

Calibrating photometric redshifts with intensity mapping observationsDavid Alonso,¹ Pedro G. Ferreira,¹ Matt J. Jarvis,^{1,2} and Kavilan Moodley³¹*Oxford Astrophysics, Department of Physics, Keble Road, Oxford OX1 3RH, United Kingdom*²*Department of Physics, University of the Western Cape, Bellville 7535, South Africa*³*Astrophysics and Cosmology Research Unit, School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, Durban 4041, South Africa*

(Received 7 April 2017; published 16 August 2017)

Imaging surveys of galaxies will have a high number density and angular resolution yet a poor redshift precision. Intensity maps of neutral hydrogen will have accurate redshift resolution yet will not resolve individual sources. Using this complementarity, we show how the clustering redshifts approach proposed for spectroscopic surveys can also be used in combination with intensity mapping observations to calibrate the redshift distribution of galaxies in an imaging survey and, as a result, reduce uncertainties in photometric-redshift measurements. We show how the intensity mapping surveys to be carried out with the MeerKAT, HIRAX and SKA instruments can improve photometric-redshift uncertainties to well below the requirements of DES and LSST. The effectiveness of this method as a function of instrumental parameters, foreground subtraction and other potential systematic errors is discussed in detail.

DOI: [10.1103/PhysRevD.96.043515](https://doi.org/10.1103/PhysRevD.96.043515)**I. INTRODUCTION**

Photometric-redshift surveys are an economic way of building up a detailed map of the large-scale structure of the Universe. By imaging large swathes of the sky, it is possible to construct catalogs of individually resolved galaxies with high number density (and therefore a low “shot” noise). The trade-off for such a large number of objects is the inability to obtain accurate redshift measurements for individual objects. Thus, photometric-redshift surveys are orders of magnitude less resolved in the radial direction than the sparser spectroscopic redshift surveys. The uncertainty in the individual redshifts and in the overall galaxy redshift distribution can severely degrade the constraining power of such data sets for cosmology.

Galaxies cluster to form the cosmic web, and one expects structures in the galaxy distribution to be spatially correlated with structures in any other tracer of the dark matter density. For example, if one has an imaging survey of galaxies (where redshifts are poorly resolved) and a spectroscopic catalog (where redshifts are well resolved), they should have nontrivial cross-correlations; in particular, structures in the imaging survey should be mirrored in the spectroscopic survey. A natural step is to use these cross-correlations so that the precise redshift measurements of the spectroscopic survey can be used to sharpen the photometric redshifts in the imaging survey, or at least calibrate its redshift distribution. These types of methods have been advocated in [1–6], and employed in the analysis of several data sets (e.g., [7–11]).

One does not necessarily have to use a catalog of resolved sources to follow this rationale. In particular, if one can accurately map out, in redshift, any tracer of the

dark matter, it can in principle be used to improve redshift measurements in a sister imaging survey. A notable example is that of an unresolved map of neutral hydrogen, HI, through a technique known as intensity mapping [12–22]. Radio observations at a gigahertz or below will map out the distribution of neutral hydrogen out to redshifts $z \sim 2$ and higher. The neutral hydrogen traces the large-scale structure of the dark matter and thus, inevitably, will be correlated with any other tracer. Maps of HI will be exquisitely resolved in the frequency domain and therefore will map out the density distribution, in detail, in redshift. Although intensity mapping observations will not resolve individual objects, they will be able to achieve sufficient angular resolution for cosmological studies (although this statement depends on the observing mode).

In this paper we explore the use of HI intensity mapping to calibrate photometric-redshift surveys. In particular we show that forthcoming intensity mapping experiments such as those undertaken by MeerKAT [23], HIRAX [24] and the SKA [25] can be used to reduce the uncertainties related to photo- z systematics well below the requirements currently posited by the DES and LSST surveys, thus improving final constraints on cosmological parameters. We structure this paper as follows: In Sec. II we describe the formalism and discuss, in detail, various aspects of the instrumental and observational models we are assuming. Section III presents our results as a function of experimental configuration and foreground uncertainty. In Sec. IV we discuss the prospects of using such a method and compare it with other proposals currently being developed. The appendixes present a number of calculations which are essential for the models considered here.

II. FORMALISM

A. Clustering-based photo- z calibration

Consider two galaxy samples with redshift distributions $\phi_i(z)$ ($i = \{1, 2\}$), and let $a_{\ell m}^i$ be the harmonic coefficients of their projected overdensity of counts on the sky. Their cross-correlation is given by

$$\langle a_{\ell m}^i a_{\ell m}^{j*} \rangle = N_{\ell}^{ij} + S_{\ell}^{ij} \quad (1)$$

$$S_{\ell}^{ij} = \frac{2}{\pi} \int dz \int dz' \phi_i(z) \phi_j(z') \\ \times \int dk k^2 b_i(z) b_j(z') P_m(k, z, z') j_{\ell}(k\chi(z)) j_{\ell}(k\chi(z')), \quad (2)$$

where P_m is the matter power spectrum, χ is the radial comoving distance, $j_{\ell}(x)$ is a spherical Bessel function, N_{ℓ}^{ij} is the cross-noise power spectrum between samples i and j , b_i is the linear bias of the i th sample and we have neglected redshift-space distortions and all other subdominant contributions to the observed power spectrum. In the Limber approximation—where $j_{\ell}(x) \rightarrow \sqrt{\pi/(2\ell+1)} \delta^{\mathcal{D}}(\ell+1/2-x)$ —this expression simplifies to

$$S_{\ell}^{ij} = \int dk P_m(k, z_{\ell}) \frac{H^2(z_{\ell}) b^i(z_{\ell}) b^j(z_{\ell})}{\ell+1/2} \phi_i(z_{\ell}) \phi_j(z_{\ell}), \quad (3)$$

where $\chi(z_{\ell}) \equiv (\ell+1/2)/k$.

For the purposes of this discussion, the most important feature of Eq. (3) is the fact that the amplitude of the cross-correlation is proportional to the overlap between the redshift distributions of those samples. This is especially relevant if one of the samples has good radial resolution, in which case it can be split into narrow bins of redshift. The cross-correlations of all narrow bins with the other sample will therefore trace the amplitude of its redshift distribution, and can effectively be used to constrain it. This is illustrated in Fig. 1, which shows the cross-power spectrum between a Gaussian photo- z bin of width $\sigma = 0.05$ and a set of narrow redshift bins ($\delta z \sim 0.002$). Note that we have only used the Limber approximation in Eq. (3) to more easily illustrate the motivation of the clustering redshifts approach. However, all the power spectra computed for this work, with the exception of the discussion in Sec. III D, were computed without adopting this approximation.

Note also that Eq. (3) implies that the redshift distribution and the redshift-dependent galaxy bias of the photometric sample are completely degenerate in this method, and therefore additional information is needed in order to separate both quantities (e.g., including prior information or lensing data). Since this is an inherent problem of the method, and not specific to the case of intensity mapping, we will simply assume that $b(z)$ is a sufficiently smoothly varying function of z , thus treating intensity mapping (IM) and spectroscopic surveys on an equal footing. The more complicated biasing scheme that arises on small scales also prevents the use of those modes to constrain $\phi(z)$ [4], and therefore one must be conservative when deciding the range of scales to include in the analysis.

Different recipes have been formulated to carry out this kind of analysis, such as the optimal quadratic estimator

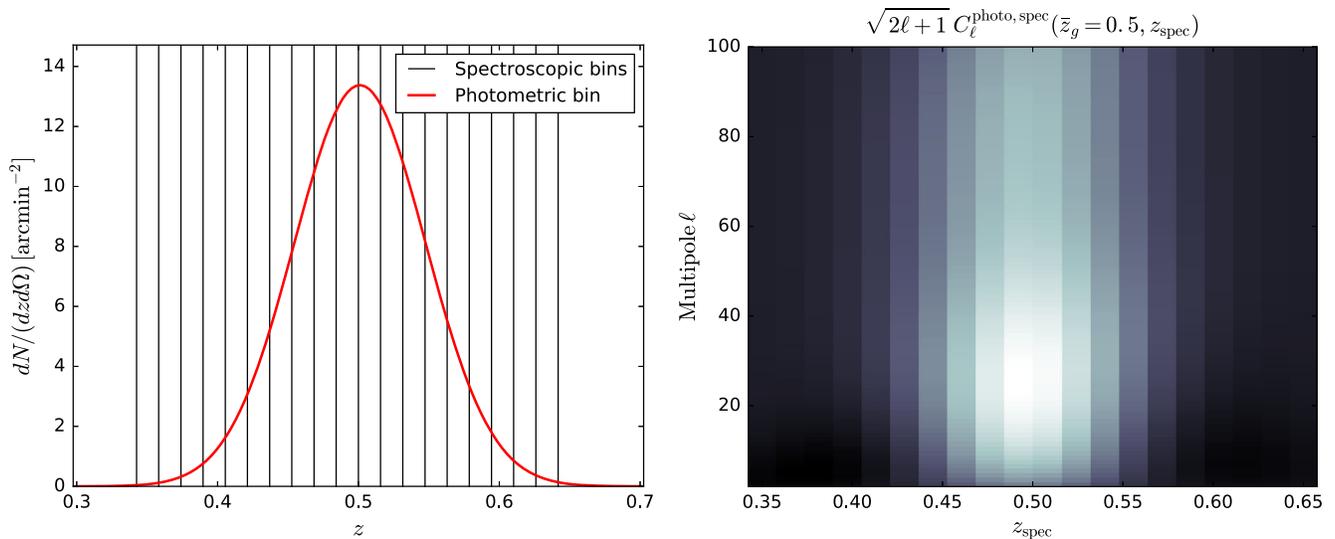


FIG. 1. (Left panel) Example of a redshift bin for a photometric survey and the redshift bins chosen for an overlapping spectroscopic survey. (Right panel) Amplitude of the cross-correlation with an overlapping spectroscopic survey as a function of spectroscopic redshift bin (x axis) and angular scale (y axis). The amplitude of the cross-correlation traces the redshift distribution, and can therefore be used to constrain it.

method of [26]. The forecasts presented here will interpret the redshift distribution (in a parametric or nonparametric form) as a set of extra nuisance parameters, on which we will carry out the Fisher matrix analysis described in Sec. II E. Thus, even though our results will be optimistic in as much as the Fisher matrix saturates the Rao-Cramer bound, they will account for all correlations between redshift distribution parameters and with the cosmological parameters, as well as the presence of redshift-space distortions and magnification bias (effects that have been overseen in previous works).

For the purposes of estimating the ability of future surveys to calibrate photometric-redshift distributions through cross-correlations, we will always consider an individual redshift bin for a photometric sample with unknown distribution, together with a set of overlapping narrow redshift bins of spectroscopic galaxies or intensity mapping observations. Let $N^p(z)$ be the overall true redshift distribution of the photometric sample, and let $p(z_{\text{ph}}|z)$ be the conditional distribution for a photo- z z_{ph} given the true redshift z . Then, the redshift distribution in a photo- z redshift bin b with bounds $z_b^i < z_{\text{ph}} < z_b^f$ is given by

$$\phi_b(z) \propto N^p(z) \int_{z_b^i}^{z_b^f} dz_{\text{ph}} p(z_{\text{ph}}|z). \quad (4)$$

In what follows we will consider 2 degrees of complexity in terms of describing the unknown redshift distribution:

- (1) We will assume Gaussian photo- z s with a given variance (σ_z^2) and bias Δz :

$$\begin{aligned} p(z_{\text{ph}}|z) &\equiv \mathcal{N}(z_{\text{ph}} - \Delta z; z, \sigma_z) \\ &\equiv \frac{\exp\left[-\frac{1}{2} \frac{(z_{\text{ph}} - z - \Delta z)^2}{\sigma_z^2}\right]}{\sqrt{2\pi\sigma_z}}, \end{aligned} \quad (5)$$

and we will assume that the uncertainty in the redshift distribution is fully described by Δz and σ_z .

- (2) We will use a nonparametric form for $\phi_b(z)$, given as a piecewise function with a free amplitude for each spectroscopic redshift bin.

Our assumed fiducial value for Δz and σ_z , as well as the binning scheme used, are described in Sec. II B.

We finish this section by noting that the use of cross-correlations with spectroscopic surveys or intensity mapping observations for photo- z calibration is not limited to the measurement of the redshift distribution of a given galaxy sample, but that they can also be used to improve the precision of photometric-redshift estimates for individual galaxies (e.g., [27]). Although we leave the discussion of this possibility for future work, we describe a Bayesian formalism for this task in Appendix A.

B. Photometric-redshift surveys

This section describes the model used here for a LSST-like photometric-redshift survey. As in [28], we base our

description of the number density of sources and their magnification bias on the measurements of the luminosity function of [29], with k corrections computed with KCORRECT [30]. We assume a magnitude cut of 25.3 in the i band, corresponding to the so-called ‘‘gold’’ sample [31]. Unlike [28], and for simplicity, we will consider a single galaxy population, instead of splitting it into ‘‘red’’ and ‘‘blue’’ sources. The resulting redshift distribution is shown by the solid black line in Fig. 2.

We model the linear galaxy bias as a function of redshift as $b(z) = 1 + 0.84z$, based on the simulations of [32], and quoted in the LSST science book [31].

The photometric-redshift requirements for the gold sample as stated in the LSST science book are $\sigma_z/(1+z) < 0.05$, with a goal of 0.02. Here we have taken a conservative estimate, assuming a standard deviation $\sigma_z = 0.03(1+z)$. We then split the full sample into redshift bins with a width given by $3 \times \bar{\sigma}_z$, where $\bar{\sigma}_z$ is the photo- z variance at the bin center. This binning scheme is chosen to reduce the correlation between bins induced by the tails of the photo- z distribution, and results in the 15 redshift bins shown in Fig. 2 [where the redshift distributions are computed with Eq. (4)]. Our fiducial photo- z model will assume biased Gaussian distributions, fully determined by σ_z and Δz .

C. Intensity mapping

IM is a novel observational technique that circumvents the long integration times needed to obtain reliable spectroscopic redshifts for individual objects through an approach that is transverse to that used by photometric surveys. The idea [13,15,17,21] is to observe the unresolved combined emission of many line-emitting sources in a relatively wide pixel at different frequencies. The

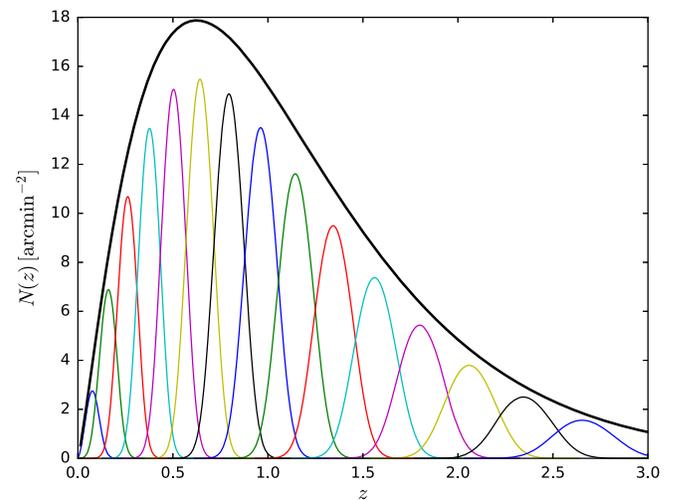


FIG. 2. Angular number density of galaxies as a function of redshift for the LSST gold sample (solid black line). The colored lines show the window functions of the 15 redshift bins considered here.

signal-to-noise ratio of the corresponding line emission is much stronger than that of the individual sources, and thus, combining the intensity measured across the sky and relating the intensity observed at a given frequency to the rest-frame wavelength of the emission line it is possible to produce three-dimensional maps of the density of the line-emitting species. This technique is particularly appealing for isolated spectral lines, as is the case of the 21 cm line caused by the spin-flip transition in HI atoms, and thus HI intensity mapping has been proposed as an ideal method to cover vast volumes at relatively low cost.

A number of experiments have been proposed to carry out IM measurements of the baryon acoustic oscillation scale, such as BINGO [33], CHIME [34], FAST [35], HIRAX [24], SKA [25] and Tianlai [36]. The different instrumental approaches to IM can be broadly classified into two camps:

- (a) Interferometers. The sky emission is measured by a set of antennas, and the measurements of pairs of antennas separated by a given baseline \mathbf{d} are cross-correlated to produce the measurement of an angular Fourier mode with scale $\ell \sim 2\pi\mathbf{d}/\lambda$ (where λ is the observed wavelength). The intensity map is then reconstructed by combining pairs with different baselines.
- (b) Single-dish. In this case the sky emission is measured and autocorrelated by individual antennas. A band-limited intensity map with a resolution $\delta\theta \sim \lambda/D_{\text{dish}}$ is then produced by varying the antenna pointing, where D_{dish} is the antenna diameter.

The expressions for the noise power spectrum for both cases are derived in Appendix C, and can be summarized as

$$N_{\ell}^{\nu} = \frac{T_{\text{sys}}^2 4\pi f_{\text{sky}}}{\eta^2 \Delta\nu t_{\text{tot}}} \begin{cases} \frac{1}{N_{\text{dish}} B^2(\ell)}, & \text{single dish} \\ \frac{\Omega_p}{N_d(\mathbf{d}=\ell\lambda/(2\pi))\lambda^2}, & \text{interferometer.} \end{cases} \quad (6)$$

Here T_{sys} is the system temperature, given as a combination of instrumental and sky temperature (see Appendix C), f_{sky} is the sky fraction covered by the observations, η^2 is the antenna efficiency,¹ $\Delta\nu$ is the bandwidth in that channel, t_{tot} is the total observation time for the survey, N_{dish} is the number of dishes, $B(\ell)$ is the harmonic transform of the antenna beam, $N_d(\mathbf{d})$ is the distribution of baselines and Ω_p is the solid angle covered per pointing. For all experiments discussed here we will assume $\eta = 1$, Gaussian beams so that $B(\ell) = \exp[-\ell(\ell + 1)\theta_{\text{FWHM}}^2/(16 \log 2)]$, and $\Omega_p = \theta_{\text{FWHM}}^2$, where θ_{FWHM} is the beam full width at half maximum, which will approximate as $\theta_{\text{FWHM}} = 1.22\lambda/D_{\text{dish}}$. Note that the baseline distribution N_d is normalized such that

¹ η is defined as the ratio of the effective to real antenna area.

TABLE I. Experimental specifications assumed for SKA, MeerKAT and HIRAX. The baseline distributions for each experiment are described in Sec. II C. Note that the frequency ranges above correspond to the ultrahigh-frequency (UHF) band of SKA and MeerKAT.

Experiment	SKA	MeerKAT	HIRAX
T_{inst}	25 K	25 K	50 K
t_{tot}	10000 h	4000 h	2.8×10^4 h
N_{dish}	197	64	1024 (32×32)
D_{dish}	15 m	13.5 m	6 m
ν range	350–1050 MHz	600–1050 MHz	400–800 MHz
f_{sky}	0.4	0.1	0.4

$$\frac{N_{\text{dish}}(N_{\text{dish}} - 1)}{2} = \int d\mathbf{d}^2 N_d(\mathbf{d}), \quad (7)$$

where $N_{\text{dish}}(N_{\text{dish}} - 1)/2$ is the total number of independent baselines.

Given their expected full overlap with LSST, we will consider here the two main currently envisaged Southern Hemisphere intensity mapping experiments: SKA (and its pathfinder, MeerKAT) and HIRAX. The specifications of these experiments are summarized in Table I.

1. MeerKAT and the SKA

MeerKAT is the 64-dish precursor to the midfrequency component of the SKA. MeerKAT is comprised of 13.5 m dishes and will operate between ~ 550 MHz and 3 GHz using three separate receivers. Although it will predominantly be used as an interferometer, and as such only be sensitive to relatively small spatial scales, there is a proposed project to use MeerKAT in single-dish mode [23]. If such a mode of operation is viable, then MeerKAT will become an extremely efficient intensity mapping facility operating at frequencies that allows the detection of HI to $z \sim 1.5$. Indeed, a proposed open-time survey would provide a several thousand square degree sky survey over the Dark Energy Survey and/or Kilo-Degree Survey areas, which will provide excellent visible wavelength coverage.

In the 2020s, MeerKAT will be enhanced by the addition of 130 15 m dishes to form the midfrequency SKA. Operating at similar frequencies to MeerKAT, the additional 130 dishes will provide much more sensitivity for all science aims, and is capable of carrying out a $\sim 10,000$ deg² intensity mapping survey [25].

As such, both MeerKAT and the SKA will provide a unique view on the HI Universe, and as we will show, can enhance the cosmological science with the LSST with cross-correlations.

2. HIRAX

The Hydrogen Intensity mapping and Real-time Analysis eXperiment (HIRAX) is a proposed close-packed

radio array comprising 1024 6 m dishes disposed in a 32×32 grid and operating at 400–800 MHz. The telescope will be located on the South African Karoo site, which has very low levels of radio-frequency interference in this band, and provides an ideal location to overlap in sky coverage with other planned southern sky cosmological surveys. The large collecting area and field of view provide excellent sensitivity and mapping speed, with the high density of short baselines allowing for sensitive measurements of the baryon acoustic oscillation (BAO) scale in the cosmic HI distribution from redshift ~ 0.8 to 2.5, which in turn will provide competitive constraints on dark energy [24]. HIRAX will make high signal-to-noise maps of 21 cm intensity fluctuations over $15,000 \text{ deg}^2$ (taken to overlap fully with LSST) on cosmological scales of interest, with the relatively high frequency resolution (1024 channels over the 400 MHz bandwidth) allowing for accurate redshift calibration of 21 cm intensity maps. This makes it ideal for calibration of LSST photometric redshifts through the cross-correlation technique.

3. Generic IM experiment

Besides SKA and HIRAX we will also explore the capabilities of a generic intensity mapping experiment in terms of photo- z calibration. The performance of a given experiment is roughly determined by three quantities:

- The range of angular scales over which the noise power spectrum is low enough to probe the cosmological HI emission. This range can be characterized by the minimum and maximum baselines d_{\min} and d_{\max} .
- The noise level (normalized by the bandwidth $\Delta\nu$) σ_T on this range of scales. For a fixed integration time, this is determined by the system temperature T_{sys} and the observed sky area f_{sky} .

Here we will model the effects of the minimum and maximum baselines as a sharp and an inverse-Gaussian cutoff respectively. Thus, our model for the angular noise power spectrum is

$$N_\ell^\nu = \frac{\sigma_T^2}{\Delta\nu} \left[\Theta\left(\frac{\ell\lambda}{2\pi}, d_{\min}\right) \right]^{-1} \exp\left[\ell(\ell+1) \frac{\theta_{\text{beam}}^2}{8 \log 2}\right], \quad (8)$$

where $\theta_{\text{beam}} \equiv 1.22\lambda/d_{\max}$ and $\Theta(x, x_i)$ is 1 if $x_i < x$ and 0 otherwise. Note that by definition σ_T has units of $[\text{mK rad MHz}^{1/2}]$. For comparison, the equivalent values of these parameters that roughly reproduce the noise curves for HIRAX are

$$d_{\min}^{\text{HIRAX}} = 6 \text{ m}, \quad d_{\max}^{\text{HIRAX}} \sim 300 \text{ m}, \\ \sigma_T^{\text{HIRAX}} \sim 10^{-3} \text{ mK rad MHz}^{1/2}.$$

4. Foregrounds

One of the main obstacles that HI intensity mapping must overcome to become a useful cosmological tool is the presence of galactic and extragalactic foregrounds several orders of magnitude larger than the 21 cm cosmological signal [37,38]. Under the assumption that foregrounds are coherent in frequency (as opposed to the cosmic signal tracing the density inhomogeneities along the line of sight), these foreground sources can be in principle efficiently removed using component-separation methods [38–40]. However, instrumental imperfections, such as frequency-dependent beams or polarization leakage, can generate foreground residuals with a nontrivial frequency structure that could strongly bias cosmological constraints from 21 cm data alone. In any case, the removal of frequency-smooth components will introduce large uncertainties on the large-scale radial modes of the 21 cm fluctuations.

Here we have introduced the effect of foregrounds by including an extra component, f , in the sky model for HI accounting for foreground residuals. Thus we will assume that the measured harmonic coefficients at a given frequency ν are given by

$$a_{\ell m}^\nu = s_{\ell m}^\nu + f_{\ell m}^\nu + n_{\ell m}^\nu, \quad (9)$$

where $s_{\ell m}^\nu$ and $n_{\ell m}^\nu$ are the true cosmological signal and the instrumental noise contribution. We will model f as an almost-correlated component with a cross-frequency power spectrum given by

$$C_{f, \ell m}^{\nu\nu'} \equiv \langle f_{\ell m}^\nu f_{\ell m}^{\nu'*} \rangle \\ = A_{\text{FG}} \left(\frac{\ell}{\ell_*}\right)^\beta \left(\frac{\nu\nu'}{\nu_*^2}\right)^\alpha \exp\left[-\frac{\log^2(\nu/\nu')}{2\xi^2}\right]. \quad (10)$$

Here A_{FG} and β parametrize the amplitude of the foreground residuals and their distribution on different angular scales, and α describes their mean frequency dependence. Finally, ξ parametrizes the characteristic frequency scale over which foregrounds are decorrelated. When including the effects of foregrounds (Sec. III C) we will also marginalize over $(A_{\text{FG}}, \alpha, \beta, \xi)$. For α and β we will use the fiducial values $\alpha = -2.7$ and $\beta = -2.4$, corresponding to galactic synchrotron emission [37,41], and we will set $A_{\text{FG}} = 1 \text{ mK}^2$, large enough for the residuals to dominate the equal- ν power spectrum.² We will study the final constraints as a function of ξ . Unless otherwise stated (e.g., in Sec. III C), all our results for intensity mapping experiments assume a fiducial $\xi = 10$, corresponding to the case where the only modes removed are mostly those associated with a fully correlated foreground component.

²We use a pivot scale $\ell_* = 1000$ and a pivot frequency $\nu_* = 130 \text{ MHz}$, as in [37].

Effectively, this extra component cancels the constraining power of all radial modes of the 21 cm fluctuations with comoving radial wave numbers k_{\parallel} below a scale $k_{\parallel}^{\text{FG}}$, related to ξ through

$$k_{\parallel}^{\text{FG}} \sim \frac{\pi H(z)}{c(1+z)\xi}. \quad (11)$$

The effect of foregrounds on the ability to constrain redshift distributions can be readily understood as loss of information in the $k_{\parallel} - k_{\perp}$ space. In the flat-sky approximation, and on linear scales, the angular power spectrum between two tracers i and j of the matter density can be computed as

$$C_{k_{\perp}}^{ij} = \int \frac{dk_{\parallel}}{2\pi} P(k_{\parallel}, k_{\perp}) W^i(k_{\parallel}) W^{j*}(k_{\parallel}), \quad (12)$$

where we have again ignored the effect of redshift-space distortions and

$$W^k(k_{\parallel}) \equiv \int dx_{\parallel} \phi^k(x_{\parallel}) b^k(x_{\parallel}) D(x_{\parallel}) e^{ix_{\parallel} k_{\parallel}}. \quad (13)$$

Here D is the linear growth factor and b^k and ϕ^k are the linear bias and selection function for the k th tracer. Let us assume that i is a photometric-redshift bin and j is a narrow intensity mapping frequency shell with comoving width $\delta\chi$ centered at χ_* . Assuming D and the bias b^j to be slowly varying functions of χ we obtain

$$C_{k_{\perp}}^{ij} = \int \frac{dk_{\parallel}}{2\pi} P(k_{\parallel}, k_{\perp}) W^i(k_{\parallel}) D(\chi_*) b^j(\chi_*) j_0(k_{\parallel} \delta\chi/2). \quad (14)$$

Now, assuming that ϕ^i has support over a wide range of redshifts, corresponding to a comoving width $\Delta\chi$, its Fourier transform ($\sim W^i$) will only have support over wave numbers $k_{\parallel} \lesssim 1/\Delta\chi$. Since the Bessel function j_0 provides support over all values of $k_{\parallel} \lesssim 1/\delta\chi$, and under the assumption that $\delta\chi < \Delta\chi$, the total number of modes that contribute to C^{ij} is bound by $\sim 1/\Delta\chi$. Since foreground contamination will mostly affect large radial modes, eventually a large fraction of this k_{\parallel} range becomes dominated by foreground uncertainties and stops contributing efficiently to the overall signal-to-noise ratio, thus degrading the final constraints on any model parameter.

We finish this section by noting that, as described in Sec. II A, the main constraining power for photo- z calibration comes from the cross-correlation of the photometric and spectroscopic samples. Since the photometric sample would not suffer from foreground contamination, these

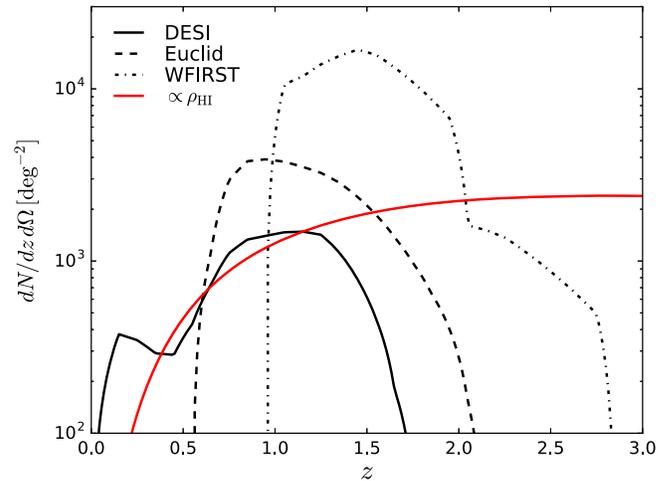


FIG. 3. Angular number density of galaxies as a function of redshift for the three spectroscopic surveys considered here: DESI (solid), Euclid (dashed) and WFIRST (dot dashed). The red solid line shows the redshift dependence of the mean HI density (arbitrarily normalized) for comparison.

cross-correlations are very robust against foreground biasing, which makes photo- z calibration an ideal application of IM.

D. Spectroscopic surveys

In order to showcase the possibility of calibrating redshift distributions through cross-correlation with future intensity mapping experiments we will compare their forecast performance against that of the most relevant future spectroscopic surveys:

- (a) The Dark Energy Spectroscopic Instrument (DESI) [42] is a spectroscopic galaxy survey planned to cover $\sim 14000 \text{ deg}^2$ from its Northern Hemisphere location at Kitt Peak National Observatory. We assume an area overlap of $f_{\text{sky}} = 0.2$ with LSST, and we model the number density and clustering bias of the two galaxy samples considered here (luminous red galaxies and emission line galaxies) as done in [43].
- (b) The Euclid galaxy survey [44] is a space-borne infrared spectrograph that will aim to detect $\sim 5 \times 10^7$ H α -emitting galaxies in the redshift range $0.65 < z < 2$ over $\sim 15000 \text{ deg}^2$. We assume full overlap with LSST, and we model the number density and bias as in [43].
- (c) The Wide Field Infrared Survey Telescope (WFIRST) [45] is a future space observatory in the infrared that will measure redshifts for $\sim 2.6 \times 10^7$ objects over $\sim 2000 \text{ deg}^2$. The deep nature of WFIRST will make it ideal to calibrate the LSST redshift distribution at high redshifts. We model the number density and bias of the WFIRST sample as in [43], and we assume a full overlap with LSST ($f_{\text{sky}} = 0.05$).

Figure 3 shows the redshift distributions for galaxies detected by these three experiments.

E. Forecasting formalism

Our formalism will distinguish between two types of tracers of the density field:

- (a) Spectroscopic. Tracers whose redshift distribution is well known. This would correspond to tracers with good radial resolution such as a narrow redshift bin of spectroscopic sources or an intensity map in a narrow frequency band, as well as other tracers with a well-known window function, such as a cosmic microwave background (CMB) lensing map.
- (b) Photometric. Tracers whose redshift distribution is unknown or uncertain. This would correspond to e.g., a photometric-redshift bin, a radio continuum survey or a map of the cosmic infrared background.

Let us start by considering a set of sky maps corresponding to a number of tracers, and let \mathbf{a} be the corresponding vector of maps expressed in a given basis. In the following sections we will assume that \mathbf{a} is stored in terms of spherical harmonic coefficients and that it takes the form $\mathbf{a}_{\ell m} = (p_{\ell m}, s_{\ell m}^1, \dots, s_{\ell m}^{N_s})$, where $p_{\ell m}$ is a photometric tracer and $s_{\ell m}^i$ is a set of spectroscopic tracers. For the moment, however, we will keep the discussion general.

Assuming that \mathbf{a} is Gaussianly distributed with zero mean and covariance $\hat{\mathbf{C}} \equiv \langle \mathbf{a}\mathbf{a}^\dagger \rangle$, its log-likelihood is given by

$$\mathcal{L} \equiv -2 \log p(\mathbf{a}) = \mathbf{a}^\dagger \hat{\mathbf{C}}^{-1} \mathbf{a} + \log(\det(2\pi\hat{\mathbf{C}})). \quad (15)$$

Now let q_i be a set of parameters modeling $\hat{\mathbf{C}}$, including (but not limited to) the parameters describing the photometric-redshift distribution. A maximum-likelihood estimator for q_i can be defined by using an iterative Newton-Raphson method to minimize Eq. (15). This is described in [26,46,47], and yields the iterative algorithm

$$q_i^n = q_i^{n-1} + [\hat{\mathbf{F}}^{-1}]_{ij} [\mathbf{a}^\dagger \hat{\mathbf{C}}^{-1} \hat{\mathbf{C}}_j \hat{\mathbf{C}}^{-1} \mathbf{a} - \text{Tr}(\hat{\mathbf{C}}_j \hat{\mathbf{C}}^{-1})],$$

$$\hat{\mathbf{F}}_{ij} \equiv \left\langle \frac{\partial^2 \mathcal{L}}{\partial q_i \partial q_j} \right\rangle = \text{Tr}(\hat{\mathbf{C}}^{-1} \hat{\mathbf{C}}_i \hat{\mathbf{C}}^{-1} \hat{\mathbf{C}}_j), \quad (16)$$

where, in Eq. (16) there is an implicit summation over j , the subindex, i implies differentiation with respect to q_i , $\hat{\mathbf{F}}$ is the Fisher matrix, q_i^n is the n th iteration of the solution for q_i and the previous iteration q_i^{n-1} is used to compute $\hat{\mathbf{C}}$ and $\hat{\mathbf{C}}_i$ in the second term. Note that we have simplified a pure Newton-Raphson iteration by taking the ensemble average of the likelihood Hessian (i.e., the Fisher matrix). Furthermore, in the case where the likelihood is well approximated by a Gaussian, $\hat{\mathbf{F}}^{-1}$ is the covariance matrix of the q_i . Equation (16) is the basis of the method proposed in [26] (with a number of simplifications) and used in [11] to constrain the redshift distribution of galaxies in the KiDS survey.

In our case, we mainly care about the uncertainty in the redshift distribution parameters included in the q_i , and therefore we will simply estimate the Fisher matrix $\hat{\mathbf{F}}$. In the case where \mathbf{a} is a set of spherical harmonic coefficients with power spectrum $\langle \mathbf{a}_{\ell m} \mathbf{a}_{\ell' m'}^\dagger \rangle = \delta_{\ell\ell'} \delta_{mm'} \hat{\mathbf{C}}_\ell$, $\hat{\mathbf{F}}$ is given by

$$\hat{\mathbf{F}}_{ij} = \sum_{\ell=2}^{\ell_{\max}} f_{\text{sky}}(\ell + 1/2) \text{Tr}(\hat{\mathbf{C}}_\ell^{-1} \hat{\mathbf{C}}_{\ell,i} \hat{\mathbf{C}}_\ell^{-1} \hat{\mathbf{C}}_{\ell,j}), \quad (17)$$

where we have approximated the effects of a partial sky coverage by scaling the number of independent modes per ℓ by the sky fraction f_{sky} . The form of the power spectra $\hat{\mathbf{C}}_\ell$ for the different tracers considered in this work is given in Appendix B.

As explicitly shown in Eq. (17), smaller-scale modes carry a higher statistical weight (proportional to $\sim \ell$), and would in principle dominate the redshift distribution constraints. The smallest scales are, however, dominated by theoretical uncertainties from nonlinearities in the evolution of the density field and the galaxy-halo connection, and therefore a multipole cutoff ℓ_{\max} must be used to contain the constraining power of systematics-dominated modes. In this paper we use a redshift-dependent cutoff defined as follows. Let z be the mean redshift of a given redshift bin, and let $\sigma^2(k_*)$ be the variance of the linear density field at that redshift on modes with wave number $k < k_*$:

$$\sigma^2(k_*, z) \equiv \frac{1}{2\pi^2} \int_0^{k_*} dk k^2 P_m(k, z). \quad (18)$$

We then define the cutoff scale as $\ell_{\max}(z) = \chi(z) k_{\max}(z)$, where $k_{\max}(z)$ satisfies $\sigma(k_{\max}, z) = \sigma_{\text{thr}}$ for some choice of σ_{thr} . In what follows we will use a fiducial threshold $\sigma_{\text{thr}} = 1$, corresponding to $k_{\max}(z=0) \simeq 0.3 h \text{ Mpc}^{-1}$, and we will study the dependence of our results on this choice. Besides this choice of ℓ_{\max} , we will also impose a hard cutoff for all galaxy-survey and intensity mapping tracers of $\ell < 2000$ [thus, in reality, $\ell_{\max} = \min(\chi k_{\max}, 2000)$].

III. RESULTS

In order to forecast the ability of future experiments to constrain photometric-redshift distributions, in the following sections we will use the formalism described in Sec. II E with a data vector given by $\mathbf{a}_{\ell m} = (p_{\ell m}, s_{\ell m}^1, \dots, s_{\ell m}^{N_s})$, where p is a photometric-redshift bin and s^i are a set of overlapping redshift bins for a spectroscopic tracer (either an intensity mapping experiment or a spectroscopic galaxy survey). The number N_s , width and redshift range of the spectroscopic redshift bins is chosen in order to adequately sample the changes in the photometric-redshift distribution. We choose the redshift bin width to be 33% of

the photo- z uncertainty σ_z , which governs the variability of the redshift distribution (i.e., each redshift interval of σ_z is sampled in three points). In order to sample the tails of the distribution we then define the redshift range of the set of spectroscopic bins as $[z_b^i - 3\sigma_z, z_b^f + 3\sigma_z]$, where z_b^i and z_b^f are the edges of the photometric-redshift bin. The number of spectroscopic redshift bins N_s is then defined in terms of these numbers.

The model parameters q_i in the following sections will be given by

- All of the parameters needed to determine the redshift distribution [σ_z , Δz or the amplitude $N(z)$ in different spectroscopic bins, depending on the case].
- Two overall clustering bias parameters, b_p and b_s , corresponding to the bias of the photometric and spectroscopic tracers.
- We will also include two cosmological parameters: the fractional matter density Ω_M and the amplitude of scalar perturbations A_s , in q_i in order to account for the possible cosmology dependence of the results.

We will change this setup in Sec. III E, where we will explore the impact of the achieved constraints on the photo- z parameters on the final cosmological constraints. In this section \mathbf{a} will correspond to the 15 photometric-redshift bins for LSST, for both galaxy clustering and weak lensing (i.e., 30 sets of spherical harmonics). Likewise q_i will contain the cosmological parameters ($\omega_c, \omega_b, h, w_0, w_a, A_s, n_s, \tau_{\text{reio}}$) as well as all the baseline photo- z parameters (Δz and σ_z for all redshift bins), with

priors corresponding to the constraints found in the preceding sections.

A. Baseline forecasts

Using the formalism described above, and in the simplified scenario of Gaussian photo- z 's, we present, in the left panel of Fig. 4, the forecast constraints on the photo- z bias (Δz) and variance (σ_z) for the key intensity mapping experiments introduced in Sec. II C. In this and all subsequent plots, the thin black solid line shows the LSST requirement of $\sigma(\Delta z, \sigma_z) \lesssim 10^{-3}(1+z)$ [48,49], while the thin dashed line corresponds to the nominal requirement for the Dark Energy Survey (2×10^{-3}) [50]. The colored bars in these and all subsequent plots show the redshift ranges corresponding to the proposed frequency bands of the three IM experiments explored here.

Two key features must be noted in this figure: first, the uncertainties grow steeply at low redshifts. This is due to the reduced number of modes available in that regime, associated with the smaller comoving volume and the impact of nonlinearities on lower values of k . The latter effect is especially severe for HIRAX, given its inability to measure angular modes larger than its beam size. Note, however, that this regime lies outside the proposed frequency ranges for both HIRAX ($0.8 \lesssim z \lesssim 2.5$) and SKA ($0.35 \lesssim z \lesssim 3$).

Secondly, the ratio between $\sigma(\sigma_z)$ and $\sigma(\Delta z)$ stays roughly constant (~ 1.4). This is compatible with the expected ratio between the uncertainties associated with

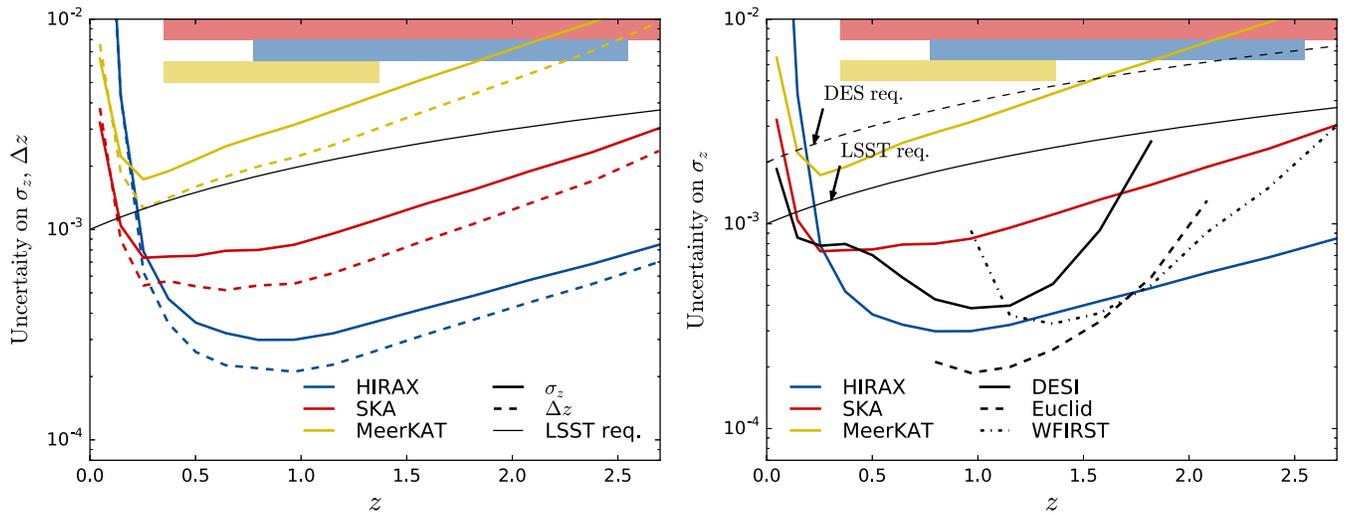


FIG. 4. (Left panel) Forecast 1σ constraints on the photo- z scatter σ_z (solid lines) and bias Δz (dashed lines) for the three IM experiments under consideration: HIRAX (blue), SKA (red) and MeerKAT (yellow). (Right panel) Comparison of the previous three intensity mapping experiments, in terms of the forecast constraints on σ_z , with three future spectroscopic surveys: DESI (black solid), Euclid (black dashed) and WFIRST (black dot dashed). In both panels, the thin solid line shows the photo- z calibration requirement on both σ_z and Δz for LSST, with the corresponding requirement for DES shown as a thin dashed line in the right panel. The colored bands in the upper part of all plots show the proposed frequency ranges for the three 21 cm experiments (same color code). We have assumed using the UHF band for SKA and MeerKAT. The L band would be able to cover all redshifts below $z \sim 0.35$. The colored dashed lines show the predictions in the presence of correlated ($\xi = 10$) foreground residuals (see Sec. III C for further details).

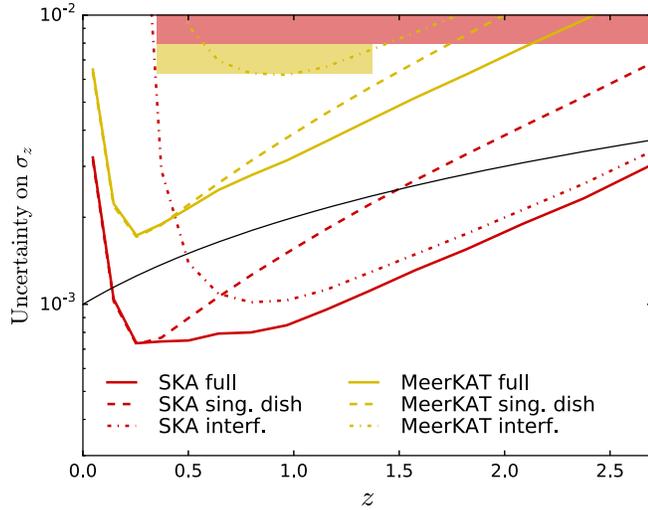


FIG. 5. Forecast constraints on the LSST photo- z scatter σ_z for SKA (red) and MeerKAT (yellow) assuming only interferometric observations (dot-dashed lines), single-dish observations (dashed lines) and both simultaneously (solid lines).

the maximum-likelihood estimates of the mean and standard deviation of a Gaussian distribution from a finite number of samples [$\sigma(\sigma)/\sigma(\mu) = \sqrt{2}$]. This result holds for most of the cases explored here (see Sec. III C for an exception), and thus we have omitted the curves for $\sigma(\Delta z)$ in most of the subsequent figures.

The right panel of Fig. 4 compares the constraints achievable by IM experiments with those forecast for the spectroscopic surveys described in Sec. II D. We see that both SKA and HIRAX would be able to satisfy the LSST requirements over the redshift range of interest. The SKA precursor MeerKAT would fall short except at low redshifts. However, the shorter-term timeline of MeerKAT (2018 onward) would make it an ideal experiment to prove the viability of this technique in cross-correlation with the Dark Energy Survey (DES) [50], particularly in the light of the proposed intensity mapping surveys [23] targeting a full overlap with DES.³

As discussed in [22], the dish size of SKA is not ideal for cosmological observations, since it is not large/small enough to resolve the angular BAO scale sufficiently well in either single-dish or interferometric modes, although single-dish observations are able to address important science cases such as primordial non-Gaussianity [28,51]. Small scales carry a larger statistical weight, however, and it is not clear that a single-dish strategy would also be ideal for the purposes of photo- z calibration. This is explored in Fig. 5, which shows the constraints on σ_z achievable with single-dish (dashed lines) and interferometric (dash-dotted

³Note that the photo- z calibration requirements, defined in terms of the degradation of the final constraints, should be less stringent for DES.

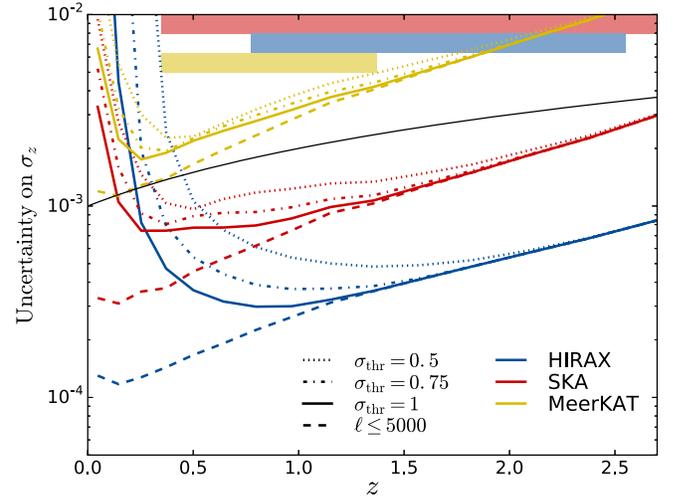


FIG. 6. Dependence of the constraints on the LSST photo- z scatter σ_z on the overdensity variance threshold used to filter out nonlinear scales [see Eq. (18)]. The results are shown for HIRAX (blue), SKA (red) and MeerKAT (yellow). The fiducial value used in this analysis is shown as solid lines, while a more optimistic scenario where all multipoles up to $\ell = 5000$ are included at all redshifts is shown as a dashed line. This case would mostly benefit interferometric observations, given their higher sensitivity on small angular scales.

lines) observations for SKA (red) and MeerKAT (yellow). The constraints from a joint auto- and cross-correlation analysis are shown as solid lines, and correspond to the results reported here. We see that, in the case of SKA, the single-dish mode outperforms the interferometer up to $z \sim 1.4$, when a sufficiently large number of usable modes enter the regime probed by the latter. This suggests that, if simultaneous single-dish and interferometric observations proved to be unfeasible, the photo- z calibration requirements could still be met by using either mode in different redshift ranges.

The performance of this method at low redshift depends crucially on the prescription used to isolate the effect of nonlinearities. Here we have done this in terms of the threshold rms variance σ_{thr} defined in Eq. (18) for a fiducial value of $\sigma_{\text{thr}} = 1$, corresponding to $k_{\text{max}} \sim 0.3 \text{ Mpc}^{-1} h$ at $z = 0$. Figure 6 shows the result of relaxing or tightening this criterion. The effect on SKA and MeerKAT is only moderate, since these experiments gather most of their sensitivity from the large, linear scales in autocorrelation mode. HIRAX, on the other hand, loses sensitivity more rapidly as the scale of nonlinearities removes a larger fraction of the available modes. Nevertheless, even for $\sigma_{\text{thr}} = 0.5$ (corresponding to $k_{\text{max}} = 0.1 \text{ Mpc}^{-1} h$ at $z = 0$) the LSST calibration requirements are satisfied in the redshift range corresponding to the HIRAX frequency band. It is also worth noting that the performance at high redshift is insensitive to the high- ℓ cut, due to the fact that nonlinear scales are noise dominated beyond $z \sim 1.5$ for

all the experiments described here (with the exception of the interferometric parts of SKA and MeerKAT).

B. Dependence on experimental parameters

We have so far quantified the potential of currently proposed experimental configurations for photo- z calibration. The aim of this section is to identify the optimal instrumental specifications for this task.

We start by exploring the balance between noise level and sky fraction, varying the total overlapping area of the three experiments explored in the previous section (SKA, MeerKAT and HIRAX) keeping the total observation time fixed. The result is presented in Fig. 7, which shows the achievable constraints on σ_z when reducing the sky area by successive factors of 2. We find that, although the results are almost insensitive to the reduction in f_{sky} , larger sky areas are always preferred, a reflection of the fact that the constraints are dominated by cosmic variance rather than noise.

As we have discussed in the previous section, the key drawback of single-dish experiments is their inability to access angular scales smaller than the beam size (with their higher statistical weight). Conversely, interferometers are unable to cover scales larger than that probed by their smallest baseline, and therefore they have access to a limited number of reliable (i.e., mildly nonlinear) modes. Using the generic instrument parametrization given by Eq. (8), and the fiducial parameters corresponding to HIRAX, Fig. 8 explores these issues.

The upper panel shows the constraints achievable by a single-dish experiment for different dish sizes. Dishes of at least 15 m (corresponding to the SKA case) are needed in

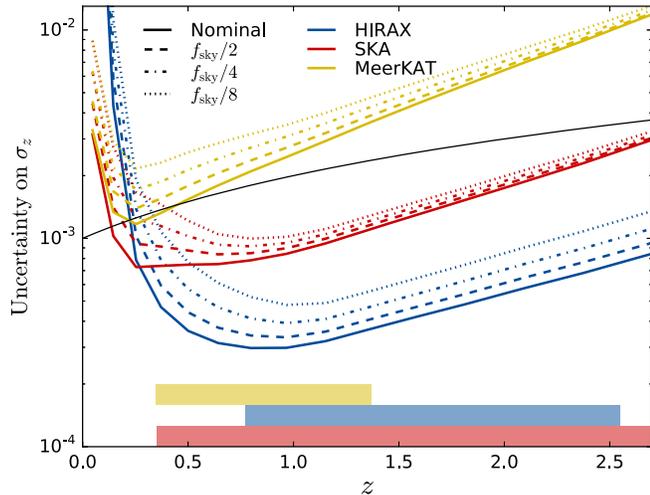


FIG. 7. Dependence of the forecast constraints on the LSST photo- z scatter σ_z on the overlap sky fraction for constant total observation time. In all cases the constraints are mostly insensitive to the trade-off between sky area and noise level, although larger areas are marginally preferred, which reflects the fact that these measurements are mostly dominated by cosmic variance and not noise.

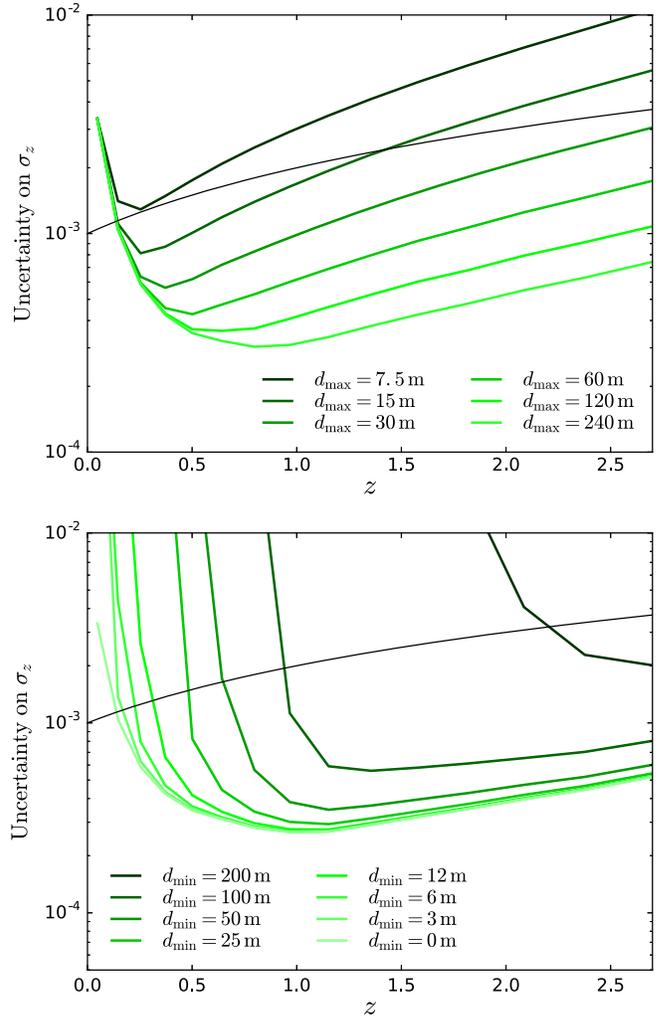


FIG. 8. (Upper panel) Dependence of the forecast constraints on the LSST photo- z scatter σ_z on the dish size for single-dish IM observations. A dish size of at least ~ 15 m is needed to match the LSST requirements throughout the whole redshift range. (Lower panel) Dependence on the minimum baseline for interferometers. Baselines of at most ~ 12 m are needed to successfully constrain photo- z systematics below $z \sim 0.5$.

order to achieve the LSST requirements at all redshifts, while a ~ 100 m dish (e.g., such as the Green Bank Telescope [20]) would be able to achieve constraints similar to those of HIRAX. The largest currently planned experiment is FAST [35], with a dish size of 500 m.

The lower panel of the figure shows the performance of an interferometer as a function of the minimum baseline d_{min} . In this case the main effect is the loss, at lower redshifts, of the large, mildly nonlinear scales. A maximum baseline of at most 12 m is needed in order to calibrate redshift distributions below $z \sim 0.5$ with sufficient accuracy.

C. Foregrounds

As we discussed in Sec. II C 4, the main effect of radio foregrounds for 21 cm observations is to make large radial

scales (i.e., modes with k_{\parallel} smaller than some $k_{\parallel}^{\text{FG}}$) useless. We have parametrized this by introducing an extra component corresponding to foreground residuals characterized by an amplitude A_{FG} and a frequency correlation length ξ [see Sec. II C 4 and Eq. (10) for a full description]. We set A_{FG} to a value large enough to dominate the emission on radial scales larger than the comoving length corresponding to ξ [see Eq. (11)], and study the final constraints as a function of ξ . We have further verified that our results do not vary significantly after increasing A_{FG} by an order of magnitude with respect to its fiducial value.

Figure 9 shows the result of this analysis: while sufficiently coherent foregrounds ($\xi \gtrsim 1$) do not qualitatively modify the final results for the three experiments

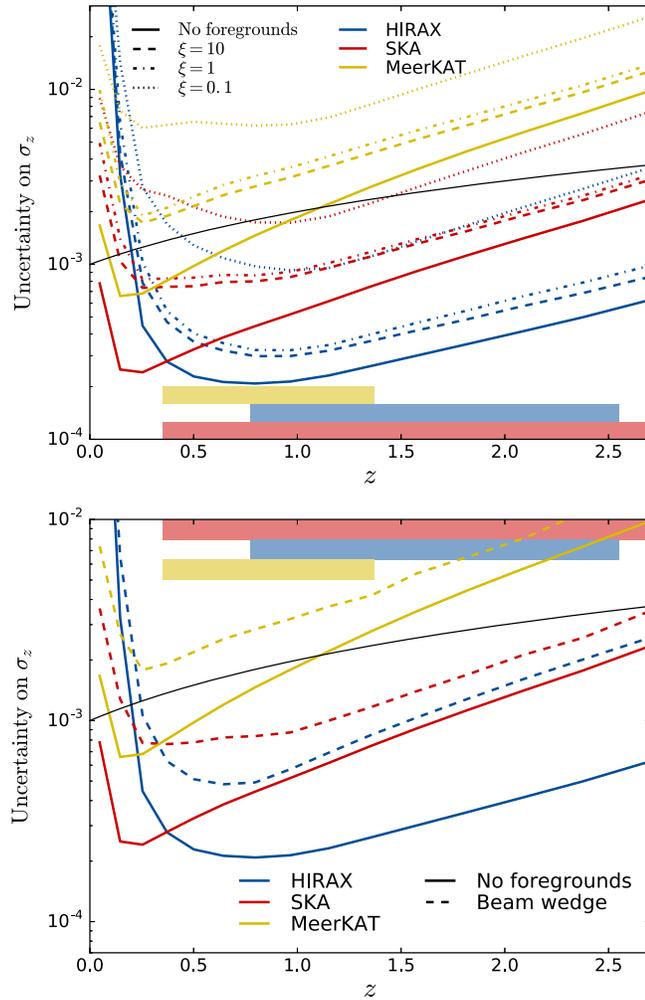


FIG. 9. (Upper panel) Dependence of the forecast constraints on the LSST photo- z scatter on the frequency correlation length of foreground residuals (a measure of the coherence of these residuals across frequencies). Correlation lengths above $\xi \sim 0.1$ (corresponding to radial scales $k_{\parallel} \sim 0.01 h \text{ Mpc}^{-1}$) are necessary to limit the degradation of the photo- z calibration. (Lower panel) Impact of the foreground wedge on the final constraints on the photo- z scatter.

under consideration, correlation lengths smaller than $\xi \sim 0.1$ would result in a fast degradation of the performance for photo- z calibration. In particular, the associated loss of k -space volume would prevent MeerKAT and SKA from reaching the calibration requirements for DES and LSST. The performance of HIRAX would also be severely compromised by foreground contamination, although it would still be able to reach the required constraints within its proposed frequency range.

It is worth noting that, even in the presence of perfectly correlated foreground residuals ($\xi \rightarrow \infty$), their smooth frequency dependence will make them dominate over a finite interval of k_{\parallel} , and this is the cause for the noticeable difference between our predictions for $\xi = 10$ and those in the absence of foregrounds (dashed and solid lines in the upper panel of Fig. 9 respectively). Smaller values of ξ will then enlarge this lost range of k_{\parallel} , and our results show that this only becomes relevant for $\xi \lesssim 1$.

These results pose the question of how uncorrelated we can expect foreground residual to be. For reference, raw foreground components are constrained to have correlation lengths of $\xi \sim 1\text{--}10$ [37]. On the other hand, more complicated residuals arising from leaked polarized synchrotron would exhibit a much richer frequency structure caused by Faraday rotation, with correlation lengths $\xi \sim 0.1$ at high frequencies ($\sim 800 \text{ MHz}$) decreasing for longer wavelengths [39,40,52]. An exquisite instrumental calibration will therefore be necessary in order to optimize the scientific output of 21 cm experiments. It is worth noting that the comoving scale corresponding to $\xi \sim 0.1$ ($k_{\parallel} \sim 0.01 h \text{ Mpc}^{-1}$ at $z \sim 1$) is similar to the cut suggested by [39], although the exponential form assumed here for the power spectrum of the foreground residuals [Eq. (10)] extends the degrading effect into larger wave numbers.

A further complication comes in the form of the so-called ‘‘foreground wedge,’’ produced by the long time-delay contribution of foregrounds from antennas with far sidelobe responses [53–56]. As proven in [57,58], this effect makes the region of k space defined by

$$k_{\parallel}^{\text{hor}} < \frac{\chi(z)H(z)}{c(1+z)} k_{\perp} \quad (19)$$

liable to foreground contamination. This is the so-called ‘‘horizon’’ wedge, and corresponds to the case where foreground contamination can be caused by the sources in the horizon picked up by very distant sidelobes. Under optimistic assumptions, however, we can consider the case where this effect extends only up to foreground sources located in the outer fringes of the primary beam, in which case the size of the wedge is reduced to the so-called primary-beam wedge [58], given by $k_{\parallel}^{\text{pb}} = \sin(\theta_{\text{FWHM}}/2)k_{\parallel}^{\text{hor}}$. The lower panel of Fig. 9 shows that the LSST photo- z calibration requirements are still met after accounting for the loss of k space to the primary-beam wedge.

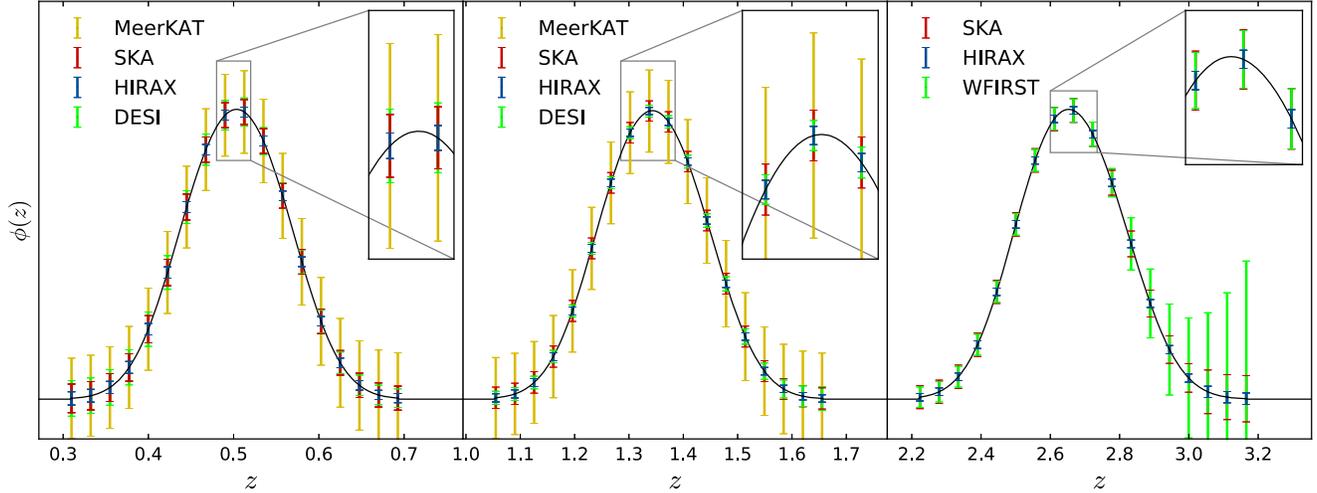


FIG. 10. Nonparametric reconstruction of the redshift distribution for four different redshift bins. The points with error bars show the 1σ constraints on $\phi(z)$ achievable with HIRAX (blue), SKA (red) and MeerKAT (yellow), as well as two spectroscopic surveys (green), DESI and WFIRST, at low and high redshifts respectively. Note that, given the proposed frequency band for HIRAX, it would not be able to calibrate the low- and high-redshift bins shown here.

D. Generalized redshift distributions

Even though the simple parametrization of photo- z systematic uncertainties in terms of Δz and σ_z allows us to easily compare the performance of different experiments in terms of photo- z calibration, in a realistic scenario we would like to calibrate generic redshift distributions without assuming a particular parametrization.

This is typically done by promoting the amplitude of the redshift distribution of the photometric sample in each narrow redshift interval to a free parameter that can be constrained from the cross-correlation with the spectroscopic survey. In this section we explore this scenario for the same redshift bins considered in the previous sections.

For this we use the method proposed by [26]. In essence, this method is equivalent to the formalism outlined in Sec. II E, where the free parameters q_i considered are the amplitudes of the photometric-redshift distribution. The method is further simplified in [26] to make it applicable to the analysis of real data using the following approximations:

- (1) All power spectra are computed using the Limber approximation. This implies (among other things) that all cross-correlations between nonoverlapping redshift bins are neglected.
- (2) The contributions from redshift-space distortions and magnification bias are not included in the model for the angular power spectra.
- (3) No marginalization over cosmological or other nuisance parameters is carried out.
- (4) The autocorrelation of photometric sources does not contain information about their redshift distribution.

We have adopted these same assumptions here to simplify the discussion.

Figure 10 shows the constraints achievable by different IM experiments and spectroscopic surveys on the

generalized form of the redshift distribution for three photometric-redshift bins centered around $z \sim 0.5, 1.35$ and 2.6 .⁴ The constraining power displayed in this figure matches the results shown in the right panel of Fig. 4.

Note that the uncertainties on the amplitude of the redshift distribution found this way can be translated into uncertainties on the two parameters Δz and σ_z used to characterize this distribution in the previous section by performing a simple two-parameter likelihood analysis for the model in Eqs. (4) and (5). We find that, using this procedure, we recover constraints on Δz and σ_z that are a factor ~ 2 to 3 worse than in the optimal case. This can be understood in terms of the simplifying assumptions adopted here, such as neglecting the information encoded in cross-bin correlations and in the autocorrelation of the photometric sample. In all cases, however, we recover the same relative performance between different experiments in terms of $\sigma(\sigma_z, \Delta z)$.

E. Impact on cosmological constraints

In order to study the impact of photo- z calibration on the final cosmological constraints, we have run a Fisher matrix analysis using the formalism described in Sec. II E with the specifications for LSST described in Sec. II B. In this case we consider a set of 54 free parameters:

- (a) Three nuisance parameters for each of the 15 redshift bins: the galaxy bias b , the photo- z bias Δz and the photo- z scatter σ_z .
- (b) Nine cosmological parameters: the fractional density of cold dark matter $\Omega_c h^2$, the fractional density in

⁴Note that the first and last bins would lie outside the proposed frequency range of HIRAX.

baryons $\Omega_b h^2$, the normalized local expansion rate h , the amplitude and tilt of primordial scalar perturbations A_s and n_s , the optical depth to reionization τ , the equation of state of dark energy in the w_0 - w_a parametrization and the sum of neutrino masses $\sum m_\nu$.

In order to pin down the early-universe parameters, we also include information from a hypothetical ground-based stage-4 CMB experiment [59] using the specifications assumed in [60] and complemented by Planck at low multipoles.

For the photo- z nuisance parameters we then add Gaussian priors corresponding to the 1σ uncertainties on σ_z and Δz found using the procedure described in the previous sections for the different experiments considered in this paper. The results of this exercise are displayed in Fig. 11, which shows the constraints on the most relevant late-universe parameters: the dark energy equation of state parameters, w_0 and w_a , and the sum of neutrino masses. The results shown correspond to the 1σ contours in the absence of photo- z uncertainties (dashed black line), with photo- z systematics constrained through cross-correlation with a 21 cm experiment, in this case HIRAX (red ellipse) and in the absence of external data for photo- z calibration (blue ellipse).

It is important to stress that the overall constraints on these parameters forecast here depend heavily on the survey specifications assumed (e.g., photo- z model and uncertainties), as well as the scales included in the analysis,

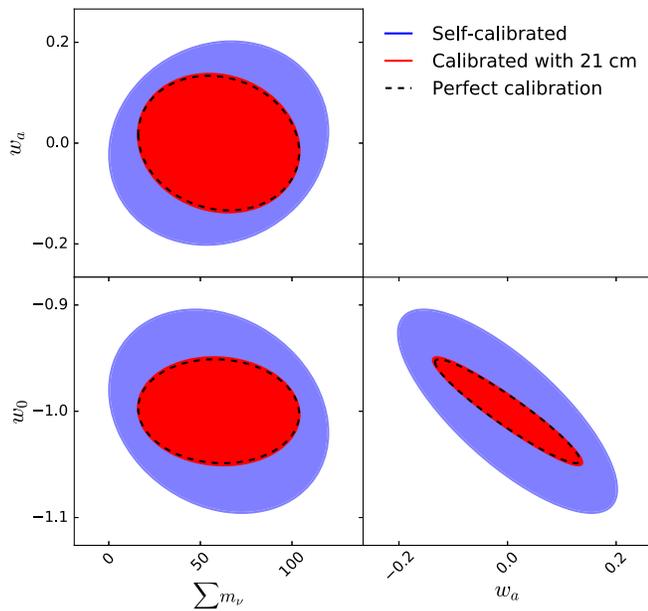


FIG. 11. Constraints on the equation of state of dark energy and the sum of neutrino masses in the absence of external photo- z calibration (blue ellipses), with redshift distributions calibrated through cross-correlation with a HIRAX-like 21 cm experiment (red ellipses) and in the case of perfect calibration (dashed lines). The constraints also include early-universe information from a stage-4 CMB experiment.

a point where we have tried to be very conservative. Thus, the results shown in Fig. 11 should not be taken to represent the final constraints achievable by LSST. The main result shown in this figure is the relative improvement on the final constraints after photo- z calibration, which should be more robust to these considerations.

Photo- z calibration improves the constraints on each of these parameters by a factor ~ 1.5 – 2 , and the dark energy figure of merit by a factor $\gtrsim 5$. Furthermore, we find that the level of calibration achievable through cross-correlations with intensity mapping experiments is nearly equivalent to the case of perfect calibration. Equivalent results are found for SKA and HIRAX, as well as for the combination of DESI, Euclid and WFIRST, as could be expected from the results displayed in the right panel of Fig. 4.

IV. DISCUSSION

In this paper we have shown that intensity maps of the HI emission can be used to improve the scientific output of photometric-redshift surveys. By exploiting the cross-correlations between imaging surveys of galaxies and HI maps, we find that it is possible to optimally calibrate the photometric-redshift distributions. This is made clear when assessing improvements in constraints on cosmological parameters: in Fig. 11 we see that the figure of merit using this method is effectively as good as having perfect calibration. This also means that, with the aid of future IM experiments it should be possible to, for example, improve the LSST equation of state figure of merit by approximately a factor of 5.

This approach is promising, but it is important to highlight some of the limitations that have to be overcome if IM is to be used successfully in this context. For a start, it will be important to be able to deal with foreground contamination. In essence, as we have discussed, one can model the effect of foregrounds as a source of noise that cancels the information contained in long-wavelength radial modes. This effect is therefore greatly dependent on how coherent foreground residuals are in frequency. We can expect foregrounds to be highly correlated in intensity, with correlation lengths of order $\xi \sim 1$ – 10 . However, instrumental effects such as polarization leakage or frequency-dependent beams could spoil this coherence and lead to significant losses in the coverage of the k_{\parallel} – k_{\perp} plane. We estimate that residuals with correlation lengths $\xi \lesssim 0.1$ (as expected for polarized synchrotron emission, and corresponding to scales $k_{\parallel} \gtrsim 0.01 h \text{ Mpc}^{-1}$ at $z \sim 1$) would significantly degrade the ability of 21 cm maps to calibrate redshift distributions efficiently. We have furthermore shown that the calibration requirements of future photometric surveys can also be matched after accounting for the so-called foreground wedge. Finally, it is worth noting that one of the main strengths of the method is its reliance on cross-correlations between the spectroscopic and photometric samples, and that this cross-correlation

should be very robust against systematic biases caused by foreground contamination.

If we are to use IM to calibrate future redshift surveys, then we need to make sure that the observational setup satisfies minimum requirements and to do so, we have explored its dependence on experimental parameters. We have found that, for the noise levels of currently proposed experiments, we are primarily limited by cosmic variance and therefore there is no advantage in gaining depth at the cost of sky area: it is preferable to maximize the overlap between the HI maps and the galaxy imaging survey. Furthermore, if we are to accurately capture the longer wave-length modes, we need to resort to single-dish observations, and therefore deal with the problems associated with correlated noise in autocorrelation measurements. In the long term, these “ $1/f$ ” noise contributions will have to be controlled, and our analysis shows that the dishes must have a minimum size of ~ 15 m. If, on the other hand, we are to use interferometric observations (a method which is more tried and tested) then we need to ensure a minimum baseline of ~ 12 m to capture the large-scale angular modes. We have shown how MeerKAT, HIRAX and SKA fall well within these experimental parameter constraints.

We must also note that our analysis has been conservative in terms of the ranges of scales that add up to the constraints on photo- z parameters, only including angular scales in the regime where nonlinear structure formation is believed to be well understood. It may, however, be possible to use even smaller scales for the purposes of photo- z calibration [5], in which case some of the conclusions drawn from this analysis could vary. In particular, the relative performance of HIRAX and SKA in terms of photo- z calibration could be significantly different, owing to the higher sensitivity of the SKA interferometer to small angular scales (see Fig. 13).

It is interesting to consider alternative approaches to sharpening photometric-redshift measurements. Gravitational lensing of the CMB has recently been advocated as a promising approach, given its perfectly well-determined radial kernel (e.g., see [61]). We can explore this possibility by considering correlations between LSST and a CMB lensing convergence κ_{CMB} map as in Sec. II E. We consider two different generations of CMB experiments: an ongoing “stage-3” experiment, characterized by a rms noise level of $\sigma_T \sim 8 \mu\text{K arcmin}$, and a future “stage-4” experiment with $\sigma_T = 1 \mu\text{K arcmin}$. In both cases we assume very optimistic configurations, with a beam FWHM of 1 arcmin and using all angular scales $\ell \in [2, 3000]$. We also fix all cosmological and bias parameters, and only consider varying Δz and σ_z for each LSST redshift bin. Figure 12 shows the constraint on the photo- z bias Δz in the absence of CMB lensing data (green solid line), and using the cross-correlation with κ_{CMB} as measured by S3 (purple) and S4 (cyan) experiments. We clearly see that, even under these overly optimistic

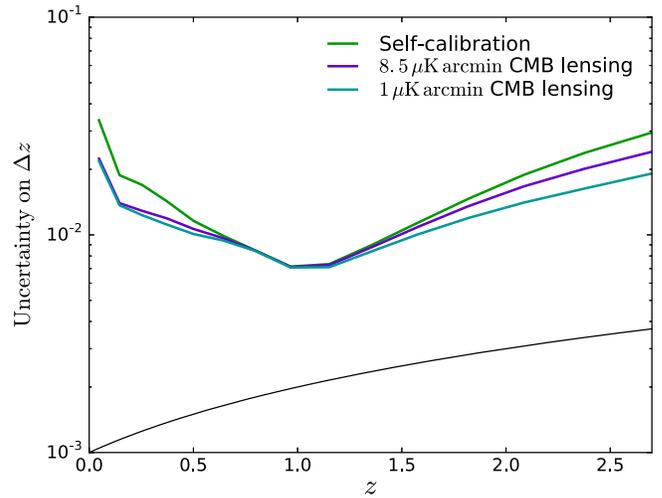


FIG. 12. Forecast constraints on the LSST photo- z bias Δz in the absence of external data sets (green), and adding CMB lensing data from stage-3 (purple) and stage-4 (cyan) CMB experiments.

assumptions, adding CMB lensing information does not significantly help calibrate photo- z uncertainties.

Experiments to undertake intensity mapping of HI are an ongoing effort, and will give us an extremely promising new window on the Universe. We have argued that they will not only contribute in their own right to the further understanding of the large-scale structure of the Universe, but will also help improve the scientific returns from a plethora of up and coming optical surveys.

ACKNOWLEDGMENTS

The authors would like to thank Elisabeth Krause, Mario Santos and Anže Slosar for the useful comments and discussions. D. A. is supported by the Science and Technology Facilities Council and the Leverhulme and Beecroft Trusts. P. G. F. acknowledges support from the Science Technology Facilities Council, the European Research Council and the Beecroft Trust. M. J. J. acknowledges support from the South African Square Kilometre Array project and the Oxford Centre for Astrophysical Surveys, which is funded through generous support from the Hintze Family Charitable Foundation. K. M. acknowledges support from the National Research Foundation of South Africa (Grant No. 98957).

APPENDIX A: INDIVIDUAL CLUSTERING REDSHIFTS

The idea of using clustering information to constrain the redshifts of individual objects of a given sample has been considered before in the literature, and shown to yield interesting results even in the absence of spectroscopic data [27]. Here we outline the steps that should be taken to include intensity mapping information in this formalism.

Our aim is to find the most general expression for the posterior distribution of the true redshifts of a set of galaxies for which we only have photometric data and an overlapping intensity mapping survey. We start by considering a data vector \mathbf{d} consisting of

- (a) $\hat{\mathbf{n}}$: galaxy positions.
- (b) \mathbf{m} : galaxy magnitudes.
- (c) δ_{HI} : a map of the perturbation in the HI density across angles and redshift.

For each galaxy we want to estimate a redshift z_i , so let \mathbf{z} be a vector containing all those redshifts. We want to study the posterior probability $p(\mathbf{z}|\delta_{\text{HI}}, \hat{\mathbf{n}}, \mathbf{m})$. Let us start by noting that, in the standard models of large-scale structure, both δ_{HI} and the galaxy distribution can be thought of as being biased and noisy representations of the underlying matter overdensity δ_{M} . Sampling the galaxy redshifts could then also give us information about δ_{M} , and therefore it is worth considering the joint distribution $p(\mathbf{z}, \delta_{\text{M}}|\delta_{\text{HI}}, \hat{\mathbf{n}}, \mathbf{m})$.

One can study this distribution by iteratively sampling the two conditional distributions:

$$\delta_{\text{M}}^{i+1} \leftarrow p(\delta_{\text{M}}|\mathbf{z}^i, \delta_{\text{HI}}, \mathbf{m}, \hat{\mathbf{n}}), \quad \mathbf{z}^{i+1} \leftarrow p(\mathbf{z}|\delta_{\text{M}}^{i+1}, \delta_{\text{HI}}, \mathbf{m}, \hat{\mathbf{n}}). \quad (\text{A1})$$

We outline these two steps below.

- (a) Conditional density distribution. We start by noting that, if the true redshifts \mathbf{z} are known, then the photometric redshifts given by the magnitudes \mathbf{m} are of no use in constraining the overdensity field, and therefore

$$p(\delta_{\text{M}}|\delta_{\text{HI}}, \mathbf{z}, \hat{\mathbf{n}}, \mathbf{m}) = p(\delta_{\text{M}}|\delta_{\text{HI}}, \mathbf{z}, \hat{\mathbf{n}}) \quad (\text{A2})$$

$$= p(\delta_{\text{M}}|\delta_{\text{HI}}, \delta_{\text{g}}), \quad (\text{A3})$$

where $\delta_{\text{g}}(\mathbf{z}, \hat{\mathbf{n}})$ is the galaxy overdensity uniquely defined by the galaxy angular coordinates and redshifts. $p(\delta_{\text{M}}|\delta_{\text{HI}}, \delta_{\text{g}})$ can then be decomposed using Bayes's theorem as

$$p(\delta_{\text{M}}|\delta_{\text{HI}}, \delta_{\text{g}}) \propto p(\delta_{\text{g}}|\delta_{\text{M}})p(\delta_{\text{HI}}|\delta_{\text{M}})p(\delta_{\text{M}}), \quad (\text{A4})$$

where, following the same philosophy as above, we have considered that $p(\delta_{\text{g}}|\delta_{\text{M}}, \delta_{\text{HI}}) = p(\delta_{\text{g}}|\delta_{\text{M}})$, since δ_{HI} is just a noisy and biased realization of δ_{M} . All that remains is then to model the conditional distributions $p(\delta_{\text{g}}|\delta_{\text{M}})$ and $p(\delta_{\text{HI}}|\delta_{\text{M}})$, which is by no means a cursory matter, but something that can certainly be accomplished in the regime of validity of structure formation models.

- (b) Conditional redshift distribution. Under the assumptions that galaxies are Poisson distributed over the (biased) matter overdensity, and that the photometric-redshift errors are independent of the environmental density, it is possible to show (e.g., [27]) that the

galaxy redshifts can be sampled individually with the distribution

$$p(z|\delta, \hat{\mathbf{n}}, m) \propto [1 + b_{\text{g}}(\delta_{\text{M}})(z, \hat{\mathbf{n}})]p(z|m), \quad (\text{A5})$$

where $b_{\text{g}}(\delta_{\text{M}})(z, \hat{\mathbf{n}})$ is the galaxy overdensity along the angular direction of each galaxy, and $p(z|m)$ is the prior photo- z distribution.

APPENDIX B: ANGULAR POWER SPECTRA

This section describes the theoretical models used for the angular power spectra entering the computation of the Fisher matrix [Eq. (17)]. The cross-power spectrum between two tracers of the cosmic density field, a and b , can be estimated as

$$C_{\ell}^{ab} = 4\pi \int_0^{\infty} \frac{dk}{k} \mathcal{P}_{\Phi}(k) W_{\ell}^a(k) W_{\ell}^b(k), \quad (\text{B1})$$

where $\mathcal{P}_{\Phi}(k)$ is the power spectrum of the primordial curvature perturbations and $W_{\ell}^a(k)$ is the window function for tracer a , containing information about the different contributions to the total anisotropy in that tracer and about its redshift distribution.

In the case of galaxy clustering and intensity mapping, and neglecting contributions from magnification bias and large-scale relativistic effects, W^a is given by

$$W_{\ell}^a(k) = \int_0^{\infty} dz \phi_a(z) \left[b_a(z) T_{\delta}(k, z) j_{\ell}(k\chi(z)) + \frac{1+z}{H(z)} T_{\theta}(k, z) j_{\ell}''(k\chi(z)) \right], \quad (\text{B2})$$

where $H(z)$ and $\chi(z)$ are the expansion rate and radial comoving distance at redshift z respectively, $\phi_a(z)$ is the source redshift distribution, and T_{δ} and T_{θ} are the transfer functions of the matter overdensity and velocity divergence fields. Note that, even though we include the effect of nonlinearities using the nonlinear transfer function for δ (through the prescription of [62]), we only introduce the effect of redshift-space distortions at the linear level, and only consider a deterministic linear bias $b_a(z)$. This is, nevertheless, a more rigorous treatment than has been used in the literature, and the procedure used to mitigate the effect of nonlinearities described in Sec. II E should minimize the corresponding impact on the forecasts presented here.

For galaxy shear tracers of weak lensing, the expression for the window function is

$$W_{\ell}^a(k) = -\frac{1}{2} \sqrt{\frac{(\ell+2)!}{(\ell-2)!}} \int_0^{\infty} \frac{dz}{H(z)} \int_z^{\infty} dz' \phi_a(z') \times \frac{\chi(z') - \chi(z)}{\chi(z')\chi(z)} T_{\phi+\psi}(k, z) j_{\ell}(k\chi(z)), \quad (\text{B3})$$

where $T_{\phi+\psi}$ is the transfer function for the sum of the two metric potentials in the Newtonian gauge.

The computation of Eq. (B1) was carried out using a modified version of the Boltzmann code CLASS [63,64].

APPENDIX C: NOISE POWER SPECTRUM FOR INTENSITY MAPPING EXPERIMENTS

This section derives the expression for the noise power spectra of single-dish experiments and interferometers presented in Eq. (6). Similar derivations have been provided before in the literature (e.g., [22]), but we include this calculation here for completeness.

Throughout this section we will use a flat-sky approach, where angles on the sky are represented by a 2D Cartesian vector \mathbf{x} . In this approximation, the spherical harmonic transform of a field becomes a simple 2D Fourier transform:

$$f_{\ell m} \equiv \sum_{\ell m} f(\hat{\mathbf{n}}) Y_{\ell m}(\hat{\mathbf{n}}) \rightarrow f_{\ell} \equiv \int \frac{(dx)^2}{2\pi} e^{i\mathbf{x}\cdot\boldsymbol{\ell}} f(\mathbf{x}). \quad (\text{C1})$$

We will also simplify the derivation by writing integrals as Riemann sums. For instance, the Fourier transform above will read

$$f_{\ell} = \sum_{\mathbf{x}} \frac{(\Delta x)^2}{2\pi} e^{i\mathbf{x}\cdot\boldsymbol{\ell}} f(\mathbf{x}). \quad (\text{C2})$$

Note that, with this normalization, the definition for the power spectrum P_f of a stochastic field f is

$$\langle f_{\ell} f_{\ell'}^* \rangle \equiv \frac{\delta_{\ell, \ell'}}{(\Delta l)^2} P_f(\ell), \quad (\text{C3})$$

where $\Delta l \equiv 2\pi/\Delta x$.

1. Single dish

The flux at angular position \mathbf{x} measured by a single dish is the sky intensity I integrated over the instrumental beam B :

$$S(\mathbf{x}) = N_B \sum_{\mathbf{x}'} (\Delta x)^2 I(\mathbf{x}') B(\mathbf{x} - \mathbf{x}'), \quad (\text{C4})$$

where $N_B \equiv 1/B(0)$. Inserting the definition (C2) in the expression above, and using the orthogonality relation $\sum_{\mathbf{x}} (\Delta x)^2 \exp[i\mathbf{x}\cdot(\boldsymbol{\ell} - \boldsymbol{\ell}')] = \delta_{\ell, \ell'} (2\pi/\Delta l)^2$, one can relate the Fourier components of S and I as

$$I_{\ell} = S_{\ell}/[N_B B(\ell)], \quad \text{where } B(\ell) \equiv \sum_{\mathbf{x}} \frac{(\Delta x)^2}{(2\pi)^2} e^{i\mathbf{x}\cdot\boldsymbol{\ell}}. \quad (\text{C5})$$

The power spectrum for I is then related to that of $S(\mathbf{s})$ as $P_I(\ell) = P_S(\ell)/[N_B B(\ell)]^2$. Assume now that S is purely

white noise with a per-pointing rms flux σ_S , such that its power spectrum is simply flat with an amplitude

$$P_S = \sigma_S^2 (\Delta x)^2 = \left(\frac{2k_B T_{\text{sys}}}{A_e \sqrt{\Delta\nu t_p}} \right)^2 (\Delta x)^2 = \left(\frac{2k_B T_{\text{sys}}}{A_e} \right)^2 \frac{\Omega_{\text{obs}}}{\Delta\nu t_{\text{tot}}}, \quad (\text{C6})$$

where Δx is the angular separation between pointings, T_{sys} is the per-pointing rms temperature fluctuation, A_e is the effective collecting area of the dish, $\Delta\nu$ is the channel frequency bandwidth, t_p is the integration time per pointing, Ω_{tot} is the total observed sky area and t_{tot} is the total integration time.

Substituting this into the expression for P_I and relating the intensity I to a brightness temperature T through the Rayleigh-Jeans law ($I = 2k_B T/\lambda^2$), we obtain the temperature noise power spectrum:

$$P_T(\ell) = \frac{T_{\text{sys}}^2 4\pi f_{\text{sky}}}{\eta^2 N_{\text{dish}} \Delta\nu t_{\text{tot}}} B^{-2}(\ell), \quad (\text{C7})$$

where f_{sky} is the observed sky fraction, we have considered the possibility of having N_{dish} independent dishes and we have defined the quantity $\eta \equiv A_e N_B/\lambda^2$. Note that, for a circular aperture telescope, $N_B = 4\lambda^2/(\pi d^2)$, where d is the dish diameter, and therefore $\eta = A_e/[\pi(d/2)^2]$ is the ratio of the effective to total dish area, which we have labeled “efficiency” in Eq. (6).

2. Interferometers

The visibility observed by a pair of antennas separated by a baseline $\mathbf{d} \equiv \lambda \mathbf{u}$ is

$$\begin{aligned} V(\mathbf{u}) &\equiv \sum_{\mathbf{x}} (\Delta x)^2 T(\mathbf{x}) B(\mathbf{x}) e^{2\pi i \mathbf{u}\cdot\mathbf{x}} \rightarrow T(\mathbf{x}) B(\mathbf{x}) \\ &= \sum_{\mathbf{u}} (\Delta u)^2 e^{-2\pi i \mathbf{u}\cdot\mathbf{x}}. \end{aligned} \quad (\text{C8})$$

Transforming this to Fourier space we find

$$\begin{aligned} T_{\ell} &= \sum_{\mathbf{u}} (\sqrt{2\pi} \Delta u)^2 V(\mathbf{u}) \sum_{\mathbf{x}} \left(\frac{\Delta x}{2\pi} \right)^2 \frac{e^{i\mathbf{x}\cdot(\boldsymbol{\ell} - 2\pi \mathbf{u})}}{B(\ell)} \\ &= \frac{V(\ell/(2\pi))}{2\pi}, \end{aligned} \quad (\text{C9})$$

where, in the last step, we have used the small-angle approximation $B = 1$. The variance of T_{ℓ} is therefore given by

$$\langle |T_{\ell}|^2 \rangle = \frac{\langle |V(\ell/(2\pi))|^2 \rangle}{(2\pi)^2 n(\ell/(2\pi)) (\Delta u)^2}, \quad (\text{C10})$$

where $n(\mathbf{u})(\Delta u)^2$ is the number of baselines in a volume $(\Delta u)^2$ in \mathbf{u} space.

In temperature units, the noise variance per visibility is given by $\langle |V(\mathbf{u})|^2 \rangle = [\lambda^2 T_{\text{sys}}^2 / (A_e \sqrt{\Delta \nu t_p})]^2$. Relating baselines to Fourier coefficients as $\mathbf{u} = \ell / (2\pi)$ and recalling the definition of a power spectrum [Eq. (C3)], the noise power spectrum in temperature is then given by

$$P_T(\ell) = \frac{\lambda^2 T_{\text{sys}}^2 N_p}{A_e^2 \Delta \nu t_{\text{tot}} n(\mathbf{u} = \ell / (2\pi))}, \quad (\text{C11})$$

where N_p is the total number of pointings. Relating $n(\mathbf{u})$ to the number density of physical baselines, and defining $N_p \Omega_p \equiv 4\pi f_{\text{sky}}$, we recover the expression for the noise power spectrum of interferometers in Eq. (6).

3. Comparison with spectroscopic surveys

Converting the angular maps in different frequency channels into a three-dimensional map of the HI overdensity, we can relate the 3D noise power spectrum to its angular counterpart as

$$P_{3D}(\mathbf{k}_{\parallel}, \mathbf{k}_{\perp}) = \frac{c[(1+z)r(z)]^2}{\nu_{21} H(z) T_{\text{HI}}^2(z)} P_T(\ell \equiv r\mathbf{k}_{\perp}), \quad (\text{C12})$$

where r is the comoving angular diameter distance to redshift z and T_{HI} is the average 21 cm temperature. This can then be directly compared with the shot-noise power spectrum of spectroscopic surveys, given by $P^{3D} = 1/\bar{n}$,

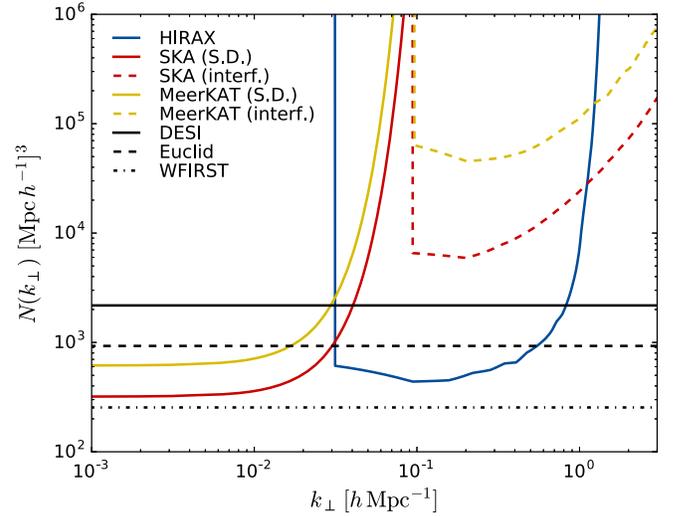


FIG. 13. Noise power spectrum at $z = 1.2$ as a function of transverse wave number k_{\perp} for HIRAX (blue), SKA (red) and MeerKAT (yellow). The curves for SKA and MeerKAT are separated into single dish (solid) and interferometer (dashed). The shot-noise levels for the spectroscopic surveys DESI, Euclid and WFIRST at the same redshift are shown as black solid, dashed and dot dashed respectively.

where \bar{n} is the 3D density of sources. The left panel of Fig. 13 shows the 3D noise power spectrum at $z \sim 1.2$ as a function of the transverse wave number \mathbf{k}_{\perp} for the three IM experiments (HIRAX, SKA and MeerKAT) and the three next-generation spectroscopic surveys (DESI, Euclid and WFIRST) considered here.

-
- [1] J. A. Newman, Calibrating redshift distributions beyond spectroscopic limits with cross-correlations, *Astrophys. J.* **684**, 88 (2008).
- [2] J. Benjamin, L. van Waerbeke, B. Ménard, and M. Kilbinger, Photometric redshifts: estimating their contamination and distribution using clustering information, *Mon. Not. R. Astron. Soc.* **408**, 1168 (2010).
- [3] D. J. Matthews and J. A. Newman, Reconstructing redshift distributions with cross-correlations: Tests and an optimized recipe, *Astrophys. J.* **721**, 456 (2010).
- [4] S. J. Schmidt, B. Ménard, R. Scranton, C. Morrison, and C. K. McBride, Recovering redshift distributions with cross-correlations: Pushing the boundaries, *Mon. Not. R. Astron. Soc.* **431**, 3307 (2013).
- [5] B. Ménard, R. Scranton, S. Schmidt, C. Morrison, D. Jeong, T. Budavari, and M. Rahman, Clustering-based redshift estimation: Method and application to data, [arXiv:1303.4722](https://arxiv.org/abs/1303.4722).
- [6] M. P. van Daalen and M. White, A cross-correlation-based estimate of the galaxy luminosity function, [arXiv:1703.05326](https://arxiv.org/abs/1703.05326).
- [7] M. Rahman, A. J. Mendez, B. Ménard, R. Scranton, S. J. Schmidt, C. B. Morrison, and T. Budavári, Exploring the SDSS photometric galaxies with clustering redshifts, *Mon. Not. R. Astron. Soc.* **460**, 163 (2016).
- [8] A. Choi, C. Heymans, C. Blake, H. Hildebrandt, C. A. J. Duncan, T. Erben, R. Nakajima, L. Van Waerbeke, and M. Viola, CFHTLenS and RCSLenS: Testing photometric redshift distributions using angular cross-correlations with spectroscopic galaxy surveys, *Mon. Not. R. Astron. Soc.* **463**, 3737 (2016).
- [9] V. Scottez, Y. Mellier, B. R. Granett, T. Moutard, M. Kilbinger *et al.*, Clustering-based redshift estimation: Application to VIPERS/CFHTLS, *Mon. Not. R. Astron. Soc.* **462**, 1683 (2016).
- [10] M. Rahman, B. Ménard, and R. Scranton, Exploring the 2MASS extended and point source catalogs with clustering redshifts, *Mon. Not. R. Astron. Soc.* **457**, 3912 (2016).
- [11] A. Johnson, C. Blake, A. Amon, T. Erben, K. Glazebrook *et al.*, 2dFLenS and KiDS: Determining source redshift

- distributions with cross-correlations, *Mon. Not. R. Astron. Soc.* **465**, 4118 (2017).
- [12] R. A. Battye, R. D. Davies, and J. Weller, Neutral hydrogen surveys for high-redshift galaxy clusters and protoclusters, *Mon. Not. R. Astron. Soc.* **355**, 1339 (2004).
- [13] M. McQuinn, O. Zahn, M. Zaldarriaga, L. Hernquist, and S. R. Furlanetto, Cosmological parameter estimation using 21 cm radiation from the epoch of reionization, *Astrophys. J.* **653**, 815 (2006).
- [14] T.-C. Chang, U.-L. Pen, J. B. Peterson, and P. McDonald, Baryon Acoustic Oscillation Intensity Mapping of Dark Energy, *Phys. Rev. Lett.* **100**, 091303 (2008).
- [15] J. S. B. Wyithe and A. Loeb, Fluctuations in 21-cm emission after reionization, *Mon. Not. R. Astron. Soc.* **383**, 606 (2008).
- [16] A. Loeb and J. S. B. Wyithe, Possibility of Precise Measurement of the Cosmological Power Spectrum With a Dedicated 21 cm Survey After Reionization, *Phys. Rev. Lett.* **100**, 161301 (2008).
- [17] J. B. Peterson, R. Aleksan, R. Ansari *et al.*, 21 cm intensity mapping, [arXiv:0902.3091](https://arxiv.org/abs/0902.3091).
- [18] J. Bagla, N. Khandai, and K. K. Datta, HI as a probe of the large-scale structure in the post-reionization universe, *Mon. Not. R. Astron. Soc.* **407**, 567 (2010).
- [19] R. A. Battye, I. W. A. Browne, C. Dickinson, G. Heron, B. Maffei, and A. Pourtsidou, HI intensity mapping: a single dish approach, *Mon. Not. R. Astron. Soc.* **434**, 1239 (2013).
- [20] K. W. Masui, E. R. Switzer, N. Banavar, K. Bandura, C. Blake *et al.*, Measurement of 21 cm brightness fluctuations at $z \sim 0.8$ in cross-correlation, *Astrophys. J. Lett.* **763**, L20 (2013).
- [21] E. R. Switzer, K. W. Masui, K. Bandura, L.-M. Calin, T.-C. Chang *et al.*, Determination of $z \sim 0.8$ neutral hydrogen fluctuations using the 21 cm intensity mapping autocorrelation, *Mon. Not. R. Astron. Soc.* **434**, L46 (2013).
- [22] P. Bull, P. G. Ferreira, P. Patel, and M. G. Santos, Late-time cosmology with 21 cm intensity mapping experiments, *Astrophys. J.* **803**, 21 (2015).
- [23] J. Fonseca, R. Maartens, and M. G. Santos, Probing the primordial Universe with MeerKAT and DES, *Mon. Not. R. Astron. Soc.* **466**, 2780 (2017).
- [24] L. B. Newburgh, K. Bandura, M. A. Bucher, T.-C. Chang, H. C. Chiang *et al.*, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, SPIE Proceedings Vol. 9906 (SPIE—International Society for Optical Engineering, Bellingham, WA, 2016), p. 99065X.
- [25] M. G. Santos, P. Bull, D. Alonso, S. Camera, P. Ferreira *et al.*, Cosmology with a SKA HI intensity mapping survey, *Proc. Sci.*, AASKA14 (2015) 019 [[arXiv:1501.03989](https://arxiv.org/abs/1501.03989)].
- [26] M. McQuinn and M. White, On using angular cross-correlations to determine source redshift distributions, *Mon. Not. R. Astron. Soc.* **433**, 2857 (2013).
- [27] J. Jasche and B. D. Wandelt, Bayesian inference from photometric redshift surveys, *Mon. Not. R. Astron. Soc.* **425**, 1042 (2012).
- [28] D. Alonso and P. G. Ferreira, Constraining ultra large-scale cosmology with multiple tracers in optical and radio surveys, *Phys. Rev. D* **92**, 063525 (2015).
- [29] A. Gabasch, U. Hopp, G. Feulner, R. Bender, S. Seitz *et al.*, The evolution of the luminosity functions in the FORS deep field from low to high redshift, *Astron. Astrophys.* **448**, 101 (2006).
- [30] M. R. Blanton and S. Roweis, K-corrections and filter transformations in the ultraviolet, optical, and near-infrared, *Astron. J.* **133**, 734 (2007).
- [31] LSST Collaboration, LSST science book, version 2.0, [arXiv:0912.0201](https://arxiv.org/abs/0912.0201).
- [32] D. H. Weinberg, R. Davé, N. Katz, and L. Hernquist, Galaxy Clustering and Galaxy Bias in a Λ CDM Universe, *Astrophys. J.* **601**, 1 (2004).
- [33] R. A. Battye, M. L. Brown, I. W. A. Browne, R. J. Davis, P. Dewdney, C. Dickinson, G. Heron, B. Maffei, A. Pourtsidou, and P. N. Wilkinson, BINGO: A single dish approach to 21 cm intensity mapping, [arXiv:1209.1041](https://arxiv.org/abs/1209.1041).
- [34] L. B. Newburgh, G. E. Addison, M. Amiri, K. Bandura, J. R. Bond *et al.*, in *Ground-Based and Airborne Telescopes V*, SPIE Proceedings Vol. 9145 (SPIE—International Society for Optical Engineering, Bellingham, WA, 2014), p. 91454V.
- [35] M.-A. Bigot-Sazy, Y.-Z. Ma, R. A. Battye, I. W. A. Browne, T. Chen *et al.*, in *Proceedings of the Frontiers in Radio Astronomy and FAST Early Sciences Symposium, Guiyang, China, 2015*, Astronomical Society of the Pacific Conference Series Vol. 502, edited by L. Qain and D. Li (Astronomical Society of the Pacific, San Francisco, 2016), p. 41.
- [36] X. Chen, Radio detection of dark energy—The Tianlai project, *Sci. Sin. Phys. Mechan. Astron.* **41**, 1358 (2011).
- [37] M. G. Santos, A. Cooray, and L. Knox, Multifrequency analysis of 21 centimeter fluctuations from the era of reionization, *Astrophys. J.* **625**, 575 (2005).
- [38] L. Wolz, F. B. Abdalla, C. Blake, J. R. Shaw, E. Chapman, and S. Rawlings, The effect of foreground subtraction on cosmological measurements from intensity mapping, *Mon. Not. R. Astron. Soc.* **441**, 3271 (2014).
- [39] J. R. Shaw, K. Sigurdson, M. Sitwell, A. Stebbins, and U.-L. Pen, Coaxing cosmic 21 cm fluctuations from the polarized sky using m -mode analysis, *Phys. Rev. D* **91**, 083514 (2015).
- [40] D. Alonso, P. Bull, P. G. Ferreira, and M. G. Santos, Blind foreground subtraction for intensity mapping experiments, *Mon. Not. R. Astron. Soc.* **447**, 400 (2015).
- [41] Planck Collaboration, Planck 2015 results. X. Diffuse component separation: Foreground maps, *Astron. Astrophys.* **594**, A10 (2016).
- [42] M. Levi, C. Bebek, T. Beers, R. Blum, R. Cahn *et al.*, The DESI Experiment, a whitepaper for Snowmass 2013, [arXiv:1308.0847](https://arxiv.org/abs/1308.0847).
- [43] A. Font-Ribera, P. McDonald, N. Mostek, B. A. Reid, H.-J. Seo, and A. Slosar, DESI and other dark energy experiments in the era of neutrino mass measurements, *J. Cosmol. Astropart. Phys.* **05** (2014) 023.
- [44] R. Laureijs, J. Amiaux, S. Arduini, J.-L. Auguères, J. Brinchmann *et al.*, Euclid definition study report, [arXiv:1110.3193](https://arxiv.org/abs/1110.3193).
- [45] D. Spergel, N. Gehrels, J. Breckinridge, M. Donahue, A. Dressler *et al.*, Wide-Field InfraRed Survey Telescope—Astrophysics Focused Telescope Assets WFIRST-AFTA final report, [arXiv:1305.5422](https://arxiv.org/abs/1305.5422).
- [46] M. Tegmark, A. J. S. Hamilton, M. A. Strauss, M. S. Vogeley, and A. S. Szalay, Measuring the galaxy power spectrum with future redshift surveys, *Astrophys. J.* **499**, 555 (1998).

- [47] J. R. Bond, A. H. Jaffe, and L. Knox, Estimating the power spectrum of the cosmic microwave background, *Phys. Rev. D* **57**, 2117 (1998).
- [48] J. A. Newman, A. Abate, F. B. Abdalla, S. Allam, S. W. Allen *et al.*, Spectroscopic needs for imaging dark energy experiments: Photometric redshift training and calibration, *Astropart. Phys.* **63**, 81 (2015); **65**, 112(E) (2015).
- [49] R. de Putter, O. Doré, and S. Das, Using cross-correlations to calibrate lensing source redshift distributions: Improving cosmological constraints from upcoming weak lensing surveys, *Astrophys. J.* **780**, 185 (2014).
- [50] Dark Energy Survey Collaboration, The Dark Energy Survey, [arXiv:astro-ph/0510346](https://arxiv.org/abs/astro-ph/0510346).
- [51] S. Camera, M. G. Santos, P. G. Ferreira, and L. Ferramacho, Cosmology on Ultralarge Scales with Intensity Mapping of the Neutral Hydrogen 21 cm Emission: Limits on Primordial Non-Gaussianity, *Phys. Rev. Lett.* **111**, 171302 (2013).
- [52] V. Jelić, S. Zaroubi, P. Labropoulos, G. Bernardi, A. G. de Bruyn, and L. V. E. Koopmans, Realistic simulations of the galactic polarized foreground: Consequences for 21-cm reionization detection experiments, *Mon. Not. R. Astron. Soc.* **409**, 1647 (2010).
- [53] A. Datta, J. D. Bowman, and C. L. Carilli, Bright source subtraction requirements for redshifted 21 cm measurements, *Astrophys. J.* **724**, 526 (2010).
- [54] H. Vedantham, N. Udaya Shankar, and R. Subrahmanyam, IMAGING the epoch of reionization: limitations from foreground confusion and imaging algorithms, *Astrophys. J.* **745**, 176 (2012).
- [55] A. Liu, A. R. Parsons, and C. M. Trott, The epoch of reionization window. II. Statistical methods for foreground wedge reduction, *Phys. Rev. D* **90**, 023019 (2014).
- [56] H.-J. Seo and C. M. Hirata, The foreground wedge and 21 cm BAO surveys, *Mon. Not. R. Astron. Soc.* **456**, 3142 (2016).
- [57] A. R. Parsons, J. C. Pober, J. E. Aguirre, C. L. Carilli, D. C. Jacobs, and D. F. Moore, A per-baseline, delay-spectrum technique for accessing the 21 cm cosmic reionization signature, *Astrophys. J.* **756**, 165 (2012).
- [58] J. C. Pober, A. Liu, J. S. Dillon, J. E. Aguirre, J. D. Bowman, R. F. Bradley, C. L. Carilli, D. R. DeBoer, J. N. Hewitt, D. C. Jacobs, M. McQuinn, M. F. Morales, A. R. Parsons, M. Tegmark, and D. J. Werthimer, What next-generation 21 cm power spectrum measurements can teach us about the epoch of reionization, *Astrophys. J.* **782**, 66 (2014).
- [59] K. N. Abazajian, P. Adshead, Z. Ahmed, S. W. Allen, D. Alonso *et al.*, CMB-S4 science book, first edition, [arXiv:1610.02743](https://arxiv.org/abs/1610.02743).
- [60] E. Calabrese, D. Alonso, and J. Dunkley, Complementing the ground-based CMB-S4 experiment on large scales with the PIXIE satellite, *Phys. Rev. D* **95**, 063504 (2017).
- [61] E. J. Baxter, J. Clampitt, T. Giannantonio, S. Dodelson, B. Jain *et al.*, Joint measurement of lensing-galaxy correlations using SPT and DES SV data, *Mon. Not. R. Astron. Soc.* **461**, 4099 (2016).
- [62] R. Takahashi, M. Sato, T. Nishimichi, A. Taruya, and M. Oguri, Revising the halofit model for the nonlinear matter power spectrum, *Astrophys. J.* **761**, 152 (2012).
- [63] D. Blas, J. Lesgourgues, and T. Tram, The Cosmic Linear Anisotropy Solving System (CLASS) II: Approximation schemes, *J. Cosmol. Astropart. Phys.* **07** (2011) 034.
- [64] E. Di Dio, F. Montanari, J. Lesgourgues, and R. Durrer, The CLASSgal code for relativistic cosmological large scale structure, *J. Cosmol. Astropart. Phys.* **11** (2013) 044.