


Impact of Foregrounds on HI Intensity Mapping Cross-Correlations with Optical Surveys

Steven Cunningham¹^{*}, Laura Wolz², Alkistis Pourtsidou^{3,4}, David Bacon¹

¹*Institute of Cosmology & Gravitation, University of Portsmouth, Dennis Sciama Building, Burnaby Road, Portsmouth PO1 3FX, UK*

²*School of Physics, University of Melbourne, Parkville, VIC 3010, Australia*

³*School of Physics and Astronomy, Queen Mary University of London, Mile End Road, London E1 4NS, UK*

⁴*Department of Physics & Astronomy, University of the Western Cape, Cape Town 7535, South Africa*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

The future of precision cosmology could benefit from cross-correlations between intensity maps of unresolved neutral hydrogen (HI) and more conventional optical galaxy surveys. A major challenge that needs to be overcome is removing the 21cm foreground emission that contaminates the cosmological HI signal. Using N -body simulations we simulate HI intensity maps and optical catalogues which share the same underlying cosmology. Adding simulated foreground contamination and using state-of-the-art reconstruction techniques we investigate the impacts that 21cm foregrounds and other systematics have on these cross-correlations. We find that the impact a FASTICA 21cm foreground clean has on the cross-correlations with spectroscopic optical surveys with well-constrained redshifts is minimal. However, problems arise when photometric surveys are considered: we find that a redshift uncertainty $\sigma_z \geq 0.04$ causes significant degradation in the cross power spectrum signal. We diagnose the main root of these problems, which relates to arbitrary amplitude changes along the line-of-sight in the intensity maps caused by the foreground clean and suggest solutions which should be applicable to real data. These solutions involve a reconstruction of the line-of-sight temperature means using the available overlapping optical data along with an artificial extension to the HI data through redshift to address edge effects. We then put these solutions through a further test in a mock experiment that uses a clustering-based redshift estimation technique to constrain the photometric redshifts of the optical sample. We find that with our suggested reconstruction, cross-correlations can be utilized to make an accurate prediction of the optical redshift distribution.

Key words: large-scale structure of Universe – distances and redshifts – cosmology: observations – techniques: spectroscopic – photometric – radio lines: galaxies

1 INTRODUCTION

Conventional optical galaxy surveys map large-scale cosmic structure by using resolved galaxies as tracers of the underlying dark matter field, which dominates the overall matter density. Their distribution contains information about the expansion history and the growth of cosmic structure. Using cosmological probes such as baryon acoustic oscillations we can measure the Universe’s expansion history and constrain dark energy (Percival et al. 2001). Measuring the large-scale distribution of matter also reveals information on the primordial state of the Universe (Slosar et al. 2008) which has the potential to constrain models of the early Universe. Fur-

thermore, probing structure on cosmic horizon scales will test General Relativity and indicate if modifications to this theory of gravity are required.

Methods involving detection of galaxies to trace large-scale structure are reliable providing that the galaxy samples obtained by a survey have a sufficient number density. If not, the measurements will suffer from significant statistical errors due to Poisson shot noise. Obtaining a large number of resolved galaxies with precise redshifts is expensive; spectroscopic redshifts with a redshift uncertainty $\sigma_z \sim 0.001$ rely on long integration times making this a slow process. Photometric redshifts offer a less precise alternative but can be obtained much more quickly allowing dense catalogues of galaxies to be built (Bolzonella et al. 2000; Fernandez-Soto et al. 2001). It is for this reason that future stage-IV surveys

^{*} E-mail: steve.cunnington@port.ac.uk

such as *Euclid*¹ (Amendola et al. 2018) will heavily rely on photometric redshifts, and the Large Synoptic Survey Telescope (LSST)² (Marshall et al. 2017) will be entirely reliant on them.

As an alternative, radio intensity mapping techniques, which do not rely on resolving individual sources, offer the prospect of more complete tracer maps with the redshift precision of a spectroscopic survey. In complete contrast to optical surveys, intensity mapping provides excellent constraints along the radial line-of-sight but poor angular resolution. This complementarity, together with the fact that cross-correlations are expected to alleviate survey-specific systematic effects, makes synergies between intensity mapping and optical galaxy surveys mutually beneficial.

Arguably the most appealing source of emission for cosmology with intensity mapping comes from the neutral hydrogen (HI) gas residing within galaxies. The signature of HI that we aim to detect comes from the hyperfine transition of hydrogen's single electron. When the electron drops to a lower energy state it emits a photon with a rest frequency of 1420 MHz, or 21cm of wavelength, hence the synonymous name 21cm intensity mapping. The observed frequency of these emitted photons places the signal in the radio part of the electromagnetic spectrum. Therefore radio dishes are the conventional choice of receiver for detecting these photons at low redshifts of $z < 3$. First detections using the HI intensity mapping technique have already been achieved in cross-correlation with optical galaxies in Pen et al. (2009), Masui et al. (2013) and Anderson et al. (2017).

The most prominent example of a next generation radio observatory is the Square Kilometre Array (SKA)³ (SKA Cosmology SWG. et al. 2018). The mid-frequency instrument, SKA1-MID (where 1 stands for Phase 1), will be an array of 197 dish receivers that can operate in interferometer and single-dish mode. The low frequency instrument, SKA1-LOW, will probe the high redshift Universe, targeting the Epoch of Reionisation. As with any interferometer, it is the largest separation (or baseline) which determines the resolution of the instrument; hence baselines of up to 150 km are proposed to maximize resolution. Conversely, it is the smallest baselines between receivers which determines the largest scales that can be probed. The SKA1-MID instrument aims to perform a wide ($\sim 20,000 \text{ deg}^2$) HI intensity mapping survey in single-dish mode. This compromises angular resolution but probes the large scales needed for cosmology.

The redshifted 21cm line signal from HI benefits from being particularly isolated in frequency, and there are few examples of spectral lines that could lead to potential line confusion, making HI intensity mapping particularly robust for redshift experiments. However, a major challenge for HI intensity mapping comes from foreground emission (e.g. synchrotron radiation), which can be orders of magnitude larger than the cosmological signal. Foregrounds are spectrally smooth signals which emit in the same range as the redshifted HI. Blind foreground removal techniques, which require no prior knowledge or templates of the foregrounds,

| 21cm IM Survey | | Photo-z Survey | f_{sky} | z_{min} | z_{max} |
|----------------|---|----------------|------------------|------------------|------------------|
| MeerKAT | × | DES | 0.1 | 0 | 1.45 |
| TIANLAI | × | DECaLS | 0.15 | 0 | 1.5 |
| SKA1-MID | × | <i>Euclid</i> | 0.2 | 0.35 | 2 |
| SKA1-MID | × | LSST | 0.4 | 0.35 | 3 |
| HIRAX | × | <i>Euclid</i> | 0.2 | 0.8 | 2 |
| HIRAX | × | LSST | 0.5 | 0.8 | 2 |
| CHIME | × | <i>Euclid</i> | 0.35 | 0.8 | 2 |
| CHIME | × | LSST | 0.5 | 0.8 | 2.5 |

Table 1. Examples of cross-correlation opportunities between 21cm intensity mapping surveys and optical photometric redshift surveys, with (approximate) estimates for their sky and redshift overlap. f_{sky} refers to the fraction of full sky for which these surveys can overlap. z_{min} and z_{max} represent the common redshift overlap range.

can be used to exploit the smooth form of most foreground signals along the line-of-sight to isolate and remove them.

In this work we investigate how foreground removal can impact important cosmological measurements. Several studies have investigated how foreground removal can be carried out without detrimental impact on the HI auto-correlation power spectrum recovery (Wolz et al. 2014; Shaw et al. 2014; Alonso et al. 2015). In this work we aim to place particular emphasis on the foreground removal's impact on cross-correlation measurements with optical galaxy surveys. Examples of some future optical-21cm cross-correlation possibilities are outlined in Table 1. In order to investigate the impact of foregrounds on cross-correlations, we utilize mock galaxy catalogues built from N -body simulations of dark matter particles. This approach allows for both optical and HI intensity map data to share the same underlying simulated cosmology, with realistic parameters (such as number density of galaxies) corresponding to the specifications of current and forthcoming surveys.

The plan of the paper is as follows: In Section 2 we describe how we simulate the cosmological signals, both our resolved optical galaxy number density maps and the overlapping HI intensity maps. Section 3 explains how we simulate the 21cm foregrounds, which are then added into our HI cosmological signal to contaminate the intensity maps. Section 4 then explains the processes used for removing these foregrounds and details the FASTICA approach that we opt to use on our simulations. In Section 5 we analyze our results and demonstrate what impact foreground cleaning can have on a cross-correlation power spectrum. In Section 6 we extend these findings and apply them to a practical experiment which utilizes these cross-correlations to constrain photometric redshifts using HI intensity maps. We conclude and discuss in Section 7.

2 COSMOLOGICAL SIGNALS & THEIR SIMULATION

In order to probe the large-scale cosmic structure and map the matter over-density δ , we rely on luminous sources which trace the underlying matter density. In optical galaxy redshift surveys we use number density fields $n_g(\vec{\theta}, z)$ where resolved galaxies can be counted in voxels (3-dimensional pix-

¹ www.euclid-ec.org/

² www.lsst.org/

³ www.skatelescope.org/

els) at angular position $\vec{\theta}$ with a redshift z which is used for defining the line-of-sight (LoS) distance. We can then calculate the over-density of galaxies δ_g , which we assume is a linearly biased tracer of the matter over-density δ , i.e.

$$\delta_g(\vec{\theta}, z) \equiv \frac{n_g(\vec{\theta}, z) - \bar{n}_g(z)}{\bar{n}_g(z)} = b_g(z)\delta(\vec{\theta}, z), \quad (1)$$

where a barred quantity represents a mean average and b_g is the (linear) galaxy bias.

For HI intensity maps there are no resolved luminous sources, only combined brightness temperatures in a given voxel. Assuming that HI is also a biased tracer of the underlying matter density we can write

$$\delta_{\text{HI}}(\vec{\theta}, z) \equiv \frac{T_{\text{HI}}(\vec{\theta}, z) - \bar{T}_{\text{HI}}(z)}{\bar{T}_{\text{HI}}(z)} = b_{\text{HI}}(z)\delta(\vec{\theta}, z), \quad (2)$$

where the linear bias factor is now b_{HI} . Note that the mean brightness temperature \bar{T}_{HI} is also an unknown quantity, degenerate with b_{HI} . Since the observable is a temperature fluctuation, it is customary to work with temperature fluctuation maps where

$$\delta T_{\text{HI}}(\vec{\theta}, z) \equiv T_{\text{HI}}(\vec{\theta}, z) - \bar{T}_{\text{HI}}(z) = \bar{T}_{\text{HI}}(z)b_{\text{HI}}(z)\delta(\vec{\theta}, z). \quad (3)$$

It is these quantities, δ_g and δT_{HI} , which can be used to make cosmological measurements e.g. auto-power spectra $P_{\text{gg}} \sim \langle |\tilde{\delta}_g|^2 \rangle$ or cross-power spectra $P_{g,\text{HI}} \sim \langle \text{Re}[\tilde{\delta}_g \tilde{\delta}_{\text{HI}}^*] \rangle$. Here the tilde notation $\tilde{\delta}$ represents the Fourier transform of the matter over-density. Hence, in order to use simulations to study the impact 21cm foregrounds can have on cosmological measurements using cross-correlations, we require a simulation which includes HI emission and resolved optical galaxies.

In many 21cm studies it is sufficient to simulate wide continuous intensity maps through Gaussian realizations of a HI power spectrum. However, for this work we need an optical galaxy catalogue which shares the same underlying cosmology as the HI intensity maps, since we are looking to exploit a shared clustering signal between resolved optical galaxies and HI emission for cross-correlated measurements. It is also preferable to have the optical galaxy simulation as a resolved catalogue of sources so that $N(z)$ distributions can be built precisely from individual galaxy redshifts. We can then choose to degrade the redshift accuracy in order to emulate a photometric imaging survey.

In order to achieve this, we use existing N -body galaxy simulations and exploit certain components of them, e.g. HI mass or halo mass to simulate HI brightness temperatures which we can build intensity maps from. Utilizing N -body simulations also allows for a more robust representation of a survey catalogue than Gaussian realized signals. With this in mind we ideally require a catalogue which has the following features:

- low halo-mass resolution ($\approx 10^9 h^{-1} M_\odot$) so that intensity maps include integrated HI emission from faint galaxies;
- HI information for each galaxy for simulating realistic intensity maps;
- deep redshift and wide sky coverage ($0 < z < 3$, $\sim 20,000 \text{ deg}^2$) to allow for testing low resolutions associated with the typical beam size of a SKA-like intensity mapping experiment;

- simulated photometry for optically resolved galaxies so that realistic cross-correlation forecasts can be made between intensity maps and photometric galaxy surveys.

A simulation including all of the above is not currently available, and is unlikely to be available in the near future. This is largely due to the fact that low halo mass resolution with sufficient galaxy number densities over large sky volumes would require N -body simulations that would be exceptionally computationally expensive.

In this work we therefore utilize two simulated catalogues with differing characteristics. One catalogue contains HI signal with a low halo mass resolution and simulated HI masses for every galaxy. The other is a more conventional optical survey catalogue with simulated photometry but which is not as resolved in halo mass. We will now describe the two catalogues in detail.

• GAEA Simulation⁴

We make use of the GAEA semi-analytic model (Zoldan et al. 2017; Xie et al. 2017; Hirschmann et al. 2016). The catalogue was built using the Millennium Simulation (Springel et al. 2005), which is a cosmological N -body simulation that used $N = 2160^3$ particles of mass $m_p = 8.6 \times 10^8 h^{-1} M_\odot$ within a comoving box of size $500^3 \text{ (Mpc}/h)^3$ with a cosmology consistent with WMAP1 data. In particular, the values of the adopted cosmological parameters are: $\Omega_b = 0.045$, $\Omega_m = 0.25$, $\Omega_\Lambda = 0.75$, $H_0 = 100h \text{ Mpc}^{-1} \text{ km s}^{-1}$, $h = 0.73$, $\sigma_8 = 0.9$ and $n_s = 1$. The GAEA catalogue is built replicating this same box, but selecting galaxies from the nearest snapshot corresponding to its co-moving distance from the observer.

GAEA used an algorithm to identify halos which allowed for a halo mass resolution of $1.7 \times 10^{10} M_\odot h^{-1}$ which resulted in just over 2×10^8 galaxies with a continuous sky coverage $f_{\text{sky}} = 0.5$. Redshifts are limited to $0 < z < 0.5$ which means we will only be able to study cross-correlations within this small range, but this should still allow for multiple redshift/frequency bins given the completeness within this range. GAEA also includes simulated HI masses for all its galaxies, which can be used to generate realistic HI brightness temperatures. We discuss this further in Section 2.1.

• MICE Simulation⁵

We also make use of the MICECATv2.0 simulation released as part of the MICE-Grand Challenge Galaxy and Halo Light-cone Catalogue (Fosalba et al. 2015b; Crocce et al. 2015; Fosalba et al. 2015a; Carretero et al. 2015; Hoffmann et al. 2015), which is a cosmological N -body dark matter only simulation containing 4096^3 dark-matter particles of mass $m_p = 2.93 \times 10^{10} h^{-1} M_\odot$ in a box-size of $3072^3 \text{ (Mpc}/h)^3$. They resolved halos down to a few $10^{11} M_\odot h^{-1}$ and used a hybrid Halo Occupation Distribution (HOD) and Halo Abundance Matching (HAM) technique for galaxy modelling resulting in just under 5×10^8 galaxies. The simulation's sky footprint is $90 \times 90 \text{ deg}^2$ filling an octant of sky ($f_{\text{sky}} = 0.125$) up to a redshift $z = 1.4$. The assumed cosmology is a flat concordance Λ CDM model with $\Omega_m = 0.25$, $\Omega_\Lambda = 0.75$, $\Omega_b = 0.044$, $n_s = 0.95$, $\sigma_8 = 0.8$ and $h = 0.7$ consistent with WMAP 5-year data.

⁴ <http://astrosims.flatironinstitute.org/gaea>

⁵ <http://maia.ice.cat/mice/>

| Catalogue | Box Size [Mpc/h] ³ | m_p [M_\odot/h] | N_{gal} | f_{sky} | z_{max} |
|-----------|----------------------------------|--------------------------|-------------------|------------------|------------------|
| GAEA | 500 | 8.6×10^8 | 201×10^6 | 0.5 | 0.5 |
| MICE | 3072 | 2.9×10^{10} | 497×10^6 | 0.125 | 1.4 |

Table 2. Summary of the two different mock galaxy catalogues we will be using. Both are built from N -body simulations for which we provide the box size and particle mass m_p .

Since the MICE catalogue does not have simulated HI masses for each galaxy, we must derive our own. We therefore take each central galaxy’s halo mass as simulated by MICE and convert this into a predicted HI mass by following the redshift dependent prescription laid out in [Padmanabhan & Kulkarni \(2017\)](#)

$$M_{\text{HI}} = 2N_1 M \left[\left(\frac{M}{M_1} \right)^{-b_1} + \left(\frac{M}{M_1} \right)^{y_1} \right]^{-1}, \quad (4)$$

where M is the galaxy’s halo mass; M_1 , N_1 , b_1 and y_1 are all free parameters with redshift dependence tuned to provide a best fit; we refer the reader to [Padmanabhan & Kulkarni \(2017\)](#) for details. Each central galaxy then has a HI mass from which we can generate a HI brightness temperature signal. While this prescription would not be ideal for small scale studies of HI distribution, since we are assuming that all HI lies within central galaxies, it suits our purposes because we will be smoothing out any small scale imprecisions when we simulate the effect of an intensity mapping beam.

From these catalogues, which we summarize in Table 2, we will produce both our HI intensity maps (Section 2.1) and our detected optical galaxy catalogue (Section 2.2), which will share the same underlying dark-matter distribution. It is this shared clustering signal which we will look to utilize in our cross-correlation tests. We emphasize once more that we use these two separate N -body simulations since each one has unique advantages. For example the semi-analytical GAEA has replication of the particle box sample which delivers larger sky sizes and also has HI masses for each galaxy at lower mass resolution. Both of these features contribute to delivering more robust simulations of large-beam HI intensity maps. In contrast MICE uses a HOD/HAM approach over a larger box size, so is arguably more realistic in its cosmological signal in that no replication is required, but perhaps less realistic in that it distributes synthetic galaxies into simulated halos using a statistical approach rather than simulating baryonic process to drive galaxy evolution, as performed in semi-analytic models. MICE also includes some simulated photometric redshifts which we can utilize for forecasting the impacts of HI foregrounds in cross-correlations with a photometric survey.

2.1 HI Intensity Map Simulation

We aim to express our HI intensity map data T_{HI} in the form of a brightness temperature with two angular dimensions (θ_{ra} and θ_{dec} , jointly represented by $\vec{\theta}$) and a radial dimension, the redshift z . To do this we follow the same

recipe laid out in [Cunnington et al. \(2019\)](#) which we repeat here for completeness.

To construct T_{HI} we start with the HI mass M_{HI} of each galaxy, which is given in the GAEA catalogue and generated using halo masses and Equation (4) for MICE. We then place our galaxies into a data cube with coordinates $(\theta_{\text{ra}}, \theta_{\text{dec}}, z)$ by binning each galaxy’s HI mass into its relevant voxel so we end up with a gridded HI mass map $M_{\text{HI}}(\vec{\theta}, z_c)$.

We can then convert this into an intensity field for a frequency width of $\delta\nu$ subtending a solid angle $\delta\Omega$ (which is effectively our pixel size)

$$I_{\text{HI}}(\vec{\theta}, z) = \frac{3h_P A_{12}}{16\pi m_h} \frac{1}{[(1+z)\chi(z)]^2} \frac{M_{\text{HI}}(\vec{\theta}, z)}{\delta\nu\delta\Omega} \nu_{21}, \quad (5)$$

where h_P is the Planck constant, A_{12} the Einstein coefficient which quantifies the rate of spontaneous photon emission by the hydrogen atom, m_h is the mass of the hydrogen atom, ν_{21} the rest frequency of the 21cm emission and $\chi(z)$ is the comoving distance out to redshift z (we will assume a flat universe).

It is conventional in radio astronomy, in particular intensity mapping, to use brightness temperature which can be defined as the flux density per unit solid angle of a source measured in units of equivalent black body temperature. Hence, our intensity $I_{\text{HI}}(\vec{\theta}, z)$ can be written in terms of a black-body temperature in the Rayleigh-Jeans approximation $T = I c^2 / (2k_B \nu^2)$ where k_B is the Boltzmann constant. Using this we can estimate the brightness temperature at redshift z

$$T_{\text{HI}}(\vec{\theta}, z) = \frac{3h_P c^2 A_{12}}{32\pi m_h k_B \nu_{21}} \frac{1}{[(1+z)\chi(z)]^2} \frac{M_{\text{HI}}(\vec{\theta}, z)}{\delta\nu\delta\Omega}. \quad (6)$$

For cosmology studies one aims to make measurements at different redshifts. We therefore choose to slice the intensity maps into thin tomographic redshift bins and collapse these to a 2D slice which can be auto-correlated or cross-correlated with another survey map. We will often discuss binning by frequency (ν) or redshift (z). To clarify, these are interchangeable expressions since $z+1 = \nu_{21}/\nu_{\text{obs}}$. The width of each tomographic redshift bin needs to be thin enough that we can make certain thin bin assumptions, yet wide enough that we allow for sufficient structure to obtain a strong cross-correlation signal. By thin bin assumptions we are referring to cosmological quantities such as the bias, which we assume to be constant within the width of our bin ($\Delta z = 0.02, 0.05$ for GAEA and MICE respectively).

In order to ensure our HI intensity map amplitudes are in agreement with what is theoretically predicted, we choose to rescale each redshift bin so that it agrees with a model average temperature \bar{T}_{HI} . For example [Battye et al. \(2013\)](#) gives this average temperature as

$$\bar{T}_{\text{HI}}(z) = 180 \Omega_{\text{HI}}(z) h \frac{(1+z)^2}{H(z)/H_0} \text{ mK} \quad (7)$$

where Ω_{HI} is the HI density (abundance). In principle Ω_{HI} can be measured using the auto-correlation HI power spectrum with redshift space distortions, assuming a fixed fiducial cosmology ([Masui et al. 2013](#); [Pourtsidou et al. 2017](#)). For this work we use a fit for the HI density ([SKA Cosmology SWG. et al. 2018](#))

$$\Omega_{\text{HI}}(z) = 0.00048 + 0.00039z - 0.000065z^2. \quad (8)$$

In radio HI intensity mapping the observable signals detected by a telescope are brightness temperature fluctuations,

$$\delta T_{\text{HI}}(\vec{\theta}, z) = T_{\text{HI}}(\vec{\theta}, z) - \bar{T}_{\text{HI}}(z). \quad (9)$$

We will therefore convert all our intensity maps to these quantities.

2.1.1 Receiver Noise

As we are aiming to simulate realistic observations, we need to include the effects of instrumental (thermal) noise. For the case of a single-dish intensity mapping experiment instrumental noise can be modelled as uncorrelated Gaussian fluctuations. Following Alonso et al. (2015) and Santos et al. (2015) we add onto our observable maps a Gaussian random field with rms

$$\sigma_{\text{noise}} = T_{\text{sys}} \sqrt{\frac{4\pi f_{\text{sky}}}{\Omega_{\text{beam}} N_{\text{dish}} t_{\text{obs}} \Delta\nu}}. \quad (10)$$

Here T_{sys} is the total system temperature which is the sum of the sky and receiver noise, $T_{\text{sys}} = T_{\text{rcvr}} + T_{\text{sky}}$ with $T_{\text{rcvr}} = 0.1T_{\text{sky}} + T_{\text{inst}}$ and $T_{\text{sky}}(\nu) \approx 60(300\text{MHz}/\nu)^{2.55}\text{K}$. We set $T_{\text{inst}} = 20\text{K}$ which is representative of SKA1-MID for the redshift range $0 < z < 0.58$. $\Omega_{\text{beam}} = 1.133\theta_{\text{beam}}^2$ is the solid angle for the intensity mapping beam. We also assume SKA1-MID-like values for the remaining variables in the noise model with the fraction of sky $f_{\text{sky}} = 0.41$ (representative of an SKA-LSST overlap), the number of dishes $N_{\text{dish}} = 197$ and the total observation time $t_{\text{obs}} = 10,000$ hours. Lastly, $\Delta\nu$ is the frequency bandwidth for a particular redshift bin. Figure 1 shows the level of this noise in relation to the cosmological HI signal. We can see that the noise only begins to dominate when the signal has the telescope beam effects (discussed in next section) included, and this is only at small scales (high ℓ).

A complete noise simulation would require the inclusion of red noise (or $1/f$ noise) which originates from time correlated gain fluctuations unique to radio receivers (Harper et al. 2018). Here we assume that using an appropriate scan speed ($\sim 1\text{deg s}^{-1}$), as well as component separation techniques, this noise can be removed (Harper et al. 2018). There is also an argument to include the effects of cross-shot noise caused by HI emitting galaxies, which provide signal in the intensity maps, also being present in the optical galaxy sample (Wolz et al. 2018). We assume these additional noise effects are sub-dominant at the scales of interest and do not include them in our simulations.

2.1.2 Beam Resolution

To model the low angular resolution of an intensity map, we convolve δT_{HI} with a telescope beam in Fourier space making use of the convolution theorem. Our telescope beam is modelled as a symmetric, two-dimensional Gaussian function with a full width half maximum of θ_{beam} acting only in the directions perpendicular to the LoS (as the frequency/redshift resolution is excellent). The beam size can be determined by the dimensions of the radio receiver and the frequency which is being probed (Alonso et al. 2017):

$$\theta_{\text{beam}} = \frac{1.22c}{\nu D_{\text{max}}}, \quad (11)$$

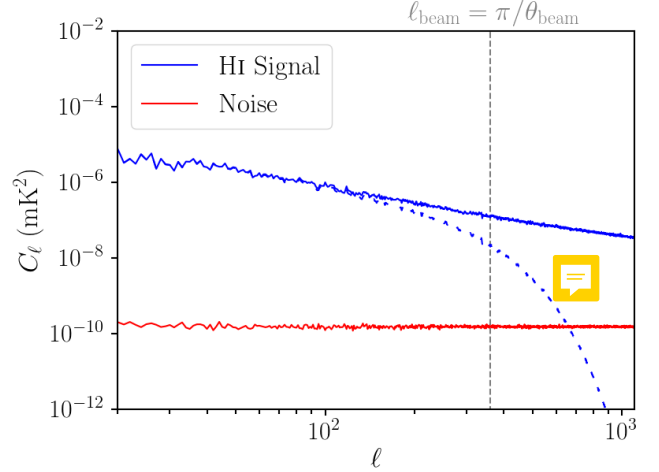


Figure 1. Angular power spectrum at a redshift of $z = 0.25$ ($\nu = 1136\text{ MHz}$) for both the cosmological signal (blue solid line) for a HI intensity map produced using the GAEA catalogue, and instrumental noise (red solid line). Also included is the effect of a $\theta_{\text{beam}} = 0.5^\circ$ Gaussian convolution (blue dashed line) which shows a degradation in the cosmological HI signal on smaller scales (high ℓ). The grey vertical dashed line shows the angular scale of this beam. We see that instrumental noise begins to dominate at around $\ell > 700$.

where D_{max} is the maximum baseline of the radio telescope; for a single dish receiver, D_{max} is given by the dish diameter. The GAEA redshift range of $0 < z < 0.5$ would mean we are looking at beam sizes of $0.99^\circ < \theta_{\text{beam}} < 1.45^\circ$ for our intensity maps, where we have assumed a maximum baseline of $D_{\text{max}} = 15\text{ m}$ which is representative of the SKA1-MID dishes (SKA Cosmology SWG. et al. 2018). The MICE catalogue, which extends to redshifts of $z = 1.4$ will reach even larger beam sizes of $\theta_{\text{beam}} = 2.36^\circ$. Cunningham et al. (2019) showed that the error associated with an estimator which relies on the intensity mapping correlation function will be proportional to the beam size:

$$\sigma \propto \frac{\theta_{\text{beam}}}{\sqrt{A}}. \quad (12)$$

Here A is the sky area covered. This shows that if a beam size increases without compensating with a larger sky area, then noise from the beam will become more dominant. Figure 1 shows how the beam effect can present additional challenges in that it causes instrumental noise to dominate at small scales and potentially destroys information there. We will include the beam scale in terms of multipole ℓ_{beam} on all our power spectra plots (as done in Figure 1) as this is one of the most dominant effects on our results and on HI intensity mapping power spectra in general.

An example of a completed intensity map tomographically sliced and collapsed into a 2D angular map is shown in Figure 2. For all our full-sky maps we use HEALPix maps (Gorski et al. 2005) where the pixelization ensures that each pixel covers the same surface area as every other pixel. We handle the maps in HEALPix RING ordering scheme with resolution $\text{nside} = 512$, which corresponds to $12 \times 512^2 = 3,145,728$ pixels across the sky.

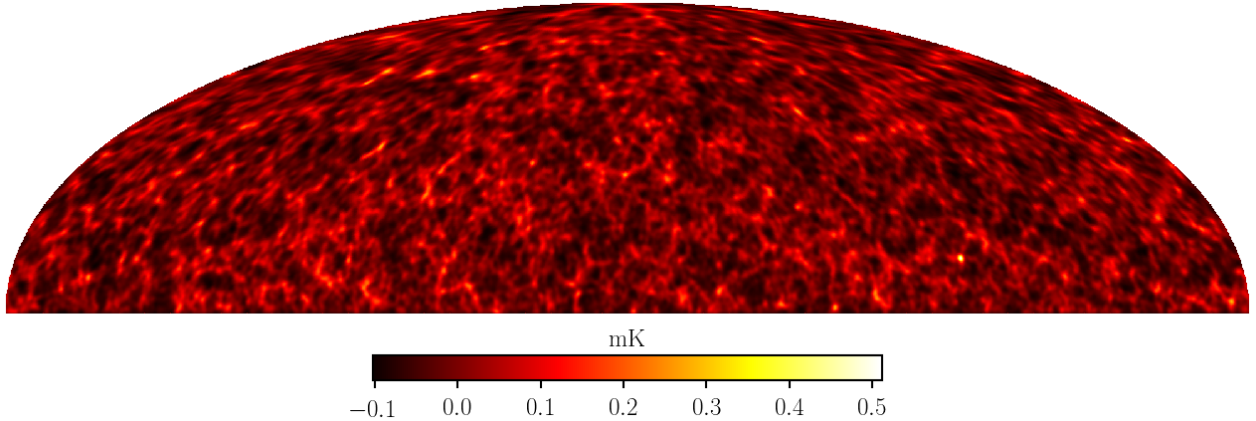


Figure 2. δT_{HI} intensity map at redshift $z = 0.25$ ($\nu = 1136\text{MHz}$ with bin width $\Delta\nu \sim 18\text{MHz}$). This includes the effects of SKA-like noise and beam, outlined in Sections 2.1.1 and 2.1.2 respectively. At this frequency the beam size is approximately $\theta_{\text{beam}} = 1.23^\circ$. This example is done with the GAEA catalogue covering half of the sky ($f_{\text{sky}} = 0.5$). This example does not include any foreground contamination.

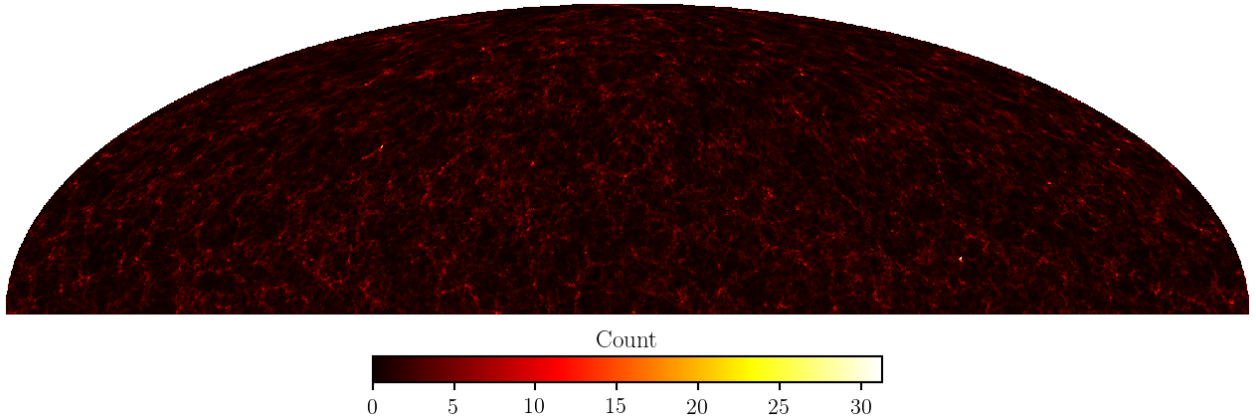


Figure 3. n_g optical galaxy number density field with galaxies binned by true redshift at $z = 0.25$. Unlike the intensity map in Figure 2, this map has no beam smoothing since it represents observations by an optical telescope. However, for this demonstration map only, we have downgraded the HEALPix resolution to $n_{\text{side}} = 128$. This is to make the shared structure between this and the intensity map at the same redshift more apparent.

2.2 Optical Galaxy Catalogue Simulation

For probing large-scale cosmic structure with resolved optical galaxies we use number density fields.

While we ideally require a simulated catalogue with high number density and completeness for the HI intensity maps, it would be unrealistic to expect every one of the low mass galaxies to be resolved and detected by a conventional wide area optical survey. **Therefore to make this a realistic test we need to introduce some detection threshold which results in only the brightest galaxies being included in our optical sample.** We also desire to have realistic $N(z)$ redshift distributions which tail off at higher redshifts where resolved detection becomes more difficult. The way this is all achieved is by invoking a model redshift distribution, given by

$$\frac{dN_g}{dz} = z^\beta \exp(-(z\alpha/z_m)^\gamma) \quad (13)$$

where we use $\alpha = \sqrt{2}$, $\beta = 2$ and $\gamma = 1.5$ (Harrison et al. 2016) which are values typical of stage-IV optical large scale structure survey such as LSST or *Euclid*. z_m is the mid-redshift

for the particular simulated catalogue we are applying this to e.g. for MICE this would be $z_m = 0.7$. We make the optical samples conform to this distribution by ordering galaxies by stellar mass in each redshift bin. Here we are using stellar mass as a crude approximation of optical brightness which for our purposes will be sufficient. We then pick the ‘brightest’ galaxies in each redshift bin until the model redshift distribution is achieved. This process gives final galaxy catalogues with 2.67×10^7 galaxies for GAEA, which is an average density of around 54 galaxies per square degree for each of the 24 redshift bins we use. For MICE we achieve a much denser catalogue with 3.97×10^8 galaxies over a smaller sky area giving 3.2×10^3 galaxies per square degree.

Our optical sample makes no consideration of any classifications of galaxies. All are treated as point-like and either ‘observed’ or not. More investigation could be taken into certain classifications e.g. by colour; red and blue galaxies are expected to cluster differently and have different densities at different redshifts. This could plausibly have an effect on our studies and bias the correlation function; this has

been touched upon in a recent cross-correlation study using Parkes HI intensity maps and 2dF galaxies (Anderson et al. 2017).

Figure 3 shows an example of a final over-density field for our optical data. This has been made using the GAEA catalogue at $z = 0.25$; similarities between this and Figure 2 should be apparent since these are both for the same dataset at the same redshift. We have shown this map with $n_{\text{side}} = 128$ to make the clustering pattern more obvious.

3 21CM FOREGROUNDS & THEIR SIMULATION

We test the effects on HI intensity maps of four main foregrounds:

- (i) Galactic synchrotron
- (ii) Extragalactic point sources
- (iii) Galactic free-free emission
- (iv) Extragalactic free-free emission

Each of these processes emit signals in the frequency region of the redshifted HI signal i.e. $\sim 1420/(1+z)$ MHz. Each of them are dominant over the HI signal which is inherently weak. In some cases, such as galactic synchrotron, the foregrounds can be several orders of magnitude higher in observed brightness temperature. It is therefore immediately apparent that this a major challenge for the success of the HI intensity mapping technique.

Extragalactic point source foregrounds (ii) are caused by objects beyond our own galaxy emitting signals with wavelengths similar to the redshifted 21cm signal, a typical example being AGNs. (iii) & (iv) represent free-free emission which is caused by free electrons scattering off ions without being captured and remaining free after the interaction. In this weak-scattering interaction low-energy photons are produced which can enter the $21(1+z)$ cm wavelength window we are interested in. These free-free interaction signals can be detected both within (galactic free-free) and beyond (extragalactic free-free) our own galaxy.

Lastly the synchrotron emission (i) occurs when high-energy electrons are subject to an acceleration perpendicular to their velocity by the application of a magnetic field. This foreground is typically caused by relativistic cosmic ray electrons accelerated by the galactic magnetic field. It is the galactic synchrotron which is by far the most dominant of the foreground types and is therefore the one we would like to concentrate most on removing.

3.1 Galactic Synchrotron

While it would be fairly straightforward to simulate Gaussian realizations of galactic synchrotron from a model power spectrum, it is far more robust to make use of existing data and use this to emulate the shape of the emission on the sky. This also allows us to study the impact of subtracting a foreground which has wide structures, potentially eliminating information at large angular scales.

Unfortunately, foregrounds within the frequency range of the redshifted 21cm signal ($400 \text{ MHz} < \nu < 1420 \text{ MHz}$) are less well studied than other foregrounds, for example

| Foreground | A | β | α | ξ |
|--------------------------|-------|---------|----------|-------|
| Galactic synchrotron | 700 | 2.4 | 2.80 | 4.0 |
| Point sources | 57 | 1.1 | 2.07 | 1.0 |
| Galactic free-free | 0.088 | 3.0 | 2.15 | 35 |
| Extra-galactic free-free | 0.014 | 1.0 | 2.10 | 35 |

Table 3. Parameter values for foreground C_ℓ 's (see Equation (16) with amplitude A given in mK^2 . Pivot values used are $\ell_{\text{ref}} = 1000$ and $\nu_{\text{ref}} = 130 \text{ MHz}$ as per Santos et al. (2005).

those which impact the microwave background emission at higher frequencies ($\nu > 10 \text{ GHz}$). Therefore, obtaining actual data maps of galactic emission at regular frequency intervals in the range we are interested is challenging.

Following a method which has been used in similar HI foreground studies (Shaw et al. 2014; Wolz et al. 2014; Alonso et al. 2014) we use the Global Sky Model (GSM) (Zheng et al. 2017) to generate maps $T_{1420}(\vec{\theta})$ and $T_{400}(\vec{\theta})$ which are emission maps at 1420 MHz and 400 MHz, then use these to construct a full-sky spectral index given by

$$\alpha(\vec{\theta}) = \frac{\ln T_{1420}(\vec{\theta}) - \ln T_{400}(\vec{\theta})}{\ln 1420 - \ln 400}. \quad (14)$$

This is then used to extrapolate the Haslam map (Haslam et al. 1982), which is one of few all-sky maps for galaxy emission around these frequencies,

$$T_0(\vec{\theta}, \nu) = T_{\text{Haslam}}(\vec{\theta}) \left(\frac{\nu}{408 \text{ MHz}} \right)^{\alpha(\vec{\theta})}. \quad (15)$$

This can now be used to simulate a map of the sky at any desired frequency. However, since the Haslam map does not provide information beyond its own resolution ($\sim 1^\circ$), we need a further process to improve the resolution of these maps for any meaningful investigation of small scales.

We add in this additional small scale information through Gaussian realizations of an angular power spectrum which models galactic synchrotron emission. Following Santos et al. (2005) we make this construction using the angular power spectrum

$$C_\ell(\nu_1, \nu_2) = A \left(\frac{\ell_{\text{ref}}}{\ell} \right)^\beta \left(\frac{\nu_{\text{ref}}^2}{\nu_1 \nu_2} \right)^\alpha \exp \left(- \frac{\log^2(\nu_1/\nu_2)}{2\xi^2} \right), \quad (16)$$

where ξ is a parameter which regulates the spectral smoothness of the foreground such that smaller ξ cases are less smooth in frequency and are therefore more of a challenge to disentangle from the cosmological signal. The rest of the parameters are defined in Table 3. Figure 4(i) shows a full-sky map of the simulated galactic synchrotron emission for a frequency slice.

Galactic synchrotron has the added complication of being partially linearly polarized. This polarized portion can undergo Faraday rotation which changes the polarization angle of the radiation. The consequences for the HI signal have been studied in Jelic et al. (2010, 2008); Moore et al. (2013). Generally speaking this polarization response tends to erode the spectral smoothness of the signal, since it is a frequency dependent effect, and the induced spectral structure is problematic for separating the foreground from the cosmological HI signal. This requires excellent instrumental calibration to avoid leakages of the polarization effects. The simulation of such polarization leakage is complex and

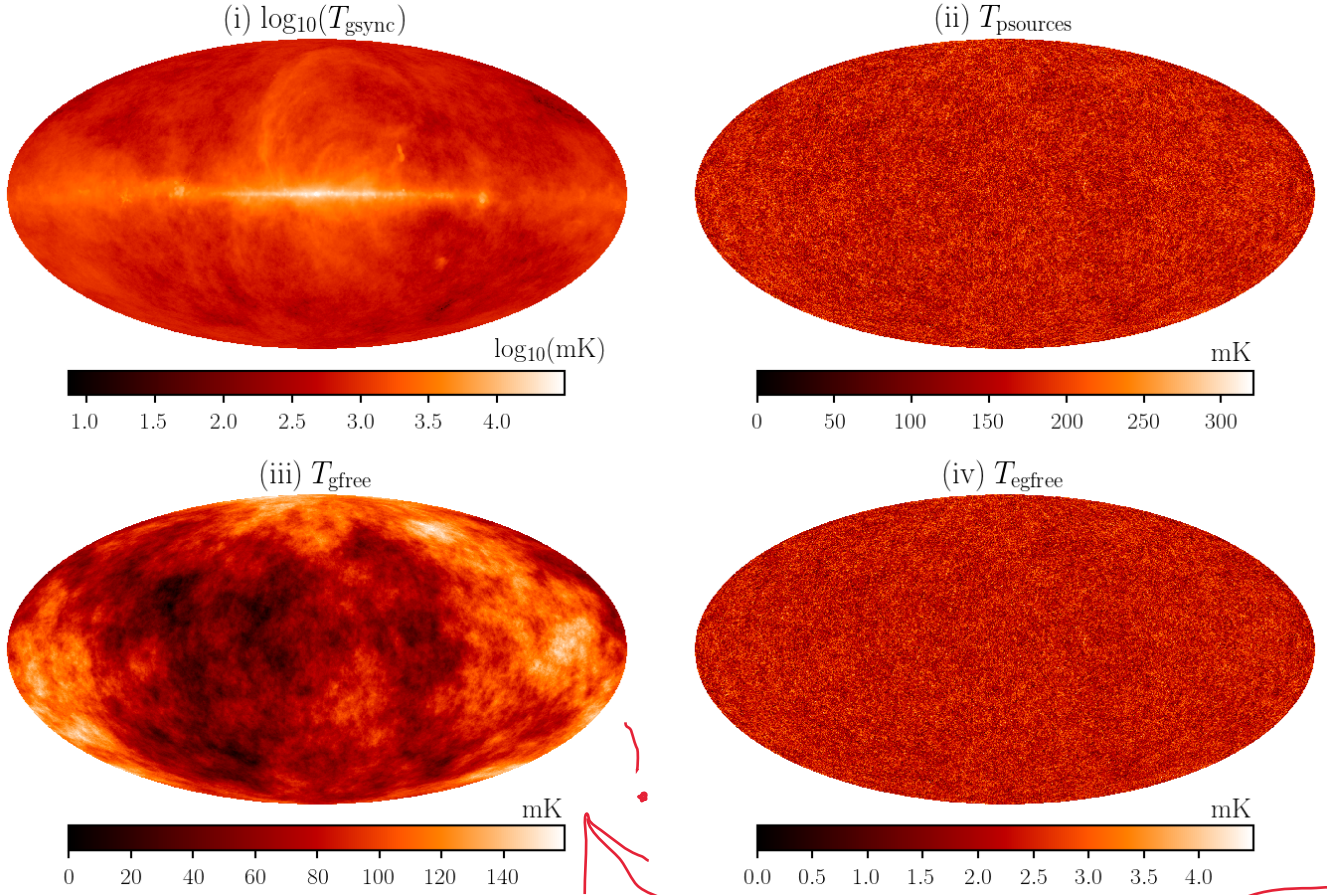


Figure 4. Full sky maps of each simulated foreground at a frequency of 1136 MHz ($z = 0.25$). These examples do not include noise or beam smoothing. All temperatures are in mK but the galactic synchrotron map (i) shows the logarithm of the temperatures.

instrument specific. For this work we do not simulate any polarization of the synchrotron emission, but we do opt to convolve all our maps at differing frequencies to a common resolution based on the maximum size of the instrument beam. This is thought to be an active step in mitigating the effects of polarization leakage and is something that is carried out in the Green Bank Telescope HI intensity mapping data analysis in [Switzer et al. \(2013\)](#).

3.2 Point Sources & Free-Free Emission

While galactic synchrotron dominates over all other HI foregrounds, it is still important to consider these additional contaminants since they still dominate over the HI signal. **Extragalactic point sources and extragalactic free-free emission are isotropic in nature, since they are sources beyond our own galaxy. Therefore it is realistic to simulate them with full-sky Gaussian realizations of the angular power spectrum we laid out in Equation (16) using parameters from Table 3.** This makes the assumption that the source of these foregrounds is Gaussian and also that there is no angular correlation between point sources and HI emitting galaxies. While point sources will cluster with the underlying matter density, the continuum signals they emit mean that in any one

redshift bin, angular correlation between point source signal and HI is likely to be small.

Galactic free-free emission is not expected to be perfectly isotropic and will have some galactic latitude dependence. However, because it has a low amplitude and very smooth frequency dependence, this will not be a difficult foreground to subtract and we therefore do not believe a more robust modelling is needed here.

For these three foregrounds, point sources, galactic free-free and extra-galactic free-free, we therefore use Equation (16) and the parameters from Table 3 for our simulations. Figures 4(ii),(iii) and (iv) shows maps of these three different foregrounds using the isotropic Gaussian realization approach we have outlined. The lack of galactic latitude dependence is immediately apparent in contrast to the galactic synchrotron map in Figure 4(i).

To complete this discussion on HI foregrounds we include the angular power spectra measured for each of the produced foregrounds in Figure 5 along with the cosmological signal. This immediately highlights the challenge faced when attempting foreground subtraction as it demonstrates how dominant all the foregrounds are over the cosmological signal we are trying to extract.

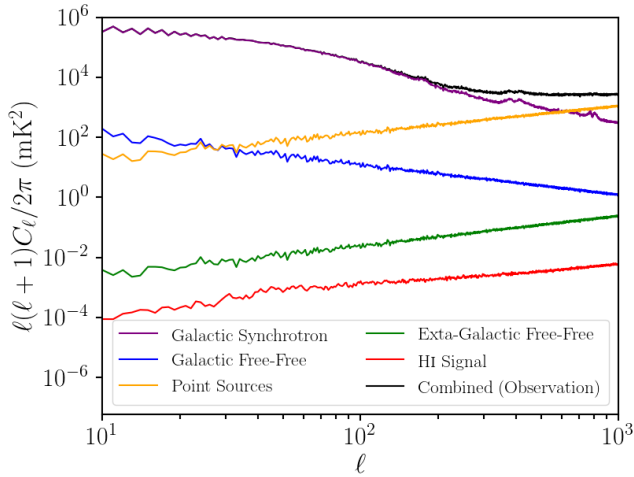


Figure 5. Angular power spectra for all the different simulated foregrounds, and the HI cosmological signal produced using the GAFA catalogue. The black solid line represents the combined signal from all foregrounds and the HI cosmological signal. All are at a frequency of 1136 MHz ($z = 0.25$) and noise free with no beam effects added.

3.3 Simulated Observable Signal

To summarize, our simulated sky signal is a composition of maps at certain frequencies (equivalently, redshifts) which can be described by

$$\delta T_{\text{sky}}(\nu) = \delta T_{\text{HI}}(\nu) + \sum_i \delta T_i^{\text{FG}}(\nu) \quad (17)$$

where the first term comes from the signal described in Section 2.1 and the second term is the contribution from all the different foregrounds outlined previously. Once these maps are combined we smooth the total temperature map δT_{sky} using the Gaussian beam given by Equation (11). We then add the simulated random noise from Equation (10) to emulate basic instrumental systematics, resulting in our final simulated observation δT_{obs} :

$$\delta T_{\text{obs}}(\nu) = \mathbf{S}_{\text{beam}} \left(\delta T_{\text{HI}}(\nu) + \sum_i \delta T_i^{\text{FG}}(\nu) \right) + \delta T_{\text{noise}}(\nu) \quad (18)$$

where \mathbf{S}_{beam} is the smoothing (or convolution) function.

4 FOREGROUND REMOVAL

While foregrounds pose a huge problem for the prospects of exploring cosmology with HI intensity mapping data, there are some features that help distinguish them from the cosmological 21cm signal. We can utilize the spectral smoothness of the foregrounds to separate them from the HI, which fluctuates with frequency. Figure 6 shows that along a LoS, the foregrounds are very smooth, whereas the expected signal from HI has a strong frequency dependence. It is this property that is utilized in a class of methods referred to as blind foreground subtraction. Less general ‘non-blind’ approaches would involve precise modelling of the foreground contamination. Given the lack of data for these foreground signals at the relevant frequencies, this approach is not currently viable.

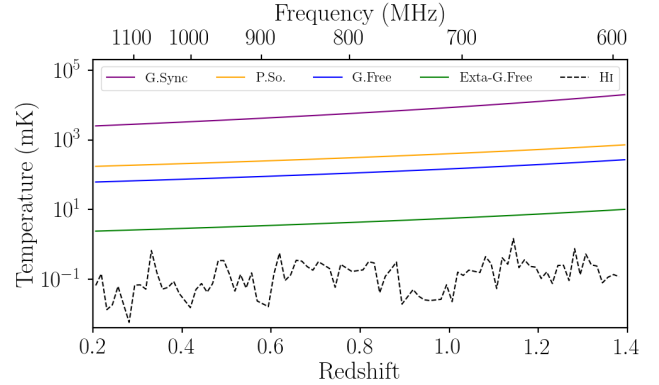


Figure 6. Observed brightness temperatures along a chosen LoS through frequency (or redshift). This is presented for the MICE catalogue with 100 redshift bins to show a large frequency range. The plot demonstrates the foreground smoothness in frequency (coloured solid lines), in contrast to the highly oscillatory fluctuations of the HI signal (black dashed line).

It is apparent however, that a foreground clean based on this distinguishing spectral smoothness would be more successful for small wavelength radial modes, whereas for larger wavelength radial modes the HI signal is more similar to the foregrounds. Hence these types of foreground cleans can render large Fourier radial modes (or small k_{\parallel}) useless. Removing large scale modes from HI intensity maps is therefore an expected effect of a foreground clean and was used as a toy model to emulate the effects of foreground cleaning in Cunningham et al. (2019). In the present work we extend our foreground investigation by directly contaminating the maps with the foregrounds we outlined in Section 3, and then use state-of-the-art foreground removal techniques to recover our HI input data and study the impact this will have on fundamental cosmological measurements.

There are several blind foreground removal techniques, for example principle component analysis (PCA) and independent component analysis (ICA) whose distinctions are outlined in Alonso et al. (2015). Further blind component separation methods include Generalized Morphological Component Analysis (GMCA) (Chapman et al. 2013) and Generalized Internal Linear Combination (GnILC) (Remazeilles et al. 2011). For this work we examine the FASTICA method (Hyvärinen 1999; Chapman et al. 2012; Wolz et al. 2014, 2017), which we describe in the following section. Alonso et al. (2015) found there to be very little distinction between a PCA and ICA approach to foreground cleaning, so our choice of FASTICA as a foreground removal process should not affect the generality of our conclusions.

4.1 FASTICA Formalism

Here we introduce the basic principles of the Fast Independent Component Analysis (FASTICA) technique, which we will utilize for foreground removal. For a more complete derivation and discussion we refer the reader to (Hyvärinen 1999). In a blind foreground removal problem we assume that a raw observed signal, such as that outlined in Equation (18), can be generalized into a linear equation where the

elements making up the signal are statistically independent. That is

$$\mathbf{x} = \mathbf{A}\mathbf{s}. \quad (19)$$

The dimensions and basic description of each term in this equation are given as:

$\mathbf{x} [N_z, 1]$: combined observed signal
 $\mathbf{A} [N_z, m]$: mixing matrix - determines the amplitudes of \mathbf{s}
 $\mathbf{s} [m, 1]$: independent components (containing foregrounds)

Practically this system will have some trace residuals which have some frequency dependence which will include instrumental noise, any residual foregrounds which cannot be classified into an independent component (IC), and the cosmological H_I signal. FASTICA aims to solve Equation (19) and identify each IC so that from the remaining residual, the H_I signal can be reconstructed. For each LoS, sorted into N_z redshift bins and assuming m independent components (ICs) are present, FASTICA assumes

$$\mathbf{x} = \mathbf{A}\mathbf{s} + \epsilon = \sum_{i=1}^{N_{IC}=m} \mathbf{a}_i s_i + \epsilon, \quad (20)$$

with $\epsilon [N_z, 1]$ the residual (containing H_I signal and noise).

Under the assumption that each independent component s_i is statistically independent, FASTICA attempts to solve Equation (20) by utilizing the central limit theorem. This states that the greater the number of independent variables in a distribution, the more Gaussian that distribution will be i.e. the probability density function (PDF) of several independent variables is always more Gaussian than that of a single variable. Hence, if we can maximize any statistical quantity that measures non-Gaussianity, then we can identify statistical independence and form a prediction for \mathbf{a}_i and s_i .

The parameter m must be pre-specified before calculations. This is the number of ICs that can be described by unique non-Gaussian descriptions and is not necessarily the number of different foregrounds one is aiming to find. It is typically assumed that $m \approx 4$ and FASTICA then works by obtaining 4 data vectors which are as statistically independent as possible. With FASTICA, going to a higher number of ICs than is required converges to the same result. However, the computational cost is increased so for efficiency, the lowest value for m which gives the best possible result is sought.

The FASTICA process considers all LoS simultaneously. Therefore for its calculations on maps with a number of pixels given by N_{pix} , the ICs \mathbf{s} in Equation (20) are actually maps, and hence an array with size $[m, N_{\text{pix}}]$, while \mathbf{x} and ϵ are arrays of size $[N_z, N_{\text{pix}}]$. Furthermore, as we will further explain below, FASTICA involves some expectation value calculations which rely on a number of samples and for this it uses the N_{pix} different LoS.

To obtain \mathbf{s} we start by inverting Equation (20), ignoring the residual term ϵ which will just be left over from signal not contained within the ICs. We can therefore write

$$\mathbf{s} = \mathbf{W}\mathbf{x}, \quad (21)$$

here \mathbf{W} is the weighting matrix, defined as the inverse of \mathbf{A}

in Equation (19). Under the assumption that the elements \mathbf{s} are as statistically independent as possible, FASTICA then begins maximizing the non-Gaussianity. For a measure of Gaussianity it uses negentropy $J(y)$, which for a variable y , is based on typical entropy $H(y)$ defined as

$$H(y) = - \sum_i P(y = a_i) \log P(y = a_i), \quad (22)$$

where $P(y = a_i)$ is the probability that y equals a possible value a_i . The modification made to obtain the negentropy $J(y)$ is

$$J(y) = H(y_G) - H(y), \quad (23)$$

where y_G is a unit-variance Gaussian random variable. In practice, negentropy is computationally hard to calculate and requires numerous realizations to obtain information on probability distributions. However, using the maximum entropy principle, we can write

$$J(y) \approx - \sum_i k_i [\langle G_i(y) \rangle_\theta - \langle G_i(y_G) \rangle_\theta], \quad (24)$$

where k_i are positive constants, G_i is referred to as the contrast function, and all pixels are utilized by averaging over them (this is denoted by $\langle \rangle_\theta$). For the contrast function, whilst practically any non-quadratic function will work, FASTICA mainly uses

$$G_1(y) = \frac{1}{a_1} \log \cosh(a_1 y), \quad G_2(y) = -\frac{1}{a_2} \exp(-a_2 y^2/2), \quad (25)$$

where $1 \leq a_1 \leq 2$ and $a_2 \approx 1$.

In a nutshell, FASTICA delivers a method of reconstructing the foreground signals as m ICs and then the residual ϵ between this reconstruction and the raw observed input map is the recovered cosmological H_I signal plus any receiver noise and residual foreground contaminants. A final point to include is that the mean temperature of the H_I cosmological signal is a smooth function of frequency and is therefore incorporated into the ICs of the analysis. This information is therefore lost and the residual maps are required to be renormalised to some model mean temperature or treated as δT observables as in Equation (9).

4.2 FASTICA Results

Here we seek to validate the FASTICA reconstruction process introduced in the previous Section 4.1 by presenting results from our simulations outlined in Section 3. Figure 7 shows the IC maps found after FASTICA has been applied. It is interesting to note that the third and fourth ICs clearly seem to pick up the galactic synchrotron angular shape whereas the second IC shows structure across the sky. The first IC is largely contained in the top half of the map, where the H_I cosmological signal lies for the GAEA catalogue. This suggests that it is this component which is collecting large radial modes which belong to the cosmological signal along with the $\bar{T}_{\text{H I}}$ average which smoothly fluctuates and therefore is removed. Despite trying a number of different values of m (the number of ICs) it appears that it is always the case that some cosmological signal will be removed. These

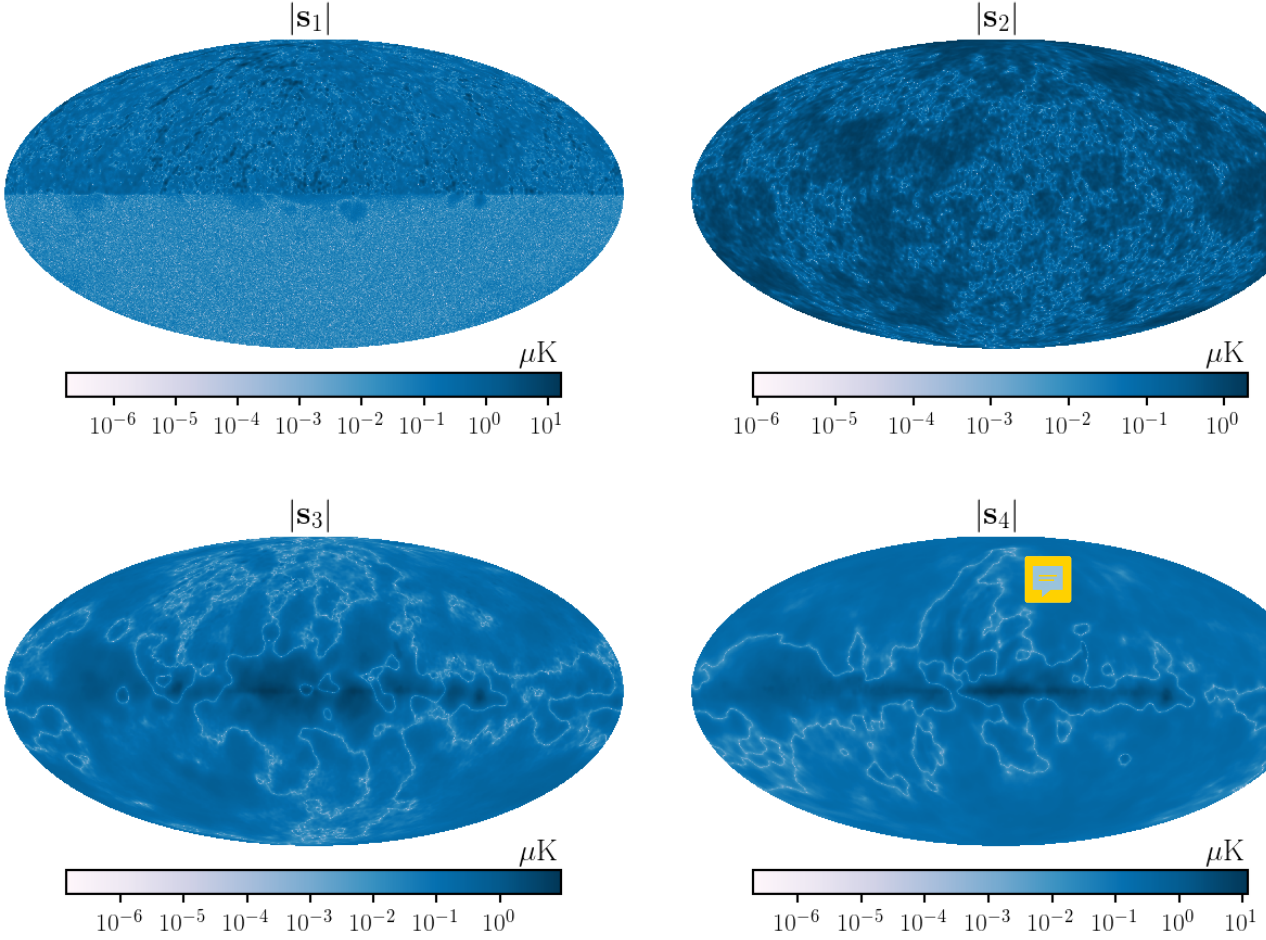


Figure 7. Independent component maps found using FASTICA with $m = 4$ on the GAEA simulation contaminated with foregrounds. This is for a constant beam of $\theta_{\text{beam}} = 0.5^\circ$ at all frequencies. Temperature fluctuations are given in μK but the true amplitudes for the estimated foregrounds are determined by their combination with the mixing matrix.

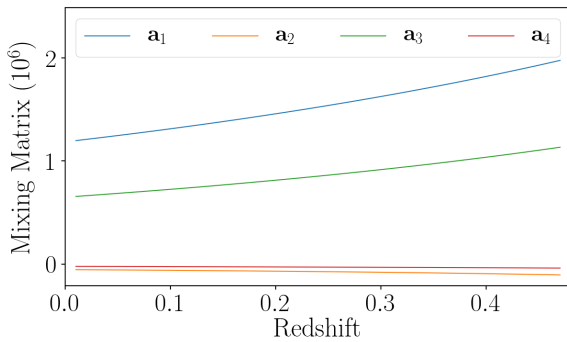


Figure 8. Mixing matrix elements as outlined by Equation (20). Combination of these with the independent components in Figure 7 determines the subtraction to be made from the combined observed signal at each frequency.

ICs from Figure 7 are then combined with the mixing matrix (displayed in Figure 8) as described in equation (20).

Figure 9 shows a pixel-by-pixel comparison between original values in the δT_{HI} intensity maps and the cleaned values for some selected redshift bins in our GAEA simulation. For a perfectly performing reconstruction we would

obtain all values along the red diagonal line, i.e. all values would match their originals. We can see that this is not the case but largely FASTICA is performing reasonably well with a Pearson correlation coefficient of $\rho \geq 0.93$ for most redshifts. We expect a value of $\rho = 1$ for a perfectly performing foreground clean indicating perfect correlation between original and cleaned maps. Figure 9 also shows that FASTICA performs better at the mid-ranges of redshift. This is not a redshift specific effect since we also see similar results in the MICE model where the best agreement is at redshift $z \sim 0.8$ which is the mid-redshift for its range. This suggests that there are some edge effects in the FASTICA process causing it to be less effective at the extreme radial ends of the input data, a result previously noted e.g. Wolz et al. (2014).

Figure 10 indicates how well the HI auto power spectrum can be recovered with FASTICA and shows how varying the number of ICs affects the recovery. We show results from both simulations, and it is interesting to note the difference between the two. We see that with GAEA only 3 ICs are needed for a successful reconstruction, however for MICE even 4 ICs is not sufficient for good agreement at large scales (small- ℓ). We tested a larger number of ICs with little improvement. The difference in results is probably due to the fact that MICE has a smaller sky coverage (25% of GAEA)

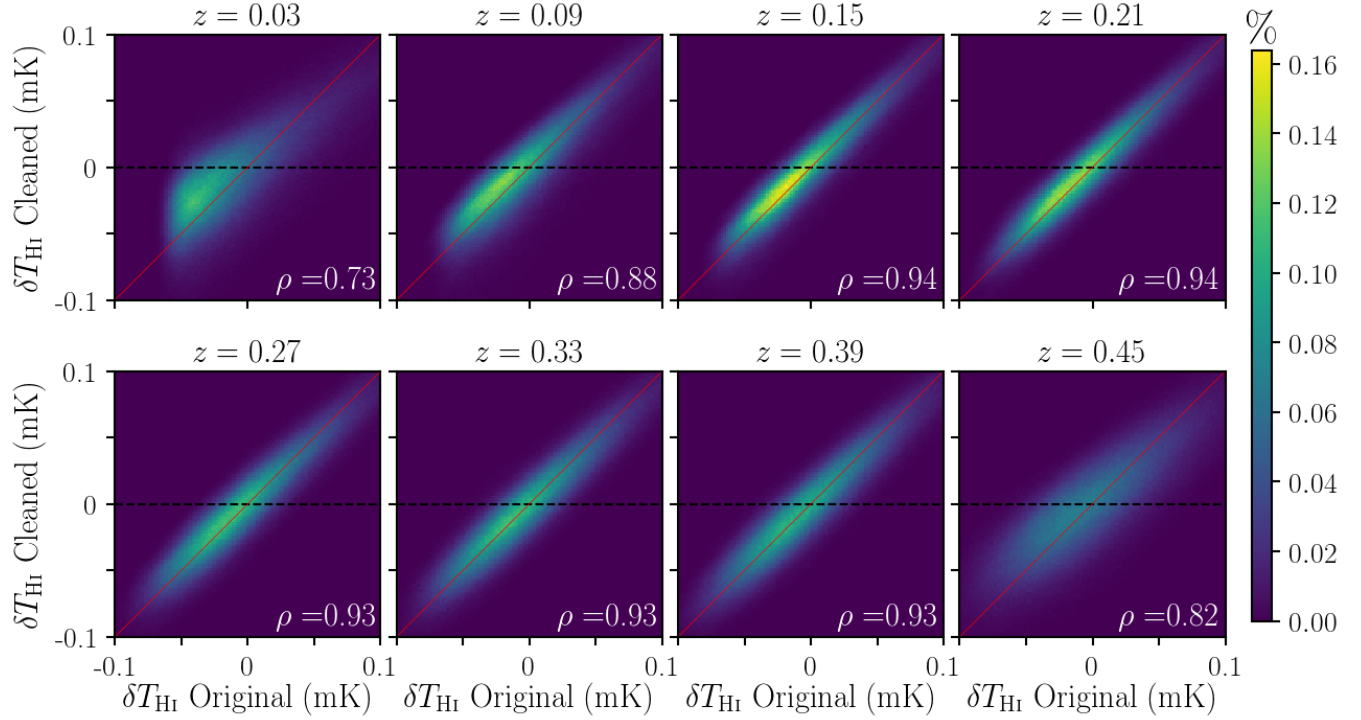


Figure 9. Histogram showing the original HI temperature against the FASTICA reconstructed value for each pixel in a range of redshift bins for the GAEA model. Each histogram has been normalized such that the histogram values sum to 100%. We also include the Pearson correlation coefficient ρ for each redshift to quantify the agreement. For a perfectly working foreground clean we would expect an entirely one-to-one ($\rho = 1$) agreement along the thin diagonal red line. We can see how FASTICA is less effective at extreme ends of redshift range with a wider dispersion of values.

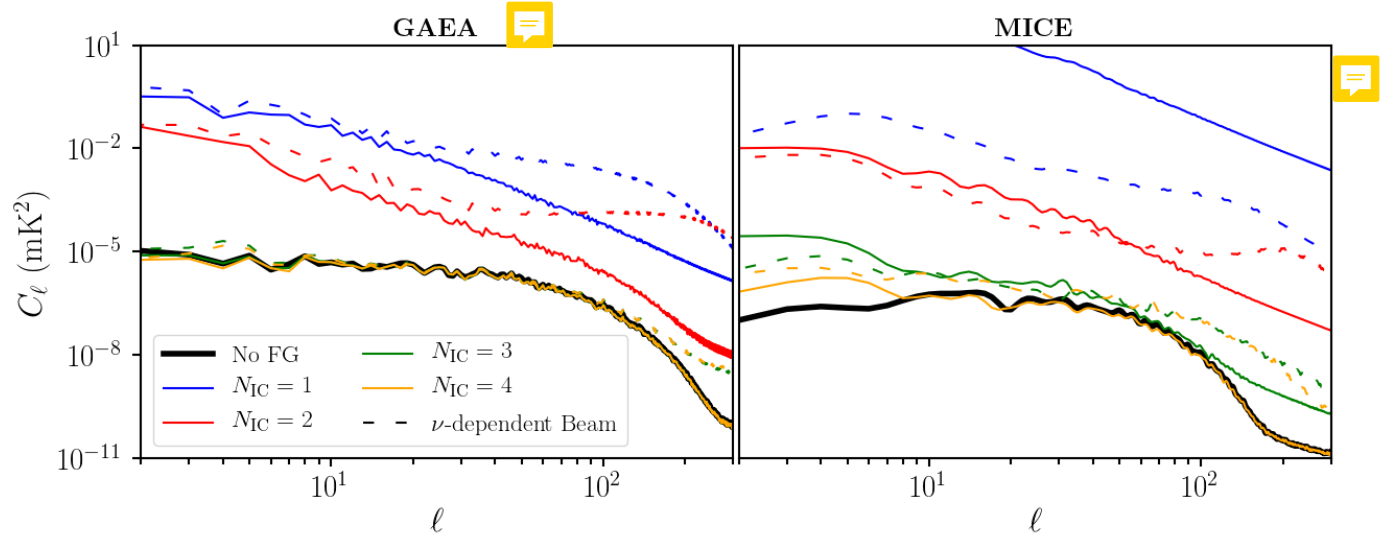


Figure 10. Impact of foregrounds on the HI auto-power spectrum for both the GAEA and MICE catalogues. Thick solid black line shows the original HI signal with no foregrounds. The coloured lines then show different values of m used i.e. the number of independent components assumed in the FASTICA process. We also include the results from using a beam which varies with frequency (dashed lines) and how this damages performance. For GAEA it appears that there is little need to go beyond 3 independent components for a successful reconstruction in the constant beam size case. Whereas for MICE even going to 4 independent components there is still some disagreement at large scales (small- ℓ). The disagreement we see is caused by residual foreground which FASTICA has failed to remove. Similar results have been previously shown in e.g. Wolz et al. (2014). These results are for mid-range redshifts for each catalogue with $z = 0.25$ for GAEA and $z = 0.825$ for MICE.

which means less samples to average over for negentropy estimation in equation (24). Furthermore, since MICE has a deeper redshift range (extending to $z = 1.4$ whereas GAEA is only up to $z = 0.5$) the constant beam size that we convolve with is much larger for MICE, $\theta_{\text{beam}} = 2.36^\circ$ against GAEA's $\theta_{\text{beam}} = 1.46^\circ$. This difference in beam size is also evident from the scales at which the power spectrum seems to degrade. Due to its larger beam, the MICE power spectrum begins to tail off at lower- ℓ than GAEA. Lastly this plot also includes results where each tomographic slice has been smoothed by a varying amount due to the frequency dependence of the beam. This is shown as the dashed lines, and it is evident that results are much worse when compared with the constant beam case. This is discussed further in the following section.

4.2.1 Frequency Dependent Beam Size

As previously outlined in Section 2.1.2, the intensity maps at different frequencies will have different beam sizes defined by equation (11), meaning intensity maps at lower redshift have less degradation of angular scales. However, since FASTICA finds m IC maps and then subtracts these from the total observation based on the mixing matrix A , trying to obtain e.g. 4 IC maps based on N_z intensity maps with different resolutions for each will cause problems because the IC map resolution will not match each of the intensity maps. This is exactly why we see poorer performance in Figure 10 in the case where there is a frequency dependent beam size (dashed lines) especially at smaller scales (large- ℓ) where the beam has a more dominant effect.

The way we resolve this issue is by carrying out a further convolution on the intensity maps such that each tomographic slice is smoothed to the same resolution. We therefore take the maximum beam for the particular redshift range and smooth over all maps with this constant beam size. FASTICA then finds IC maps which, when subtracted from the observed signal, prove more effective for reconstructing the original HI signal as shown by the solid lines in Figure 10.

Artificially re-smoothing over all our intensity maps may appear to be a wasteful process in terms of loss of large- ℓ modes, but it is necessary for a successful FASTICA reconstruction. In fact, this has additional benefits when dealing with real data, as an effective way of mitigating the effects of polarization leakage. For example, Switzer et al. (2013) include a constant convolution over their data, reducing their resolution to below that induced by the beam alone, which limits the contamination caused by polarization leakage, which can degrade the spectral smoothness of the foregrounds.

4.2.2 Increasing the Number of Frequency Bins

For both our GAEA and MICE simulations we are only using 24 redshift (frequency) bins with the bin width determined by a constant separation in redshift Δz . This may be seen as quite a low number of bins to be using in an intensity mapping simulation which uses an ICA process. This is largely out of necessity due to the choice of simulation approach: since we are using N -body simulations we have a finite number of galaxies to use from which to build intensity maps.

By using bins which are too thin we risk under-sampling the intensity maps and making them an unrealistic emulation of a continuous field of emission.

In practice when using real data, the typical approach would be to perform the FASTICA method on more maps (> 100 frequency channels), then re-stack these into fewer bins for cosmological analysis and cross-correlations with optical data. We trialed this with our MICE catalogue using 240 bins, and found that it made no improvement on the FASTICA foreground removal, hence justifying our choice of using 24 frequency bins in all our analysis. Furthermore, in cross-correlation with optical imaging surveys, the intensity maps would likely need stacking into thicker bins to match the thick bin widths needed for the poorly constrained photometric redshifts.

5 HI \times OPTICAL COSMOLOGY WITH FOREGROUNDS

An important measurement in cosmology is the angular clustering of a matter density tracer. In order to apply this with HI intensity maps, we measure the angular power spectrum by decomposing the temperature fluctuations into spherical harmonics $Y_\ell^m(\hat{n})$:

$$\delta T_{\text{HI}}(\hat{n}, \nu) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m}(\nu) Y_\ell^m(\hat{n}). \quad (26)$$

The harmonic coefficients $a_{\ell m}(\nu)$ describe the amplitudes of the fluctuations in spherical harmonic space; we can then define the angular power spectrum between tracers X and Y as

$$C_\ell^{XY}(\nu_1, \nu_2) = \langle a_{\ell m}^X(\nu_1) a_{\ell m}^{Y*}(\nu_2) \rangle. \quad (27)$$

This can either represent the HI intensity map auto-correlation where $X = Y = \text{HI}$, or the cross-correlation with the optical galaxies where $X = g$ and $Y = \text{HI}$. If these tracer fields are Gaussian, the power spectrum is a complete statistical representation of the fields. We have already used these quantities to examine both the foreground dominance over the HI signal in Figure 5 and to validate the FASTICA reconstruction process in auto-correlations in Figure 10. In this section we investigate the impact that HI foreground contamination and removal with FASTICA have on the cross-correlation power spectra with our simulated optical catalogues.

In recent work, Blake (2019) has developed a framework which models observational effects on 3D power spectra for HI-optical cross-correlations. This framework can be extended to include the effects of foreground removal and photometric redshift uncertainty. By doing this one could analytically model the foreground removal effects, as well as the photometric redshift effects, as a loss of small and large k_\parallel modes respectively, and attempt quantitative corrections accordingly. In this work we aim to use our simulations to investigate what corrections can be made to the data to extract the most information from these cosmological measurements.

To begin exploring how HI foregrounds can impact cross-correlations with optical surveys we first perform a best-case scenario test and cross-correlate with an optical

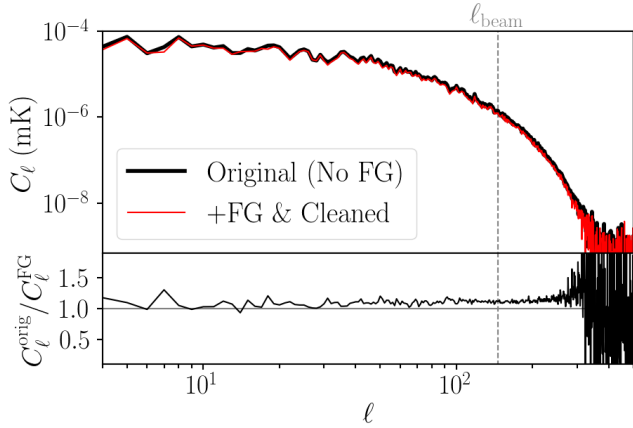


Figure 11. Cross-correlation angular power spectrum between the HI intensity map at redshift $z = 0.25$ and the optical galaxies binned using their true redshifts. This is representative of a scenario in which spectroscopic redshifts are used in the optical survey. The original result with no foregrounds is shown as the black thick line and the case where foregrounds have been included then removed by FASTICA is shown as the red thin line. The bottom panel shows the ratio of the two spectra. This test was carried out on the GAEA simulation where the HI intensity maps have been re-smoothed with a constant maximum beam of $\theta_{\text{beam}} = 1.46^\circ$.

survey which we assume has very well constrained redshifts; Figure 11 shows the result of this cross-correlation. Here we bin the optical galaxies from the GAEA simulation by their true redshift so we have a sample of optical galaxies at $z = 0.25$ to cross-correlate with an HI intensity map at the same redshift. This shows that foregrounds should have little impact on optical spectroscopic cross-correlations. The bottom panel shows a small bias which in principle could be corrected for by constructing a foreground cleaning transfer function (Switzer et al. 2015), but it is encouraging that our initial efforts have already reconstructed the cross-power to within 8.5% at scales below those unaffected by the beam ($\ell < \ell_{\text{beam}}$). It is only at higher ℓ , way below the resolution of the beam (ℓ_{beam}), that we start to have large errors on C_ℓ . This is unsurprising since this is going beyond the scales of the radio instrument’s resolution. We experimented with smoothing the optical field to replicate the HI intensity map resolution but find no mitigation of the noise we see at $\ell > 250$.

5.1 Optical Redshift Uncertainty

Future optical galaxy redshift surveys such as LSST and *Euclid* will rely on using photometric redshifts for estimating the radial position of each galaxy (note that *Euclid* will also perform a wide spectroscopic survey). It is therefore important to forecast the cross-correlation potential between HI intensity maps and photometric galaxy redshift surveys, taking into account foreground removal effects. The higher uncertainty on redshift measurement inherent in these photometric surveys, equates to a degradation in radial mode measurement on small scales. Since foreground removal also impacts radial modes but on larger scales, it is unclear whether

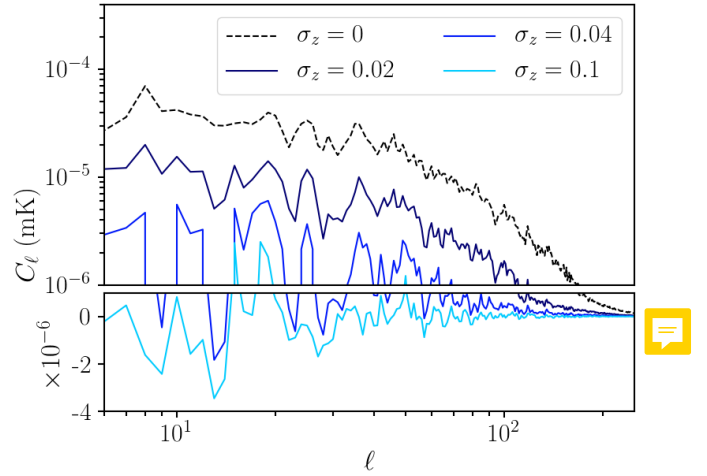


Figure 12. Cross-correlation between HI intensity maps with FASTICA reconstruction and an optical survey using GAEA. We degrade the constraints on the optical galaxy redshifts by increasing the redshift error σ_z shown by going from dark to lighter blue. In other words we go from cross-correlating intensity maps with a spectroscopic-like ($\sigma_z \sim 0$) survey, to a photometric-like survey where there is significant uncertainty on the optical galaxy redshifts. This strongly affects the measured cross-correlation power spectrum. Plot includes a hybrid log-linear y-axis to fully demonstrate the degradation in power.

combining these two effects will leave enough useful modes for a cross-correlation signal (Witzemann et al. 2018).

To investigate this we can begin by simply introducing a Gaussian error on our optical redshifts for each galaxy and cross-correlate with foreground contaminated intensity maps. Figure 12 shows the effect on the cross-power spectrum when we introduce a Gaussian photo- z error σ_z into each of the optical galaxies. We can see how increasing the uncertainty in redshift (dark to light blue lines) rapidly degrades the agreement with the original (black-dashed line) where no redshift error is applied. Abell et al. (2009) suggests a fiducial model of $\sigma_z = \sigma_{z0}(1+z)$ is appropriate for an LSST-like instrument, where $\sigma_{z0} = 0.05$. Therefore, the fact that Figure 12 suggests the cross-power spectra signal-to-noise will be damaged for $\sigma_z \sim 0.1$, which would correspond to LSST’s photo- z error at $z = 1$, is cause for concern.

We can further explore this with the use of some more robust photometric redshift simulations and compare to foreground free cross-correlations. Realistic photometry for a number of optical surveys is included within the MICEv2 simulation, for example the Dark Energy Survey (DES)⁶. We thus make use of the DES-like photometric redshifts available to make a more robust forecast of the cross-correlation between a photometric survey and HI intensity maps. We refer the reader to the MICE website⁷ for more details on how these DES-like photometric redshifts were simulated.

Figure 13 shows the results when we include these simulated DES-like photometric redshifts in our simulations. The dashed lines show the cross-correlation power spectrum with

⁶ www.darkenergysurvey.org/

⁷ <http://maia.ice.cat/mice/>

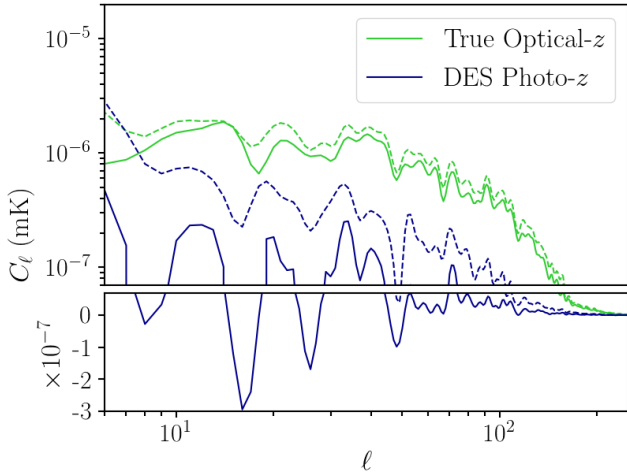


Figure 13. Cross-correlation between HI intensity maps with MICE optical galaxies. Dashed lines show the cases without HI foregrounds, solid lines show the impact of including them. We use the DES-like photometric redshifts available in MICE for the photo-z forecasts shown in blue and compare these with using ideal true redshifts (green). While a drop in signal is inevitable when using less constrained redshifts, including the effects of HI foregrounds (solid lines) degrades the signal further in the photo-z case. These tests have been performed at redshift $z = 0.725$.

the original HI intensity map with no foreground contamination. The solid lines then show the inclusion of foregrounds and a FASTICA reconstruction. What is clear from this plot is that while we still get a degradation in signal from using photometric redshifts (blue line) compared with true redshifts (green line), the signal deterioration accelerates in the case where HI foregrounds are included in the simulation.

The conclusion from our GAEA simulation using Gaussian photometric redshifts and MICE using DES-like photometric redshifts appears to be the same and both forecast damaging signal loss when FASTICA reconstructed intensity maps are cross-correlated with photometric redshift surveys.

5.2 Mitigating the Effects of FASTICA

Here we begin investigating the precise reasons why combining the effects of HI foregrounds and the poor redshift constraints from photometric galaxy surveys is so detrimental to the cross-correlation signal. Generally, it can be considered unsurprising that combining an effect that removes information at large radial modes, with a survey which has poor constraints at small radial modes, can damp the amplitude of projected angular power spectra, as we see in Figures 12 and 13. Our aim here is to quantify this explanation with the hope of being able to provide a solution.

It is interesting to look at the effects FASTICA has along the line-of-sight (LoS) of the HI intensity mapping data. It is known that large radial modes are removed since this is where the contamination from foregrounds lies due to their smooth variation in frequency. Figure 14 shows the specific effect this has and illustrates how FASTICA removes all information on the mean temperature along the LoS. Our simulations are arranged such that the transverse mean of each

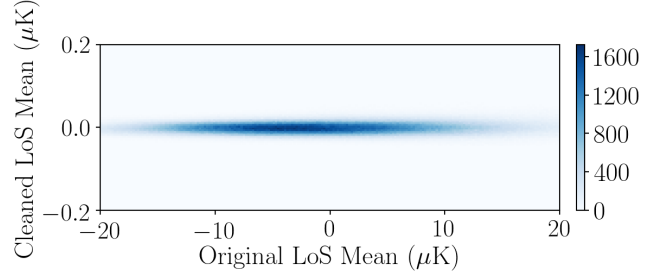


Figure 14. The mean δT temperatures along the line-of-sight (LoS) for the original HI intensity map against one with which has undergone a FASTICA foreground clean. This is shown for all available LoS in the GAEA simulation which for $f_{\text{sky}} = 0.5$ and $n_{\text{side}} = 512$ equates to over 1.5 million pixels (or LoS). The plot shows how FASTICA essentially removes any non-zero LoS mean present in the original HI signal and collapses it to zero.

map is zero but even with this setup it is of course still possible to have a large range of values for the LoS mean temperatures, which is what we see in Figure 14. However, we can see that the large range of LoS mean values present in the original HI signal (shown on the x-axis) are removed after the foreground clean to a much narrower range (shown on the y-axis). It is worth pointing out that the y-axis range is two orders of magnitude smaller than the x-axis. So essentially a FASTICA reconstruction destroys any non-zero mean along the LoS. The original line-of-sight means have a slight skewness away from zero and centre at around $-4 \mu\text{K}$. This is caused by the presence of some dominant bright pixels which, when setting transverse means in each map to zero, can result in there being more negative temperatures than positive ones.

It is conceivable that an increase in the number of redshift bins could effect this LoS result, so we therefore conducted a test using the MICE catalogue and extended to 240 redshift bins following the same procedure. Even with this more realistic number of redshift bins we still find a similar result to Figure 14 suggesting that this is not a feature of the relatively low number of redshift bins we are using.

In summary, the problem is that while FASTICA reconstructs the shape of the LoS signal, unfortunately it changes the amplitudes in an unpredictable manner based on the original LoS mean. The further from zero a particular LoS mean is, the greater the change in amplitude for pixels along this LoS. We attempt to model this by hypothesising that in a FASTICA foreground clean the main resulting change is given by

$$\delta T_{\text{clean}}(\vec{\theta}, \nu) \sim \delta T_{\text{orig}}(\vec{\theta}, \nu) - \overline{\delta T}_{\text{LoS}}(\vec{\theta}), \quad (28)$$

where $\overline{\delta T}_{\text{LoS}}(\vec{\theta})$ is the mean fluctuation along a LoS for a pixel at position $\vec{\theta}$,

$$\overline{\delta T}_{\text{LoS}}(\vec{\theta}) = \frac{\sum_i \delta T_{\text{orig}}(\vec{\theta}, \nu_i)}{N_z}, \quad (29)$$

where the summation is over the N_z number of frequency (or redshift) bins.

Figure 15 shows the impact along the LoS resulting from the effect outlined in Equation (28). We have chosen two pixels and show their δT values through redshift, taking two

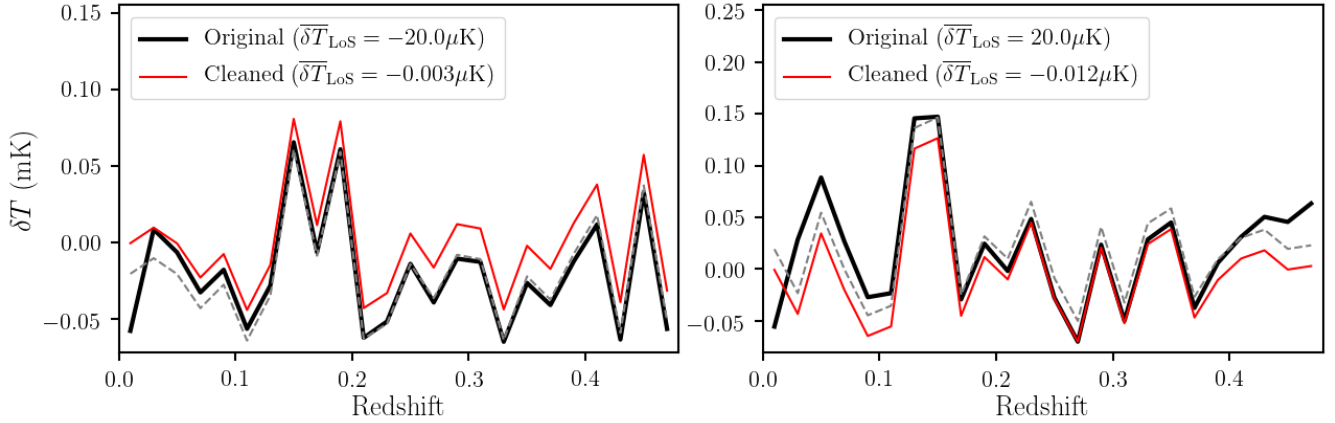


Figure 15. GAEA δT amplitudes along chosen lines-of-sight (LoS). Original mean values along the LoS are given in the legend along with the cleaned ones. The thick black line shows the original amplitude and the red solid line shows the impact of a foreground contamination and FASTICA foreground clean. The grey dashed line shows the amplitude with the LoS mean added back on as outlined in Equation (30).

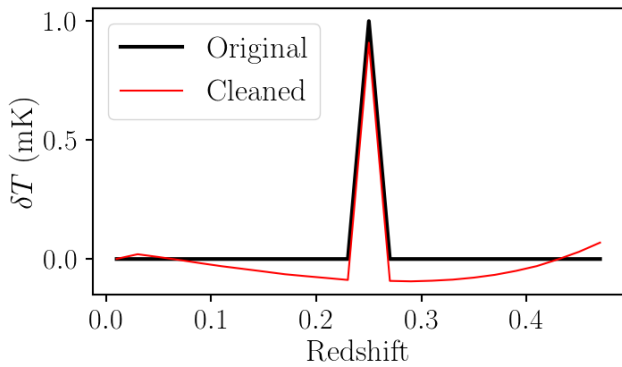


Figure 16. Effect of FASTICA on a test response function. For the GAEA model, all values along a chosen LoS have been set to 0 except one at $z = 0.25$ which is set to 1. This data is then subject to a FASTICA clean. An amplitude change from the LoS mean removal is apparent and there are also under-dense side lobes either side of the temperature spike.

extreme examples for demonstrative purposes. The plot on the left is for a pixel where the original LoS mean $\overline{\delta T}_{\text{LoS}}(\bar{\theta})$ is from the extreme low end from Figure 14. The plot on the right is for a pixel with a high $\overline{\delta T}_{\text{LoS}}$. In both cases their LoS means are collapsed to zero for the reasons discussed above and the impact this has on the agreement between individual values through redshift is evident.

We can demonstrate that this is the main impact of FASTICA by reversing the effect, i.e. adding back in the original LoS mean to each foreground-removed pixel:

$$\delta T_{\text{Hi}}(\bar{\theta}, \nu) = \delta T_{\text{clean}}(\bar{\theta}, \nu) + \overline{\delta T}_{\text{LoS}}(\bar{\theta}). \quad (30)$$

The corrected δT_{Hi} should agree with the original signal δT_{orig} . We have tested this and find this to be the case and show the results of this approach in Figure 15, where we included the reconstructed LoS based on Equation (30) shown by the gray dashed line.

Unfortunately, this LoS HI mean reconstruction is chal-

lenging in reality. The original $\overline{\delta T}_{\text{LoS}}$ will be information buried in the foreground contaminated maps, and which is then lost after the foreground clean. So performing the process outlined in equation (30) would require some extra information to reconstruct these LoS means.

In a similar demonstration to Figure 15, we also analyse the FASTICA result on a test response function in the form of a Dirac-delta spike in temperature, shown in Figure 16. By manipulating the GAEA data such that all pixels along a chosen LoS are set to 0 except for one which is set to 1, we can gain a deeper insight into the effects of a FASTICA foreground clean. The large side lobes which form either side of the temperature spike can explain why the cross-correlation with photometric galaxy data is performing so badly. A galaxy at $z = 0.25$ with high measured redshift uncertainty, is likely to cross-correlate with the false under-temperature regions. This effect, compounded over many galaxies and temperature spikes, could cause signal loss.

As an additional problem, we also find that FASTICA foreground removal is less successful at the extremes of the redshift range (something already concluded from Figure 9). Therefore reconstructing the LoS means will not be a sufficient correction on its own at the redshift edges of the data.

All this highlights the problems for the future success of HI intensity mapping cross-correlations with photometric galaxies. Nevertheless, photometric galaxy surveys are an important choice of probe to cross-correlate with given their complementary strengths, i.e. good angular resolution for optical and good radial resolution for HI intensity maps. We therefore suggest potential methods to mitigate the effects which a FASTICA foreground clean has on HI intensity maps. These not only serve to drastically improve cross-correlations with photometric optical data, but also provide additional improvements in cross-correlations with spectroscopic galaxy surveys, as well as HI intensity mapping auto-correlations. The two methods we propose are:

- **LoS Mean Reconstruction:** This is theoretically possible using optical galaxies which measure density

along the LoS. By relating the optical over-density to the HI temperature we can make a prediction for the LoS mean HI temperature that has been removed and reverse the effect of this loss of information.

- **Artificial Extension to Redshift Range:** Introducing a buffer at either end of the data sets in the redshift (or frequency) direction will limit edge effects and as we will demonstrate, improves the general agreement with the original data.

We discuss both of these methods in more detail in the following sub-sections.

5.2.1 Line-of-Sight Mean Reconstruction

While recovering the exact LoS means from the intensity map data is not possible (they are inaccessible before the clean, and removed after it) we can make predictions of what they are from other data. Then by measuring the power spectrum of the LoS mean predictions, we can reverse the effects of the mean removal. Making use of our hypothesis in Equation (30) we can write

$$\langle \delta_g \delta T_{\text{HI}} \rangle = \langle \delta_g \delta T_{\text{clean}} \rangle + \langle \delta_g \overline{\delta T}_{\text{LoS}} \rangle, \quad (31)$$

and similarly for the auto-correlation we have

$$\langle \delta T_{\text{HI}} \delta T_{\text{HI}} \rangle = \langle \delta T_{\text{clean}} \delta T_{\text{clean}} \rangle + 2 \langle \delta T_{\text{clean}} \overline{\delta T}_{\text{LoS}} \rangle + \langle \overline{\delta T}_{\text{LoS}} \overline{\delta T}_{\text{LoS}} \rangle. \quad (32)$$

Therefore, for a cross-correlation we require the correction term $\langle \delta_g \overline{\delta T}_{\text{LoS}} \rangle$ and for an auto-correlation we require $2 \langle \delta T_{\text{clean}} \overline{\delta T}_{\text{LoS}} \rangle + \langle \overline{\delta T}_{\text{LoS}} \overline{\delta T}_{\text{LoS}} \rangle$. We can utilise the optical number density fields to make estimates for these terms. This is because we can relate the optical over-density $\delta_g = b_g \delta_M$ to temperature fluctuations $\delta T_{\text{HI}} = \bar{T}_{\text{HI}} b_{\text{HI}} \delta_M$ through

$$\delta T_{\text{orig}}(z_i) = \frac{\bar{T}_{\text{HI}}(z_i) b_{\text{HI}}(z_i)}{b_g(z_i)} \delta_g(z_i). \quad (33)$$

Then we relate this to each LoS mean by

$$\overline{\delta T}_{\text{LoS}} = \frac{1}{N_z} \sum_i \frac{\bar{T}_{\text{HI}}(z_i) b_{\text{HI}}(z_i)}{b_g(z_i)} \delta_g(z_i). \quad (34)$$

This is all that is required to construct the correction terms for the cross- and auto-correlations outlined by Equations (31) and (32). This approach does not require precise optical redshift information for the $\delta_g(z)$. It is sufficient to use the poorly constrained photometric redshifts since the error on these should not heavily impact on the slowly varying summation kernel $\bar{T}_{\text{HI}}(z) b_{\text{HI}}(z) / b_g(z)$.

5.2.2 Artificial Extension to Redshift Range

While the reconstruction of the LoS means works reasonably well for the mid-range redshifts, improvements can still be made especially to the edge effects caused by a FASTICA foreground clean. These edge effects have been previously noted and suggestions have been made to exclude these contaminated regions Wolz et al. (2014, 2015). One simple solution to mitigate this effect and limit the data excluded, is to extend the range of the data with the idea that the new

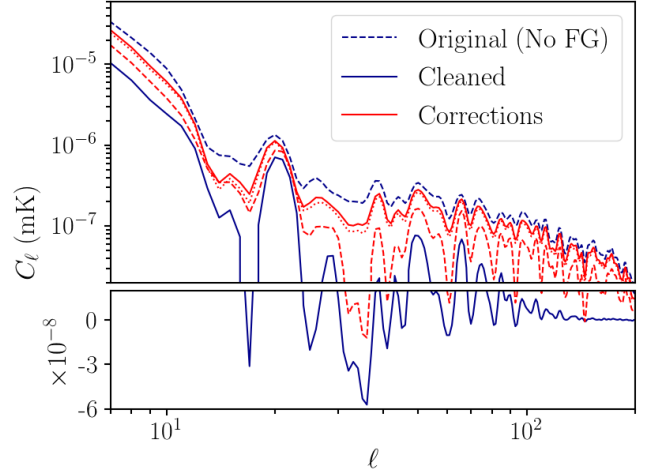


Figure 17. Cross-correlation angular power spectrum for the MICE simulation at redshift of $z = 1.075$ and like Figure 13 we have used the DES-like photometric redshifts available in MICE. Again the impact from foregrounds is visible in the difference between the blue dashed line and blue solid line. However, the effectiveness of the corrective techniques that we outlined in Section 5.2, shown by the red lines, is encouraging. The dashed red line is for the LoS mean correction, the dotted red line represents the extended- z correction and the red solid line represents both corrections applied.



artificial edges suffer the edge effect problems, but can then be removed from the rest of the data. We therefore pad the redshift bins such that for N original redshift bins we have

$$[z_1, z_2, \dots, z_N] \Rightarrow [-z_N, \dots, -z_1, z_1, \dots, z_N, \dots, z_{2N}].$$

The biggest challenge with this method is ensuring the padded data includes continuous foregrounds since this is what FASTICA needs to utilise to remove them. We therefore extrapolate the data by replicating and concatenating the observed signal through redshift before and after the main data.

Figure 17 shows the performance of these corrections on the MICE catalogue. We have shown this at a redshift of $z = 1.075$ which is closer to the extreme end of the redshift range for MICE and therefore has more need for correction. The solid blue line which shows the cross-correlation signal for FASTICA foreground cleaned map demonstrates how poor the signal is without any correction. The solid red line then shows that with the artificial extension to the redshift ranges and the LoS mean corrections to the power spectrum outlined by Equation (31), the signal is significantly recovered and approaches the original signal with no foregrounds (blue dashed line).

We also demonstrate the more general improvement made across all redshift bins with Figure 18 which is for the GAFA simulation. Using the relative difference between original and clean power spectra as a gauge of performance (stated above the colour-bar), this shows how improved the signal is across all redshifts and scales with the corrections in place. We still see some poor disagreement in the very first redshift bin and slightly poorer performance for the last few

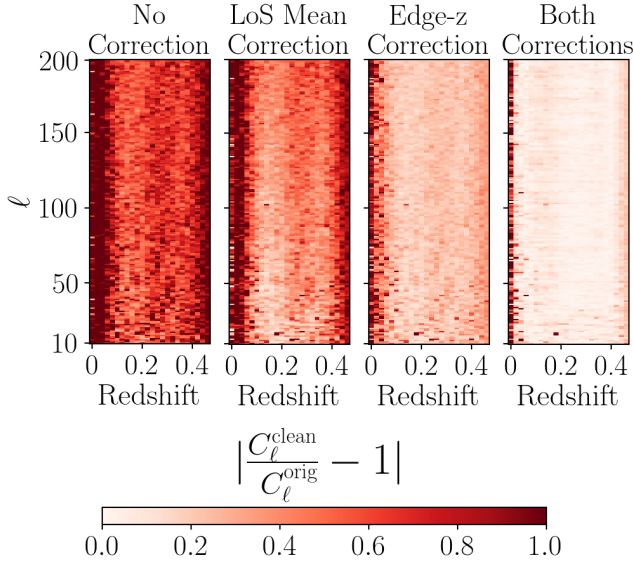


Figure 18. Demonstration of improvement on cross-correlation by including the corrections to the data outlined in Section 5.2. This is for the GAEA data-set and shows relative differences for cross-correlation of optical photometric-like data with HI intensity maps for the original (no foregrounds) and cleaned cases. For the optical sample we used a catalogue with redshift error of $\sigma_z = 0.06$.

bins, but the catastrophic discrepancies that we were seeing previously have been addressed.

These results are encouraging and suggest that with further refinement and understanding, cross-correlations between foreground cleaned intensity maps and photometric imaging surveys should be a useful probe of cosmology. We stress that our suggested corrections need further testing, preferably alongside real data to ensure they are reliable.

6 CLUSTERING-BASED REDSHIFT ESTIMATION

As a direct example of the potential impact that foreground removal can have on cross-correlations with photometric redshift surveys, we now aim to use our simulations to see if a photometric calibration method using such cross-correlations is still viable. This method utilises the shared clustering signal between photometric optical galaxies and overlapping HI intensity maps. This clustering-based redshift estimation process has previously been studied in Alonso et al. (2017) and Cunningham et al. (2019), but a full analysis including simulated foreground contamination has not yet been conducted.

Given the difficulties outlined in Section 5, such a method represents a stern test since the intensity maps are correlated with a population of optical galaxies where little redshift information is assumed. The only assumption made is that the optical galaxies are within the redshift range covered by the reference intensity maps. This is applicable to weak-lensing probes where wide redshift bins may be used, and where the aim is to obtain the source distribution which is required for precise measurements of cosmological shear.

This wide redshift binning would mean huge degradation in small-scale radial modes, which is a major obstacle for this method given the increased noise due to the redshift uncertainty as outlined in Figure 12.

6.1 HI Clustering- z Method

In order to make a prediction for the redshift distribution of optical galaxies we require an estimator which utilises the shared clustering signal between the opticals and the HI intensity maps. We use the following estimator and refer the reader to Cunningham et al. (2019) where a full derivation is given:

$$\frac{dN_g}{dz}(z) = \frac{w_{g,H_I}(z)}{w_{H_I,H_I}(z)} \bar{T}_{H_I}(z) \frac{b_{H_I}(z)}{b_g(z)} \frac{1}{\Delta z}. \quad (35)$$

Here we use angular correlations functions w where $w_{g,H_I}(z)$ is the cross-correlation between all the optical galaxies and an HI intensity map at a redshift z . Similarly, $w_{H_I,H_I}(z)$ is the auto-correlation between two intensity maps at redshift z . An effective test of this estimator given the contamination of foregrounds is to use information from the C_ℓ power spectra since this is a measurement of angular clustering which is what we want to utilise for estimating dN_g/dz . An effective measurement for the angular correlation functions, which closely follows previous clustering redshift work (Menard et al. 2013) is given by

$$w_{XY}(z) = \int_{\ell_{\min}}^{\ell_{\max}} W(\ell) C_\ell^{XY}(z) d\ell, \quad (36)$$

where $W(\ell)$ is a weight function which can be tuned to certain scales. For our purposes $W(\ell) = \ell$ is sufficient to give weight to smaller scales where more useful matching is expected to exist. As previously, the indexes X and Y can either be chosen to represent the HI intensity map auto-correlation where $X = Y = \text{HI}$ or the cross-correlation with the optical where $X = g$ and $Y = \text{HI}$.

As before, \bar{T}_{H_I} is the average brightness temperature which is known in our simulations. In reality however, the observable is a temperature fluctuation and \bar{T}_{H_I} requires modelling as explained previously in equation (7). For our purposes here we assume an accurate modelling of \bar{T}_{H_I} has been achieved, i.e. we simply measure the quantity in our simulations.

Finally, the estimator in Equation (35) also requires the bias ratio b_{H_I}/b_g . We can find this from the angular auto-correlation power spectra for the two samples:

$$\frac{b_{H_I}(z)}{b_g(z)} = \frac{1}{\bar{T}_{H_I}(z)} \sqrt{\frac{C_{H_I H_I}(\ell, z)}{C_{gg}(\ell, z)}}. \quad (37)$$

However this relies on binning the galaxies by true redshift to measure the bias at that redshift. But we choose to assume that the optical sample has very poorly known (effectively unconstrained) redshifts, since it will be these surveys where redshift calibration is most in demand, so obtaining $C_{gg}(z)$ accurately is not possible. For this study we therefore rely on fiducial models of the individual biases. We take

$$b_g(z) = 1 + 0.84z, \quad (38)$$

which was estimated from simulation results in Weinberg et al. (2004) and used in the LSST Science book (Abell et al.

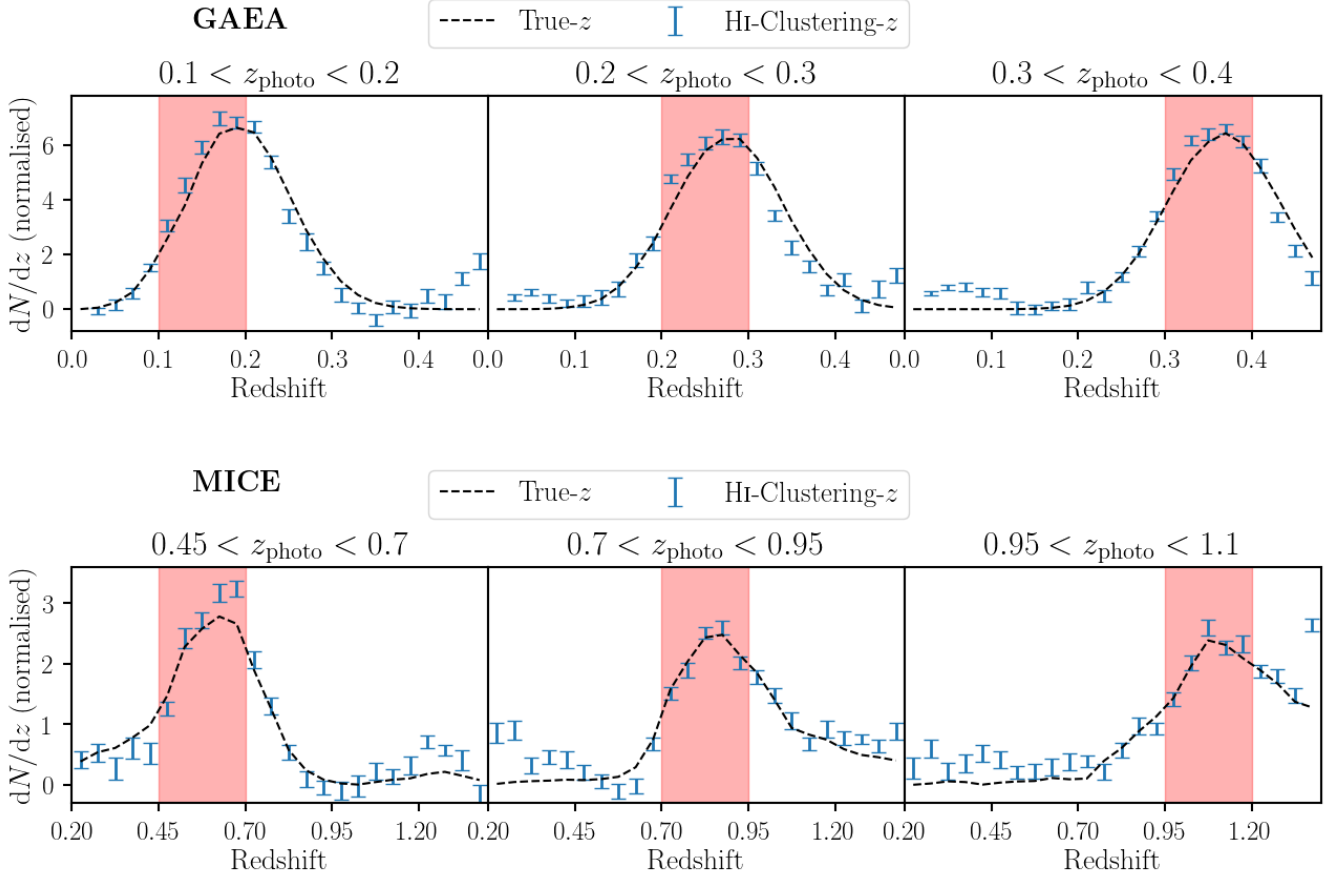


Figure 19. Clustering-based redshift estimation results using both the GAEA and MICE catalogues. The pink shaded regions represent the optical sample chosen as galaxies whose photometric redshift lies within the targeted redshift ranges. The black dashed lines show the true redshift distributions of these galaxies. The blue data points give the estimated redshift distributions based on cross-correlations with HI intensity maps and using the estimator in Equation (35). Intensity maps have foregrounds added and then removed with the FASTICA process with corrections made (Section 5.2). The GAEA model uses a beam size of $\theta_{\text{beam}} = 1.46^\circ$, representative of an SKA-like beam for that redshift range. However, for MICE we have used a smaller beam size of $\theta_{\text{beam}} = 1^\circ$ because of its smaller sky coverage.

2009). Following SKA Cosmology SWG. et al. (2018) we model the HI bias as

$$b_{\text{HI}}(z) = 0.67 + 0.18z + 0.05z^2. \quad (39)$$

6.2 HI Clustering- z Results

We are now ready to present a simple test of our HI clustering-based redshift estimation method and demonstrate its capability of recovering a redshift distribution using the HI intensity maps discussed in Section 2.1 for our simulated optical photometric sample with a detection threshold applied as discussed in Section 2.2. This is all in the presence of 21cm foreground contamination which has been cleaned using a FASTICA process. We also apply our corrections as outlined in Section 5.2.

We test this approach on both our GAEA and MICE based simulations and Figure 19 shows the results. In both cases we select optical galaxies based on their photometric redshifts in targeted redshift ranges shown as the pink shaded regions on the plots. Because these galaxies have been selected using their poorly constrained photometric

redshifts, the true redshift distribution (black dashed line) extends way beyond these ranges. By cross-correlating with HI intensity maps and using the estimator outlined by Equation (35) we can make a prediction of this true redshift distribution, shown by the blue data points. We obtain the error bars for dN/dz using a jackknifing technique, gridding the maps into an array of 25 smaller sub-samples. We then measure our estimator on the map but omit one of the 25 sub-samples. We repeat the procedure, averaging over the estimators obtained from omitting sub-samples, and obtain a standard deviation.

These results are very encouraging for the future of using shared clustering signals from HI intensity maps to calibrate photometric redshifts. Small discrepancies tend to exist at the extreme ends of the redshift distribution. When the true redshift distribution at these edges should be close to zero, often the estimator predicts a non-zero quantity. These are due to residual edge effects not fully mitigated by the correction outlined in Section 5.2.2. Because of this, it is difficult to place quantitative interpretation on the results in Figure 19 without these edge discrepancies skewing the measurement. We calculate the Median Absolute Deviation

(MAD) for the differences between the true and estimated distributions i.e. $dN/dz_{\text{true}} - dN/dz_{\text{est}}$, since this measurement will not be too sensitive to the incorrect estimations near the edges. We find that for the three GAEA distributions shown in Figure 19 these MAD values are 0.426, 0.304 and 0.386 for $0.1 < z_{\text{photo}} < 0.2$, $0.2 < z_{\text{photo}} < 0.3$ and $0.3 < z_{\text{photo}} < 0.4$ respectively. For MICE, the MAD values for the differences in true and estimated distributions are 0.149, 0.226 and 0.260. The similar values in each simulation demonstrate that the redshift prediction method is behaving consistently. The relatively low MAD values, under 10% of the normalised dN/dz peak value, also suggest the discrepancies between true and estimated distributions are mostly small and is indicative of the estimator's precision. This represents an excellent test of cross-correlations between foreground affected H I intensity maps and photometric surveys. This is because this method relies on sufficient cross-signal existing for poorly constrained optical redshifts over wide redshift ranges. The relative success of this method suggests that the problems outlined in Section 5.1 will be surmountable.

A key factor regarding the success of the clustering-based redshift estimation method using H I intensity maps is the size of the instrumental beam. Cunningham et al. (2019) found that the error on the estimation is directly proportional to the beam size and can be approximated by

$$\sigma_{N(z)} \propto \frac{\theta_{\text{beam}}}{\sqrt{A}}, \quad (40)$$

where A is the area of the sky covered. Due to the smaller sky coverage in the MICE simulation we found that we were unable to use a constant beam size of $\theta_{\text{beam}} = 2.36^\circ$ which would be representative of an SKA-like beam probing redshifts up to $z = 1.4$. Instead we have only smoothed with a 1° beam. However, having larger sky coverage in future simulations would mitigate this issue.

7 SUMMARY & CONCLUSIONS

Forthcoming H I intensity mapping experiments will be able to contribute to cosmological studies through H I auto-correlations as well as cross-correlations with optical galaxy surveys. To ensure that H I intensity mapping is a competitive technique, it is important to understand 21cm foreground contamination, and the effects of foreground removal on the measurements.

In this work we have taken a simulations-based approach to investigate these issues, focusing on the foreground removal effects on H I intensity mapping cross-correlations with photometric galaxy surveys. By using existing N -body simulations and the galaxy catalogues produced from them, we constructed both optical galaxy catalogue data and H I intensity map data with the same underlying cosmological clustering signal. We then simulated the relevant 21cm foreground signals that are expected to contaminate the H I intensity maps, and used a state-of-the-art blind foreground removal process known as FASTICA. This approach allowed us to then examine what impact the FASTICA foreground removal process has on cosmological probes such as the clustering measured by the angular power spectrum C_ℓ .

Our main conclusions are as follows:

- We have shown evidence that a FASTICA reconstruction will successfully allow accurate auto-correlation measurements. Figure 10 showcases the results for both our simulations, GAEA and MICE. The better result obtained for the GAEA model is likely due to its larger sky size allowing for more samples to average over in negentropy calculations.
- The auto-correlation tests we performed strongly suggest that a **frequency dependent beam size will cause problems** for independent component-like methods (demonstrated in Figure 10). A solution to this is to re-smooth the intensity maps to match the beam size for the highest redshift when using these foreground removal techniques.
- FASTICA also delivers good results in cross-correlation with optical galaxy data where the redshifts for the opticals are very well constrained as they would be in a spectroscopic-like survey. In Figure 11 we used optical galaxies with true redshifts in cross-correlation with H I intensity maps. The figure shows the excellent agreement between using the original (no foregrounds included) intensity maps and the foreground cleaned ones.
- We find that further treatment is needed when cross-correlating foreground cleaned H I intensity maps with photometric-like optical galaxy surveys with poor redshift constraints. Figures 12 and 13 show the impact of combining foreground cleaned intensity maps with an imaging galaxy survey which has poorly constrained redshifts. This poor result is unsurprising and can be generally explained by the combination of eroded large-radial modes caused by the foreground cleaning, with eroded small radial-modes caused by the uncertainty in the photometric redshifts Witzemann et al. (2018).
- More specifically, we find that a cause of the poor results when considering $\text{H I} \times \text{Photo-}z$ is the loss of LoS mean information when conducting the foreground clean. Figure 14 shows how any prior off-zero LoS means are collapsed to zero which has the effect of unpredictably changing pixel values in the transverse maps, as demonstrated in Figure 15. As a possible treatment for this unwanted effect we proposed a LoS reconstruction that uses information from the optical galaxies as outlined by Equation (30). This, coupled with artificially extending the redshift range to mitigate the edge effects caused by the foreground clean, improves results as shown by Figures 17 and 18.
- Finally, we conducted a comprehensive test of these methods by attempting to use foreground contaminated intensity maps for clustering-based redshift estimation of a photometric optical sample. By using FASTICA and our additional corrections we were able to accurately predict the redshift distributions for mock optical catalogues in both our models (Figure 19).

This work used two independent N -body simulations, where one (GAEA) used a semi-analytical approach to con-

structuring a galaxy catalogue and the other (MICE) used a HOD/HAM hybrid method. The resulting catalogues formed the basis for constructing the optical and HI intensity map mock data. This means we can be confident that the conclusions we have made are unlikely to be specific to these simulations.

A limitation in using existing mock galaxy catalogues to generate HI intensity maps however comes from the finite number of galaxies available to sample in the map. The great advantage of HI intensity mapping is the frequency resolution which allows for numerous tomographic bins. While the catalogues we use are large ($>10^8$ galaxies), this finite number means care was needed when going to large numbers of tomographic bins. If the bin is too thin, it will contain a low number of galaxies (sparse galaxy density), and therefore a sparse signal in each pixel. This is not an accurate emulation of an intensity map which should provide a near continuous emission profile. Tests were carried out with a higher number of bins in some cases. For example we used 240 redshift bins for the MICE catalogue and tested if we still see the LoS mean destruction demonstrated by Figure 14. Indeed we find that even with this more realistic number of bins, we find similar results but cannot be certain that these are accurate simulations of combined emission maps since the number density of simulated galaxies becomes low (~ 5 per voxel) at this fine radial resolution. This is why we used relatively thick tomographic bins in this work ($\Delta z = 0.02$ for GAIA and $\Delta z = 0.05$ for MICE). Furthermore, it is likely that intensity maps would need to be integrated to this size of bin when using cross-correlations with photometric surveys to measure power spectra or measure cosmological parameters. This is because the redshift uncertainty on the optical would demand a thick tomographic bin to ensure enough signal-to-noise.

Throughout this work we have made assumptions that parameters such as the mean HI temperature (\bar{T}_{HI}) can be precisely obtained. While we use a model for this parameter in our analysis, this same model was used in the construction of the HI intensity map signal, therefore its success is unsurprising. However, other parameters such as the clustering bias terms (b_g and b_{HI}) are not directly fed into our simulated signals, so the success of modelling these as scale-independent biases in our clustering-based redshift estimation is encouraging.

Note that in this work we have not simulated any foreground polarisation leakage effects. However, in many frequency channels we have smoothed our maps more than is required to simulate the instrument beam, which is a treatment previously used in real data to mitigate these effects (Switzer et al. 2013). However, for a follow-up study with simulations of realistic observations including polarisation leakage will be an important step.

Furthermore, in this work we did not consider the clustering of point source foregrounds, which one could argue has potential to bias cosmological clustering measurements. Nor in our simulations did we simulate the anisotropy of galactic free-free emission which is expected to be stronger in the galactic plane. However, neither of these subtle features are likely to affect the frequency coherence of the signals which FASTICA uses to isolate them.

In future work we plan to include a further analysis into the effects of foreground removal on cosmological

measurements including the 3D correlation function $\xi(s)$ and power spectrum $P(k)$ multipoles, extending the work of Blake (2019).

As HI intensity mapping data becomes available alongside the plethora of high precision optical datasets, we will be able to confirm conclusions derived from simulated mocks using real observations. Future measurements of HI \times Photo- z data, for example from MeerKAT and DES (Pourtsidou 2018; SKA Cosmology SWG. et al. 2018) or TIANLAI (Chen 2012) and DECaLS⁸ (Blum et al. 2016), will be an excellent test for our claims in this paper. We have demonstrated the potential of such experiments with our example of how cross-correlations can be used for photometric redshift calibration. This is a major challenge for forthcoming Stage-IV instruments utilising photometric optical samples, such as LSST and *Euclid*. We believe that photometric redshift calibration using HI intensity mapping data is an alternative method with great promise for tackling this challenge.

To summarise, we have shown evidence that a method such as FASTICA performs excellently at reconstructing the inherently weak HI signal in the presence of dominant 21cm foreground contamination. Even in cross-correlation with optical data with poorly constrained redshifts, with our suggested corrections it is possible to make good measurements of the cosmological signal. We have introduced a LoS mean reconstruction as a treatment for foreground cleaned intensity mapping signal loss, which improves the fidelity of cross-correlation measurements but which will benefit from further investigation and refinement. Foreground contamination is a challenge for HI intensity mapping, but this work alongside others demonstrates that it is a surmountable one. We look forward to providing even more realistic simulations, and testing our proposed methods with real data, in the near future.

ACKNOWLEDGEMENTS

We are grateful to Chris Blake, Emma Chapman, Pablo Fosalba, Ian Harrison, Mario Santos and Anna Zoldan for many useful discussions and feedback. SC is supported by the University of Portsmouth. AP is supported by STFC Grant ST/S000437/1. DB is supported by STFC Grant ST/N000668/1.

This work has made use of CosmoHub. CosmoHub has been developed by the Port d'Informacio Cientifica (PIC), maintained through a collaboration of the Institut de Fisica d'Altes Energies (IFAE) and the Centro de Investigaciones Energeticas, Medioambientales y Tecnologicas (CIEMAT), and was partially funded by the "Plan Estatal de Investigacion Cientifica y Tecnica y de Innovacion" program of the Spanish government (Carretero et al. 2017).

Numerical computations for this research were done on the Sciama High Performance Compute (HPC) cluster which is supported by the ICG, SEPNet, and the University of Portsmouth.

⁸ <http://legacysurvey.org/>

REFERENCES

- Abell P. A., et al., 2009
- Alonso D., Ferreira P. G., Santos M. G., 2014, *Mon. Not. Roy. Astron. Soc.*, 444, 3183
- Alonso D., Bull P., Ferreira P. G., Santos M. G., 2015, *Mon. Not. Roy. Astron. Soc.*, 447, 400
- Alonso D., Ferreira P. G., Jarvis M. J., Moodley K., 2017, *Phys. Rev.*, D96, 043515
- Amendola L., et al., 2018, *Living Rev. Rel.*, 21, 2
- Anderson C. J., et al., 2017, [10.1093/mnras/sty346](https://arxiv.org/abs/10.1093/mnras/sty346)
- Battye R. A., Browne I. W. A., Dickinson C., Heron G., Maffei B., Pourtsidou A., 2013, *Mon. Not. Roy. Astron. Soc.*, 434, 1239
- Blake C., 2019
- Blum R. D., et al., 2016, in *American Astronomical Society Meeting Abstracts #228*. p. 317.01
- Bolzonella M., Miralles J.-M., Pello' R., 2000, *Astron. Astrophys.*, 363, 476
- Carretero J., Castander F. J., Gaztanaga E., Crocce M., Fosalba P., 2015, *MNRAS*, 447, 646
- Carretero J., et al., 2017, *PoS, EPS-HEP2017*, 488
- Chapman E., et al., 2012, *Mon. Not. Roy. Astron. Soc.*, 423, 2518
- Chapman E., et al., 2013, *Mon. Not. Roy. Astron. Soc.*, 429, 165
- Chen X., 2012, *International Journal of Modern Physics: Conference Series*, 12, 256
- Crocce M., Castander F. J., Gaztanaga E., Fosalba P., Carretero J., 2015, *Mon. Not. Roy. Astron. Soc.*, 453, 1513
- Cunnington S., Harrison I., Pourtsidou A., Bacon D., 2019, *Mon. Not. Roy. Astron. Soc.*, 482, 3341
- Fernandez-Soto A., Lanzetta K. M., Chen H.-W., Pascarella S. M., Yahata N., 2001, *Astrophys. J. Suppl.*, 135, 41
- Fosalba P., Gaztañaga E., Castander F. J., Crocce M., 2015a, *Mon. Not. Roy. Astron. Soc.*, 447, 1319
- Fosalba P., Crocce M., Gaztañaga E., Castander F. J., 2015b, *Mon. Not. Roy. Astron. Soc.*, 448, 2987
- Gorski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelman M., 2005, *Astrophys. J.*, 622, 759
- Harper S., Dickinson C., Battye R., Roychowdhury S., Browne I., Ma Y.-Z., Olivari L., Chen T., 2018, *Mon. Not. Roy. Astron. Soc.*, 478, 2416
- Harrison I., Camera S., Zuntz J., Brown M. L., 2016, *Mon. Not. Roy. Astron. Soc.*, 463, 3674
- Haslam C. G. T., Salter C. J., Stoffel H., Wilson W. E., 1982, *A&AS*, 47, 1
- Hirschmann M., De Lucia G., Fontanot F., 2016, *MNRAS*, 461, 1760
- Hoffmann K., Bel J., Gaztanaga E., Crocce M., Fosalba P., Castander F. J., 2015, *MNRAS*, 447, 1724
- Hyvärinen A., 1999, *IEEE transactions on neural networks*, 10, 3, 626
- Jelic V., et al., 2008, *Mon. Not. Roy. Astron. Soc.*, 389, 1319
- Jelic V., Zaroubi S., Labropoulos P., Bernardi G., de Bruyn A. G., Koopmans L. V. E., 2010, *Mon. Not. Roy. Astron. Soc.*, 409, 1647
- Marshall P., et al., 2017, [10.5281/zenodo.842713](https://arxiv.org/abs/10.5281/zenodo.842713)
- Masui K. W., et al., 2013, *Astrophys. J.*, 763, L20
- Menard B., Scranton R., Schmidt S., Morrison C., Jeong D., Budavari T., Rahman M., 2013
- Moore D. F., Aguirre J. E., Parsons A. R., Jacobs D. C., Pober J. C., 2013, *Astrophys. J.*, 769, 154
- Padmanabhan H., Kulkarni G., 2017, *Mon. Not. Roy. Astron. Soc.*, 470, 340
- Pen U.-L., Staveley-Smith L., Peterson J., Chang T.-C., 2009, *Mon. Not. Roy. Astron. Soc.*, 394, 6
- Percival W. J., et al., 2001, *Mon. Not. Roy. Astron. Soc.*, 327, 1297
- Pourtsidou A., 2018, *PoS, MeerKAT2016*, 037
- Pourtsidou A., Bacon D., Crittenden R., 2017, *Mon. Not. Roy. Astron. Soc.*, 470, 4251
- Remazeilles M., Delabrouille J., Cardoso J.-F., 2011, *Mon. Not. Roy. Astron. Soc.*, 418, 467
- SKA Cosmology SWG. Bacon D. J., et al., 2018, Submitted to: *Publ. Astron. Soc. Austral.*
- Santos M. G., Cooray A., Knox L., 2005, *Astrophys. J.*, 625, 575
- Santos M. G., et al., 2015
- Shaw J. R., Sigurdson K., Pen U.-L., Stebbins A., Sitwell M., 2014, *Astrophys. J.*, 781, 57
- Slosar A., Hirata C., Seljak U., Ho S., Padmanabhan N., 2008, *JCAP*, 0808, 031
- Springel V., et al., 2005, *Nature*, 435, 629
- Switzer E. R., et al., 2013, *Mon. Not. Roy. Astron. Soc.*, 434, L46
- Switzer E. R., Chang T.-C., Masui K. W., Pen U.-L., Voytek T. C., 2015, *Astrophys. J.*, 815, 51
- Weinberg D. H., Dave R., Katz N., Hernquist L., 2004, *Astrophys. J.*, 601, 1
- Witzemann A., Alonso D., Fonseca J., Santos M. G., 2018
- Wolz L., Abdalla F. B., Blake C., Shaw J. R., Chapman E., Rawlings S., 2014, *Mon. Not. Roy. Astron. Soc.*, 441, 3271
- Wolz L., et al., 2015, *PoS, AASKA14*, 035
- Wolz L., et al., 2017, *Mon. Not. Roy. Astron. Soc.*, 464, 4938
- Wolz L., Murray S. G., Blake C., Wyithe J. S., 2018
- Xie L., De Lucia G., Hirschmann M., Fontanot F., Zoldan A., 2017, *MNRAS*, 469, 968
- Zheng H., et al., 2017, *Mon. Not. Roy. Astron. Soc.*, 464, 3486
- Zoldan A., De Lucia G., Xie L., Fontanot F., Hirschmann M., 2017, *MNRAS*, 465, 2236

This paper has been typeset from a \LaTeX file prepared by the author.