

THE LINE-OF-SIGHT PROXIMITY EFFECT AND THE MASS OF QUASAR HOST HALOS

CLAUDE-ANDRÉ FAUCHER-GIGUÈRE¹, ADAM LIDZ¹, MATIAS ZALDARRIAGA^{1,2}, LARS HERNQUIST¹*Draft*

ABSTRACT

We show that the Lyman- α (Ly α) optical depth statistics in the proximity regions of quasar spectra depend on the mass of the dark matter halos hosting the quasars. This is owing to both the overdensity around the quasars and the associated infall of gas toward them. For a fiducial quasar host halo mass of $3.0 \pm 1.6 h^{-1} \times 10^{12} M_{\odot}$, as inferred by Croom et al. from clustering in the 2dF QSO Redshift Survey, we show that estimates of the ionizing background (Γ^{bkg}) from proximity effect measurements could be biased high by a factor of ≈ 2.5 at $z = 3$ owing to neglecting these effects alone. The clustering of galaxies and other active galactic nuclei around the proximity effect quasars enhances the local background, but is not expected to skew measurements by more than a few percent. Assuming the measurements of Γ^{bkg} based on the mean flux decrement in the Ly α forest to be free of bias, we demonstrate how the proximity effect analysis can be inverted to measure the mass of the dark matter halos hosting quasars. In ideal conditions, such a measurement could be made with a precision comparable to the best clustering constraints to date from a modest sample of only about 100 spectra. We discuss observational difficulties, including continuum flux estimation, quasar systematic redshift determination, and quasar variability, which make accurate proximity effect measurements challenging in practice. These are also likely to contribute to the discrepancies between existing proximity effect and flux decrement measurements of Γ^{bkg} .

Subject headings: Cosmology: theory, diffuse radiation — methods: numerical, statistical — galaxies: halos — quasars: general, absorption lines

1. INTRODUCTION

As the geometrical properties and initial conditions of the Universe are becoming fairly well known (e.g., Spergel et al. 2006), an increasingly important aspect of cosmology concerns the emergence and evolution of complex, non-linear structures. Light sources — stellar and quasistellar — are tracers of these structures and it is thus of prime interest to study their nature and evolution. A measure of the stellar and quasistellar activity is given by the amount and distribution of photons with energy above the hydrogen ionization potential in the intergalactic medium (IGM). The relevant integral quantity is the redshift (z)-dependent background hydrogen photoionization rate, $\Gamma^{bkg}(z)$, which measures the integrated contributions from all sources of ionizing photons. Several measurements of Γ^{bkg} exist in the literature (see Figure 1, complemented by the numerical values tabulated in Table 2 of Appendix A). Unfortunately, the agreement between various studies using different methods is poor.

Two main methods have been used to estimate Γ^{bkg} : the flux decrement method and the line-of-sight proximity effect method. In the flux decrement method (Rauch et al. 1997; Songaila et al. 1999; McDonald & Miralda-Escudé 2001; Meiksin & White 2003; Tytler et al. 2004a; Bolton et al. 2005; Kirkman et al. 2005; Jena et al. 2005), one considers the portion of the Lyman- α (Ly α) forest along lines of sight to quasars (QSOs) away from the latter. The parameter $\mu \propto \Omega_b^2 h^3 / \Gamma^{bkg}$, where Ω_b is the

cosmic baryon density in units of the critical density and $h = H_0/100 \text{ km s}^{-1} \text{ Mpc}^{-1}$, is obtained by requiring the mean flux decrement ($D \equiv \langle 1 - e^{-\tau} \rangle$) in a cosmological simulation to agree with the mean decrement observed in quasar spectra. In the proximity effect method (Carswell et al. 1987; Bajtlik et al. 1988; Lu et al. 1991; Kulkarni & Fall 1993; Williger et al. 1994; Cristiani et al. 1995; Giallongo et al. 1996; Srianand & Khare 1996; Scott et al. 2000), in which we will be most interested here, one focuses on the “proximity region” near the quasar, where the ionizing flux emitted by the quasar itself is comparable to the background. (For related work on the *transverse* proximity effect see, e.g., Schirber et al. 2004; Croft 2004; Adelberger 2004; Hennawi & Prochaska 2006.)

In the simplest proximity effect model, the quasar lies in a random region of the Universe and the hydrogen gas is in photoionization equilibrium. Neglecting all motion of the gas, the Ly α optical depth at any point is then given by

$$\tau^{prox} = \frac{\tau^{off}}{1 + \omega(r)}, \quad (1)$$

where $\omega(r) \equiv \Gamma^{QSO}(r)/\Gamma^{bkg}$, $\Gamma^{QSO}(r)$ is the contribution to the photoionization rate owing to the quasar itself at proper distance r from it, and τ^{off} is the optical depth that would be obtained if the quasar were turned off. Near the quasar, $\Gamma^{QSO} \propto r^{-2}$ (and hence ω) is large, causing a statistical decrease in observed Ly α absorption. Authors have generally fitted a model, introduced by Bajtlik et al. (1988) (hereafter BDO), for the variation of the density of absorption lines near the quasars from spectra to estimate Γ^{bkg} . Integration over known sources of radiation (e.g., Steidel et al. 2001; Hopkins et al. 2006d) can also be used to estimate

¹ Department of Astronomy, Harvard University, Cambridge, MA, 02138, USA; cgiguere@cfa.harvard.edu.

² Jefferson Physical Laboratory, Harvard University, Cambridge, MA, 02138, USA.

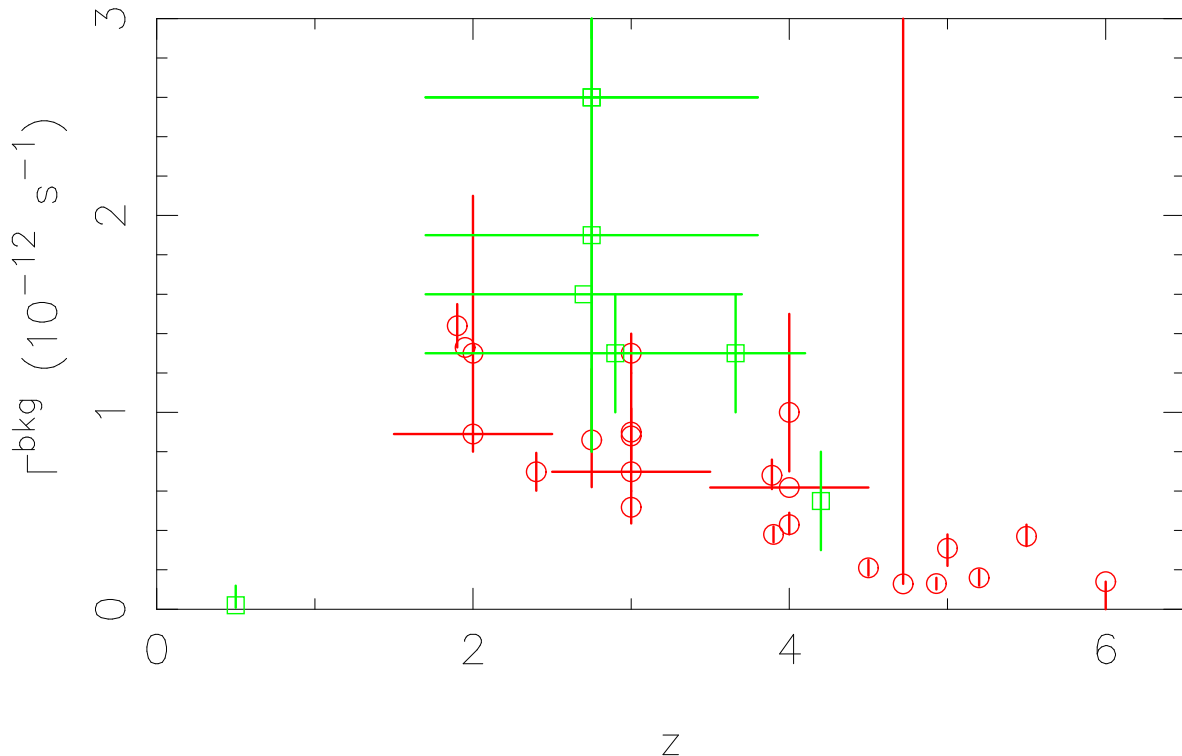


FIG. 1.— Existing measurements of Γ^{bkg} using the flux decrement (red, circles) and the proximity effect methods (green, squares). The horizontal error bars, where present, indicate the redshift range over which the measurement applies. The vertical bars show the reported uncertainties. The early proximity effect estimate $\Gamma^{bkg} \gtrsim 7.8 \times 10^{12} \text{ s}^{-1}$ at $z = 3.75$ of Carswell et al. (1987) is not shown. In the redshift range where measurement have been made using both methods, the proximity effect estimates are generally higher by a factor of 2–3. The scatter between the different proximity effect measurements is also larger than in the flux decrement case. Numerical values, references, and additional comments are provided in Table 2.

Γ^{bkg} . However, this method provides only lower bounds and much of the interest in radiation backgrounds lies in determining whether the measured values agree with the summed contribution of resolved sources. Thus, Γ^{bkg} should be measured independently of such estimates and we will not consider them further in this paper.

The proximity effect method has generally yielded measurements of Γ^{bkg} higher than flux decrement estimates (Figure 1), suggesting unmodeled biases in one or both methods. There are indeed potential sources of bias in the proximity effect method, mostly owing to the peculiar environments in which quasars are thought to reside, that have not generally been modeled in proximity effect analyses (although see Loeb & Eisenstein 1995, Rollinde et al. 2005, Guimarães et al. 2007, and §9 for a comparison with previous work). These potential biases are a focus of this paper.

In cold dark matter (CDM) models of structure formation, quasars are expected to preferentially form in overdense regions of the universe (e.g., Springel et al. 2005), violating the assumption that they lie in random locations. The overdense environments of quasars not only affect the local density of the absorbing gas, but also its peculiar velocity field. General infall of the gas toward the density peaks where quasars reside is expected (Loeb & Eisenstein 1995; Barkana 2004) and in fact may have been directly detected (Barkana & Loeb 2003). Moreover, galaxies and other active galactic nuclei (AGN) should preferentially form in overdense regions (Bond et al. 1991) and therefore cluster around quasars. The local “background” flux near the quasar may be

more intense than average. Thermal broadening owing to the finite temperature of the gas also causes redshift-space distortions, increasing the equivalent width of absorption lines.

The flux decrement method is also not guaranteed to be unbiased, most importantly as it requires estimation of the unabsorbed continuum level. This is difficult to do reliably, especially at high redshifts where unabsorbed portions become increasingly rare, if existent at all, in quasar spectra. Alternatively, one can estimate the unabsorbed continuum level by extrapolating from redward of Ly- α (e.g., Press & Rybicki 1993). However, there may be a break in the quasar continuum close to Ly- α (Zheng et al. 1997; Telfer et al. 2002), which could bias resulting estimates of the flux decrement (Seljak et al. 2003). For further discussion of flux decrement estimation and related issues see, e.g., Seljak et al. 2003, Tytler et al. 2004b, and Lidz et al. 2006a.

The flux decrement method is also sensitive to assumed cosmological parameters, most notably Ω_b , through the factor $\Omega_b h^2$ in μ . If an independent measurement of Γ^{bkg} is available, then the flux decrement may be used to infer Ω_b . The proximity effect provides such a measurement and, as integration over known sources can only provide lower bounds on Γ^{bkg} , is the only known way to use the flux decrement to obtain a full measurement of Ω_b .

Integration over sources from a luminosity at rest-frame frequencies below the Lyman limit requires knowledge of the fraction of ionizing photons which escape the immediate environment of the emitters and actually have

an effect on the surrounding IGM. As recently shown by Shapley et al. (2006), the escape fraction of Lyman break galaxies (LBG) at $z \sim 3$, whose contribution may well dominate Γ^{bkg} at these redshifts (Steidel et al. 2001), is highly uncertain. Agreement between measurements of Ω_b from light element abundances interpreted in the context of Big Bang nucleosynthesis (e.g., Burles et al. 2001), cosmic microwave background anisotropies (e.g., Spergel et al. 2006), and the flux decrement would provide a powerful consistency check of the standard cosmological model as well as of the gravitational paradigm for the Ly α forest (Hui et al. 2002).

A detailed understanding of the proximity effect would also provide insight into the astrophysics of quasar environments and in turn could be used to study these. For example, if the dark matter (DM) halos in which quasars lie can be shown to induce significant biases in measurements of Γ^{bkg} , then one could parameterize these biases with respect to the quasar host halo mass, M_{DM} . Knowledge of the quasar host halo mass and constraints on its dependence on quasar luminosity would be tremendously useful in determining how activity is triggered in AGN. This is especially true in a picture in which activity is related to mergers (e.g., Hopkins et al. 2006a; Lidz et al. 2006b), as such events inevitably increase the host halo mass.

Recent clustering measurements suggest a universal quasar host halo mass. For example, Croom et al. (2005) find that quasars lie in dark matter halos of mass $3.0 \pm 1.6 h^{-1} \times 10^{12} M_\odot$ based on an analysis of the quasar two-point correlation function as measured from 2dF QSO Redshift Survey (2QZ) quasars, only weakly dependent on redshift and with no evidence for a luminosity dependence. If confirmed, this result would provide constraints on quasar evolution and call for an explanation of the physical origin of this critical mass for nuclear activity. Given that halos continuously grow through mergers, a universal mass for AGN activity also constrains quasar lifetimes. Combined with a relation between the mass of the central black hole (M_{BH}) and the mass of its host halo (e.g., Ferrarese 2002), a universal halo mass would also suggest a universal black hole mass, which may be at odds with the variety of nuclear black hole masses observed in the local universe (e.g., Kormendy & Richstone 1995) if most galactic nuclei have gone through an active phase, although it is possible that the $M_{BH} - M_{DM}$ relation evolves with redshift (Hopkins et al. 2006c, Hopkins et al., in preparation). If the universal quasar host halo mass is correct, it would provide evidence for a redshift-dependent $M_{BH} - M_{DM}$ relation. As practically all the evidence for the quasar host halo mass universality is inferred from clustering measurements, independent quasar host halo mass estimates are clearly desirable.

In this work, we quantify the effects of quasar host halos on the Ly α statistics in the proximity regions of quasars and the biases they induce in proximity effect measurements of Γ^{bkg} (§4 and §5), develop a Monte Carlo-based method for making unbiased Γ^{bkg} measurements (§6), and demonstrate how the analysis can be inverted to use the proximity effect to measure the masses of dark matter halos hosting quasars (§7). Throughout most of the paper, we assume that the analyses are per-

formed in ideal quasar spectra. In §8, we consider observational difficulties which make accurate proximity effect measurements challenging in practice and which may also contribute to the discrepancies between proximity effect and flux decrement measurements of Γ^{bkg} . We compare with previous work in §9 and conclude in §10. We begin by summarizing conventions used throughout the paper in §2 and describing our numerical framework in §3.

Most previous studies of the proximity effect were based on a picture of the Ly α forest as arising from absorption by discrete gas clouds and on counting absorption lines. Here, we adopt the modern picture of the forest as arising from absorption by smooth density fluctuations imposed on the warm photoionized IGM as a natural consequence of hierarchical structure formation within CDM models. This picture, based on a combination of detailed hydrodynamical simulations (Cen et al. 1994; Zhang et al. 1995; Hernquist et al. 1996; Katz et al. 1996; Miralda-Escude et al. 1996; Theuns et al. 1998; Davé et al. 1999) and analyses of high-resolution, high signal-to-noise quasar spectra (Lu et al. 1996; Kirkman & Tytler 1997; Kim et al. 2002) with matching properties, is strongly supported by all available evidence.

As we were finalizing this work we became aware of a related study by Kim & Croft (2007). These authors also consider constraining the mass of quasar host halos from the Ly α forest. They, however, focus on using quasar pairs and work under the hypothesis (supported by observations) that the *transverse* proximity effect is negligible. We, on the other hand, focus on the *line-of-sight* proximity effect and model it in detail. The two works, while broadly consistent where there is overlap, are hence complementary.

2. CONVENTIONS

In this paper, \hat{X} denotes a random variable, X a realized value, and $f_{\hat{X}}(X)$ the probability density function (PDF) of \hat{X} evaluated at X . The linear overdensity δ is defined such that $\rho/\bar{\rho} = 1 + \delta$, where ρ is the local mass density and $\bar{\rho}$ is the mean density of the Universe. We assume a flat Λ CDM cosmology similar to the one inferred from *Wilkinson Microwave Anisotropy Probe* (WMAP) measurements of the cosmic microwave background (Spergel et al. 2006); the exact parameters assumed in our cosmological simulations are given in Table 1. We select “typical” quasars to have spectral indexes ($L_\nu \propto \nu^{-\alpha}$) $\alpha = 1.57$ blueward of Ly α . For the typical specific luminosity at the Lyman limit, we take the median $\log_{10}(L_{912}/\text{erg s}^{-1} \text{ cm}^{-2} \text{ Hz}^{-1}) = 31.1$ for the quasars in the proximity effect analysis of Scott et al. (2000). When we refer to quasars with luminosity one standard deviation from typical, we approximate (generously) the standard deviation in Scott et al. (2000) to $\sigma_{\log_{10} L_{912}} = 1$.

3. NUMERICAL SIMULATIONS AND MOCK SPECTRA

To study the effects of various potential sources of bias on the measurement of Γ^{bkg} from the proximity effect, we will make use of mock quasar spectra generated from cosmological simulations. In generating the mock spectra, we will turn on and off different effects, examining their impact on the derived Γ^{bkg} . In this section, we describe the simulations that we use and how basic mock

TABLE 1
SIMULATION PARAMETERS

<i>N</i> -body Dark Matter Simulation	
Parameter	Value
Code	GADGET-2 ^a
Box Side Length	100 h^{-1} comoving Mpc
Boundary Conditions	Periodic
Number of Particles	256^3
Particle Mass	$4.96 \times 10^9 h^{-1} M_\odot$
Ω_m	0.3
Ω_Λ	0.7
Ω_b	0.04
h	0.7
σ_8	0.9
n_s	1
Friends-of-Friends Halo Finder	
Linking Length	$0.2 \langle \Delta r_p \rangle^b$
Minimum Number of Particles	31
Grid	
Grid Interpolation	CIC on 256^3 grid points
Line of Sight Interpolation	TSC

^a Springel (2005), ^b Mean interparticle spacing.
spectra are constructed from them. The reader who is not concerned with such numerical details is encouraged to skip to §4, where we begin our discussion of the effects of overdensities and redshift-space distortions.

3.1. Cosmological Dark Matter Simulations

The first step in our modeling of the Ly α forest is to compute the dark matter density and velocity fields in a simulation box at different redshifts. We do so using the GADGET-2 (Springel 2005) cosmological code. The simulation parameters are given in Table 1. When generating density and velocity profiles and constructing optical depth PDFs in the following sections, significant numbers of halos in different mass ranges are required. This motivates our choice of a large simulation box, with side length $100 h^{-1}$ comoving Mpc. In order to compare our results with existing proximity effect analyses, it is also necessary to run the simulation to redshifts at least as low as $z \approx 2$. We use *N*-body outputs with 256^3 particles at $z = 2, 3$, and 4; the least massive resolved halos have mass $4.96 h^{-1} \times 10^9 M_\odot$. Ideally, for a detailed comparison with Ly α forest data, we would use a higher resolution, fully hydrodynamic simulation (e.g., Viel et al. 2006). Our present *N*-body approach should, however, be adequate for investigating the *relative* impact of quasar environments on Ly α absorption statistics.

3.2. Mock Quasar Spectra from Dark Matter Simulations

In the simplest model, the neutral hydrogen (HI) density at any given point in the Universe is given by photoionization equilibrium. Let Γ^{tot} (s^{-1}) be the total HI photoionization rate per atom, $R(T)$ ($cm^3 s^{-1}$) the recombination rate, n_{HI} the HI number density, n_{HII} the ionized hydrogen (HII) number density, and n_e be the free electron number density in proper units. Then equilibrium requires

$$\Gamma^{tot} n_{HI} = R(T) n_e n_{HII}. \quad (2)$$

For a small ionized fraction $x_i \equiv n_{HII}/n_{HI} \ll 1$ (which is certainly true in the IGM at the redshifts of interest here,

$z \lesssim 6$, as indicated by the absence of a Gunn-Peterson in the known quasars at these redshifts),

$$n_{HI} = \frac{R(T) n_{tot}^2}{\Gamma^{tot}}, \quad (3)$$

where $n_{tot} = n_{HI} + n_{HII} \approx \Omega_b X_H \rho_c / m_p$ is the total hydrogen number density. We take the cosmic hydrogen mass fraction $X_H = 0.75$ (e.g., Burles et al. 1999). The above expression for n_{HI} neglects the helium contribution to the free electron density. For fully singly ionized helium, we make an error of 8% on n_e and hence on n_{HI} ; for fully doubly ionized helium, the error is doubled. Simplifying approximations such as this one have no impact on our discussion, which should be relatively independent of the details of the cosmology. For the recombination rate, we use the approximate expression

$$R(T) = 4.2 \times 10^{-13} \left(\frac{T}{10^4 \text{ K}} \right)^{-0.7} cm^3 s^{-1}, \quad (4)$$

(Hui & Gnedin 1997) where T is the local gas temperature.

For $\delta \lesssim 5$, Hui & Gnedin (1997) derived the equation of state

$$T = T_0 (1 + \delta)^\beta, \quad (5)$$

for the IGM, where $T_0(z)$ is the temperature of a fluid element which remains at the cosmic mean density and $\beta(z)$ parameterizes the density dependence. Detailed expressions for T_0 and β , which we use in our numerical calculations, are given in Appendix B. To order of magnitude, $T_0 \sim 10^4$ K and in the limit of early reionization ($z_{reion} \gtrsim 10$, consistent with recent measurements of the Thomson optical depth to reionization by Page et al. 2006), $\beta \rightarrow 0.62$. Temperature measurements from the $z \sim 3$ Ly α forest favor slightly larger temperatures and a slightly less steep density dependence (e.g., McDonald et al. 2001), possibly owing to recent HeII reionization. Our basic conclusions should, however, not depend on the precise temperature-density relation assumed.

For a quasar with isotropic specific luminosity $L_\nu = L_Q (\nu/\nu_{Ly})^{-\alpha_Q}$, where ν_{Ly} is the hydrogen ionization frequency in its ground state (the Lyman limit) and α_Q is the spectral index, $\Gamma^{tot} = \Gamma^{QSO}(L_Q, \alpha_Q; r) + \Gamma^{bkg}(z)$, where

$$\Gamma^{QSO}(L_Q, \alpha_Q; r) \approx \int_{\nu_{Ly}}^{\infty} d\nu \sigma(\nu) \frac{L_\nu}{4\pi r^2 h_P \nu} \quad (6)$$

$$= \frac{A_0 L_Q}{4\pi h_P (\alpha_Q + 3) r^2}. \quad (7)$$

Here, h_P is Planck's constant and the hydrogen photoionization cross section is approximately given by

$$\sigma(\nu) = A_0 \left(\frac{\nu}{\nu_0} \right)^{-3} \quad (8)$$

for ν near ν_{Ly} , where $A_0 = 6.30 \times 10^{-18} cm^2$ (Osterbrock & Ferland 2006). In equation (6), we assume that the proper distance to the quasar, r , is sufficiently small that the point is effectively at the same redshift as the quasar, $z \approx z_Q$.

The *N*-body dark matter simulation provides the position and velocity of each particle in the box. These are

interpolated on a uniform grid using a cloud-in-cell (CIC) algorithm (Hockney & Eastwood 1988) to construct density and velocity fields. Triangular-shaped cloud (TSC) interpolation is then used to smoothly project the fields onto sight lines which are drawn through the box with randomly selected orientations. We obtain fields $\delta(r)$ and $v_{\parallel}(r)$ ($v_{\parallel} > 0$ indicates motion toward the observer) for the density and velocity along the line of sight, respectively, as a function of proper distance from the quasar. From the the density field, we compute the corresponding temperature field using the equation of state (5). Given quasar properties $Q = \{z_Q, L_Q, \alpha_Q\}$, Γ^{QSO} is computed at each pixel using equation 6. For simplicity, in §5 – §7 all simulated quasars are given the median luminosity from the Scott et al. (2000) sample (see §2).

We assume that the baryons trace the dark matter.³ Although no Jeans-scale smoothing filter is explicitly applied, our simulation has resolution comparable to the Jeans scale ($\sim 40 \text{ km s}^{-1}$ and $\sim 20 \text{ km s}^{-1}$, respectively, at $z = 3$) and we do model thermal broadening in our mock spectra (see §4.2). The HI number density along the line of sight is then given by equation (3). Ignoring redshift-space distortions for the moment and for a negligible line width, the Ly α optical depth at any given point is

$$\tau = \frac{\pi e^2 f_{Ly\alpha}}{m_e H_0 \nu_{Ly\alpha}} \frac{n_{\text{HI}}}{\sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda}}, \quad (9)$$

where $f_{Ly\alpha}$ is the oscillator strength and $\nu_{Ly\alpha}$ is the frequency of the Ly α transition. Here, the optical depth as a function of distance from the quasar, $\tau(r)$ (equivalently, the transmission coefficient, $e^{-\tau}$), is the quantity of interest and this is what we will refer to as a “spectrum.”

Figure 2 shows examples of mock spectra constructed from the simulation. Figure 3 shows the transmission coefficient as a function of distance from the quasar averaged over ensembles of 1000 quasars at $z = 2, 3$, and 4. The increase in the mean transmission near the quasars illustrates the proximity effect. We note that the effect is more pronounced at higher redshifts, even for fixed Γ^{bkg} and Γ^{QSO} , owing to the increase in the neutral hydrogen density attributed to the cosmological expansion. At low redshifts, the IGM is more dilute and hence the transmission is high even away from quasars. This suggests, as corroborated by our analysis in §5.5, that it is easier to obtain constraints on Γ^{bkg} from the proximity effect at higher redshifts. On the other hand, bright quasars are rarer, and continuum fitting is more difficult, at high redshift.

4. OVERDENSITIES, INFALL, AND CLUSTERING AROUND QUASARS

After setting up the basis of our numerical framework in the previous section, we are now ready to discuss the effects of overdense quasar environments on mock spectra. We describe how we model matter overdensities in §4.1 and the redshift-space distortions owing to gas infall in §4.2. We then pause to take a look at how the

³ See e.g., Meiksin & White (2001) and Viel et al. (2006), for comparisons between Ly α forest statistics extracted from dark matter, pseudo-hydrodynamic, and fully hydrodynamic simulations.

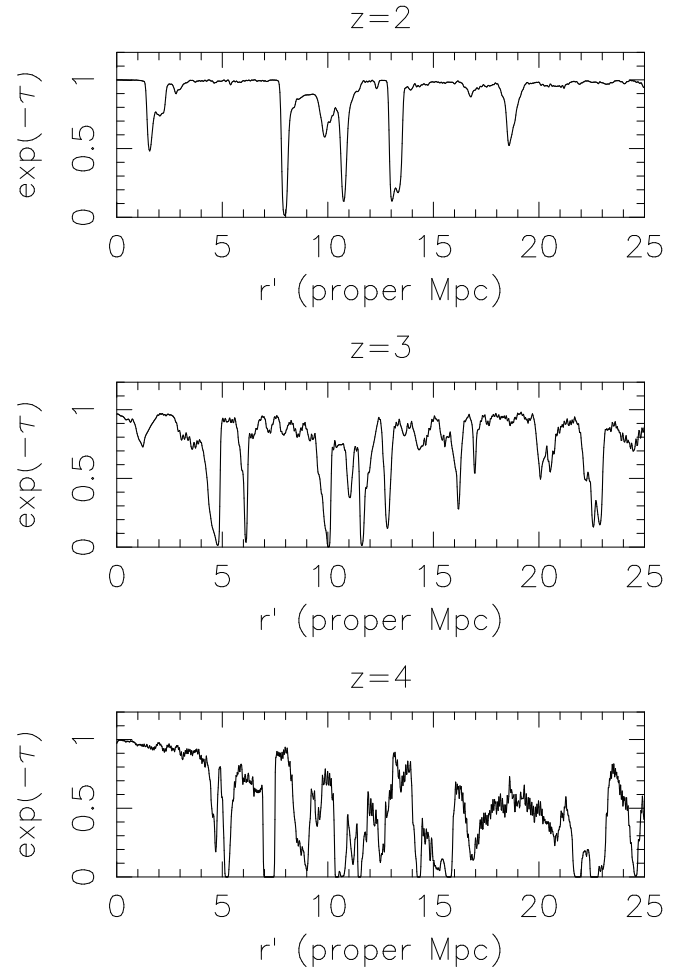


FIG. 2.— Examples of mock spectra (shown as the transmission coefficient as a function of the redshift-space distance from the quasar) constructed from the simulation at $z = 2, 3$, and 4 (top, middle, and bottom panel). In each case, the quasar has typical luminosity and $\Gamma^{bkg} = 10^{-12} \text{ s}^{-1}$. The local overdensities in which the quasars reside and the redshift-space distortions are modeled as in §4.1 and §4.2.

optical depth PDFs are affected by the overdensities and redshift-space distortions associated with massive dark matter host halos in §4.3. The biases they induce on measurements of Γ^{bkg} are given full consideration later in §5, after describing our likelihood formalism. We estimate the effect clustering of galaxies and other AGN around proximity effect quasars on measurements of Γ^{bkg} in §4.4, concluding that that it is relatively unimportant.

4.1. Overdensities

Structures in the Universe such as galaxies and quasars form in the collapse of regions of the Universe with mean density exceeding the critical value necessary to overcome the Hubble flow. Such regions preferentially arise in larger-scale overdensities (e.g., Bond et al. 1991). In addition, the density enhancement around collapsed objects generally extends well beyond their virial radius (Barkana 2004; Prada et al. 2006). As a consequence, quasars are expected to be surrounded by an excess of absorbing gas with respect to a random location in the Universe. In fact, there is strong and growing evidence from clustering measurements that quasars reside in massive dark matter halos. Using AGN-AGN clustering mea-

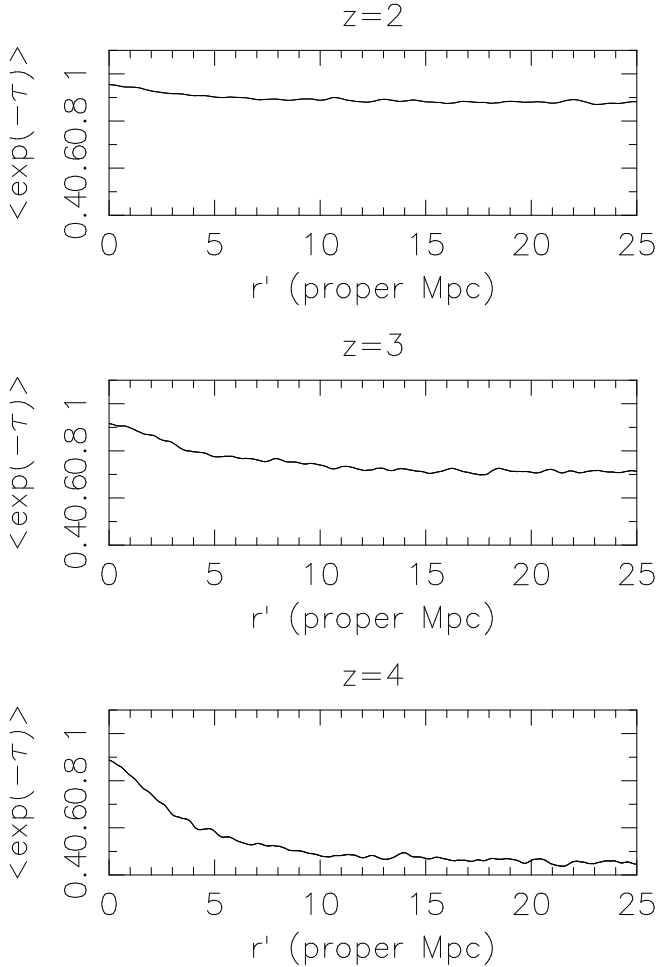


FIG. 3.— Transmission coefficient as a function of redshift-space distance from the quasar averaged over samples of 1000 mock spectra for quasars of typical luminosity and $\Gamma^{bg} = 10^{-12} \text{ s}^{-1}$ at $z = 2, 3$, and 4 (top, middle, and bottom panel). The proximity effect, seen as an increase in the mean transmission near the quasars, is stronger at higher redshifts where the absorbing gas is denser owing to the cosmological expansion. The local overdensities in which the quasars reside and the redshift-space distortions are modeled as in §4.1 and §4.2. In each case, the curves were smoothed with a boxcar of width 0.2 comoving Mpc.

measurements from the 2dF QSO Redshift Survey (2QZ) at $0.8 < z < 2.1$, Porciani et al. (2004) find a minimum mass for a DM halo to host a quasar of $10^{12} M_{\odot}$, with a characteristic mass $\sim 10^{13} M_{\odot}$. This result has been recently confirmed by Porciani & Norberg (2006).

Similarly, Wake et al. (2004) conclude, from a sample of narrow-line AGN from the Sloan Digital Sky Survey (SDSS) Data Release 1, that the minimum host DM halo mass is $2 \times 10^{12} M_{\odot}$. From 2QZ AGN-AGN clustering data at $0.3 < z < 2.2$, Croom et al. (2005) find that quasars lie in DM halos of mass $3.0 \pm 1.6 h^{-1} \times 10^{12} M_{\odot}$, regardless of luminosity. Coil et al. (2006) measure the AGN-galaxy cross-correlation function at $0.7 < z < 1.4$, using data from the SDSS and Deep Extragalactic Evolutionary Probe (DEEP) 2 survey and find a minimum DM halo mass $\sim 5 \times 10^{11} M_{\odot}$, with mean $3 \times 10^{12} M_{\odot}$, again with no evidence for luminosity dependence. The lack of dependence of clustering (and hence of DM halo mass) on quasar luminosity has been confirmed at higher redshifts by Adelberger & Steidel (2005a), who show that

the AGN-galaxy cross-correlation length at $1.8 < z < 3.5$ is constant over a range of 10 optical magnitudes. Finally, Barkana & Loeb (2003) estimate the DM host halo masses of the bright quasars SDSS1122 – 0229 at $z = 4.795 \pm 0.004$ and SDSS1030+0524 at $z = 6.28 \pm 0.02$ to be $2.5 \times 10^{12} M_{\odot}$ and $4.0 \times 10^{12} M_{\odot}$, respectively, independently of clustering based on the spectral signature of gas infall. Hopkins et al. (2006c) compare clustering of quasars and galaxies as a function of luminosity, redshift, and color and find that the clustering of local ellipticals is in accord with models which associate quasar activity with the formation of spheroids (Hopkins et al. 2006b).

We model the overdensities in which quasars reside by putting each mock quasar at the center of mass of a DM halo with mass M_{DM} in a specified range $[M_{min}, M_{max}]$ randomly chosen in the simulation box. To identify halos in the box, we use a friends-of-friends algorithm (e.g., Davis et al. 1985; Springel & Hernquist 2003) with linking length $b = 0.2 \times \langle \Delta r_p \rangle$, where the mean interparticle spacing $\langle \Delta r_p \rangle \equiv L/N$ for a box with side length L and N^3 particles. The mass of the halo is then the sum of the particle masses.

Using this definition, Jenkins et al. (2001) obtained a universal mass function. Assuming that halos are isothermal spheres, this corresponds to a mean density inside the halo of 180 times the background density, although in practice there is a large scatter about this value (White 2002). For our $100 h^{-1}$ comoving Mpc box with 256^3 particles, $b \approx 0.08 h^{-1}$ comoving Mpc. We require that each halo contains at least 31 particles. Figure 4 shows the mean overdensity profiles of dark matter halos hosting quasars for the fiducial mass range $3.0 \pm 1.6 h^{-1} \times 10^{12} M_{\odot}$ inferred from clustering measurements by Croom et al. (2005), with the corresponding standard deviation at each point. Also shown are analogous curves for some of the least (fiducial mass range divided by ten, $3.0 \pm 1.6 h^{-1} \times 10^{11} M_{\odot}$) and most massive halos (fiducial mass range multiplied by two, $6.0 \pm 3.2 h^{-1} \times 10^{12} M_{\odot}$) in the simulation box. For comparison, the virial radius of a halo is given by⁴

$$r_{vir} \approx 42 \left(\frac{M_{DM}}{10^{12} h^{-1} M_{\odot}} \right)^{1/3} \left(\frac{1+z}{4} \right)^{-1} \text{ kpc} \quad (10)$$

(Barkana & Loeb 2001), which in all cases is much smaller than the radius up to which the halo profiles have a significant impact on the density field, given that most of the Ly α forest arises from fluctuations of order unity. As we show the next section, the halo infall regions extend to even larger distances.

4.2. Infall and Redshift-Space Distortions

In general, neither the quasar nor the absorbing gas is at rest with respect to the Hubble flow. In particular, quasar host halos grow via the infall of matter toward their centers. This is illustrated in Figure 5, in which we show profiles of relative velocity between the quasar and the absorbing gas computed from the simulations. The peculiar velocities have the consequence of shifting the effective redshift of Ly α absorption of gas parcels through the Doppler effect.

⁴ This expression is valid at $z \gtrsim 1$, where the effect of a cosmological constant can be neglected.

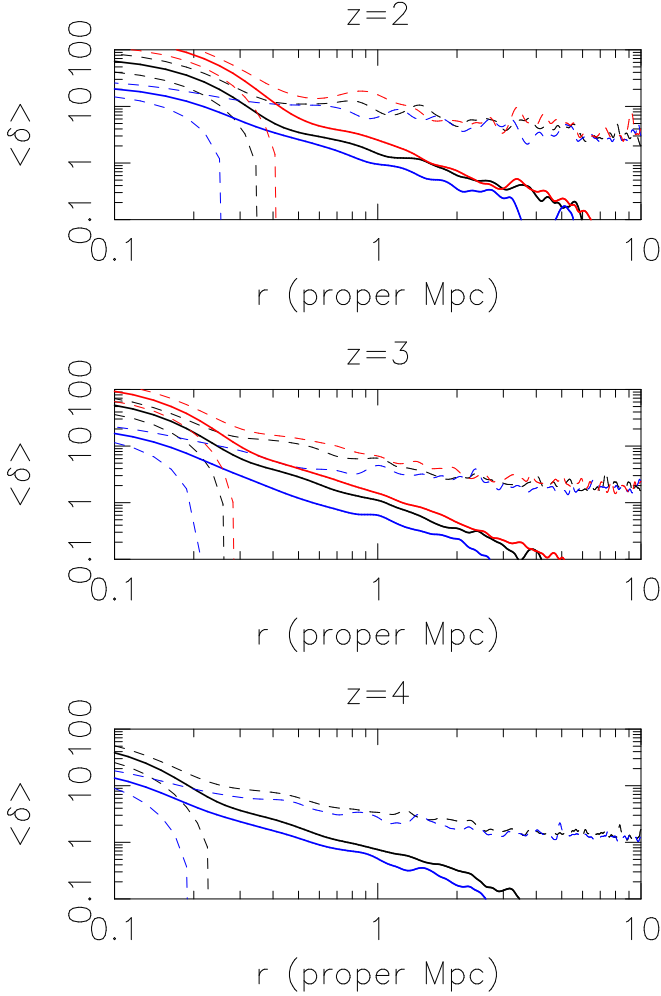


FIG. 4.— Profiles of linear overdensity of dark matter halos at $z = 2, 3$, and 4 (top, middle, and bottom). In each case, 100 lines of sight were drawn from each of 100 halos with mass in the ranges $3.0 \pm 1.6 h^{-1} \times 10^{11} M_{\odot}$ (light, blue), $3.0 \pm 1.6 h^{-1} \times 10^{12} M_{\odot}$ (Croom et al. 2005, black), and $6.0 \pm 3.2 h^{-1} \times 10^{12} M_{\odot}$ (heavy, red) randomly selected in the simulation box. The solid curves show the profiles averaged over all lines of sight and the bounding dashed curves show the sample standard deviations of the density at each point. The profiles extend significantly above the mean cosmic density ($\delta \equiv 0$) at distances greatly exceeding the halo virial radii (~ 50 proper kpc). We do not show the case of the most massive halos at $z = 4$ because the simulation box contains too few halos in this mass range.

Consider the case where the quasar follows the Hubble flow and a gas parcel along the line of sight at actual proper distance r from the quasar has peculiar velocity v_{\parallel} (> 0 for motion toward the observer) along the line of sight. This case is sufficiently general for our purpose, since the peculiar motion of the quasar is simply taken into account by considering its redshift as determined from spectroscopy. The important quantity is the relative velocity of the gas with respect to the quasar along the line of sight; in practice, the velocity of each mock quasar is identified with the velocity of the center of mass of the halo in which it lies.

To first order, the gas parcel will absorb gas that has Ly α frequency (in the quasar rest frame) at proper distance $r' = r + \Delta r$ from the quasar, where $\Delta r = v_{\parallel}/H$. We call r' the “redshift-space” distance. Inserting reasonable numbers ($z_Q = 3$, $v_{\parallel} = 300 \text{ km s}^{-1}$; see Figure

5), we obtain $\Delta r \approx 1$ proper Mpc. The Doppler shift

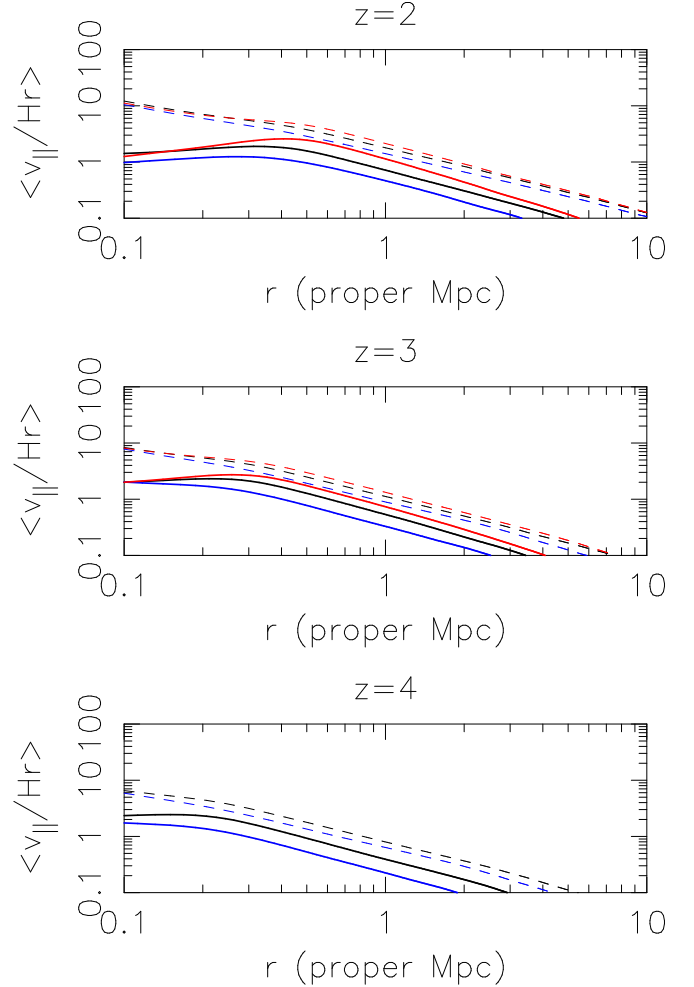


FIG. 5.— Profiles of relative velocity along the line of sight between the gas and the quasar for halos at $z = 2, 3$, and 4 (top, middle, and bottom). In each case, 100 lines of sight were drawn from each of 100 halos with mass in the ranges $3.0 \pm 1.6 h^{-1} \times 10^{11} M_{\odot}$ (light, blue), $3.0 \pm 1.6 h^{-1} \times 10^{12} M_{\odot}$ (Croom et al. 2005, black), and $6.0 \pm 3.2 h^{-1} \times 10^{12} M_{\odot}$ (heavy, red) randomly selected in the simulation box. The solid curves show the profiles averaged over all lines of sight and the bounding dashed curves show the sample standard deviation of the velocity at each point. Shown are the absolute values of the true mean velocities, which are negative owing to systematic infall of the gas toward the halo centers. The velocities are given in units of Hr because in linear theory $v_{\parallel}/Hr \sim \delta$ (c.f. Figure 4) and this is the quantity with which the expected Γ^{bkg} bias in proximity effect analyses scales (§5.4). We do not show the case of the most massive halos at $z = 4$ because the simulation box contains too few halos in this mass range.

effect is thus potentially very significant. We incorporate the effect of peculiar velocities in the usual manner, which we review here briefly for completeness (see e.g., Hui et al. 1997 for more details). Since the ionized fraction of a given gas parcel depends on the photoionization rate at its actual proper distance r from the quasar, we first calculate the neutral hydrogen density field ignoring redshift-space distortions as in §3.2. Because of discretization, we obtain a sequence of neutral densities $\{\rho_0, \rho_1, \dots, \rho_N\}$, with corresponding peculiar velocities $\{v_{\parallel,0}, v_{\parallel,1}, \dots, v_{\parallel,N}\}$ at proper distances

$\{r_0, r_1, \dots, r_N\}$ from the quasar. For each i , we calculate the redshift-space distance $r'_i = r_i + \Delta r_i$, with $\Delta r_i = v_{\parallel,i}/H$, for the gas parcel. The redshift-space density field $\{\rho'_0, \rho'_1, \dots, \rho'_N\}$ is then constructed such that

$$\rho'_i = \sum_{r'_j > 0 \text{ is nearest to } r_i} \rho_j \quad (11)$$

i.e. at each point i of the discretized line-of-sight field we sum the densities of the gas parcels whose redshift-space coordinate is positive (the absorption occurs along the line of sight to the quasar) and nearest to r_i .

The thermal motion of the gas particles also causes redshift-space distortions, broadening the absorption lines owing to the random velocities. The characteristic temperature of the IGM is $T_{IGM} \approx 2 \times 10^4$ K (Ricotti et al. 2000; Schaye et al. 2000; McDonald et al. 2001), corresponding to typical Doppler thermal velocities $v_{rms} \approx \sqrt{3k_B T_{IGM}/m_p} \approx 20 \text{ km s}^{-1}$. This is an order of magnitude less than the typical bulk velocities of the gas parcels owing to peculiar velocities, suggesting that this effect has little direct impact on Γ^{bkg} measurements from the proximity effect. However, it is important to simulate thermal broadening to accurately reproduce the statistical properties of the Ly α forest.

The Doppler profile owing to thermal broadening is given by

$$\phi(\nu) = \frac{1}{\sqrt{\pi}\Delta\nu_D} \exp\left[-\frac{(\nu - \nu_{Ly\alpha})^2}{(\Delta\nu_D)^2}\right], \quad (12)$$

with

$$\Delta\nu_D = \frac{\nu_{Ly\alpha}}{c} \sqrt{\frac{2k_B T}{m_p}} \quad (13)$$

for Ly α absorption by hydrogen atoms (e.g., Rybicki & Lightman 1979). A shift $\Delta\nu$ in frequency corresponds to a shift $\Delta r = c\Delta\nu/H\nu$ in effective proper position of absorption, so that in proper real-space the profile becomes

$$\phi(\Delta r) = \frac{Hc\nu_{Ly\alpha}}{\sqrt{\pi}\Delta\nu_D} \exp\left[-\frac{(Hc\nu_{Ly\alpha}\Delta r)^2}{(\Delta\nu_D)^2}\right]. \quad (14)$$

For each gas parcel we use the temperature T as given by the equation of state (5) for the undistorted density field. For each j , we then assign a fraction $\phi(r_j - r_i)\Delta r_i$ of the neutral density of parcel i to the cell at position r_j , where Δr_i is the proper spacing between neighboring cells. When modeling both peculiar velocities and thermal broadening simultaneously, the numbers of cells by which each partial gas parcel is shifted owing to peculiar motion and thermal broadening are added. Note that we assume a thermal profile here, and hence ignore the natural line width, but this is accurate for the low column density Ly- α forest (e.g., Hui et al. 1997).

4.3. Optical Depth PDF vs. Halo Mass

At this point, it is interesting to take a look at how the optical depth PDF varies with halo mass at different distances from the quasars when the overdensity and gas infall associated with dark matter host halos are modeled as above. This is illustrated in Figure 6, in which we show results for dark

matter halos of mass $3.0 \pm 1.6 h^{-1} \times 10^{11} M_\odot$ (light), $3.0 \pm 1.6 h^{-1} \times 10^{12} M_\odot$ (Croom et al. 2005), and $6.0 \pm 3.2 h^{-1} \times 10^{12} M_\odot$ (heavy) at $z = 3$, along with fits to the trend of mean $\ln \tau$ with halo mass for the range of masses probed by our simulation. We also show the case in which quasars lie in random locations, which can be clearly distinguished from the massive halos. The optical depth PDFs are shifted to higher values with increasing halo mass, which is expected since the matter overdensity should increase with halo mass. The significant effect of quasar environments on the optical depth statistics in the proximity regions of quasars suggests that existing proximity effect analyses which have neglected these will not necessarily recover the correct Γ^{bkg} . We make analytic estimates of the expected biases in §5.3 and quantify these more accurately with mock spectra in §5.5. The environmental dependence of the optical depth statistics also raises the possibility of using the latter to probe quasar environments, in particular measure the mass of the host halos. We focus attention to this exciting possibility in §7. For the moment, we note that the effect of halos on the optical depth statistics is to “separate” the PDFs with mass. This is non-degenerate with varying Γ^{bkg} , which just shifts PDFs of $\ln \tau$ by the same amount, regardless of the halo mass⁵. We must note that although the optical depth statistics near quasars differ by large *factors* for different host halo masses, the absolute optical depth differences are small, of order $\Delta\tau \sim 0.01$. Measuring these differences in actual spectra, with noise and for which the continuum level must be estimated from the data, may be challenging in practice, as we discuss in §8.

4.4. Clustering

A possible bias related to overdensities is the clustering of galaxies and other quasars around proximity effect quasars. If quasars indeed form in large-scale overdensities, then the probability of other quasars and galaxies forming in the vicinity is increased with respect to average regions of the Universe, because there it is easier for local density peaks (“peaks within peaks”) to cross the threshold for collapse (Bond et al. 1991). The clustering of galaxies around quasars has in fact been both predicted in simulations of galaxy and quasar formation (e.g., Kauffmann & Haehnelt 2002) and observationally measured over a wide redshift interval (e.g., Croom et al. 2004 at $z < 0.3$; Coil et al. 2006 at $z \sim 1$; Adelberger & Steidel 2005a at $1.5 \lesssim z \lesssim 3.5$). Quasars, which are generally thought to be hosted by galaxies, have been shown to exhibit similar clustering among themselves (e.g., Croom et al. 2001, 2004; Wake et al. 2004; Croom et al. 2005).

We have thus far identified Γ^{bkg} with the total photoionization rate near quasars contributed by all *other* light sources, implicitly assuming that this quantity is equal to the total photoionization rate averaged over large regions of the Universe (the “true” background rate). Since galaxies and quasars (which together presumably dominate the contribution to the photoionizing background) cluster around quasars, we in fact expect Γ^{bkg} near quasars to be higher than the true background

⁵ Ignoring redshift-space distortions, $\tau \propto n_{\text{HI}} \propto (\Gamma^{tot})^{-1}$ (§3.2).

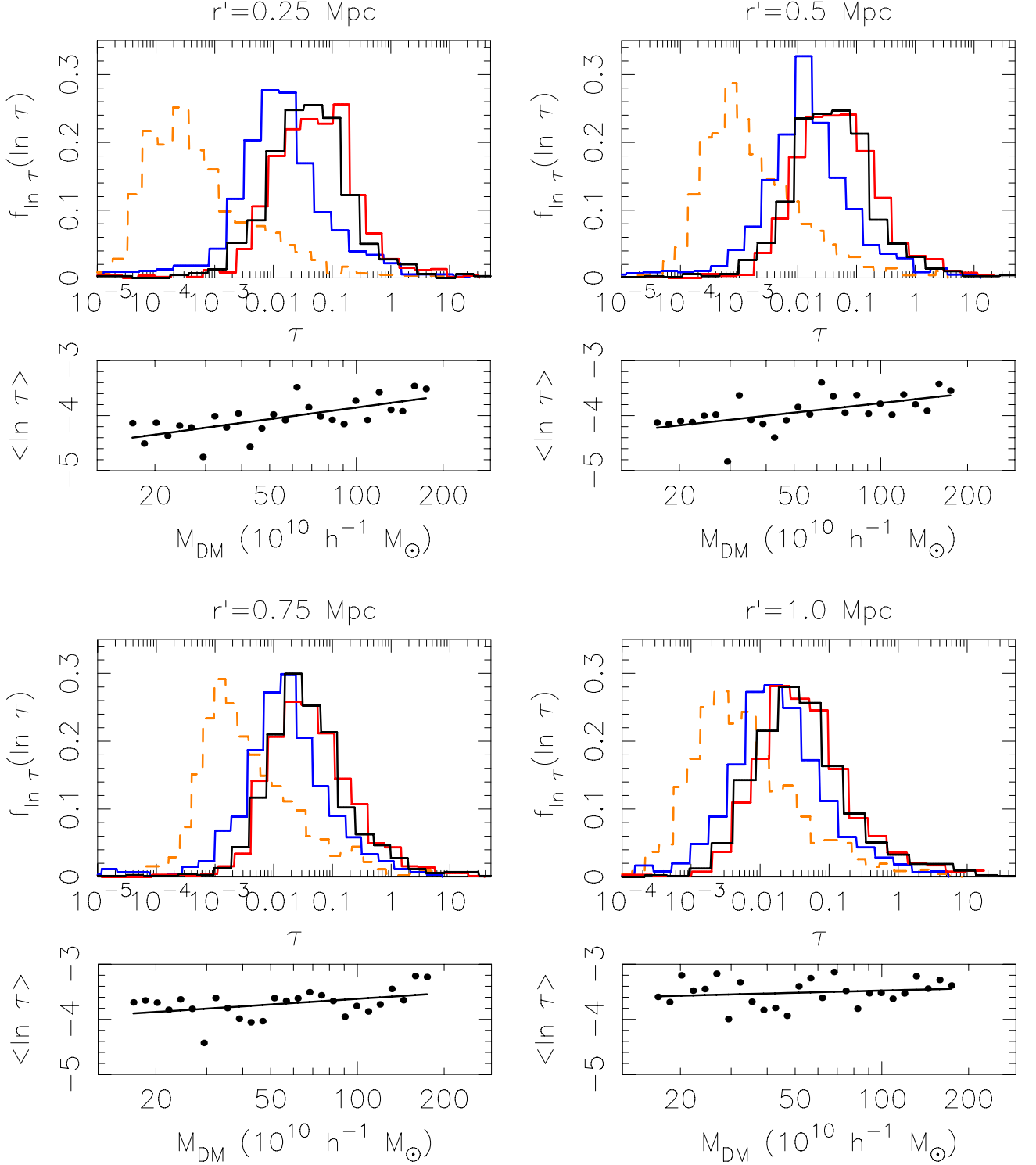


FIG. 6.— Optical depth PDFs at increasing redshift-space proper distance from quasars of typical luminosity at $z = 3$. In each case, the solid-line histograms show PDFs approximated from 10 lines of sight drawn from each of 100 halos with masses in the ranges $3.0 \pm 1.6 h^{-1} \times 10^{11} M_{\odot}$ (light, blue), $3.0 \pm 1.6 h^{-1} \times 10^{12} M_{\odot}$ (Croom et al. 2005, black), and $6.0 \pm 3.2 h^{-1} \times 10^{12} M_{\odot}$ (heavy, red) randomly selected in the simulation box. The dashed orange PDFs show the case of quasars lying in random locations in the box. This situation can be clearly distinguished from the cases where the quasars lie in massive dark matter halos. Note, however, that although the optical depth statistics near quasars differ by large *factors* for different host halo masses, the absolute optical depth differences are small, of order $\Delta\tau \sim 0.01$. We discuss in §8 observational difficulties which make such differences challenging to measure in practice. In general, the optical depth PDFs are shifted to higher values as the halo mass is increased. Below each panel, we show how the sample mean $\ln \tau$ varies with mass for 10 lines of sight from each of 10 halos in bins of logarithmic width $\Delta \ln M_{DM} = 0.05$ (10 times finer than the fiducial Croom et al. mass range) along with the least-squares log-log linear fit. The slope of the linear fit compared to the scatter between the dots provides an estimate of how well halos of different masses can be distinguished with 100 lines of sight, using only one data point at the given distance from each spectrum. The results of full mock mass likelihood analyses are presented in §7.

Γ^{bkg} . In this section, we quantitatively estimate this effect.

Let us first suppose that the ionizing background is dominated by emission from star-forming LBGs, as may be the case at $z \sim 3$ (e.g., Steidel et al. 2001). Consider the photoionizing flux at a given position owing to all sources of light in the Universe *other* than the quasar in an idealized model of N_G isotropic sources, each with specific luminosity $L_{\nu,G}$, distributed in a spherical volume $V = 4\pi R^3/3$ with quasar-galaxy correlation function

$$\zeta_{QG}(r) = \left(\frac{r}{r_0}\right)^{-\gamma}. \quad (15)$$

Cosmological expansion can be neglected in the argument, since for $\gtrsim 2$ the radiation is largely local, i.e. sources at higher redshifts are absorbed and can be neglected (Madau et al. 1999). Let n_G be the mean number density of sources and $f_{\nu,G} = L_{\nu,G}/4\pi r^2$ be the specific flux at distance r from each source. Suppose that a quasar is located at the origin of the volume. The total specific flux owing to the other sources at a point displaced by the vector \mathbf{r} from the quasar is

$$f_{\nu,G}^{other} = \sum_{i=1}^{N_G} f_{\nu}^i = \frac{L_{\nu,G}}{4\pi} \sum_{i=1}^{N_G} |\mathbf{r} - \mathbf{r}_i|^{-2}, \quad (16)$$

where the sums are over the sources other than the quasar and \mathbf{r}_i is the position of source i . Since we will eventually use this quantity for the sole purpose of calculating a photoionization rate, we need not worry that light rays incident at \mathbf{r} are not necessarily parallel, i.e. the scalar sum of the specific fluxes contains all the relevant information. To first order, we have, averaging over \mathbf{r}_i ,

$$\langle f_{\nu,G}^{other} \rangle(r) = \frac{N_G L_{\nu,G}}{4\pi} \langle |\mathbf{r} - \mathbf{r}_i|^{-2} \rangle, \quad (17)$$

where

$$\langle |\mathbf{r} - \mathbf{r}_i|^{-2} \rangle = V^{-1} \int_0^R \int_0^\pi \int_0^{2\pi} d(|\mathbf{r} - \mathbf{r}_i|) d\theta d\phi \sin \theta \quad (18)$$

$$\times [1 + \zeta_{QG}(r_i)]. \quad (19)$$

The integral above is over a sphere of radius R centered on \mathbf{r} : the $|\mathbf{r} - \mathbf{r}_i|^2$ factor of the spherical volume element cancels the $|\mathbf{r} - \mathbf{r}_i|^{-2}$ term that is integrated over. Using $r_i^2 = r^2 + |\mathbf{r} - \mathbf{r}_i|^2 - 2r|\mathbf{r} - \mathbf{r}_i|\cos\theta$ (for a suitable choice for the orientation of the coordinate system), we may evaluate

$$\langle f_{\nu,G}^{other} \rangle(r) = \langle f_{\nu,G}^{other,smooth} \rangle + \langle f_{\nu,G}^{other,clust} \rangle(r), \quad (20)$$

where

$$\langle f_{\nu,G}^{other,smooth} \rangle = n_G L_{\nu,G} R \quad (21)$$

is the contribution of the smooth background component (arising from the factor 1 in eq. 18) and

$$\langle f_{\nu,G}^{other,clust} \rangle = \frac{n_G L_{\nu,G}}{2} \int_0^R \int_0^\pi d(|\mathbf{r} - \mathbf{r}_i|) d\theta \sin \theta \zeta_{QG}(r_i) \quad (22)$$

is the component (arising from the ζ_{QG} term in eq. 18) owing to the clustering of galaxies around the quasar. For both the smooth and clustering components, we define

the photoionization rates as calculated from the average contributions:

$$\Gamma^{bkg,x} \equiv \int_{\nu_{Ly}}^\infty d\nu \sigma(\nu) \frac{\langle f_{\nu,G}^{other,x} \rangle}{h\nu}. \quad (23)$$

Are the local fluctuations in the photoionizing flux owing to the quasar-galaxy correlations important when attempting to measure Γ^{bkg} using the proximity effect? To answer this question, we note that the proximity effect is sensitive to $\Gamma^{bkg} = \Gamma^{bkg,smooth} + \Gamma^{bkg,clust}$ and consider the ratio:

$$\frac{\Gamma^{bkg,clust}}{\Gamma^{bkg,smooth}}(R; r) = \frac{\int_0^R \int_0^\pi d(|\mathbf{r} - \mathbf{r}_i|) d\theta \sin \theta \zeta_{QG}(r_i)}{2R}. \quad (24)$$

Note that this ratio is purely geometrical, depending on only the attenuation radius and the quasar-galaxy correlation function. In particular, it is independent of the abundance and brightness of the galaxies. The cancellations occur because we are assuming that the smooth background and the fluctuations are produced by the same sources.

In the top panel, Figure 7 shows $\Gamma^{bkg,smooth}$, $\Gamma^{bkg,fluct}$, and Γ^{QSO} as a function of distance from the quasar for the order-of-magnitude estimate $\Gamma^{bkg} = 10^{-12} \text{ s}^{-1}$ (see Figure 1) and quasars of typical luminosities at $z = 3$. For the quasar-galaxy correlation function, we adopt $r_0 = 5 \text{ h}^{-1}$ comoving Mpc and $\gamma = 1.6$, consistent with the results of Adelberger & Steidel (2005b) for LBGs at $z \sim 3$. The characteristic value of the R is the mean free path l_{mfp} of ionizing photons, i.e. the radius beyond which incoming ionizing photons are exponentially suppressed. Madau et al. (1999) calculate $l_{mfp} \approx 33[(1+z)/4]^{-4.5}$ proper Mpc for Lyman limit photons, which we take as characteristic of ionizing soft galactic emission.

In the bottom panel, the ratios of $\Gamma^{bkg,fluct}$ to $\Gamma^{bkg,smooth}$ and Γ^{QSO} are shown. It is seen that $\Gamma^{bkg,fluct} \ll \Gamma^{bkg,smooth}$ ($< 20\%$ for $r > 1$ Mpc and $< 5\%$ for $r > 7$ Mpc, which is much less than the factor $\gtrsim 2$ discrepancy between proximity effect and flux decrement measurements of Γ^{bkg}), except at very small distances from the quasar, where it diverges. This divergence is most likely benign, as it occurs on a scale much smaller than the proximity region (i.e., the region where $\Gamma^{bkg} \sim \Gamma^{QSO}$, which contains significant information about Γ^{bkg}) and in any case is an artifact of assuming that the correlation function (eq. 15) maintains its power-law behavior as $r \rightarrow 0$. In reality, $\Gamma^{bkg,fluct}$ is expected to be bounded above on physical grounds. Moreover, note that in the region where $\Gamma^{bkg,fluct} \sim \Gamma^{bkg,smooth}$, the quasar flux already dominates over the locally-enhanced background flux by a large factor. We thus conclude that clustering of galaxies around quasars is unlikely to significantly affect measurements of Γ^{bkg} using the proximity effect, at least at $z = 3$ for a background which is dominated by the contribution from LBGs.

What if quasars contribute significantly to Γ^{bkg} and how is the situation modified at different redshifts? Consider the case of a background dominated by quasars. Then we must consider the quasar-quasar correlation function, but it is in fact similar to the quasar-LBG one (e.g., Croom et al. 2005). Let us denote it generically by $\zeta(r)$. The ratio in equation (24) is then only modified by

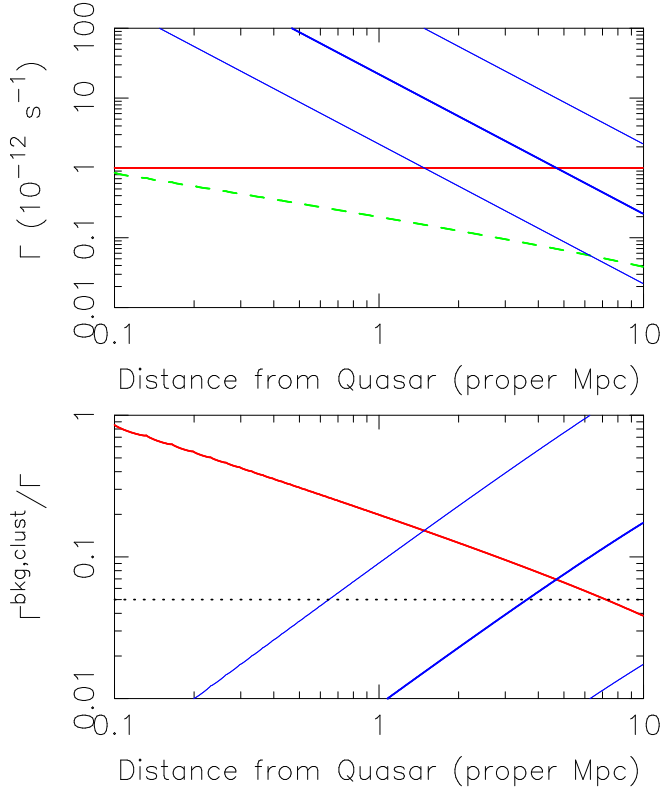


FIG. 7.— Top: Comparison of the contributions to the total hydrogen photoionizing rate owing to the background ($\Gamma^{bkg} = 10^{-12} \text{ s}^{-1}$; horizontal red line), the quasar (solid blue curves; the central thick curve corresponds to the mean typical quasar luminosity and the bounding thin curves indicates ranges of one standard deviation), and the local excess of galaxies owing to AGN-galaxy clustering (assuming that the background is dominated by emission from LBGs; green dashed curve) as a function of distance from the quasar at $z = 3$. Bottom: Ratios of the local contributions from the quasar (blue) and the clustering of galaxies around the quasar (red) to the true photoionizing background. The horizontal dotted line shows the 5% level.

the attenuation length R . For a galaxy-dominated background, we took R to be the mean-free path to Lyman-limit photons, but the appropriate R for quasars will be larger owing to their harder spectra and the longer mean free paths for high energy quasar photons. Since $\zeta(r)$ is a strictly decreasing function of r ,

$$\left(\frac{\Gamma^{bkg,clust}}{\Gamma^{bkg,smooth}} \right)_{QQ} < \left(\frac{\Gamma^{bkg,clust}}{\Gamma^{bkg,smooth}} \right)_{QG}, \quad (25)$$

where the subscripts QQ and QG refer to the quasar- and galaxy-dominated cases, respectively. The clustering effect is therefore even less important in the quasar-dominated case. For the general case in which both galaxies and quasars contribute significantly to the background, Γ^{bkg} is a linear superposition of the galaxy and quasar contributions and so the total clustering contribution is again negligible. For redshifts $z < 3$, the attenuation length is increased because the number density of absorbers is decreased by cosmic expansion, so the result still holds. In the limit of large redshifts, $R \rightarrow 0$ and $\Gamma^{bkg,clust}/\Gamma^{bkg,smooth} \rightarrow \infty$, so that the result must eventually break down. However, we are here mainly concerned with explaining the discrepancy between proximity effect and flux decrement measurements of Γ^{bkg} at $z \lesssim 3$ (Figure 1). The argument given so far indicates

that clustering has a negligible effect and we are therefore not compelled to model it and to pursue a further analysis of the issue.

5. BIASES IN PROXIMITY EFFECT MEASUREMENTS OF Γ^{bkg}

To understand why the proximity effect measurements of Γ^{bkg} are systematically higher than the flux decrement measurements of the same quantity, we first formulate a statistical method for estimating Γ^{bkg} from the proximity effect, with physical assumptions mimicking those made in previous proximity effect measurements of Γ^{bkg} . In contrast to most previous analyses, we consider the optical depth statistics in the proximity region of quasar spectra instead of the differential number of absorption lines as a function of redshift, dN/dz . Following the seminal work of Bajtlik et al. (1988), the differential number of absorption lines has been used in almost all studies of the proximity effect to date. However, absorption lines arising from discrete absorbers are no longer consistent with the modern picture of the Ly α forest as arising from smooth density fluctuations in the IGM. Moreover, the counting of absorption lines is a somewhat ill-defined and uncertain procedure, making it difficult to rigorously treat statistically. On the other hand, once the quasar continuum is estimated, the optical depth at any point can be simply measured as $\tau = -\ln F^{obs}/F^{cont}$ ⁶, where F^{obs} is the observed flux and F^{cont} is the continuum level. As we will see, it is also easy to treat the optical depth in a statistically sound framework.

To estimate the biases introduced by the local overdensities around and the infall of gas toward the quasars (the two effects which are most likely to play an important role according to our discussion in §3.2) in analyses which have neglected those complications, we apply our measurement method to mock spectra (with a known value of Γ^{bkg}) in which those effects are modeled. We begin with some remarks on the optical depth probability density function (PDF) in the general Ly α forest in §5.1 and formulate our statistical formalism in §5.2. In §5.3, we provide order-of-magnitude analytic estimates of the expected biases, before quantifying those more accurately using mock spectra in §5.5.

5.1. Lyman- α Forest Optical Depth PDF

Before proceeding with defining a method for inferring Γ^{bkg} from optical depth statistics in the proximity region, it is useful to consider the optical depth PDF in the general Ly α forest. Coles & Jones (1991) first used a lognormal PDF to describe the dark matter density field and showed that it follows plausibly from the assumption of initial linear Gaussian and velocity fluctuations. Bi et al. (1992) and Bi & Davidse (1997) applied this approach to baryons also and showed that the resulting Ly α forest matched the observations well. Several observational studies, e.g. Becker et al. (2006) and Desjacques et al. (2007), have also found that the optical depth PDF in the Ly α forest is close to lognormal.

Figure 8 shows normalized histograms of the optical depth constructed from random lines of sight through

⁶ Because of the redshift-space distortions, this quantity is not directly proportional to HI density as in equation (9). It is best interpreted as a redshift-space effective optical depth, or taken as the definition of the optical depth in this work.

our simulation box at $z = 2, 3$, and 4 , along with the corresponding lognormal distribution as estimated from the mean logarithm of the optical depth and its standard deviation. Although there are noticeable deviations – es-

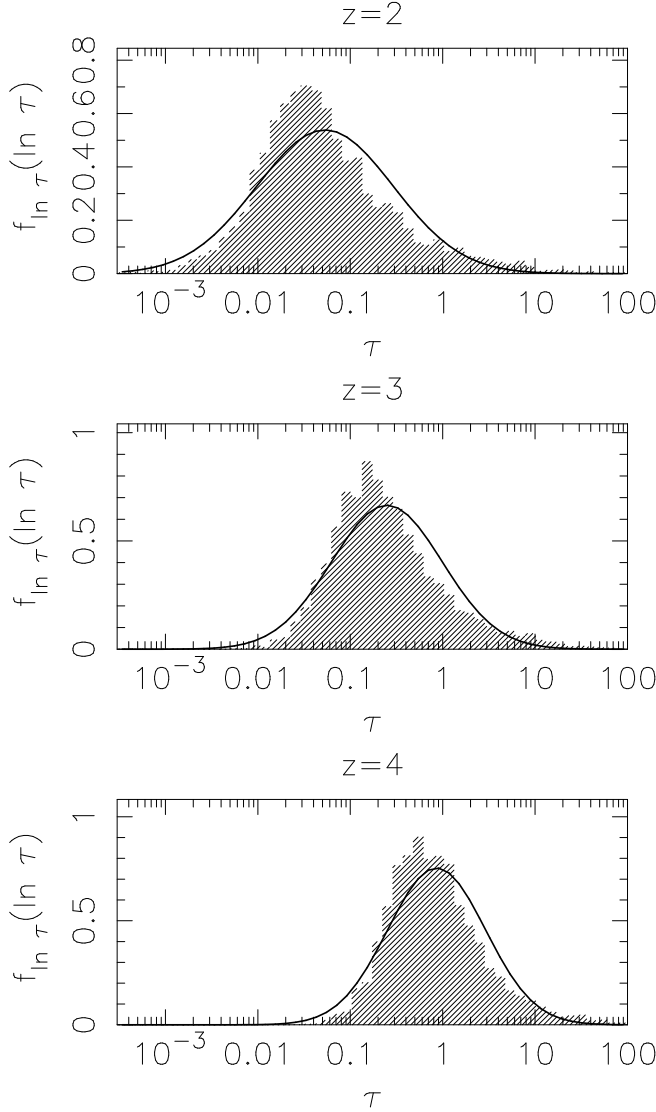


FIG. 8.— Histograms of the optical depth as constructed from random lines of sight in the simulation box at $z = 2, 3$, and 4 (top, middle, and bottom panels). Redshift-spaces distortions have been modeled as in §4.2. The solid curves show the corresponding lognormal distribution as estimated from the mean logarithm of the optical depth and its standard deviation.

pecially at low redshifts - of the simulated optical depth distribution from lognormal shape, the main features of the distributions are well-captured by a lognormal. The lognormal provides a convenient analytic form, significantly simplifying the discussion, and so we will use it in making estimates of the biases that arise when naively measuring Γ^{bkg} from the proximity effect. Our results should be robust to this assumption as, in this analysis, the information about Γ^{bkg} is contained in the *scaling* equation (1).

5.2. Maximum Likelihood Method for Γ^{bkg}

We now formulate a maximum likelihood method to measure Γ^{bkg} from optical depth statistics in the proximity regions of Ly α absorption spectra. While the statistical formalism differs from the BDO line-counting method, we make essentially the same physical assumptions. Thus, our conclusions should also provide insight in the validity of studies which used the line-counting approach, i.e. virtually all existing proximity effect analyses. Mathematically, the explicit assumption is that that the Ly α optical depth in the presence of the quasar, τ^{prox} , is related to the optical depth if the quasar were turned off, τ^{off} , by

$$\tau^{prox} = \frac{\tau^{off}}{1 + \omega(Q; r)}, \quad (26)$$

where $\omega(Q; r) = \Gamma^{QSO}(Q; r)/\Gamma^{bkg}(z_Q)$, $\Gamma^{QSO}(Q; r)$ is the contribution to the total photoionization rate of the quasar at distance r from it, and $Q = \{z_Q, L_Q, \alpha_Q\}$ specifies the redshift, luminosity and spectral index of the quasar under consideration. We call this method the “ τ -scaling” likelihood analysis. Owing to the small size of the proximity region, r can be taken to be the proper distance. Equation (26) is just what would be obtained from photoionization equilibrium assuming that the quasar lies in a random location and neglecting all redshift-space distortions. In what follows, we simplify the notation and use τ to denote τ^{off} .

Using equation (26), we can relate the optical depth statistics in the Ly α forest away from the quasar, but at approximately equal redshift ($z \approx z_Q$), to those inside the proximity region. For a lognormal τ distribution with mean logarithm $\langle \ln \tau \rangle$ and standard deviation, also in the logarithm, $\sigma_{\ln \tau}$, we have:

$$f_{\hat{\tau}}(z_Q; \tau) = \frac{1}{\sqrt{2\pi}\sigma_{\ln \tau}} \exp \left[-\frac{(\langle \ln \tau \rangle - \ln \tau)^2}{2\sigma_{\ln \tau}^2} \right], \quad (27)$$

which translates to lognormal distribution in the proximity region given by:

$$f_{\hat{\tau}^{prox}}(Q, r; \tau^{prox}) = \frac{1 + \omega(Q; r)}{\sqrt{2\pi}\sigma_{\ln \tau} \tau^{prox}} \times \exp \left[-\frac{(\langle \ln \tau \rangle - \ln[(1 + \omega(Q; r))\tau^{prox}])^2}{2\sigma_{\ln \tau}^2} \right]. \quad (28)$$

We define a correlation length, r_{corr} , such that points separated by this distance can be assumed to have independent optical depths. Then, for a given quasar spectrum, we estimate $\langle \ln \tau \rangle$ and $\sigma_{\ln \tau}$ using the usual unbiased estimators

$$\widehat{\langle \ln \tau \rangle} = \frac{1}{N} \sum_{i=1}^N \ln \tau_i \quad (29)$$

and

$$\widehat{\sigma_{\ln \tau}^2} = \frac{1}{N-1} \sum_{i=1}^N \left(\ln \tau_i - \widehat{\langle \ln \tau \rangle} \right)^2, \quad (30)$$

where τ_i are optical depths at points separated by r_{corr} outside the proximity region.

We may now construct the likelihood function

$$\mathcal{L}[\Gamma^{bkg}(z)] \equiv \prod_Q \left\{ \prod_{prox \ pts} f_{\hat{\tau}^{prox}}(\tau^{prox}) \right\}, \quad (31)$$

where $f_{\tau^{prox}}(\tau^{prox})$ has the same parameter dependence as in equation (28). Here, the outer product is over the different quasars in a sample. The inner product is over data points in the proximity region of quasar Q (defined to be between r_{min} and r_{max} from the quasar), again separated by a proper distance r_{corr} . Each factor is proportional to the probability of obtaining a realization of the optical depth given the adopted model for Γ^{bkg} . The terms in the products are assumed to be evaluated at independent points, so that the likelihood is proportional to the probability of obtaining all the measured optical depths given the model. By Bayes's theorem, this is proportional to the probability that the assumed model is correct. After normalization, the likelihood function gives the PDF for a given model for Γ^{bkg} to be correct. The likelihood estimator in equation 31 is suboptimal, since it does not make use of correlated data points. The estimated likelihood may therefore be wider (less constraining) than could be achieved in principle. The design of an optimal estimator is outside of the scope of this paper. In §8, we discuss practical considerations arising with actual quasar spectra of finite resolution and signal-to-noise ratio, which are likely to be more important than the optimality of the likelihood estimator.

5.3. Analytic Estimates of Biases

Before applying the τ -scaling likelihood to mock spectra, we make order-of-magnitude estimates for the biases that may be expected when the method is applied to spectra with quasars lying in overdense regions and with redshift-space distortions. Our estimates are based on the fact that the likelihood analysis is sensitive to the $(1 + \omega(Q; r))^{-1}$ scaling assumed for τ in the proximity region (eq. 26). The local overdensities and infall regions around quasars distort this scaling, leading to an incorrect estimate of ω . Assuming that $\Gamma^{QSO}(r)$ is known (e.g., from measurements of the magnitude and spectrum of the quasar), this leads to an incorrect estimate of Γ^{bkg} .

5.3.1. Bias Owing to Local Overdensity

Consider first the case where the quasar lies in an overdense region, but where redshift-space distortions are ignored. Let us denote by $\omega^{true}(r)$ the true ratio $\Gamma^{QSO}(Q; r)/\Gamma^{bkg}$ and by $\omega^{app}(r)$ the ratio as inferred when assuming the scaling given by equation (26). Denote the overdensity $\Delta(r) \equiv \rho(r)/\langle\rho\rangle$. Then, using $\tau \propto n_{HI} \propto \Delta^{2-0.7\beta}$ (eq. 3, 4, and 5), we expect the optical depth scaling to be approximately modified to

$$\tau^{prox} = \frac{\Delta^{2-0.7\beta}\tau}{1 + \omega}. \quad (32)$$

Assuming the scaling of equation (26) amounts to incorrectly assuming that $\Delta \equiv 1$, from which we obtain the relation

$$\Delta^{2-0.7\beta}(1 + \omega^{true})^{-1} = (1 + \omega^{app})^{-1}, \quad (33)$$

and hence

$$\frac{\Gamma^{bkg,true}}{\Gamma^{bkg,app}} = \frac{\omega^{app}}{\omega^{true}} = \Delta^{0.7\beta-2}(1 + \omega^{true-1}) - \omega^{true-1}. \quad (34)$$

The typical distance r_{typ} from the quasar at which information about Γ^{bkg} is encoded in the proximity effect

is such that the photoionizing rate owing to the quasar is comparable to the background. We take it to satisfy $\omega^{true}(r_{typ}) = 1$, so that

$$\frac{\Gamma^{bkg,true}}{\Gamma^{bkg,app}} = 2\Delta^{0.7\beta-2} - 1. \quad (35)$$

It will be useful to express the bias in terms of the linear overdensity $\delta = \Delta - 1$; taking the reciprocal of equation (35), we have the bias owing to the local overdensity to first order in δ :

$$B^{od} \equiv \frac{\Gamma^{bkg,app}}{\Gamma^{bkg,true}} \approx 1 + 3.1\delta(r_{typ}) \quad (36)$$

for $\beta = 0.62$, the late reionization limit.

5.3.2. Bias Owing to Gas Infall

Consider now the bias owing to the redshift-space distortions caused by gas infall toward the centers of halos. To simplify the estimate, we ignore the local overdensity here. As explained in §4.2, a gas parcel at proper distance r from the quasar and with velocity $v_{||}$ along the line of sight will, owing to the Doppler effect, appear to be absorbing from a distance $r' = r + \Delta r$ from the quasar, with $\Delta r = v_{||}/H$. In this case, assuming the scaling of equation (26) results in the incorrect identification

$$\omega^{true}(r) = \omega^{app}(r'), \quad (37)$$

which gives, using $\Gamma^{QSO}(r) \propto r^{-2}$, the bias owing to gas infall

$$B^{infall} \equiv \frac{\Gamma^{bkg,app}}{\Gamma^{bkg,true}} = \left(\frac{r}{r'}\right)^2 \approx 1 - 2\frac{v_{||}(r_{typ})}{Hr_{typ}}. \quad (38)$$

5.4. Relation Between Overdensity and Infall Biases

The density and velocity fields are related by mass conversion through the continuity equation. We thus expect B^{od} and B^{infall} to also be simply related. In this section, we establish this connection. For $\delta \ll 1$, we may use the linear theory continuity equation

$$\frac{\partial\delta}{\partial t} + \nabla \cdot \mathbf{v} \approx \frac{\partial\delta}{\partial t} + \frac{\partial v_{||}}{\partial r} = 0, \quad (39)$$

where t is proper time, and the spatial derivatives are proper-coordinate derivatives. For the second equality, we have assumed pure radial infall toward the center of the halo and neglected spherical-geometry corrections to the divergence which are important only at small radii. Within the spherical collapse model (Gunn & Gott 1972) and for given initial conditions, one could write an exact relation between δ and $v_{||}$ from equation (39). However, a simple expression for this relation is not available. We thus simply proceed with an order-of-magnitude estimate, expressing derivatives in terms of characteristic scales at r_{typ} :

$$\frac{|\delta|}{t_{ff}} \sim \frac{|v_{||}|}{r_{typ}}, \quad (40)$$

where the free-fall time for a spherically symmetric mass configuration of interior mean density $\bar{\rho}$ is given by

$$t_{ff} = \frac{1}{4}\sqrt{\frac{3\pi}{2G\bar{\rho}}}. \quad (41)$$

In fact the interior mean density exceeds the mean cosmic density around a halo, for otherwise it would not

have collapsed. In spite of this, we use the cosmic mean density in the denominator of equation (41), since using the full overdensity would result only in a second order correction to $|\delta|/t_{ff}$ (eq. 40).

Solving for $|v_{\parallel}|$ in equation (40) and substituting in equation (38) for the bias owing to infall, we obtain

$$B^{infall} \approx 1 + 1.57f\delta(r_{typ}). \quad (42)$$

The factor f , of order unity, has been introduced to allow for the approximations made. In deriving equation (42), we have made use of the Friedmann equation which expresses the Hubble parameter in terms of $G\bar{\rho}$,

$$H = \sqrt{\frac{8\pi G\bar{\rho}}{3}} \quad (43)$$

(again neglecting the cosmological constant, which is valid at the redshifts of interest), and of the fact that $v_{\parallel} < 0$. Comparing the expressions for B^{od} and B^{infall} (eq. 36 and 42) in terms of $\delta(r_{typ})$, we see that both biases scale similarly with $\delta(r_{typ})$. The total bias can be estimated at $B^{tot} \approx B^{od} \times B^{infall} \approx 1 + 4.67\delta(r_{typ})$ (for $f = 1$). At $z = 3$, $r_{typ} = 5$ proper Mpc for a quasar of typical luminosity (Figure 7) and $\langle\delta(r_{typ})\rangle = 0.1$, with a large point-to-point dispersion of $\sigma_{\delta} = 2$. One may thus reasonably choose $\delta = 0.32$, obtaining $B^{tot} \approx 2.5$, and quantitatively make sense of the results of the detailed calculations of the next section.

5.5. Biased Γ^{bkg} Likelihoods

To more accurately estimate the biases that could be present in proximity effect measurements of Γ^{bkg} , we perform the full likelihood analysis of §5.2 on mock spectra with known Γ^{bkg} and quasars which reside at the centers of dark matter halos with masses in the fiducial range $3.0 \pm 1.6 h^{-1} \times 10^{12} M_{\odot}$. We consider both the case in which only the local overdensity is modeled in the mock spectra and the realistic case in which the redshift-space distortions owing to gas infall and thermal broadening are also simulated. For the likelihood analysis, we choose the parameters $(r_{min}, r_{max}, r_{corr}) = (1, 20, 1)$ proper Mpc so as to include the region where the proximity effect is important for typical quasars (e.g., Figure 3). The correlation length is selected based on measurements of the flux correlation function and power spectrum in the Ly α forest, which indicate small correlations on approximately this scale (McDonald et al. 2000, 2006). In Figure 9, we show the ratio of the correlation function of $\ln \tau$ to its variance, ϵ , computed from the simulation box (at random locations) at $z = 2, 3$, and 4. The ratio is $< 15\%$ at $z = 4$ and 3, and $< 30\%$ at $z = 2$ for $r' > 1$ proper Mpc. In Appendix D, we develop a toy model to estimate the effect of correlations on estimates of the likelihood that ignore them, as we calculate. We show there that the error on the width of estimated likelihood, with respect to the true likelihood (which would take into account the correlations), is a factor $\approx (1 - \epsilon/2)$ (for $\epsilon \ll 1$, neglecting correlations between nonadjacent points), independent of the total number of (independent) spectra used.

The resulting likelihood functions, computed on mock data sets of 100 typical spectra at $z = 2, 3$, and 4, are shown in Figure 10. In each case, the maximum-

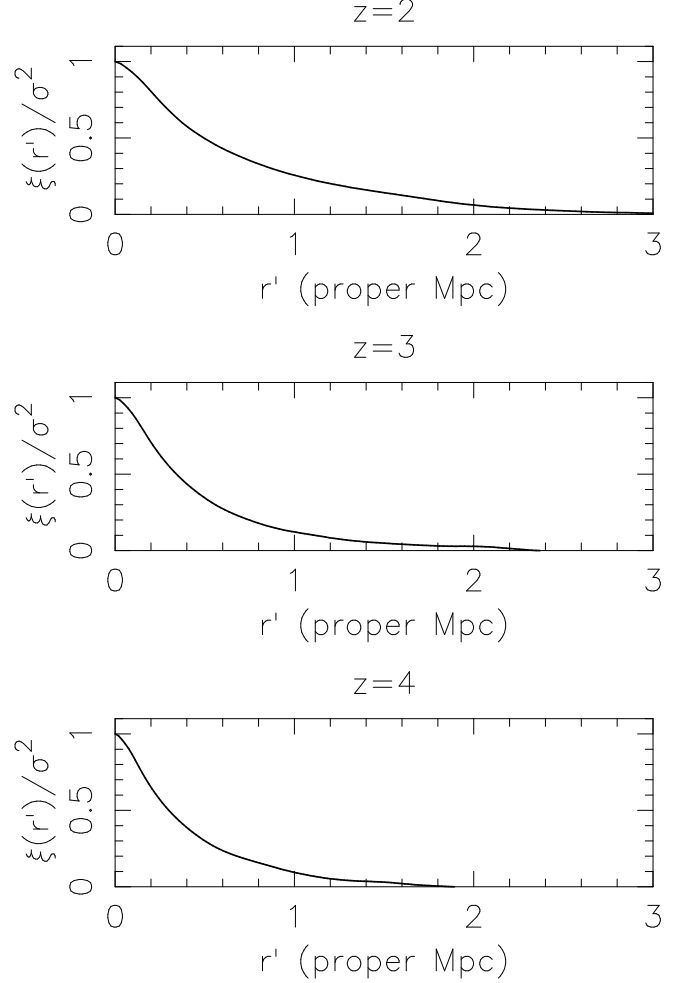


FIG. 9.— Ratio of the correlation function of $\ln \tau$ to its variance computed from the simulation at $z = 2, 3$, and 4 (top, middle, and bottom). The ratio is $< 15\%$ at $z = 2$ and 3, and $< 30\%$ at $z = 4$ for $r' > 1$ proper Mpc.

likelihood Γ^{bkg} overestimates the true Γ^{bkg} , as qualitatively expected from the analytic estimates of §5.3. Moreover, the total bias in the central ($z = 3$) panel agrees quantitatively with the analytic estimate for that redshift. The “typical” $\delta(r_{typ})$ used for this estimate was in fact adjusted to quantitatively reproduce the order-of-unity bias shown here; the essential point is that the bias can be understood to order of magnitude. Most importantly, the above shows that the fact that quasars lie in massive host halos, associated with local matter overdensities and gas infall, *alone* could bias proximity effect measurements of Γ^{bkg} high by a factor ~ 2.5 at $z \approx 3$. These effects should thus be taken into account in future proximity effect analyses.

6. UNBIASED MONTE CARLO LIKELIHOOD FOR Γ^{bkg}

6.1. Method

The likelihood analysis in the previous section is biased because the scaling of equation (26) ignores the local matter overdensity and redshift-space distortions induced by quasar host halos. The likelihood method can however be modified to take these effects into account and yield an unbiased measurement of Γ^{bkg} . To do so, it suffices to replace expression (28) for $f_{\hat{\tau}prox}(\tau^{prox})$ in

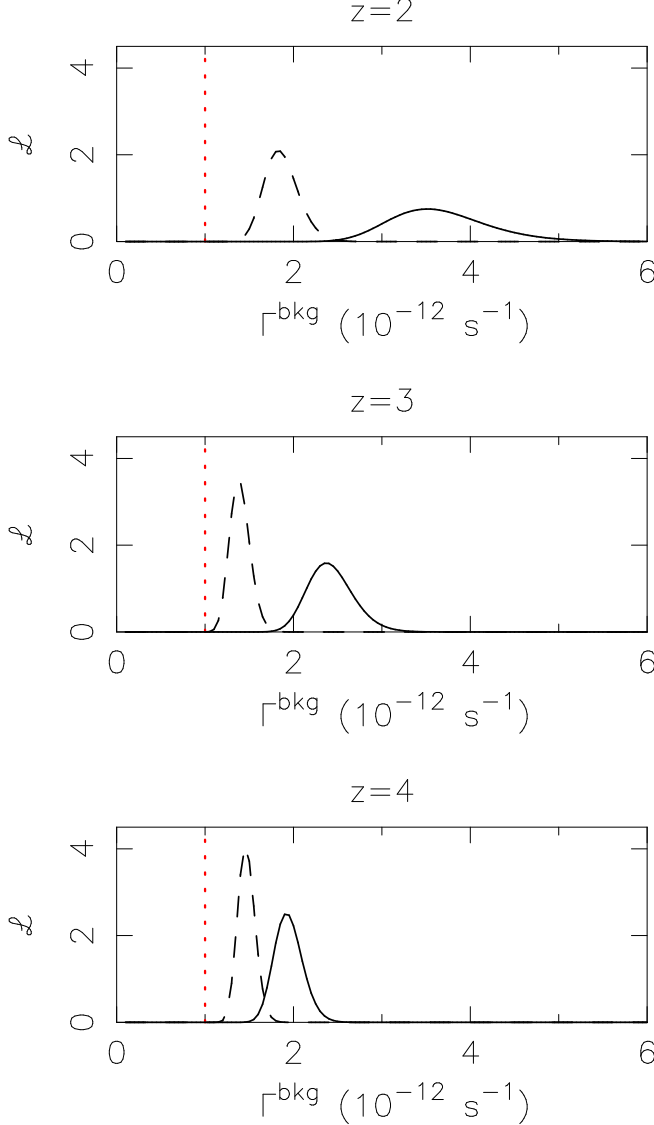


FIG. 10.— Likelihoods for Γ^{bkg} assuming the mapping $\tau^{prox} = \tau(1 + \omega)^{-1}$ computed on samples of 100 mock spectra at $z = 2, 3$, and 4 (top, middle, and bottom) for quasars of typical luminosity lying at the centers of mass of dark matter halos of mass in the range $3.0 \pm 1.6 \times 10^{12} h^{-1} M_{\odot}$. The dashed curves show the cases where the overdensities around quasars have been modeled in the mock spectra, but with no redshift-space distortions. The solid curves show the realistic case where the mock spectra include both the overdensities and the redshift-space distortions owing to gas infall and thermal broadening. The red vertical dotted line indicates the known Γ^{bkg} in the simulation (10^{-12} s^{-1}). Deviations of the solid likelihood peaks from this value are estimates of the bias which is introduced in analyses, e.g. BDO-type, which ignore overdensities and redshift-space distortions. At $z = 3$, this total bias is ≈ 2.5 , in quantitative agreement with the discrepancies between existing proximity effect and flux decrement measurements of Γ^{bkg} (c.f. Figure 1). The overdensities and redshift-space distortions are seen to contribute approximately equally to the total bias.

the likelihood function (31) by PDFs numerically constructed from mock spectra with quasars placed in halos of the correct mass and with redshift-space distortions properly modeled. The PDFs constructed in this way will match those expected in reality for the correct value of Γ^{bkg} . The likelihood method will in this case be unbiased. In this section, we demonstrate how this can be

achieved. Because the PDFs used in this method are generated by Monte Carlo, we refer to this analysis as the “Monte Carlo” (or MC) likelihood.

The first step is to construct optical depth PDFs from the mock spectra as a function of redshift-space proper distance r' from the quasar. We assume here that the mass range of halos hosting quasars, $[M_{min}, M_{max}]$, is known (e.g., from clustering measurements). For any given set of quasar parameters $\{z_Q, L_Q, \alpha_Q\}$, we generate mock spectra for quasars at redshift z_Q in halos in the mass range $[M_{min}, M_{max}]$, with spectral energy distribution parameterized by L_Q and α_Q ⁷. We do so for each value of Γ^{bkg} at which we wish to evaluate the likelihood. In each case, 10 lines of sight are drawn from each of 100 halos randomly chosen in the simulation box. This gives a total of 1000 lines of sight which should capture well the variance between halos of similar masses.

Since τ is nearly lognormally distributed, it is best to first work in logarithmic space and then calculate the PDF of τ itself from the transformation

$$f_{\hat{\tau}}(\tau) = \frac{1}{\tau} f_{\ln \hat{\tau}}(\ln \tau). \quad (44)$$

For $r' \in \{r_{min}, r_{min} + r_{corr}, \dots, r_{max}\}$, the optical depths at redshift-space distance r' from the quasar are tabulated for the 10^4 mock spectra. Using the first four moments (mean, standard deviation, skewness, and kurtosis), we obtain a very good approximate analytic expression for $f_{\ln \hat{\tau}}(\ln \tau)$ using the Edgeworth expansion, described in Appendix C. Figure 11 shows a comparison of the Edgeworth approximation to the data from which it is computed for a representative example of the PDF of the optical depth in the proximity regions of quasars of typical luminosity lying in halos in the fiducial mass range at $z = 3$. After grids of such PDFs are constructed, likelihood analysis is performed as before, simply using these instead of the incorrectly scaled PDFs of equation (28). The Edgeworth expansion accurately approximates the PDF of the data, as illustrated in Figure 11. For 10^4 data points at redshift-space distance $r' = 0.5$ proper Mpc from quasars of typical luminosity in halos in the fiducial mass range $3.0 \pm 1.6 \times 10^{12} h^{-1} M_{\odot}$ at $z = 3$, the reduced χ^2 between the data and its Edgeworth approximation is $\chi^2 = 0.87$. In §7.3, we compute likelihoods using this approximation on mock spectra and recover the correct halo mass, thus confirming that the approximation does not introduce significant biases.

6.2. Test of the Monte Carlo Likelihood

To test the MC likelihood method, we proceed as in §5.5 for the τ -scaling method: we again apply the analysis to mock data sets with known Γ^{bkg} and see if the correct value is recovered. We test the MC likelihood on spectra with both the local overdensity and redshift-space distortions simulated as in §5.5. The results are shown in Figure 12 for 100 spectra at $z = 2, 3$, and 4 . As expected, the correct value of Γ^{bkg} is accurately recovered.

6.3. Relationship to Flux Decrement Method

⁷ In fact, it is not necessary to know L_Q and α_Q separately; it is sufficient to give the integral quantity Γ^{QSO} at some distance from the quasar.

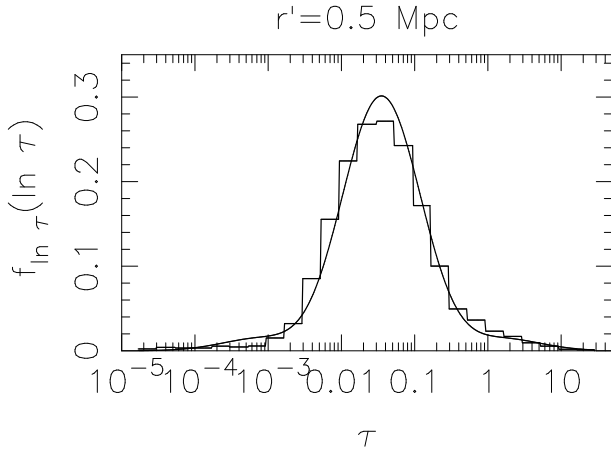


FIG. 11.— Example of the Edgeworth approximation to the optical depth PDF. The normalized histogram shows 10^4 data points at redshift-space distance $r' = 0.5$ proper Mpc from quasars of typical luminosity in halos in the fiducial mass range $3.0 \pm 1.6 \times 10^{12} h^{-1} M_\odot$ at $z = 3$ in the simulation. The Edgeworth expansion, calculated from the first four moments of the data, is shown by the solid curve. Neglecting bins with less 500 data points (which reduce the χ^2 even further since the tails are best approximated) and accounting for the four degrees of freedom estimated from the data, the reduced $\chi^2 = 0.87$, i.e. the Edgeworth approximation is a very good fit to the data.

Unlike the τ -scaling method, which is sensitive only to the scaling of the optical depth in the proximity region with respect to far away from the quasar⁸, the Monte Carlo likelihood method is sensitive to the absolute level of absorption, both close to and far from the quasar. This is because the MC PDFs at each distance from the quasar depend on this absolute level. In fact, in the limit where we ignore points in the proximity regions of the quasars and extend the analysis to greater distances ($r_{prox} \ll r_{min} < r_{max}$), the method essentially reduces to the flux decrement approach. This proximity analysis thus, although in principle unbiased, no longer presents clear advantages over the flux decrement method, such as being independent of Ω_b .

7. MASS OF QUASAR HOST HALOS FROM THE PROXIMITY EFFECT

In the MC likelihood method of the previous section, we assumed that M_{DM} was known and maximized the likelihood function with respect to Γ^{bkg} . In this section, we assume we have an independent measure of Γ^{bkg} , coming either from the flux decrement method or any other reliable measurement. We then proceed as in the previous section, but instead parameterize the likelihood function by the mass of quasar host halos and try to constrain it from the data. In what follows, we test this idea on mock spectra.

7.1. Method

PDFs are numerically constructed as in §6.1, but this time Γ^{bkg} is fixed and the mass range $[M_{min}, M_{max}]$ is varied. For each mass range, we again draw 10 lines of sight from 100 randomly selected halos. Examples of resulting PDFs for different redshift-space distances from the quasars and host halo masses at $z = 3$ are

⁸ The optical depth statistics away from each quasar are estimated from the data in the analysis.

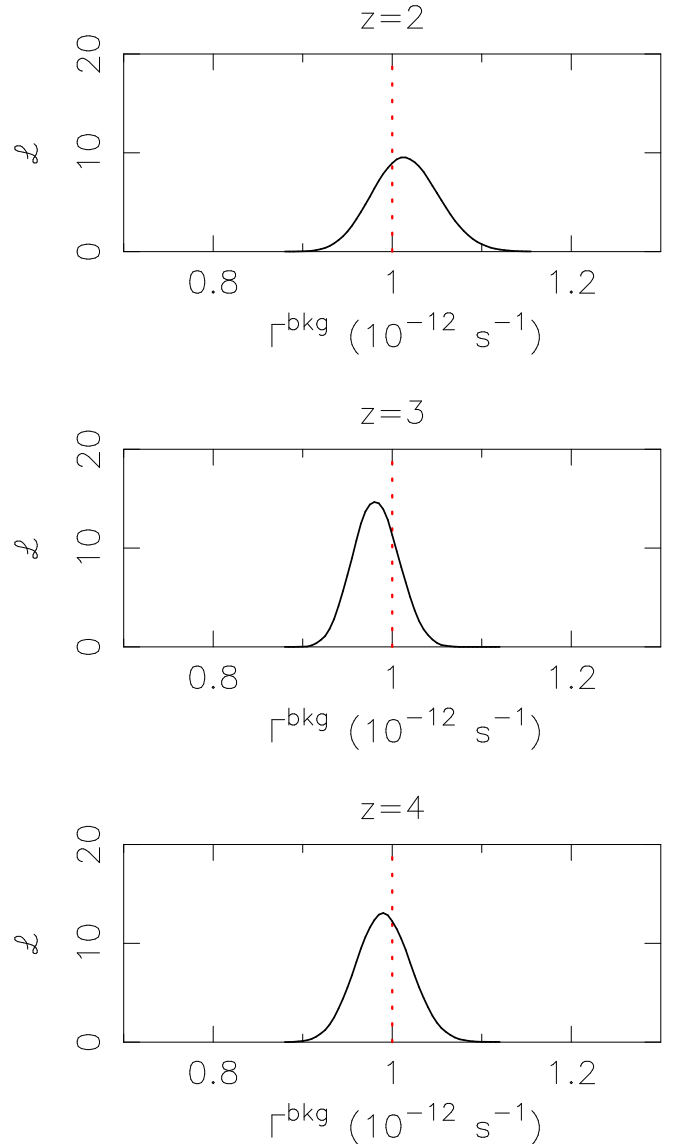


FIG. 12.— Likelihoods for Γ^{bkg} computed from Monte Carlo optical depth PDFs on samples of 100 mock spectra at $z = 2, 3$, and 4 (top, middle, and bottom) for quasars of typical luminosity lying at the centers of mass of dark matter halos of mass in the range $3.0 \pm 1.6 \times 10^{12} h^{-1} M_\odot$, with redshift-space distortions and overdensities fully modeled. As expected, the simulated Γ^{bkg} (10^{-12} s^{-1} , indicated by the vertical red dotted lines) is accurately recovered.

shown in Figure 6. The number of halos employed in constructing the PDFs imposes limitations on the width of the mass ranges that can be used, since the simulation box contains a finite number of halos in any given mass interval. Moreover, the width of the optical depth PDF for any given mass range depends on the width of the range. We thus consider mass ranges of fixed width in the logarithm, equal to that of the fiducial range $3.0 \pm 1.6 h^{-1} \times 10^{12} M_\odot$, covering the range of halo masses represented in the simulation and requiring that each mass range contains at least 100 halos. The mass ranges are indexed by the central mass in logarithmic space, $M_{DM} = \exp[0.5(\ln M_{min} + \ln M_{max})]$. This choice is consistent with the exponential nature of the mass function (e.g., Jenkins et al. 2001) and avoids an

excessive pile-up of halos to the left of the mass index.

7.2. Location of Mass Information

In the case of Γ^{bkg} , the proximity effect optical depth statistics are expected to convey significant information in the entire region where Γ^{QSO} and Γ^{bkg} are comparable. This region can extend to $\gtrsim 20$ proper Mpc for bright quasars at $z = 3$ (Figure 7). However, the halo overdensity and infall velocity (in units of Hr) profiles drop greatly on scales $\lesssim 1$ proper Mpc (Figures 4 and 5). The information about the dark matter halos is contained in these signatures and it is a priori unclear how far away from a quasar one can go and still learn about its host halo. We address this question in this section, before proceeding with testing the likelihood method for recovering M_{DM} in the next.

To quantify the mass information content as a function of distance from a quasar, we use the Kolmogorov-Smirnov (K-S) test (Press et al. 1992). This statistical test compares two random samples and quantifies the significance that the realized values were drawn from different distributions via the K-S P-value, P_{K-S} . If $P_{K-S} < \alpha$, then a deviation as large as observed between the two random samples would have probability $< \alpha$ to occur by chance if the two samples were drawn from the same distributions.

Here, we fix $\{z_Q, L_Q, \alpha_Q\}$, the known host halo mass, M_{DM}^{true} , and the number of spectra N_Q in the mock data set and consider points at increasing distance from the quasar. At each redshift-space distance r' , we vary M_{DM} . For each trial value, we compare the sample of N_Q optical depths at distance r' in the mock spectra to a sample of 1000 optical depths for M_{DM}^{true} . If the resulting P_{K-S} is small, then the optical depth statistics can significantly distinguish between M_{DM} and M_{DM}^{true} .

Figure 13 shows the results of this K-S analysis at $z = 2, 3$, and 4. Because the realized P_{K-S} values vary with the random samples, we show the mean value (dot) and the standard deviation (error bar) for each mass, computed from ensembles of 50 sets of mock spectra. At each redshift, the mass information appears mostly contained at points $r' \lesssim 1$ proper Mpc (note that this is still $\sim 20r_{vir}$ – see eq. 10), although this conclusion is somewhat dependent on the range of masses probed by our simulation. Even if P_{K-S} drops only marginally for masses differing from M_{DM}^{true} for the masses probed, the analysis would presumably significantly rule out much smaller or much larger masses. This is supported by Figure 6, which shows that the optical depth PDFs for halos located in random locations (effectively, $M_{DM} = 0$) strongly differ from those of the massive halos resolved in the simulation.

7.3. Test of the Halo Mass Likelihood

Figure 14 shows likelihoods for M_{DM} computed on samples of 100 mock spectra for typical quasars lying in halos in the fiducial mass range $3.0 \pm 1.6 \times 10^{12} h^{-1} M_\odot$. For these likelihood calculations, we use $(r_{min}, r_{max}) = (0.1, 5.1)$ proper Mpc and vary the correlation length, r_{corr} , assumed in the computation of the likelihoods from 1.0 proper Mpc to 0.5 and 0.1 proper Mpc. Better apparent constraints are obtained with smaller r_{corr} , as more data points are used. However, the width of the

estimated likelihoods is underestimated by an increasing factor as more correlated points are assumed independent (see Appendix D)). The correct mass range is recovered with a precision comparable to the width of the true mass interval in the case $r_{corr} = 1.0$ proper Mpc, in which case the correlations are small (see Figure 9). This shows that, in principle, it is possible to measure quasar host halo masses from the proximity effect with a data set of modest size. In fact, a data set of 100 spectra is two orders of magnitude smaller than those used in state-of-the-art clustering analyses such as the Croom et al. (2005) one. It is true, however, that the clustering analyses have relatively low spectroscopic requirements. As we discuss in §8, important challenges must be addressed before the level of precision estimated here can be achieved in practice. In fact, realistic spectra have finite resolution and signal-to-noise ratio. Moreover, the continuum level is not given and must be estimated, and systematic quasar redshifts are often not precisely known.

8. PRACTICAL CONSIDERATIONS

Throughout this paper, we have optimistically assumed ideal data sets. In particular, we have assumed spectra with infinite resolution and signal-to-noise ratio (SN), and perfect redshift determinations for the quasars. In this section, we briefly discuss the complications associated with realistic data. As we will argue, *if* the continuum flux and redshift of each quasar are known accurately, finite resolution and signal to noise can be accounted for and do not present serious problems. However, our ability to estimate a quasar’s continuum flux is limited by the resolution and noise in the spectra and accurate quasar systemic redshifts are challenging to obtain. These difficulties are likely to be the most important in applying the method presented in this paper to measure quasar host halo masses to actual data.

8.1. Finite Resolution and Signal-to-Noise Ratio

Let us first suppose that the continuum flux and redshift of each quasar are given, and consider the effects of finite resolution and signal-to-noise ratio.

Finite resolution alters the optical depth PDF by smoothing out the small-scale fluctuations. This can in principle be simply accounted for by smoothing the mock spectra from which the model PDF are constructed to the same resolution as the actual data.

A finite signal-to-noise ratio requires more spectra to detect intrinsic fluctuations in the IGM of a fixed amplitude. The total variance on the pointwise transmission $F \equiv e^{-\tau}$ is a sum of the variance due to intrinsic IGM fluctuations, σ_F^2 , and the variance in the observed flux due to the finite SN , $\sigma_{SN}^2 \equiv 1/SN^2$. Thus, the error on the standard estimator for the mean F estimated from N spectra is

$$\sigma_{\langle F \rangle} = \frac{\sqrt{\sigma_F^2 + \sigma_{SN}^2}}{\sqrt{N}} \quad (45)$$

The PDF shown on Figure 6 suggests that to distinguish between different halo masses requires measuring mean F at the percent level. Thus, one needs approximately

$$N = \frac{\sigma_F^2 + 1/SN^2}{0.01^2}. \quad (46)$$

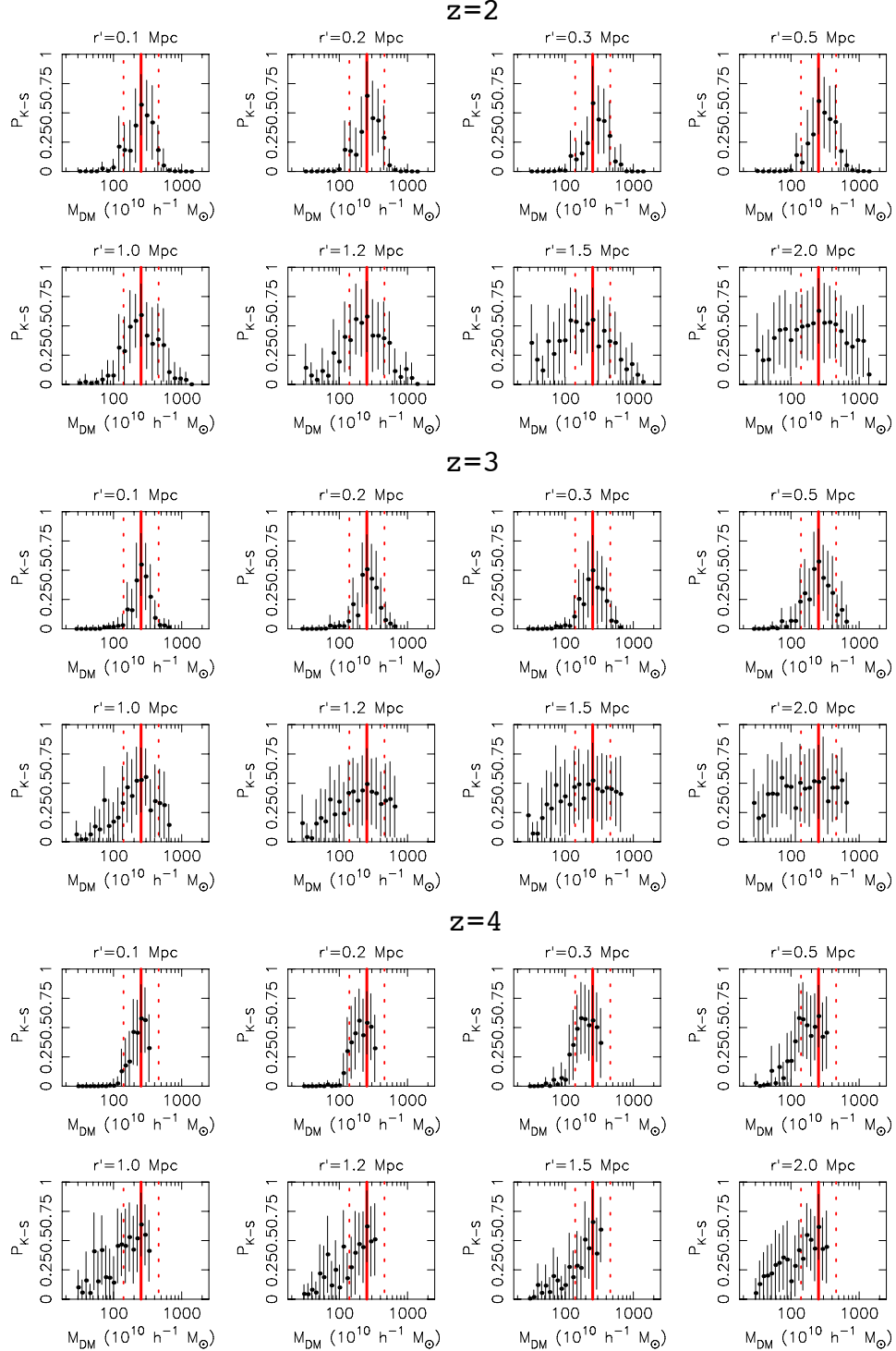


FIG. 13.— Kolmogorov-Smirnov P-value for the significance of the difference between the true ($M_{DM} = 3.0 \pm 1.6 h^{-1} \times 10^{12} M_{\odot}$) and trial optical depth PDFs as a function of the quasar host dark matter halo mass range, at increasing apparent distances from the quasar. Each trial mass range has the same logarithmic width as the true one (indicated by the vertical red lines) and is labeled by the central mass in logarithmic space. In each case, the P-value is averaged over 50 samples of 100 typical quasar spectra at $z = 4$ and the vertical error bars indicate the standard deviation. At $z = 4$, the K-S points do not appear to fall on the right side of the true mass range simply because there are too few massive halos in the simulation box to probe mass ranges much larger than the fiducial one.

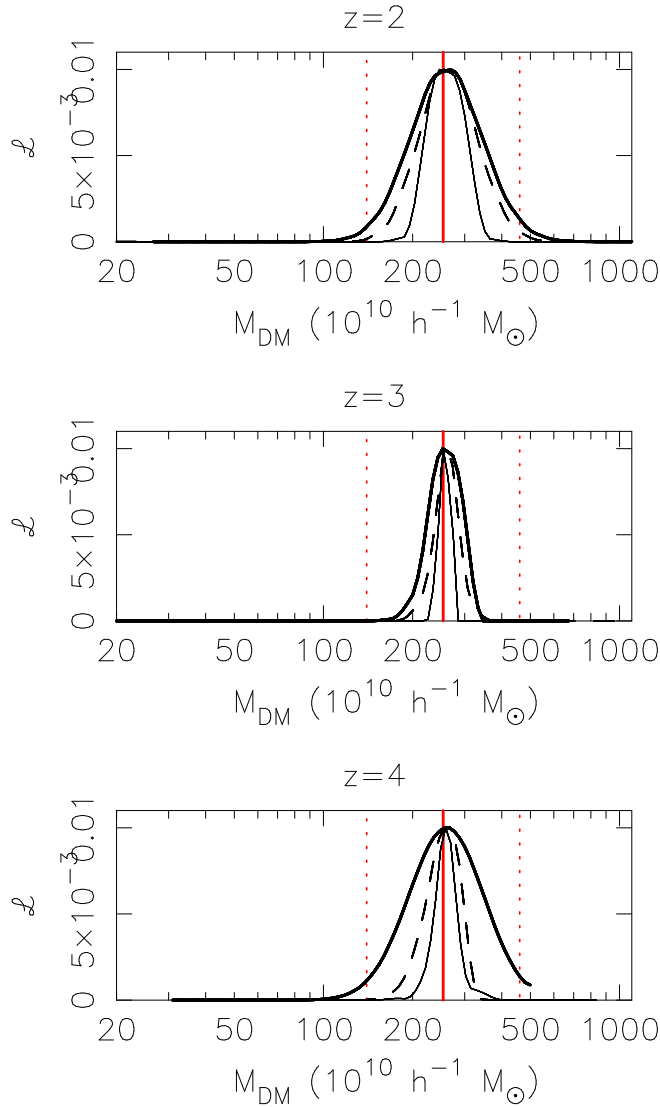


FIG. 14.— Likelihoods for M_{DM} computed on samples of 100 mock spectra at $z = 2, 3$, and 4 (top, middle, and bottom) for quasars of typical luminosity lying at the centers of mass of dark matter halos of mass in the range $3.0 \pm 1.6 \times 10^{12} h^{-1} M_{\odot}$ (indicated by the vertical red lines). $\Gamma^{bkg} = 10^{-12} s^{-1}$ is assumed to be known exactly. The correlation length, r_{corr} , assumed in the computation of the likelihoods is 1.0 proper Mpc (thick solid curves), 0.5 proper Mpc (dashed curves), and 0.1 proper Mpc (thin solid curves). Better apparent constraints are obtained with smaller r_{corr} , as more data points are used. However, the width of the estimated likelihoods is underestimated by an increasing factor as more correlated points are assumed independent (Appendix D). The correct mass range is recovered with a precision comparable to the width of the true mass interval in the case $r_{corr} = 1.0$ proper Mpc, in which case the correlations are small (Figure 9). The left-most value on the horizontal axis corresponds to the smallest-mass halos resolved in our simulation. The $z = 4$ likelihood is truncated at high masses simply because the simulation box contains too few halos with mass much larger than the fiducial range to evaluate the likelihood at those points. The likelihoods on this figure were normalized to peak at 0.01.

spectra to distinguish between different realistic halo masses. At $z \sim 3$, typical intrinsic fluctuations have $\sigma_F \approx 0.2$, thus requiring $SN \approx 5$ to dominate over the noise. This, by itself, is a modest requirement. In that case, of order 10^3 spectra are necessary to distinguish between different halo masses. Of course, valuable mass information may also be encoded in the higher moments

of the transmission, in particular in the incidence of large (~ 1) optical depths. Several data points from each spectrum can also in principle be used (e.g., as in the likelihood method described in §5.2), alleviating the data requirements.

8.2. Continuum Fitting

An indirect, more severe effect of finite resolution and SN is to make accurate continuum fitting more difficult. In the case of finite SN , the “peaks” in the observed spectra which are usually fitted to estimate the continua in general do not exactly correspond to a transmission of unity, owing to the noise contribution. In the case of finite resolution, the peaks are smoothed and their height may therefore be underestimated. There is also a fundamental limit to how precisely quasar spectra can be continuum fitted, since peaks of unity transmission become increasingly rare, if extant at all, at high ($\gtrsim 4$) redshifts. Continuum errors are of special concern because continuum fitting is a somewhat subjective procedure and they are likely to be affected by systematics to some extent. It may thus be fruitful to devise more sophisticated methods which bypass continuum estimation, along the lines suggested by Lidz et al. (2006a) for the analysis of the matter power spectrum.

8.3. Redshift Errors

In practice, the uncertainties on quasar redshifts are substantial. The redshifts determined from broad emission lines, which are affected by inflows and outflows of material, can differ a lot from the actual systemic redshift of the quasar (Gaskell 1982; Tytler & Fan 1992; Vanden Berk et al. 2001; Richards et al. 2002). For example, Richards et al. (2002) find that the broad C IV emission line has a median blueshift of 824 km s^{-1} with respect to the narrow Mg II line, with a dispersion about the median of 511 km s^{-1} . At $z = 3$, this corresponds to a physical scale of a few Mpc, comparable to the radius which is substantially affected by the host halo of the quasar. Narrow emission lines, such as [O III] 5007 Å and Mg II 2798 Å, which are associated with the host galaxy, can provide systemic redshifts to a precision $\approx 50 - 300 \text{ km s}^{-1}$ (J. Hennawi, private communication; see also Vanden Berk et al. 2001 and Richards et al. 2002a). At $z \gtrsim 2$, these narrow lines however require measurements extending in the near infrared. The importance of accurate redshifts has been highlighted in the context of the transverse proximity effect by Hennawi et al. (2006) and Hennawi & Prochaska (2006). Methods to obtain robust systemic quasar redshifts, like that developed by Hennawi et al. (2006) and Shen et al. (2007), are therefore likely to be necessary for accurate proximity effect work.

9. COMPARISON WITH OTHER WORK

Loeb & Eisenstein (1995) first estimated the bias in proximity effect measurements of Γ^{bkg} arising from their overdense environments. They found that the bias could be up to a factor ~ 3 , consistent with our results. Their analytical analysis, very different from ours, was based on counting the Ly α absorption lines Doppler shifted beyond the quasar redshift owing to gas infall. Rollinde et al. (2005) also proposed using the proximity effect to study the density structure around quasars.

They however incorrectly neglected the effects of gas infall toward the quasars, which we have shown to be very important. Instead of using dark matter halos from their simulation, they simply modeled the density enhancement around quasars with an analytic function of distance multiplying the density field. This approach neglects the fact that the density fluctuations are not necessarily linearly scaled as a quasar lying in a halo is approached, e.g. the fluctuations and their standard deviation need not be scaled by the same factor. As Figure 4 shows, the ratio of $\langle \delta \rangle$ to the pointwise standard deviation of δ varies strongly with distance, violating the authors' assumption. This approach also has the drawback of not relating the recovered density profiles to the mass of the quasar host halos. Guimarães et al. (2007) have applied the Rollinde et al. (2005) model on a sample of 45 quasars at $z > 4$ and measured an average halo profile extending to $\gtrsim 40$ proper Mpc, an order of magnitude farther than expected from simulations of structure formation (c.f., Figure 4), casting doubt on the validity of their approach. This anomalous result is likely an artifact of an inaccurate modeling of the ionizing effect of the quasars (the standard proximity effect), which is dominant and must be subtracted in order to measure the host halo mass. For instance, these authors neglect redshift-space distortions altogether and their analytical model is partly heuristic and mathematically inexact.

10. CONCLUSION

In this paper, we have shown that the Ly α optical depth statistics in the proximity regions of quasar spectra are significantly affected by the massive halos which they are expected to occupy. In general, the mean optical depth increases with halo mass. We have then quantified the biases induced in line-of-sight proximity effect measurements of the background photoionizing rate Γ^{bkg} that neglect the effects of quasar host halos assuming ideal data, i.e. that the quasar spectra have infinite resolution and signal-to-noise ratio and perfectly known continuum level and systemic redshift. The local matter overdensity around and the infall of gas toward the quasars contribute approximately equally to the total upward bias. At $z \approx 3$, where most proximity effect measurements of Γ^{bkg} have been made, the proximity effect Γ^{bkg} bias for host halo masses in the range $3.0 \pm 1.6 h^{-1} \times 10^{12} M_{\odot}$, as inferred by Croom et al. (2005) from clustering measurements, could be ≈ 2.5 , enough to bring in agreement the existing proximity effect and flux decrement measurements. The existing proximity effect measurements were however made on data of finite quality, with continua and redshifts estimated from the data. These observational difficulties may also affect the validity of those proximity effect measurements, beyond the effects investigated here. The fact that quasars lying in overdense regions of the Universe introduces very significant biases is nonetheless a robust conclusion of our study, and should be taken into account in future analyses.

The clustering of galaxies and other AGN around proximity effect quasars has a small effect on the local magnitude of the background ionizing flux, at least at $z \lesssim 3$, and therefore is not expected to significantly bias Γ^{bkg} measurements.

By constructing optical depth PDFs by Monte Carlo from realistic mock spectra with quasars in host halos of prescribed mass, one can perform a likelihood analysis for Γ^{bkg} which is unbiased. However, this method is sensitive to the absolute level of Ly α absorption, both close to and far from the quasars. In the limit where only points far from the quasars are considered in the analysis, it essentially reduces to the flux decrement method. As such, the method does not possess the main advantage of “true” proximity effect analyses (which are only sensitive to the *change* in absorption statistics near the quasars) that they are relatively free of cosmological parameter assumptions. In particular, this method requires knowledge of Ω_b and so a Γ^{bkg} value obtained with it cannot be used in conjunction with the flux decrement to constrain the baryon density.

Of perhaps greatest interest, we have demonstrated how, given a measurement of Γ^{bkg} (e.g., from the flux decrement), the proximity effect analysis can in principle be inverted to probe the environments of quasars. In particular, we have shown that the masses of dark matter halos hosting quasars could be measured using the optical depth statistics in the proximity regions of quasars.

As mentioned, we have in this work side-stepped a number of observational issues. For instance, we have assumed that the continuum flux and redshift of each quasar were known exactly; in practice, these are plagued by uncertainties. Moreover, real spectra have finite resolution and signal-to-noise ratio. From our discussion of these complications, we have concluded that accurate continuum fitting and quasar redshift determination are likely to be the most important challenges to using our method to measure halo masses. Further work in developing techniques to either improve or bypass these measurements would thus be highly desirable. Our theoretical picture is also simplified in some respects. For instance, we have assumed that the low-density Hui & Gnedin (1997) equation of state for the IGM holds all the way to the centers of the quasar host halos, which may not be the case. In particular, if helium is doubly-ionized by quasars at $z \sim 3$ (e.g. Madau et al. 1999), the temperature should be enhanced close to quasar sources. Whether such thermal effects are important in the vicinity of quasars could be investigated by studying how the small-scale power spectrum in the Ly α forest varies as the quasars are approached, as was proposed by Zaldarriaga (2002) for the general forest.

We thank Suvendra Dutta for assistance with the simulations used in this work. We also acknowledge enlightening discussions with Joe Hennawi, Jason Prochaska, and Scott Burles on practical issues in the analysis of quasar spectra. CAFG is supported by a NSERC Postgraduate Fellowship and Harvard University grants. This work was supported in part by NSF grants ACI 96-19019, AST 00-71019, AST 02-06299, AST 03-07690, and AST 05-06556, and NASA ATP grants NAG5-12140, NAG5-13292, NAG5-13381, and NNG-05GJ40G. Further support was provided by the David and Lucile Packard, the Alfred P. Sloan, and the John D. and Catherine T. MacArthur Foundations.

TABLE 2
EXISTING FLUX DECREMENT AND PROXIMITY EFFECT MEASUREMENTS OF THE BACKGROUND
PHOTOIONIZING FLUX

Redshift	Flux Decrement		Reference
	J_{-21} ergs s ⁻¹ cm ⁻² Hz ⁻¹ sr ⁻¹	Γ^{bkg} 10 ⁻¹² s ⁻¹	
1.95		1.33	Jena et al. (2005)
3		1.3 ± 0.1	Kirkman et al. (2005)
2		1.3 ^{+0.8} _{-0.5}	Bolton et al. (2005)
3		0.9 ± 0.3	...
4		1.0 ^{+0.5} _{-0.3}	...
1.9		1.44 ± 0.11	Tytler et al. (2004a)
2.75		0.86 ^{+0.36} _{-0.24}	Meiksin & White (2004)
3.0		0.88 ^{+0.14} _{-0.12}	...
3.89		0.68 ^{+0.08} _{-0.07}	...
4.0		0.43 ^{+0.06} _{-0.05}	...
5.0		0.31 ^{+0.07} _{-0.09}	...
5.5		0.37 ^{+0.06} _{-0.05}	...
6.0		< 0.14	...
2.4		0.698 ± 0.096	McDonald & Miralda-Escudé (2001)
3		0.518 ± 0.083	...
3.9		0.380 ± 0.04	...
4.5		0.21 ± 0.04	...
4.93		0.13 ± 0.03	...
5.2		0.16 ± 0.04	...
4.72	≫ 0.04 ^a	≫ 0.129	Songaila et al. (1999)
[1.5, 2.5]		0.890	Rauch et al. (1997) ^b
[2.5, 3.5]		0.698	...
[3.5, 4.5]		0.618	...
Proximity Effect			
[1.7, 3.8]	0.7 ^{+0.34} _{-0.44}	1.9 ^{+1.2} _{-1.0} ^{c,d}	Scott et al. (2000)
[2.0, 4.5]	1.0 ^{+0.5} _{-0.3}	2.6 ^{+1.3} _{-0.8}	Cooke et al. (1997)
[1.7, 3.7]	0.6	1.6 [*]	Srianand & Khare (1996)
[1.7, 4.1]	0.5 ± 0.1	1.3 ± 0.3 [*]	Giallongo et al. (1996)
3.66	0.5 ^e	1.3 [*]	Cristiani et al. (1995)
≈ 4.2	≈ 0.1 – 0.3 ^f	0.3 – 0.8 [*]	Williger et al. (1994)
≈ 0.5	≈ 0.006 ^g	0.02 [*]	Kulkarni & Fall (1993)
[1.7, 3.8]	1 ⁺² _{-0.7}	2.6 ^{+4.2} _{-1.8} [*]	Lu et al. (1991)
[1.7, 3.8]	1 ^{+3.2} _{-0.7}	2.6 ^{+8.3} _{-1.8} [*]	Bajtlik et al. (1988)
3.75	≥ 3 ^h	≥ 7.8 [*]	Carswell et al. (1987)

^a Assuming a spectral index of 0.7 for the background flux. ^b For their Λ CDM cosmology. ^c The authors claim that the presence of lines on the saturated part of the curve of growth could cause their estimate to be overestimated by a factor 2–3. ^d Contrary to other proximity effect analyses, this value does not assume a spectral index for the background; the authors repeated their analysis solving directly for Γ^{bkg} . ^e From a single quasar, QSO 0055–269. ^f From a single quasar, QSO BR 1033–0327. ^g The uncertainties are large; at the 1 σ could be lower by a factor of 3 or higher by a factor of 6. ^h From a single quasar, PKS 2000–330. ^{*} Calculated from $J_{\nu Ly}$ assuming a spectral index for the background flux equal to the typical value for radio-quiet quasars, $\alpha = 1.57$ (Telfer et al. 2002). These authors assumed that the background flux had the same spectral index as the quasars shortward of the Lyman limit in their analyses, from which they inferred $J_{\nu Ly}$.

APPENDIX

A. EXISTING FLUX DECREMENT AND PROXIMITY EFFECT MEASUREMENTS OF THE BACKGROUND PHOTOIONIZING FLUX

In this appendix, we tabulate the existing measurements of Γ^{bkg} from the flux decrement and proximity effect methods. This serves as a quantitative complement to Figure 1.

B. ACCURATE EXPRESSIONS

In some of our calculations, we have used approximate expressions for the physical properties of the IGM in order to improve computational efficiency. These approximations are not expected to have any significant effect on our results, as they were applied consistently. However, when the methods presented in this work are to be applied to real spectra, it may be important to use more accurate expressions, in order to be consistent with the physical processes as they occur in the Universe. Some lengthy formulae were also omitted in the main body in order to improve readability. In this Appendix, based on that of Hui & Gnedin (1997), we collect explicit accurate expressions missing in the text.

Hydrogen photoionization cross section:

$$\sigma(E) = 5.475 \times 10^{-14} \text{ cm}^2 (E/0.4298 \text{ eV} - 1)^2 \frac{(E/0.4298 \text{ eV})^{-4.018}}{(1 + \sqrt{E/14.13 \text{ eV}})^{2.963}} \quad (\text{B1})$$

(Verner et al. 1996; accurate to 10% from the ionization thresholds to 5 keV).

Hydrogen recombination coefficient:

$$R(T) = 1.269 \times 10^{-13} \text{ cm}^{-3} \text{ s}^{-1} \frac{(2 \times 157,807 \text{ K}/T)^{1.503}}{[1.0 + (2 \times 157,807 \text{ K}/0.522T)^{0.470}]^{1.923}} \quad (\text{B2})$$

(fit by Hui & Gnedin 1997 to Ferland et al. 1992 data; accurate to 2% from 3 to 10^9 K).

T_0 coefficient in IGM equation of state 5:

$$T_0^{1.7} = \left(\frac{a_{reion}^2 T_{reion}}{a^2} \right)^{1.7} + \frac{1.7}{1.9} a^{-[3/2+(a-0.25)/a]} \left(1 - \frac{a_{reion}^{1.9}}{a} \right) B, \quad (\text{B3})$$

where

$$B = \frac{\rho_b}{H_0 m_p \sqrt{\Omega_m}} T_{reion} R(T_{reion}), \quad (\text{B4})$$

$T_{reion} = 24,000 \text{ K}$, and we take $z_{reion} = 1/(1 + a_{reion}) = 10$.

IGM equation of state exponent β :

$$\beta = \frac{1}{1.7} \left[1 - \left(\frac{a_{reion}^2 T_{reion}}{a^2 T_0} \right)^{1.7} \frac{a_{reion}}{a} \right] + \left(\frac{2}{3} - \frac{1}{1.7} \right) (a^2 T_0)^{-1.7} C, \quad (\text{B5})$$

where

$$C = \frac{1.7}{1.9} B \left\{ \frac{a^{1.9}}{2.9} \left[1 - \left(\frac{a_{reion}}{a} \right)^{2.9} \right] - a_{reion}^{1.9} \left(1 - \frac{a_{reion}}{a} \right) \right\} + (a_{reion}^2 T_{reion})^{1.7} \left(1 - \frac{a_{reion}}{a} \right). \quad (\text{B6})$$

C. THE EDGEWORTH EXPANSION

The Edgeworth expansion (e.g., Juszkiewicz et al. 1995; Blinnikov & Moessner 1998) is an expansion of a nearly Gaussian distribution in terms of its moments. Since the Ly α optical depth statistics are nearly lognormal, their PDF is well approximated by an Edgeworth expansion in the first four moments:

$$f_{\ln \hat{\tau}}(\ln \tau) \approx \frac{1}{\sigma_{\ln \tau}} \left| \left\{ 1 + \frac{1}{3!} \text{skew}(\ln \tau) H_3(\eta) + \frac{1}{4!} \text{kurt}(\ln \tau) H_4(\eta) + \frac{10}{6!} [\text{skew}(\ln \tau)]^2 H_6(\eta) \right\} \phi(\eta) \right|_{\eta = \frac{\ln \tau - \langle \ln \tau \rangle}{\sigma_{\ln \tau}}}, \quad (\text{C1})$$

where $\text{skew}(\ln \tau)$ and $\text{kurt}(\ln \tau)$ are the skewness and kurtosis of $f_{\ln \hat{\tau}}$, respectively, the $H_n(x)$ are Hermite polynomials (e.g., Abramowitz & Stegun 1965), and

$$\phi(\eta) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\eta^2}{2}\right) \quad (\text{C2})$$

is a standard Gaussian PDF. Unbiased estimators of the skewness and kurtosis are given by

$$\widehat{\text{skew}(\ln \tau)} = \frac{N\sqrt{N-1}}{N-2} \frac{\sum_{i=1}^N (\ln \tau_i - \widehat{\langle \ln \tau \rangle})^3}{[\sum_{i=1}^N (\ln \tau_i - \widehat{\langle \ln \tau \rangle})^2]^{3/2}} \quad (\text{C3})$$

and

$$\widehat{\text{kurt}(\ln \tau)} = \frac{(N+1)N(N-1)}{(N-2)(N-3)} \frac{\sum_{i=1}^N (\ln \tau_i - \widehat{\langle \ln \tau \rangle})^4}{[\sum_{i=1}^N (\ln \tau_i - \widehat{\langle \ln \tau \rangle})^2]^2} - 3 \frac{(N-1)^2}{(N-2)(N-3)}. \quad (\text{C4})$$

The Edgeworth expansion is used to approximate PDFs of $\ln \tau$ in sections 6 and 7.

D. EFFECT OF CORRELATIONS ON THE LIKELIHOOD FUNCTION

In this section, we present a toy model to understand the effect of correlations on the likelihood function. Specifically, we seek to quantify the error that is made if neighboring data points are correlated, but a likelihood estimator assuming independence is used. For example, in this paper we pretended that points separated by a “correlation length” were perfectly independent (e.g., in §5.2). We would like to know how well our estimate of the likelihood obtained in this way represents the true likelihood.

Consider N independent pairs of Gaussian random variables $X_{i,1}$ and $X_{i,2}$, each with mean μ and variance σ^2 . The Gaussian functional form is well-motivated, since $\ln \tau$ is approximately normally distributed. $X_{i,1}$ and $X_{i,2}$ are correlated with one another, with covariance matrix

$$\Sigma = \begin{pmatrix} \sigma^2 & \zeta \\ \zeta & \sigma^2 \end{pmatrix}, \quad (\text{D1})$$

so that their joint PDF is given by

$$f(\mathbf{x}) = \frac{1}{2\pi|\Sigma|^{1/2}} \exp [(\mathbf{x} - \mu)\Sigma^{-1}(\mathbf{x} - \mu)^T], \quad (\text{D2})$$

where $\mathbf{x} \equiv (x_1, x_2)$. Let $\mathcal{L}^{i,est}$ be the likelihood for μ estimated from $X_{i,1}$ and $X_{i,2}$, ignoring the correlation between the two random variables, and $\mathcal{L}^{i,true}$ be the “true” likelihood. Then:

$$\ln \mathcal{L}^{i,est} = -\ln [2\pi\sigma] - \frac{1}{2\sigma^2} [(x_1 - \mu)^2 + (x_2 - \mu)^2] \quad (\text{D3})$$

and

$$\ln \mathcal{L}^{i,true} = -\ln [2\pi\sigma(1 - \epsilon^2)^{1/2}] - \frac{1}{2\sigma^2(1 - \epsilon^2)} [(x_1 - \mu)^2 - 2\epsilon(x_1 - \mu)(x_2 - \mu) + (x_2 - \mu)^2], \quad (\text{D4})$$

where $\epsilon \equiv \zeta/\sigma^2$. Given that $(X_{i,1}, X_{i,2})$ is distributed as in equation D2, we may compute the expectation values of $\ln \mathcal{L}^{i,est}$ and $\ln \mathcal{L}^{i,true}$:

$$\langle \ln \mathcal{L}^{i,est} \rangle = -\ln [2\pi\sigma^2] - \left[1 + \left(\frac{\Delta\mu}{\sigma} \right)^2 \right] \quad (\text{D5})$$

and

$$\langle \ln \mathcal{L}^{i,true} \rangle = -\ln [2\pi\sigma^2(1 - \epsilon^2)^{1/2}] - \frac{1}{1 - \epsilon^2} \left[1 + \left(\frac{\Delta\mu}{\sigma} \right)^2 \right] + \frac{\epsilon}{1 - \epsilon^2} \left[\epsilon + \left(\frac{\Delta\mu}{\sigma} \right)^2 \right], \quad (\text{D6})$$

where $\Delta\mu$ is the difference between the true mean and the value at which the likelihood is evaluated. Since $\langle \ln \mathcal{L}^{i,est} \rangle$ is maximum for $\Delta\mu = 0$, the maximum likelihood estimate for μ is unbiased, even if the estimator ignores the correlations between the random variables. An estimate for the width of $\langle \ln \mathcal{L} \rangle$ is given by the curvature at the maximum likelihood value, $\tilde{\sigma}(\mathcal{L}) \equiv -[\sqrt{\partial^2 \langle \ln \mathcal{L} \rangle / \partial (\Delta\mu)^2}]^{-1}$. This expression is exactly equal to the standard deviation of the likelihood function when the latter is Gaussian. Explicitly,

$$\tilde{\sigma}(\mathcal{L}^{i,est}) = \frac{\sigma}{\sqrt{2}} \quad \text{and} \quad \tilde{\sigma}(\mathcal{L}^{i,true}) = \sqrt{1 + \epsilon} \frac{\sigma}{\sqrt{2}}. \quad (\text{D7})$$

Let us now consider the total likelihood for μ , as estimated from the N pairs $(X_{i,1}, X_{i,2})$. By independence,

$$\mathcal{L}^{tot} = \prod_{i=1}^N \mathcal{L}^i, \quad (\text{D8})$$

so that

$$\ln \mathcal{L}^{tot} = \sum_{i=1}^N \ln \mathcal{L}^i. \quad (\text{D9})$$

Therefore,

$$\langle \ln \mathcal{L}^{tot} \rangle = N \langle \ln \mathcal{L}^i \rangle, \quad (\text{D10})$$

and

$$\tilde{\sigma}(\mathcal{L}^{i,est}) = \frac{\sigma}{\sqrt{2N}} \quad \text{and} \quad \tilde{\sigma}(\mathcal{L}^{i,true}) = \sqrt{1 + \epsilon} \frac{\sigma}{\sqrt{2N}}. \quad (\text{D11})$$

Thus, the width of the estimated likelihood, when correlations are ignored, is underestimated by a factor $1/\sqrt{1 + \epsilon}$ ($\approx 1 - \epsilon/2$ for $\epsilon \ll 1$). In particular, the error on the width of the likelihood is independent of the number of independent pairs of correlated data. In the limit $\epsilon \rightarrow 0$ (no correlation), $\tilde{\sigma}(\mathcal{L}^{i,est}) \rightarrow \tilde{\sigma}(\mathcal{L}^{i,true})$, and the standard result for the error on the mean is recovered. In the limit $\epsilon \rightarrow 1$ (the two data points from each pair are the same), $\tilde{\sigma}(\mathcal{L}^{i,est}) \rightarrow \tilde{\sigma}(\mathcal{L}^{i,true})/\sqrt{2}$, again as expected, since there are half as many data independent data points as the number of points from which the mean is estimated.

REFERENCES

- Abramowitz, M., & Stegun, I. A. 1965, Handbook of mathematical functions with formulas, graphs, and mathematical tables (Dover Books on Advanced Mathematics, New York: Dover, —c1965, Corrected edition, edited by Abramowitz, Milton; Stegun, Irene A.)
- Adelberger, K. L. 2004, ApJ, 612, 706
- Adelberger, K. L., & Steidel, C. C. 2005a, ApJ, 630, 50
- , 2005b, ApJ, 627, L1
- Bajtlik, S., Duncan, R. C., & Ostriker, J. P. 1988, ApJ, 327, 570
- Barkana, R. 2004, MNRAS, 347, 59
- Barkana, R., & Loeb, A. 2001, Phys. Rep., 349, 125
- , 2003, Nature, 421, 341
- Becker, G. D., Rauch, M., & Sargent, W. L. W. 2006, ArXiv Astrophysics e-prints
- Bi, H., & Davidsen, A. F. 1997, ApJ, 479, 523
- Bi, H. G., Boerner, G., & Chu, Y. 1992, A&A, 266, 1
- Blinnikov, S., & Moessner, R. 1998, A&AS, 130, 193
- Bolton, J. S., Haehnelt, M. G., Viel, M., & Springel, V. 2005, MNRAS, 357, 1178
- Bond, J. R., Cole, S., Efstathiou, G., & Kaiser, N. 1991, ApJ, 379, 440

- Burles, S., Nollett, K. M., Truran, J. W., & Turner, M. S. 1999, *Physical Review Letters*, 82, 4176
- Burles, S., Nollett, K. M., & Turner, M. S. 2001, *ApJ*, 552, L1
- Carswell, R. F., Webb, J. K., Baldwin, J. A., & Atwood, B. 1987, *ApJ*, 319, 709
- Cen, R., Miralda-Escudé, J., Ostriker, J. P., & Rauch, M. 1994, *ApJ*, 437, L9
- Coil, A. L., Newman, J. A., Cooper, M. C., Davis, M., Faber, S. M., Koo, D. C., & Willmer, C. N. A. 2006, *ApJ*, 644, 671
- Coles, P., & Jones, B. 1991, *MNRAS*, 248, 1
- Cooke, A. J., Espey, B., & Carswell, R. F. 1997, *MNRAS*, 284, 552
- Cristiani, S., D’Odorico, S., Fontana, A., Giallongo, E., & Savaglio, S. 1995, *MNRAS*, 273, 1016
- Croft, R. A. C. 2004, *ApJ*, 610, 642
- Croom, S., Boyle, B., Shanks, T., Outram, P., Smith, R., Miller, L., Loaring, N., Kenyon, S., & Couch, W. 2004, in *ASP Conf. Ser. 311: AGN Physics with the Sloan Digital Sky Survey*, ed. G. T. Richards & P. B. Hall, 457–+
- Croom, S. M., Boyle, B. J., Shanks, T., Smith, R. J., Miller, L., Outram, P. J., Loaring, N. S., Hoyle, F., & da Ângela, J. 2005, *MNRAS*, 356, 415
- Croom, S. M., Shanks, T., Boyle, B. J., Smith, R. J., Miller, L., Loaring, N. S., & Hoyle, F. 2001, *MNRAS*, 325, 483
- Davé, R., Hernquist, L., Katz, N., & Weinberg, D. H. 1999, *ApJ*, 511, 521
- Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, *ApJ*, 292, 371
- Desjacques, V., Nusser, A., & Sheth, R. K. 2007, *MNRAS*, 374, 206
- Ferland, G. J., Peterson, B. M., Horne, K., Welsh, W. F., & Nahar, S. N. 1992, *ApJ*, 387, 95
- Ferrarese, L. 2002, *ApJ*, 578, 90
- Gaskell, C. M. 1982, *ApJ*, 263, 79
- Giallongo, E., Cristiani, S., D’Odorico, S., Fontana, A., & Savaglio, S. 1996, *ApJ*, 466, 46
- Guimarães, R., Petitjean, P., Rollinde, E., de Carvalho, R. R., Djorgovski, S. G., Srianand, R., Aghaee, A., & Castro, S. 2007, *MNRAS*, 238
- Gunn, J. E., & Gott, J. R. I. 1972, *ApJ*, 176, 1
- Gunn, J. E., & Peterson, B. A. 1965, *ApJ*, 142, 1633
- Hennawi, J. F., & Prochaska, J. X. 2006, *ArXiv Astrophysics e-prints*
- Hennawi, J. F., Prochaska, J. X., Burles, S., Strauss, M. A., Richards, G. T., Schlegel, D. J., Fan, X., Schneider, D. P., Zakamska, N. L., Oguri, M., Gunn, J. E., Lupton, R. H., & Brinkmann, J. 2006, *ApJ*, 651, 61
- Hernquist, L., Katz, N., Weinberg, D. H., & Miralda-Escudé, J. 1996, *ApJ*, 457, L51+
- Hockney, R. W., & Eastwood, J. W. 1988, *Computer simulation using particles* (Bristol: Hilger, 1988)
- Hopkins, P. F., Hernquist, L., Cox, T. J., Di Matteo, T., Robertson, B., & Springel, V. 2006a, *ApJS*, 163, 1
- Hopkins, P. F., Hernquist, L., Cox, T. J., Robertson, B., & Springel, V. 2006b, *ApJS*, 163, 50
- Hopkins, P. F., Lidz, A., Hernquist, L., Coil, A. L., Myers, A. D., Cox, T. J., & Spergel, D. N. 2006c, *ArXiv Astrophysics e-prints*
- Hopkins, P. F., Richards, G. T., & Hernquist, L. 2006d, *ArXiv Astrophysics e-prints*
- Hui, L., & Gnedin, N. Y. 1997, *MNRAS*, 292, 27
- Hui, L., Gnedin, N. Y., & Zhang, Y. 1997, *ApJ*, 486, 599
- Hui, L., Haiman, Z., Zaldarriaga, M., & Alexander, T. 2002, *ApJ*, 564, 525
- Jena, T., Norman, M. L., Tytler, D., Kirkman, D., Suzuki, N., Chapman, A., Melis, C., Paschos, P., O’Shea, B., So, G., Lubin, D., Lin, W.-C., Reimers, D., Janknecht, E., & Fechner, C. 2005, *MNRAS*, 361, 70
- Jenkins, A., Frenk, C. S., White, S. D. M., Colberg, J. M., Cole, S., Evrard, A. E., Couchman, H. M. P., & Yoshida, N. 2001, *MNRAS*, 321, 372
- Juszkiewicz, R., Weinberg, D. H., Amsterdamski, P., Chodorowski, M., & Bouchet, F. 1995, *ApJ*, 442, 39
- Katz, N., Weinberg, D. H., Hernquist, L., & Miralda-Escudé, J. 1996, *ApJ*, 457, L57+
- Kauffmann, G., & Haehnelt, M. G. 2002, *MNRAS*, 332, 529
- Kim, T.-S., Carswell, R. F., Cristiani, S., D’Odorico, S., & Giallongo, E. 2002, *MNRAS*, 335, 555
- Kim, Y.-R., & Croft, R. 2007, *ArXiv Astrophysics e-prints*
- Kirkman, D., & Tytler, D. 1997, *ApJ*, 484, 672
- Kirkman, D., Tytler, D., Suzuki, N., Melis, C., Hollywood, S., James, K., So, G., Lubin, D., Jena, T., Norman, M. L., & Paschos, P. 2005, *MNRAS*, 360, 1373
- Kormendy, J., & Richstone, D. 1995, *ARA&A*, 33, 581
- Kulkarni, V. P., & Fall, S. M. 1993, *ApJ*, 413, L63
- Lidz, A., Heitmann, K., Hui, L., Habib, S., Rauch, M., & Sargent, W. L. W. 2006a, *ApJ*, 638, 27
- Lidz, A., Hopkins, P. F., Cox, T. J., Hernquist, L., & Robertson, B. 2006b, *ApJ*, 641, 41
- Loeb, A., & Eisenstein, D. J. 1995, *ApJ*, 448, 17
- Lu, L., Sargent, W. L. W., Womble, D. S., & Takada-Hidai, M. 1996, *ApJ*, 472, 509
- Lu, L., Wolfe, A. M., & Turnshek, D. A. 1991, *ApJ*, 367, 19
- Madau, P., Haardt, F., & Rees, M. J. 1999, *ApJ*, 514, 648
- McDonald, P., & Miralda-Escudé, J. 2001, *ApJ*, 549, L11
- McDonald, P., Miralda-Escudé, J., Rauch, M., Sargent, W. L. W., Barlow, T. A., & Cen, R. 2001, *ApJ*, 562, 52
- McDonald, P., Miralda-Escudé, J., Rauch, M., Sargent, W. L. W., Barlow, T. A., Cen, R., & Ostriker, J. P. 2000, *ApJ*, 543, 1
- McDonald, P., Seljak, U., Burles, S., Schlegel, D. J., Weinberg, D. H., Cen, R., Shih, D., J., S., Schneider, D. P., Bahcall, N. A., Briggs, J. W., Brinkmann, J., Brunner, R. J., Fukugita, M., Gunn, J. E., Ivezić, Ž., Kent, S., Lupton, R. H., & Vanden Berk, D. E. 2006, *ApJS*, 163, 80
- Meiksin, A., & White, M. 2001, *MNRAS*, 324, 141
- , 2003, *MNRAS*, 342, 1205
- , 2004, *MNRAS*, 350, 1107
- Miralda-Escudé, J., Cen, R., Ostriker, J. P., & Rauch, M. 1996, *ApJ*, 471, 582
- Osterbrock, D. E., & Ferland, G. J. 2006, *Astrophysics of gaseous nebulae and active galactic nuclei* (Astrophysics of gaseous nebulae and active galactic nuclei, 2nd ed. by D.E. Osterbrock and G.J. Ferland. Sausalito, CA: University Science Books, 2006)
- Page, L., Hinshaw, G., Komatsu, E., Nolte, M. R., Spergel, D. N., Bennett, C. L., Barnes, C., Bean, R., Dore, O., Halpern, M., Hill, R. S., Jarosik, N., Kogut, A., Limon, M., Meyer, S. S., Odegard, N., Peiris, H. V., Tucker, G. S., Verde, L., Weiland, J. L., Wollack, E., & Wright, E. L. 2006, *ArXiv Astrophysics e-prints*
- Porciani, C., Magliocchetti, M., & Norberg, P. 2004, *MNRAS*, 355, 1010
- Porciani, C., & Norberg, P. 2006
- Prada, F., Klypin, A. A., Simonneau, E., Betancort-Rijo, J., Patiri, S., Gottlöber, S., & Sanchez-Conde, M. A. 2006, *ApJ*, 645, 1001
- Press, W. H., & Rybicki, G. B. 1993, *ApJ*, 418, 585
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, *Numerical recipes in C. The art of scientific computing* (Cambridge: University Press, —c1992, 2nd ed.)
- Rauch, M., Miralda-Escudé, J., Sargent, W. L. W., Barlow, T. A., Weinberg, D. H., Hernquist, L., Katz, N., Cen, R., & Ostriker, J. P. 1997, *ApJ*, 489, 7
- Richards, G. T., Vanden Berk, D. E., Reichard, T. A., Hall, P. B., Schneider, D. P., SubbaRao, M., Thakar, A. R., & York, D. G. 2002, *AJ*, 124, 1
- Ricotti, M., Gnedin, N. Y., & Shull, J. M. 2000, *ApJ*, 534, 41
- Rollinde, E., Srianand, R., Theuns, T., Petitjean, P., & Chand, H. 2005, *MNRAS*, 361, 1015
- Rybicki, G. B., & Lightman, A. P. 1979, *Radiative processes in astrophysics* (New York, Wiley-Interscience, 1979. 393 p.)
- Schaye, J., Theuns, T., Rauch, M., Efstathiou, G., & Sargent, W. L. W. 2000, *MNRAS*, 318, 817
- Schirber, M., Miralda-Escudé, J., & McDonald, P. 2004, *ApJ*, 610, 105
- Scott, J., Bechtold, J., Dobrzycki, A., & Kulkarni, V. P. 2000, *ApJS*, 130, 67
- Seljak, U., McDonald, P., & Makarov, A. 2003, *MNRAS*, 342, L79
- Shapley, A. E., Steidel, C. C., Pettini, M., Adelberger, K. L., & Erb, D. K. 2006, *ArXiv Astrophysics e-prints*
- Shen, Y., Strauss, M. A., Oguri, M., Hennawi, J. F., Fan, X., Richards, G. T., Hall, P. B., Gunn, J. E., Schneider, D. P., Szalay, A. S., Thakar, A. R., Vanden Berk, D. E., Anderson, S. F., Bahcall, N. A., Connolly, A. J., & Knapp, G. R. 2007, *AJ*, 133, 2222
- Songaila, A., Hu, E. M., Cowie, L. L., & McMahon, R. G. 1999, *ApJ*, 525, L5

- Spergel, D. N., Bean, R., Dore, O., Nolte, M. R., Bennett, C. L., Hinshaw, G., Jarosik, N., Komatsu, E., Page, L., Peiris, H. V., Verde, L., Barnes, C., Halpern, M., Hill, R. S., Kogut, A., Limon, M., Meyer, S. S., Odegard, N., Tucker, G. S., Weiland, J. L., Wollack, E., & Wright, E. L. 2006, *ArXiv Astrophysics e-prints*
- Springel, V. 2005, *MNRAS*, 364, 1105
- Springel, V., & Hernquist, L. 2003, *MNRAS*, 339, 312
- Springel, V., White, S. D. M., Jenkins, A., Frenk, C. S., Yoshida, N., Gao, L., Navarro, J., Thacker, R., Croton, D., Helly, J., Peacock, J. A., Cole, S., Thomas, P., Couchman, H., Evrard, A., Colberg, J., & Pearce, F. 2005, *Nature*, 435, 629
- Srianand, R., & Khare, P. 1996, *MNRAS*, 280, 767
- Steidel, C. C., Pettini, M., & Adelberger, K. L. 2001, *ApJ*, 546, 665
- Telfer, R. C., Zheng, W., Kriss, G. A., & Davidsen, A. F. 2002, *ApJ*, 565, 773
- Theuns, T., Leonard, A., Efstathiou, G., Pearce, F. R., & Thomas, P. A. 1998, *MNRAS*, 301, 478
- Tytler, D., & Fan, X.-M. 1992, *ApJS*, 79, 1
- Tytler, D., Kirkman, D., O'Meara, J. M., Suzuki, N., Orin, A., Lubin, D., Paschos, P., Jena, T., Lin, W.-C., Norman, M. L., & Meiksin, A. 2004a, *ApJ*, 617, 1
- Tytler, D., O'Meara, J. M., Suzuki, N., Kirkman, D., Lubin, D., & Orin, A. 2004b, *AJ*, 128, 1058
- Vanden Berk, D. E., Richards, G. T., Bauer, A., Strauss, M. A., Schneider, D. P., Heckman, T. M., York, D. G., Hall, P. B., Fan, X., Knapp, G. R., Anderson, S. F., Annis, J., Bahcall, N. A., Bernardi, M., Briggs, J. W., Brinkmann, J., Brunner, R., Burles, S., Carey, L., Castander, F. J., Connolly, A. J., Crocker, J. H., Csabai, I., Doi, M., Finkbeiner, D., Friedman, S., Frieman, J. A., Fukugita, M., Gunn, J. E., Hennessy, G. S., Ivezić, Ž., Kent, S., Kunszt, P. Z., Lamb, D. Q., Leger, R. F., Long, D. C., Loveday, J., Lupton, R. H., Meiksin, A., Merelli, A., Munn, J. A., Newberg, H. J., Newcomb, M., Nichol, R. C., Owen, R., Pier, J. R., Pope, A., Rockosi, C. M., Schlegel, D. J., Siegmund, W. A., Smee, S., Snir, Y., Stoughton, C., Stubbs, C., SubbaRao, M., Szalay, A. S., Szokoly, G. P., Tremonti, C., Uomoto, A., Waddell, P., Yanny, B., & Zheng, W. 2001, *AJ*, 122, 549
- Verner, D. A., Ferland, G. J., Korista, K. T., & Yakovlev, D. G. 1996, *ApJ*, 465, 487
- Viel, M., Haehnelt, M. G., & Springel, V. 2006, *MNRAS*, 367, 1655
- Wake, D. A., Miller, C. J., Di Matteo, T., Nichol, R. C., Pope, A., Szalay, A. S., Gray, A., Schneider, D. P., & York, D. G. 2004, *ApJ*, 610, L85
- White, M. 2002, *ApJS*, 143, 241
- Williger, G. M., Baldwin, J. A., Carswell, R. F., Cooke, A. J., Hazard, C., Irwin, M. J., McMahon, R. G., & Storrie-Lombardi, L. J. 1994, *ApJ*, 428, 574
- Zaldarriaga, M. 2002, *ApJ*, 564, 153
- Zhang, Y., Anninos, P., & Norman, M. L. 1995, *ApJ*, 453, L57+
- Zheng, W., Kriss, G. A., Telfer, R. C., Grimes, J. P., & Davidsen, A. F. 1997, *ApJ*, 475, 469