

A first quantification of the effects of absorption for HI intensity mapping experiments

Sambit Roychowdhury (সম্বিত রায়চৌধুরী)^{1,2,3}, Clive Dickinson¹, and Ian W. A. Browne¹

¹ Jodrell Bank Centre for Astrophysics, Alan Turing Building, Department of Physics & Astronomy, School of Natural Sciences, The University of Manchester, Oxford Road, Manchester M13 9PL, UK

² Institut d'Astrophysique Spatiale, CNRS, Université Paris-Sud, Université Paris-Saclay, Bât. 121, 91405 Orsay Cedex, France

³ Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Hawthorn, VIC 3122, Australia
e-mail: sroychowdhury@swin.edu.au

Received 20 June 2019 / Accepted 17 September 2019

ABSTRACT

Context. HI intensity mapping (IM) will be used to do precision cosmology, using many existing and upcoming radio observatories. It will measure the integrated HI 21 cm emission signal from “voxels” of the sky at different redshifts. The signal will be contaminated due to absorption, the largest component of which will be the flux absorbed by the HI emitting sources themselves from the potentially bright flux incident on them from background radio continuum sources.

Aims. We, for the first time, provide a quantitative estimate of the magnitude of the absorbed flux compared to the emitted HI flux. The ratio of the two fluxes was calculated for various voxels placed at redshifts between 0.1 and 2.5.

Methods. We used a cosmological sky simulation of the atomic HI emission line, and summed over the emitted and absorbed fluxes for all sources within voxels at different redshifts. In order to determine the absorbed flux, for each HI source the flux incident from background radio continuum sources was estimated by determining the numbers, sizes, and redshift distribution of radio continuum sources that lie behind it, based on existing observations and simulations. The amount of this incident flux that is absorbed by each HI source was calculated using a relation between integrated optical depth with HI column density determined using observations of damped Lyman- α systems (DLAs) and sub-DLAs.

Results. We find that for the same co-moving volume of sky, the HI emission decreases quickly with increasing redshift, while the absorption varies much less with redshift and follows the redshift distribution of faint sources that dominate the number counts of radio continuum sources. This results in the fraction of absorption compared to emission to be negligible in the nearby Universe (up to a redshift of ~ 0.5), increases to about 10% at a redshift of one, and continues to increase to about 30% up to a redshift of 2.5. These numbers can vary significantly due to the uncertainty on the exact form of the following relations: firstly, the number counts of radio continuum sources at sub-mJy flux densities; secondly, the relation between integrated optical depth and HI column density of HI sources; and thirdly, the redshift distribution of radio continuum sources up to the highest redshifts.

Conclusions. Absorption of the flux incident from background radio continuum sources might become an important contaminant to HI IM signals beyond redshifts of 0.5. The impact of absorption needs to be quantified more accurately using inputs from upcoming deep surveys of radio continuum sources, HI absorption, and HI emission with the Square Kilometre Array and its precursors.

Key words. cosmology: observations – large-scale structure of Universe – galaxies: ISM – radio lines: ISM – galaxies: statistics – radio continuum: galaxies

1. Introduction

HI intensity mapping (IM) refers to a novel method of doing precision cosmology using the integrated HI 21 cm emission signal from volumes or “voxels” of the sky at different redshifts, starting immediately after reionization right up to the present day (Battye et al. 2004; Peterson et al. 2006; Loeb & Wyithe 2008; Shaw et al. 2014; Santos et al. 2014; Bull et al. 2015; Wolz et al. 2017; Kovetz et al. 2017; Villaescusa-Navarro et al. 2018). The cosmic HI signal is weak and needs to be integrated over large volumes of sky for it to be detectable as a fluctuation, over and above, the statistical and instrumental noise. There are also much brighter “foregrounds” that need to be removed for the HI signal to become detectable – radio synchrotron and free-free emissions from the Galaxy itself, radio continuum emission from extragalactic sources, instrumental effects like $1/f$ noise, etc. In spite of these challenges HI IM promises to be one of the most important tools to study the evolution of large scale struc-

ture and consequently cosmology. There are multiple ongoing efforts to use existing (Pen et al. 2009; Chang et al. 2010; Masui et al. 2010, 2013; Santos et al. 2014; Anderson et al. 2018) and upcoming radio observatories (Chang et al. 2008; Chen 2012; Battye et al. 2013, 2016; Bandura et al. 2014; Bigot-Sazy et al. 2016; Newburgh et al. 2016; Smoot & Debono 2017) to carry out HI IM. Cosmology using HI IM is one of the main science objectives of the Square Kilometre Array (SKA, Santos et al. 2015; Villaescusa-Navarro et al. 2017; Square Kilometre Array Cosmology Science Working Group et al. 2018), and there is a proposal to perform an HI IM survey with the SKA precursor, MeerKAT (Santos et al. 2017).

The success of HI IM will depend on precisely measuring integrated signals from the sky. We therefore need to understand and estimate all possible contributions that can contribute to the integrated signal, their comparative strengths, and work out methods to extract the HI emission signal from the observed integrated signal. There exist estimates of the measured flux

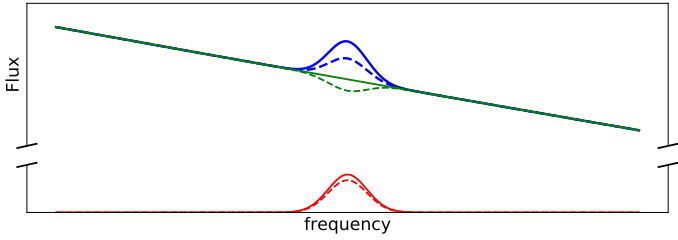


Fig. 1. Schematic representation of how absorption will affect the signal that an HI IM experiment aims to detect. The total HI emission from sources within a given volume of the sky is represented by the bold red line, where the frequency spread corresponds to the redshift spread of the HI emitting sources within the volume. The dashed red line represents the HI emission after correcting for any potential absorption within the emitting sources themselves and within sources in the foreground. The bold green line represents the broadband spectrum of the radio continuum emission from sources in the foreground as well as the background of the HI emitting sources. We note that the flux level of this continuum emission is likely to be much higher than the HI emission one aims to detect. The dashed green line represents the radio continuum emission corrected for the absorption of a fraction of the emission incident from background radio continuum sources by the HI emitting sources themselves. The bold blue line represents the total of the HI emission and radio continuum emission (sum of the bold red and green lines), where the HI emission can be thought to be a positive fluctuation on top of a spectrally smooth broadband radio continuum “foreground”. The dashed blue line represents how the actual combined HI and radio continuum emission signal will look like, after correcting for the negative fluctuation due to the various absorptions (sum of the dashed red and green lines).

that will be detected by HI IM experiments for observations centred at various redshifts with varying configurations (e.g. Bigot-Sazy et al. 2015; Harper et al. 2018). These estimates take into account various “foregrounds” as well as instrumental effects, and show that sky and instrumental foreground signals are expected to dominate the integrated signal. The foreground signals correlate in contrast to the HI signal (Ansari et al. 2012; Liu & Tegmark 2012), and various methods have been developed to exploit this fact and “subtract” the foreground signals (e.g. Liu & Tegmark 2011; Masui et al. 2013; Switzer et al. 2013, 2015; Shaw et al. 2014; Wolz et al. 2014; Alonso et al. 2015; Olivari et al. 2016), and thus extract the HI emission signal.

IM experiments aim to detect the combined HI 21 cm emission from individual HI rich sources within a volume of sky. This emission from the (unresolved) HI sources can be visualized as a fluctuation on top of a base flux level having a smooth broadband spectrum due to the various “foregrounds”, as shown in the schematic Fig. 1. The astronomical “foregrounds” in HI IM will consist of radio continuum emission from our Galaxy, and radio continuum emission from extragalactic sources in front of the HI emitting sources we are interested in (i.e. $z < z_{\text{IM}}$ where z_{IM} is the redshift of an HI emitting source), as well as extragalactic sources in the background (i.e. $z > z_{\text{IM}}$). A part of the HI flux being emitted will be absorbed within the HI emitting sources themselves, and also by sources in the foreground including the Galaxy. More importantly, the HI in the sources of interest will absorb a part of the emission incident on them from background radio continuum sources. Therefore, along with the positive fluctuation due to HI emission there will be some negative fluctuation due to the combined effect of these various absorptions. This negative fluctuation will also enter the combined signal being picked up by the IM experiment, similar to what is shown in Fig. 1. To date there has been no attempt to quantify the amount

of absorption, the assumption being that the absorption would be low compared to the emission.

In order to understand the problem at hand, the basic equation to consider is the brightness temperature T_B of an isothermal cloud of atomic hydrogen in front of a background source of brightness temperature T_C ,

$$T_B(\nu) = T_S(1 - e^{-\tau(\nu)}) + T_C(\nu)e^{-\tau(\nu)}, \quad (1)$$

where $\tau(\nu)$ is the optical depth and T_S is the spin temperature of the hydrogen in the cloud. In practice, the brightest radio continuum sources in the field would be flagged based on radio continuum source catalogues. For the fainter uncatalogued radio continuum sources, their base flux level would be determined and removed during “foreground subtraction” using the smooth spectral slope of radio continuum sources. Thus at the redshifted frequency of HI 21 cm emission from a given HI source being observed by the IM experiment, the base flux level calculated for any background source is what it would have been without any absorption by the HI source itself. Therefore the flux absorbed by the HI source itself becomes part of the HI “signal”. Therefore effectively the brightness temperature that will be measured is,

$$T_B(\nu) = T_S(1 - e^{-\tau(\nu)}) - T_C(\nu)(1 - e^{-\tau(\nu)}). \quad (2)$$

Analytical estimates of the brightness temperature T_B of the HI emission that would be observable by IM experiments (e.g. as in Battye et al. 2013) ignore the term related to T_C . The spin temperature for hyperfine 21 cm emission line of HI in astrophysical situations is set by collisional excitation of the hydrogen atoms. Thus working under the assumption that radiative excitation of hydrogen atoms is negligible, the physics of radiative transfer for stimulated emission can be used to find an expression for $T_S(N_{\text{HI}})$, where N_{HI} is the column density of atomic hydrogen in the cloud (Field 1958; Dickey et al. 1978). Analytical estimates of the brightness temperature also assume that the HI emitting gas is optically thin. Equation (2) therefore becomes,

$$T_B(\nu) = T_S(N_{\text{HI}})\tau(\nu). \quad (3)$$

Estimates based on the latest hydrodynamical simulations (e.g. Villaescusa-Navarro et al. 2018) start with modelling the amount and distribution of HI in detail using state-of-the-art prescriptions. But when converting the column densities of HI thus derived into a brightness temperature, they also work under the assumption that the HI emitting gas is optically thin, and do not take into account the terms related to T_C in Eq. (2). The validity of the assumption about the HI being optically thin and the potential correction for HI self-absorption are discussed in Appendix A.1.1.

Analytical estimates show that the rms fluctuation corresponding to the emission term $T_S(1 - e^{-\tau(\nu)})/T_S\tau(\nu)$, within the power spectrum that will be measured by an HI IM experiment will be ~ 0.1 mK (Battye et al. 2013). Correspondingly, the rms fluctuation due to radio continuum sources (with sources > 10 mJy flagged), or T_C , will be ~ 10 mK. Therefore it is important to confirm whether in Eq. (2) the factor due to the optical depth offered by the HI sources $(1 - e^{-\tau(\nu)})$ will be small enough so that the second term on the right hand side is negligible compared to the first. We will focus exclusively on this type of absorption in this paper. Other absorptions are possible but expected to have negligible effect. See Appendix A for a more complete discussion.

It has been proposed that the large number of HI absorption systems to be detected by upcoming HI IM surveys can potentially be used to statistically measure the Sandage-Loeb effect

by observing the systems over a long period of time, and thus directly measure the acceleration of the universe (Yu et al. 2014; Jiao et al. 2019). In this paper though we focus on how significant the flux absorbed from background radio continuum sources by the HI emitting sources is, when compared to the emitted HI flux that HI IM experiments ultimately aim to detect. The paper is organized as follows. In Sect. 2 we detail the method and the underlying assumptions that we use to calculate the amount of such absorption. In Sect. 3 we present our estimates for the absorption flux compared to the emission flux for all sources integrated over a voxel in the sky. We discuss the significance of our results, and how the various assumptions might affect our results. Finally in Sect. 4 we summarize our results and what needs to be done in the future to fully understand the effects of absorption on HI IM experiments.

2. Method

From this point in the paper we focus on what is likely to be the most significant type of absorption – that of incident radiation from background continuum sources intercepted in the HI clouds (sources) responsible for the HI emission we are interested in detecting. For a given IM experiment observing a given voxel in the sky, our aim was to calculate the ratio of the total flux absorbed from the incident background continuum, to the HI 21 cm flux emitted by the HI sources within the voxel. This ratio will provide a quantitative measure of the significance of absorption in HI IM experiments, as in its denominator is the cosmological signal one is actually interested in measuring. We used the cosmological sky simulation of the atomic HI emission line by Obreschkow et al. (2009a,b), hereafter the O9 simulation/catalogue, for our estimates. The simulation is built by post-processing a semi-analytic virtual galaxy catalogue (De Lucia & Blaizot 2007) built on the Millenium Simulation of cosmic structure. In the O9 simulations the cold gas is split into HI and H₂, spatial distribution and velocity profiles are assigned separately to the atomic and molecular gas components, and the resulting galaxies compared against the observed mass functions and various scaling relations. This is important as we want to focus exclusively on the HI properties of the sources. These simulations are ideal for large-scale cosmological predictions where individual galaxies are represented by a handful of global properties. As we detail below, a simple representation of the HI disks in galaxies is what we used for our present estimates, as any further details become redundant given the large uncertainties that exist for other observables and properties that need to be factored into our estimates. Various parameters for HI sources including positions, gas contents, morphologies, velocity information, etc. are available up to a redshift of $z_{\text{max}} = 10$. For reasons detailed in Sect. 2.5, we restricted our calculation to below redshift $z = 2.5$. This is also a practical limit given that $z \sim 2.5$ is the upper limit of even the most ambitious targeted HI IM experiments, and approaching the upper limit ($z \sim 3$) of planned HI IM surveys with the SKA (Bull et al. 2015).

For each HI source at a given redshift, we estimated the number of radio continuum sources in different flux density ranges that will lie behind it. We also estimated the covering factor (fraction of the background source that is covered by the foreground source) for any radio continuum source given its size on the sky based on its flux density. The fraction of the incident flux from all these continuum sources that is absorbed by the HI source is then calculated based on the optical depth of the HI in the source. Finally, the total fluxes emitted and absorbed by all HI sources within a given voxel of sky is summed up.

The various intermediate steps required to ultimately calculate the total emitted and absorbed fluxes from the HI sources within a voxel of the O9 simulated sky are listed in the subsections below. In many of these steps we had to make some assumptions and simplifications in order to calculate the relevant quantities. We discuss the validity and shortcomings of each assumption either in this section or later in Sect. 3.

2.1. HI source parameters

We started with a couple of assumptions about the HI sources themselves. Firstly, we assumed that the HI in the source is distributed uniformly over an elliptical disk. This is a valid approximation even when nearby HI rich galaxies are observed, but at low resolution. For HI IM experiments given the scales of interest, the angular resolution would not be good enough even to resolve the disk of any HI source being observed, keeping in mind we are interested in sources at cosmological distances and not in the local universe.

Secondly, we assumed that the HI is distributed uniformly over some velocity width. This is a workable assumption given the double-horned profile of the HI emission used in the O9 simulations. We should be aware though that unlike spatial resolution, HI IM experiments would be observing with high spectral resolution. This might change the results of our calculation slightly for say a narrow channel width centred on the peak of the HI emission. In the end though, in order to have enough signal-to-noise for constraining cosmology the bandwidths would be much larger than the observing channel width (e.g. see Battye et al. 2013) – thus this assumption too would be valid.

2.2. Number of background sources per unit flux density

We needed to know the number and distribution in terms of spectral flux density, of radio continuum sources that would be covered by the HI source of interest. In order to do so, we started with the observed number count of radio sources at 1.4 GHz, the assumption being that the distribution is the same at the frequency corresponding to that of HI 21 cm emission at the redshift of the HI source. Hales et al. (2014a) provide an empirical fit to the number count of radio sources as observed in the second data release of the Australia Telescope Large Area Survey (ATLAS) at 1.4 GHz done with the Australia telescope Compact Array (ATCA), using the parametric form for the fit to the Phoenix Deep Survey and Faint Images of the Radio Sky at Twenty cm (FIRST) survey data by Hopkins et al. (2003) while accounting for the observed excess in the number of radio sources below 300 μJy :

$$\log \left(S^{2.5} \frac{dN}{dS} \right) = \begin{cases} \sum_{j=0}^6 a_j [\log(S)]^j & \text{if } S \geq 300 \mu\text{Jy} \\ \sum_{j=0}^6 a_j [\log(0.3)]^j & \text{if } 30 \mu\text{Jy} \leq S < 300 \mu\text{Jy} \\ \sum_{j=0}^6 a_j [\log(10S)]^j & \text{if } S < 30 \mu\text{Jy} \end{cases} \quad (4)$$

where the flux density S is in mJy, and $a_0 = 0.859$, $a_1 = 0.508$, $a_2 = 0.376$, $a_3 = -0.049$, $a_4 = -0.121$, $a_5 = 0.057$, $a_6 = -0.008$.

The distribution of sources given by Eq. (4) is shown in Fig. 2. Given the angular size of the HI source, using Eq. (4) we can determine the number distribution of radio continuum sources as a function of spectral flux density that would overlap with the HI source on the sky. The NRAO VLA Sky Survey (NVSS) and Sydney University Molonglo Sky Survey (SUMSS) are radio continuum surveys at 1.4 GHz ($\sim 45''$ resolution), which in between them cover the entire sky and are complete to < 10 mJy (Condon et al. 1998; Mauch et al. 2003). Data

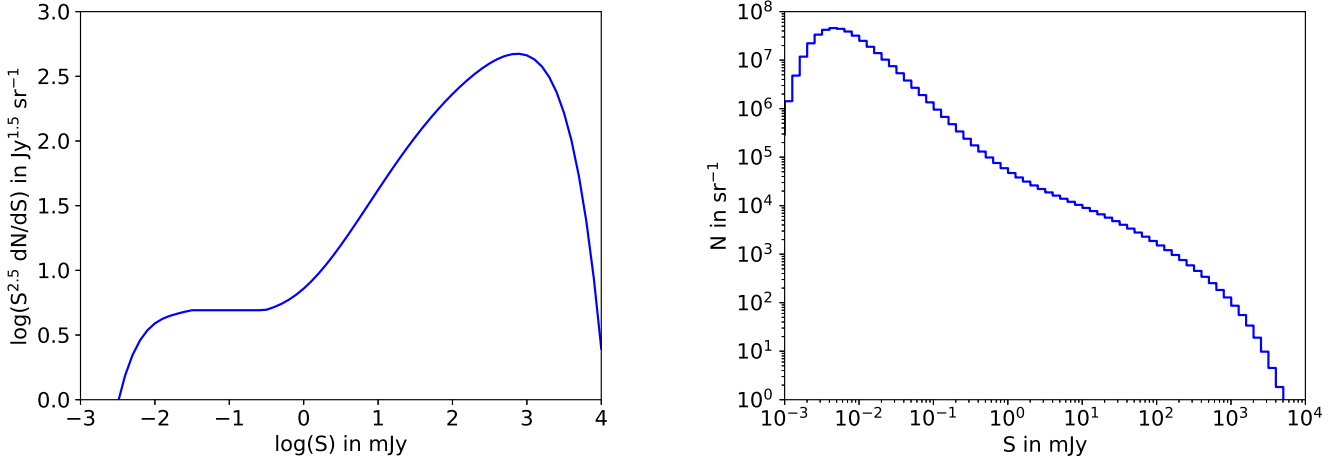


Fig. 2. Number count of radio continuum sources at 1.4 GHz (Hales et al. 2014a): Euclidean normalized differential (*left panel*) and total (*right panel*) as functions of flux density.

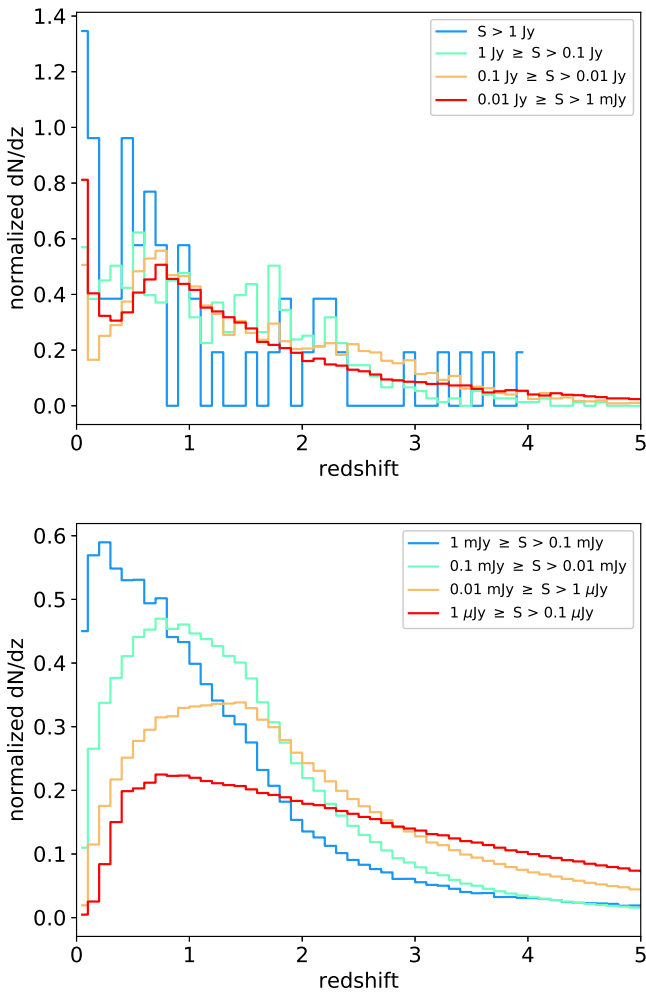


Fig. 3. Simulated redshift distribution of radio continuum sources in various flux density bins, from Wilman et al. (2008). *Upper panel:* fluxes above 1 mJy and has a y-axis range which is more than twice as large as the *lower panel*.

from these surveys would be used to identify the bright radio continuum sources in future HI IM experiments, which will then be masked out from the images (this would be done before “foreground subtraction” of radio continuum sources). Therefore we

set a conservative upper limit of 100 mJy for radio continuum sources which will contribute to the absorption that we aimed to quantify. In reality the HI flux absorbed will be slightly less than what we calculate here as radio continuum sources with lower flux densities will also be masked out given the completeness limits of NVSS and SUMSS. But the reduction in absorption will be small, as sources with flux densities between 10 mJy and 100 mJy do not contribute significantly to the total absorption (see Fig. 6).

2.3. Distribution of background sources with redshift

Of the radio continuum sources, only those that are at a redshift higher than the HI source redshift will contribute to the HI absorption. Therefore in order to determine the number of radio continuum sources in the background of the voxel of sky being observed, we require the redshift distribution of radio continuum sources. We used the redshift distribution of radio continuum sources from the simulated radio continuum sky of Wilman et al. (2008). We needed to use results from simulations because even the deepest observational studies of the redshift distribution of radio continuum sources are complete only to 10 mJy in flux density (see e.g. Brookes et al. 2008). In Sect. 3 we discuss how our final results are affected if the redshift distribution of sources varies as compared to what we assumed here.

Wilman et al. (2008) provides a catalogue of ~ 320 million simulated radio sources down to a flux density limit of 10 nJy at five different frequencies. We used the catalogued values for 1.4 GHz. The simulation uses a semi-empirical approach designed with SKA goals of galaxy evolution and large scale structure in mind. Sources are drawn from observed (extrapolated when needed) luminosity functions placed on top of an underlying dark matter density field with biases which reflect the measured large-scale clustering of radio sources. The radio sources are of five distinct types: radio loud and radio quiet active galactic nuclei of both type I and type II Fanaroff-Riley structural classes, as well as star-forming galaxies which in turn are split into quiescent and starbursting galaxies. We determined the redshift distributions of the radio sources from the catalogue after binning them in flux density. The normalized versions of these redshift distributions are shown in Fig. 3. It is of interest to note that for radio sources with flux density below 1 mJy, the distribution starts to have a more pronounced tail at high redshifts. We used these normalized distributions to calculate the fraction,

$f_{z,\text{HI}}$, of the total number of radio continuum sources that would lie behind our HI source of interest, based on its redshift.

2.4. Covering fraction for background sources

We also needed to calculate what fraction of a background source would be incident on the HI source, given that some of the background sources may be larger in size than the HI source. We used the relation between median angular sizes of radio sources at 1.4 GHz and their flux densities found by Windhorst et al. (1990) for the flux density range between 0.1 mJy and 10 Jy:

$$d_{\text{med}} = 2'' S^{0.3}, \quad (5)$$

to define a flux density S_{eq} for which the background source's angular size (Ω_b in sr) matches that of the HI source (Ω_{HI} in sr). More recent deep radio continuum surveys at 1.4 GHz find results consistent with the above relation (e.g. Bondi et al. 2003; Prandoni et al. 2018). Thus the incident flux density on the HI source from a background source with flux density S is $f_c S$, where

$$f_c = \begin{cases} 1 & \text{if } S < S_{\text{eq}} \\ \frac{\Omega_{\text{HI}}}{\Omega_b} & \text{if } S \geq S_{\text{eq}}. \end{cases} \quad (6)$$

We can expect radio source sizes to vary with redshift, and thus the distribution of source sizes is coupled to the redshift distribution of sources (Fig. 3). Windhorst et al. (1990) considered all sources from multiple surveys irrespective of redshift, and thus any such coupling has in principle been incorporated in Eq. (5).

We note though that the way we take covering fraction into account is an over-simplification, born out of necessity as there is very little information to fall back upon. The limited information on the variation of radio continuum source sizes compared to HI source sizes, especially beyond the very local Universe where the HI distribution can be mapped in emission, also restricts us from quantitatively determining the error on our final results due to our assumed covering fractions. This information, like many others discussed in Sect. 3, needs to come from future deep surveys of radio continuum and HI in emission by the SKA.

2.5. Optical depth for a certain HI column density

The amount of flux absorbed is proportional to the optical depth offered by the HI column density of the absorber. Specifically, the quantity of interest is the “integrated” optical depth over the full width of the absorption. This relation between integrated optical depth and column density is at the heart of our calculations, and therefore we needed to base it on observations as opposed to expectations based on theoretical calculations or simulations. Empirically this relation can be quantified in HI sources up to intermediate redshifts by using observations of Damped Lyman- α Systems (DLAs) which contain most of the neutral hydrogen in the Universe (Wolfe et al. 2005). Ultraviolet/optical spectroscopic observations through which DLAs are identified directly provide a measure of their HI column densities, and follow-up HI 21 cm absorption observations are needed to measure their integrated optical depths.

We compiled an up-to-date list of DLAs and sub-DLAs (systems with HI column densities just below the DLA limit of 2×10^{20} atoms cm^{-2}) for which the HI column density has been measured, and the integrated optical depth has either been measured or has a robust upper limit. Of the ones with measured integrated optical depths, one is at a redshift of 3.39, and the

rest span a redshift range of 0.09–2.35. Kanekar et al. (2014) find statistically significant evidence for an increase in spin temperatures beyond $z \approx 2.4$ by combining observed upper limits and detections of HI absorption in of DLAs and sub-DLAs. An increased spin temperature would imply lower integrated optical depths for the same HI column density. Considering this fact combined with the paucity of measurements beyond $z > 2.35$, we restricted our estimates to a redshift of $z = 2.5$ in this paper.

All the DLAs and sub-DLAs from our compilation with redshift $z < 2.5$ are listed in Table 1. Column 1 lists the quasar in whose sightline the Lyman- α absorption is observed, Col. 2 gives the redshift at which the absorption is detected, Col. 3 lists whether the integrated optical depth (τ_{dv}) was measured or only an upper limit was obtained using follow-up HI absorption measurements. Column 4 lists whether the reference from which the measured values have been taken is Kanekar et al. (2014) – otherwise they were provided by Nissim Kanekar (priv. comm.).

In Fig. 4 the measured integrated optical depths and upper limits of the DLAs and sub-DLAs mentioned above are plotted against their column densities. From the figure we can see that for similar HI column densities, on average DLAs have lower integrated optical depths compared to the values in our own Galaxy. We can also see that there is no obvious trend with redshift for the relationship between integrated optical depth and HI column density. We fitted the following relation using the Akritas-Theil-Sen estimator for censored data (Akritas et al. 1995), with the Turnbull estimate of intercept, using the CENKEN function in R programming language:

$$\log \left(\frac{\int \tau_{\text{dv}}}{\text{km s}^{-1}} \right) = 0.9 \log \left(\frac{N_{\text{HI}}}{\text{atoms cm}^{-2}} \right) - 19.1, \quad (7)$$

plotted as the bold black line in Fig. 4. In order to estimate the uncertainty on the above relation, we bootstrapped the data to create multiple datasets by varying each measurement within the 1σ error on either side of the respective measurement (upper limits remain unchanged). The fits to the bootstrapped censored datasets is shown as the grey shaded region in Fig. 4. We used the above relation to arrive at our main results, and the uncertainty on the fitted relation to estimate one of the major uncertainties in our result.

2.6. Total flux emitted and absorbed by a HI source

Finally, we needed to combine the various steps listed above to estimate the total flux absorbed within a given voxel, and compare it to the total flux emitted from the same voxel. We used the catalogued properties for HI sources from the O9 catalogue to calculate the emitted flux F_{emit} . We also used the assumption that the emission is spread uniformly across the velocity width. We note that column definitions from the O9 catalogue are quoted hereon when describing parameters used in our calculations. The “velocity-integrated line flux of the HI line” (in Jy km s^{-1}) of a particular HI source is divided by the source's HI velocity width as defined by the “line width at 20% of peak luminosity density (already corrected for the galaxy inclination)” (in km s^{-1}). This value is then multiplied by Δv to obtain F_{emit} . Δv (in km s^{-1}) is equal to the observation channel width Δv_{cw} , when $\Delta v_{\text{cw}} < \Delta v_{\text{HI}}$, and Δv is equal to Δv_{HI} when $\Delta v_{\text{cw}} \geq \Delta v_{\text{HI}}$.

To calculate the flux absorbed by each such source in the O9 catalogue, we started with the catalogued value for “HI radius where Σ_{HI} is at 10% of its maximum value” (in arcseconds) to be the semi-major axis of the elliptical disk. We used the catalogued value for the “minor axis/major axis for HI” to calculate the size

Table 1. Sample of DLAs and sub-DLAs used to determine the integrated optical depth – column density relation.

QSO	z_{abs}	τ_{dv}	Reference
0738+313	0.0912	Measured	1
0738+313	0.2212	Measured	1
0952+179	0.2378	Measured	1
1127–145	0.3127	Measured	1
1229–021	0.3949	Measured	1
0235+164	0.5241	Measured	1
0827+243	0.5247	Measured	1
J1431+3952	0.6039	Measured	1
1331+305	0.6921	Measured	1
2355–106	1.1727	Measured	1
J1623+0718	1.3367	Measured	1
2003–025	1.4106	Measured	1
1331+170	1.7763	Measured	1
1157+014	1.9436	Measured	1
1755+578	1.9698	Measured	2
1850+402	1.9888	Measured	2
0458–020	2.0394	Measured	1
2039+187	2.1920	Measured	1
0311+430	2.2890	Measured	1
0438–436	2.3474	Measured	1
1122–168	0.6819	Limit	1
0454+039	0.8596	Limit	1
2149+212	0.9115	Limit	1
J0407–4409	1.913	Limit	2
1230–101	1.931	Limit	2
0347–211	1.9470	Limit	1
1452+502	1.969	Limit	2
J0733+2721	1.9758	Limit	2
1215+333	1.999	Limit	2
0620+389	2.031	Limit	1
J1412+1257	2.0632	Limit	2
J0845+4257	2.0652	Limit	2
J1634+3203	2.0923	Limit	2
J0214+0632	2.1075	Limit	2
J1419+0603	2.1080	Limit	2
1645+635	2.1253	Limit	2
0149+335	2.141	Limit	1
0528–250	2.141	Limit	2
J1301+1904	2.1482	Limit	2
J1522+2119	2.1709	Limit	2
1228–113	2.193	Limit	2
J1224+1947	2.2104	Limit	2
J0934+3050	2.2143	Limit	2
J1529+1904	2.220	Limit	2
1048+347	2.2410	Limit	1
J1138+0428	2.2427	Limit	2
0432–440	2.297	Limit	2
J1138+0428	2.3292	Limit	2
J0912+4126	2.3867	Limit	2
J0934+3050	2.3883	Limit	2
0201+365	2.462	Limit	2
J0214+0157	2.4886	Limit	2
J1406+3433	2.4989	Limit	1

References. 1 – Kanekar et al. (2014), 2 – priv. comm.

of the semi-minor axis of the elliptical disk. The major and minor axis values in combination with the angular diameter distance

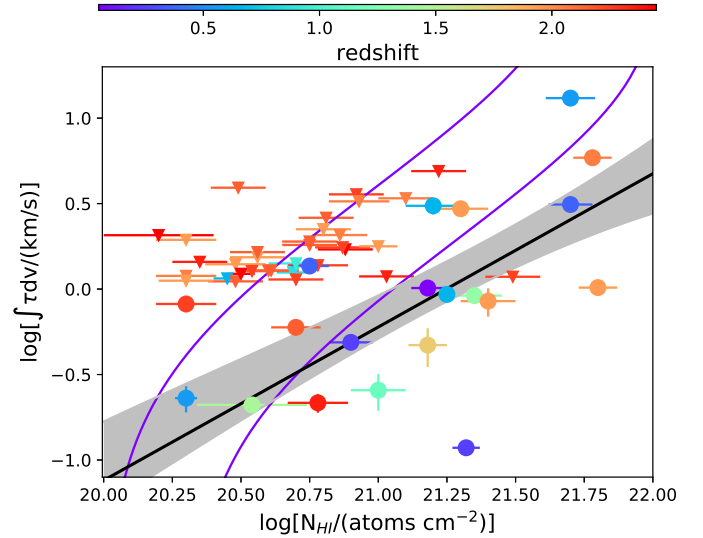


Fig. 4. Variation of the measured integrated optical depth with the measured HI column density in DLAs and sub-DLAs up to a redshift of $z = 2.5$. The points are colour-coded according to their respective redshifts following the colourbar at top. Filled circles represent data with measured optical depths, whereas downward pointing triangles represent data with upper limits on optical depth. The black bold line represents the linear fit to the censored data, with the grey shaded region representing the uncertainty on the fit determined using bootstrapping datasets (see Sect. 2.5 for details). The violet lines encompass the region occupied by observations of various sightlines within the Galaxy ($z = 0$) with both HI absorption and emission measurements (Kanekar et al. 2011).

determined from the catalogued values of “comoving distance to the object” (in Mpc) and “apparent redshift (including Doppler correction)”, is used to calculate the area of the disk. The average HI column density over the disk for a given source is calculated by dividing the “HI mass” (in M_{\odot}) of the source by the area of its elliptical disk in pc^2 . This is based on the fact that given the low spatial resolution of upcoming HI IM experiments, they will only be sensitive to the average column density over the disk of the HI source.

Thereafter we used the formalism developed in the previous sections. Given the average column density N_{HI} for the source, we calculated the $\int \tau(v)dv$ using Eq. (7). Next, we drew upon our assumption from Sect. 2.1 that the HI is distributed uniformly over some velocity width Δv_{HI} , and divided $\int \tau(v)dv$ by Δv_{HI} to determine an effective optical depth per unit velocity τ_e . We used the catalogued value of “line width at 20% of peak luminosity density (already corrected for the galaxy inclination)” (in km s^{-1}) as the HI velocity width Δv_{HI} .

If the total flux density incident on the HI source is S_0 , the total reduction in flux due to HI absorption produced by the HI source in consideration is $\int S_0(1 - e^{-\tau_e})dv = S_0(1 - e^{-\tau_e})\Delta v$, where Δv is defined in the same way as for the case of emission. In order to determine S_0 we summed over all the background sources whose flux would be absorbed by the HI source under consideration. We therefore obtained the total flux absorbed by the HI source from the incident flux of radio continuum sources behind it by

$$F_{\text{abs}} = f_{z,\text{HI}} \left[\int_{S_{\text{min}}}^{S_{\text{max}}} f_c \frac{dN}{dS} \Omega_{\text{HI}} dS \right] (1 - e^{-\tau_e}) \Delta v, \quad (8)$$

where $\frac{dN}{dS}$, $f_{z,\text{HI}}$ and f_c have been defined in Sects. 2.2–2.4 above. The observed number of radio sources at 1.4 GHz falls sharply

Table 2. Size of O9 simulation box width at various redshifts.

Redshift	Simulation box side (degrees)	No. of sources for 1 MHz wide channel
0.1	116.5	2.9×10^5
0.5	21.4	3.0×10^5
1.5	9.0	4.8×10^5
2.5	6.7	5.4×10^5

below $1 \mu\text{Jy}$ in the model we use (there are no actual observations of radio continuum sources with flux densities below $1 \mu\text{Jy}$), and therefore we chose $S_{\text{min}} = 10^{-7} \text{ Jy}$.

We repeated the above exercise for all sources within a voxel to arrive at the quantity of interest: the ratio of the total absorbed flux to the total emitted flux for all sources within the voxel, $\Sigma(F_{\text{abs}}/F_{\text{emit}})$.

2.7. Summing over sources within a voxel

The ratio of the total absorbed flux to the total emitted flux for all HI sources within a given voxel is what we wanted to calculate. In the O9 simulated sky, the full simulated sky area at any redshift is a square of side $\sim 500 h^{-1} \text{ Mpc}$. Over the large redshift range we are interested in for this work, the comoving size of the simulated sky area of the O9 simulations being constant results in a large variation in the angular size of the simulation boxes, as can be seen in Table 2. At each redshift we focused on voxels defined by equally wide channel widths in frequency, between 0.1 and 10 MHz, which cover the likely range in frequency widths to be used by upcoming HI IM surveys. Given that the same width in frequency units translate to increasingly large velocity widths with increasing redshift, we find from Table 2 that the total number of sources actually increases (within a factor of 2) with redshift. Within the channel width range of 0.1 and 10 MHz, the total number of sources vary between a few times 10^4 to a few times 10^6 . Thus for all redshifts and all channel widths we integrated over statistically large number of sources thereby determining true ‘‘average’’ values for the calculated quantities.

For each redshift and channel width combination, those sources from the O9 catalogue are considered which have some part of their HI emission lying within the channel width. Such sources are found by considering the catalogued values of ‘‘apparent redshift (including Doppler correction)’’ in combination with their HI velocity widths. The physical dimensions of each of the HI sources are calculated from their sizes in arcseconds using angular diameter distances for each source, which in turn are calculated using their catalogued values for ‘‘comoving distance’’ and ‘‘apparent redshift’’.

3. Results and discussion

3.1. Variation of absorbed-to-emitted flux with redshift

Here we present the results of our estimates for all HI sources at a particular redshift within a voxel defined by the O9 simulation box size of $\sim 500 h^{-1} \text{ Mpc}$ and a representative constant channel width of 1 MHz. We find the HI emission and absorption fluxes increase with increasing channel width as the number of sources within a voxel increases, for channel widths between 0.1 and 10 MHz. What we also find as a consequence of always sampling a large number of sources for our chosen range of

channel widths, is that the HI emission and absorption fluxes show the exact same variation with varying channel width. As we are interested in the relative variation between HI emission and absorption, hereafter we report the results only for the representative channel width.

In Fig. 5 we plot for all HI sources within each voxel, the variation with redshift of the total emitted flux, the total absorbed flux, and the ratio of the total absorbed flux to total emitted flux. The total HI emission flux in Fig. 5 shows a marked decline with increasing redshift. The decrease is mainly driven by the increasing luminosity distance with redshift. The actual variation though of the total emitted flux with redshift is driven by a complex combination of factors like source numbers varying with redshift (Table 2), and the nature of the sources themselves changing with redshift following the O9 simulation. The variation of total HI absorption flux with redshift as shown in Fig. 5 is even more complex given all the factors discussed in Sect. 2. What is worth noting is that the total HI absorption flux actually increases between $z = 0$ to $z = 1.1$, and has a shallow peak followed by a slow decline. Therefore while the total emitted flux reduces by more than two orders of magnitude between $z = 0$ and $z = 2.5$, the total absorbed flux varies very little over the same range of redshifts.

As can be seen from the right panel of Fig. 5, the differing variations of the total HI emission and absorption fluxes with redshift results in the ratio of total $F_{\text{abs}}/F_{\text{emit}}$ rising steeply from 0.027% at $z = 0.1$, 1.4% at $z = 0.5$ to 6.5% at $z = 1.0$ and continuing to increase to 21% by $z = 2.5$. It should be noted that if instead we define voxels in terms of unit solid angle/survey beam size combined with a constant channel width, the variation of the emission and absorption fluxes will be different to that shown in the left panel of Fig. 5, but the ratio of the two will vary in the same way as shown in the right panel Fig. 5. Our results are indicating that HI 21 cm absorption of flux incident from background radio continuum sources can be a major source of uncertainty in HI IM experiments, particularly those probing higher redshifts. These results should be taken to be indicative only, as there are large uncertainties on our final results and how much of an effect absorption will have on HI IM experiments is very difficult to quantify precisely. The uncertainties are a result of our limited knowledge of both the HI sources and the radio continuum sources at intermediate to high redshifts, which forces us to use various assumptions and relations based on predictions from simulations. Below we try to quantify the uncertainties that are present in our final results presented in Fig. 5 due to the intrinsic uncertainties of the various assumptions and relations we have used.

As stated in Sect. 2.1, HI IM experiments will be done at low angular resolutions given the scales of cosmological interest which are much larger than the sizes of the absorbers – the HI sources. We note therefore that the power spectrum of the absorption will be approximately flat over the scales of interest for HI IM.

3.2. Background source flux density range which affects the results most

The results in this study are for a maximum radio continuum source flux density of 100 mJy on the assumption that stronger sources are cataloged and will be removed. In order to understand which flux range of background sources affects the results most, we repeated the study for the chosen channel width of 1 MHz and with varying ranges of radio continuum source flux density. The variation of total $F_{\text{abs}}/F_{\text{emit}}$ with redshift for varying

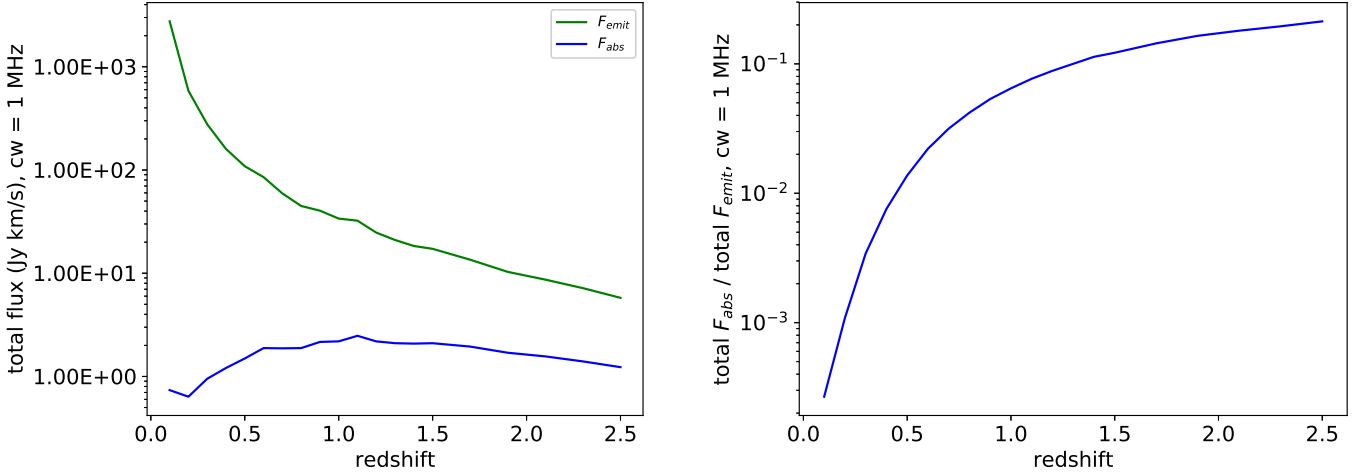


Fig. 5. Variation of total absorbed and emitted fluxes (*left panel*) and the ratio of the total absorbed-to-emitted flux (*right panel*) using the O9 simulated sky and the assumptions listed in Sect. 2. The results are for all sources within voxels defined by channel width of 1 MHz, placed at various redshifts. Note though that the ratio of the total absorbed-to-emitted flux remains unchanged for channel widths ranging between 0.1 and 10 MHz.

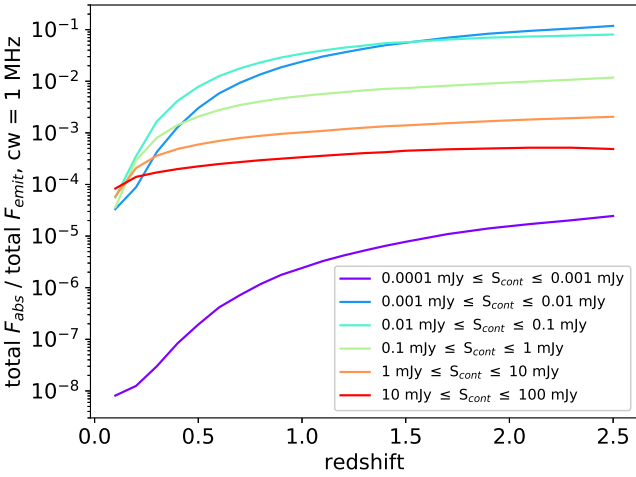


Fig. 6. Ratio of the total absorbed to total emitted fluxes for all sources within voxels defined by channel width of 1 MHz placed at various redshifts, for various ranges of background radio continuum source fluxes.

ranges of background radio continuum source flux densities is presented in Fig. 6. We find that the major contribution to the HI absorption is from the flux incident from background sources with flux densities in the range $1 \mu\text{Jy} \leq S_{\text{cont}} \leq 0.1 \text{ mJy}$, that is around the peak of the number distribution of sources when plotted against flux density as in the right panel of Fig. 2. We can now understand that the variation of the total absorbed flux with redshift in the left panel of Fig. 5 is driven by the cumulative number of radio continuum sources in the flux density range $1 \mu\text{Jy} \leq S_{\text{cont}} \leq 0.1 \text{ mJy}$ behind a voxel placed at a particular redshift (which can be calculated from the distributions shown in the lower panel of Fig. 2).

3.3. Effect of uncertainty in sub-mJy number counts of background sources

The number counts of radio continuum sources at sub-mJy flux densities remains an open area of research. As described in Sect. 2.2, we have used the relation from Hales et al. (2014a) which models the effect of an exaggerated population of faint

sources below 0.3 mJy at 1.4 GHz. Hales et al. (2014b) shows that Eq. (4) reproduces the average low ($<0.1 \text{ mJy}$) flux density source counts observed by deep surveys of the radio sky at 1.4 GHz (Biggs & Ivison 2006; Fomalont et al. 2006; Owen & Morrison 2008; Ibar et al. 2009; Morrison et al. 2010; Padovani et al. 2011). The field-to-field scatter in number counts from these and other determinations of the number counts at sub-mJy flux densities (e.g. Richards 2000; Bondi et al. 2008; Seymour et al. 2008) is large and cannot be simply explained away by sample variance (Heywood et al. 2013). It has been suggested though that the observed scatter is a result of data processing differences and calibration errors (e.g. see discussion in Ibar et al. 2009; Hales et al. 2014a). How the number counts might look if data processing and calibration errors are accounted for is beyond the scope of this work. In order to understand the effect this scatter can have on our results, we simply take into account all published number counts and consider the number counts from two studies which envelop the observed scatter in sub-mJy number counts, Owen & Morrison (2008) and Richards (2000).

In the left panel of Fig. 7 we plot the observed Euclidean normalized number counts from these two studies. We create two alternate distributions of radio source counts against flux density by deviating from the Hales et al. (2014a) fit at the low flux density end using fits to the data points from Owen & Morrison (2008) and Richards (2000), shown as red and green curves respectively in the left panel of Fig. 7. The right panel of Fig. 7 shows how the actual distribution of sources with flux density varies at the low flux density end for these fits, as compared to that given by the fits from Hales et al. (2014a). What is apparent from this panel is that even a small variation in the Euclidean normalized distribution at low flux densities has a very large impact on the total number of low flux density sources. This is reflected in how much $F_{\text{abs}}/F_{\text{emit}}$ varies depending on the distribution of radio source counts we choose, as shown in Fig. 8. When considering the radio source count distribution which is consistent with the data from Owen & Morrison (2008), total $F_{\text{abs}}/F_{\text{emit}}$ is 0.055% at $z = 0.1$, 2.4% at $z = 0.5$, 17% at $z = 1.0$, and 94% by $z = 2.5$. In contrast, when considering the radio source count distribution which is consistent with the data from Richards (2000), total $F_{\text{abs}}/F_{\text{emit}}$ is 0.018% at $z = 0.1$, 0.4% at $z = 0.5$, 1.2% at $z = 1.0$, and 2.8% by $z = 2.5$. Comparing

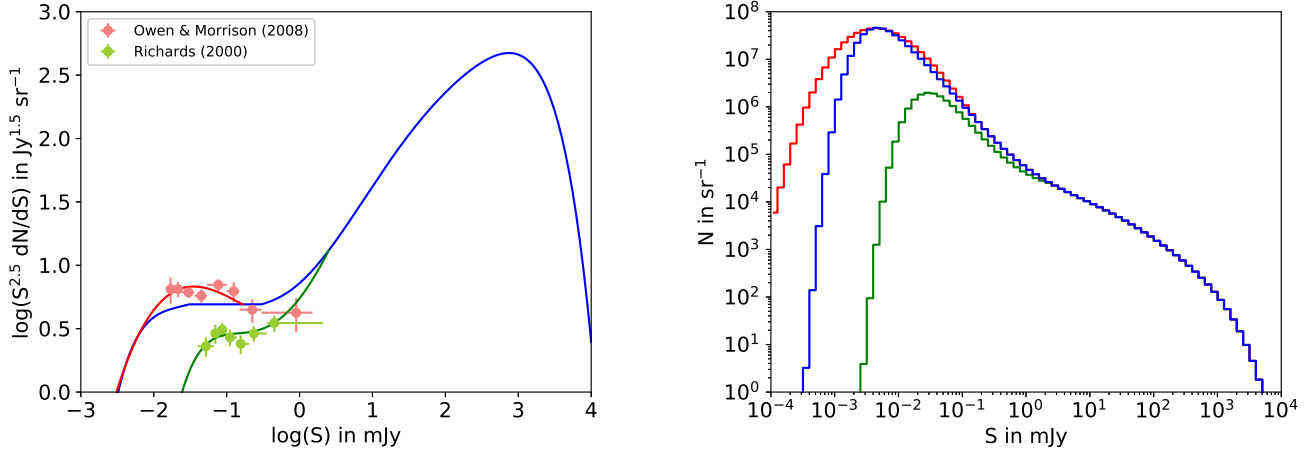


Fig. 7. Varying number counts of radio continuum sources at 1.4 GHz for sub-mJy flux densities. Euclidean normalized differential (*left panel*) and total (*right panel*) as functions of flux density. The red and green points with error bars in the *left panel* are measurements from [Owen & Morrison \(2008\)](#) and [Richards \(2000\)](#) respectively. The blue curves in both panels is the fit from [Hales et al. \(2014a\)](#) that we use to calculate our results, whereas the red and green curves are our fit to the red and green points in the *left panel*. Our fits are extensions of the [Hales et al. \(2014a\)](#) fit at the low flux ranges where the red and green points deviate from the blue curve (see text for details). We note that in the *right panel* the *x*-axis starts an order of magnitude lower as compared to the *left panel*, in order to accentuate the variation in the number count at the low flux end from different measurements of the same.

with our main result as shown in Fig. 5, we find that the scatter in the number counts of sub-mJy radio continuum sources among various measurements can vary our results by almost an order of magnitude in either direction. As the number counts for sub-mJy flux densities are uncertain, deeper radio continuum surveys are very much necessary to constrain the number counts and thus the effect of absorption on HI IM experiments. One important additional caveat is that since the low flux density sources are mostly star forming (dwarf) galaxies which are clustered, the absorption of the HI signal will be scale-dependent and can create a scale-dependent bias in the final results of an HI IM experiment.

3.4. Effect of uncertainty in optical depth vs. column density relation for HI sources

Another important relation defining the results presented in Fig. 5 is the empirical relation between integrated optical depth and HI column density discussed in Sect. 2.5. This relation is given in Eq. (7), which is based on a fit to observations with large intrinsic scatter as can be seen in Fig. 4. In Fig. 9 we show how our results change on considering the error on the fit determined by bootstrapping the data, the grey shaded region in Fig. 4. If we first consider the upper envelope of fits to bootstrapped data, the total $F_{\text{abs}}/F_{\text{emit}}$ goes from 0.042% at $z = 0.1$, 2.0% at $z = 0.5$ to 9.5% at $z = 1.0$ and continuing to increase to 31% by $z = 2.5$. On the other hand if we consider the lower envelope of fits to bootstrapped data, the total $F_{\text{abs}}/F_{\text{emit}}$ goes from 0.022% at $z = 0.1$, 1.1% at $z = 0.5$ to 5.2% at $z = 1.0$ and continuing to increase to 17% by $z = 2.5$. Therefore the uncertainty on the relation between integrated optical depth and HI column density can introduce a substantial uncertainty on the quantification of then flux absorbed in an HI IM experiment. From Fig. 4 we can see that the point-to-point scatter around the fit between the different individual measurements is even larger than the uncertainty on the fit. Therefore, given the sensitivity of the results in this work to the integrated optical depth vs. HI column density relation, we think that a detailed quantification of the relation up to intermediate redshifts is absolutely necessary for estimating the effect of absorption on HI IM experiments.

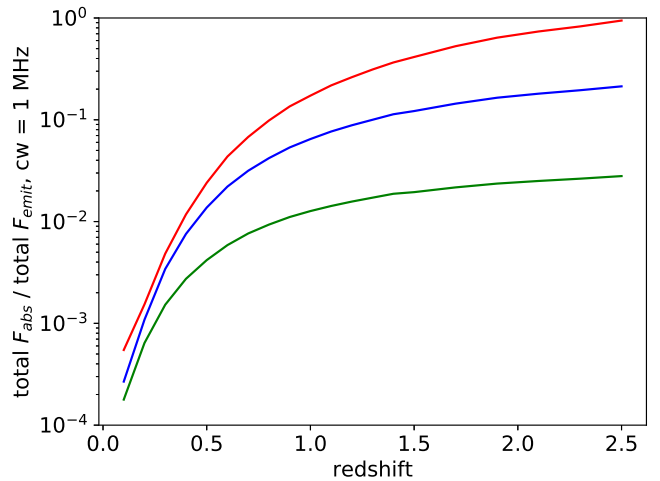


Fig. 8. Ratio of the total absorbed to total emitted fluxes for all sources within voxels defined by channel width of 1 MHz placed at various redshifts using the O9 simulated sky, for varying sub-mJy radio continuum source number counts. The solid blue line is identical to the result presented in the right panel of Fig. 5. The solid red and green lines show the results for the extreme version of the number counts as represented by [Owen & Morrison \(2008\)](#) counts and [Richards \(2000\)](#) counts respectively, which are shown in Fig. 7.

In our calculations we assumed that the optical depth of the HI source is distributed uniformly over the full HI velocity width of the source under consideration, in keeping with our primary assumptions about HI sources outlined in Sect. 2.1. This assumption is actually a conservative one when considering the strength of absorption, as typically the optical depth will have a non-uniform distribution with one or more peaks. At the velocity of the peak(s) the absorbed flux would be much higher given its exponential dependence on the optical depth compared to what it would be if the optical depth is distributed uniformly over the HI velocity width. Also at high spatial resolutions, HI absorption might start to dominate over HI emission in an IM experiment. In the future when the various relations affecting the quantification of the absorption strength are better constrained, we need to

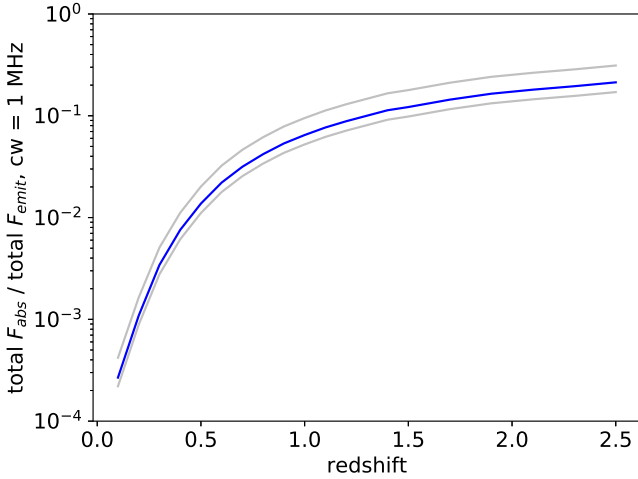


Fig. 9. Solid blue line: ratio of the total absorbed to total emitted fluxes for all sources within voxels defined by channel width of 1 MHz placed at various redshifts using the O9 simulated sky, identical to the result presented in the right panel of Fig. 5. The grey lines show the results due to the uncertainty of the relation between integrated optical depth and HI column density, i.e. the grey shaded region in Fig. 4. The scale of the plot is identical to Fig. 8.

move beyond simply comparing the strength of the absorption to the emission as is done here. We should aim to quantify what the power spectrum of HI absorption itself looks like at high spatial resolutions, possibly paving the way for HI absorption intensity mapping.

3.5. Effect of uncertainty in redshift distribution of background sources

In Sect. 2.3 we describe how we use input from simulations for the distribution of radio continuum sources with redshift. We check how our results change on using alternate distributions to those plotted in Fig. 3. We consider the two distributions shown in Fig. 10 and assume that all radio continuum sources of all fluxes follow these distributions instead of those shown in Fig. 3. Alternative distribution 1 is derived from the fit to the redshift distribution from de Zotti et al. (2010) of CENSORS sources with flux density >10 mJy (Brookes et al. 2008),

$$N(z) = 1.29 + 32.37z - 32.89z^2 + 11.13z^3 - 1.25z^4, \quad (9)$$

where z is the redshift. The normalized form of Eq. (9) is shown in cyan in Fig. 10. We find that this redshift distribution has some features which are in common to the simulated distribution shown in Fig. 3: a peak at low redshifts followed by a decline. Unlike the redshift distributions of sources with flux densities <100 mJy (which we use to calculate our results) shown in Fig. 3, alternative distribution 1 has no radio sources beyond a redshift of $z \sim 3.75$. To complement alternative 1, as alternative 2 we use an extremely hypothetical distribution where the sources are distributed uniformly up to a very high redshift – which we take to be $z = 10$. This provides a check of how an excess of radio continuum sources at high redshifts might affect our results.

Figure 11 shows how the main result of the study changes on considering these alternate distributions. We find that the redshift distribution of radio continuum sources, especially at high redshifts, can affect the amount of HI absorption significantly

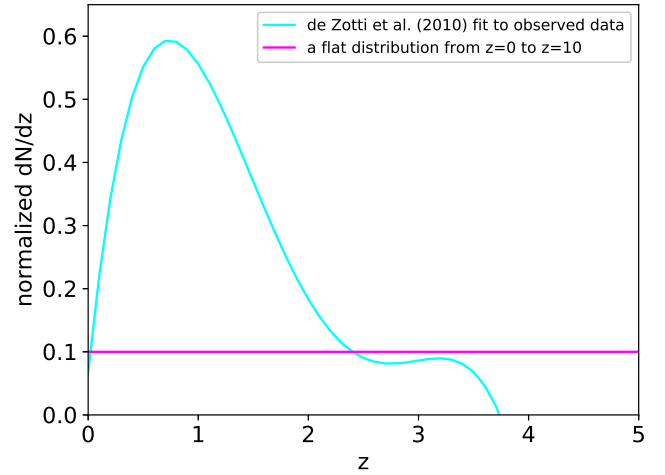


Fig. 10. Two alternate distributions of radio continuum sources with redshift (see text for details). Compare with the original redshift distributions shown in Fig. 3.

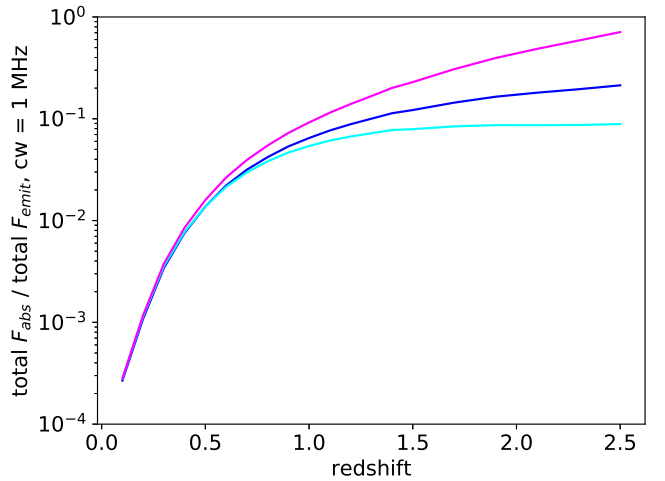


Fig. 11. Cyan and magenta lines: ratio of the total absorbed to total emitted fluxes for all sources within voxels defined by channel width of 1 MHz placed at various redshifts using the O9 simulated sky, for the two alternate redshift distributions of radio continuum sources shown in Fig. 10. While solid blue line: original result and is identical to the result presented in the right panel of Fig. 5. The scale of the plot is identical to Figs. 8 and 9.

at $z \gtrsim 0.5$. Any distribution which has a large fraction of the sources at higher redshift will expectedly make the absorption higher and increase the ratio of absorbed to emitted fluxes. As alternative distribution 1 has features similar to the distributions in Fig. 3 up to a peak around $z = 1$, the results on assuming alternative distribution 1 starts deviating appreciably from the results shown in Fig. 5 only for $z > 1$. Even when assuming a flat redshift distribution of sources as for alternative distribution 2, the results remain unaffected for $z < 0.5$, and only start to deviate appreciably from the main result around and beyond $z = 1$. Therefore it is important also to measure the redshift distribution of radio continuum sources at high redshifts, especially in the critical flux density range $0.001 \text{ mJy} \leq S_{\text{cont}} \leq 0.1 \text{ mJy}$ given the reasons mentioned above. Planned SKA surveys in combination with next generation infrared space observatories should be able to determine the redshift distribution of such sources (Padovani 2011).

3.6. Effects of other assumptions

In our study we have based the properties of all background radio continuum sources on their properties at 1.4 GHz. Thus we have inherently assumed that the spectral index of radio sources below 1.4 GHz is on average flat. There are observational indications though that the average radio spectral index below 1.4 GHz is -0.8 (de Gasperin et al. 2018) but appears to flatten below 10 mJy (e.g. Mahony et al. 2016). The spectral indices of the sub-mJy radio continuum sources below 1.4 GHz has not been constrained observationally. If they have steep spectral indices, our inherent assumption about a flat spectral index will turn out to be a conservative one. The radio continuum sources will have higher flux density at the redshifted frequency of observation and the absorbed flux will also be slightly higher. And this effect will be exacerbated with increasing redshift. Again, future radio continuum surveys at 1.4 GHz and lower frequencies will provide the answers regarding the spectral indices of sub-mJy radio sources.

Finally, the results from this study are inherently dependent on the simulated HI sky from the O9 simulations. Their simulations are based on empirical recipes and physical constraints, and have been shown to match the overall observed properties of cosmic structure. But in the end simulations are an extrapolation of existing data, and future surveys of HI sources up to intermediate-to-high redshifts will help improve the precision of both the simulations, and predictions based on simulations like in this work.

4. Conclusion

In this study we, for the first time, estimated the amount of HI 21 cm flux that would be absorbed by the same sources whose emission HI IM experiments aim to detect. We estimated the fraction of absorbed flux compared to the emitted HI flux in voxels defined by channel widths as would be used in upcoming HI IM experiments. We used the HI sky from the O9 simulations, and used multiple assumptions and relations based either on observations or input from simulations, to arrive at our result. We find (i) generally low absorption effects at $z < 0.5$, (ii) increasing absorption effects above $z = 0.5$ reaching $\sim 10\%$ of the emission at $z = 1$. (iii) Absorption effects continue to increase above $z = 1$ and can reach 30% of the emission signal at $z = 2.5$. The above numbers vary significantly, as large as an order of magnitude in either direction, depending on the uncertainties in the relations used to calculate the amount of absorption. Up to $z \sim 0.5$ the modelling uncertainties are not too large but above that they can become significant. The three main relations whose imprecision affect our results, in order of importance, are: the number counts of radio continuum sources especially at sub-mJy flux densities, the relation between integrated optical depth and HI column density of HI sources, and the redshift distribution of radio continuum sources up to the highest redshifts.

Therefore the results from multiple planned deep high-resolution surveys, specifically using the SKA, are required to properly quantify the effect of absorption on the HI IM signal. These include deep surveys of the radio continuum emission to determine the number counts, redshift distributions, and size distribution of radio sources. SKA absorption surveys which will detect HI absorption from multiple DLA sightlines can possibly improve the accuracy of the measured integrated optical depth vs. HI column density relation and its variation with redshift. SKA surveys of HI 21 cm emission sources up to intermediate redshifts will help fine-tune simulations like O9, and enable us

to make better predictions using such simulations. Therefore the interpretation of HI IM surveys will require knowledge of the effects of absorption, which could bias the results. The amount of absorption, especially at high redshifts, needs to be quantified using data from future deep surveys focusing on radio continuum, HI absorption, and HI emission from individual sources.

Acknowledgements. SR would like to thank Nissim Kanekar for detailed discussions and suggestions. We thank James Allison for helpful comments regarding material presented in Appendix B. We thank Stuart Harper for helpful comments and suggestions. SR and CD acknowledge support from ERC Starting Grant No. 307209. SR acknowledges that this research was partially supported by the Australian Government through the Australian Research Council funding (FT150100269). CD is also supported by an STFC Consolidated Grant (ST/P000649/1).

References

- Aditya, J. N. H. S. 2019, *MNRAS*, **482**, 5597
 Aditya, J. N. H. S., & Kanekar, N. 2018a, *MNRAS*, **473**, 59
 Aditya, J. N. H. S., & Kanekar, N. 2018b, *MNRAS*, **481**, 1578
 Akritas, M., Murphy, S., & LaValley, M. 1995, *J. Am. Stat. Assoc.*, **90**, 170
 Allison, J. R., Curran, S. J., Emonts, B. H. C., et al. 2012, *MNRAS*, **423**, 2601
 Allison, J. R., Sadler, E. M., & Meekin, A. M. 2014, *MNRAS*, **440**, 696
 Alonso, D., Bull, P., Ferreira, P. G., Maartens, R., & Santos, M. G. 2015, *ApJ*, **814**, 145
 Anderson, C. J., Luciw, N. J., Li, Y.-C., et al. 2018, *MNRAS*, **476**, 3382
 Ansari, R., Campagne, J. E., Colom, P., et al. 2012, *A&A*, **540**, A129
 Bandura, K., Addison, G. E., Amiri, M., et al. 2014, in *Ground-based and Airborne Telescopes V*, Proc. SPIE, 9145, 914522
 Battye, R. A., Davies, R. D., & Weller, J. 2004, *MNRAS*, **355**, 1339
 Battye, R. A., Browne, I. W. A., Dickinson, C., et al. 2013, *MNRAS*, **434**, 1239
 Battye, R., Browne, I., Chen, T., et al. 2016, ArXiv e-prints [arXiv:1610.06826]
 Biggs, A. D., & Ivison, R. J. 2006, *MNRAS*, **371**, 963
 Bigot-Sazy, M.-A., Dickinson, C., Battye, R. A., et al. 2015, *MNRAS*, **454**, 3240
 Bigot-Sazy, M. A., Ma, Y. Z., Battye, R. A., et al. 2016, in *Frontiers in Radio Astronomy and FAST Early Sciences Symposium 2015*, eds. L. Qain, & D. Li, *ASP Conf. Ser.*, **502**, 41
 Bondi, M., Ciliegi, P., Zamorani, G., et al. 2003, *A&A*, **403**, 857
 Bondi, M., Ciliegi, P., Schinnerer, E., et al. 2008, *ApJ*, **681**, 1129
 Braun, R. 2012, *ApJ*, **749**, 87
 Brookes, M. H., Best, P. N., Peacock, J. A., Röttgering, H. J. A., & Dunlop, J. S. 2008, *MNRAS*, **385**, 1297
 Brown, R. L., Lockman, F. J., & Knapp, G. R. 1978, *ARA&A*, **16**, 445
 Bull, P., Ferreira, P. G., Patel, P., & Santos, M. G. 2015, *ApJ*, **803**, 21
 Carilli, C. L., Menten, K. M., Reid, M. J., Rupen, M. P., & Yun, M. S. 1998, *ApJ*, **494**, 175
 Chandola, Y., Sirothia, S. K., & Saikia, D. J. 2011, *MNRAS*, **418**, 1787
 Chandola, Y., Gupta, N., & Saikia, D. J. 2013, *MNRAS*, **429**, 2380
 Chang, T.-C., Pen, U.-L., Peterson, J. B., & McDonald, P. 2008, *Phys. Rev. Lett.*, **100**, 091303
 Chang, T.-C., Pen, U.-L., Bandura, K., & Peterson, J. B. 2010, *Nature*, **466**, 463
 Chen, X. 2012, *Int. J. Mod. Phys. Conf. Ser.*, **12**, 256
 Condon, J. J., Cotton, W. D., Greisen, E. W., et al. 1998, *AJ*, **115**, 1693
 Darling, J., Macdonald, E. P., Haynes, M. P., & Giovanelli, R. 2011, *ApJ*, **742**, 60
 de Gasperin, F., Intema, H. T., & Frail, D. A. 2018, *MNRAS*, **474**, 5008
 De Lucia, G., & Blaizot, J. 2007, *MNRAS*, **375**, 2
 de Zotti, G., Massardi, M., Negrello, M., & Wall, J. 2010, *A&ARv*, **18**, 1
 Dickey, J. M. 1986, *ApJ*, **300**, 190
 Dickey, J. M., Salpeter, E. E., & Terzian, Y. 1978, *ApJS*, **36**, 77
 Field, G. B. 1958, *Proc. IRE*, **46**, 240
 Fomalont, E. B., Kellermann, K. I., Cowie, L. L., et al. 2006, *ApJS*, **167**, 103
 Gallimore, J. F., Baum, S. A., O'Dea, C. P., Pedlar, A., & Brinks, E. 1999, *ApJ*, **524**, 684
 Geréb, K., Maccagni, F. M., Morganti, R., & Oosterloo, T. A. 2015, *A&A*, **575**, A44
 Gupta, N., Salter, C. J., Saikia, D. J., Ghosh, T., & Jeyakumar, S. 2006, *MNRAS*, **373**, 972
 Hales, C. A., Norris, R. P., Gaensler, B. M., et al. 2014a, *MNRAS*, **441**, 2555
 Hales, C. A., Norris, R. P., Gaensler, B. M., & Middelberg, E. 2014b, *MNRAS*, **440**, 3113
 Harper, S. E., Dickinson, C., Battye, R. A., et al. 2018, *MNRAS*, **478**, 2416
 Heywood, I., Jarvis, M. J., & Condon, J. J. 2013, *MNRAS*, **432**, 2625
 Hopkins, A. M., Afonso, J., Chan, B., et al. 2003, *AJ*, **125**, 465

- Hunt, L. K., & Hirashita, H. 2009, *A&A*, **507**, 1327
- Ibar, E., Ivison, R. J., Biggs, A. D., et al. 2009, *MNRAS*, **397**, 281
- Jiao, K., Zhang, J. C., Zhang, T. J., et al. 2019, ArXiv e-prints [arXiv:1905.01184]
- Kanekar, N., Braun, R., & Roy, N. 2011, *ApJ*, **737**, L33
- Kanekar, N., Prochaska, J. X., Smette, A., et al. 2014, *MNRAS*, **438**, 2131
- Kovetz, E. D., Viero, M. P., Lidz, A., et al. 2017, ArXiv e-prints [arXiv:1709.09066]
- Liu, A., & Tegmark, M. 2011, *Phys. Rev. D*, **83**, 103006
- Liu, A., & Tegmark, M. 2012, *MNRAS*, **419**, 3491
- Loeb, A., & Wyithe, J. S. B. 2008, *Phys. Rev. Lett.*, **100**, 161301
- Maccagni, F. M., Morganti, R., Oosterloo, T. A., Geréb, K., & Maddox, N. 2017, *A&A*, **604**, A43
- Mahony, E. K., Morganti, R., Prandoni, I., et al. 2016, *MNRAS*, **463**, 2997
- Masui, K. W., McDonald, P., & Pen, U.-L. 2010, *Phys. Rev. D*, **81**, 103527
- Masui, K. W., Switzer, E. R., Banavar, N., et al. 2013, *ApJ*, **763**, L20
- Mauch, T., & Sadler, E. M. 2007, *MNRAS*, **375**, 931
- Mauch, T., Murphy, T., Buttery, H. J., et al. 2003, *MNRAS*, **342**, 1117
- Moore, C. B., Carilli, C. L., & Menten, K. M. 1999, *ApJ*, **510**, L87
- Morganti, R., & Oosterloo, T. 2018, *ARA&A*, **26**, 4
- Morganti, R., Tadhunter, C. N., & Oosterloo, T. A. 2005, *A&A*, **444**, L9
- Morrison, G. E., Owen, F. N., Dickinson, M., Ivison, R. J., & Ibar, E. 2010, *ApJS*, **188**, 178
- Newburgh, L. B., Bandura, K., Bucher, M. A., et al. 2016, in *Ground-based and Airborne Telescopes VI*, Proc. SPIE, 9906, 99065X
- Obreschkow, D., Croton, D., De Lucia, G., Khochfar, S., & Rawlings, S. 2009a, *ApJ*, **698**, 1467
- Obreschkow, D., Klöckner, H.-R., Heywood, I., Levrier, F., & Rawlings, S. 2009b, *ApJ*, **703**, 1890
- Olivari, L. C., Remazeilles, M., & Dickinson, C. 2016, *MNRAS*, **456**, 2749
- Owen, F. N., & Morrison, G. E. 2008, *AJ*, **136**, 1889
- Padovani, P. 2011, *MNRAS*, **411**, 1547
- Padovani, P., Miller, N., Kellermann, K. I., et al. 2011, *ApJ*, **740**, 20
- Pen, U.-L., Staveley-Smith, L., Peterson, J. B., & Chang, T.-C. 2009, *MNRAS*, **394**, L6
- Peterson, J. B., Bandura, K., & Pen, U. L. 2006, ArXiv e-prints [arXiv:astro-ph/0606104]
- Pihlström, Y. M., Conway, J. E., & Vermeulen, R. C. 2003, *A&A*, **404**, 871
- Prandoni, I., Guglielmino, G., Morganti, R., et al. 2018, *MNRAS*, **481**, 4548
- Richards, E. A. 2000, *ApJ*, **533**, 611
- Roberts, M. S. 1970, *ApJ*, **161**, L9
- Santos, M. G., Alonso, D., Bull, P., Camera, S., & Ferreira, P. G. 2014, in *Statistical Challenges in 21st Century Cosmology*, eds. A. Heavens, J. L. Starck, & A. Krone-Martins, *IAU Symp.*, **306**, 165
- Santos, M., Bull, P., & Alonso, D. 2015, *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, 19
- Santos, M. G., Cluver, M., Hilton, M., et al. 2017, ArXiv e-prints [arXiv:1709.06099]
- Seymour, N., Dwelly, T., Moss, D., et al. 2008, *MNRAS*, **386**, 1695
- Shaw, J. R., Sigurdson, K., Pen, U.-L., Stebbins, A., & Sitwell, M. 2014, *ApJ*, **781**, 57
- Smoot, G. F., & Debono, I. 2017, *A&A*, **597**, A136
- Square Kilometre Array Cosmology Science Working Group, Bacon, D. J., Battye, R. A., et al. 2018, ArXiv e-prints [arXiv:1811.02743]
- Switzer, E. R., Masui, K. W., Bandura, K., et al. 2013, *MNRAS*, **434**, L46
- Switzer, E. R., Chang, T.-C., Masui, K. W., Pen, U.-L., & Voytek, T. C. 2015, *ApJ*, **815**, 51
- van Gorkom, J. H., Knapp, G. R., Ekers, R. D., et al. 1989, *AJ*, **97**, 708
- Vermeulen, R. C., Pihlström, Y. M., Tschager, W., et al. 2003, *A&A*, **404**, 861
- Villaescusa-Navarro, F., Alonso, D., & Viel, M. 2017, *MNRAS*, **466**, 2736
- Villaescusa-Navarro, F., Genel, S., Castorina, E., et al. 2018, *ApJ*, **866**, 135
- Wilman, R. J., Miller, L., Jarvis, M. J., et al. 2008, *MNRAS*, **388**, 1335
- Windhorst, R., Mathis, D., & Neuschaefer, L. 1990, in *Evolution of the Universe of Galaxies*, ed. R. G. Kron, *ASP Conf. Ser.*, **10**, 389
- Wolfe, A. M., Gawiser, E., & Prochaska, J. X. 2005, *ARA&A*, **43**, 861
- Wolz, L., Abdalla, F. B., Blake, C., et al. 2014, *MNRAS*, **441**, 3271
- Wolz, L., Blake, C., & Wyithe, J. S. B. 2017, *MNRAS*, **470**, 3220
- Yu, H.-R., Zhang, T.-J., & Pen, U.-L. 2014, *Phys. Rev. Lett.*, **113**, 041303

Appendix A: Other absorptions in HI IM

Here we discuss absorptions that will be imprinted in the signal picked up by an HI IM experiment, other than absorption of flux incident from background radio continuum sources by the HI emitting sources themselves (the main topic of this paper). The strength of these absorptions are likely to be insignificant compared to the main type of absorption discussed in this paper, but are listed here for completeness.

A.1. Within the HI emitting sources

A.1.1. HI self-absorption

HI self-absorption occurs when the HI responsible for emission has very high optical depths ($\tau > 1$). The self-absorption can be quite high in the highest column density regions of nearby disk galaxies (Braun 2012). But the volume filling factor of high column density HI is insignificant when considering the entire disk of a gas-rich galaxy ($\sim 1\%$), which is the typical HI emitting source. HI IM in the foreseeable future will be done (at best) at resolutions of few arcminutes, because of which the studies will only be sensitive to the average HI column densities over the disks of the HI emitting sources. Therefore self-absorption is unlikely to be an important effect on scales relevant for HI IM experiments, and the assumption made regarding the HI emitting gas being optically thin is justifiable.

A.1.2. Absorption by electrons

A secondary internal source of absorption of the HI emission from a given source are electrons, the same ones which are responsible for synchrotron and free-free continuum emission from the source itself. Such emission again has a low volume filling factor when integrated over the entire disk of a galaxy, because the radio continuum is dominated by localized dense HII regions and the electron density in such regions drops sharply with their size (Hunt & Hirashita 2009).

A.2. Intervening absorption

A set of absorptions will occur due to absorbing media situated in between the HI emission source and the observer, including our Galaxy.

A.2.1. HI absorption by other sources in voxel

The HI flux emitted by any individual source within a voxel being observed by an HI IM experiment is subject to possible absorption by the HI and the electrons responsible for radio continuum emission (see Appendix A.1.2) of intervening HI sources within the voxel itself. Based on the arguments listed in Appendix A.1.2, the absorption by the electrons producing the radio continuum is likely to be negligible. As for the absorption by HI in the intervening sources within the voxel, in addition to the low probability of HI sources within a voxel being observed to spatially overlap each other, the optical depth responsible for the absorption will be low because of the coarse beam size. Given the emitted HI flux that is being absorbed is itself minuscule compared to say the radio continuum flux from background sources, such absorption is likely to be significantly sub-dominant to the absorption of flux from background radio continuum sources by the HI sources under consideration.

A.2.2. Absorption by electrons in intervening sources including our Galaxy

Regarding the absorption that can occur in sources between $z = 0$ and the voxel being observed, there can be no absorption by the HI in such intervening sources including the Galaxy given velocity considerations. The velocity spread of the HI is such that HI IM experiments beyond the very local ($z < 0.005$) Universe will remain unaffected.

There will be absorption though by the electrons in the intervening sources responsible for radio continuum emission, especially so in the Galaxy itself. Such absorption can occur at all frequencies below 1.4 GHz to which the HI 21 cm emission signal from the sources being observed might be redshifted to, even as low as ~ 400 MHz for HI IM experiments observing the Universe at $z = 2.5$. At these low frequencies though the amount of absorption by electrons producing synchrotron and free-free emission is likely to be small. In intervening sources the emission will become insignificant given the small volume filling factors of the HII regions which dominates the radio continuum emission (see Appendix A.1.2). As for our Galaxy, HI IM experiments will observe the sky out of the plane of the Galaxy in order to avoid the regions with the strongest synchrotron and free-free emission.

A.2.3. Absorption by other species

There is also the possibility of absorption in intervening sources and the Galaxy by other species with rest frame wavelengths which happen to match that of the redshifted HI 21 cm emission incident on them. Specifically, a number of hydrogen recombination lines have frequencies in this range. The optical depth of such absorption is negligible except at very high densities, as is found in the densest HII regions (Brown et al. 1978). Therefore as in Appendices A.1.2 and A.2.2, this type of absorption will be negligible due to volume filling factor considerations.

A.3. Associated HI absorption

There is a special type of absorption which will also be included in the signal picked up by HI IM experiments, but which does not arise along the sightline of the HI emitting sources discussed here or in Appendix A. We are referring to “associated” (as opposed to “intervening”) HI absorption within radio continuum sources which happen to be in the voxel of sky being observed, by HI clouds within the sources themselves HI clouds (see Morganti & Oosterloo 2018, for a review). HI clouds are approximately 100 pc in size, for example see Braun (2012). Therefore realistically such absorption will only be detected against sources whose radio continuum emission is dominated by a very compact core – core-dominated active galactic nuclei (AGNs) and nuclear starbursts are the obvious candidates, with edge-on viewing angles favoured. The number of detections of associated HI absorption are rare (< 100 detections) even with targeted searches, especially at $z > 0.25$ (e.g. Roberts 1970; Dickey 1986; van Gorkom et al. 1989; Carilli et al. 1998; Gallimore et al. 1999; Moore et al. 1999; Pihlström et al. 2003; Vermeulen et al. 2003; Morganti et al. 2005; Gupta et al. 2006; Chandola et al. 2011, 2013; Darling et al. 2011; Allison et al. 2012, 2014; Geréb et al. 2015; Maccagni et al. 2017; Aditya & Kanekar 2018a,b; Aditya 2019). The bright continuum sources against which such absorption might be detectable will in any case be flagged in an HI IM experiment (conversely see Appendix B for an estimate of how many “detections” can be expected for a given

HI IM experiment when considering the brighter end of radio continuum sources). As for the non-detectable radio continuum background, the fainter end of the radio continuum emission is increasingly dominated by star forming dwarf galaxies and the fraction of starbursts and AGNs decreases rapidly with flux (Padovani 2011). Moreover given the fact that such absorption is spatially localized, the low spatial resolution of upcoming HI experiments will dramatically reduce the amount of absorption that will be picked up. Therefore associated HI absorption is unlikely to be of significance for upcoming HI IM experiments, though a detailed quantification of its contribution especially for high resolution HI IM experiments should ideally be undertaken in the future.

Appendix B: Detections of associated HI absorption

In Appendix A we mentioned how associated HI absorption against the bright radio emission from the centres of radio continuum sources lying within the voxel of sky being observed will be picked up by HI IM experiments. We also reasoned why such absorption from unflagged radio continuum sources in the data is likely to be insignificant. Here we reverse the question and ask how many associated HI absorption detections will be possible for a given HI IM experiment. Thus the brightest radio continuum sources that would otherwise be flagged for HI IM, become potential targets for new associated HI absorption detections. Given that detections of associated HI absorption are limited, and HI IM experiments will observe large areas of the sky much of which has not been explored in previous searches of associated absorption using existing facilities, data from HI IM experiments can potentially be used to search for new associated HI absorption detections.

Number of detections for a test case: BINGO

Here we calculate the potential number of detections of associated HI absorption, given the parameters of an actual HI IM experiment viz. BAO from Integrated Neutral Gas Observations (BINGO, Battye et al. 2016). Upcoming HI IM experiments will have beam sizes between few arcminutes to tens of arcminutes. For BINGO the beam size will be 40'. Therefore we based our calculations on the detection rates achieved by associated HI absorption searches using single-dishes.

Allison et al. (2014) searched in archival HI Parks All Sky Survey (HIPASS) data for associated HI absorption within a sample of 204 radio and star-forming galaxies within a volume bounded by a maximum redshift of 0.042 and covering 34 432 deg² in the sky. HIPASS has a rms noise of ~13 mJy per 15.5' beam, with channel width of 13.2 km s⁻¹ at $z = 0$. The rms noise can potentially vary significantly across HIPASS. Therefore to ensure a $>5\sigma$ detection of a peak absorption of 30%, the continuum flux density limit was set to ~250 mJy and only sources above that limit (at either 843 MHz or 1.4 GHz) were targeted. Absorption was detected in four sources, of which one is a compact starburst, two are compact AGNs when compared to the respective stellar disks, and the fourth one is an extended AGN. The detection rates were 2% for the entire sample, as well as for the star-forming and AGN dominated samples separately. It is notable that their determinations of the total number of galaxies which are either star-forming or are AGN dominated, match quite well with the numbers predicted using the local radio luminosity functions for the two classes

Table B.1. Associated HI absorptions that will be detectable using the HI IM survey with BINGO covering 5000 deg² of sky.

\bar{z}_{mean}	5σ (mJy)	$S_{\text{cont}}^{\text{limit}}$ (mJy)	Vol. (Gpc ³)	No. of SF galaxies	No. of AGNs	Total detections
0.15	4.9	16.3	0.262	5	1334	27
0.25	4.1	13.7	0.643	<1	2247	45
0.35	3.5	11.7	1.127	≪1	2598	52
0.45	3.1	10.3	1.666	≪1	2512	50

of galaxies from Mauch & Sadler (2007). Allison et al. (2014) also compare their detection rates with those from the Arcibo Legacy Fast ALFA Survey (ALFALFA) pilot survey (Darling et al. 2011), which is a survey of 517 deg² of the sky in the redshift range $0 < z < 0.058$ (with 10 km s⁻¹ velocity resolution and a beam size of 3.3' × 3.8'). The flux density limit for continuum sources against which absorption is expected to be detected is 42 mJy, considering a 5σ detection of at least 30% peak absorption. There was only a single detection of (strong) absorption, among 29 potential detectable candidates as calculated using the radio luminosity functions from Mauch & Sadler (2007), which leads to ~3% detection rate which Allison et al. (2014) claim to be consistent with their number. But Allison et al. (2014) did not take into account the 10% of redshift range found unusable by Darling et al. (2011), which would move the detection rate closer to 4%. Therefore there is an indication that the detection rate of associated HI absorption increases as the flux density limit for continuum sources is lowered. Nevertheless, because of small number statistics we do not base our estimates on Darling et al. (2011) but on the detection rate from Allison et al. (2014).

BINGO will survey a 5000 deg² patch of the sky, lying in between HIPASS and the ALFALFA Pilot Survey coverages. It will have much better sensitivity than either, while at the same time have a worse resolution (40' beam) than either. The redshift range covered by it will be $0.12 < z < 0.48$, higher than that covered by either survey. This also implies that the potential number of AGNs will be larger than what will be calculated using the luminosity functions from Mauch & Sadler (2007), as their fits were based on the Six-degree Field Galaxy Survey (6dFGS) matched with NVSS, and the galaxies have a median redshift of 0.043. We calculated the number of expected detections of HI absorption for 5σ detection of at least 30% peak absorption against background continuum sources – in order to base our calculations on the numbers from Allison et al. (2014). Therefore we are conservative regarding our estimates given the potential increase in detection rates with better continuum flux sensitivity discussed earlier. As BINGO will cover a large redshift range and the sensitivity of the instrument will vary across the range, we do the calculations for four $\Delta z = 0.1$ redshift slices.

Table B.1 presents the results of calculating the number of expected detections for BINGO for the different redshift slices. Column 1 lists the mean redshift for each of these slices, and quantities like the luminosity distance for sources in each bin is calculated based on this mean redshift. Although the actual thermal noise target for BINGO is lower, the relevant rms noise to consider is that due to the integrated HI emission from sources within the voxel being observed (the signal the experiment aims to detect), which is expected to be of the order of 0.1 mK, for which the corresponding 5σ limits are given in Col. 2. The continuum flux limits assuming that the absorption will be detected at 5σ when having a minimum τ_{peak} of 0.3, are given in Col. 3 of Table 1. We note that other continuum sources in the beam will

not change these numbers – they will merely raise the level of the total continuum emission measured. Columns 4 and 5 give the expected number of continuum sources above the flux limit for star forming galaxies and radio-loud AGNs respectively. They are calculated using Eqs. (5) and (6) from [Mauch & Sadler \(2007\)](#). As expected at these redshifts for the high power end of the radio luminosity function, AGNs dominate the number counts. Finally, Col. 6 gives the expected number of detections in each redshift bin for a 2% detection rate.

As we can see from Table 1, even with a conservative 2% detection rate, and estimating source numbers based on a radio luminosity function which is biased towards low redshift sources and therefore does not take into account higher occurrences of AGNs with increasing redshift, we expect a large number of associated HI absorption detections with BINGO. A point to note here is given the typical velocity widths of 100 km s^{-1} or

more for these absorption features, one should aim for channel widths no smaller than 10 km s^{-1} . At the higher redshift end of the band this corresponds to around 30 KHz. For detailed science regarding determining inflow/outflow velocities and source morphologies, each detection should ideally be followed up by high resolution interferometric observations.

According to the classification of HI IM experiments by [Bull et al. \(2015\)](#), the targeted IM experiment BINGO is a Stage I setup. Such experiments are to be superseded by more sensitive experiments with wider sky coverage, through Stage II and III experiments. With conservative assumptions, we expect BINGO to more than double the number of known detections of associated HI absorptions. Therefore detection and follow-up studies of associated HI absorption will be an obvious avenue for commensal science using upcoming HI IM experiments.