NEUTRAL GAS DENSITY IN DAMPED Lyα SYSTEMS

M. TRENTI AND M. STIAVELLI
Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218; Trenti@stsci.edu, mstiavel@stsci.edu

Received 2006 May 31; accepted 2006 June 28

ABSTRACT

We estimate the intrinsic neutral gas density in damped Lyα (DLA) systems ($\Omega_{\text{HI}}^{\text{DLA}}$) in the redshift range 2.2 ≤ z ≤ 5 from the SDSS DR3 DLA sample of optically selected quasars. We take into account self-consistently the obscuration on background quasars due to the dust present in damped Lyα systems. We model the column density and redshift distribution of these systems by using both a nonparametric and a parametric approach. Under conservative assumptions on the dust content of DLA systems, we show that selection effects lead to underestimating the intrinsic neutral gas density by at least 15% with respect to the observed neutral gas density. Over the redshift range 2.2–5.5, we find $\Omega_{\text{HI}}^{\text{DLA}} = 0.97\pm0.08\pm0.28 \times 10^{-3}$, where the first set of error bars gives the 1 σ random errors and the second set gives the modeling uncertainty dependent on the fraction of metals in dust, from 0% to 50%. This value compares with $\Omega_{\text{HI}}^{\text{DLA}} = 0.82\pm0.05 \times 10^{-3}$ (1 σ error bars), which is obtained when no correction for dust is introduced. In the model with half of the metals mass in dust, we cannot constraint $\Omega_{\text{HI}}^{\text{DLA}}$ at a confidence level higher than 90%. In this case there is indeed a probability of about 10% that the intrinsic column density distribution of DLA systems is a power law $f(N_{\text{HI}}) \propto 1/N_{\text{HI}}^{1.5}$. In contrast, with 25% of the metals in dust, the most realistic estimate, a power law is ruled out at a 99.5% confidence level.

Subject headings: dust, extinction — galaxies: high-redshift — galaxies: ISM — intergalactic medium

Online material: color figures

1. INTRODUCTION

Damped Lyα systems (hereafter DLA systems) are quasar absorption systems with a column density above $2 \times 10^{20} \text{cm}^{-2}$ and represent the high end of the distribution of absorption systems starting from the Lyα forest at $N_{\text{HI}} \gtrsim 10^{14} \text{cm}^{-2}$. DLA systems represent the most significant reservoir of neutral hydrogen available for star formation in the universe. These systems are considered to be either (cold) massive rotating disks, the progenitors of today’s disk galaxies (Prochaska & Wolfe 1997; Wolfe et al. 2005), or compact protogalactic clumps (Haehnelt et al. 1998; Nagamine et al. 2004).

In the era of precision cosmology, an accurate measure of the total mass density of neutral gas as a function of the redshift represents an important constraint for galaxy formation models. Ground observations are able to identify DLA absorption features in the spectra of quasars from $z_{\text{abs}} \gtrsim 1.8$, where the absorption lines enter the atmospheric window, up to $z_{\text{abs}} \approx 5.5$. With an all-sky survey like the Sloan Digital Sky Survey, spectra of several thousands of quasars with enough resolution for DLA detection have been acquired and the observed density of neutral gas in DLA systems is now measured with errors below 10% (Prochaska et al. 2005).

This measurement must be interpreted with some caution, as the presence of DLA systems along a line of sight leads to a potential obscuration due to the dust that they host; the observed gas density is a biased estimator of the intrinsic density unless the dust effects are accurately quantified. Several papers, starting from Ostriker & Heisler (1984) have attempted to model the influence of dust along the line of sight, often with conflicting results.

A detailed analysis framework for the obscuration of quasars has been developed by Fall & Pei (1993, 1989; see also Fall et al. 1989; Pei et al. 1991; Pei & Fall 1995) and applied to the quasars sample of Lanzetta et al. (1991). Their study highlighted a potentially severe effect of the dust bias that prevented them from setting an upper limit on the intrinsic density of DLA systems. In fact, absorbers with high column densities and/or with high dust-to-gas ratio represent essentially “bricks” along the line of sight to a quasar and are very likely to be missed in optically selected surveys. Additional evidence for the dust obscuration came in the form of a detected preferential reddening in the spectra of quasars with DLA absorption with respect to a control sample without detection of these systems (Pei et al. 1991).

More recent investigations revised these earlier results on the dust content in DLA systems and on their reddening of background objects (Murphy & Liske 2004; Ellison et al. 2005), finding in particular no robust evidence for the reddening of $z \approx 3$ quasars with DLA features in their spectra; at 3 σ Murphy & Liske (2004) find $E(B-V) < 0.02$ mag, while Ellison et al. (2005) have $E(B-V) < 0.04$ mag. At the same time, radio-selected quasar surveys (Ellison et al. 2001), with complete optical follow-up detection, have provided the first bias-free constraints on the intrinsic distribution of DLA systems.

Taking advantage of these recent measurements for the number density of DLA systems, we have previously characterized (Trenti & Stiavelli 2006) the dust absorption along random lines of sight by means of a Monte Carlo code, finding that, on average, the deviations from unit transmission are effectively modest \([\exp(-\tau)] \gtrsim 0.99\) at an emitted wavelength $\lambda_e = 0.14\mu\text{m}$ over all the redshift range] and of limited impact on most observations. However, this result does not preclude the presence of a small fraction of lines of sight (of the order of a few percent) through the most massive and/or the most metal-rich DLA systems that might have a large optical depth. Indeed, Wild & Hewett (2005) and Wild et al. (2006) find significant evidence of reddening in DLA systems with Ca ii absorption lines at moderate redshift ($z_{\text{abs}} \approx 1$; $E(B-V) \gtrsim 0.1$). Similarly, York et al. (2006) measure $E(B-V)$ up to 0.085 for Mg ii selected DLA systems at $z_{\text{abs}} \approx 1.0$. As the determination of the gas density in DLA systems is dominated by these most massive absorbers, the potential bias in this measure, correctly stressed by Fall & Pei...
(1993), must not be dismissed by the recent evidence of a very modest average deviation from unity transmission.

In this paper, we take advantage of the large sample of DLA systems identified in Sloan quasars (Prochaska et al. 2005), and we investigate the relation between the observed and intrinsic density of neutral gas in these systems. The Sloan sample that we consider has three main advantages over the sample used by Fall & Pei (1993): (1) it is larger by a factor 30; (2) the quasars have been selected within a luminosity limit in the $I$ band, significantly less sensitive to dust obscuration than the $B$ band used for the older sample; and (3) the color selection algorithm has a good sensitivity to red quasars (Richards et al. 2002). In fact, even if a quasar with a DLA absorption system is not dropped below the $I$-band flux limit, its color can be changed so that it ends up lying outside the color selection box used to identify quasar candidates for follow-up spectroscopy. Thanks to the precision of the Sloan photometry, which allows a clear separation of the stellar locus in the color space, and to the use of an extended color-selection box, the loss of completeness due to this latter effect is only marginal (see Richards et al. 2002 for a detailed discussion of the Sloan color selection algorithm and for completeness tests) and is therefore not considered in this work.

This paper is organized as follows: in § 2 we characterize the obscuration bias in a magnitude-limited survey; in § 3 we describe the data set that we are using; in § 4 we present our analysis for the parametric estimation of the intrinsic comoving density of neutral gas, whose uncertainties are quantified in §§ 5 and 6 by means of Monte Carlo simulations of synthetic observations; in § 7 we discuss the accuracy of a nonparametric estimator for the neutral gas density; and we summarize our findings in § 8.

2. DUST EFFECTS IN MAGNITUDE-LIMITED SURVEYS

In a magnitude-limited survey, obscuration along the line of sight leads to the potential loss of some lines of sight. Following Fall & Pei (1993), to quantify the effect let us consider an intrinsic quasar luminosity function $\Phi_i(L)$ in the interval $[L_{\text{min}}, +\infty]$, where $L_{\text{min}}$ is the luminosity above which the sample is complete. The intrinsic number of objects above the completeness threshold is

$$N_i = \int_{L_{\text{min}}}^{+\infty} \Phi_i(L) \, dL.$$  \hspace{1cm} (1)

The presence of dust introduces an optical depth $\tau$ along the line of sight, with a corresponding transmission coefficient $\exp(-\tau)$. Under this condition, the number of objects with observed luminosity above $L_{\text{min}}$ is

$$N_o = \int_{L_{\text{min}}}^{+\infty} \Phi_i(L) \, dL \cdot \exp(-\tau).$$  \hspace{1cm} (2)

The dust obscuration leads to a fraction of missing objects given by $N_o/N_i$:

$$\frac{N_o}{N_i} = \frac{\int_{L_{\text{min}}}^{+\infty} \Phi_i(L) \, dL \cdot \exp(-\tau)}{\int_{L_{\text{min}}}^{+\infty} \Phi_i(L) \, dL}.$$  \hspace{1cm} (3)

In general, the precise value of $N_o/N_i$ for a given $\tau$ depends on the sensitivity limit of the survey, $L_{\text{min}}$, and on the form of the luminosity function $\Phi_i(L)$. Under the assumption that the luminosity function is a power law $\Phi_i(L) \propto L^{-\beta - 1}$ (in the luminosity range from $L_{\text{min}}$ to $+\infty$), the ratio $N_o/N_i$ can be easily computed and is independent of $L_{\text{min}}$:

$$\frac{N_o}{N_i} = \left( \frac{L_{\text{min}}}{L_{\text{min}} + \tau} \right)^{\beta \tau} = e^{-\beta \tau}.$$  \hspace{1cm} (4)

This was already noted by Fall & Pei (1993).

The luminosity function for quasars is modeled in terms of a gamma function and/or of a double power law (e.g., see Pei 1995). However the SDSS spectroscopic quasar sample is shallower, for $z_{\text{phot}} \gtrsim 2.2$, than the knee of the distribution so that the luminosity function for the quasars that we are considering in this paper can be effectively treated, up to $L_{\text{min}}$, as a single power law (see Richards et al. 2006), simplifying significantly our analysis by use of equation (4).

In the following analysis, we assume a redshift-dependent luminosity function (Richards et al. 2006),

$$\Phi_i(L_i, z) = A(z) L_i^{-\beta(z)-1},$$  \hspace{1cm} (5)

where $L_i$ is the luminosity in the $I$ band\(^1\) (where the sensitivity limit for detection of quasars in SDSS is given) and $\beta(z)$ is a slowly evolving function of the redshift with $\beta$ varying from 2.2 to 1.1 in the redshift range 2–5 (Richards et al. 2006):

$$\beta(z) = 2.1 - 0.275(z - 2.45)\theta(z - 2.45),$$  \hspace{1cm} (6)

where $\theta(x)$ is the step function defined as $\theta(x) = 1$ if $x \geq 0$ and $\theta(x) = 0$ otherwise. The SDSS DR3 DLA sample of quasars considered in this paper has an average redshift of 2.97 with a standard deviation of 0.65. The average slope can be approximated as

$$\langle \beta \rangle = 1.95 \approx 2.$$  \hspace{1cm} (7)

In case of absorption due to a dusty DLA system at redshift $z_a$ with column density $N_{\text{HI}}$, the optical depth at an observed wavelength $\lambda_o$ can be written as

$$\tau(\lambda_o, z_a, N_{\text{HI}}, k) = kN_{\text{HI}} \xi(\lambda_o/(1 + z_a)),$$  \hspace{1cm} (8)

where $k$ is the dust-to-gas ratio and $\xi$ is the relative extinction curve normalized to the absorption cross section in the $B$ band $\sigma(\lambda_B)$ (e.g., see Pei 1992):

$$\xi(\lambda) = \sigma(\lambda)/\sigma(\lambda_B).$$  \hspace{1cm} (9)

If $N_{\text{HI}}$ is expressed in units of $10^{21}$ cm$^{-2}$, the absorption cross section in the $B$ band is $kN_{\text{HI}}$, with $k = 0.8$ for Galactic dust. The value of $k$ for DLA systems depends on their metallicity $Z$ and on the fraction of metals in dust, i.e., on the dust-to-metals ratio. DLA systems are generally characterized by a low metallicity and by a moderate evolution of their properties with the redshift (Wolfe et al. 2005). We approximate the observed average (H i column density weighted) redshift-metallicity relation (based on a number of observations, e.g., Prochaska et al. 2003, Dessauges-Zavadsky et al. 2004, 2006, Akerman et al. 2005) with a linear function for $\log[Z_{\text{HI}}(z)]$. This provides a good agreement with the data (e.g., see Fig. 13 in the compilation by Kulkarni et al. 2005) in the redshift range $2 \leq z \lesssim 5$:

$$Z_{\text{HI}}(z)/Z_{\odot} = 0.2 \times 10^{-0.2z}.$$  \hspace{1cm} (10)

\(^1\) In this paper we assume an average wavelength for the band of $\langle \lambda_B \rangle = 0.8$ μm.
The average metallicity in equation (10) translates, for a Milky Way dust-to-metals ratio (i.e., 50% of the metals in dust grains), into a Milky Way average dust-to-gas ratio of

\[ k_{MW}(z) = 0.16 \times 10^{-0.2z}. \] (11)

To account for different fractions of metals in the form of dust grains, we introduce a correction factor \( \alpha_c \) for the intrinsic dust-to-gas ratio \( k_i(z) \),

\[ k_i(z) = \alpha_c k_{MW}(z). \] (12)

DLA systems are considered to have a smaller fraction of metals in dust than our Galaxy (with roughly one quarter of the total metals content in dust), so realistic values for \( \alpha_c \) are expected to be around \( \alpha_c = 0.5 \) (Pettini et al. 1997; Vladilo 2002; but see Pei et al. 1999, where the Milky Way dust-to-metals ratio is used). In the following sections we present our analysis using a range of dust-to-metals ratios, highlighting the dependence of the dust bias on this quantity. We assume a reference value \( \alpha_c = 0.5 \) (25% of metals in dust, the value derived by Pettini et al. 1997), but we include a grid of models with \( \alpha_c \in [0, 1] \) and with \( \Delta \alpha_c = 0.125 \), discussing in depth also the case \( \alpha_c = 1 \) (50% of metals in dust).

Our analysis is carried out by assuming that all DLA at a given redshift have a fixed metallicity, as well as a fixed dust-to-metals ratio. A detailed modeling of the metallicity and dust-to-metals ratio distributions for the purpose of determining \( \Omega_{DLA}^{(\text{H}_1)} \) is extremely challenging, as only a small subsample of DLA systems have measured metallicities. Fortunately, the value of \( \Omega_{DLA}^{(\text{H}_1)} \) is not expected to significantly depend on the scatter in the dust-to-gas ratio \( k_i \) of the absorbers (see Fall & Pei 1993, Appendix A). To ensure that this is the case, in §6 we validate our scatterless approach by analyzing synthetic observations with a variety of dust-to-gas distributions.

3. THE DATA: THE SDSS DR3 DLA AND THE CORALS SURVEYS

The data used in this paper are primarily taken from the DLA survey by Prochaska et al. (2005). We consider all the quasars from their Table 1 and all the DLA systems reported in their Table 3. The sample consists of 525 DLA systems found in 4568 spectra of quasars with a minimum signal-to-noise ratio of 4.

In addition, we consider the DLA systems detected in the radio-selected CORALS survey (Ellison et al. 2001), intrinsically free from dust bias. This survey consists of 19 DLA systems detected in 66 spectra of quasars (see their Table 3). The statistical sample of the survey (see Ellison et al. 2001) that we consider for the analysis is restricted to 17 DLA in the redshift range 1.8–3.5.

4. PARAMETRIC ESTIMATION OF \( \Omega_{DLA}^{(\text{H}_1)} \) FROM SDSS DR3

Following the standard practice (e.g., Lanzetta et al. 1991), we define the number of observed DLA systems in the intervals \([N_{\text{H}_1}, N_{\text{H}_1} + dN_{\text{H}_1}]\) and \([X, X + dX] \) as

\[ f_0(N_{\text{H}_1}, X) dN_{\text{H}_1}, dX, \] (13)

where \( f_0 \) is the observed frequency distribution and \( dX \) the absorption distance:

\[ dX = \frac{H_0}{H(z)} (1 + z)^2 dz. \] (14)

In this paper, we adopt the standard \( \Lambda \)CDM cosmology with \( H_0 = 70 \) km s\(^{-1}\) Mpc\(^{-1}\), \( \Omega_{\Lambda} = 0.7 \), and \( \Omega_M = 0.3 \), so that

\[ dX = \frac{(1 + z)^2}{\sqrt{0.7 + 0.3(1 + z)^2}} dz. \] (15)

The data from SDSS DR3 DLA span over a total integrated absorption path length \( \Delta X \approx 7333 \) Mpc.

Considering the obscuration bias discussed in §2 for the Sloan survey (see also Fall & Pei 1993) and assuming that the distribution of absorbers along the line of sight is not correlated,\(^2\) we can write the relation between the intrinsic and the observed frequency distribution of DLA systems as

\[ f_0(N_{\text{H}_1}, z) = f_0(N_{\text{H}_1}, X(z)) \exp(-[\beta_0 \Delta \alpha_N(z) N_{\text{H}_1} f(z)/(1+z)], \] (16)

where we assume the average slope \( \langle \beta \rangle = 2 \) for the quasar luminosity function. Following Prochaska et al. (2005), we parameterize \( f_0 \) using a gamma function

\[ f_0(N_{\text{H}_1}, X) = \eta_1 \frac{N_{\text{H}_1}^{m_1}}{\Gamma(m_1)} \exp\left(-\frac{N_{\text{H}_1}}{\eta_1}\right), \] (17)

and a power law

\[ f_0(N_{\text{H}_1}, X) = \eta_2 (N_{\text{H}_1})^{m_2}. \] (18)

The use of the gamma function for the intrinsic distribution \( f_0 \) has the advantage that, given equation (16), the observed distribution \( f_0 \) also remains a gamma function. In contrast, an intrinsic power law is mapped into an observed gamma function by the effect of dust.

If we assume that the intrinsic distribution of DLA systems follows Poisson statistics, the observed distribution will also follow Poisson statistics. The likelihood function to be maximized is, therefore (see also eq. [27] in Fall & Pei 1993),

\[ \mathcal{L} = \exp \left[ -\sum_{n} \int_{z_{\text{inf}}}^{z_{\text{sup}}} dN_{\text{H}_1} f_0(N_{\text{H}_1}, z) \frac{dX}{dz} \right] \prod_{n} f_0(N_{\text{H}_1}, z_n), \] (19)

where the sum extends over all the samples of quasars that have been searched for DLA systems in the redshift interval \([z_{\text{inf}}, z_{\text{sup}}]\), the signal-to-noise ratio is above 4, and the product is over all the DLA systems detected in those quasars. In equation (19) the observed distribution function \( f_0 \) is expressed in terms of the intrinsic distribution \( f_0 \) through equation (16). \( N_{\text{DLA}} = 2 \times 10^{20} \) cm\(^{-2}\) is the assumed lower limit for the column density of a DLA system. Before presenting the results from the maximum likelihood analysis, we investigate the properties of \( f_0 \) with the aim to discuss upper limits on the dust-to-gas ratio and on \( \Omega_{DLA}^{(\text{H}_1)} \).

4.1. Basic Considerations from the Properties of \( f_0 \)

Figure 1 shows the observed column density distribution of DLA systems, averaged over redshift. The observed data are well fitted by a gamma function (Prochaska et al. 2005, see also eq. [17] and Table 1) with a knee at \( (N_{\text{H}_1})_0 = 3 \times 10^{21} \) cm\(^{-2}\) and a slope \( m_1 = -1.8 \).

The origin of the observed knee at \( (N_{\text{H}_1})_0 \) can in principle be either due to an intrinsic decrease of the distribution at high

\(^2\) We are essentially neglecting clustering along the line of sight.
column densities, i.e., due to \(N_o\), or due to the effect of obscuration bias, which introduces in the data an exponential-like decrease. We can obtain a limit on the dust-to-gas (expressed in terms of \(\alpha_c\)) ratio by assuming that the observed knee is entirely due to dust bias. From equation (16) we can write

\[
\langle \beta \rangle \alpha_c \kappa_{\text{MW}}(z) \xi \left[ \lambda_i/(1 + z) \right] \leq 1/(N_o),
\]
so that, considering that the average redshift of DLA absorbers in our sample is \(\langle z \rangle = 3.1\), we estimate a limit on \(\alpha_c\) to be

\[
\alpha_c \leq \frac{1}{(N_o)} \frac{1}{\langle \beta \rangle \kappa_{\text{MW}}(z) \xi \left[ \lambda_i/(1 + z) \right]} \leq \frac{1}{3(2)(0.038)(2.47)} \leq 1.75.
\]

### Table 1

| \(\alpha_c\) | \(\eta_1\) | \(m_1\) | \(N_o\) | \(\log(\mathcal{L})\) | \(\Omega_{H_1}^{(\text{DLA})}\) |
|----------------|---------|---------|-----------|----------------|------------------|
| 0.000           | 3.03 \times 10^{-3} | -1.80   | 3.00  | 0.00 | 0.82 \pm 0.05 \times 10^{-3} |
| 0.125           | 2.61 \times 10^{-3} | -1.80   | 3.23  | 0.16 | 0.85 \pm 0.05 \times 10^{-3} |
| 0.250           | 2.49 \times 10^{-3} | -1.78   | 3.40  | 0.32 | 0.89 \pm 0.05 \times 10^{-3} |
| 0.375           | 2.21 \times 10^{-3} | -1.78   | 3.61  | 0.46 | 0.91 \pm 0.05 \times 10^{-3} |
| 0.500           | 1.70 \times 10^{-3} | -1.79   | 4.15  | 0.66 | 0.97 \pm 0.05 \times 10^{-3} |
| 0.625           | 1.47 \times 10^{-3} | -1.79   | 4.51  | 0.83 | 1.01 \pm 0.05 \times 10^{-3} |
| 0.750           | 1.10 \times 10^{-3} | -1.79   | 5.27  | 1.00 | 1.05 \pm 0.05 \times 10^{-3} |
| 0.875           | 1.02 \times 10^{-3} | -1.78   | 5.60  | 1.18 | 1.13 \pm 0.05 \times 10^{-3} |
| 1.000           | 6.24 \times 10^{-4} | -1.80   | 7.22  | 1.33 | 1.25 \pm 0.05 \times 10^{-3} |

**Notes.** Summary table with the maximum likelihood parameters for the intrinsic distribution of DLA systems fitted with a gamma function and using different values for \(\alpha_c\). The parameters \(\eta_1\) and \(N_o\) are given in units of \(10^{21}\) cm\(^{-2}\) for \(N_{\text{DLA}}\). The likelihood \(\mathcal{L}\) is normalized to the value of the best-fitting model (gamma function) with no dust.

As this value implies that about 90% of metals are in dust, which appears extremely improbable for DLA systems, the presence of an observed knee in the column density distribution of the gas is likely an intrinsic feature and not induced by the dust bias alone. This implies that the expected difference between the intrinsic and the observed density of neutral gas in DLA systems is limited for the most realistic fraction of metals in dust (25% i.e., \(\alpha_c = 0.5\)). But in the case of a larger fraction of metals in dust (i.e., \(\alpha_c \approx 1\)) there can be the possibility of a significant dust bias, as, in that case, an acceptable model of the intrinsic distribution can still be obtained in terms of a power law. In § 4.2, we quantify more precisely the dust bias by studying the likelihood \(\mathcal{L}\) for \(f_i\) at increasing \(\Omega_{H_1}^{(\text{DLA})}\) for different values of \(\alpha_c\).

### 4.2. Whole Sample Analysis

By aggregating all the data in SDSS DR3 DLA, the results of our maximum likelihood analysis are reported in Tables 1 and 2 and in Figure 2. The results for the fit with an intrinsic gamma function are reported in terms of the comoving density of gas in DLA systems (\(\Omega_{H_1}^{(\text{DLA})}\)), which is the first moment of \(f_i\):

\[
\Omega_{H_1}^{(\text{DLA})} dX = \frac{\mu m_H \rho_c}{c^2} \int_{N_{\text{DLA}}}^{\infty} dN_{\text{H}_1} f(N_{\text{H}_1}, z) dX,
\]

where \(m_H\) is the mass of the hydrogen atom, \(\mu = 1.3\) is a correction factor for the composition of the gas, \(c\) is the speed of light, \(\rho_c\) is the critical density, and \(N_{\text{DLA}} = 2 \times 10^{20}\) cm\(^{-2}\).

We obtain the results:

\(\alpha_c = 0\).—If we neglect the effect of dust absorption, so that \(f_i = f_{i,0}\), we rederive the parameters for the gamma distribution reported in Prochaska et al. (2005; see our Table 1). All our likelihood values for the Sloan data are presented in log units normalized to the likelihood value for the best-fitting dust-free model. For \(\alpha_c = 0\), the likelihood is sharply peaked around the best value for \(\Omega_{H_1}^{(\text{DLA})}\). However, when we introduce the effect of the dust (\(\alpha_c > 0\)), there is an increasingly strong degeneration toward high values for \(\Omega_{H_1}^{(\text{DLA})}\) (see Fig. 2).

\(\alpha_c = 0.5\).—For our standard scenario, the likelihood has a maximum at \(\Omega_{H_1}^{(\text{DLA})} = 0.97 \times 10^{-3}\), to be compared with the value \((\Omega_{H_1}^{(\text{DLA})})_0 = 0.82 \times 10^{-3}\) given by the maximum likelihood analysis with \(\alpha_c = 0\) (no dust). In our modeling of the column
density distribution the metallicity is introduced as an external input, based on independent observations. An additional reason to motivate the need for taking into account the obscuration bias comes from the fact that the likelihood for \( \alpha_e = 0.5 \) is higher than for \( \alpha_e = 0 \).

\( \alpha_e = 1 \).—The bias induced by the dust increases as larger \( \alpha_e \) are considered. As expected qualitatively by the simple considerations on the shape of \( f_m \) presented in § 4.1, the maximum likelihood analysis for our large dust-to-metal ratio scenario (\( \alpha_e = 1 \)) is unable to set a strong upper limit to \( \Omega_i^{\text{DLA}} \) as the likelihood function remains high for large neutral gas densities. At \( \alpha_e = 1 \), a power law \( f_i \) with slope \( m_i = -1.95 \) has a likelihood ratio over the best gamma solution of \( R = 2 \log (L_0/L_{\gamma}) = 2.1 \). A likelihood ratio test with this value implies that a power-law solution can be ruled out only at about an 85% confidence level if we assume that \( R \) is distributed as a \( \chi^2 \) with one degree of freedom. However this is in general valid only when the model is linear in the parameters and the errors follow a Gaussian distribution (Lupton 1993). As our model is strongly nonlinear, in § 5 we estimate the confidence level by means of Monte Carlo simulations.

\( \alpha_e > 1 \).—If we consider fits for \( f_i \) at \( \alpha_e > 0 \) (see Table 2), we obtain an increasingly better modeling in terms of power laws up to \( \alpha_e \approx 2 \). As expected from the preliminary examination of \( f_m \), eventually the likelihood decreases at higher dust-to-gas ratio starting from \( \alpha_e \approx 2 \). We recall that \( \alpha_e > 2 \) is formally unphysical, as it implies more than 100% of the metals are in dust; however, due to the fact that we have adopted a simple fit to the observed metallicity, a high \( \alpha_e \) value also means that the intrinsic metallicity content of DLA system is underestimated in our model. As it appears unlikely that the metallicity measurements do significantly underestimate the dust content of DLA systems, we expect that realistic values of \( \alpha_e \leq 1 \).

**CORALS.**—For comparison we have repeated the maximum likelihood analysis on the radio-selected CORALS quasars (Ellison et al. 2001). The results are reported in Table 3 and in Figure 3. For these data the measure is guaranteed to be free from dust bias, but the sample size is insufficient to accurately constrain the parameters of \( f_i \) (the likelihood curve in Fig. 3 is relatively flat around the maximum). The maximum is at \( \Omega_i^{\text{DLA}} = 1.2 \times 10^{-3} \) with extremely large 1 \( \sigma \) uncertainties (of the order of 50%) due to the flatness of the likelihood function. In addition, the maximum likelihood for the gamma function model is only negligibly better than the one for a power law (\( R \approx 0.1 \)).

### 5. ERRORS ESTIMATE

To estimate the error on the determination of \( \Omega_i^{\text{DLA}} \), we have performed a bootstrapping analysis of the data. We have randomly extracted samples of lines of sight with uniform probability from the SDSS DLA survey. The samples have the same number of lines as the original data. For each of these simulated samples we have performed a maximum likelihood analysis to determine \( \Omega_i^{\text{DLA}} \). In the case of \( \alpha_e = 1 \), the results obtained with 400 random realizations are presented in Figure 4 and show that \( \Omega_i^{\text{DLA}} \) is indeed distributed around the best value derived from the original data (\( 1.25 \times 10^{-3} \)), which coincides with the median of the distribution. The average value of the distribution is 1.27 \( \times 10^{-3} \). At a 68% confidence level, the measurements lie in the interval 1.10–1.41 \( \times 10^{-3} \), which coincides with the median of the distribution. The average value of the distribution is 1.27 \( \times 10^{-3} \). At a 90% confidence level \( \Omega_i^{\text{DLA}} \leq 1.48 \times 10^{-3} \). The bootstrapping procedure has been repeated for all the scenarios with different \( \alpha_e \), resulting in at least 100 simulations for each value of the dust-to-metals ratio considered. The results are reported as standard 1 \( \sigma \) error in Table 1. The case \( \alpha_e = 0.5 \) is shown in Figure 5 (using 200 simulations in this case).

As a additional test to investigate the uncertainties in our analysis, we have simulated a number of synthetic samples of data using the Monte Carlo code developed in Trenti & Stiavelli (2006). These samples have then been processed using the same method applied to the real observed data, to test its efficiency in retrieving the input parameters and to quantify the typical error.

Each random realization of a simulated observation is characterized by the same observed path length of the SDSS sample (\( \Delta X = 7333.1 \)) but consists of lines of sight with DLA distribution (and detection) in the redshift range 2.2–4. The optical depth along each intrinsic line of sight is computed from the realized DLA distribution and then a test on the obscuration probability (eq. [4]) is performed to accept or reject the intrinsic

---

**TABLE 3**

*Maximum Likelihood Parameters for the CORALS Survey*

| \( \eta \) | \( m \) | \( N_i \) | \( \log \mathcal{L} \) | \( \Omega_i^{\text{DLA}} \) |
| --- | --- | --- | --- | --- |
| 6.33 \( \times 10^{-6} \) | -2.13 | 39.09 | 0 | 1.2 \( \times 10^{-3} \) |

**Notes:** Summary table with the maximum likelihood parameters for the CORALS survey modeling. The likelihood value is normalized to the value given by the best-fitting CORALS model. Other units are like in Table 1.
line of sight as an observed one. The only source of uncertainty in these synthetic observations comes from the discrete sampling of \( f_i \) and from the obscuration probability test. The model does not include observational errors.

We have simulated two models: (1) an intrinsic gamma function and (2) an intrinsic power law, with the parameters given by the maximum likelihood analysis from the SDSS data (see Table 1 for the gamma function and Table 2 for the power law).

Figure 4 shows, for \( \alpha_e = 1 \), the distribution of \( \Omega_{\text{HI}}^{(\text{DLA})} \) as measured from the simulated observations (80 different realizations) of an intrinsic gamma function. It is reassuring that the distribution basically agrees with the one obtained using the bootstrap method. The mean value for \( \Omega_{\text{HI}}^{(\text{DLA})} \) is \( 1.29 \times 10^{-3} \), with a standard deviation of \( 0.23 \times 10^{-3} \). As we used an input value \( f_i^{(\text{DLA})} = 1.25 \times 10^{-3} \) to generate the synthetic observations,
the fitting procedure is able to recover the intrinsic density of neutral gas introducing only a negligible bias, which is much smaller than the $1\sigma$ random uncertainty.

The analysis of the simulated observations generated from an intrinsic power-law distribution allows us to quantify the probability that the identification of a gamma function as the best-fitting model for the Sloan data is only due to the combined effect of discrete sampling and of the additional degree of freedom of the gamma function over a power law. For our reference scenario ($\alpha_e = 0.5$), we have generated 400 random realizations of observations starting from an intrinsic power law with unconstrained neutral gas density ($\eta_2 = 1.54 \times 10^{-2}$, $m_2 = -2.05$). The likelihood ratio (gamma over power law) distribution shows only two realizations that have $R$ greater than the one measured from the SDSS DLA data. This means that an intrinsic power law can be ruled out at a 99.5% confidence level. Repeating the experiment in the case of $\alpha_e = 1$, $\eta_2 = 1.77 \times 10^{-2}$, and $m_2 = -1.95$, we find that about 8% of the realizations have a value of $R$ greater than the one measured from the SDSS DLA data. This means that the confidence level at which a power-law distribution (with unconstrained upper cut off) can be ruled out is, in this case, 92%.

## 6. SYNTHETIC OBSERVATIONS WITH DUST-TO-GAS RATIO SCATTER

In the maximum likelihood applied in this work, we assume that all the absorbers have a fixed dust-to-metals ratio $\alpha_e$ and a metallicity dependent only on the redshift. In reality, a significant scatter around the mean values is expected for both these quantities and has indeed been observed (e.g., see Kulkarni et al. 2005). Fall & Pei (1993) have already noted that the scatter in the dust-to-gas ratio is not introducing significant bias, assuming that there is no correlation between $k_i$ and the column density $N_{HI}$.

To confirm the results of Fall & Pei (1993) and to extend the analysis of the bias to the case where a correlation between $k_i$ and $N_{HI}$ is present, we have simulated a number of synthetic samples of data adopting an approach similar to the one described in §5. In this case, we compute the optical depth of each absorber by drawing the dust-to-gas ratio $k_i$ from a given distribution, rather than adopting a fixed value depending only on $z$.

When there is no correlation between $k_i$ and $N_{HI}$, we use (based on eq. 12)

$$k_i(z) = \xi\alpha_e k_{MW}(z),$$

where $\xi$ is a random variable with a given probability distribution $p(\xi)$ such that the expected value of $\xi$ under $p$ is $E_p(\xi) = 1$. We have considered the following form for $p(\xi)$:

1. A uniform distribution around the average value: $p(\xi) = 1$ if $\xi \in [0.5, 1.5]$ and $p(\xi) = 0$ otherwise.

2. A lognormal distribution with variance parameter $\sigma_\xi$, as studied by Fall & Pei (1993): $p(\xi) = \exp\left(-\log^2[\xi \exp(\sigma^2/2)] / (2\sigma^2)\right) / (2\pi\sigma^2)^{1/2}$. We adopt $\sigma_\xi = 0.5, 1,$ and $1.5$, inspired by the dispersion in the dust-to-gas ratio observed for galaxies in the local universe (Pei 1992).

In addition, we have generated synthetic observations with a correlation between $k_i$ and $z_i$, adopting the prescription

$$k_i(z) = \xi N_{HI} dN_{DLA} f(N_{HI}) f(N_{HI}) \times \alpha_e k_{MW}(z),$$

where $\gamma = -0.8, -0.4, 0.4,$ and $-0.8$ (i.e., we explore both an increasing and decreasing trend of $k_i$ vs. $N_{HI}$) and $\xi$ is a random variable with the same properties described above. The normalization adopted is such that the $HI$ column density averaged value of $k_i$ is $\alpha_e k_{MW}(z)$.

For each combination of the parameters considered (see Table 4), we have generated 60 synthetic SDSS-like samples of data (as in §5), adopting a gamma-function model for $f(N_{HI})$ ($N_{HI} = 4.15 \times 10^{21}$ cm$^{-2}$, $m_1 = -1.79$, $\Delta N_{DLA} = 0.97 \times 10^{-5}$). The first series of models employs an uncorrelated dispersion in $k_i$ (uniform and lognormal, see §6), the second series also assumes a correlation between $k_i$ and $N_{HI}$. For each model we have analyzed 60 random realizations to recover $\Omega_{DLA}$, reporting the mean value obtained from the maximum likelihood analysis with average $k_i(z)$ (second column), the standard deviation (third column), and the bias, i.e., the average value minus the nominal model value (fourth column).

### Table 4

| Model | $10^3 \Omega_{DLA}$ | $10^3 \sigma_{\Omega_{DLA}}$ | $10^3 \Delta \Omega_{DLA}$ |
|-------|---------------------|-----------------------------|----------------------------|
| Uniform | 1.00 | 0.11 | 0.03 |
| Lognormal | $\sigma_\xi = 0.5$ | 0.97 | 0.08 | 0.00 |
| Lognormal | $\sigma_\xi = 1$ | 0.99 | 0.09 | 0.02 |
| Lognormal | $\sigma_\xi = 1.5$ | 1.03 | 0.10 | 0.05 |

**Notes.**—Summary of the synthetic observations with a scatter in the dust-to-gas ratio $k_i$. The underlying column density distribution is the best-fitting SDSS gamma function when $\alpha_e = 0.5$ ($N_{HI} = 4.15 \times 10^{21}$ cm$^{-2}$, $m_1 = -1.79$, $\Delta N_{DLA} = 0.97 \times 10^{-5}$). The first series of models employs an uncorrelated dispersion in $k_i$ (uniform and lognormal, see §6), the second series also assumes a correlation between $k_i$ and $N_{HI}$. For each model we have analyzed 60 random realizations to recover $\Omega_{DLA}$, reporting the mean value obtained from the maximum likelihood analysis with average $k_i(z)$ (second column), the standard deviation (third column), and the bias, i.e., the average value minus the nominal model value (fourth column).

Uncorrelated scatter in $k_i$.—The maximum likelihood analysis is able to correctly recover the value of $\Omega_{DLA}$ used to generate the observations (see Fig. 6 and Table 4). There is only a small bias (at the percent level) toward obtaining output values marginally larger than the input, even when the dust obscuration is such that the observed column density averaged dust-to-gas ratio is smaller by 15% with respect to the intrinsic value, as happens in our lognormal model with $\sigma_\xi = 1$. This demonstrates that our analysis in terms of the average dust-to-gas ratio is robust with respect to uncorrelated scatter in $k_i$. Interestingly, we also note that the difference between the observed and the intrinsic dust-to-gas ratio for the lognormal model with $\sigma_\xi = 1$ is in agreement with the difference between the metallicity in optically and radio-selected samples of DLA systems (see Akerman et al. 2005).

**Correlated scatter in $k_i$.**—In this scenario, the analysis in terms of an average dust-to-gas ratio systematically misestimates the dust content at both ends of the column density distributions of the absorbers. In particular, if $k_i$ increases with $N_{HI}$, then absorbers with the largest column densities will be more dust rich than assumed by our analysis. This means that these absorbers are more likely to be dropped by the sample than is assumed in...

---

4 Note, however, that the difference in the metallicities measured by Akerman et al. (2005) is not particularly robust from a statistical point of view, as it is within the $1\sigma$ error bar.
the modeling, so \( \Omega_{\text{HI}}^{(\text{DLA})} \) is underestimated. When \( k_i \) decreases with \( N_\text{i} \), the opposite happens; the densest absorbers are less dusty than assumed and \( \Omega_{\text{HI}}^{(\text{DLA})} \) is overestimated. Due to the skewness of the lognormal distribution, the bias introduced by the correlation is expected to be larger when \( k_i \) is a decreasing function of \( N_\text{H}_\text{i} \). This is indeed what we observe in the analysis of our synthetic observations (see Fig. 7 and Table 4). The bias introduced is, however, relatively modest (at most 15% in \( \Omega_{\text{HI}}^{(\text{DLA})} \)), even assuming a rather strong dependence of \( k_i \) on \( N_\text{H}_\text{i} \) (i.e., \( k_i \propto N_\text{H}_\text{i}^{-0.8} \)). The bias could be avoided by taking into account the precise form of the correlation in the modeling of the dust-to-gas ratio. Unfortunately, the \( k_i \) versus \( N_\text{H}_\text{i} \) correlation appears hard to quantify with the present data. Only a small subset of the known DLA systems have a measured metallicity and highly dusty and dense absorbers are preferentially missed from the sample. In addition, there is some evidence that the fraction of metals depleted in dust depends on the column density of the absorbers (Welty et al. 1997). Therefore, plots of the measured gas phase metallicity versus the hydrogen column density, such as those presented in Boisse et al. (1998) and Savaglio (2001), cannot be easily interpreted to extract the \( k_i(N_\text{H}_\text{i}) \) relation.

7. NONPARAMETRIC ESTIMATION OF \( \Omega_{\text{HI}}^{(\text{DLA})} \) FROM SDSS DR3

The comoving density of DLA systems can be also estimated using a nonparametric approach by taking the discrete limit of equation (22) combined with equation (16):

\[
\Omega_{\text{HI}}^{(\text{DLA})} = \frac{\mu m_\text{HI}H_0}{c\rho_c} \sum_x N_\text{H}_\text{x} e^{\beta_x} (\Delta X)_x. \quad (25)
\]

The sum is done over all the DLA systems in the sample, while the intrinsic total path length \( (\Delta X)_x \), is given by summing over all the individual path lengths for the quasars in the survey, with a weight depending on the optical depth \( \tau_n \) along the line of sight to the \( n \)th quasar and on the redshift-dependent slope of the quasar luminosity function \( \beta(z) \):

\[
(\Delta X)_x = \sum_n (dX)_n e^{\beta X_n}. \quad (26)
\]

The optical depth \( \tau_n \) is given by

\[
\tau_n = \alpha_\kappa \sum_p k_{\text{MW}}(z_p) N_{\text{H}_\text{i}} N_\text{v} [\lambda_i/(1 + z_p)]^\beta, \quad (27)
\]

where the index \( p \) is running over all the DLA systems along the \( n \)th line of sight.

The random uncertainty on the measure has been estimated by using a “jackknife” analysis (Lupton 1993); we derive \( \Omega_{\text{HI}}^{(\text{DLA})} \) for \( N = 22 \) subsamples of quasars by ignoring \( N_\text{tot}/N \) quasars each time. The uncertainty on \( \Omega_{\text{HI}}^{(\text{DLA})} \) is then given by

\[
\sigma^2 = \frac{N - 1}{N} \sum_x (\Omega_x - \langle \Omega_x \rangle)^2, \quad (28)
\]

where \( \Omega_x \) stands for \( \Omega_{\text{HI}}^{(\text{DLA})} \). We obtain the following results:

\( \alpha_\kappa = 1 \).—Equation (25), applied to the whole sample with \( \alpha_\kappa = 1 \), gives \( \Omega_{\text{HI}}^{(\text{DLA})} = 1.04 \times 10^{-3} \) with an error of \( 8 \times 10^{-5} \) (see Table 5). The obscuration bias based on this estimate is about 30%. This represents a lower limit on the bias, as from Figure 8 it is evident that with \( \alpha_\kappa = 1 \), the dust bias is so severe that the discrete evaluation of \( \Omega_{\text{HI}}^{(\text{DLA})} \) has not converged. On the top of the random errors, estimated with the jackknife method,
this nonparametric estimation suffers from systematic effects of the order of at least 10%–20%.

\( \alpha_s = 0.5 \).—The values of \( \Omega_{\text{DLA}}^{i} \) obtained in our reference model with \( \alpha_s = 0.5 \) for different redshift intervals are reported in Table 5. The systematic corrections due to the dust bias are relatively modest (of the order of 10%) but still represent a contribution to the total uncertainty that cannot be neglected. Like in the case discussed above, a systematic effect is also present. Here the magnitude of the effect is smaller, of the order of 5%, as can be evaluated from the difference in \( \Omega_{\text{DLA}}^{i} \) between this method and the maximum likelihood analysis of \( \delta 4.2 \).

**CORALS.**—We also apply the same analysis on the DLA detections in the CORALS survey (Ellison et al. 2001), obtaining \( \Omega_{\text{DLA}}^{i} = (1.2 \pm 0.5) \times 10^{-3} \) (1 \( \sigma \) error bars; see Table 5). Note that our analysis leads to a different value for \( \Omega_{\text{DLA}}^{i} \) because in Ellison et al. (2001) the analysis was performed with a different cosmology (\( \Omega_M = 1, \Omega_{\Lambda} = 0 \)). The main limit of the survey is its modest path length extension (\( \Delta X = 200.8, \) to be compared with the Sloan path length \( \Delta X = 7333.1 \)), which is translated into a large random error on \( \Omega_{\text{DLA}}^{i} \). In addition, potential systematic errors due to incompleteness may bias the CORALS measure, much like dust introduces a bias in optically selected surveys. In fact, the likelihood modeling of the data is inconclusive, allowing a variety of different forms for \( f_i \) (see Fig. 3).

**8. CONCLUSIONS**

In this paper we have improved the analysis of the column density distribution of DLA systems in the SDSS DR3 DLA survey (Prochaska et al. 2005) to take into account the bias due to dust obscuration along the line of sight. A first modeling of the bias was constructed by Ostriker & Heisler (1984) and improved by Fall & Pei (1993). These earlier estimates expected a severe effect of obscuration, with the observed gas density of DLA systems being up to several times smaller than the intrinsic one. The best estimate for \( \Omega_{\text{DLA}}^{i} \), given by Fall & Pei (1993), is \( 4.9 \times 10^{-3} \) (obtained using a cosmology with \( \Omega_M = 1 \) and \( H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1} \)) and their upper limit is \( \Omega_{\text{DLA}}^{i} \leq 3 \times 10^{-2} \), obtained from a standard big bang nucleosynthesis abundance model. More recent works (Murphy & Liske 2004; Ellison et al. 2005) tend to dismiss the issue of obscuration bias based on the absence of systematic redening in the spectra of quasars with DLA absorption features.

Here we show that the effect of obscuration, while not being as severe as predicted by Fall & Pei (1993), does indeed play an important effect on the precise measurement of \( \Omega_{\text{DLA}}^{i} \). In the era of precision cosmology, where the observed density \( \Omega_{\text{DLA}}^{i} \) is constrained with errors below 10%, the systematic effects are not to be underestimated. With the typical amount of metals present in DLA systems, the observed density \( \Omega_{\text{DLA}}^{i} \) derived from shallow magnitude-limited surveys of quasars underestimates the intrinsic density \( \Omega_{\text{DLA}}^{i} \) by about 15%, assuming dust-poor DLA systems (i.e., systems with a fraction of metals in dust of 25%). Our best estimation for \( z = [2.2, 5.5] \) gives an intrinsic neutral gas density \( \Omega_{\text{DLA}}^{i} = 0.97^{+0.08}_{-0.06} \times 10^{-3} \) (1 \( \sigma \) error bars) to be compared with the observed gas density \( \Omega_{\text{DLA}}^{i} = 0.81^{+0.05}_{-0.025} \times 10^{-3} \) derived by Prochaska et al. (2005). If we leave the dust-to-metal ratio parameter \( \alpha_s \) free to vary over the relevant range (from 0 to 1), we find \( \Omega_{\text{DLA}}^{i} = 1.06^{+0.27}_{-0.20} \times 10^{-3} \), where the first set of error bars gives the 1 \( \sigma \) random errors and the second set

### Table 5

**Discrete Evaluation of \( \Omega_{\text{H}_1}^{i} \)**

| \( z_{\text{min}} \) | \( z_{\text{max}} \) | \( dX_\alpha \) | \( N_{\text{DLA}} \) | \( 10^3 \Omega_{\text{H}_1}^{i} \) | \( 10^3 \Omega_{\text{H}_1}^{i} \) | \( 10^3 \Omega_{\text{H}_1}^{i} \) |
|------------------------|------------------------|------------------|------------------|------------------|------------------|------------------|
| SDSS                   |                        |                  |                  |                  |                  |                  |
| 2.2                    | 5.5                    | 7333.1           | 525              | 0.04 \( \pm \) 0.08 | 0.91 \( \pm \) 0.07 | 0.82 \( \pm \) 0.06 |
| 2.2                    | 2.5                    | 1589.2           | 83               | 0.53 \( \pm \) 0.07 | 0.48 \( \pm \) 0.06 | 0.45 \( \pm \) 0.06 |
| 2.5                    | 3.0                    | 2697.6           | 189              | 0.64 \( \pm \) 0.14 | 0.90 \( \pm \) 0.11 | 0.79 \( \pm \) 0.09 |
| 3.0                    | 3.5                    | 1836.3           | 152              | 1.27 \( \pm \) 0.18 | 1.13 \( \pm \) 0.15 | 1.01 \( \pm \) 0.13 |
| 3.5                    | 4.0                    | 877.3            | 69               | 1.47 \( \pm \) 0.30 | 1.27 \( \pm \) 0.24 | 1.11 \( \pm \) 0.19 |
| 4.0                    | 5.5                    | 332.6            | 33               | 1.13 \( \pm \) 0.31 | 1.06 \( \pm \) 0.27 | 1.00 \( \pm \) 0.24 |
| CORALS                 |                        |                  |                  |                  |                  |                  |
| 1.8                    | 3.5                    | 200.8            | 17               | ...              | ...              | ...              |
|                        |                        |                  |                  |                  | 1.20 \( \pm \) 0.50 |                  |

Notes.—\( \Omega_{\text{DLA}}^{i} \) evaluated in the discrete limit for different redshift intervals from the SDSS and CORALS data. The first two columns give the redshift interval considered, the third the corresponding path length probed by the data (\( dX_\alpha \)). \( N_{\text{DLA}} \) is the number of systems identified in the interval, \( \Omega_{\text{H}_1}^{i} \) the intrinsic density of neutral gas (evaluated using \( \alpha_s = 1, \alpha_s = 0.5 \), and \( \alpha_s = 0 \)). The uncertainty is evaluated with a jackknife analysis.

![Fig. 8.—Cumulative neutral gas density in DLA systems vs. column density for the SDSS DR3 DLA (observed, dotted line; intrinsic with \( \alpha_s = 1 \), solid line). Superimposed onto the intrinsic curve, we show the profile from the best-fitting gamma function \( f_i \) (short-dashed line) and from a power law \( f_i \) (long-dashed line) with slope \( -1.95 \); the discrete evaluation for the intrinsic distribution does not extend to sufficiently high column densities to allow convergence. [See the electronic edition of the Journal for a color version of this figure.]](image-url)
gives the modeling uncertainty dependent on the fraction of metals in dust. The obscuration bias therefore represents the main source of uncertainties for the determination of the intrinsic neutral gas content in DLA systems.

Our analysis has been carried out assuming that all the DLA absorbers at a given redshift have the same dust-to-gas ratio. By means of Monte Carlo simulations of synthetic observations, we show that this assumption does not introduce a significant bias as long as the dust-to-gas ratio is not correlated with the hydrogen column density. A correlation of the form \( k_i \propto N_i^\gamma \) introduces a systematic error of the order of \( 14\% \) in \( \Omega_{\text{H}i}^{(\text{DLA})} \) for \( \gamma = -0.8 \) and of \( -3\% \) for \( \gamma = +0.8 \) (the bias is reduced to \( +8\% \) and \( 0\% \) for \( \gamma = \pm 0.4 \), respectively).

Caution is also needed when taking the result from the maximum likelihood at face value. In our reference scenario, we find that a power-law function for column density distribution of DLA systems can be ruled out at a confidence level no greater than 99.5%. If the dust content in DLA systems is higher, the dust bias becomes more significant. For 50% of the metals in dust grains, the present data do not allow us to put an upper limit on the neutral gas content with confidence level greater than 90%. We show in fact that there is a probability of about 10% that the SDSS DR3 DLA data are consistent with an intrinsic power-law distribution in column density adopting \( \alpha_c = 1 \). The slope of this power law is \( m_2 = -1.95 \) and an upper cutoff cannot be derived from the data. In optically selected surveys, absorbers with high column densities of neutral gas and with high metallicity are missed, letting systematic uncertainties go easily out of control.

Radio-selected quasar samples would represent an elegant, bias-free solution to the measurements of \( \Omega_{\text{H}i}^{(\text{DLA})} \). Unfortunately, with the present extension, these surveys constraint \( \Omega_{\text{H}i}^{(\text{DLA})} \) with a random uncertainty on the order of 40% (plus potential systematic effects due to the limited number of observed DLA systems). To reduce the errors to within the 10% level, without the need to assume a modeling for the dust bias, it is therefore necessary to increase the path length of radio surveys by a factor of 10, at least. The UCSD radio survey (Jorgenson et al. 2006) has recently made promising progresses in this direction.

It is a pleasure to thank Michael S. Fall for stimulating discussions and Jason X. Prochaska for useful suggestions and comments on a draft of the paper. We are grateful to the referee for constructive suggestions. This work was supported in part by NASA JWST IDS grant NAG5-12458.

## REFERENCES

Akerman, C. J., Ellison, S. L., Pettini, M., & Steidel, C. C. 2005, A&A, 440, 499
Boisse, P., Le Brun, V., Bergeron, J., & Deharveng, J.-M. 1998, A&A, 333, 841
Dessauges-Zavadsky, M., Calura, F., Prochaska, J. X., D’Odorico, S., & Matteucci, F. 2004, A&A, 416, 79
Dessauges-Zavadsky, M., Prochaska, J. X., D’Odorico, S., Calura, F., & Matteucci, F. 2006, A&A, 445, 93
Ellison, S. L., Hall, P. B., & Lira, P. 2005, AJ, 130, 1345
Ellison, S. L., Yan, L., Hook, I. M., Wall, J. V., & Shaver, P. 2001, A&A, 379, 393
Fall, S. M., & Pei, Y. C. 1989, ApJ, 337, 7
———. 1993, ApJ, 402, 479
Fall, S. M., Pei, Y. C., & McMahon, M. G. 1989, ApJ, 341, L5
Hahnelt, D. G., Steinmetz, M., & Rauch, M. 1998, ApJ, 495, 647
Jorgenson, R. A., Wolfe, A. M., Prochaska, J. X., Lu, L., Howk, J. C., Cooke, J., Gawiser, E., & Gelino, D. M. 2006, ApJ, 646, 730
Kulkarni, V. P., Fall, S. M., Lauroesch, J. T., York, D. G., Welty, D. E., Khare, P., & Truran, J. W. 2005, ApJ, 618, 68
Lanzetta, K. M., McMahon, R. G., Wolfe, A. M., Turnshek, D. A., Hazard, C., & Lu, L. 1991, ApJS, 77, 1
Lupton, R. 1993, Statistics in Theory and Practice (Princeton: Princeton Univ. Press)
Murphy, M. T., & Liske, J. 2004, MNRAS, 354, L31
Nagamine, K., Springel, V., & Hernquist, L. 2004, MNRAS, 348, 421
Ostriker, J. P., & Heisler, J. 1984, ApJ, 278, 1
Pei, Y. C. 1992, ApJ, 395, 130
———. 1995, ApJ, 438, 623
Pei, Y. C., & Fall, S. M. 1995, ApJ, 454, 69
Pei, Y. C., Fall, S. M., & Bechtold, J. 1991, ApJ, 378, 6
Pei, Y. C., Fall, S. M., & Hauser, M. G. 1999, ApJ, 522, 604
Pettini, M., King, D. L., Smith, L. J., & Hunstead, R. W. 1997, ApJ, 478, 536
Prochaska, J. X., Gawiser, E., Wolfe, A. M., Cooke, J., & Gelino, D. 2003, ApJS, 147, 227
Prochaska, J. X., Herbert-Fort, S., & Wolfe, A. M. 2005, ApJ, 635, 123
Prochaska, J. X., & Wolfe, A. M. 1997, ApJ, 487, 73
Richards, G. T., et al. 2002, AJ, 123, 2945
———. 2006, AJ, 131, 2766
Savaglio, S. 2001, in IAU Symp. 204, The Extragalactic Infrared Background and Its Cosmological Implications, ed. M. Harwit (Cambridge: Cambridge Univ. Press), 307
Trenti, M., & Stiavelli, M. 2006, ApJ, in press
Vladilo, G. 2002, A&A, 391, 407
Welty, D. E., Lauroesch, J. T., Blades, J. C., Hobbs, L. M., & York, D. G. 1997, ApJ, 489, 672
Wild, V., Hewett, P. C. 2005, MNRAS, 361, L30
Wild, V., Hewett, P. C., & Pettini, M. 2006, MNRAS, 367, 211
Wolfe, A. M., Gawiser, E., & Prochaska, J. X. 2005, ARA&A, 43, 861
York, D. G., et al. 2006, MNRAS, 367, 945