Constraining a halo model for cosmological neutral hydrogen

Hamsa Padmanabhan* and Alexandre Refregier†
Institute for Astronomy, ETH Zurich, Wolfgang-Pauli-Strasse 27, CH-8093 Zurich, Switzerland

Accepted —. Received —; in original form —

ABSTRACT

We describe a combined halo model to constrain the distribution of neutral hydrogen (HI) in the post-reionization universe. We combine constraints from the various probes of HI at different redshifts: the low-redshift 21-cm emission line surveys, intensity mapping experiments at intermediate redshifts, and the Damped Lyman-Alpha (DLA) observations at higher redshifts. We use a Markov Chain Monte Carlo (MCMC) approach to combine the observations and place constraints on the free parameters in the model. Our best-fit model involves a relation between neutral hydrogen mass $M_{\text{HI}}$ and halo mass $M$ with a non-unit slope, and an upper and a lower cutoff. We find that the model fits all the observables but leads to an underprediction of the bias parameter of DLAs at $z \sim 2.3$. We also find indications of a possible tension between the HI column density distribution and the mass function of HI-selected galaxies at $z \sim 0$. We provide the central values of the parameters of the best-fit model so derived. We also provide a fitting form for the derived evolution of the concentration parameter of HI in dark matter haloes, and discuss the implications for the redshift evolution of the HI-halo mass relation.

Key words: cosmology: observations – radio lines: galaxies – galaxies: evolution

1 INTRODUCTION

Studying the evolution of neutral hydrogen (HI) in the post-reionization universe offers key insights into cosmology and galaxy formation. Intensity mapping, in which the HI intensity fluctuations are mapped out without the resolution of individual galaxies, is a novel technique which allows the study of HI on large scales and promises unprecedented constraints on galaxy formation and evolution, cosmology (Chang et al. 2010; Masui et al. 2013; Switzer et al. 2013; Santos et al. 2015; Bull et al. 2015) and stringent constraints on models of dark energy (e.g., Chang et al. 2008) and modified gravity (e.g., Hall et al. 2013). The 21-cm intensity power spectrum is also an important tracer of the underlying large-scale structure in the post-reionization universe ($z \sim 0$), due to the absence of the complicated reionization astrophysics at these epochs.

The key astrophysical ingredients in the estimation of the 21-cm intensity mapping power spectrum are the density parameter of HI, $\Omega_{\text{HI}}$, and its bias $b_{\text{HI}}$, which measures how the HI is clustered with respect to the underlying dark matter. Constraints on these parameters come from a variety of astrophysical probes (a detailed summary is available in Padmanabhan et al. (2015), hereafter Paper 1). At low redshifts, HI gas in galaxies and their environments is primarily studied through blind 21-cm emission line surveys of nearby galaxies like the HI Parkes All Sky Survey (HIPASS; Zwaan et al. 2005a) and the Arecibo Fast Legacy ALFA (ALFALFA) Survey (Martin et al. 2010). These provide measurements of the mass function of HI-selected galaxies, from which the density parameter of HI in galaxies, $\Omega_{\text{HI,gal}}$, can be derived. The clustering of the HI-selected galaxies, which constrains the galaxy bias $b_{\text{HI,gal}}$, can also be measured from these surveys (Martin et al. 2012). At intermediate redshifts ($z \sim 0$), intensity mapping experiments typically constrain the combination $\Omega_{\text{HI}}/b_{\text{HI}}$ (e.g., Switzer et al. 2013).

At higher redshifts, the major reservoirs of HI gas in the post-reionization universe are the Damped Lyman-Alpha (DLA) systems. These systems have very high column densities ($>10^{20.3}$ cm$^{-2}$) and are found to contain more than 80 % of the neutral hydrogen at redshifts 2-5 (e.g., Lanzetta et al. 1991). DLAs have been studied through line-of-sight absorption against bright background quasars (Noterdaeme et al. 2012; Prochaska & Wolfe 2009), as well as direct imaging surveys (Fumagalli et al. 2014, 2015). The study of DLA systems enables the measurements of the column density distribution, $f_{\text{HI}}(N_{\text{HI}})$ of the DLAs at various redshifts, the incidence $dN/dX$ of the DLAs per unit absorption distance interval, and the density parameter of HI in DLAs, $\Omega_{\text{DLA}}$. Recently, the large-scale bias parameter of DLA systems, $b_{\text{DLA}}$ has also been constrained through a cross-correlation study with the Lyman-$\alpha$ forest (Font-Ribera et al. 2012).

Analytical and simulation techniques have typically been used to model the HI radio observations at lower redshifts (e.g., Davé et al. 2013; Bagla et al. 2010; Guha Sarkar et al. 2012; Kim et al. 2016; Villaescusa-Navarro et al. 2015) and the DLA observations at higher redshifts (Fumagalli et al. 2011; Bird et al. 2014; Barnes & Haehnelt 2014) separately. Modelling the the 21-cm based observables requires knowledge of the HI-halo mass relation, $M_{\text{HI}}(M)$ which is combined with the radial distribution of HI (the HI den-
In this section, we describe the details of the modelling of H\textsubscript{I} 21-cm intensity fluctuations, as also for the connections between \(M_{\text{HI}}-M\) parameter halo model which is a generalization of the H\textsubscript{I} in galaxies. Besides being of direct relevance in determining the bias factor of neutral hydrogen (as seen in e.g., Paper 2), the two cutoffs are also related to physical processes that influence the masses of the dark matter haloes that are able to host H\textsubscript{I}. It is found as a result of simulations (Pontzen et al. 2008), that halos with virial velocities of less than 30 km/s and greater than 200 km/s preferentially do not host neutral hydrogen. However, these cutoffs, especially the upper cutoff \(v_{c1}\), are also directly linked to \(\beta\), which provides a measure of the effective weights given to the range of dark matter halo masses that host H\textsubscript{I}. Hence, these two cutoffs are also important in the contexts of galaxy formation and feedback processes leading to the generation and destruction of H\textsubscript{I}.

Together with the above H\textsubscript{I}-halo mass relation, we also use a radial density profile, \(\rho_{\text{HI}}(r)\) to describe the distribution of H\textsubscript{I} in the dark matter halo. This is modelled following Maller & Bullock (2004), the form being motivated by the simulation constraints on cooling in multi-phase gas in halos:

\[
\rho_{\text{HI}}(r) = \frac{\rho_0 r_\text{s}^3}{(r + 0.75 r_\text{s})(r + r_\text{s})^2}
\]  

(2)

The form is an altered Navarro-Frenk-White (NFW; Navarro et al. 1997) profile, where \(r_\text{s}\) is the scale radius of the halo, defined through \(r_\text{s} = R_c(M)/c(M, z)\), with \(R_c(M)\) being the virial radius. The H\textsubscript{I} concentration parameter, \(c(M, z)\) is given by:

\[
c(M, z) = c_{\text{HI}} \left( \frac{M}{10^{11} M_\odot} \right)^{-0.109} \frac{4}{(1 + z)^2}
\]  

(3)

The \(M_{\text{HI}}-M\) relation is modelled as:

\[
M_{\text{HI}}(M) = \alpha f_{H,e} M \left( \frac{M}{10^{11} h^{-1} M_\odot} \right)^\beta \exp \left[ -\left( \frac{v_{c0}}{v_c(M)} \right)^3 \right] \times \exp \left[ -\left( \frac{v_c(M)}{v_{c1}} \right)^3 \right]
\]

(1)

which is a generalization of the HI-halo mass relations used previously in the literature (Barnes & Haehnelt 2014, and Paper 2). The relation involves the following quantities:

(i) the multiplicative constant \(\alpha\), the fraction of H\textsubscript{I} relative to the cosmic \(f_{H,e}\)

(ii) the excess (logarithmic) slope, \(\beta\), of the \(M_{\text{HI}}-M\) relation, taken to be zero in previous studies (so that the overall power of \(M\) was unity), the relation is pivoted at \(10^{11} h^{-1} M_\odot\); (iii) the two cutoffs in the virial velocity \(v_c(M)\) of the halo; \(v_{c0}\) (the lower cutoff) and \(v_{c1}\) (the upper cutoff), used previously in Paper 2; the lower cutoff alone has been used in DLA-based analyses such as Barnes & Haehnelt (2014).

Each of the factors in the above \(M_{\text{HI}}-M\) relation has a physical implication. The multiplicative constant, \(\alpha\), denotes the fraction of hydrogen relative to cosmic and is directly related to the redshift evolution of H\textsubscript{I} and its depletion due to star-formation and feedback processes. From observational and simulation studies of the neutral hydrogen evolution, this value is expected to be fairly constant or slowly evolving over redshifts (e.g., Crighton et al. 2015; Davé et al. 2013).

The value of \(\beta\) represents the correction to unity of the logarithmic slope of the H\textsubscript{I}-halo mass relation. It is seen from the results of abundance matching (e.g., Papastergis et al. 2013) that the observationally favoured value of \(\beta\) is less than zero, such that the power of the halo mass in the H\textsubscript{I}-halo mass relation is less than unity. The two cutoffs, \(v_{c0}\) (the lower cutoff) and \(v_{c1}\) (the upper cutoff) determine, at different redshifts, the range of dark matter halo masses that contribute chiefly as hosts of H\textsubscript{I} in galaxies. Besides being of direct relevance in determining the bias factor of neutral hydrogen (as seen in e.g., Paper 2), the two cutoffs are also related to physical processes that influence the masses of the dark matter haloes that are able to host H\textsubscript{I}. It is found as a result of simulations (Pontzen et al. 2008), that halos with virial velocities of less than 30 km/s and greater than 200 km/s preferentially do not host neutral hydrogen. However, these cutoffs, especially the upper cutoff \(v_{c1}\), are also directly linked to \(\beta\), which provides a measure of the effective weights given to the range of dark matter halo masses that host H\textsubscript{I}. Hence, these two cutoffs are also important in the contexts of galaxy formation and feedback processes leading to the generation and destruction of H\textsubscript{I}.

In this section, we describe the details of the modelling of H\textsubscript{I} and its depletion due to star-formation and feedback processes. From observational and simulation studies of the neutral hydrogen evolution, this value is expected to be fairly constant or slowly evolving over redshifts (e.g., Crighton et al. 2015; Davé et al. 2013).

The value of \(\beta\) represents the correction to unity of the logarithmic slope of the H\textsubscript{I}-halo mass relation. It is seen from the results of abundance matching (e.g., Papastergis et al. 2013) that the observationally favoured value of \(\beta\) is less than zero, such that the power of the halo mass in the H\textsubscript{I}-halo mass relation is less than unity.

The two cutoffs, \(v_{c0}\) (the lower cutoff) and \(v_{c1}\) (the upper cutoff) determine, at different redshifts, the range of dark matter halo masses that contribute chiefly as hosts of H\textsubscript{I} in galaxies. Besides being of direct relevance in determining the bias factor of neutral hydrogen (as seen in e.g., Paper 2), the two cutoffs are also related to physical processes that influence the masses of the dark matter haloes that are able to host H\textsubscript{I}. It is found as a result of simulations (Pontzen et al. 2008), that halos with virial velocities of less than 30 km/s and greater than 200 km/s preferentially do not host neutral hydrogen. However, these cutoffs, especially the upper cutoff \(v_{c1}\), are also directly linked to \(\beta\), which provides a measure of the effective weights given to the range of dark matter halo masses that host H\textsubscript{I}. Hence, these two cutoffs are also important in the contexts of galaxy formation and feedback processes leading to the generation and destruction of H\textsubscript{I}.

Together with the above H\textsubscript{I}-halo mass relation, we also use a radial density profile, \(\rho_{\text{HI}}(r)\) to describe the distribution of H\textsubscript{I} in the dark matter halo. This is modelled following Maller & Bullock (2004), the form being motivated by the simulation constraints on cooling in multi-phase gas in halos:

\[
\rho_{\text{HI}}(r) = \frac{\rho_0 r_\text{s}^3}{(r + 0.75 r_\text{s})(r + r_\text{s})^2}
\]

(2)

The form is an altered Navarro-Frenk-White (NFW; Navarro et al. 1997) profile, where \(r_\text{s}\) is the scale radius of the halo, defined through \(r_\text{s} = R_c(M)/c(M, z)\), with \(R_c(M)\) being the virial radius. The H\textsubscript{I} concentration parameter, \(c(M, z)\) is given by:

\[
c(M, z) = c_{\text{HI}} \left( \frac{M}{10^{11} M_\odot} \right)^{-0.109} \frac{4}{(1 + z)^2}
\]

(3)
In the above relation, $c_{HI}$ is a free parameter, analogous to the dark matter halo concentration $c_0 = 3.4$ in, e.g., Macciò et al. (2007). This form of the profile is found to reproduce well the observed properties of high-redshift DLAs, though it is generally observed (Barnes & Haehnelt 2010, 2014, Paper 2) that the value of $c_{HI}$ is higher than $c_0$. In other words, the HI gas is more tightly concentrated than the dark matter. The value of $\gamma$ in the above relation was taken to be unity in previous analyses, including the case for dark matter alone. In the present analysis, it signifies the extent of the evolution correction due to the redshift evolution of the parameter $c_{HI}$, i.e., the value $\gamma = 1$ translates into no redshift evolution of the concentration parameter $c_{HI}$.

The value of $\rho_0$ in Eq. (2) is determined by normalization:

$$\int_0^{R_{IM}(M)} 4\pi r^2 \rho_{HI}(r)dr = M_{HI}(M)$$

with the total mass of HI in the halo. Hence, all the free parameters in the model are involved in the description of the HI profile.

Once the $M_{HI} - M$ relation and the form of the HI profile are known, the various quantities related to the HI observations can be derived analytically (a detailed description of the method is available in Paper 2):

(i) The column density as a function of the impact parameter $s$ of a line-of-sight through a DLA system:

$$N_{HI}(s) = \frac{2}{\pi} \int_0^{R_{IM}(M)} \rho_{HI}(r)dr$$

where $M_{HI}$ is the proton mass, and the cross section is defined by $\sigma_{DLA}(M) = \pi s^2$ where $s$ is the impact parameter corresponding to the column density reaching the DLA limit: $N_{HI}(s_*) = 10^{20.4} \text{cm}^{-2}$.

(ii) The incidence, $dN/dX$ and the bias of the DLAs, $b_{DLA}$, both of which can be derived from the cross section:

$$\frac{dN}{dX} = \frac{c}{H_0} \int_0^{\infty} n(M, z)\sigma_{DLA}(M, z) dM$$

$$b_{DLA}(z) = \frac{\int_0^{\infty} dM n(M, z) b(M, z)\sigma_{DLA}(M, z)}{\int_0^{\infty} dM n(M, z)\sigma_{DLA}(M, z)}$$

(iii) The column density distribution $f_{HI}(N_{HI}, z)$, given by:

$$f_{HI}(N_{HI}, z) = \frac{c}{H_0} \int_0^{\infty} n(M, z) dM$$

with the dark matter halo bias $b(M, z)$ following, e.g., Scoccimarro et al. (2001).

We can use the available data to constrain the free parameters in the model above. The available data are summarized in the next section and include the column density distribution at redshifts 0, 1 and 2.3, the large-scale bias of HI selected galaxies from ALFALFA at $z \sim 0$, the bias of DLAs estimated at $z \sim 2.3$, and the DLA incidence $dN/dX$ from other observations at these redshifts. We use a Markov Chain Monte Carlo approach to constrain the free parameters $c_{HI}$, $\alpha$, $\beta$, $\gamma$, $v_1$, and $v_2$, from the observations, as summarized in Sec. 4.

3 DATA

In this section, we provide a brief summary of the data at different redshifts used to fit to the model and estimate the best-fitting values of the free parameters (a detailed summary of the data related to HI observations across redshifts are in Papers 1 and 2):

(i) The column density distribution of DLAs at redshifts 0 from the WHISP survey (Zwaan et al. 2005b), from redshift 1 (Rao et al. 2006), and the observations at redshift 2.3 (Noterdaeme et al. 2012),

(ii) the DLA incidence (Zwaan et al. 2005b; Rao et al. 2006; Braun 2012; Zafar et al. 2013) at the above redshifts, the value of $\Omega_{HI} M_{z} \sim 2.3$ from Zafar et al. (2013),

(iii) the HI mass function at $z \sim 0$ from the HIPASS survey (Zwaan et al. 2005a),

(iv) the product $\Omega_{HI} M_{HI}$ at $z \sim 1$ from the results of intensity mapping (Switzer et al. 2013), and

(v) the clustering of the DLA systems at $z \sim 2.3$ (Font-Ribera et al. 2012).

Of these data, (i), (ii), (iv) and (v) were also used in the previous analysis comparing different models of the HI and DLA observations (Paper 2). The HI mass function, added in to the present data, is not a monotonic function, the contributions from the two sides of the turnover point are added when computing the term $dM/dM_{HI}$. In practice, it is found that the contribution from masses higher than the maximum value of $M_{HI}(M)$ is negligible, so we assume a sharp cutoff (instead of exponential) at the maximum value of $M_{HI}(M)$ for the computation.

Note that if the $M_{HI}(M)$ is not a monotonic function, the contributions from the two sides of the turnover point are added when computing the term $dM/dM_{HI}$. In practice, it is found that the contribution from masses higher than the maximum value of $M_{HI}(M)$ is negligible, so we assume a sharp cutoff (instead of exponential) at the maximum value of $M_{HI}(M)$ for the computation.
4 COMBINING CONSTRAINTS: MCMC ANALYSIS

We use a Markov Chain Monte Carlo (MCMC) analysis with the COSMOHAMMER routine (Akeret et al. 2013) to explore the available parameter space, combine the constraints and sample the likelihood to find the best-fitting values and errors on the free parameters. The cosmological parameters used are consistent with the earlier analysis in Paper 2: a flat Λ CDM cosmology with $\Omega_m = 0.281$, $\Omega_{\Lambda} = 0.719$, $\Omega_b = 0.0462$, $\sigma_8 = 0.81$, $n_s = 0.963$, and $h = 0.71$.

4.1 Parameters and priors

We parametrize the model by the six free parameters: $c_{\text{HI}}$, $\alpha$, $\beta$, $\gamma$, $v_{c,0}$ and $v_{c,1}$. The model assumes flat priors for each of the parameters, with the ranges for the priors indicated in the second column of Table 1.

4.2 Likelihood

The likelihood is defined by:

$$L = \exp \left( -\frac{\chi^2}{2} \right)$$  \hspace{1cm} (13)

where the $\chi^2$ is calculated by:

$$\chi^2 = \sum \frac{(g_{i,\text{obs}} - g_i)^2}{\sigma_i^2}$$  \hspace{1cm} (14)

In the above equation, the $g_i$’s are the model predictions for the different quantities, the $g_{i,\text{obs}}$ are the observed values, and the $\sigma_i$’s are the errors on the observations which are assumed independent. We note that the observations of the column density distribution and the HI mass function at $z \sim 0$ provide only the statistical errors. To obtain an indicative measure of the systematic effects on the derived parameter values, we:

(a) use the systematic uncertainties on the parameters of the best-fit Schechter form of the HIPASS mass function at $z \sim 0$ (Zwaan et al. 2005a) and add these in quadrature to the statistical uncertainties; the resulting error band is shown in grey in the second panel of Fig. 1, and

(b) perform the analysis separately with the WHISP (Zwaan et al. 2005b) and the Braun (2012) measurements of the column density distribution at $z \sim 0$, in order to provide a measure of the effect of the systematic uncertainty on the derived best-fit parameters. We describe the results of fitting to the WHISP sample in the main text, the comparison to the Braun (2012) data is provided in the Appendix.

4.3 Results

We sample the likelihood using 100 chains each of 10 random walkers for each of the 6 parameters, making a total of 6000 chains. 4 The best-fitting values and the uncertainties so derived for the parameters are in the last two columns of Table 1. For the quoted best-fit values and errors, we use the mean and 1σ uncertainties of the values from the MCMC analysis.

A comparison of the best-fit model to the data is provided in the panels of Figs. 1, 2 and 3, for the redshifts 0, 1 and 2.4 respectively. We note that, at low redshifts, the systematic uncertainties in the reported measurements of the column density distribution and the mass function may be significant compared to the statistical uncertainties. To provide an indicative measure of the systematic effects, we have plotted multiple measurements where available, for example, the Braun (2012) and the Zwaan et al. (2005b) measurements of the column density distribution, and the Zwaan et al. (2005a) and Martin et al. (2010) measurements of the mass function at $z \sim 0$. To explore the possibility of the observed tension being alleviated by the systematics in the measurements at $z \sim 0$, especially in the case of $f_{\text{HI}}$, we repeat the analysis with the Braun (2012) data (in place of the WHISP data) in the Appendix. We find, as a result of this analysis, that the systematics in the data may not fully alleviate the observed tension.

The posterior probabilities from the MCMC analysis, and the parameter space are shown in Fig. 4. The resultant form of $M_{\text{HI}}(M)$ at $z = 0$ is shown in the left panel of Fig. 5, along with the results of other studies in the literature (Marin et al. 2010; Davé et al. 2013). The evolution of the $M_{\text{HI}}(M)$ relation for redshifts 0-4 is shown in the right panel of Fig. 5, along with the relation for $M_{\text{HI}} \propto M$ (shown by the dashed line). It can be seen that the $M_{\text{HI}} - M$ relation derived from the best-fitting model has a shallower slope in all cases, illustrating the effect of negative values for $\beta$.

There is no redshift evolution in the present $M_{\text{HI}}(M)$, apart from that coming from the evolution of the virial velocity to halo mass relation. A possible evolution in the parameter $\beta$ may be needed to match the observed clustering bias of DLAs at $z \sim 2.3$, which we explore in greater detail in future work. The evolution of the $M_{\text{HI}} - M$ relation can be constrained by current and future intensity mapping experiments which measure the large-scale

---

Table 1. Summary of the free parameters, their flat prior ranges, the best-fitting values and their uncertainties obtained from the model.

| Parameter | Range for prior | Best fit | Error (1σ) |
|-----------|----------------|----------|------------|
| $c_{\text{HI}}$ | [20, 400] | 113.80 | 14.03 |
| $\alpha$ | [0.05, 0.5] | 0.17 | 0.02 |
| $\log v_{c,0}$ (km/s) | [1.30, 1.90] | 1.57 | 0.03 |
| $\log v_{c,1}$ (km/s) | [2.1, 6.5] | 4.39 | 0.75 |
| $\beta$ | [-1.3] | -0.55 | 0.12 |
| $\gamma$ | [-0.9,2] | 0.22 | 0.67 |

Note that, at low redshifts, the systematic uncertainties in the reported measurements of the column density distribution and the mass function may be significant compared to the statistical uncertainties. To provide an indicative measure of the systematic effects, we have plotted multiple measurements where available, for example, the Braun (2012) and the Zwaan et al. (2005b) measurements of the column density distribution, and the Zwaan et al. (2005a) and Martin et al. (2010) measurements of the mass function at $z \sim 0$. To explore the possibility of the observed tension being alleviated by the systematics in the measurements at $z \sim 0$, especially in the case of $f_{\text{HI}}$, we repeat the analysis with the Braun (2012) data (in place of the WHISP data) in the Appendix. We find, as a result of this analysis, that the systematics in the data may not fully alleviate the observed tension.

4 We find the likelihood to be adequately converged for the above number of iterations. For computational efficiency, we restrict the fitting of the column density distribution to column densities up to $N_{\text{HI}} \sim 22 \text{ cm}^{-2}$, however, since the function is monotonic with little or no structure, we find that the best-fitting model is an adequate match at higher column densities as well.
A halo model for neutral hydrogen

\[ \log_{10}(N_{HI}/\text{cm}^{-2}) \]

\[ -28 \quad -26 \quad -24 \quad -22 \quad -20 \]

\[ \log_{10}(f_{HI}/\text{cm}^{-2}) \]

This work

WHISP - Zwaan+ (2005)

Braun (2012)

\[ \log_{10}(M_{HI}/(M_{\odot})) \]

\[ -4 \quad -2 \]

\[ \log_{10}(\phi(M_{HI})/\text{cm}^{-3} \text{pc}^{-1}) \]

This work

HIPASS - Zwaan+ (2005)

Martin+ (2010)

\[ dN/dX \]

This work

Rao+ (2006)

\[ \Omega_{HI} \]

This work

Rao+ (2006)

\[ \Omega_{DLA} \]

This work

Rao+ (2006)

Figure 1. The data and analysis for \( z \sim 0 \). The data sets for the column density distribution \( f_{HI} \) (top panel) include (i) WHISP (Zwaan et al. 2005b), which was used in the fitting, and (ii) the results from Braun (2012), from the measurements of Hi in local galaxies. The mass function \( \phi(M_{HI}) \) of Hi selected galaxies (second panel) is from the HIPASS survey (Zwaan et al. 2005a). The grey band indicates the overall observational error estimated by adding the systematic and statistical uncertainties in quadrature. The mass function from the ALFALFA survey (Martin et al. 2010) is also shown for comparison. The bias function of Hi-selected galaxies from ALFALFA (Martin et al. 2012) is shown in the third panel. The values of the incidence, \( dN/dX \) (fourth panel) are shown from both the WHISP and the Braun (2012) surveys.

Figure 2. The data and analysis for \( z \sim 1 \). The data points with error bars include the column density distribution \( f_{HI} \) (top panel; Rao et al. 2006). The second panel shows the product \( \Omega_{HI} b_{HI} \) from the results of intensity mapping (Switzer et al. 2013). The derived parameters: the DLA incidence \( dN/dX \) (third panel) and the DLA density parameter \( \Omega_{DLA} \) (fourth panel) are also shown.

power spectrum \( P_{HI}(k) \). There are also studies that infer the mean host halo mass of DLA systems at high redshift through clustering and cross-correlation studies (Bouché et al. 2005; Font-Ribera et al. 2012). Cross-correlations of the 21-cm intensity maps with a large sample of optical galaxies (e.g. Chang et al. 2010; Masui...
Figure 3. The data and analysis for $z \sim 2.3$. The data points include the column density distribution $f_{\text{HI}}$ (top panel) from the results of Noterdaeme et al. (2012), and the clustering of DLAs, $b_{\text{DLA}}$, from the analysis of Font-Ribera et al. (2012, second panel). Also shown are the density parameter $\Omega_{\text{DLA}}$ (third panel) and incidence $dN/dX$ (fourth panel) from Zafar et al. (2013).

Finally, we plot the model predictions for the evolution of the bias, both for HI and DLAs, and the incidence $dN/dX$ as a function of redshift. The errors on these are again determined by propagating in quadrature the uncertainties in the free parameters. These predictions are shown along with the data points in Figs. 6 and 7 respectively. The predictions for $b_{\text{HI}}$ at intermediate and high redshifts are important in the context of current and upcoming intensity mapping experiments.

We thus find that the additional input of the HI mass function at low-$z$, constrains the value of $\beta$ to less than zero, thereby making the logarithmic slope of the HI-halo mass relation lower than unity. This slope is directly related also to the high-mass cutoff in the HI-halo mass relation (the value of $v_{\text{HI}}$). A smaller slope allows a higher value of $v_{\text{HI}}$ to be consistent with the measurements, since it results in a lower weight being given to the high mass end of the HI-halo mass relation. We thus see that the introduction of the parameter $\beta$ is essential for obtaining a good fit to the low-redshift HI mass function and contributes towards constraining the overall "shape" of the HI-halo mass relation.

The parameter $\gamma$, used in the redshift evolution of the HI profile, is also found to converge to a value less than unity. This quantifies the redshift evolution of the HI concentration parameter compared to that expected from dark matter alone. If we define an "effective" concentration parameter $c_{\text{HI,eff}} = c_{\text{HI}}(1 + z)^{-\gamma}$, we thus find evidence for a non-trivial redshift evolution of the $c_{\text{HI,eff}}$, with the remaining evolution in the profile being identical to that of the dark matter. This evolution of the effective concentration of HI is shown in Fig. 8 with its error band, which indicates a mild increase of the best-fitting HI concentration with increasing redshift.

5 CONCLUSIONS

In this paper, we have carried out a detailed statistical study of the evolution of HI across redshifts by utilizing all the available data, both from the 21-cm based as well as the DLA-based observables. We build upon the results of our previous fitting to the observations (Paper 2) by generalizing the form of $\mathcal{M}_{\text{HI}}(M)$ and the HI profile from older analyses, thus introducing extra degrees of freedom. We thus explore how well we can use the data available at present to develop a halo-model based approach towards analyzing HI. The model free parameters are constrained using a full MCMC analysis so that the posterior distributions and the correlations between parameters can be effectively explored. We find that a model which is reasonably consistent with all available data underpredicts the value of DLA bias at $z \sim 2.3$, as shown in Figs. 3 and 6. This was also observed with the previous fitting analysis (Paper 2). The model predictions are nevertheless consistent with the results of DLA imaging studies, which favour the association of high-redshift DLAs with faint dwarf galaxies (Cooke et al. 2015; Fumagalli et al. 2014, 2015), and the results of hydrodynamical simulations (e.g., Rahmati & Schaye 2014; Davé et al. 2013).

The introduction of the new parameters $\beta$ and $\gamma$ constrains the shape of the HI-halo mass relation and also the form and evolution of the HI density profile. Constraining $\beta < 0$ results in a smaller weight being given to the higher range of host halo masses, thus allowing the upper cutoff $v_{\text{HI}}$ to be higher than for the case with $\beta = 0$. It is also seen from the results of abundance matching (Papastergis et al. 2013) that a flatter $\mathcal{M}_{\text{HI}} - M$ relation is favoured by the observations of ALFALFA HI galaxies. However, there are indications of a possible tension between the HI column density distribution and the mass function of HI-selected galaxies at $z \sim 0$. This tension may not be fully alleviated by the systematic effects in the measurements of the column density distribution and the mass function at $z \sim 0$. It is possible that the form of the profile at low redshifts needs to be modified in order to fit the observed column density distribution. We hope to consider other forms of the profile...
such as an exponential surface density profile, as well as a possible variation of the form of the profile across redshifts in future work.

With the introduction of the parameter $\gamma$, the redshift evolution of the concentration $c_{\text{HI}}$ can be effectively explored, and it can be seen that a mild increase in H$\text{I}$ concentration with redshift is favoured by the best-fitting model.

In future work, it would be interesting to extend the results of this halo model framework to include small-scale clustering, measured from surveys of H$\text{I}$ selected galaxies and DLA systems. This serves to provide additional constraints on the profile parameters, as well as their evolution across redshifts. This extension would also enable comparison to the results of hydrodynamical simulations, and thus shed light on the connections between H$\text{I}$ and galaxy evolution over cosmic time. Finally, such a combined halo model would have important consequences for constraining the expected signal in the H$\text{I}$ power spectrum, to be measured with current and future intensity mapping experiments.

**ACKNOWLEDGEMENTS**

We thank Tirthankar Roy Choudhury, Adam Amara, Phil Bull, Girish Kulkarni, Supratik Pal, Sebastian Seehars and Ali Rahmati for useful discussions and comments on the manuscript, and Joel Akeret for help and support with the CosmoHammer package.
Figure 5. Left panel: the best-fitting $M_{\text{HI}} - M$ relation at $z \sim 0$, along with the results of simulations in the literature (Marín et al. 2010; Davé et al. 2013). Right panel: The evolution of the best-fitting $M_{\text{HI}} - M$ relation across redshifts (plots of the relation at five redshifts are shown). A plot of the relation with logarithmic slope unity, i.e. $M_{\text{HI}} \propto M$, is also shown for comparison (dashed line).

Figure 6. The evolution of the bias, both for $\text{HI}$ as well as for DLAs, as a function of redshift in the best-fitting model. The error band indicates the range allowed by the uncertainties in the model parameters. The data include the bias of $\text{HI}$-selected galaxies at $z \sim 0$ (Martin et al. 2012) and the bias parameter of the DLAs at $z \sim 2.3$ (Font-Ribera et al. 2012).

Figure 7. The evolution of the incidence, $dN/dX$ across redshifts with the best-fitting model, with the error band indicating the range of allowed values from the model uncertainties. The data points are the results of (i) at redshift 0: Zwaan et al. (2005b) and Braun (2012), (ii) at redshifts $\sim 1$: Rao et al. (2006); (iii) at higher redshifts: Zafar et al. (2013).
Figure 8. The evolution of the "effective" concentration parameter of the H\textsc{i}, $c_{\text{HI,eff}}$, as a function of redshift. The model uncertainties are indicated by the error band. The effective evolution of the concentration parameter scales approximately as $(1 + z)^{0.5}$.

We thank the anonymous referee for a detailed and helpful report that improved the content and presentation. The research of HP is supported by the Tomalla Foundation.

REFERENCES

Akeret J., Seehars S., Amara A., Refregier A., Csillaghy A., 2013, Astronomy and Computing, 2, 27
Bagla J. S., Khandai N., Datta K. K., 2010, MNRAS, 407, 567
Barnes L. A., Haehnelt M. G., 2010, MNRAS, 403, 870
Barnes L. A., Haehnelt M. G., 2014, MNRAS, 440, 2313
Bird S., Vogelsberger M., Haehnelt M., Stajic D., Genel S., Torrey P., Springel V., Hernquist L., 2014, MNRAS, 445, 2313
Bouché N., Gardner J. P., Katz N., Weinberg D. H., Davé R., Lowenthal J. D., 2005, ApJ, 628, 89
Braun R., 2012, ApJ, 749, 87
Bull P., Ferreira P. G., Patel P., Santos M. G., 2015, ApJ, 803, 21
Chang T.-C., Pen U.-L., Peterson J. B., McDonald P., 2008, Physical Review Letters, 100, 091303
Chang T.-C., Pen U.-L., Bandura K., Peterson J. B., 2010, Nature, 466, 463
Cooke R. J., Pettini M., Jorgenson R. A., 2015, ApJ, 800, 12
Crighton N. H. M., et al., 2015, MNRAS, 452, 217
Davé R., Katz N., Oppenheimer B. D., Kollmeier J. A., Weinberg D. H., 2013, MNRAS, 434, 2645
Font-Ribera A., et al., 2012, J. Cosmology Astropart. Phys., 11, 59
Fumagalli M., Prochaska J. X., Kasen D., Dekel A., Ceverino D., Primack J. R., 2011, MNRAS, 418, 1796
Fumagalli M., O'Meara J. M., Prochaska J. X., Kanekar N., Wolfe A. M., 2014, MNRAS, 444, 1282
Fumagalli M., O'Meara J. M., Prochaska J. X., Rafelski M., Kanekar N., 2015, MNRAS, 446, 3178
Guha Sarkar T., Mitra S., Majumdar S., Choudhury T. R., 2012, MNRAS, 421, 3570
Hall A., Bonvin C., Challinor A., 2013, Phys. Rev. D, 87, 064026

Kim H.-S., Wyithe J. S. B., Baugh C. M., Lagos C. d. P., Power C., Park J., 2016, preprint, (arXiv:1603.02383)
Lanzetta K. M., Wolfe A. M., Turnshek D. A., Lu L., McMahon R. G., Hazard C., 1991, ApJS, 77, 1
Macciò A. V., Dutton A. A., van den Bosch F. C., Moore B., Potter D., Stadel J., 2007, MNRAS, 378, 55
Maller A. H., Bullock J. S., 2004, MNRAS, 355, 694
Marin F. A., Niedin N. Y., Seo H.-J., Vallinotto A., 2010, ApJ, 718, 972
Martin A. M., Papastergis E., GiovANELLI R., Haynes M. P., Springob C. M., Stierwalt S., 2010, ApJ, 723, 1359
Martin A. M., Giovanelli R., Haynes M. P., Guzzo L., 2012, ApJ, 750, 38
Masai K. W., et al., 2013, ApJ, 763, L20
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Netzer A. M., et al., 2012, A&A, 547, L1
Obreschkow D., Klockner H.-R., Heywood I., Levrier F., Rawlings S., 2009, ApJ, 703, 1890
Padmanabhan H., Choudhury T. R., Refregier A., 2015, MNRAS, 447, 3745
Padmanabhan H., Choudhury T. R., Refregier A., 2016, MNRAS, 458, 781
Papastergis E., GiovANELLI R., Haynes M. P., Rodriguez-Puebla A., Jones M. G., 2013, ApJ, 776, 43
Pontzen A., et al., 2008, MNRAS, 390, 1349
Prochaska J. X., Wolfe A. M., 2009, ApJ, 696, 1543
Rahmati A., Schaye J., 2014, MNRAS, 438, 529
Rao S. M., Turnshek D. A., Nestor D. B., 2006, ApJ, 636, 610
Santos M., Alonso D., Bull P., Silva M. B., Yahya S., 2015, Advancing Astrophysics with the Square Kilometre Array (AASKA14), p. 21
Scoccimarro R., Sheth R. K., Hui L., Jain B., 2001, ApJ, 546, 20
Sheth R. K., Tormen G., 2002, MNRAS, 329, 61
Switzer E. R., et al., 2013, MNRAS, 434, L46
Villaescusa-Navarro F., Bull P., Viel M., 2015, ApJ, 814, 146
Wang J., et al., 2014, MNRAS, 441, 2159
Zafar T., Péroux C., Popping A., Milliard B., Deharveng J.-M., Frank S., 2013, A&A, 556, A141
Zwaan M. A., Meyer M. J., Staveley-Smith L., Webster R. L., 2005a, MNRAS, 359, 1349
Zwaan M. A., et al., 2005b, MNRAS, 364, 1467
Zwaan M. A., van der Hulst J. M., Briggs F. H., Verheijen M. A. W., Ryan-Weber E. V., 2005b, MNRAS, 364, 1467

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

APPENDIX A: IMPACT OF SYSTEMATIC UNCERTAINTIES ON THE FITTED PARAMETERS

In this appendix, we provide an indicative measure of the impact of systematic uncertainties in the measurement of the column density distribution at $z \sim 0$ by repeating the analysis in the main text for the Braun (2012) data. The resulting corner plot is shown by the green curves in Fig. A1 and is overplotted on the corresponding corner plot in the main text (shown in blue). The comparison to the analysis in the main text is summarized in the first four columns of Table A1.

An indicative measure of the systematic effects can also be provided by considering the difference between two observational datasets to provide an estimate of systematic errors. We consider the effect of using the difference between the two column density distribution measurements (from the WHISP and Braun (2012)), and also the two mass function measurements (from HIPASS; Zwaan et al. (2005a) and ALFALFA; Martin et al. (2010)) as measures of the systematic uncertainties in the observables. The systematic uncertainty in each case is then added to the statistical uncertainty as measures of the systematic uncertainties in the measurement of the column density distribution between the two column density distribution measurements (from HIPASS; Zwaan et al. (2005a) and ALFALFA; Martin et al. (2010)).
Table A1. Comparison of the best-fitting values and their uncertainties obtained from (a) fitting to the $z \sim 0$ WHISP column density distribution, (b) fitting to the $z \sim 0$ Braun (2012) column density distribution, and (c) fitting to a combined dataset with systematic errors assumed to be the differences in the measurements of the mass function and the column density distribution at $z \sim 0$.

| Parameter | Best fit (WHISP) | Error (1σ; WHISP) | Best fit (Braun) | Error (1σ; Braun) | Best fit (combined) | Error (1σ; combined) |
|-----------|------------------|--------------------|------------------|-------------------|---------------------|-----------------------|
| $c_{\text{HI}}$ | 113.80          | 14.03              | 128.05           | 12.63             | 123.79              | 15.67                 |
| $\alpha$  | 0.17             | 0.02               | 0.16             | 0.01              | 0.16                | 0.02                  |
| $\log v_{c0}$ (km/s) | 1.57          | 0.03               | 1.61             | 0.03              | 1.61                | 0.05                  |
| $\log v_{c1}$ (km/s) | 4.39          | 0.75               | 4.93             | 0.68              | 4.64                | 0.72                  |
| $\beta$   | -0.55            | 0.12               | -0.66            | 0.06              | -0.63               | 0.07                  |
| $\gamma$  | 0.22             | 0.67               | 0.55             | 0.20              | 0.58                | 0.37                  |

Figure A1. Same as Fig. 4, but now repeated for the column density distribution from Braun (2012). The analysis is in green, and overplotted on the original corner plot (blue).