HIR4: Cosmological signatures imprinted on the cross correlation between 21cm map and galaxy clustering

Feng Shi1, Yong-Seon Song1, Jacobo Asorey1,2, David Parkinson1, Kyungjin Ahn3, Jian Yao5, Le Zhang4,5 and Shifan Zuo6,7

1 Korea Astronomy and Space Science Institute, Yuseong-gu, Daejeon 34055, Korea
2 Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Av. Complutense, 40, 28040 Madrid, Spain
3 Department of Earth Sciences, Chosun University, Gwangju 61452, Korea
4 School of Physics and Astronomy, Sun Yat-Sen University, 2 Daxue Road, Zhuhai, 519082, P.R. China
5 Department of Astronomy, Shanghai Jiao Tong University, Shanghai, 200240, P.R. China
6 Key Laboratory of Computational Astrophysics, National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, P. R. China
7 Department of Astronomy and Tsinghua Center for Astrophysics, Tsinghua University, Beijing 100084, P. R. China

ABSTRACT
We explore the cosmological multi-tracer synergies between a DESI-like Emission Line Galaxy (ELG) distribution and a Tianlai-like 21cm intensity map. We use simulated maps generated from a particle simulation in the light-cone volume (Horizon Run 4), sky-trimmed and including the effects of foreground contamination, its removal, and instrument noise. We first validate how the foreground residual affects the recovered 21cm signal by putting different levels of foreground contamination into the 21cm maps. We find that the contamination cannot be ignored in the angular auto-power spectra of HI even when it is small, but has no influence on the accuracy of the angular cross-power spectra between HI and galaxies. In the foreground-cleaned map case, as information is lost in the cleaning procedure, there is also a bias in the cross-power spectrum. However, we found that the bias from the cross-power spectrum is scale-independent, which is easily parameterised as part of the model, while the offset in the HI auto-power spectrum is non-linear. In particular, we tested that the cross power also benefits from the cancellation of the bias in the power spectrum measurement that is induced by the instrument noise, which changes the shape of the auto-power spectra but leaves the cross-power unaffected. We then modelled the angular cross-power spectra to fit the BAO feature in broadband shape of the angular cross power spectrum, including contamination from the residual foreground and the effect of instrument noise. We forecast a constraint on the angular diameter distance $D_A$ for the Tianlai Pathfinder redshift $0.775 < z < 1.03$, giving a distance measurement with a precision of 3.3% at that redshift.

Key words: cosmology: theory, dark energy, large-scale structure of the Universe

1 INTRODUCTION
Understanding the nature of the accelerated expansion of the Universe is one of the most important currently outstanding problems in cosmology. In the last few decades, cosmological observations from galaxy surveys, the Cosmic Microwave Background (CMB), Baryon Acoustic Oscillations (BAO), weak gravitational lensing shear (WL), gravitational wave standard candles, and other observations, have all made significant contributions in understanding our Universe. But there is still much uncertainty in our knowledge about the dark energy and the physics of the expansion of the Universe, such as the $H_0$ tension between CMB and low redshift measurement (Bernal et al. 2016), and the varying dark energy (Zhao et al. 2017), that requires further tests. In fact, the intermediate timeline of our cosmic distances and expansion rates from observation has not yet been systematically surveyed.

The intensity mapping of 21cm is a technique for surveying the large-scale structure of the Universe, by using
the 21cm spectral line that arises from hyperfine ‘spin flip’ transition of neutral hydrogen (HI). It measures the integrated emission lines that originate from many unresolved galaxies that trace the HI gas, which follows fluctuations in the underlying cosmic density field. As the frequency of the emission line is redshifted by the expansion of the Universe, one can detect the underlying clustering signal as a function of redshift. This is in principal similar to the traditional galaxy redshift survey, but with an important distinction that 21cm intensity mapping is sensitive to all sources of emission, rather than just cataloging the brightest galaxies. As high angular resolution is not required, 21cm intensity mapping can cover large sky areas in a limited observing time. It can explore larger volumes and measure BAO continuously from low to high redshift, such as CHIME (Newburgh et al. 2014), Tianlai (Chen 2011; Xu et al. 2015; Das et al. 2018), SKA (Santos et al. 2015) and HIRAX (Newburgh et al. 2016), which are currently being built and taking data with the goal of measuring the BAO scale to $z = 2.5$ with unprecedented precision. This is ideal for testing the time-variation of the dark-energy equation of state (Dinda et al. 2018), and the problem of the Hubble parameter inconsistency between CMB and local measurements, with a single tracer across a quite wide redshift range (Kovetz et al. 2019).

However, a key problem of the endeavor of such a successful measurement is that 21cm signal is contaminated by several sources of foreground radiation, which are many orders of magnitude brighter than the cosmological signal. Great efforts have been made to study the foreground removal, such as principal component analysis (PCA, de Oliveira-Costa et al. 2008) and fast independent component analysis (FastICA, Chapman et al. 2012; Wolz et al. 2014) method. However, since the foreground noise is tangled with the 21cm signal, there would be always residual foreground on the map, together with the signal information loss by the cleaning procedure.

One possible way to mitigate the contamination by the residuals is to do the cross correlation. So far, the detection of the cross correlation between the large-scale structure and 21cm intensity maps was reported in Chang et al. (2010), based on the data from Green Bank Telescope and DEEP2 galaxy survey at $z \sim 0.8$, and Anderson et al. (2018), based on the data from Parkes radio telescope and the 2dF galaxy survey at $0.057 < z < 0.098$. Based on the simulation data, great efforts have been made not only to understand the HI clustering from auto or cross correlation, but also to extract cosmological information (e.g. Xu et al. 2015; Cunnington et al. 2019; Witzemann et al. 2019; Padmanabhan et al. 2019; Hu et al. 2020; Cunnington et al. 2020).

In order to be more realistic, our HIR4 (HI with Horizon Run 4) project aims to investigate the prospects for probing the HI clustering from future 21cm intensity maps in a manner closer to observation. In our first paper (Asorey et al. 2020), we simulated the future survey of 21cm using the Horizon Run 4 (HR4) (Kim et al. 2015) cosmological N-body simulation in the light cone volume. We generated HI intensity maps from the halo catalogue, and combined with foreground radio emission maps from the Global Sky Model, to create accurate simulations over the entire sky. We simulated the HI sky for the frequency range 700-800 MHz, matching the sensitivity of the Tianlai pathfinder.

In this paper, we focus on testing how the residual foreground affects the clustering signal imprinted in the 21cm intensity map, and validate the detectability of HI clustering from the 21cm auto and 21cm $\times$ galaxy cross angular power spectrum when using the foreground-cleaned 21cm maps. We also forecast the angular diameter distance constraint from the cross correlation between Dark Energy Spectroscopic Instrument (DESI) (DESI Collaboration et al. 2016) galaxy survey and the Tianlai 21cm intensity mapping. The Tianlai project is designed to demonstrate the feasibility of using wide field view radio interferometers to map the density of neutral hydrogen, which is divided into three stages: Pathfinder, Pathfinder+, and Full Array. At present, construction of the Pathfinder was completed and it is now undergoing the calibration process before starting the survey in the near future. The DESI is a new instrument for conducting a spectroscopic survey of distant galaxies, which achieved its first light test in October 2019, and begun survey validation with the completed instrument in early 2020. Both of the two surveys have a quite big overlap in sky coverage and redshift, which provides a great opportunity to measure the cross-correlation in the coming years.

This paper is organized as follows: Section 2 introduces the mock data to simulate the galaxy distribution and 21cm intensity mapping including foreground contamination and instrument noise. In Section 3, we test the detectability of HI clustering from the foreground-removed 21cm intensity maps. Modeling the angular cross-power spectra and fitting the broadband BAO feature is presented in Section 4. Finally, we summarize our conclusion in Section 5.

2 MOCK CATALOGUE

The deepest and widest large scale structure experiments will commence in the near future, designed to observe various targets, such as luminous red galaxies, emission line galaxies, the shear or the magnification of galaxy morphology caused by gravitational lensing, and 21cm emission of neutral atomic hydrogen HI. Both spectroscopic surveys and 21cm emission observations on large scale are planned to scan nearly the same region of sky over the northern hemisphere, and it will be possible to open a new window by cross-correlating them to reveal the underlying matter fluctuations in a different way. While each experiment suffers from systematic uncertainties, such as the fibre assignment contamination for the spectroscopic survey and foreground contamination for the 21cm intensity mapping survey, the cross-correlated measurement is immune to those errors, and provides an independent representation of the invisible large scale structure of cold dark matter. The detectability of underlying clustering by exploiting cross-correlating methodology is verified by analyzing the mock catalogue in which both targets are generated from the same realization of an N-body simulation.

2.1 The planned surveys

The Dark Energy Spectroscopic Instrument (hereafter DESI) has been constructed to probe the signature of dark energy on the expansion of the universe, by obtaining spectroscopy measurements of numerous galaxies to construct a

MNRA 000, 1-14 (2019)
Figure 1. (Left panel) The target angular distribution of the ELG (blue) and Tianlai (cyan) after applying the corresponding mask scheme. (Right panel) The redshift distribution of the DESI expected number density of ELG (blue line) and LRG (red line), which is obtained from DESI Collaboration et al. (2016). The cyan band corresponds to the redshift range of Tianlai Pathfinder.

ELGs are selected using optical color-selection techniques, slicing in optical color-color space to effectively isolate the population of 0.6 < z < 1.7 galaxies, a method that has been confirmed by multiple experiments. A high redshift success rate is expected for ELGs with integrated OII emission-line strengths of 8 \times 10^{-17} \text{erg/s/cm}^2, corresponding to a limiting star-formation rate of 1.5, 5 and 15 M_☉/yr at z ~ 0.6, 1 and 1.6, respectively, which can be well identified at z > 0.6 (DESI Collaboration et al. 2016). The separation of galaxies with redshift lower and higher than z ~ 0.6 is possible by the spectrum blue-ward of the Balmer break moving toward the z-band filter, which causes g−r and r−z colors to be relatively blue at higher redshifts. After applying those specs, we estimate the ELG target distribution in redshift presented in the blue region in the left panel of Figure 1.

Tianlai is planned to demonstrate the feasibility of using HI emission from Tianlai and the DESI target survey area is presented in the green region in the left panel of Figure 1, which is well-overlapped with DESI ELG target sky patch. When both are cross-correlated, most DESI targets observed at the northern hemisphere contribute to the cross-correlation computation.

The Tianlai Pathfinder will be operated at the frequency range of 700-800MHz, which corresponds to the redshift spanning of 0.78 < z < 1.03, computed using HI emission frequency at the rest frame. According to the DESI predicted redshift distribution of number density (DESI Collaboration et al. 2016), shown in right panel of Figure 1, DESI ELG targets keep larger and more complete number density in this redshift range while DESI LRG targets are much less common at that high redshift. It is therefore ideal to cross-correlate between HI emission from Tianlai and the DESI ELG observations.

2.2 Creating ELG and HI mock catalogs

The Horizon Run4 (HR4) (Kim et al. 2015) simulation is used to generate both HI intensity mapping and galaxy distribution mocks. This particle mock is a high-resolution ΛCDM simulation that evolves the distribution of 6300^3 dark matter particles in a periodic cubic box of 3150 h⁻¹ Mpc on a side. It adopted a standard ΛCDM cosmology in concordance with WMAP 5-year results, and each particle has a mass of 9 \times 10^5 M_☉. Specifically, the matter, baryonic matter, and dark energy densities are Ω_m = 0.26, Ω_b = 0.044, and Ω_λ = 0.74, respectively. The current Hubble expansion is...
Toy model for testing the angular power spectrum using the 21cm maps contaminated by different level foregrounds. The contamination level is parameterized by the ratios of the temperature variance between foreground and 21cm signal (see Equation (9)).

In the upper panels, the red squares and blue circles correspond to HI × HI and HI × ELG of realistic temperature map with mixture of HI and foreground noise. For comparison, the black dashed and solid lines show the corresponding pure power spectrum with no foreground noise on the map. Note that the pure power spectrum are same in the different panels. In the lower panels, the blue circles show the normalized difference of $C_\ell$ between the contaminated and the pure HI × ELG. The black dashed lines indicate the normalized 1σ error range of the pure cross power spectrum. Note that the units for auto and cross $C_\ell$ are different, as indicated in the legend.

$H_0 = 72\text{ km s}^{-1}\text{Mpc}^{-1}$ and the clustering amplitude of matter on scales of $8\ h^{-1}\text{Mpc}$ is $\sigma_8 = 1/1.26$. In addition to the box volume, HR4 team also built a past lightcone space data of halos that covers the all-sky up to $z \approx 1.5$. Below we present the way to make ELG galaxy and HI intensity mocks.

The statistics for distribution of dark matter particles is assumed to be linked to those of galaxies. The statistical methods to match the observed galaxy properties to those of dark matter halos have been developed, and are quantitatively expressed by the halo mass function and the halo bias tracing. While galaxies are gravitationally bound to dark matter halos and their evolution is tightly correlated with host halo, there are a plenty of missing elements which are not presented in the halo model alone. Some, but not all, galaxies exhibit star formation quenching to be passive, which causes a bimodal blue or red distribution in the galaxy population, with few found in between. Those galaxies increase with host halo mass and become more frequent at the present time. In this manuscript, we adapt the constraints on relation between host dark matter halos and galaxies which are measured by observation of peak location of the stellar to halo mass ratio using the combination of CFHTLenS and VIPERS (Coupon et al. 2015). Among central and satellite galaxies populated for a given host halo, we identify the satellite galaxies with their host halo masses above the threshold bound $10^{12} \ h^{-1}\text{M}_\odot$ as DESI ELG. The redshift profile is presented in right panel of Figure 1 (blue), which is made by trimming it to fit to the DESI forecast of the ELG target distribution.

The HI distribution is simulated by assigning HI to the halo according to the halo model, which estimates the mass of neutral hydrogen from the host halo mass (Barnes & Haehnelt 2015; Padmanabhan et al. 2016; Padmanabhan & Refregier 2017). The 21cm brightness temperature, $T_h$, can then be modeled from the HI density field (Battye et al. 2013; Bull et al. 2015). In practice, we stack the hydrogen mass hosted by the corresponding halo mass in each cube defined by an angular pixel and a redshift bin. The corresponding hydrogen mass, $M_H$, is used to generate the temperature maps. A suite of foreground maps for each frequency bin of our mock catalogues have to be added, for which we used he Global Sky Model (hereafter GSM) de Oliveira-Costa et al. (2008); Zheng et al. (2017). The GSM model maps include information from five different foregrounds: synchrotron, free-free, CMB, warm dust and cold dust.

Finally, the instrument noise is estimated by assuming uncorrelated thermal noise across all baselines and frequencies. The noise level (hereafter RMS) is represented in units of $\text{K}$. Below we present the way to make ELG galaxy and HI intensity mocks.

$$\sigma_{ij} = \left( N_{ij} \right)^{1/2} = \frac{T_{\text{sys}}}{\Delta \nu \Delta t_{ij}} \left( \frac{I^2}{A_e} \right),$$  

where $\Delta t_{ij}$ is the total integration time of baseline $ij$, $T_{\text{sys}}$ is the system temperature, $A_e$ is the effective area of antenna,
\( \lambda \) is the observing frequency, and \( \Delta \nu \) is the width of the frequency channel. The system temperature is the sum of the sky brightness and the analog receiver noise temperature, \( T_{\text{sys}} = T_{\text{sky}} + T_{\text{rec}} \). At the frequency of interest of 700–800MHz, the Tianlai array would be expected to achieve a total system temperature of 50 – 100 K, and thus we assume \( T_{\text{sys}} = 50 \text{K} \) in this study. We also assume two full years of observation for the Tianlai pathfinder survey. The effective antenna area \( A_e \) is calculated by \( A_e \Omega = \lambda^2 \), where the beam solid angle \( \Omega \) is well approximated by \( \Omega = 0.1 \) for the current Tianlai cylinder array. Now the observed temperature is given as the one that combines the cosmological signal from the simulation with the foregrounds and the observational noise, 

\[
T_{\text{obs}}(\hat{\lambda}) = T_b^{\text{HI}}(\hat{\lambda}) + T_b^{\text{foreground}}(\hat{\lambda}) + T_b^{\text{noise}}(\hat{\lambda}).
\]  

(2)

The foreground signal computed using the GSM is much bigger than the cosmological signal \( T_{21} \), and so the two cannot be simply decomposed. Instead reconstruction methods are developed to split all different types of foreground emissions caused by diverse origins. The local signal and that of cosmological origin will have distinguishable frequency dependence, which can be reconstructed using, for example, methods such as fast independent component analysis (FastICA), principal component analysis (hereafter PCA), and log-polynomial fitting. In this manuscript, we exploit the PCA and FastICA methods to decompose the foreground noise from the cosmological signals.

Our descriptions of these two methods match those in Asorey et al. (2020), but to summarise again:

**PCA:** The pre-whitening method is applied to subtract the mean of simulated data. The covariance matrix of the data, after the pre-whitening procedure, is computed among different data bins in frequency space. This matrix in frequency space is decomposed into eigenvectors with distinct eigenvalues. It is assumed that the foregrounds dominate the eigenmodes with the highest eigenvalues, as those are highest amplitude components of the power in the maps. It is observed that most foreground power exhibits a smooth curve in frequency space that is normally described by a combination of a few leading eigenmodes. In the PCA approach, those leading principal components with the largest eigenvalues in frequency space are projected from every spatial pixels to obtain foreground cleaned maps. There is a caveat that some correlations are cosmologically introduced as well, despite the smallness of the cosmological signal, which causes this cleaning process to affect the 21cm signal slightly.

**FastICA:** The fastICA method assumes that the maps can be decomposed into a set of signals with some non-Gaussian distribution and some Gaussian noise. The non-Gaussian components should correspond to the foregrounds, which should be well behaved and continuous in frequency space. In contrast, the intensity of the cosmological 21cm emission depends on the mass of neutral hydrogen present in the ‘voxel’, which is a stochastic quantity with a Gaussian distribution, and so resembles the noise in a fastICA reconstruction process. Using the implementation of fastICA as part of the scikit-learn python machine learning package (Pedregosa et al. 2011), we maximise the negentropy, defined by \( J(y) = H(y_{\text{gauss}}) - H(y) \), assuming the negentropy is approximated by a log cosh(y) function. As a measure of distance from gaussianity for the negentropyn functions, maximising it with respect to the components should remove the foreground signal, leaving behind the Gaussian cosmological signal.

### 3 LARGE-SCALE HI DISTRIBUTION

The mock HI signal embedded on large scale structure is probed more efficiently using a cross-correlation with a galaxy sample, over the same redshift range. We will describe measuring the clustering using the angular power spectrum, and investigate how well the HI clustering signal can be recovered from the foreground-cleaned map, and focus on
making a comparison between auto-power and cross-power spectrum.

### 3.1 Angular power spectrum

For the galaxies distribution, we use number density field, \( n_g \), where resolved galaxies can be counted in pixels within a redshift bin, and then calculate the over-density, \( \delta_g \), as,

\[
\delta_g(a, \delta) = \frac{n_g(a, \delta)}{\bar{n}_g} - 1. \tag{3}
\]

Here \( a \) and \( \delta \) respectively are the right ascension and declination in the equatorial coordinate system, and \( \bar{n}_g \) is the averaged galaxies number value over the map. For the 21cm maps, we compute the temperature fluctuations as,

\[
\delta T_a(a, \delta) = \bar{T}_a - T_a. \tag{4}
\]

Next, we measure the angular power spectrum by decomposing the fluctuations into spherical harmonics in this way,

\[
\delta(h) = \sum_{\ell=m}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(h). \tag{5}
\]

The harmonics coefficients \( a_{\ell m} \) describe the amplitude of the fluctuations in spherical harmonics space. Note that \( \delta(h) \) corresponds to \( \delta_g(a, \delta) \) or \( \delta T_a(a, \delta) \). Then, the angular power spectrum is calculated by,

\[
C_{\ell}^{XY} = \frac{1}{2\ell + 1} \sum_{m=0}^{\ell} \left| a_{\ell m}^{X} \right| \left| a_{\ell m}^{Y} \right|^2, \tag{6}
\]

where \( X \) and \( Y \) are 'g' and 'H' corresponding to the galaxy and HI, respectively. Meanwhile, the error can be estimated by,

\[
\Delta C_{\ell}^{XY} = \sqrt{\frac{1}{2\ell + 1} \Delta f_{sky} \left( C_{\ell}^{XY} \right)^2 + C_{\ell}^{XX} C_{\ell}^{YY}}^{1/2}, \tag{7}
\]

where \( \Delta f_{sky} \) is the total observed sky fraction. Note that for auto power spectrum, the error equation is reduced to

\[
\Delta C_{\ell}^{XX} = \sqrt{\frac{2}{2\ell + 1} \Delta f_{sky} C_{\ell}^{XX}}, \tag{8}
\]

which is verified by computing dispersion of the corresponding azimuthal modes with the given \( l \). To measure the angular power spectrum we used the NaMaster\(^1\) code (Alonso et al. 2019) with \( \Delta f_{sky} = 25 \), which uses the pseudo-Cl (also known as MASTER) approach including the effect of the sky mask. In the measurements we also find that our results are essentially insensitive to the bin width for reasonable choices.

### 3.2 Threshold limit for residual foreground noise

The foreground signal dominates over the HI temperature by 5 orders of magnitude, rendering the 21cm clustering signal to be nearly invisible. Those foreground sources can be cleaned to the level that is close to the HI signal. However, there would be residual foreground on the map, since those sources are tangled with the 21cm signal in a nearly inseparable manner. Before applying the foreground removal method for the mock maps, we would like to validate how the foreground noise affects the 21cm clustering signal, and seek a guideline threshold level for the cleaning procedure.

The residual foreground noises are assumed to be remaining mixed with HI signal. The HI mock maps with full foreground noises is repainted with the sub-level residual noises. Here the variance of noise temperature variance is simply reduced coherently. The contamination level is parameterized by the ratio of the temperature variance between foreground and 21cm signal as,

\[
\beta = \frac{\langle (T_f - H)^2 \rangle}{\langle (T_H - H)^2 \rangle}, \tag{9}
\]

where, note that, \( \beta = 0 \) corresponds to the pure HI map with no foreground noises. Figure 2 shows the comparison between the auto-correlation of HI × HI and the cross-correlation of HI × ELG. Here color symbols represent \( C_l \) power spectrum of realistic temperature map with mixture of HI and foreground noise, and black curves represent the case with no foreground noise limit. Here we have four different-level contaminated maps, corresponding to \( \beta = [0.5, 1, 5, 10] \), as indicated in different columns. Note that the power spectrum with no foreground noise are same in the different panels.

The HI × HI power spectra with foreground noise exhibit the biased result against the power spectrum at no foreground noise limit. Although the foreground noise is mixed with much smaller sub-level than HI signal of \( \beta = 0.5 \), the contamination is not ignorable at large scale. The effect of non-trivial auto correlation of foreground noise is clearly visible so that we are not able to access to underlying structure formation of particles. It is not likely that the foreground noise would be cleaned under the level of \( \beta = 0.5 \), and it makes us difficult to probe cosmological information using this auto-correlation.

The cross-power spectra of HI × ELG exhibit the better performance to probe HI signal than auto HI × HI spectra, with the remaining residual foreground contamination. While the foreground noise is strongly self-correlated in HI × HI, it does not correlate with galaxy sample collected much higher redshift in which HI signal is radiated. Because foreground noise does not trace the underlying particle density fields which both HI signal and galaxies do, no contribution of the cross term between foreground noise and galaxy is presented in HI × ELG. To see precisely the performance, we plot the normalized difference of \( C_l \) between the contaminated and pure HI × ELG in the lower panels of Figure 2. The black lines indicate the normalized 1σ error range of the pure cross power spectrum, where error is computed by equation (7). One can see that the contaminated keeps consistent to the pure within 1σ level in \( \beta = 0.5 \) & 1 cases. In \( \beta = 5 \) & 10 cases, although the contaminated becomes to fluctuate outside 1σ level, there is no obvious bias to the pure cross. This indicates that, in the case of HI × galaxy cross-correlation, the foreground contamination could have no influence on the accuracy but impacts on the precision only.

In order to evaluate the precision influence by the foreground in further step, we compute the signal-to-noise ratio

---

\(^1\) Downloaded from https://github.com/LSSTDESC/NaMaster.
(SNR) as a function of $\beta$ by,

$$S/N(\beta) = \frac{1}{\sum_{\ell_{\text{min}}}^{\ell_{\text{max}}} \left( \frac{C_\ell(\beta = 0)}{\Delta C_\ell(\beta)} \right)^2},$$

where the $C_\ell(\beta = 0)$ is the power spectra measured from the HI map with no foreground noise, $\Delta C_\ell(\beta)$ is the error of the power spectra measured from the HI map contaminated by $\beta$-level foreground noise, and the summation is computed from $\ell_{\text{min}} = 0$ to $\ell_{\text{max}} = 300$. In the left panel of Figure 3, we plot the SNR as a function of $\beta$ for both of the HI $\times$ HI (dashed) and the HI $\times$ ELG (solid) power spectra. As expected, the HI $\times$ ELG retains a larger SNR than the HI $\times$ HI after around $\beta = 1$, since the cross correlation could remove most of the non-correlated noise and keep more information inside. Meanwhile, we found that the SNR is going down rapidly when increasing $\beta$ from 1 to 5. This indicates that it should be possible to significantly increase the sensitivity of detecting HI clustering when cleaning the foreground to $\beta = 1$ level.

### 3.3 Foreground cleaning

The threshold limit of mixture level of foreground noise is presented in the previous subsection. We found that the HI signal is accurately measured even with the residual foreground noise remaining around $\beta \sim 1$, but only when using the HI $\times$ ELG cross-correlation. It is plausible that one might be developed, but presently there is no perfectly selective foreground cleaning that will leave no damage to the HI signal. The information of HI signal and foreground noise is entangled in a manner to not be perfectly separable. It is expected that foreground cleaning to reduce $\beta$ will cause some significant damage to the HI signal. In this subsection, the practical cleaning methods such as PCA and FastICA are applied to the simulated map, to investigate the level of information loss and the impact on the detectability of HI signal. In addition, the instrument noise is added on to make the prediction more realistic.

The frequency range of Tianlai experiment will be $f = 700 – 800\text{MHz}$ with the frequency resolution of $df = 1\text{MHz}$. It is split into 10 frequency bins, and PCA method is applied on this 10 principal frequency components. In the right panel of Figure 3, the eigenvalues of $10 \times 10$ covariance matrix of those principal components are presented.

As shown, the first two modes capture more than 99 per cent of the information, dominating the whole sky. Such modes can be regarded as the foreground-dominated modes, so that a removal of them should mostly leave the foreground noise out and keep the 21cm-dominated signal in maps with high fidelity. In addition, to make sure the foreground is removed at the right level, we recovered the 21cm distribution by removing different number modes from the original map, and compare the recovered power spectrum to the cases with no foreground. Results are shown in Figure 4. Red and blue symbols represent the recovered HI $\times$ HI and HI $\times$ ELG,
Figure 5. The measured angular power spectrum as a function of multipole number $\ell$ from the Tianlai simulated and reconstructed 21cm maps with two modes removed. The first two panels are measured from maps which do not include the Tianlai instrument noise, while the third and fourth do include it. (Upper panels) Angular power spectrum in the maps with two modes removed. The first two panels are measured from maps which do not include the Tianlai instrument noise, while other cases results in a systematic error in the inferred map. The same conclusions are also found in upper panels of Figure 5, we plot the recovered intensity maps, with foregrounds removed using the first two modes only. In upper panels of Figure 5, we plot the recovered power spectra are closest to the ones with no foreground when removing the first two or three modes. Just removing the first mode leaves too much residual foreground on the sky, which makes the recovered HI × HI to be over-estimated and much larger than the true spectra, and the recovered HI × ELG cross-correlation is also too large, and has fluctuations. Removing the first four modes removes too much of the 21cm signal, resulting in a significantly under-estimated power spectrum, compared to the true values. We thus conclude that the PCA can adequately recover the 21cm distribution when removing the first two or three modes, while other cases results in a systematic error in the inferred map. The same conclusions are also found using FastICA.

Our analyses below are based on the reconstructed 21cm intensity maps, with foregrounds removed using the first two modes only. In upper panels of Figure 5, we plot the recovered HI × HI (red square) and HI × ELG (blue circle) power spectrum, compared to the true ones with no foreground noise (lines). We first present the results without including Tianlai instrument noise in the first two columns, which are based on the PCA and the FastICA, respectively. Both recovered HI × HI and HI × ELG are quite close to their corresponding true cases, indicating most of the foreground noise is already removed from the map. Though, there is a systematic bias in both recovered cases. To see clearly the bias level, the normalized difference of $C_\ell$ between the recovered and the true power spectrum are shown in the lower panels of Figure 5. The black dashed lines are shown for the $1\sigma$ error ranges (computed by Eq. (7)) of the true HI × HI, similar with the one for HI × HI. Clearly, the bias level is significantly out of $1\sigma$ range for both recovered power spectrum. This is expected, since the the foreground and the 21cm signal is tangled together and there would always be a signal loss in the foreground cleaning procedure. In this case, we couldn’t expect the cross correlation would be immune to the bias by the residual foreground, like that in the toy model. There is an inescapable bias caused by the signal loss in the cross case as well.

However, on the other hand, the pattern of bias introduced by the contamination is obviously different in the recovered auto and cross power spectrum. We can see clearly that the amount of offset between the true and measured values is changing with scale in the auto power spectra, but not in the case of cross-correlation. To quantify the amount of offset as a function of the scale, we then defined a bias factor relating the recovered to the true power spectra by,

$$C_\ell^{\text{Recover}} = b^2 C_\ell^{\text{True}}.$$  

where the true $C_\ell$ gives the power spectra without any noise.
lists the best-fitting bias and the corresponding reduced χ^2 when no instrument noise added. As shown in the cross case, the reduced χ^2 values for the best-fitting $b$ are quite small and close to 1. This means the recovered cross power spectra could match the true one using a linear bias. However, in the auto-power spectra case, the reduced χ^2 values are much larger, indicating the recovered auto-power spectra cannot be related to the true one by a linear bias, and the offset should depend on the scale.

It seems from these simulations that the recovered HI × ELG is just linearly biased to the true cross power spectra. This would be a very useful result for the modeling of the cross-power spectra, as we already have the linear galaxy clustering bias as part of the model for that kind of data. Thus, the cross correlation offset caused by the signal loss can be easily parameterized as a linear bias, which can be merged into the linear clustering bias parameter. However, for modeling the recovered HI × HI auto power spectra, the scale-dependent bias would be a difficult nuisance parameter.

In addition, to make the prediction more realistic, we add the Tianlai instrumental thermal noise on the map by assuming two-years observation (see Section 2.2 for details). We repeat the PCA and FastICA on the map including the instrument noise to reconstruct the 21cm distribution, and measuring the recovered HI × HI and HI × ELG power spectra. Results are shown in the last two columns of Figure 5. We also repeat the fitting using Eq. (11), where the best-fitting bias with its reduced χ^2 is shown in the last two rows of Table 1. As shown, the HI × HI auto power spectra totally deviates from the true one, and so we expect to be unable to extract any useful clustering information. In the cross case, although the recovered HI × ELG suffers more fluctuations, we could still see its overall clustering pattern, following the true one clearly. This is expected since the instrumental thermal noise is also uncorrelated with the galaxies. More interestingly, the quite small χ^2 value for the best-fitting of the cross indicates that the bias here keeps not depending on the scale. This leads a positive indication that we could still have a modeling of the recovered cross power spectra even though the instrument noise existing.

Table 1. SNR, bias fitting and χ^2 per dof for foreground-cleaned map

|                  | Auto      | Cross     |
|------------------|-----------|-----------|
|                  | S/N b χ^2 | S/N b χ^2 |
| PCA (no instr. noise) | 226 1.14 26 162 0.90 0.69 |
| FasctICA (no instr. noise) | 301 0.98 29 190 0.77 1.01 |
| PCA (with instr. noise) | 81 1.52 322 64 0.89 1.80 |
| FasctICA (with instr. noise) | 83 1.53 306 68 0.76 1.79 |

4 DETERMINATION OF COSMIC DISTANCE

The baryon acoustic oscillations (hereafter BAO) are an acoustic peak structure caused by the tension between gravitational infall and outward radiative pressure of the baryon-photon fluid that was imprinted on the large scale structure at the surface of last scattering. The peak structure appearing on correlation function and 3D power spectrum can be (and has already been) exploited to measure cosmic distance dubbed as standard ruler of the Universe. These peaks are less visible in 2D projected power spectrum (when the redshift bins are wide), as modes along the line of sight are integrated to smooth out peak structure. Although it is a less model-independent feature to measure cosmic distance, it still provides useful tool (as shown, for example, in Seo et al. (2012)). This broadband BAO fitting requires modelling the non-linear clustering and galaxy biases, and marginalising over the common coherent effects in terms of nuisance parameters. In this section, we explain the appropriate treatment to measure cosmic distance through HI × ELG cross-correlation.

4.1 Theoretical formulation for 2D broadband power spectrum

We now turn to fit the BAO feature by modeling the angular power spectrum. First, we consider the model for the tracer galaxy as a general case, which is easily applied to other tracers. The angular power spectrum of galaxies can be computed by,

$$ C_{XY}^{\ell} = \frac{2}{\pi} \int d\ell k^2 \left[ W_X^Y(k) \right] \left[ W_Y^X(k) \right]. $$

where $X$ and $Y$ denotes tracers, such as 21cm or ELG galaxy, and the kernel function is,

$$ W_X^Y(k) = \int dz \phi(z) \sqrt{P_{XX}(k,z)J_0(kr)} \int dz \delta_X D(z) \sqrt{P_{XX}(k,z)} J_0(kr). $$

where $P_{XX}(k)$ is the 3D power spectrum of tracer $X$. Here $r$ is the comoving distance along our past light cone, $\phi(z)$ is the radial selection function, $D(z)$ is the linear growth factor relative to $z = 0$ and $b(z)$ is the galaxies bias factor.

Both 21cm signal and ELG galaxy trace the underlying particle clustering, and the tracing pattern can be formulated using local bias model exploiting non-linear and non-local halo density bias model $\delta_X = \delta_X(\delta_m)$. This functional form is expanded as (McDonald & Roy 2009),

$$ \delta_X(\delta_m) = b_1 \delta + \frac{1}{2} b_2 \sigma^2 + \text{higher order terms}, $$

where $b_1$ is the linear bias parameter, $b_2$ is the second-order local bias parameter and the term $\sigma_2$ is introduced to ensure the condition $\langle \delta_X \rangle = 0$. In this test, the higher order bias model is simplified to ignore non-local bias parts. Then power spectrum of both tracers are written as,

$$ P_{XX}(k) = \frac{P_{\delta \delta}^2(k)}{P_{\delta_m}^2(k)} \left[ b_1^2 P_{\delta \delta}^2(k) + b_2^2 P_{\delta_m}^2(k) \right] P_{\delta \delta}(k). $$
Here $P_{\delta\delta}$ is the auto power spectra of dark matter density, $\delta$, and $P_{\delta\sigma}$ is the tracers-$\delta$ cross power spectra. $b_1$ is the linear bias parameter and $b_2$ is the second-order local bias parameter. Then, Equation (13) changes to,

$$W^X_\ell = \int dz \phi(z) D(z) \left[ b_1^X(z) \sqrt{P_{\delta\delta}(k)} + b_2^X(z) \sqrt{P_{\delta\sigma}(k)} \right] j_\ell(kr),$$

where $P_{\delta\sigma}(k) = P_{\delta\sigma}^2(k)/P_{\delta\delta}(k)$. Combined with Equation (12), we have,

$$C^X_\ell = C^{XY}_{b_1} + C^{XY}_{b_2}$$

where,

$$C^{XY}_{b_1} = \frac{2}{\pi} \int dk k^2 P_{\delta}(k) \left[ W^X_{\ell}(b_1(k)) \right]$$

$$W^X_{\ell}(k) = \int dz \phi(z) D(z) b_1^X(z) j_\ell(kr)$$

and,

$$C^{XY}_{b_2} = \frac{2}{\pi} \int dk k^2 P_{\delta\sigma}(k) \left[ W^X_{\ell}(b_2(k)) \right]$$

$$W^X_{\ell}(k) = \int dz \phi(z) D(z) b_2^X(z) j_\ell(kr)$$

Thus, we have a non-linear bias modeling for the angular auto power spectrum of galaxies. However, when applying this model to the angular cross power spectrum between HI and galaxies, we have an additional procedure that needs to be considered. As the 21cm observation is to measure the aggregate emission from many unresolved galaxies, we need to have a smoothing window function in Equation (17). We then multiply $C^X_{\ell b_1}$ and $C^X_{\ell b_2}$ by a Gaussian smoothing kernel $W_\ell(\theta_{sm}) = \exp \left[-(\theta_{sm})^2/2\right]$, where $\theta_{sm}$ is the smoothing angle radius. Since we are working with pixelised HI and galaxy data, in the form of HEALPIX maps (Górski et al. 2005), it is highly likely that this smoothing scale will be related to the pixelisation scale.

In left panel of Figure 6, we make a comparison of HI × ELG between the measurement and the model. Note that, as discussed below, the angular power spectrum are normalized to the mean temperature of the reconstructed 21cm maps. The blue circles are the measurements obtained by using the 21cm map without adding the foreground and instrument noise. The red solid line is the prediction of the model, by Equation (17), which is fitted to the measurement on scale 50 < $\ell$ < 767. As shown, they have a great agreement with each other from large to small scale. To show the performance of the smoothing, we also plot the predictions of the model based on the linear (green dash-dot) and non-linear (black dashed) power spectrum of dark matter without smoothing. The nonlinear one fails prediction on both large and small scales. Although the linear one recovers well on large scale, it still overestimates the values on small scale. Therefore, in our paper, we apply the model based on the non-linear dark matter power spectrum with the additional smoothing procedure to fit the cross power spectra HI × ELG, and focus on the cosmic distance constraint.

4.2 Probing broadband BAO

The BAO in the radial and tangential directions provide measurements of the Hubble parameter and angular diameter distance, respectively. Here we focus on using angular power spectrum to constrain the angular diameter distance.
Figure 7. Fitting to the recovered ELG × HI. Different panels are for the results in 10 redshift bins, as the median redshift of each bin indicated on each panel. The blue circles are the measurements of the angular cross power spectrum between ELG and the recovered 21cm, where filled (empty) circles indicate positive (negative) values of Cl. The red lines are the best-fitted model.

Figure 8. Best fit and projected distribution of the parameters in two- or one-dimensional space for the redshift bin $z = 1.014$. The red and black contours in the two-dimensional plane correspond to the boundaries of 68% and 95% confidence levels, respectively. The one-dimensional distributions are the marginalized distributions of individual parameters. The vertical black lines indicate the best-fitting values and the 68% confidence region.
The subtended angle, $\theta(z)$, of the sound horizon length is determined by,

$$\theta = \frac{r_s(z_{\text{drag}})}{(1+z)D_A(z)},$$

(20)

where $D_A(z)$ is the proper angular diameter distance, and $r_s(z_{\text{drag}})$ is the sound horizon at the drag epoch which we fix as the fiducial value here. Then, the length scale of the polynomials of degree, $\ell$, in the angular power spectrum is related to the angular diameter distance by,

$$\ell \propto \frac{1}{\theta} = \frac{(1+z)D_A(z)}{r_s(z_{\text{drag}})},$$

(21)

Consequently, assuming an incorrect $D_A(z)$ can result in systematic errors in the polynomials of degree, $\ell$, which turn out to shrink the shape of the angular power spectrum. We can then model such shrinking deviations and constrain $D_A(z)$ by searching for the model that matches the measurement best. Then, we define the rescale parameter as,

$$(1 + \Delta \alpha)\ell \propto (1 + \Delta \alpha)D_A(z),$$

(22)

where we rescale $D_A(z)$ by a factor $1 + \Delta \alpha$ to introduce the error coming from the changes of $D_A(z)$ into the $C_\ell$. As $\ell$ is direct proportion to $D_A(z)$, $\ell$ is also changed by the same rescaling level. We then derive constraints on $D_A(z)$ by searching for the best-fitting $\Delta \alpha$. In order to see how the angular power spectrum is affected by the rescale parameter $\Delta \alpha$, we apply $\Delta \alpha = 0.2, 0.0 \& -0.2$ in the models, which are plotted with red, black and green lines respectively in Figure 6. It is reassuring that the theoretical $C_\ell$ best matches the measurement for $\Delta \alpha = 0$, which is the value that corresponds to the fiducial $D_A(z)$ in the simulation. Adopting $\Delta \alpha$ = 0.2 or -0.2, corresponding to 20% difference with the true $D_A(z)$, results in an inferred bias for the $C_\ell$ that is systematically too high or too low.

We now derive constraints on the related parameters, $\Delta \alpha$, as well as the bias parameters, $b_1$, $b_2$ and $\theta_{am}$. The biases are defined as,

$$b_1 = b_2 g_8 \sqrt{\frac{\Delta HI^{\text{ELG}}}{2}},$$

$$b_2 = 2 g_8 \sqrt{\frac{\Delta HI^{\text{ELG}}}{2}},$$

(23)

where $b_2 g_8$ denotes the coherent offset caused by information loss in foreground cleaning procedure, and $g_8$ denotes the growth factor of particle clustering. Here we do not separately fit ($b_1, b_2, g_8$), because all are coherent parameters not to be distinguishable. Also for keeping the clustering be in the same level in different redshifts, we normalized the angular power spectrum to the mean temperature of the reconstructed 21cm maps in the whole Tialnai redshift range. We then use a Monte Carlo Markov chain (MCMC) method to explore the likelihood function in the multidimensional space. We use the affine-invariant ensemble sampler, known as MCMC hammer and described in Foreman-Mackey et al. (2013) to sample the parameter space. The corresponding $\chi^2$ is defined as,

$$\chi^2 = \sum_{\ell} \left( \frac{C_{\ell}^{\text{obs}}(z) - C_{\ell}^{\text{fid}}(z)/T_m}{\sigma_{\ell}^2(z)} \right)^2,$$

(24)

where $C_{\ell}^{\text{mod}}(z)$ is the power spectrum measured from the map at redshift $z$, with $\sigma_{\ell}(z)$ for the error obtained from Equation (7), $C_{\ell}^{\text{mod}}(z)$ is the model one computed by using

2 The code emcee can be found at https://github.com/dfm/emcee.
Equation (17), with $\ell$ multiplied by the factor $1 + \Delta(\alpha)$, and $T_m$ is mean temperature of the reconstructed 21cm intensity map in the Tianlai redshift range.

As an illustration, Figure 7 show the comparison between the MCMC best-fitted models (red line) and the measurements (blue circles) for the 10 redshift bins, as the median redshift of each bin indicated on each panel. One can see they agree quite well with each other. Figure 8 shows the projected two-dimensional boundaries in the parameter space for the first redshift bin, to which other bins are quite similar. The red and black contour indicate the 68% and 95% confidence levels, respectively. In the top panels, it shows also the marginalized, one-dimensional distribution for each parameter, with vertical dashed lines indicating the mean and 68% confidence regions. Finally, Figure 9 shows the best-fit $\Delta(\alpha)\%$ for each of the 10 redshift bins, with error bars indicating the 68% confidence regions. As one can see, the results are in agreements with the expected values within 1$\sigma$ level. However, the results are not only suffering from the precision error introduced by the residual foreground and instrument noise, but also the accuracy problem by the cosmic variance in different redshift slices. At last, we forecast a result for the whole redshift bin by averaging the results from the 10 bins, as shown with the red star in Figure 9. Our mean results imply that $\Delta z = 0.008 \pm 0.033$, indicating that the constraint on $D_A$ can be to 3.3% level.

Moreover, in order to show the ability of constraint from the cross correlation, in Figure 10, we compare our forecast on $D_A$ to the measurements from previous galaxy survey, based on the galaxy clustering only. These include the results from BOSS DR12 (Alam et al. 2017), CMASS DR12 (Chuang et al. 2017), eBOSS DR14 (Bautista et al. 2018), DES photometric redshift survey (The Dark Energy Survey Collaboration et al. 2017), and DECam Legacy Survey (DECaLS) (Sridhar et al. 2020). The solid black line corresponds to the theoretical predictions as a function of redshift obtained using the cosmological parameters from (Planck Collaboration et al. 2018). We then conclude that, by using the angular cross-power spectrum between the ELG and the recovered HI, we would be able to put successful constraints on the angular diameter distance $D_A$ to quite good level in the whole redshift bin.

5 CONCLUSIONS

In this paper, we presented a forecast of 21cm clustering measurements for the Tianlai Pathfinder observations, and the effectiveness of cosmological measurements by cross-correlating the 21cm map with the positions of DESI galaxies. Based on the HR4 simulation in the light-cone volume, we generate the mock Tianlai Pathfinder 21cm intensity maps at the frequency range of 700-800MHz, corresponding the redshift spanning of 0.78 $< z < 1.03$. We then added simulated foregrounds and instrument noise, and performed a reconstruction to produce the foreground-cleaned 21cm maps, all of which is described in Asorey et al. (2020). To perform the cross-correlation forecast, we also generated an ELG sample from the same simulation (HR4), which has the same redshift and footprint distribution as is planned for the DESI survey.

We first made a toy model to add different level foregrounds on to the 21cm maps, and validated how the foreground affected the recovered 21cm signal in the HI $\times$ HI auto and HI $\times$ ELG cross power spectrum. Results shows that the contamination is not ignorable in the auto-power spectra at large scales even with a small level of foreground noise, such as amplitude of $\beta \approx 0.5$ times that of signal. However, in the case of the HI $\times$ ELG cross correlation, it has no influence on the accuracy but impacts on the precision only.

In the more realistic case, based on the reconstructed 21cm maps without instrument noise, the bias due to contamination appears in both the HI $\times$ HI and HI $\times$ ELG power spectra, as there are residual foregrounds and signal loss in the cleaning procedure. However, we found that the bias pattern is not same in the two cases, as the recovered HI $\times$ HI is biased a scale-dependent manner, while the recovered HI $\times$ ELG signal is found to be related to the true one by a linear offset. This linear offset in the cross-power spectrum caused by the residual foreground is a useful feature, as it therefore can be easily parameterized as a linear bias in the model. Moreover, when considering the Tianlai instrument noise, although it leaves more fluctuations in the recovered HI $\times$ ELG measurement, we could still clearly see the clustering trend following the true one in a linear pattern. This is expected since the instrument thermal noise map is uncorrelated with the galaxy distribution. However, the recovered HI $\times$ HI measurement is non-linearly changed by both the residual foreground and instrument noise, making it very difficult to obtain useful clustering information.

We then developed a method to model the angular cross-power spectra by considering the linear bias $b_1$, the second order bias $b_2$, and the smoothing parameter $\theta_{sm}$. Based on this model, we were able to recover the HI $\times$ ELG angular cross power spectrum quite well, even though it is contaminated by the residual foreground and instrument noise.

Finally, we applied this model to the mock data to fit the BAO feature in broadband shape of the angular cross power spectrum, with a focus on measuring the angular diameter distance in 10 redshift bins over the Tianlai redshift range $0.775 < z < 1.03$. The results show that the distances can be recovered well at 1$\sigma$ level in all redshift bins. We forecast a constraint of the angular diameter distance for the whole redshift bin by averaging the results from the 10 bins, giving a distance measurement with a precision of 3.3% at that redshift.

ACKNOWLEDGEMENTS

We would like to thank Sungwook Hong, Benjamin L’Huillier, Juhan Kim, and Changbom Park for multiple discussions and for providing the HR4 simulations. We would also like to thank Xuelei Chen and Yidong Xu for providing us details on the Tianlai instrument. F. Shi would like to thank Jiajun Zhang for useful discussions. F. Shi, Y-S. Song, J. Asorey, and D. Parkinson are supported by the project 우주자대구조를 이용한 암흑우주 연구 ("Understanding Dark Universe Using Large Scale Structure of the Universe"), funded by the Ministry of Science. K. Ahn is supported by NRF-2016R1D1A1B04935414 and NRF-2016R1A5A1013277. L. Zhang is supported by the National Science Foundation of China (11621303, 11653003, 11773021, 11890691), the National Key R&D Program
of China (2018YFA0404601, 2018YFA0404504), the 111 project, and the CAS Interdisciplinary Innovation Team (JCTD-2019-05).

This research made use of Astropy,\(^3\) a community-developed core Python package for Astronomy (Astropy Collaboration et al. 2013; Price-Whelan et al. 2018), the HEALPix and Healpy package (Górski et al. 2005; Zonca et al. 2019), the Numpy package Oliphant (2006), the Scipy package (Virtanen et al. 2019) and Matplotlib package (Hunter 2007).

REFERENCES

Alam S., et al., 2017, MNRAS, 470, 2617
Alonso D., Sanchez J., Slosar A., 2019, MNRAS, 484, 4127
Anderson C. J., et al., 2018, MNRAS, 476, 3382
Asorey J., et al., 2020, MNRAS, 495, 1788
Astropy Collaboration et al., 2013, A&A, 558, A33
Barnes L. A., Haehnelt M. G., 2015, MNRAS, 454, 218
Battye R. A., Browne I. W. A., Dickinson C., Heron G., Maffei B., Pourtsidou A., 2013, MNRAS, 434, 1239
Bautista J. E., et al., 2018, ApJ, 863, 110
Bernal J. L., Verde L., Riess A. G., 2016, J. Cosmology Astropart. Phys., 2016, 019
Bull P., Ferreira P. G., Patel P., Santos M. G., 2015, ApJ, 803, 21
Chang T.-C., Pen U.-L., Bandura K., Peterson J. B., 2010, Nature, 466, 463
Chapman E., et al., 2012, MNRAS, 423, 2518
Chen X., 2011, Scientia Sinica Physica, Mechanica &amp; Astronomica, 41, 1358
Chuang C.-H., et al., 2017, MNRAS, 471, 2370
Coupon J., et al., 2015, MNRAS, 449, 1352
Cunnington S., Wolk L., Pourtsidou A., Bacon D., 2019, MNRAS, 488, 5452
Cunnington S., Pourtsidou A., Soares P. S., Blake C., Bacon D., 2020, arXiv e-prints, p. arXiv:2002.05626
DESI Collaboration et al., 2016, arXiv e-prints, p. arXiv:1611.00036
Das S., et al., 2018, in Proc. SPIE. p. 1070836 [arXiv:1806.04698], doi:10.1117/12.2313031
Dinda B. R., Sen A. A., Choudhury T. R., 2018, arXiv e-prints, p. arXiv:1804.11137
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, ApJ, 622, 759
Hu W., Wang X., Wu F., Wang Y., Zhang P., Chen X., 2020, MNRAS, 493, 5854
Hunter J. D., 2007, Computing in Science & Engineering, 9, 90
Kim J., Park C., L’Huillier B., Hong S. E., 2015, Journal of Korean Astronomical Society, 48, 213
Kovetz E., et al., 2019, BAAS, 51, 101
McDonald P., Roy A., 2009, J. Cosmology Astropart. Phys., 2009, 020
Newburgh L. B., et al., 2014, Calibrating CHIME: a new radio interferometer to probe dark energy. p. 91454V, doi:10.1117/12.2055962
Newburgh L. B., et al., 2016, HIRAX: a probe of dark energy and radio transients. p. 99065X, doi:10.1117/12.2234286
Oliphant T., 2006, Guide to NumPy
Padmanabhan H., Choudhury T. R., Refregier A., 2016, MNRAS, 458, 781
Padmanabhan H., Refregier A., Amara A., 2019, arXiv e-prints, p. arXiv:1909.11104
Pedregosa F., et al., 2011, Journal of Machine Learning Research, 12, 2825
Planck Collaboration et al., 2018, arXiv e-prints, p. arXiv:1807.06209
Price-Whelan A. M., et al., 2018, AJ, 156, 123
Santos M., et al., 2015, in Advancing Astrophysics with the Square Kilometre Array (AASKA14). p. 19 [arXiv:1501.03969]
Seo H.-J., et al., 2012, ApJ, 761, 13
Sridhar S., et al., 2020, arXiv e-prints, p. arXiv:2005.13126
The Dark Energy Survey Collaboration et al., 2017, arXiv e-prints, p. arXiv:1712.06209
Thompson A. R., Moran J. M., Swenson George W. J., 2001, Interferometry and Synthesis in Radio Astronomy, 2nd Edition
Virtanen A., Alonso D., Fonseca J., Santos M. G., 2019, MNRAS, 485, 5519
Wolk L., Abdalla F. B., Blake C., Shaw J. R., Chapman E., Rawlings S., 2014, MNRAS, 441, 3271
Xu Y., Wang X., Chen X., 2015, ApJ, 798, 40
Zhao G.-B., et al., 2017, Nature Astronomy, 1, 627
Zheng H., et al., 2017, MNRAS, 464, 3486
Zonca A., Singer L., Lenz D., Reinecke M., Rosset C., Hivon E., Górski K., 2019, Journal of Open Source Software, 4, 1298
de Oliveira-Costa A., Tegmark M., Gaensler B. M., Jonas J., Landecker T. L., Reich P., 2008, MNRAS, 398, 247

\(^3\) http://www.astropy.org