maps with Ly- α data from spectroscopic surveys will help remove some of the degeneracies. One of the important contaminants in Ly- α data inhibiting cosmological interpretations is the “high-column density systems” (HCDs), dense clouds of H I associated with the circumgalactic medium. In contrast, the Lyman-alpha forest cosmological analysis focuses on the dilute intergalactic medium. The H I intensity-mapping signal is expected to be dominated by H I gas in galaxies, and hence highly correlated with the HCDs. It should be possible to estimate the impact of HCDs on the cosmological constraints from the Ly- α forest analysis by cross-correlating the optical spectroscopic data with 21-cm intensity maps. Overall, the SKAO presents unprecedented opportunities to leverage cross-correlations between different large-scale structure probes, at least to redshift ~ 1 .

5 Conclusion

This chapter describes the progress in post-EoR H I intensity mapping with radio interferometers. At the redshift range in question, while single dish surveys are sensitive to the large angular scales, interferometers are sensitive to the smaller angular separations. Hence, observation and modelling in these scales probe the neutral gas astrophysics as well as the relationship between the gas and the host haloes. There has been significant progress in recent years in both observations and theoretical modelling.

The main highlights of the milestones from observations and systematic modelling include:

1. A tentative H I power spectrum detection reported using ~ 100 hours deep observations of MeerKAT telescope. Observations targeted redshifts centred around $z \sim 0.32$ and $z \sim 0.44$.

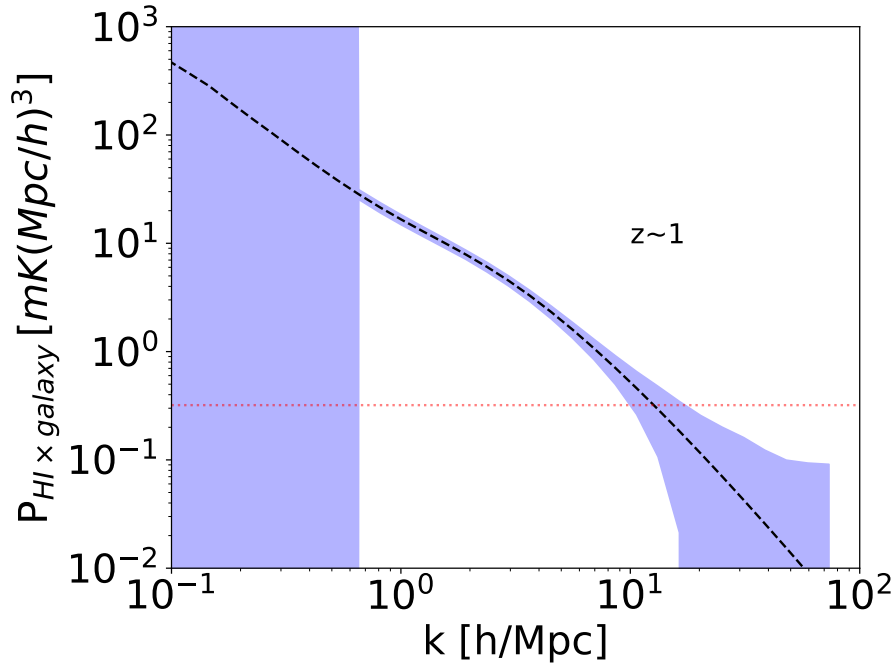


Figure 8: Forecasts on the uncertainty on the HI-galaxy cross-correlation power spectrum using galaxy number density for Euclid survey (10^{-4}) and a 1000 hour survey with the Band-1 of SKA-Mid using AA4 configuration. The grey dotted line shows the cross-power amplitude, and the red dashed line shows the shot-noise levels.

The calibration and power spectrum pipelines developed enabled robust delay-spectrum estimation directly from visibilities.

2. An upper limit on the HI power spectrum was placed using wide-area observations of the MeerKAT MIGHTEE survey. Using the incoherent averaging technique to average data from multiple pointings in the Fourier domain, this work used a redshift bin centred at $z \sim 0.44$ and demonstrated the feasibility of using wide area surveys.
3. An investigation on testing the performance of intensity mapping, using HI sources directly detected from MIGHTEE observations, is demonstrated for $z < 0.1$. It allows us to test the robustness of the methodology developed to investigate how much HI either the intensity mapping signal or galaxy detections miss. The work demonstrates using intensity mapping and galaxy detections that, in the limit of the galaxy mesh resolution, the power spectrum measurements from interferometric intensity mapping data agree well with the HI galaxy catalogue, but become unreliable at scales smaller than the resolution.
4. There have also been developments of several systematic isolation methods, such as using baseline and uv -domain flagging, using the 2D Fourier plane & development of MASS, a simulator for estimating visibilities and systematics from MeerKAT.

Parallely, there have also been developments in modelling the non-linear power spectrum and constraints on the halo model of cosmic HI. Forecasts for SKA data using a standard survey of 100

trackings of different pointings, each with 10 hrs of integration time, show promising results for recovery of the H_I power spectrum and subsequent estimation of halo model parameters and the H_I mass function. Using the halo model formalism, constraints of the H_I density Ω_{H_I} can be achieved at per-cent level accuracy at $z \sim 1.0$. Measurements of the H_I density and the H_I shot noise can be used to measure the HIMF at $z \sim 1.0$, beyond limits of conventional H_I stacking techniques. Interferometric intensity mapping with SKA-Mid will be able to open a new window into probing the cosmic H_I content and the galaxy evolution at $1.0 \lesssim z \lesssim 3.0$.

References

- M. Amiri et al. *The Astrophysical Journal*, 947(1):16, Apr. 2023. ISSN 1538-4357. doi: 10.3847/1538-4357/acb13f. URL <http://dx.doi.org/10.3847/1538-4357/acb13f>.
- S. E. I. Bosman et al. *MNRAS*, 514(1):55–76, July 2022. doi: 10.1093/mnras/stac1046.
- R. Braun et al. *arXiv e-prints*, art. arXiv:1912.12699, Dec. 2019. doi: 10.48550/arXiv.1912.12699.
- A. Chakraborty et al. *The Astrophysical Journal Letters*, 907(1):L7, jan 2021. doi: 10.3847/2041-8213/abd17a. URL <https://dx.doi.org/10.3847/2041-8213/abd17a>.
- Z. Chen, L. Wolz, M. Spinelli, and S. G. Murray. *MNRAS*, 502(4):5259–5276, Apr. 2021. doi: 10.1093/mnras/stab386.
- CHIME Collaboration et al. *arXiv e-prints*, art. arXiv:2309.04404, Sept. 2023. doi: 10.48550/arXiv.2309.04404.
- CHIME Collaboration et al. *arXiv e-prints*, art. arXiv:2511.19620, Nov. 2025. doi: 10.48550/arXiv.2511.19620.
- A. Cooray and R. Sheth. *Phys. Rep.*, 372(1):1–129, Dec. 2002. doi: 10.1016/S0370-1573(02)00276-4.
- D. Crichton et al. *Journal of Astronomical Telescopes, Instruments, and Systems*, 8:011019, Jan. 2022. doi: 10.1117/1.JATIS.8.1.011019.
- S. Cunnington&Wang et al. In *Advancing Astrophysics with the SKA – II (AASKAII)*. 2026. arXiv search: Report number AASKAII/Cunnington01.
- B. Diemer et al. *ApJSS*, 238(2):33, Oct. 2018. doi: 10.3847/1538-4365/aae387.
- K. M. A. Elahi et al. In *Advancing Astrophysics with the SKA – II (AASKAII)*. 2026. arXiv search: Report number AASKAII/Elahi01.
- Euclid Collaboration et al. *A&A*, 642:A191, Oct. 2020. doi: 10.1051/0004-6361/202038071.
- K. M. Gorski et al. *ApJ*, 622(2):759–771, apr 2005. doi: 10.1086/427976. URL <https://doi.org/10.1086%2F427976>.
- I. Heywood et al. *MNRAS*, 509(2):2150–2168, Jan. 2022. doi: 10.1093/mnras/stab3021.
- I. Heywood et al. *MNRAS*, 534(1):76–96, Oct. 2024. doi: 10.1093/mnras/stae2081.
- J. C. Jackson. *MNRAS*, 156:1P, Jan. 1972. doi: 10.1093/mnras/156.1.1P.
- M. Jarvis et al. In *Proceedings of MeerKAT Science: On the Pathway to the SKA — PoS(MeerKAT2016)*, volume 277, page 006, 2016. doi: 10.22323/1.277.0006.
- M. Jarvis et al. In *MeerKAT Science: On the Pathway to the SKA*, page 6, Jan. 2016. doi: 10.22323/1.277.0006.
- N. Kaiser. *MNRAS*, 227:1–21, July 1987. doi: 10.1093/mnras/227.1.1.

- M. R. Krumholz, C. F. McKee, and J. Tumlinson. *ApJ*, 699(1):850–856, July 2009. doi: 10.1088/0004-637X/699/1/850.
- J. U. Lange. *Monthly Notices of the Royal Astronomical Society*, 525(2):3181–3194, 08 2023. doi: 10.1093/mnras/stad2441. URL <https://doi.org/10.1093/mnras/stad2441>.
- Z. Li et al. *MNRAS*, 534(3):1801–1815, Nov. 2024. doi: 10.1093/mnras/stae2182.
- N. Maddox et al. *A&A*, 646:A35, Feb. 2021. doi: 10.1051/0004-6361/202039655.
- A. M. Matthews, J. J. Condon, W. D. Cotton, and T. Mauch. *ApJ*, 909(2):193, Mar. 2021a. doi: 10.3847/1538-4357/abdd37.
- A. M. Matthews, J. J. Condon, W. D. Cotton, and T. Mauch. *ApJ*, 914(2):126, June 2021b. doi: 10.3847/1538-4357/abfaf6.
- T. Mauch et al. *ApJ*, 888(2):61, Jan. 2020. doi: 10.3847/1538-4357/ab5d2d.
- J. Mayor et al. In *Advancing Astrophysics with the SKA – II (AASKAII)*. 2026. arXiv search: Report number AASKAII/Mayor01.
- A. Mazumder et al. *MNRAS*, 541(1):476–493, July 2025. doi: 10.1093/mnras/staf975.
- M. F. Morales. *ApJ*, 619(2):678–683, Feb. 2005. doi: 10.1086/426730.
- A. Nasirudin et al. In *Advancing Astrophysics with the SKA – II (AASKAII)*. 2026. arXiv search: Report number AASKAII/Nasirudin01.
- A. Parsons et al. *ApJ*, 753(1):81, July 2012. doi: 10.1088/0004-637X/753/1/81.
- S. Paul et al. *ApJ*, 833(2):213, Dec. 2016. doi: 10.3847/1538-4357/833/2/213.
- S. Paul, M. G. Santos, Z. Chen, and L. Wolz. *arXiv e-prints*, art. arXiv:2301.11943, Jan. 2023. doi: 10.48550/arXiv.2301.11943.
- Planck Collaboration et al. *A&A*, 641:A6, Sept. 2020. doi: 10.1051/0004-6361/201833910.
- P. Schechter. *ApJ*, 203:297–306, Jan. 1976. doi: 10.1086/154079.
- N. Thyagarajan et al. *ApJ*, 804(1):14, May 2015. doi: 10.1088/0004-637X/804/1/14.
- J. Tinker et al. *ApJ*, 688(2):709–728, Dec. 2008. doi: 10.1086/591439.
- J. L. Tinker et al. *ApJ*, 724(2):878–886, Dec. 2010. doi: 10.1088/0004-637X/724/2/878.
- J. Townsend et al. *arXiv e-prints*, art. arXiv:2602.23055, Feb. 2026.
- K. Vanderlinde et al. In *Canadian Long Range Plan for Astronomy and Astrophysics White Papers*, volume 2020, page 28, Oct. 2019. doi: 10.5281/zenodo.3765414.
- F. Villaescusa-Navarro et al. *ApJ*, 866(2):135, Oct. 2018. doi: 10.3847/1538-4357/aadba0.
- L. Wolz, C. Blake, and J. S. B. Wyithe. *MNRAS*, 470(3):3220–3226, Sept. 2017. doi: 10.1093/mnras/stx1388.
- J. Zhang et al. *ApJ*, 895(1):34, May 2020. doi: 10.3847/1538-4357/ab88b1.