

# Forecast for FAST: from galaxies survey to intensity mapping

Wenkai Hu<sup>1,2</sup>, Xin Wang<sup>3</sup>, Fengquan Wu<sup>1</sup>, Yougang Wang<sup>1</sup>, Pengjie Zhang<sup>4</sup>  
and Xuelei Chen<sup>1,2,5</sup>★

<sup>1</sup>Key Laboratory of National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012, China

<sup>2</sup>School of Astronomy and Space Science, University of Chinese Academy of Sciences, Beijing 100049, China

<sup>3</sup>School of Physics and Astronomy, Sun Yat-Sen University, Guangzhou 510297, China

<sup>4</sup>Department of Astronomy, School of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai, 200240, China

<sup>5</sup>Center of High Energy Physics, Peking University, Beijing 100871, China

Accepted 2020 February 28. Received 2020 February 28; in original form 2019 September 24

## ABSTRACT

The Five-Hundred-Meter Aperture Spherical Radio Telescope (FAST) is the largest single-dish radio telescope in the world. In this paper, we make forecast on the FAST H I large-scale structure survey by mock observations. We consider a drift scan survey with the L-band 19 beam receiver, which may be commensal with the pulsar search and Galactic H I survey. We also consider surveys at lower frequency, using either the current single feed wide-band receiver or a future multibeam phased array feed (PAF) in the UHF band. We estimate the number density of detected H I galaxies and the measurement error in positions and the precision of the surveys are evaluated using both Fisher matrix and simulated observations. The measurement error in the H I galaxy power spectrum is estimated, and we find that the error is relatively large even at moderate redshifts, as the number of positively detected galaxies drops drastically with increasing redshift. However, good cosmological measurement could be obtained with the intensity mapping technique where the large scale H I distribution is measured without resolving individual galaxies. The figure of merit for the dark energy equation of state with different observation times is estimated, and we find that with the existing L-band multibeam receiver, a good measurement of low redshift large-scale structure can be obtained, which complements the existing optical surveys. With a PAF in the UHF band, the constraint can be much stronger, reaching the level of a dark energy task force stage IV experiment.

**Key words:** galaxies: evolution – galaxies: ISM – radio lines: galaxies.

## 1 INTRODUCTION

Over the past two decades, an increasing number of optical galaxy surveys, such as the 2DFGRS (Percival et al. 2001), SDSS (Gunn & Weinberg 1995; Tegmark et al. 2004; Schlegel, White & Eisenstein 2009; Dawson et al. 2016), and WiggleZ (Glazebrook et al. 2007; Drinkwater et al. 2010) are probing increasingly large volumes of the Universe and provide large-scale structure (LSS) data for cosmological studies. For example, assuming that the observed number density of galaxies traces the total density of the matter distribution, the baryon acoustic oscillation (BAO) features in the galaxy power spectrum are measured and used as standard rulers to constrain cosmological models (Cole et al. 2005; Eisenstein et al. 2005; Percival et al. 2010; Beutler et al. 2011; Blake et al. 2011; Ross et al. 2015; Alam et al. 2017; Ata et al. 2018). While it is plausible that the galaxies formed from overdensity perturbations

and therefore trace the total density on large scales, it is vital to check this hypothesis and also understand the range of validity of it by observing the galaxies with different means. The 21-cm line of the neutral hydrogen (H I) provides a good alternative way of observation in the radio wavelength. A number of H I galaxies surveys have also been carried out, e.g. the HIPASS survey (Meyer et al. 2004; Zwaan et al. 2004), the ALFALFA survey (Giovanelli et al. 2005; Giovanelli et al. 2007; Saintonge 2007), and JVLA deep survey Jarvis et al. (2014). However, limited by the sensitivity of the telescopes, the redshift range of these surveys is much smaller than the current optical surveys.

A new generation of radio telescopes is being built or underdevelopment, including the square kilometer array (SKA) in the Southern hemisphere, and the FAST (Nan et al. 2011) in the Northern hemisphere. These radio telescopes have much better sensitivities and observe H I galaxies at larger distances. Here we consider FAST, which is about the complete its commissioning process and starts science runs. It has a very large aperture (300 m during operation) and is to be equipped with multibeam feed system and low-noise cryogenic receivers, ideal for conducting large surveys.

\* E-mail: xuelei@cosmology.bao.ac.cn

The FAST has unprecedented large effective area and high sensitivity; nevertheless, for a traditional galaxy survey (Duffy et al. 2008), the redshift at which an individual galaxy could be detected is still very limited, and its angular resolution would be insufficient to resolve the galaxy at high redshift. However, to map the LSS, in principle it is not necessary to resolve individual galaxies as traditional galaxy surveys do, instead the redshifted 21-cm line intensity can be mapped with lower angular resolution, as is done in the Epoch of Reionization (EoR) experiments.<sup>1</sup> In more general context, Chang et al. (2008) studied this mode of observation and named it the *intensity mapping* method and also proposed that a cost-effective way to survey LSS is to develop a dedicated dense array of cylinder or small dish antennas (Ansari et al. 2008; Chang et al. 2008; Seo et al. 2010; Ansari et al. 2012). Indeed, a number of such small-to-mid scale experiments are undergoing, such as those of Tianlai (Chen 2012; Xu, Wang & Chen 2015), CHIME (Bandura et al. 2014), and HIRAX (Newburgh et al. 2016), as well as the specially designed single dish experiment BINGO (Battye et al. 2012, 2016).

For the FAST itself, several studies used the Fisher matrix formalism to make simple forecasts on the constraining power of cosmological parameters by H I galaxy survey (Duffy et al. 2008) or intensity mapping surveys (Bigot-Sazy et al. 2015; Smoot & Debono 2017; Yohana, Li & Ma 2019).

In this paper, we make a more detailed investigation by simulating the observed galaxies and also compare the galaxy survey and intensity mapping. The layout of this paper is as follows. In Section 2, we describe our model of the telescope and its receiver feeds. In Section 3, we present the modelling of the H I galaxies and their observation, as well as the simulated intensity map. In Section 4, we make Fisher matrix forecasts of the precision of power spectrum measurement using both H I galaxy surveys and intensity mapping surveys and also make measurement using numerical simulation. The niche of H I galaxy survey and H I intensity mapping survey and the effect of foreground are discussed in Section 5. Finally, we summarize the results in Section 6.

## 2 THE FAST TELESCOPE

In order to study how the FAST could survey the large-scale-structure, we conduct mock observations with simulated sky. We first generate a catalogue of galaxies from simulation and then convert it into the simulated sky of H I intensity as would be observed by FAST.

### 2.1 The instrument

The diameter of the FAST reflector is 500 m; the fully illuminated aperture at any time is  $D = 300$  m since the telescope is designed to track objects. The beam size of the FAST is given by

$$\theta = 1.22 \times \frac{21 \text{ cm}(1+z)}{300 \text{ m}} = 2.94(1+z) \text{ arcmin} \quad (1)$$

for observation of the 21-cm line from redshift  $z$ .

During a drift scan, a single feed is fixed to be pointed to a particular declination in the due north or south direction, so that in a sidereal day, a ring of width  $2.94(1+z)$  arcmin centred at that

<sup>1</sup>One of us (X. C.) first realized that this mode of observation could be used with FAST to probe the large-scale structure and presented it at a meeting on the FAST science case held in May 2007 in Hangzhou, China.

**Table 1.** FAST survey receiver parameters. The  $t_{\text{sur}}$  refers to the time needed to finish a full drift scan of  $\pm 40$  deg from the centre declination of FAST. The  $T_{\text{rec}}$  is the receiver noise.

Receiver	Band (GHz)	Beams	$T_{\text{rec}}$ (K)	$t_{\text{sur}}$ (days)
L-band	1.05–1.45	19	20	220
Wide-band	0.27–1.62	1	60	1211
UHF PAF (future)	0.5–1.0	81	30	135

declination is scanned. The pointing declination can be changed so as to cover the whole observable part of the sky. The FAST site is located at a latitude of  $25^\circ 48'$  North, and the maximum zenith angle is 40 deg, allowing the observation of  $\approx 50$  per cent of the full sky or about 20,000 deg<sup>2</sup>.

The FAST is equipped with a number of different feed and receiver systems. For H I survey, the most relevant are the L-band 19-beam feed/receiver system and the wide-band receiver system. Additionally, there are also several low-frequency receivers that cover down to 70 MHz, which can be used for EoR observations. Here, we shall consider mainly the first two, which are relevant for low-or-mid redshift observations of LSS. In addition, below we shall also consider a possible future UHF phased array feed (PAF) system. We summarize the information of these receiver systems in Table 1.

#### 2.1.1 L-band 19 beam receiver system

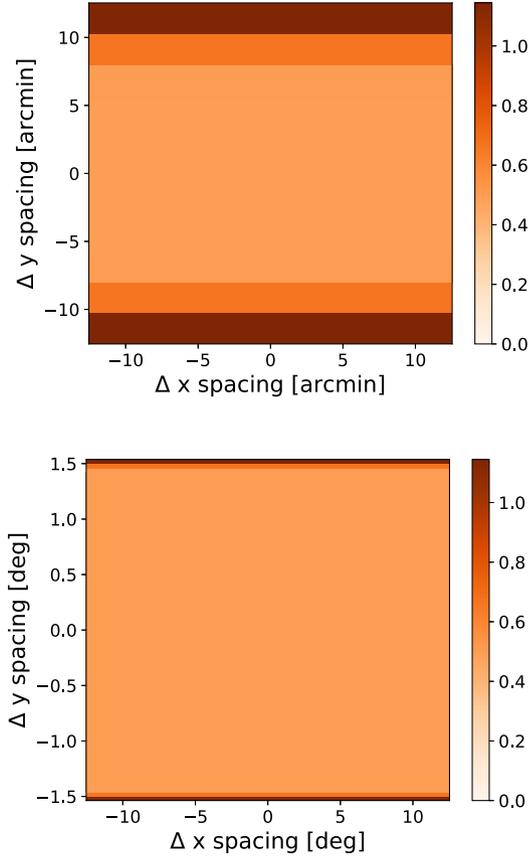
It covers the frequency range of 1.05–1.45 GHz, and the beams are arranged in two concentric hexagonal rings around the central beam. The minimum spacing between beam centres is 5.73 arcmin and is approximately constant, though for each beam the width scales roughly as  $\theta \propto (1+z)$ . In this paper, we assume that the feed array is tilted an angle of 23.4 deg with respect to the compass points to increase the area covered for each scan, as was proposed for the Commensal Radio Astronomy FasT survey (CRAFTS) (Li et al. 2018), though we note that this is not the only choice available. The whole 19 beams span 22.8 arcmins across the north–south direction at 1.42 GHz (calculated for the centre of the beam). The sky is covered by shifting the whole array in declination by 21.9 arcmin for the next scan. A drift scan of  $\pm 40$  deg from the centre declination would require about 220 strips (i.e. 220 d) to cover the region once.

#### 2.1.2 Wide-band receiver system

For higher redshift ( $z > 0.35$ ), at present the survey can be done with a single feed wide-band receiver, which covers a frequency from 0.27 GHz to 1.62 GHz. The receiver noise for wide-band receiver system is  $\approx 60$  K, to have the same noise scale, and the survey for redshift larger than 0.35 needs twice more time than the survey for redshift smaller than 0.35. The strip width in this case is  $2.9(1+z)$  arcmin, so to cover the  $\pm 40$  deg sky, it would require 1211 d with single feed strip to cover the same sky region at  $z = 0.35$ , which is much less practical due to the long observation time required.

#### 2.1.3 The PAF receiver system for UHF band

In the future, it is worthwhile to consider equipping the FAST telescope with a multibeam receiver at the lower frequency band for a survey of higher redshifts. A PAF with cryogenic receiver system would allow rapid survey of large areas of sky, and such



**Figure 1.** The top panel shows the strips of noise level produced by one scan of the 19-beam feed. The bottom panel shows the noise level from multiple scans with vertical intervals of 21.9 arcmin. The width of strip is selected to be 2.3 arcmin, which is the vertical intervals of two nearest dashed line in the top panel.

development has also been pursued for FAST (Wu et al. 2016). Here, as an illustrative example, we consider a low-frequency PAF system with 81 effective beams, 500-MHz bandwidth centred at 0.75 GHz (i.e. 0.5–1.0 GHz), a system noise about 30 K, and an aperture efficiency around 70 per cent. These beams could be positioned in square array and then a pixel in the sky will be scanned nine times if drift along a side of the square. In this way, a full drift scan of  $\pm 40$  deg could be finished in 135 d, with integration time of 291s on each pixel.

## 2.2 Integration time and noise

The sky drifts across with a speed of  $\omega_e \cos \delta$  in a drift scan survey, where  $\omega_e \approx 0.25$  arcmin/s is the angular velocity of the rotation of the Earth, and  $\delta$  is the declination of the pointing. The time for drifting across a pixel is given by

$$t_{\text{pix}} = 2.9(1+z) \text{ arcmin}/(\omega_e \cos \delta). \quad (2)$$

One circle is completed in a sidereal day, though in practice the night time data are usually of much smaller noise than the day time data. At  $z = 0$ , and  $\cos \delta \approx 1/2$  (near the zenith of FAST site), we get 24 s per beam. And because of the overlap of 19 beams in one horizontal scanning (see Fig. 1), most pixels in a 19-beam strip will be scanned twice, resulting in 48 s per beam. Within the observable part of sky, the circles with higher declination (northern part of sky) have smaller area, while the integration time per pixel is larger. The

expected thermal noise for a dual polarization single beam is

$$\sigma_{\text{noise}} = \sqrt{2} \frac{k_B T_{\text{sys}}}{A_{\text{eff}}} \frac{1}{\sqrt{\Delta \nu t}}, \quad (3)$$

where  $t$  is the total integration time,  $\Delta \nu$  is the frequency bandwidth for a channel, and  $k_B$  the Boltzmann constant. The aperture efficiency is about 70 per cent, giving effective aperture  $A_{\text{eff}} \approx 50\,000$  m<sup>2</sup>. The system temperature is  $T_{\text{sys}} = T_{\text{rec}} + T_{\text{sky}}$ , where  $T_{\text{rec}}$  is the receiver temperature and taken to be 20 K, and away from the Galactic plane, the sky temperature is modelled as

$$T_{\text{sky}} = 2.73 + 25.2 \times (0.408/\nu_{\text{GHz}})^{2.75} \text{ K}. \quad (4)$$

If we assume a velocity line width of  $\approx 5$  km s<sup>-1</sup> for the spectral line observation in HI galaxy, and 48 s integration time per beam, the instantaneous sensitivity of each beam of the FAST system will be 0.86 mJy. Below, we shall consider surveys with an average of 48 s, 96 s, 192 s, and 384 s integration time per beam, according to once, twice, three times, and four times repeat observation, respectively.

In the case of 19-beam L-band feed, approximately every pixel would be covered by several beams, effectively double the integration time. In a more careful treatment, we may estimate the noise as follows. The time stream data are related to the signal by

$$d = \mathbf{A}s + \mathbf{n}, \quad (5)$$

where the time-ordered data vector  $d$  has a dimension of  $19N_t$  where  $N_t$  is the length of the time-ordered data, the sky pixel vector  $s$  has a dimension  $N_{\text{pix}}$ , and the pointing matrix  $\mathbf{A}$  has a dimension of  $19N_t \times N_{\text{pix}}$ . The minimum variance estimator for the sky is

$$\hat{\mathbf{s}} = (\mathbf{A}^t \mathbf{N}^{-1} \mathbf{A})^{-1} \mathbf{A}^t \mathbf{N}^{-1} \mathbf{s}, \quad (6)$$

where  $\mathbf{N}$  is the covariance matrix of the noise in the time-ordered data. The sky map noise covariance matrix is then

$$\mathbf{C}_{\mathbf{N}} = (\mathbf{A}^t \mathbf{N}^{-1} \mathbf{A})^{-1}. \quad (7)$$

Using this expression, we can estimate the map noise.

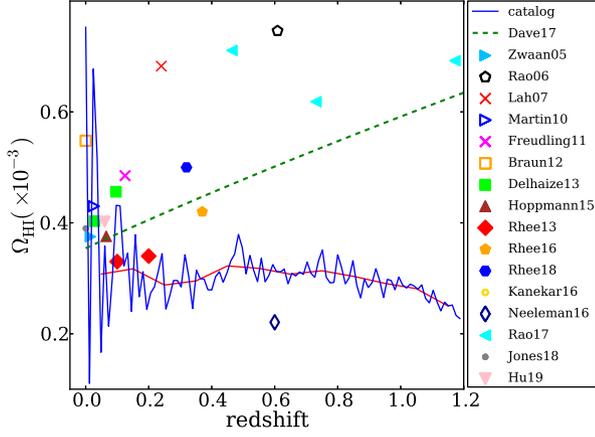
In Fig. 1, we show the estimated noise of the sky map obtained by the 19-beam receiver in units of single beam receiver. For simplicity, we assumed that the beams are identical and have a Gaussian beam within the beamwidth, though in reality there is much difference in the central and outer beams. Also, we assumed a constant system temperature, though actually the system temperature varies, as the sky temperature varies.

As one might expect, in a scan along the horizontal direction of the 19-beam receiver, pixels that are near the centre of the receivers will be scanned by more than one beam, resulting in a lower noise than others. Such inhomogeneous noise distribution is undesirable, because in the LSS measurement, it may bias the observation and induce superfluous structures. The bias may be approximately corrected by introducing selection functions, but as the real noise is varying and not accurately known, precision is hard to achieve. To reduce such effects, we need to have a relatively uniform distribution of noise in the survey regions. From Fig. 1, we see that by partially overlapping the scanning strips (with a vertical intervals of 21.9 arcmin), a large part of this inhomogeneity could be removed, making it a nearly uniform survey in the central part.

## 3 SIMULATION

### 3.1 The galaxy model

We used the catalogue from the Semi-Analytic Suite of the SKA Simulated Skies ( $S^3$ -SAX), in which the cosmic evolution of the



**Figure 2.** The  $\Omega_{\text{HI}}(z)$  from the mock catalogue (blue line and red line for the smoothed result) and the various observations. The dashed green line represents results in the MUFASA cosmological hydrodynamical simulation from:  $\Omega_{\text{HI}} = 10^{-3.45} \times (1+z)^{0.74}$ .

galaxies is tracked by semi-analytic models (De Lucia & Blaizot 2007) based on the millennium  $N$ -body simulation (Springel et al. 2005), and the amount of neutral atomic hydrogen (HI) and molecular ( $\text{H}_2$ ) hydrogen in galaxies is computed with the semi-analytical model (Obreschkow et al. 2009a,b,c). An easy-to-use mock catalogue (Obreschkow & Meyer 2014) of galaxies with detailed physical properties (position in the sky, apparent redshift, stellar mass, HI mass, effective radius, etc.) is available. The catalogue is for a cone with a field of 10-by-10 degrees and a redshift range of 0.0–1.2. It is complete down to an HI mass of  $10^8 M_\odot$ . A deficiency of this model is that this HI cutoff mass is still relatively high, which could miss a significant amount of HI in dwarf galaxies. This limit is only a minor concern when dealing with isolated direct HI detections in blind surveys, because only a tiny fraction of the total survey volume is sensitive to HI masses  $< 10^8 M_\odot$ . However, when dealing with global HI mass estimates, e.g. the intensity mapping experiment, the HI mass contained in unresolved galaxies is non-negligible. We compute  $\Omega_{\text{HI}}(z)$  from all the galaxies in the mock catalogue and compare it with the observations (Zwaan et al. 2005; Rao, Turnshek & Nestor 2006; Lah et al. 2007; Martin et al. 2010; Freudling et al. 2011; Braun 2012; Delhaize et al. 2013; Hoppmann et al. 2015; Kanekar, Sethi & Dwarakanath 2016; Neeleman et al. 2016; Rao et al. 2017; Jones et al. 2018; Rhee et al. 2013, 2016, 2018; Hu et al. 2019); the result is shown in Fig. 2. There are still quite large scatters and discrepancy in the result, but already we can see that the  $S^3$ -SAX simulation may have underestimated the amount of HI by a factor of between 1.2 and 2.0, especially at higher redshifts. The more recent MUFASA cosmological hydrodynamical simulation (Davé et al. 2017) is in better agreement with the observations. We have done most of our galaxy survey simulation with the  $S^3$ -SAX mock catalogue and for the computation of the intensity mapping, we scale all HI flux (equivalent to HI mass) with a  $z$ -dependent factor to compensate the lost HI mass.

For each galaxy in the catalogue, the HI distribution and 21-cm emission are modelled following Elson, Blyth & Baker (2016). We generate a mini data cube for each galaxy and then we re-grid the mini data cube to a full-sized cube that contains all the galaxies.

The HI mass density distribution of the galaxy is modelled as a thin axisymmetric exponential model:

$$\Sigma_{\text{HI}}(r) = \frac{\tilde{\Sigma}_{\text{H}} e^{-r/r_{\text{disk}}}}{1 + R_{\text{mol}}^c e^{-1.6r/r_{\text{disk}}}}, \quad (8)$$

where  $r$  denotes the galactocentric radius,  $r_{\text{disk}}$  refers to the scale length,  $\Sigma_{\text{HI}}(r)$  is the surface density of the total hydrogen component,  $\Sigma_{\text{H}} = M_{\text{H}}/(2\pi r_{\text{disk}}^2)$  is a normalization factor, and  $R_{\text{mol}}^c$  denotes the  $\text{H}_2/\text{HI}$  mass ratio at the galaxy centre. This is based on a list of empirically supported assumptions: (i) the cold gas of regular galaxies resides in a flat disc [see Leroy et al. (2008) for local spiral galaxies, Young (2002) for local elliptical galaxies, and Tacconi et al. (2006) for galaxies at higher redshifts]; (ii) the surface density of the total hydrogen component (HI + H<sub>2</sub>) is well described by an axisymmetric exponential profile [Leroy et al. (2008)]; and (iii) the local H<sub>2</sub>/HI mass ratio scales as a power of the gas pressure of the ISM outside molecular clouds [Blitz & Rosolowsky (2006)]. Recent observation [Serra et al. (2012); Nyland et al. (2017)] for early-type galaxies reveals that for most of the elliptical galaxies, the HI gas has a morphology that is similar in appearance to the discs of radio emission associated with SF in spiral galaxies. Most of the HI detection exhibits a large, settled HI disc or ring. The orientation of the galaxy is randomly chosen. In reality, the galaxy may have some correlation of intrinsic alignment in its orientation, but such alignment is generally a second-order effect and does not significantly affect the analysis given below. When generating a galaxy model with the parameters from the catalogue, we convert the apparent HI half-mass radius,  $R_{\text{HI}}^{\text{half}}$ , along the major axis into an exponential disc scale length  $r_{\text{disk}}$ . The galaxy is modelled out to a radius of  $3.5 r_{\text{disk}}$ . With the surface density deduced, we then convert it to the mass distribution. The circular velocity profile of the galaxy is modelled with the Polyex analytic function (Giovannelli & Haynes 2002):

$$V_{\text{PE}}(r) = V_0(1 - e^{-r/r_{\text{PE}}}) \left( 1 + \frac{\alpha r}{r_{\text{PE}}} \right), \quad (9)$$

where  $V_0$ ,  $r_{\text{PE}}$ , and  $\alpha$  determine the amplitude, exponential scale of the inner region, and the slope of the outer part of the rotation curve, respectively. These parameters are derived from the luminosity of the galaxy given in the semi-analytical model, using the empirical relations derived from nearly 2200 low redshift disc galaxies (Catinella, Giovanelli & Haynes 2006). The semi-analytical model gives R-band luminosity, while the Catinella et al. (2006) model used I-band luminosity, so we convert them by  $M_{\text{I}} = M_{\text{R}} - 0.37$  (Duffy et al. 2012).

The HI flux density is then given by

$$\frac{M_{\text{HI}}}{M_\odot} = 2.36 \times 10^5 \left( \frac{D_{\text{L}}}{\text{Mpc}} \right)^2 \frac{S_{\text{i}}}{\text{Jy km}} \frac{dv}{\text{s}^{-1}} (1+z)^{-2}, \quad (10)$$

where  $S_{\text{i}}$  is flux density in units of Jy in channel  $i$  of the mini-cube,  $dv$  is the velocity width of a channel in  $\text{km s}^{-1}$ ,  $D_{\text{L}}$  is the luminosity distance of the target galaxy in Mpc units, and  $z$  is its evaluated redshift. Note that  $dv$  here is defined in intrinsic velocity bin; if it is for the observed velocity bin,  $(1+z)^{-2}$  will be replaced by  $(1+z)^{-1}$ . For each galaxy, the data cube has  $100 \times 100$  angular pixels and  $5 \text{ km s}^{-1}$  in velocity channel width. Each voxel (volume pixel) has its HI flux density and velocity that computed, and we reposition it in a 3D data cube.

We use a light cone catalogue from the Obreschkow & Meyer (2014) simulation, which spans a field of  $10 \times 10 \text{ deg}^2$  on the sky and a redshift range of 0.0–1.2. This volume contains 19 210 309 galaxies with a total HI mass of  $2.065 \times 10^{16} M_\odot$ . We then re-

grid the mini data cubes into the full-sized synthetic cube and place it in the corresponding angular position and frequency. The final full-sized synthetic cube has a pixel width of 0.0133 deg and a fixed channel width of 0.0237 MHz, corresponding to 5 km s<sup>-1</sup> at the redshift of  $z = 0$ . To simulate the sky observed by FAST, we convolve each channel of the synthetic data cube with a redshift-dependent Gaussian point spread function, with beamwidth proportional to  $1 + z$ .

### 3.2 HI galaxy detection

To simulate the detection of HI galaxies, we first re-bin the full-sized synthetic data in RA and DEC to an angular resolution of 0.08 deg, corresponding to two times Full width at half maximum (FWHM) of the FAST beam. If we assume that the noise for each beam is Gaussian, the noise for each pixel is rescaled as

$$\sigma_{\text{noise}}^{\text{pixel}} = \sigma_{\text{noise}}^{\text{beam}} \times \sqrt{\frac{A_{\text{pixel}}}{A_{\text{beam}}}}, \quad (11)$$

where  $A_{\text{pixel}}$  is the sky area of a pixel and  $A_{\text{beam}}$  is the sky area of the beam.

In the frequency axis, the synthetic data cube is re-binned to a resolution of 0.1 MHz, corresponding to a velocity width of 20 km s<sup>-1</sup> at  $z = 0$ . Because the flux and noise scale will change with the bandwidth, if the full-sized synthetic data are smoothed to a velocity width of  $W_s$  km s<sup>-1</sup>, the signal to noise from one velocity bin is scaled as [Saintonge (2007)]

$$S/N = \frac{(F_0/N)/W_s}{\sigma_0} \times \left(\frac{W_s}{5 \text{ km s}^{-1}}\right)^{1/2}, \quad (12)$$

where  $F_0$  is the total velocity integrated HI flux of a galaxy,  $N$  is the number of the velocity bins the galaxy spans, and  $\sigma_0$  is the thermal noise with a velocity width of 5 km s<sup>-1</sup>.

With the above mock data, we may simulate galaxy detection as follows:

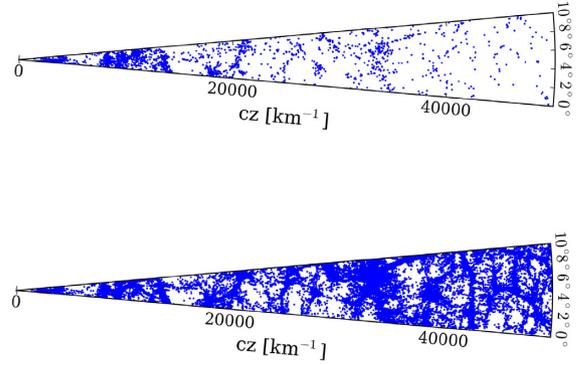
(i) *Coarse resolution search.* Re-bin the noise-filled data to an angular resolution of 0.08 deg (two times of FWHM of FAST) and a frequency resolution of 0.473 MHz (corresponding to velocity resolution of 100 km s<sup>-1</sup> at redshift 0), setting a threshold of  $3\sigma$ , and detect voxels above the threshold.

(ii) *Fine resolution fit.* For galaxies detected in the coarse search, use a finer frequency resolution (0.0236 MHz, corresponding to velocity resolution of 5 km s<sup>-1</sup> at redshift 0) to fit its spectrum in the data cube with a parametrized profile function. If a reasonable HI profile is obtained, we integrate the HI profile and the candidate is selected as a galaxy if the total flux exceeds  $5\sigma$ . About 20 per cent of the candidates found in the first step passed the second step, and other ones might be large noise.

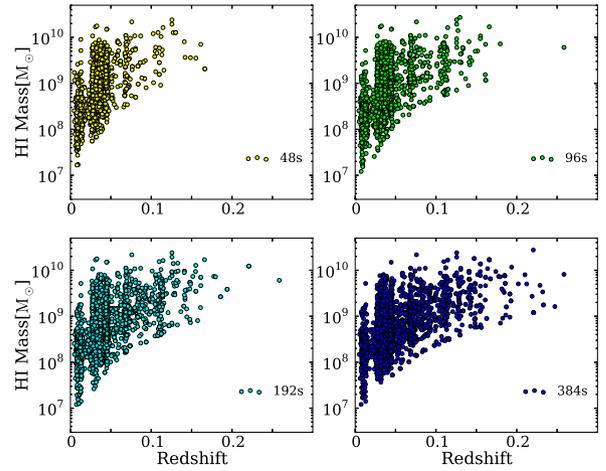
For the HI profile, we use the so-called ‘busy function’ proposed by Westmeier et al. (2014), which has great flexibility in fitting a wide range of HI profiles from the Gaussian profiles of dwarf galaxies to the broad, asymmetric double-horn profiles of spiral galaxies.

### 3.3 Galaxy distribution

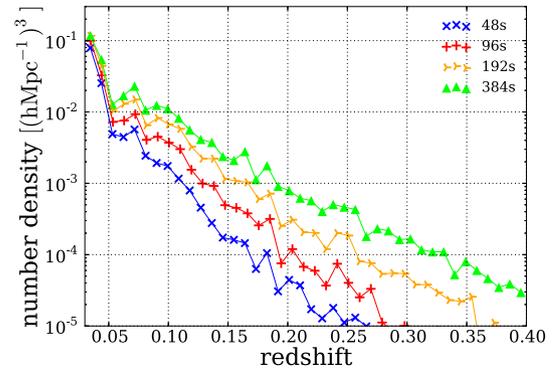
To show the capacity of surveys with different integration time, we have run our selection pipeline with four different  $\sigma_0$ : 0.86 mJy, 0.61 mJy, 0.43 mJy, and 0.31 mJy, corresponding to 48 s, 96 s, 192 s, and 384 s integration time per beam.



**Figure 3.** The top panel is the 10 Mpc/h slab of the galaxy candidates from the FAST 2-yr (or 192-s integration time) galaxy survey; the bottom panel is the distribution of all galaxies in the catalogue.

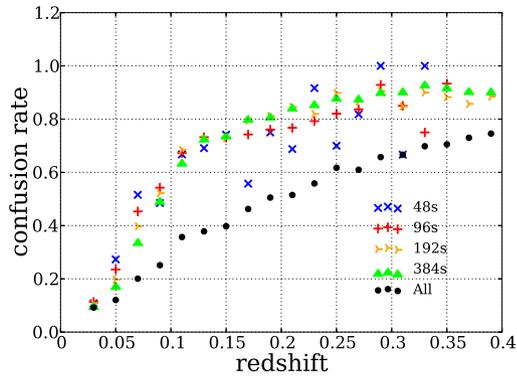


**Figure 4.** The distribution of the HI mass of the galaxy candidates from the  $15 \times 15 \times 600 \text{ h}^{-3} \text{ Mpc}^3$  comoving volume.



**Figure 5.** The comoving number density computed from the detected catalogue.

We plot the 10 Mpc/h slab of the mock galaxies detected in a 2-yr (or 192-s integration time) survey in Fig. 3; all galaxies in the simulated are also plotted for comparison. The masses and redshift of these detected galaxies are plotted in Fig. 4, and in Fig. 5, we show the number density of the detected galaxies. We can see from these figures that as redshift increases, the galaxies thinned out, and the number density decreases drastically as redshift increases. At  $z = 0.2$ , the number density already falls off by two



**Figure 6.** The confusion rate of the detected galaxies from different surveys. The minimum H I mass is set as  $10^{8.5}M_{\odot}$ .

or three orders of magnitude compared with the nearby galaxies, and only the most massive galaxies can be detected at the higher redshifts. Indeed, existing H I galaxy surveys are all limited to  $z < 0.2$ , and even FAST, the largest single dish telescope in the world, could not detect many galaxies in this mode. In the case of integration time of 48 s, the galaxy number density would drop to about  $10^{-4}(\text{Mpc}/h)^{-3}$  at  $z = 0.15$ , corresponding to roughly the galaxy number density required for BAO measurement. For the longer integration time, the distribution of the detectable galaxies extends to higher redshift. But even for the 384-s survey, the number density drops to  $10^{-4}(\text{Mpc}/h)^{-3}$  at  $z = 0.4$ .

For galaxy survey, galaxies may overlap with each other along the line of sight. The ability to uniquely identify the individual galaxy is an important evaluation of the performance of a telescope. To quantify this, we introduce the confusion rate, which is defined as the fraction of galaxies fall within the same voxel. The voxel has a length of 0.08 deg a side, which is two times of the FWHM of the FAST beam at  $z = 0$ , and a bandwidth of 1 MHz in frequency axis. We have set a limit of H I mass of  $10^{8.5}M_{\odot}$  in the mock catalogue to remove the effect induced by low-mass galaxies. For the galaxy detection algorithm we used, the confusion rate is shown in Fig. 6. As the redshift increases, the confusion rate rises rapidly. The integration time has a mild influence on the confusion rate, as the smaller galaxies are detected for longer integration time, which leads to an increase in the number of galaxies per beam and hence a higher confusion rate. This steady raise of the confusion rate shows how the galaxy count transits smoothly and naturally to 21-cm intensity mapping with the increase of redshift in the FAST observation.

For the detected galaxies, due to the finite beam size and the noise, there will also be errors in the measured positions. To quantify this effect, we compare the measured positions in the mock observation with the original position in the catalogue. The measured central position is taken to be the flux weighted average,

$$v_c = \frac{\int s(v)v dv}{\int s(v)dv}, \quad \theta_c = \frac{\int s_i \theta_i d\theta}{\int s_i d\theta}. \quad (13)$$

The frequency integration range covers about  $\pm 200 \text{ km s}^{-1}$ , with resolution of 0.0236 MHz. We sum the flux from the sub-data cube along the frequency axis, and the angular beam average is computed on a grid with spacing of 0.0133 deg and beamwidth of 0.08 deg.

The error of the position measurement in comoving coordinates is shown in Fig. 7. We see that the noise in the galaxy H I profile can induce a shift in its position. For the plotted redshift range ( $z < 0.3$ ), the shift in the direction perpendicular to the line of sight

(los) is generally smaller than that along the los direction. About 95 per cent galaxies have position shifts in the perpendicular direction lower than 0.1 Mpc/h and about 90 per cent galaxies have shift in parallel direction lower than 0.5 Mpc/h. We coloured the galaxies with their velocity-integrated flux. It shows that the galaxies with low velocity-integrated flux tend to have large shift. There are more galaxies that have large shift in survey with higher integration time, because more galaxies with low velocity-integrated flux can be detected with longer integration time. We also show the standard deviation of the position shift with blue vertical lines at different redshifts in Fig. 7. Higher redshift has larger shift because there are more galaxies with lower velocity-integrated flux.

### 3.4 H I intensity map

In Fig. 8, we show slices of H I intensity maps at six different redshifts in the range  $0 < z < 0.3$ ; the depth is 1 MHz for each slice. We also added a thermal noise corresponding to 1 MHz of channel bandwidth and 192 s of integration time. To obtain the noise-filled maps, we construct the flux cube and noise cube with a bandwidth of 1 MHz, respectively, and then combine them together. As we have mentioned above, missing the galaxies with H I mass below  $10^8 M_{\odot}$  will have a significant influence in intensity mapping experiment. We correct this by scaling all H I masses in the mock catalogue with a  $z$ -dependent factor ( $> 1$ ) to match  $\Omega_{\text{H I}}$  inferred from MUFASA cosmological hydrodynamical simulation. As can be seen from this figure, even for the relatively long integration time of 192 s, the map is still largely dominated by noise, showing the challenges one will face in H I surveys.

In Fig. 9, we show the projected H I intensity maps at  $0 < z < 0.8$ , each with a redshift interval of 0.1. The H I distribution is shown more clearly in these maps without the noise. For slices nearby, one can clearly see individual galaxies. As the distance increases, the structures become more blurred and also the intensity drops. Individual galaxies become increasingly difficult to see, but the overall structure remains, which illustrates how the intensity mapping could be used to probe the LSS. Note that the equal spacing in redshift means slightly smaller comoving distance spacing at higher redshift, but we have checked and found that the equal spacing in comoving distance generates maps pretty similar to these.

## 4 POWER SPECTRUM MEASUREMENT

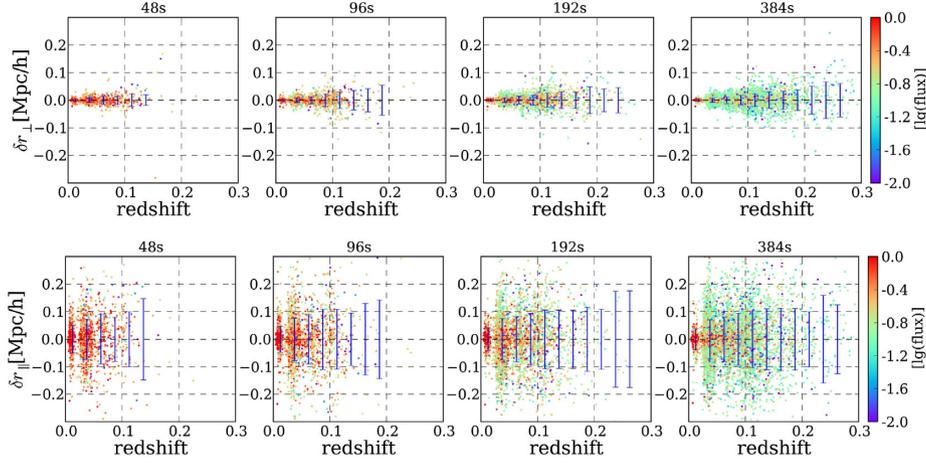
The power spectrum is the most widely used statistics for LSS. In this section, we describe its measurement and error forecast for both the H I galaxy survey and the H I intensity mapping survey with FAST.

### 4.1 H I galaxy power spectrum

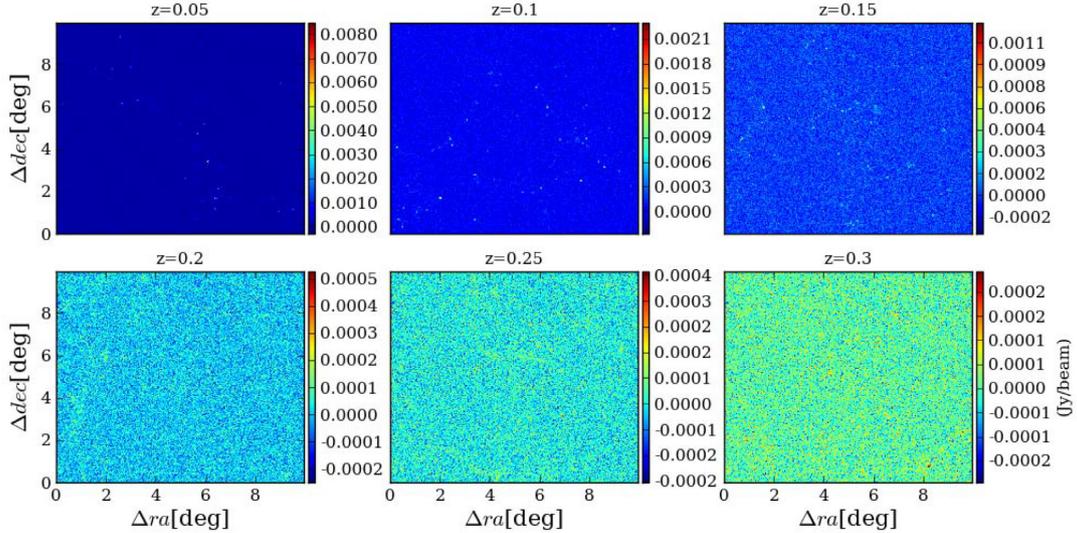
In a galaxy redshift survey with negligible error on the position of galaxies, the measurement error of the power spectrum comes from sample variance as well as shot noise. Over a  $k$  bin of width of  $\Delta k$  (Feldman, Kaiser & Peacock 1994; Duffy et al. 2008),

$$\frac{\sigma_P}{P} = \sqrt{2 \frac{(2\pi)^3}{V_{\text{eff}}} \frac{1}{4\pi k^2 \Delta k} \frac{P(k) + 1/n}{P(k)}}, \quad (14)$$

where  $V_{\text{eff}}(k) = \int [\frac{n(\vec{r})P(k)}{n(\vec{r})P(k)+1}]^2 d^3\vec{r}$  for detected galaxies  $n(r)$ . The error on galaxy position may also induce slight errors, but for the scale of interest, they are negligible.



**Figure 7.** The position error induced by thermal noise in the observation for different integration times. Top panels: the error perpendicular to the los. Bottom panels: the error along the los. The galaxies are coloured according to their integrated flux.



**Figure 8.** Noise-filled H I flux slice from different redshifts; all have the same frequency interval 1 MHz. The galaxy flux and noise both are computed with a bandwidth of 1 MHz and a integration time of 192 s per beam.

An optimal weighted estimator (known as the Feldman–Kaiser–Peacock or FKP estimator) may be formed to minimize the measurement error of the power spectrum (Feldman et al. 1994). To make the measurement in the irregular geometry of an actual survey, a mock sample of random points is generated. The detected galaxies are re-gridded into a rectangular box, the FKP estimate for the weighted density field is

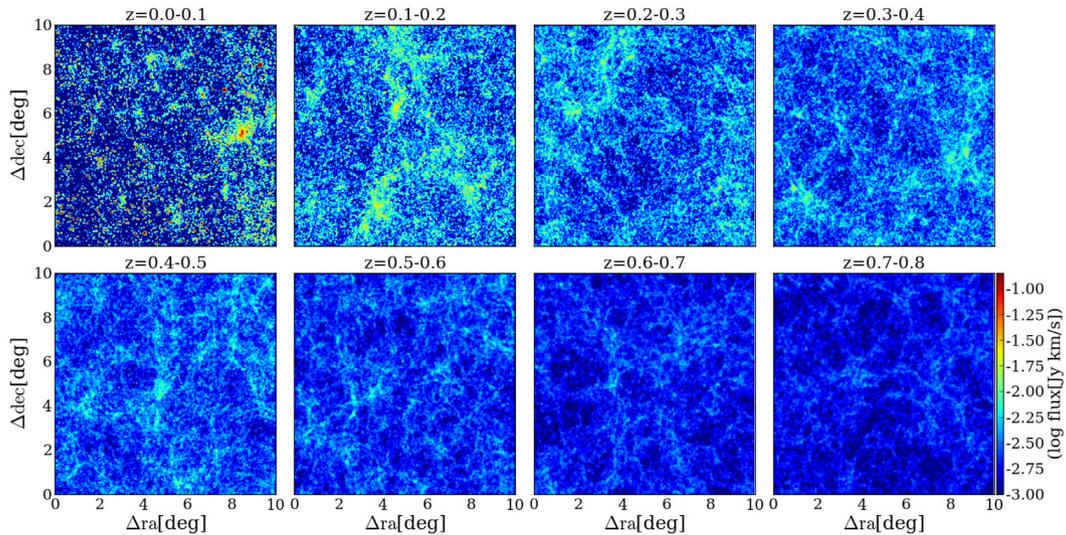
$$F(\mathbf{r}) = \frac{w(\mathbf{r})}{N} [n_g(\mathbf{r}) - \alpha n_s(\mathbf{r})], \quad (15)$$

where  $w(\mathbf{r}) = 1/(1 + \bar{n}(\mathbf{r})P)$  is the FKP weight,  $\bar{n}(\mathbf{r})$  is the selection function (i.e. mean density) at the position  $\mathbf{r}$ ,  $P$  is a prior estimate of the power  $P(k_1)$  at the scale of interest  $k_1$ ,  $n_g(\mathbf{r})$  and  $n_s(\mathbf{r})$  refer to the number density of the observed galaxy catalogue and the random mock catalogue, respectively, and  $\alpha$  is the real-to-mock ratio:  $\alpha = \sum w(\mathbf{r}_{\text{real}}) / \sum w(\mathbf{r}_{\text{random}})$ . The normalization factor  $N$  in equation (15) is given by

$$N^2 = \int d^3r \bar{n}^2(\mathbf{r}) w^2(\mathbf{r}) = \beta \sum_{\text{random}} \bar{n}(\mathbf{r}_i) w^2(\mathbf{r}_i), \quad (16)$$

where  $\beta$  is the unweighted ratio of number of galaxies in the real (in our case simulation) catalogue to that in the random mock catalogue. In order to reduce the shot noise, the mock catalogue is always set to contain much more galaxies than the real catalogue. Here, we choose  $\beta$  to be 0.02. The power spectrum can then be estimated from the Fourier Transform of the weighted overdensity field ( $|F(\mathbf{k})|^2$ ). The measurement error on power spectrum is estimated with  $\Delta k/k = 0.125$ , for which the error of different  $k$  bins can be regarded as uncorrelated (Goldberg & Strauss 1998; Meiksin, White & Peacock 1999).

Our simulation data cube is in a pencil-beam shape and the amplitude and shape of the power spectrum estimated from the cubic grid are biased,  $\langle |F(\mathbf{k})|^2 \rangle = W(k) \hat{P}(k)$ . To correct for this effect, we compute the window function for this survey geometry by producing two sets of random catalogues, one is distributed only in the pencil-beam region and the other one in a cube region that encloses the pencil-beam region, and the window function is then the ratio of the two power spectrum  $P_{\text{pencil}}(k)/P_{\text{cube}}(k)$ . We produce 10 pairs of samples and use the mean value to make the estimate. The true power spectrum is then obtained by dividing the window function.



**Figure 9.** Velocity-integrated H I flux map from different redshifts with  $\Delta z = 0.1$ .

In summary, the power spectrum is obtained with the following steps:

- (i) Compute the selection function  $\bar{n}(\mathbf{r})$  in each redshift bin.
- (ii) Produce the weighted overdensity field  $F(\mathbf{r})$  in the gridded box with grid spacing 1.0 Mpc/h using the Nearest Grid Point assignment technique.
- (iii) Fourier transform of the weighted overdensity field and compute the power spectrum  $\langle |F(\mathbf{k})|^2 \rangle$  with  $\Delta k/k = 0.125$ .
- (iv) Correct the shape effect using the final window function,  $\hat{P}(k) = \langle |F(\mathbf{k})|^2 \rangle / W(k)$ .

The L-band receiver can cover a redshift range of up to  $z = 0.35$ . In Fig. 10, we show the projected error on power spectrum from a 20 000-deg<sup>2</sup> galaxy surveys for redshift  $0.05 < z < 0.15$ ,  $0.15 < z < 0.25$ , and  $0.25 < z < 0.35$ , with integration time 48 s, 96 s, 192 s, and 384 s per beam. On larger scales, the measurement precision is limited by the available number of modes (cosmic variance), while on the smaller scale it is limited by the available number of galaxies per cell (shot noise). The best relative error in the power spectrum is achieved somewhere at  $10^{-2} \text{ h/Mpc} < k < 10^0 \text{ h/Mpc}$ . At higher redshift, the optimal point shifts towards larger scales (smaller  $k$ ), as the probed volume increases and the observed galaxy number density decreases. At the BAO scale  $k \approx 0.07 \text{ h/Mpc}$ , the signal-to-noise ratio can reach 5.0 at  $z \approx 0.2, 0.25, 0.3$ , and  $0.35$ , respectively. So in our Fisher estimation, we will use these redshift regions as the survey volume of the FAST H I galaxy survey. As can be seen from the figure, at the lower redshift, and especially on small  $k$  values, the difference in the integration time per beam does not make much difference; however, for the high redshift and large  $k$  values, the increase of integration time can significantly improve the measurement precision.

The power spectra from the simulated observations are shown in Fig. 11. From top to bottom panels are the result with 48 s, 96 s, 192 s, and 384 s, respectively, each in the redshift range of 0.05–0.15 (left), 0.15–0.25 (middle), and 0.25–0.35 (right). In each subfigure, the theoretically expected power spectrum, the mock observation power spectrum, and the projected measurement error are plotted. For the noise, we also show separately the sample variance (corresponding to the size of the simulation box) and the shot noise components. In this case, the cosmic variance is much

larger as the size of simulation box spans only a region of  $10 \times 10 \text{ deg}^2$ .

In this simulated galaxy survey, in almost all cases the measured power spectrum is larger than the theoretical power spectrum, as can be seen from Fig. 11. However, it does agree well with the projection. On small scales, the shot noise dominated, which results in a nearly flat spectrum. This is particularly obvious at higher redshifts and for shorter integration times, where the number density of the detected galaxies is too small. On larger scales, at lower redshifts and longer integration times, the shape of the power spectrum is more similar to the theoretical power spectrum, but there is still significant difference, and there are some large fluctuations at the large scales, thanks to the contribution of the cosmic variance. The overall normalization of the power is higher than the matter power spectrum (marked as theoretical), due to the fact that only the rare massive galaxies can be detected, and a clustering bias is introduced. The projected errors get larger in higher redshifts. However, we note that for a real survey with large sky area, the cosmic variance can be significantly reduced.

#### 4.2 Intensity map power spectrum

For intensity mapping survey, the measurement error of the power spectrum can be written as (Seo et al. 2010)

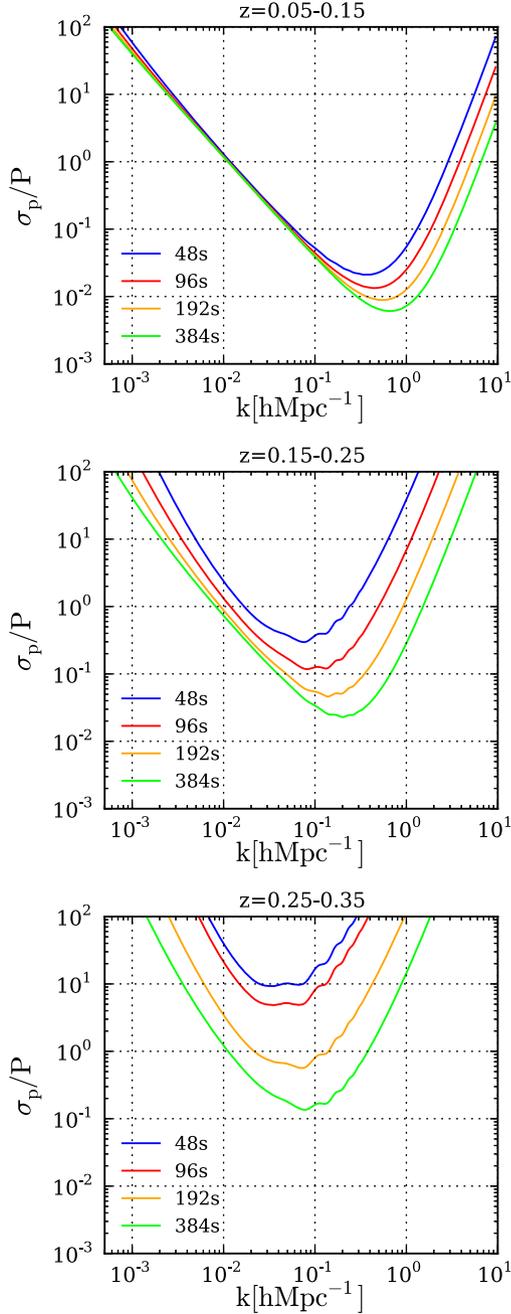
$$\frac{\sigma_P}{P} = 2\pi \sqrt{\frac{1}{V_{\text{eff}}(k)k^2 \Delta k}}, \quad (17)$$

with the  $V_{\text{eff}}$  given in this case by

$$V_{\text{eff}}(\vec{k}) = V_{\text{sur}} \left( 1 + \frac{\sigma_{\text{pix}}^2 V_{\text{pix}}}{[\bar{S}(z)]W(\vec{k})^2 P} + \frac{1/\bar{n}}{P} \right)^{-2}, \quad (18)$$

where  $\bar{S}(z)$  is the average 21-cm emission flux density and  $V_{\text{pix}}$  is the pixel volume. The first term is due to sample variance, the second term is induced by the system thermal noise, and the last term is the shot noise due to the discreteness of the H I sources, with  $1/\bar{n} \approx 100 \text{ h}^{-3} \text{ Mpc}^3$  [Seo et al. (2010)]. We model the angular resolution (the frequency resolution is much higher) as:

$$W(k) = \exp \left[ -\frac{1}{2} k^2 r(z)^2 \left( \frac{\theta_{\text{pix}}(z)}{2\sqrt{2 \ln 2}} \right)^2 \right]. \quad (19)$$



**Figure 10.** The projected error on power spectrum from a 20000 deg<sup>2</sup> galaxy survey with integration time of 48 s, 96 s, 192 s, and 384 s per beam. At  $k \approx 0.07$  h/Mpc, the S/N can reach 5.0 at  $z \approx 0.2, 0.25, 0.3,$  and  $0.35,$  respectively.

The image cube is gridded into a rectangular box with grid spacing of 2 Mpc/h; here, we take 2 Mpc/h-sized pixels as the standard sampling size and it is well below the BAO scale, which is about 150 Mpc/h and would not affect the cosmological result. The H I emission flux density field is related to the overdensity field by:

$$\delta(\mathbf{r}) = S_{\text{H I}}(\mathbf{r})/\bar{S}_{\text{H I}} - 1, \quad (20)$$

where  $S_{\text{H I}}(\mathbf{r})$  is the H I emission flux density at position  $\mathbf{r}$  and  $\bar{S}_{\text{H I}}$  is the mean flux density of H I emission. We also add a Gaussian thermal noise of 0.74 mJy, 0.52 mJy, 0.37 mJy, and 0.26 mJy per

beam, corresponding to 48 s, 96 s, 192 s, and 384 s integration time per beam. The simulated power spectrum is measured as follows:

- (i) Generating the mock data filled with Gaussian thermal noise,
- (ii) Re-gridding the image data cube to rectangular box with grid spacing of 2 Mpc/h,
- (iii) Converting the flux density field to H I mass overdensity field and Fourier transform of the field to get the power spectrum  $\langle |F(\mathbf{k})|^2 \rangle$  with  $\Delta k/k = 0.125,$
- (iv) Using the pencil beam survey window function to correct the shape effect.

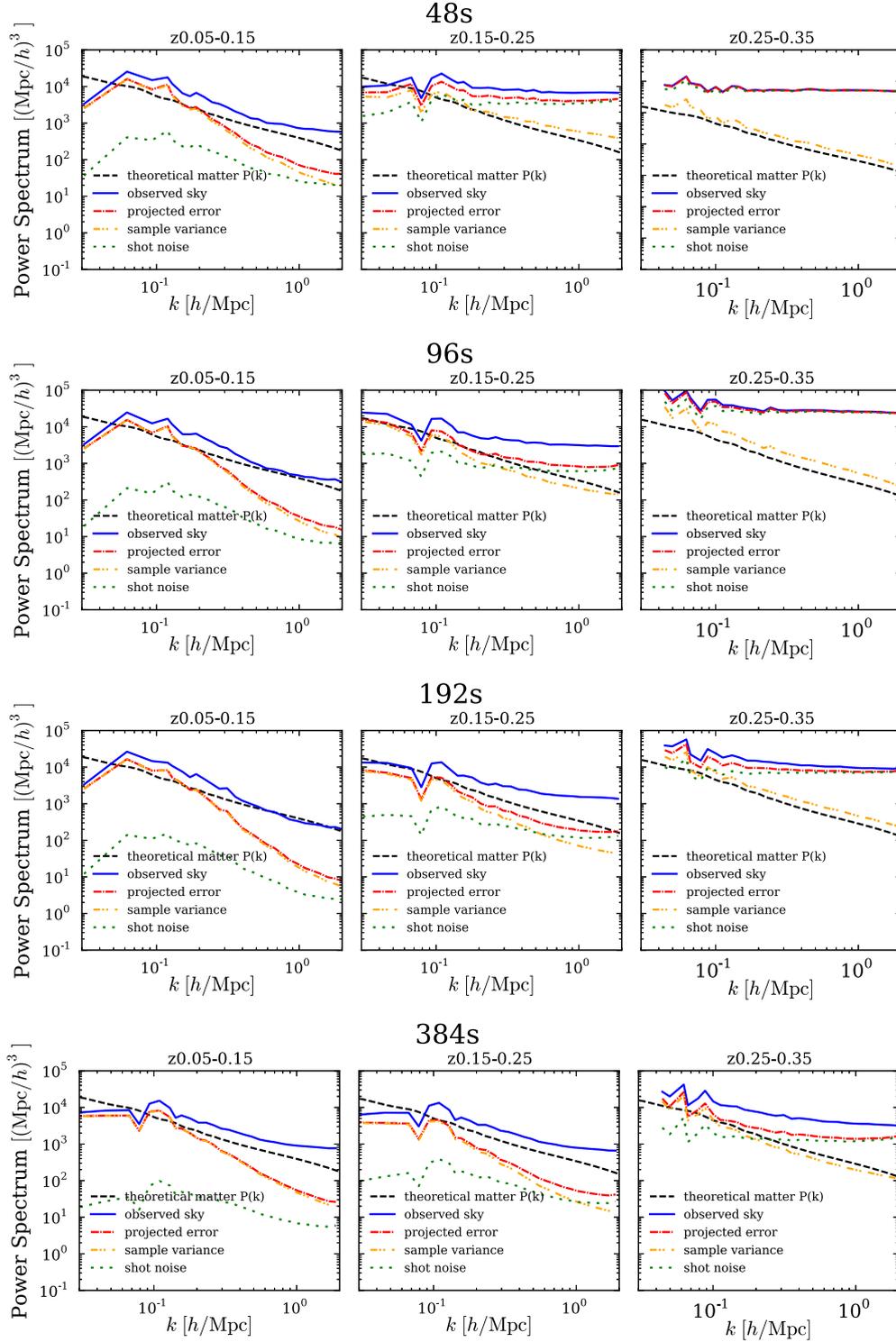
The statistical error is estimated as follows:

- (i) Generating a synthetic product with a spatial resolution same as FAST beam size and a frequency resolution of 1 MHz,  $S(r_x, r_y, z),$
- (ii) Generating a mock data filled with Gaussian thermal noise,  $N(r_x, r_y, z),$  the noise data cube has the same size and resolution as  $S(r_x, r_y, z),$
- (iii) Re-gridding the synthetic data  $S(r_x, r_y, z)$  and the noise data  $N(r_x, r_y, z)$  to a rectangular box with grid spacing of 2 Mpc/h, obtaining  $S(r'_x, r'_y, z)$  and  $N(r'_x, r'_y, z),$
- (iv) Computing the mean signal of  $S(r'_x, r'_y, z)$  and the variance of  $N(r'_x, r'_y, z)$  at different redshift, obtaining  $\bar{S}(z)$  and  $\sigma_{\text{pix}},$
- (v) Obtaining the projected error by use of equation (17).

Fig. 12 shows the statistical error of the power spectrum that can be achieved by 20000 deg<sup>2</sup> FAST intensity mapping survey with integration time 48 s, 96 s, 192 s, and 384 s per beam, respectively. Here, we ignore for the moment the frequency range of the L-band receiver system but simply plot up to much higher redshifts, which may be accomplished with the wide-band single feed receiver, or a future UHF PAF receiver introduced in Section 2.1. For simplicity and easier comparison, here we also plot the projected error for the same integration time, though to the size of the beam and the number of available beams the required total survey time would be very different at the higher redshifts.

A comparison with Fig. 10 shows that for the lowest redshift bin ( $0.05 < z < 0.15$ ), the projected errors are almost the same as the H I galaxy survey of the same integration time. We saw from Fig. 5 that even in this low redshift range, the comoving number density of the detected H I galaxies is decreasing with increasing redshift, showing that the FAST does not detect all H I galaxies. Nevertheless, the detected H I galaxy has sufficiently high number density that it gives a good representation of the underlying H I galaxy sample and total mass density. However, for all the other redshift bins, the projected error of the intensity mapping is much smaller than the H I galaxy survey. In fact, as redshift increases and the signal-to-noise ratio actually improves for a while as the survey volume increases and sample variance decreases, though eventually it begins to drop as the thermal noise becomes higher and H I signal becomes weaker. The shift in the optimal comoving scale (the minimum of the relative error in the power spectrum) at different redshifts is also much less than the H I galaxy surveys. At the BAO scale  $k \approx 0.07$  h/Mpc, the S/N can reach 5.0 until redshift of 0.35, 0.55, 0.75, and 1.05 for survey with 48 s, 96 s, 192 s, and 384 s, respectively.

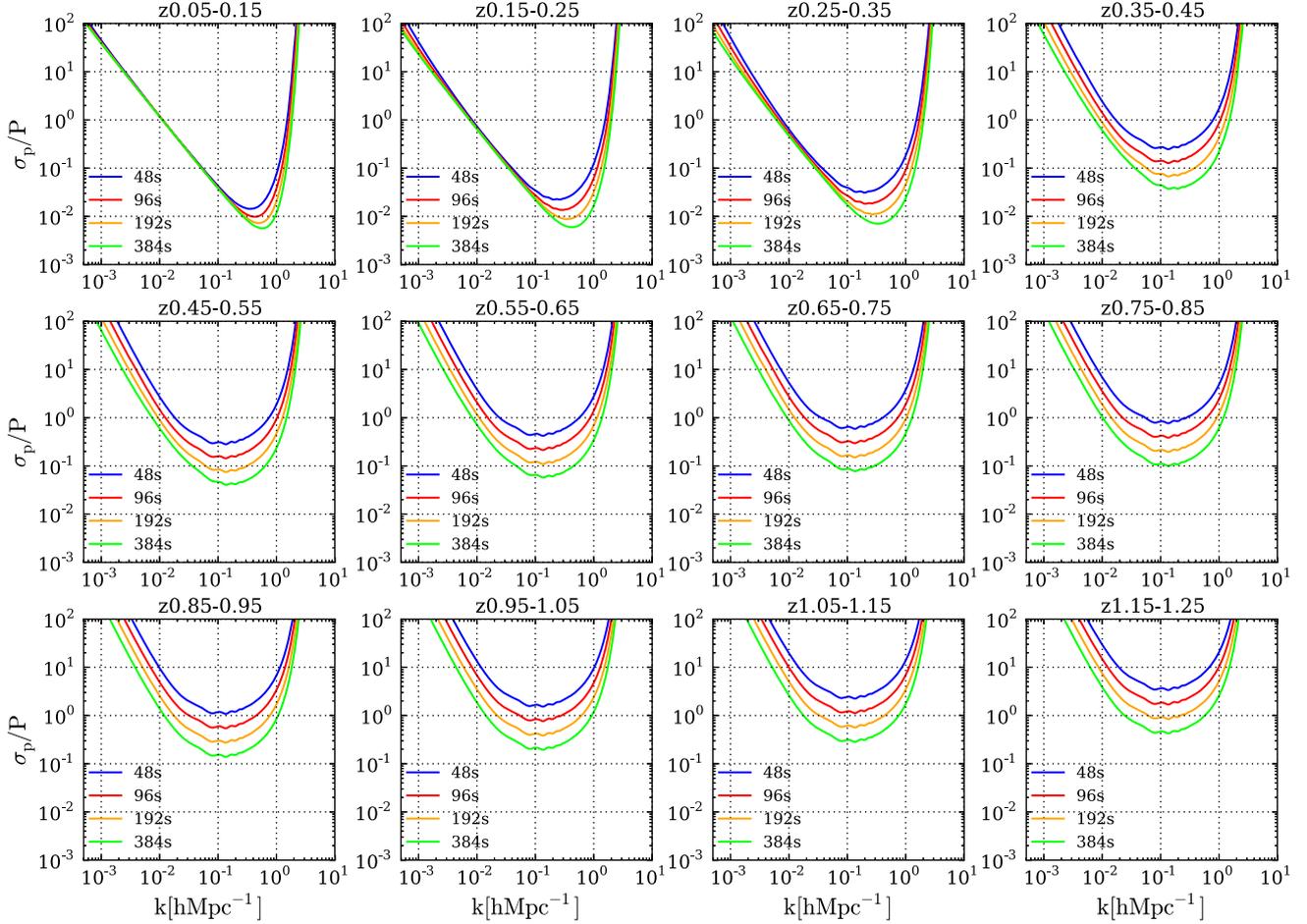
The simulated power spectrum measurement is shown in Fig. 13. The left three columns show the simulation with the same L-band receiver, while the column on the right shows a higher redshift ( $0.35 < z < 0.45$ ) with the wide-band receiver system of a single feed. Compared with Fig. 11, the thermal noise is replaced by the shot noise at the smaller scales. Within the L-band (up to  $z = 0.35$ ), it



**Figure 11.** The HI galaxy power spectrum from simulation. We show the theoretical matter power spectrum (black dashed line), the simulated measurement (blue solid line), the total projected error estimated with Fisher matrix (red dash-dotted line), the error due to sample variance (orange dash-dot-dotted line), and shot noise (green dotted line).

is sub-dominant even for the shortest integration time considered here (48 s). So, the intensity mapping can yield much nicer results than galaxy surveys on the small scales, and good precision can be achieved at higher redshifts. On the larger scales where the cosmic variance dominates, the precision of this simulated measurement is limited by the size of the simulation box, so we see some

large fluctuations, but this can be significantly reduced with larger volumes. The intensity mapping survey can efficiently map the LSS of BAO scale until  $z \approx 0.35$  even with moderate integration time. Note that the error from intensity mapping is smaller at most of scales, but at small scales ( $k \approx 1.0 h/\text{Mpc}$ ), the noise is larger than HI signal at all redshifts due to the limited resolution.



**Figure 12.** The projected error on power spectrum from FAST 20000 deg<sup>2</sup> intensity mapping experiment with integration time of 48 s, 96 s, 192 s, and 384 s per beam. At  $k \approx 0.07$  h/Mpc, the S/N can reach 5.0 until redshift of 0.35, 0.55, 0.75, and 1.05, respectively.

For redshift  $0.35 < z < 0.45$  case, i.e. the right column in Fig. 13, the wide-band receiver is assumed to be used as the corresponding frequency range is beyond that of the L-band receiver. This receiver has a much higher noise level and the system temperature is 60 K, so the survey precision for the same integration time degraded a lot. Also, as it is a single feed receiver, the required survey time for accomplishing the same integration time would be much longer. In order to make a good intensity mapping survey at the higher redshifts, one must be able to have multibeam receivers, preferably with low system noise. In the next subsection where the cosmological measurement is discussed, we shall assume that such a receiver, namely the UHF PAF receiver introduced in Section 2.1, will be available in the future for such surveys.

### 4.3 Cosmological constraints

Because of the limit of our survey volume, we estimate the induced measurement error on cosmological parameters by using the Fisher matrix formalism (Tegmark 1997; Seo & Eisenstein 2003; More et al. 2013). The Fisher matrix for parameter set  $\{p_i\}$  is given by

$$F_{ij} = \int_{k_{\min}}^{k_{\max}} \frac{\partial \ln P(\vec{k})}{\partial p_i} \frac{\partial \ln P(\vec{k})}{\partial p_j} V_{\text{eff}}(\vec{k}) \frac{d\vec{k}}{2(2\pi)^3}. \quad (21)$$

The usable range of  $k_{\min}$  and  $k_{\max}$  is assumed to be  $k_{\min} = 10^{-3}$  h/Mpc (from survey volume) and  $k_{\max} = 0.1$  h/Mpc. The

observed power spectrum  $P(\vec{k})$  is given by

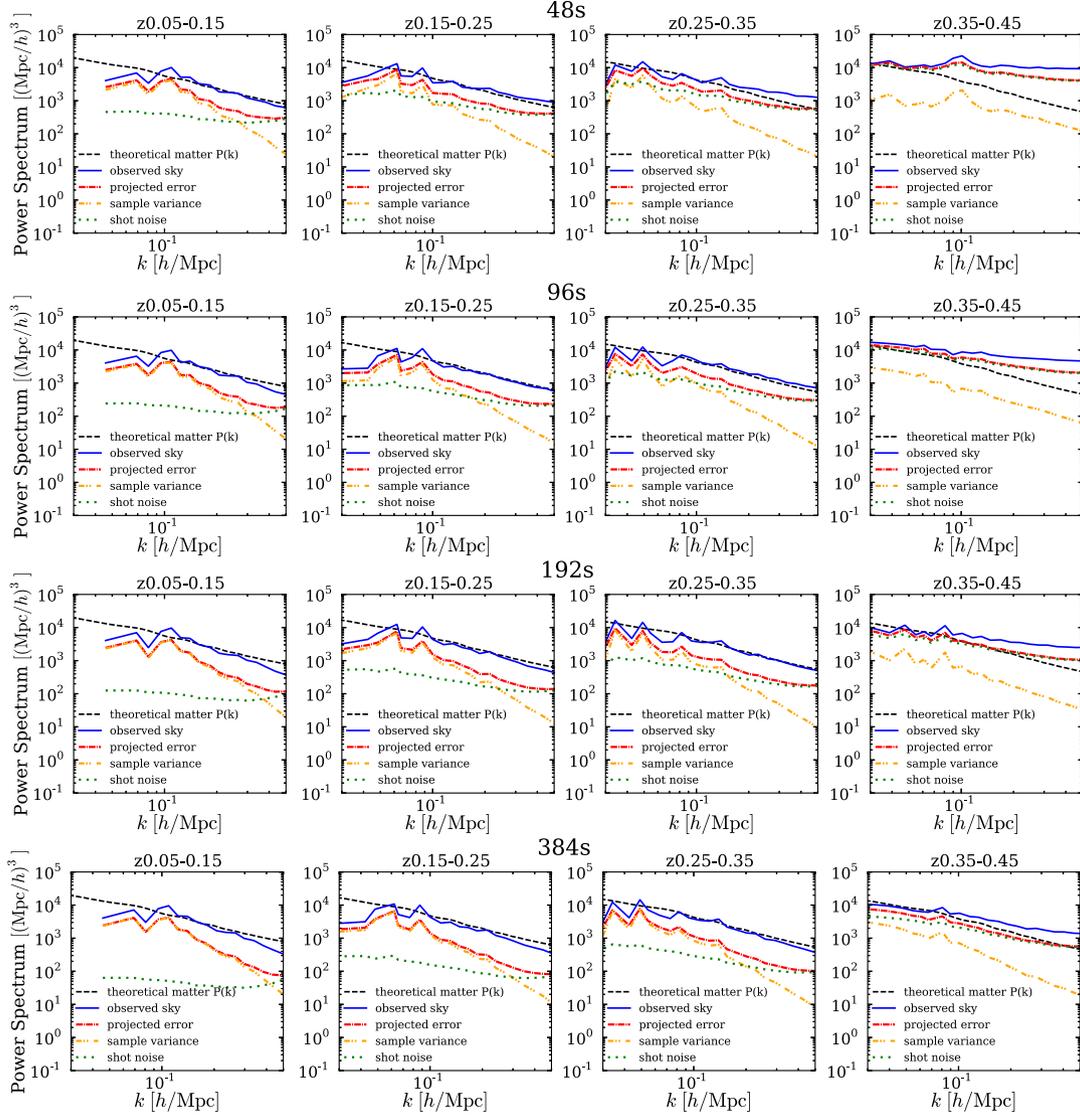
$$P_{\text{obs}}(k_{\text{ref}\perp}, k_{\text{ref}\parallel}) = \frac{D_A(z)_{\text{ref}}^2 H(z)}{D_A(z)^2 H(z)_{\text{ref}}} b_{\text{H I}}^2 \left( 1 + \beta \frac{k_{\parallel}^2}{k_{\perp}^2 + k_{\parallel}^2} \right)^2 \times \left( \frac{G(z)}{G(z=0)} \right)^2 P_{\text{m},z=0}(k) + P_{\text{shot}}, \quad (22)$$

where  $b_{\text{H I}}$  is the linear bias factor of HI gas and the redshift space distortion factor  $\beta = \Omega_m(z)^{0.6}/b_{\text{H I}}(z)$ . The Hubble parameter and the angular diameter distance can be computed for a model with dark energy equation of state parametrized in the form  $w(z) = w_0 + w_a \frac{z}{1+z}$  (Chevallier & Polarski 2001). To obtain useful constraints on cosmological parameters, it is necessary to break the degeneracy by combining the BAO data with data obtained from some other cosmological observations, e.g. CMB. The total Fisher matrix on distance parameters is given by

$$F^{\text{tot}} = F^{\text{CMB}} + \sum_i F^{\text{LSS}}(z_i), \quad (23)$$

and  $F^{\text{LSS}}(z_i)$  is the Fisher matrix derived from the  $i$ -th redshift bin of the LSS survey. For FAST HI galaxy survey and intensity mapping observations, we divide the redshift region into several bins with equal redshift interval, which we set as 0.05.

The HI gas mostly distributed in galaxies hosted by haloes after the reionization; thus, the HI bias can be modelled as the halo bias



**Figure 13.** The intensity mapping power spectrum for different integration time per beam from simulation.

weighted by HI mass hosted by these haloes:

$$b_{\text{HI}} = \frac{1}{\rho_{\text{HI}}(z)} \int_{M_{\text{min}}}^{M_{\text{max}}} dM \frac{dn}{dM}(M, z) M_{\text{HI}}(M) b(M, z), \quad (24)$$

where  $dn/dM$  is the halo mass function, for which we use the fittings of Tinker et al. (2008, 2010) of the Sheth–Tormen function (Sheth & Tormen 2002). The HI mass in a halo of mass  $M$  is given by  $M_{\text{HI}}(M) = AM^\alpha$  where the prefactor  $A$  will be cancelled in the normalization of  $\rho_{\text{HI}}(z)$ , and  $\alpha \simeq 0.6$  (Santos et al. 2015). The halo bias in this ellipsoidal collapse model is modelled as,

$$b(M, z) = \frac{1}{\delta_c(z)} \left[ v'^2 + b v'^{2(1-c)} - \frac{v'^{2c} \sqrt{a}}{v'^{2c} + b(1-c)(1-c/2)} \right],$$

where  $a = 0.707$ ,  $b = 0.5$ ,  $c = 0.6$ ,  $v' = \sqrt{a}v$ , and  $v = \frac{1.686}{D(z)\sigma_R}$ , where  $D(z)$  is the linear growth factor (Mo & White 2002).

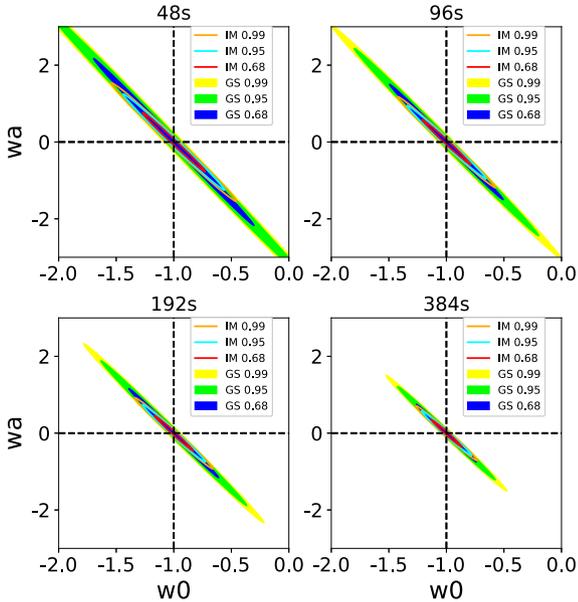
The bias for the HI galaxy survey and the HI intensity mapping survey is different. For the HI galaxy survey, the  $M_{\text{min}}$  is given by the minimum mass of detected galaxies in the survey, while for HI intensity mapping survey, it is the minimum mass, about  $\sim 10^6 M_\odot$ .

#### 4.4 Results

We consider the constraint on dark energy equation of state (EOS) parameters ( $w_a$ ,  $w_0$ ) from the HI galaxy surveys and intensity mapping surveys. The cosmological constant point  $w_0$ ,  $w_a = (-1, 0)$  is taken as the fiducial model. The L-band 19-beam receiver is assumed to be used for the redshift  $0.05 < z < 0.35$ . We then consider the cosmological constraint derived from (i) the L-band survey only; (ii) the L-band survey plus the existing wide-band single feed receiver; or (iii) the L-band plus the future PAF receiver with 81 beams in the UHF band for higher redshift intensity mapping (see Section 2.1 for descriptions). In each case, we consider that the integration time per beam on the L-band is 48 s, 96 s, 192 s, and 384 s, respectively. As discussed in Section 2.1, the 48-s integration can be completed in one scan of the 19-beam receiver. When the wide-band single feed receiver is used, for simplicity, we assume that it acquires the same amount of integration time per beam. For the PAF in UHF band, we have considered instead two integration time per beam, 216 s corresponding to (one scan) or 432 s (two scans) of the PAF receiver. These are added to the L-band 192 s and 384 s, respectively, for illustration.

**Table 2.** Constraint on dark energy EOS parameters ( $w_0$ ,  $w_a$ ) for FAST galaxy surveys (GS) and intensity mapping (IM) surveys with a total of 20000 deg<sup>2</sup> area and Planck prior. In first column, the L denotes L-band 19 feed receiver,  $w$  denotes wide-band single feed receiver, P denotes the PAF receiver in UHF band, and the number denotes integration time in seconds per beam. The last column shows the total time for the survey to be completed.

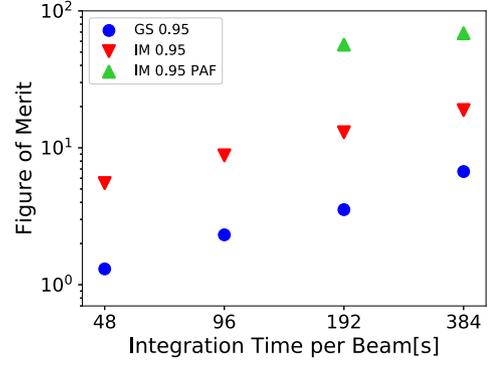
Survey	GS		IM	Observation time (day)
	$(\sigma_{w_0}, \sigma_{w_a})$	$(\sigma_{w_0}, \sigma_{w_a})$		
L 48 s	(0.46, 1.44)	(0.19, 0.53)		220
L 96 s	(0.33, 1.00)	(0.15, 0.43)		440
L 192 s	(0.25, 0.77)	(0.13, 0.36)		880
L 384 s	(0.17, 0.49)	(0.12, 0.33)		1760
(L + w) 48 s	(0.46, 1.44)	(0.18, 0.50)		220 (L) + 2422 (w)
(L + w) 96 s	(0.33, 1.00)	(0.14, 0.39)		440 (L) + 4844 (w)
(L + w) 192 s	(0.25, 0.77)	(0.11, 0.30)		880 (L) + 9688 (w)
(L + w) 384 s	(0.17, 0.49)	(0.09, 0.23)		1760 (L) + 19376 (w)
L(192 s) + P (216 s)	–	(0.05, 0.12)		880 (L) + 135 (P)
L(384 s) + P (432 s)	–	(0.04, 0.10)		1760 (L) + 270 (P)



**Figure 14.** The constraints on dark energy EOS parameters from FAST galaxies survey and intensity mapping, both combined with Planck CMB observation. The three colours are iso-probability contours for 0.68, 0.95, and 0.99, respectively. Galaxies survey (GS) is labelled as filled ellipse and intensity mapping (IM) is labelled with solid lines.

These survey configurations are listed in the first column of Table 2. The precision of the dark energy EOS parameters ( $\sigma_{w_0}$ ,  $\sigma_{w_a}$ ) is given in the second column for galaxy survey and third column for intensity mapping survey. The required total observation time corresponding to each survey is given in the fourth column of the table, where the time required for each band is given separately. Note that the time listed is for observations, not counting offline time required for calibration, maintenance, etc., so the real time required to complete the survey would be even longer.

We also show in Fig. 14 the error ellipses of the dark energy EOS parameters with the L-band 19-beam receiver and the wide-band single feed receiver. The results for the L-band + PAF surveys are not shown as their error ellipses are much smaller. The dark energy task force (DETF) figure of merit, which is defined as the



**Figure 15.** The figure of merit for dark energy EOS parameters from FAST galaxies survey and intensity mapping combined with Planck CMB observation. Galaxy surveys (GS) are labelled as filled blue circles and intensity mapping surveys (IM) are labelled with filled red down-pointing triangle for confidence limit of 0.95. The IM surveys with low-frequency PAF are shown in green up-pointing triangle symbols.

inverse of the area of the  $2\sigma$  error ellipse (Albrecht et al. 2006), is shown in Fig. 15 for the L-band + wide-band receiver as well as the L-band + UHF-band PAF.

From these, we see that the intensity mapping can achieve much higher precision in the measurement of the dark energy EOS parameters than the HI galaxy surveys. The intensity mapping survey with the shortest integration per beam (48 s) has a figure of merit comparable with the HI galaxy survey of the longest integration time per beam (384 s) but requires only 1/8 of the total observation time in the L-band. However, even with the intensity mapping survey, the figure of merit is only of order  $10^1$ , much less than the current optical surveys. This is not surprising as the L-band is limited to relatively low redshifts ( $z < 0.35$ ). However, we note that so far there has not yet been an HI survey providing constraints on the dark energy parameters, and the HI survey complements the optical survey as it uses a difference tracer, which would be valuable to reduce any possible systematic errors in the BAO measurement.

From Table 2, we also see that the constraints on the EOS parameters are only slightly improved by adding the surveys of equal integration time per beam with the wide-band receiver. This is because this receiver has a higher system temperature (60 K) over its very wide frequency coverage. Furthermore, because it has only a single feed, to achieve such integration time per beam would require very long observation time, which is quite impractical. However, if equipped with a powerful PAF receiver, the measurement can be taken to much higher redshift in reasonable time – indeed, for the PAF parameter we assumed, it would take even less time to complete than the L-band. The figure of merit could then be lifted substantially, up to a level comparable with DETF stage IV experiments (Albrecht et al. 2006). This shows that a PAF receiver at the UHF band would be a very valuable addition to the FAST telescope.

## 5 DISCUSSIONS

### 5.1 The choice of galaxy survey versus intensity mapping

Using the Fisher information, Cheng et al. (2018) developed a formalism to quantify the performance of galaxy redshift survey and intensity mapping when measuring LSSs. Under the assumption

**Table 3.** Four limiting regimes defined by the relative value of the luminosity scale where the voxels are highly susceptible to shot noise,  $L_{\text{SN}}$ , the rms noise per voxel,  $\sigma_L$ , and the characteristic luminosity for a certain luminosity function,  $l_*$ .

Number	Regime	Optimal strategy
1	$L_{\text{SN}} < \sigma_L < l_*$	Galaxy detection
2	$\sigma_L < L_{\text{SN}} < l_*$	Galaxy detection/intensity mapping <sup>a</sup>
3	$L_{\text{SN}} < l_* < \sigma_L$	Intensity mapping
4	$l_* < L_{\text{SN}}$	Intensity mapping

<sup>a</sup>Here the optimal strategy is an intermediate between the intensity mapping and galaxy detection observables.

that the galaxy population follows a Schechter function form,

$$n(L) = \phi_* \left( \frac{L}{l_*} \right)^\alpha e^{-L/l_*}, \quad (25)$$

the optimal strategy for survey can be found using the relative value of three parameters:  $\{L_{\text{SN}}, \sigma_L, l_*\}$ , where  $L_{\text{SN}}$  is the luminosity scale on which the voxels are susceptible to shot noise,  $\sigma_L$  refers to the rms noise per voxel, and  $l_*$  is the characteristic luminosity. The  $L_{\text{SN}}$  is derived with  $\sigma_{\text{SN}}(L_{\text{SN}}/l_*) = L_{\text{SN}}/l_*$ , where

$$\sigma_{\text{SN}}^2(l) = V_{\text{vox}} \phi_* \int_0^l dl' l'^{\alpha+2} e^{-l'}. \quad (26)$$

$V_{\text{vox}}$  is the comoving volume of a voxel and  $\alpha$  is the faint-end slope parameter of the luminosity function. Observations can be divided into four limiting regimes for optimal strategy, as shown in Table 3. In regime 1, the instrument noise is much smaller than  $l_*$ , confusion effect is small, and galaxy detection is optimal. In regime 2, the optimal strategy is somewhere intermediate between the intensity mapping and galaxy detection, because the voxels with  $L \geq \sigma_L$  will suffer from confusion noise. In regime 3, the instrument noise in a voxel is very large,  $l_* < \sigma_L$ , the intensity mapping will be the only choice. Regime 4 corresponds to a large effective number of galaxies per voxel, galaxy detection will suffer from large confusion noise, and intensity mapping is optimal.

For the FAST HI survey, assuming a voxel of an angular resolution of 0.08 deg and bandwidth of 200 km s<sup>-1</sup>, we compute the evolution of  $L_{\text{SN}}/l_*$  and  $\sigma_L/l_*$  in redshift and show the results in Fig. 16. We use the luminosity function given by Blanton et al. (2003), with the following parameters:  $\phi_* = 5.11 \times 10^{-3} h_{70}^3 \text{Mpc}^{-3}$ ,  $\log(L_*/L_\odot) = 10.36 + \log h_{70}$ , and  $\alpha = -1.05$ . The redshift evolution of Schechter luminosity function parameters is modelled as (Lin et al. 1999),

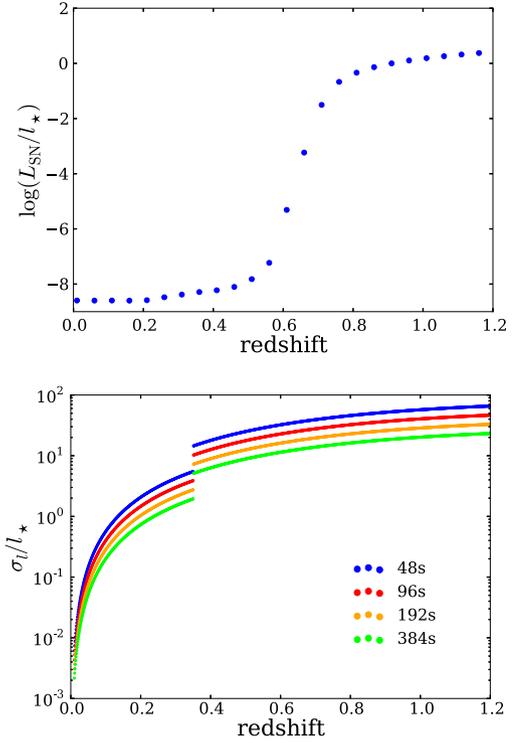
$$\alpha(z) = \alpha(z_0), \quad (27)$$

$$M^*(z) = M^*(z_0) - Q(z - z_0), \quad (28)$$

$$\varphi^*(z) = \varphi^*(z_0) 10^{0.4P(z-z_0)}, \quad (29)$$

where  $P = 1.0$  and  $Q = 1.03$  (Loveday et al. 2015) in the  $r$ -band.

Fig. 16 shows the redshift evolution of  $L_{\text{SN}}/l_*$  (top panel) and  $\sigma_L/l_*$  (bottom panel). We see  $L_{\text{SN}} \leq l_*$  at redshift  $\lesssim 0.8$  and  $\sigma_L \leq l_*$  at redshift  $\lesssim 0.13, 0.16, 0.19,$  and  $0.23$  for integration time of 48 s, 96 s, 192 s, and 384 s, respectively. Combining with Table 3, it shows for FAST that the galaxy redshift survey is the optimal strategy at redshift  $\lesssim 0.13, 0.16, 0.19,$  and  $0.23$  for survey with 48 s, 96 s, 192 s, and 384 s integration time per beam. This is in agreement with what the projected error shown in Figs 10 and 12. We note that the voxel size in the projected error calculation for the intensity



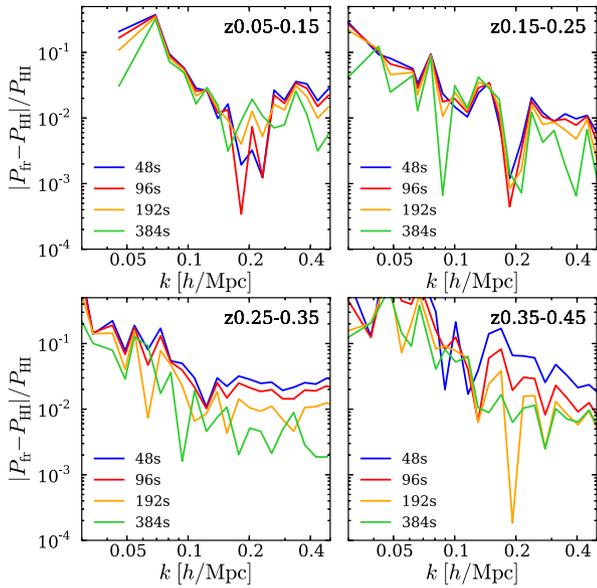
**Figure 16.** The redshift evolution of  $L_{\text{SN}}/l_*$  and  $\sigma_L/l_*$ , assuming an angular resolution of 0.08 deg and a bandwidth of 200 km s<sup>-1</sup>. The jump at redshift 0.35 in the bottom panel is due to the change of receiver system.

mapping is 2 Mpc/h, which is larger than the voxel size for galaxy detection. This makes the redshift points where intensity mapping is better than galaxy detection a little lower than the redshifts shown in Fig. 16.

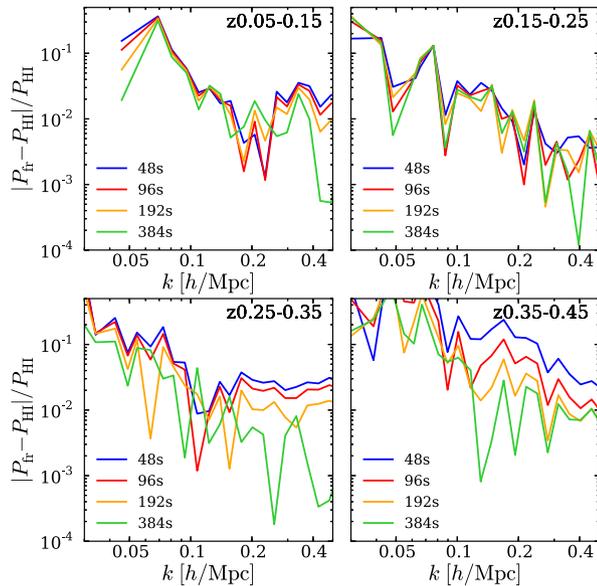
## 5.2 Foreground

One of the most challenging problems in intensity mapping experiment may be the contamination from the foreground radiation, which is several orders of magnitude larger in amplitude than the HI intensity signal. It can in principle be subtracted, and the true signal recovered, based on the fact that the frequency dependence and some statistical properties of the foreground are different from the true signal. Sophisticated mathematical methods have been developed (Liu & Tegmark 2011; Chapman et al. 2013; Switzer et al. 2013; Shaw et al. 2014; Wolz et al. 2014; Alonso et al. 2015; Zhang et al. 2016). The intensity mapping experiment with the Green Bank Telescope (Switzer et al. 2013) has shown that the foregrounds can indeed be suppressed significantly, though at present a positive detection of 21-cm auto-power spectrum is yet to be achieved. Here, we assume that after a successful foreground subtraction, the contamination can be reduced to the thermal noise level (Bigot-Sazy et al. 2015).

To investigate the impact of foreground on the FAST IM survey, we made a simple test of foreground removal in our simulation. We produce the foreground with the global sky model (de Oliveira-Costa et al. 2008; Zheng et al. 2017), convolved with a frequency-dependent beam and add to the noise-filled data cube. Then using a third-order log-log polynomial fitting, we find that the foreground can be removed effectively. The residue difference power spectra between the original 21-cm signal and the one obtained by removing the simulated foreground with the polynomial fit in the image cube



**Figure 17.** The relative error power spectra, defined by  $(P_{\text{HI}} - P_{\text{fr}})/P_{\text{HI}}$ , where  $P_{\text{HI}}$  is the true 21-cm power, and  $P_{\text{fr}}$  is the foreground removed power at different redshifts and integration time.



**Figure 18.** Same as Fig. 17, but the beam is modelled by Bessel function.", changed

are shown in Fig. 17, with different integration times per beam for a Gaussian beam. The power spectra show that the differences are small at about 1 percent level at  $k \approx 0.1$  h/Mpc for these conditions. As expected, the amplitude of this error power spectra generally decreases as integration time increases. The differences at larger scales (smaller  $k$ ) are larger, because a few large-scale modes (modes with small  $k_{\text{nl}}$ ) are removed by the foreground removal methods; in principle, such large-scale modes can be reconstructed via cosmic tidal reconstruction (Zhu et al. 2018). Using the foreground removed cube, we obtain the same constraints on the cosmological parameters in high accuracy.

Besides the Gaussian beam profile, we also made a test for the case of oscillatory side lobes. In Fig. 18 we make the same foreground subtraction excise but for a beam function modelled using the Bessel function as  $|2 \times J_1(x)/x|^2$ , where  $x = 3.23 \times \theta/\sigma_b$ . The result is qualitatively similar to the Gaussian case, though there are some differences.

The simulation of the foreground and its removal presented above may still be too simplified. In reality, the beam response could be more complicated, and we have only an imperfect knowledge, which must itself be determined from observation, and there are also irregularities in the beam and bandpass, polarization leakage, the  $1/f$  noise and variations of the system gain, etc., making it much harder to remove. So, the actual impact of the foreground could be higher than this simple estimate. Obviously, the difficulty of the foreground removal depends on the design of the telescope. The stability, regularity, and dynamic range all affect the induced foreground contamination. A detailed study of foreground subtraction requires a realistic assessment of FAST telescope response, which could be obtained only with actual observational data. This is beyond the scope of this paper. However, the high sensitivity of the FAST is advantageous for 21-cm extraction: the individual voxel signal-to-noise ratio is relatively high, so it is easier to be detected.

## 6 SUMMARY

In this paper, we make a detailed study of large area drift scan HI survey with the FAST telescope. We considered using the existing L-band 19-beam receiver, the single feed wide-band receiver for lower frequency (UHF), and also contemplated using a future UHF-band PAF receiver. We simulated observation of HI galaxies, the number density of the detectable HI galaxies decreases rapidly as the redshift increases, and also due to the larger beam size and smaller galaxy size at higher redshifts, the mean number of galaxies within each voxel, which we called confusion rate, also increases. For the detected galaxies, there are also measurement errors in both the direction of the line of sight and the direction perpendicular to it. We estimated such errors but found that the main source of HI galaxy power spectrum error still comes from the shot noise on small scales and cosmic variance on large scales. We also considered the HI intensity mapping observation.

The precision of the power spectrum measurement is forecasted using Fisher matrix for large survey areas, and we also make mock observations for both galaxy survey and intensity mapping survey using simulation. With intensity mapping, the power spectrum can be measured with high precision. We find that the FAST can effectively detect the individual galaxy till  $z \approx 0.2, 0.25, 0.3,$  and  $0.35,$  or map the LSS with intensity mapping till  $z \approx 0.35, 0.55, 0.75,$  and  $1.05,$  respectively, if we assume 48 s, 96 s, 192 s, and 384 s integration time per beam. Generally, the HI intensity mapping observations can yield much more precise measurement, though the HI galaxy survey can also achieve nearly optimal measurement at lower redshifts, with  $z \lesssim 0.13, 0.16, 0.19,$  and  $0.23$  for surveys with 48 s, 96 s, 192 s, and 384 s integration time per L-band beam.

We find that the FAST HI intensity mapping survey can produce a good measurement of the underlying power spectrum and use the BAO method to measure the dark energy equation of state parameters. Such a measurement with a radio tracer is complementary to the optical BAO measurements and reduces possible systematic errors. We also made a study of the impact of foregrounds on the measurement by simulation and found that it does not significantly affect the results, at least under the simplified conditions assumed in the study.

With existing receivers, the achievable precision is moderate, as it is largely limited to  $z < 0.35$ , and for higher redshifts, the wide-band single feed would take too long time to complete the survey. However, if equipped with a UHF-band multibeam receiver, higher redshifts can be observed more efficiently. We considered the case of a UHF-band PAF receiver with frequency coverage of  $0.5 \sim 1.0$  GHz and found that the resulting survey may yield dark energy figure of merit up to  $10^2$ , comparable with the DETF stage IV results. This shows that a state-of-the-art PAF multibeam receiver would be a very valuable addition to the FAST telescope.

## ACKNOWLEDGEMENTS

This work was first inspired by the late Prof. Rendong Nan. The parameters of the FAST receivers are provided to us by Di Li, Youling Yue, Lei Qian, and Jin Fan. We also thank Shifan Zuo and Yichao Li for discussions. This work is supported by the National Natural Science Foundation of China (NSFC) key project grant 11633004, the Chinese Academy of Sciences (CAS) Frontier Science Key Project QYZDJ-SSW-SLH017 and the CAS Interdisciplinary Innovation Team grant (JCTD-2019-05), the NSFC-Isreal Science Foundation (ISF) joint research program no. 11761141012, the NSFC grant 11773034, the Ministry of Science and Technology (MoST) grant 2016YFE0100300, the CAS Strategic Priority Research Program XDA15020200, and the National Key R&D Program 2017YFA0402603.

## REFERENCES

- Alam S. et al., 2017, *MNRAS*, 470, 2617  
 Albrecht A. et al., 2006, preprint (astro-ph/0609591)  
 Alonso D., Bull P., Ferreira P. G., Santos M. G., 2015, *MNRAS*, 447, 400  
 Ansari R. et al., 2012, *A&A*, 540, A129  
 Ansari R., Le Goff J., Magneville C., Moniez M., Palanque-Delabrouille N., Rich J., Ruhlmann-Kleider V., Yèche C., 2008, preprint (arXiv:0807.3614)  
 Ata M. et al., 2018, *MNRAS*, 473, 4773  
 Bandura K. et al., 2014, Proc. SPIE Conf. Ser. Vol. 9145, in Ground-based and Airborne Telescopes V. SPIE, Bellingham, p. 914522,  
 Battye R. et al., 2016, preprint (arXiv:1610.06826)  
 Battye R. A. et al., 2012, preprint (arXiv:1209.1041)  
 Beutler F. et al., 2011, *MNRAS*, 416, 3017  
 Bigot-Sazy M.-A. et al., 2015, *MNRAS*, 454, 3240  
 Blake C. et al., 2011, *MNRAS*, 418, 1707  
 Blanton M. R. et al., 2003, *ApJ*, 592, 819  
 Blitz L., Rosolowsky E., 2006, *ApJ*, 650, 933  
 Braun R., 2012, *ApJ*, 749, 87  
 Catinella B., Giovanelli R., Haynes M. P., 2006, *ApJ*, 640, 751  
 Chang T.-C., Pen U.-L., Peterson J. B., McDonald P., 2008, *Phys. Rev. Lett.*, 100, 091303  
 Chapman E. et al., 2013, *MNRAS*, 429, 165  
 Chen X., 2012, Int. J. Mod. Phys. Conf. Ser., 12, 256,  
 Cheng Y.-T., de Putter R., Chang T.-C., Dore O., 2018, *ApJ*, 877, 86  
 Chevallier M., Polarski D., 2001, *Int. J. Mod. Phys. D.*, 10, 213  
 Cole S. et al., 2005, *MNRAS*, 362, 505  
 Davé R., Rafieferantsoa M. H., Thompson R. J., Hopkins P. F., 2017, *MNRAS*, 467, 115  
 Dawson K. S. et al., 2016, *AJ*, 151, 44  
 De Lucia G., Blaizot J., 2007, *MNRAS*, 375, 2  
 de Oliveira-Costa A., Tegmark M., Gaensler B. M., Jonas J., Landecker T. L., Reich P., 2008, *MNRAS*, 388, 247  
 Delhaize J., Meyer M. J., Staveley-Smith L., Boyle B. J., 2013, *MNRAS*, 433, 1398  
 Drinkwater M. J. et al., 2010, *MNRAS*, 401, 1429  
 Duffy A. R., Battye R. A., Davies R. D., Moss A., Wilkinson P. N., 2008, *MNRAS*, 383, 150  
 Duffy A. R., Meyer M. J., Staveley-Smith L., Beryk M., Croton D. J., Koribalski B. S., Gerstmann D., Westerland S., 2012, *MNRAS*, 426, 3385  
 Eisenstein D. J. et al., 2005, *ApJ*, 633, 560  
 Elson E. C., Blyth S. L., Baker A. J., 2016, *MNRAS*, 460, 4366  
 Feldman H. A., Kaiser N., Peacock J. A., 1994, *ApJ*, 426, 23  
 Freudling W. et al., 2011, *ApJ*, 727, 40  
 Giovanelli R. et al., 2005, *AJ*, 130, 2598  
 Giovanelli R. et al., 2007, *AJ*, 133, 2569  
 Giovanelli R., Haynes M. P., 2002, *ApJ*, 571, L107  
 Glazebrook K. et al., 2007, in Metcalfe N., Shanks T., eds, Cosmic Frontiers, Vol. 379, Astron. Soc. Pac., San Francisco, p. 72  
 Goldberg D. M., Strauss M. A., 1998, *ApJ*, 495, 29  
 Gunn J., Weinberg D., 1995, in Maddox S. J., Aragon-Salamanca A., eds, Wide Field Spectroscopy and the Distant Universe, p. 3  
 Hoppmann L., Staveley-Smith L., Freudling W., Zwaan M. A., Minchin R. F., Calabretta M. R., 2015, *MNRAS*, 452, 3726  
 Hu W. et al., 2019, *MNRAS*, 489, 1619  
 Jarvis M. J. et al., 2014, preprint (arXiv:1401.4018)  
 Jones M. G., Haynes M. P., Giovanelli R., Moorman C., 2018, *MNRAS*, 477, 2  
 Kanekar N., Sethi S., Dwarakanath K. S., 2016, *ApJ*, 818, L28  
 Lah P. et al., 2007, *MNRAS*, 376, 1357  
 Leroy A. K., Walter F., Brinks E., Bigiel F., de Blok W. J. G., Madore B., Thornley M. D., 2008, *AJ*, 136, 2782  
 Li D. et al., 2018, *IEEE Microw. Mag.*, 19, 112  
 Lin H., Yee H. K. C., Carlberg R. G., Morris S. L., Sawicki M., Patton D. R., Wirth G., Shepherd C. W., 1999, *ApJ*, 518, 533  
 Liu A., Tegmark M., 2011, *Phys. Rev. D*, 83, 103006  
 Loveday J. et al., 2015, *MNRAS*, 451, 1540  
 Martin A. M., Papastergis E., Giovanelli R., Haynes M. P., Springob C. M., Stierwalt S., 2010, *ApJ*, 723, 1359  
 Meiksin A., White M., Peacock J. A., 1999, *MNRAS*, 304, 851  
 Meyer M. J. et al., 2004, *MNRAS*, 350, 1195  
 More S., van den Bosch F. C., Cacciato M., More A., Mo H., Yang X., 2013, *MNRAS*, 430, 747  
 Mo H. J., White S. D. M., 2002, *MNRAS*, 336, 112  
 Nan R. et al., 2011, *Int. J. Mod. Phys. D*, 20, 989  
 Neeleman M., Prochaska J. X., Ribaud J., Lehner N., Howk J. C., Rafelski M., Kanekar N., 2016, *ApJ*, 818, 113  
 Newburgh L. B. et al., 2016, Proc. SPIE Conf. Ser. Vol. 9906, Ground-based and Airborne Telescopes VI. SPIE, Bellingham, p. 99065X  
 Nyland K. et al., 2017, *MNRAS*, 464, 1029  
 Obreschkow D., Meyer M., 2014, preprint (arXiv:1406.0966)  
 Obreschkow D., Croton D., De Lucia G., Khochfar S., Rawlings S., 2009a, *ApJ*, 698, 1467  
 Obreschkow D., Heywood I., Klöckner H.-R., Rawlings S., 2009b, *ApJ*, 702, 1321  
 Obreschkow D., Klöckner H.-R., Heywood I., Levrier F., Rawlings S., 2009c, *ApJ*, 703, 1890  
 Percival W. J. et al., 2001, *MNRAS*, 327, 1297  
 Percival W. J. et al., 2010, *MNRAS*, 401, 2148  
 Rao S. M., Turnshek D. A., Nestor D. B., 2006, *ApJ*, 636, 610  
 Rao S. M., Turnshek D. A., Sardane G. M., Monier E. M., 2017, *MNRAS*, 471, 3428  
 Rhee J., Zwaan M. A., Briggs F. H., Chengalur J. N., Lah P., Oosterloo T., van der Hulst T., 2013, *MNRAS*, 435, 2693  
 Rhee J., Lah P., Chengalur J. N., Briggs F. H., Colless M., 2016, *MNRAS*, 460, 2675  
 Rhee J., Lah P., Briggs F. H., Chengalur J. N., Colless M., Willner S. P., Ashby M. L. N., Le Fèvre O., 2018, *MNRAS*, 473, 1879  
 Ross A. J., Samushia L., Howlett C., Percival W. J., Burden A., Manera M., 2015, *MNRAS*, 449, 835  
 Saintonge A., 2007, *AJ*, 133, 2087  
 Santos M. et al., 2015, Advancing Astrophysics with the Square Kilometer Array (AASKA14), SKA office, p. 19

- Schlegel D., White M., Eisenstein D., 2009, in *astro2010: The Astronomy and Astrophysics Decadal Survey*, preprint ([arXiv:0902.4680](https://arxiv.org/abs/0902.4680))
- Seo H.-J., Eisenstein D. J., 2003, *ApJ*, 598, 720
- Seo H.-J., Dodelson S., Marriner J., McGinnis D., Stebbins A., Stoughton C., Vallinotto A., 2010, *ApJ*, 721, 164
- Serra P. et al., 2012, *MNRAS*, 422, 1835
- Shaw J. R., Sigurdson K., Pen U.-L., Stebbins A., Sitwell M., 2014, *ApJ*, 781, 57
- Sheth R. K., Tormen G., 2002, *MNRAS*, 329, 61
- Smoot G. F., Debono I., 2017, *A&A*, 597, A136
- Springel V. et al., 2005, *Nature*, 435, 629
- Switzer E. R. et al., 2013, *MNRAS*, 434, L46
- Tacconi L. J. et al., 2006, *ApJ*, 640, 228
- Tegmark M., 1997, *Phys. Rev. Lett.*, 79, 3806
- Tegmark M. et al., 2004, *ApJ*, 606, 702
- Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, *ApJ*, 688, 709
- Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, *ApJ*, 724, 878
- Westmeier T., Jurek R., Obreschkow D., Koribalski B. S., Staveley-Smith L., 2014, *MNRAS*, 438, 1176
- Wolz L., Abdalla F. B., Blake C., Shaw J. R., Chapman E., Rawlings S., 2014, *MNRAS*, 441, 3271
- Wu Y., Jin C., Fan J., Zhao X., Yu L., Du B., 2016, *Development of the L-band phased array feed for the Five hundred meter Aperture Spherical Telescope*, progress in electromagnetic research symposium (PIERS), Shanghai, 1667
- Xu Y., Wang X., Chen X., 2015, *ApJ*, 798, 40
- Yohana E., Li Y.-C., Ma Y.-Z., 2019, preprint ([arXiv:1908.03024](https://arxiv.org/abs/1908.03024))
- Young L. M., 2002, *AJ*, 124, 788
- Zhang L., Bunn E. F., Karakci A., Korotkov A., Sutter P. M., Timbie P. T., Tucker G. S., Wandelt B. D., 2016, *ApJS*, 222, 3
- Zheng H. et al., 2017, *MNRAS*, 464, 3486
- Zhu H.-M., Pen U.-L., Yu Y., Chen X., 2018, *Phys. Rev. D*, 98, 043511
- Zwaan M. A. et al., 2004, *MNRAS*, 350, 1210
- Zwaan M. A., Meyer M. J., Staveley-Smith L., Webster R. L., 2005, *MNRAS*, 359, L30

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.