EXTENDED Lyα EMISSION AROUND YOUNG QUASARS: A CONSTRAINT ON GALAXY FORMATION

ZOLTÁN HAIMAN
Princeton University Observatory, Princeton, NJ 08544; zoltan@astro.princeton.edu

AND

MARTIN J. REES
Institute of Astronomy, Madingley Road, Cambridge, CB3 0HA, UK; mjr@ast.cam.ac.uk

Received 2001 March 26; accepted 2001 March 26

ABSTRACT

The early stage in the formation of a galaxy inevitably involves a spatially extended distribution of infalling, cold gas. If a central luminous quasar turned on during this phase, it would result in significant extended Lyα emission, possibly accompanied by other lines. For halos condensing at redshifts $3 \leq z \leq 8$ and having virial temperatures $2 \times 10^5 \, K \leq T_{\text{vir}} \leq 2 \times 10^6 \, K$, this emission results in a "fuzz" of characteristic angular diameter of a few arcseconds and surface brightness $\sim 10^{-18}$ to $10^{-16} \, \text{ergs s}^{-1} \, \text{cm}^{-2} \, \text{arcsec}^{-2}$. The fuzz around bright, high-redshift quasars could be detected in deep narrowband imaging with current telescopes, providing a direct constraint on galaxy formation models. The absence of detectable fuzz might suggest that most of the protogalaxy's gas settles to a self-gravitating disk before a quasar turns on. However, continued gas infall from large radii, or an on-going merger spreading cold gas over a large solid angle, during the luminous quasar phase could also result in extended Lyα emission, and can be constrained by deep narrowband imaging.

Subject heading: black hole physics — cosmology: theory — galaxies: formation — quasars: general

On-line material: color figures

1. INTRODUCTION

There has been considerable progress recently in the study of the formation of galaxies and quasars. Both the galaxy (Steidel et al. 1999) and quasar (Fan et al. 2000a) luminosity functions are now observationally determined to redshifts of $z \sim 4$, probing into the epoch when at least some of these objects are still expected to be assembling. Although the evolutions of the galaxy and quasar populations is generally expected to be connected, the detailed nature of this link is yet to be elucidated. One possibility is that quasars represent a brief phase in the early life of each galaxy. Indeed, theoretically, one might expect quasar activity to be triggered by the mergers that a galaxy is experiencing during its assembly (e.g., Cavaliere & Vittorino 2000; Kauffmann & Haehnelt 2000). This assertion has some observational support: the galaxy abundance appears to peak at a somewhat later redshift than does the quasar abundance (e.g., Madau & Pozzetti 2000), and at least infrared-selected quasars appear to preferentially reside in morphologically disturbed hosts (e.g., Baker & Clements 1997).

It is therefore interesting to consider a quasar that turns on within an assembling protogalaxy. The behavior of the gas inside a dark matter (DM) condensation during the formation of disk galaxies has been investigated in semi-analytic schemes (White & Rees 1978; Fall & Efstathiou 1980; Mo, Mao, & White 1998; van den Bosch & Dalcanton 2000), and numerical simulations (Katz 1991; Navarro & White 1994; Navarro, Frenk & White 1997, hereafter NFW; Navarro & Steinmetz 1997; Moore et al. 1999). In these studies, the bulk of the baryons in the halo cool and settle into a rotationally supported disk. In the simplest picture, the disk material originates as smooth gas, collapsing from the virial radius $R_{\text{vir}}$ to its final orbital radius of $\sim \lambda R_{\text{vir}}$, where $\lambda \sim 0.05$ is the typical spin parameter. Numerical simulations have revealed a more complex process, where a fraction of the infalling gas forms smaller clumps early on; these clumps then progressively merge together, collide, and dissipate to form larger systems.

A robust feature of galaxy formation is, at least in the early stages, a spatially extended distribution of gas. At the densities expected at $z > 2$ for this gas ($\gtrsim 100$ times the background density), the radiative cooling timescales are shorter than the typical dynamical times. In the absence of heat input from stars or a quasar, a significant fraction of this gas would therefore be cold ($T \lesssim 10^4 \, \text{K}$) and neutral (Fall & Rees 1985). As it contracts inside the DM halo, the cold gas is heated by the halo potential; this heat is dissipated largely via collisional excitation of the Lyα, and possible other (metal) lines. The resulting line radiation results in a potentially observable, extended, low surface brightness Lyα "fuzz" (Haiman, Spaans, & Quataert 2000; Fardal et al. 2000).

The presence of a bright central quasar during this phase could strongly enhance the surface brightness in the Lyα line, since a significant fraction of the quasar's ionizing radiation could be reprocessed into recombination radiation in the same line. The purpose of this paper is to quantify the expected Lyα fluxes and describe constraints that the presence or absence of extended Lyα "fuzz" around luminous quasars implies for the host galaxy. This paper is organized as follows: in § 2 we describe a simple toy model for the distribution of cold gas in a protogalaxy; in § 3 we characterize the resulting Lyα "fuzz"; in § 4 we discuss the implied constraints on galaxy formation; and in § 5, we summarize our conclusions. Throughout this work, we adopt a ΛCDM cosmology with $\Omega_m = 0.3$, $\Omega_b = 0.04$, $\Omega_\Lambda = 0.7$, $h = 0.7$, and $\sigma_8 = 0.9$.

2. THE AMOUNT OF COLD GAS: A SPHERICAL MODEL

In this section we describe a simple model for the structure of a spherically symmetric, two-phase gas in a DM halo.
without a central ionizing source. We shall assume that the gas has a centrally condensed radial profile, \( \rho(r) \), and that it consists of a two-phase medium with a cold (\( T \approx 10^4 \) K) mass fraction \( f(r) \) and a hot (\( T \approx T_{\text{vir}} \)) mass fraction \( 1 - f(r) \). The physical state of the gas in a realistic protogalaxy is, of course, likely to be much more complicated, with asymmetric infall of preexisting dense clumps and the settling of gas in the central regions into a disk. Nevertheless, this model will serve as a reference point for our discussion in § 3 and § 4 below.

For sake of concreteness, we assume that the gas initially settles to a radial profile \( \rho_{\text{gas}}(r) \) that satisfies hydrostatic equilibrium within a dark matter halo. For the profile of the halo, we follow the description in NFW with a concentration parameter \( c = 5 \). However, we assume that the average enclosed mass density is a fraction \( \Delta_c \) of the critical density, where we obtain \( \Delta_c \approx 18\pi^2 \) from the spherical top-hat collapse (rather than using the fixed value of 200). To obtain \( \rho_{\text{gas}}(r) \), we also assume that the gas is isothermal at the virial temperature of the halo (Makino, Sasaki & Suto 1998). It is useful to note that under these assumptions, the gas is centrally condensed with a flat core and has a mean internal (volume-averaged) “clumping” of \( \langle \rho_{\text{gas}}^2 \rangle \approx 2.8 \times 10^5 \rho_b^2 \) relative to the background baryon density \( \rho_b \).

In order to obtain the density \( \rho_{\text{cold}}(r) \) of neutral gas, we next assume that a mass fraction \( f(r) \) of the gas cools and condenses out into a cold phase with \( T = 10^4 \) K. Pressure balance with the remaining hot, ionized gas implies that the densities of the two components are

\[
\begin{align*}
\eta_{\text{cold}} &= \frac{\rho_{\text{cold}}}{\rho_{\text{gas}}} = f + (1-f) \frac{T_{\text{vir}}}{10^4} \geq 1, \\
\eta_{\text{hot}} &= \frac{\rho_{\text{hot}}}{\rho_{\text{gas}}} = (1-f) + f \frac{10^4}{T_{\text{vir}}} \leq 1. 
\end{align*}
\]

The value of \( f = f(r) \) is determined from the condition that the cooling time of the rarefied hot component is equal to the age of the system. The cooling time is given by \( t_{\text{cool}} = (3/2) \mu m_p c_s^2 T_{\text{vir}} / (\rho_{\text{hot}} \Lambda) \), where \( \Lambda \approx 10^{-23} \) ergs s\(^{-1}\) cm\(^{-3}\) is the cooling function at \( T_{\text{vir}} \). We conservatively adopt a metal-free cooling function (B"ohringer & Hensler 1989). For the age of the system, we adopt 20\% of the Hubble time, \( t_{\text{age}} = 0.2(6\pi G \rho_b)^{-1/2} \), where \( \rho_b \) is the total (dark matter + baryons) mean background density at redshift \( z \). This is roughly the “mass doubling” time for halos of interest in the extended Press-Schechter formalism (see Lacey & Cole 1993; and Haiman et al. 2000 for a discussion).

Illustrative results for the cold fraction \( f(r) \) as a function of radius under the above assumptions are shown in Figure 1. The dashed and solid curves describe halos at redshifts \( z = 3 \), and \( z = 5 \), respectively. At both redshifts, four different halo sizes are shown, with virial temperatures of \( T_{\text{vir}} = 2 \times 10^5 \), \( 4 \times 10^5 \), \( 10^5 \), and \( 10^7 \) K (top to bottom). At \( z = 3 \), these correspond to halo masses \( M_{\text{halo}} \approx 4 \times 10^{10}, 10^{11}, 5 \times 10^{11}, \text{and } 10^{13} M_\odot \); at \( z = 5 \), the halo masses are a factor of \( \sim 2 \) smaller.

As expected, the cold fraction is a function of radius in each case, increasing toward \( r = 0 \) where densities are higher and cooling times are shorter. The cold fraction is larger for smaller halos, which have smaller initial binding energies, and hence cool more rapidly. The cold fraction increases with redshift, because of the higher densities and shorter cooling times. Although our model is highly idealized, it captures the above scalings (which are expected to be robust, as long as the cold fraction is determined by the cooling time), and provides a conservative estimate of the amount of cold gas. For comparison, we note that Mo & Miralda-Escudé (1996) have used a different, simplified model of a two-phase medium in order to model quasar line absorption systems. In their model, nearly all of the gas is cold within the “cooling radius,” defined as the radius at which the cooling time equals the Hubble time (exceeding \( R_{\text{vir}} \) for all halos considered here).

Finally, we emphasize that our model assumptions apply only within the virial radius, where the gas has been shock heated. In the rest of this paper, we focus on the reprocessing of quasar light within protogalactic halos. However, the clumpy background gas outside the virial region can itself reprocess ionizing UV radiation into Ly\( \alpha \) emission. Indeed, Gould & Weinberg (1996) have shown that Ly\( \alpha \) absorption systems, illuminated by the UV background, can cause significant Ly\( \alpha \) fluorescence (see Bunker, Marleau, & Graham 1998 for a current status of observations). Numerical simulations suggest that the background gas near a collapsing protogalaxy is highly clumped and hence could reprocess any ionizing radiation from the protogalaxy into Ly\( \alpha \) emission. This type of reprocessing of quasar light could be the interpretation of extended Ly\( \alpha \)-emitting blobs found in the vicinity of quasars (Hu et al. 1991; Hu, McMahon, & Egami 1996). Although these blobs are likely associated with small satellite protogalaxies, their Ly\( \alpha \) emission could be dominated by reprocessed quasar light.

3. IONIZATION BY A CENTRAL SOURCE AND THE ACCOMPANYING LY\( \alpha \) FUZZ

We next consider a quasar that turns on at the center of the halo described above. In general, the cold gas will then be photoionized by the quasar’s UV radiation, out to a
radius that depends on the quasar’s ionizing luminosity. Utilizing the spatial distribution of cold gas obtained in our models, we first argue that reasonably sized quasar black holes (BHs) can keep most of the cold phase photoionized. We then compute the characteristic Lyα surface brightness of such quasar-illuminated protogalaxies.

### 3.1. Effect of an Ionizing Source

The radius of the photoionized Strömgren sphere around a central quasar embedded in a spherical halo is given by

\[
R_{H\alpha} = \left( \frac{3N_{\mathrm{ph}}}{4\pi \alpha_B \langle n_{\mathrm{HI}}^2 \rangle} \right)^{1/3}, \quad (2)
\]

where \(N_{\mathrm{ph}}\) is the ionizing photon production rate of the central source, \(\alpha_B\) is the hydrogen recombination coefficient evaluated at \(\approx 10^4\) K, and \(\langle n_{\mathrm{HI}}^2 \rangle\) is the volume averaged mean squared density of cold hydrogen within \(R_{H\alpha}\). Note that in our models, the cold gas is compressed by a factor of \(n_{\mathrm{cold}}\), which enhances the local recombination rate by \(n_{\mathrm{cold}}^2\). However, the cold gas occupies only a fraction \(f\) of the volume, and hence the compression increases the total recombination rate by an overall factor of \(f n_{\mathrm{cold}}\).

Assuming a fixed cold mass fraction \(f = 0.5\) across the halo, and assuming further that the central BH shines at the Eddington luminosity (for typical quasar spectra, this corresponds to an ionizing photon production rate of \(N_{\mathrm{ph}} \approx 6 \times 10^{47}\) photons s\(^{-1}\) per \(M_\odot\) of BH mass; see Cen & Haiman 2000), we then find the required size of the BH so that the Strömgren sphere extends all the way out to the virial radius:

\[
M_{\mathrm{bh}} \approx 6 \times 10^8 M_\odot \left( \frac{M_{\mathrm{halo}}}{10^{12} M_\odot} \right)^{5/3} \left( \frac{1 + z}{6} \right)^4. \quad (3)
\]

Here we have utilized a relation between the halo mass, radius, and virial temperature from NFW. Equation (3) can be understood by recalling the scalings \(R_{H\alpha}^3 \propto M_{\mathrm{bh}}/n_{\mathrm{HI}}^{-1} \propto M_{\mathrm{bh}} T_{\mathrm{vir}}^{-1}(1 + z)^{-6} \); \(T_{\mathrm{vir}} \propto M_{\mathrm{halo}}^{2/3} (1 + z)^{1/2}\); and \(R_{\mathrm{vir}} \propto M_{\mathrm{halo}}^{1/3} (1 + z)^{-1}\).

Equation (3) reveals that in most halos of interest (3 \(\lesssim z \lesssim 8\); \(10^{10} M_\odot \lesssim M_{\mathrm{halo}} \lesssim 10^{12} M_\odot\)), converting 2% of the gas mass into a central BH is sufficient to keep most of the cold gas photoionized. This conclusion is conservative, since we assumed a constant cold fraction \(f = 0.5\), which maximizes the internal clumping and the total recombination rate (see eq. [1]). Using the profiles of \(f(z)\) obtained in our models, we have verified that for all halos considered in this work, the cold gas can be fully ionized by still smaller BHs, in all cases with \(M_{\mathrm{bh}} \leq 0.01 \times \Omega_\Lambda \Omega_\mathrm{m} M_{\mathrm{halo}}\). Although the mass of BHs in protogalaxies at \(z \gtrsim 3\) are unknown, ratios as large as 1% of the gas mass would be consistent with the sizes of supermassive BHs found in nearby galaxies (Magorrian et al. 1998; Gebhardt et al. 2000; Ferrarese & Merritt 2000), especially since the ratio can evolve (and decrease) as a function of redshift (Menou, Haiman, & Narayanan 2000).

### 3.2. Surface Brightness and Angular Size of Lyα Fuzz

Assuming that all of the cold gas in the model halo is photoionized, we obtain the total Lyα line luminosity as

\[
L_\alpha = 0.68 \times E_a \int_{0}^{R_{\mathrm{vir}}} 4\pi r^2 dr \left( \frac{f}{n_{\mathrm{cold}}} \right) n_{\mathrm{HI}} \alpha_B, \quad (4)
\]

where \(E_a = 10.2\)eV is the energy of a Lyα photon, and the integral represents the total recombination rate of photoionized gas within the halo. We have explicitly included the volume filling factor \(f/n_{\mathrm{cold}}\) of the cold gas, whose hydrogen number density is \(n_{\mathrm{HI}} = 0.76 \rho_{\mathrm{cold}}/m_p\), and the fraction 0.68 of case B recombinations that yield a Lyα photon (Osterbrock 1989). We also note that additional recombination lines might be observable, including those of heavy elements, provided the halo gas is sufficiently preenriched (see De Breuck et al. 2000 for a review of relevant observations).

For the halos shown in Figure 1, the values of \(L_\alpha\) are 7 \(\times 10^{40}\), 3 \(\times 10^{42}\), 4 \(\times 10^{43}\), and 3 \(\times 10^{46}\) erg s\(^{-1}\) (at \(z = 3\)), and 10\(^{42}\), 5 \(\times 10^{42}\), 5 \(\times 10^{43}\), and 4 \(\times 10^{46}\) ergs s\(^{-1}\) (at \(z = 5\)). The spatial extent of the Lyα-emitting gas is a fraction of the virial radius \(R_{\mathrm{vir}} \sim 10^2\) kpc. As Figure 1 shows, this fraction varies inversely with the virial temperature of the halo. Here we define \(R_{1/2}\) as the radius at which the cold fraction is \(f = 0.5\) and assume that the Lyα radiation is emitted from within a characteristic radius, which we take to be \(R_\alpha \equiv \min (R_{1/2}, R_{\mathrm{vir}})\). The characteristic angular size \(\theta_\alpha\) of the Lyα-emitting fuzz should then be \(\sim \theta_\alpha = R_\alpha/d_A\), where \(d_A\) is the angular diameter distance.

We find that for the halos shown in Figure 1, \(\theta_\alpha \approx 2\)–3 arcsec. The angular size varies relatively little with redshift and virial temperature, and implies that the typical Lyα-emitting fuzz should appear extended and can be resolved with optical instruments.

The characteristic Lyα surface brightness within \(\theta_\alpha, F_\alpha = L_\alpha/4\pi d_A^2/\theta_\alpha^2\) (where \(d_A\) is the luminosity distance), is shown in Figure 2 as a function of \(T_{\mathrm{vir}}\). The solid curves correspond to redshifts \(z = 3, 4, 5, 6, 7, 8\) (top to bottom near \(10^7\) K). This figure reveals several interesting features. First, the
surface brightness is generally high, and all but the smallest halos should be potentially detectable with current instruments (see discussion below and in Haiman et al. 2000). Second, nearly all of the halo gas remains cold up to a virial temperature of $\sim 3 \times 10^5$ K (see Fig. 1). In this case, the Lyz fuzz extends out to the virial radius, and the surface brightness has the simple scaling $F_\alpha \propto \rho^2 V/R_{200}^3(1 + z)^4 \propto T_{\text{vir}}^{-1/2}$ $(1 + z)^{1/2}$ (here $V \propto R_{200}$ is the volume of the halo); i.e., the surface brightness weakly increases with both virial temperature and redshift. The latter result arises primarily from the strong dependence of the recombination rate on redshift, $\rho^2 \propto (1 + z)^6$ [vs. the $(1 + z)^4$ surface brightness dimming]. Third, for higher virial temperatures ($\geq 3 \times 10^5$ K), the cold fraction $f$ falls somewhat below unity, allowing the cold phase to be compressed and enhancing the recombination rate and the surface brightness. In this range of virial temperatures, the surface brightness scales approximately as $F_\alpha \propto T_{\text{vir}}^{-2}$ [note that $\Lambda(T) \propto T^{1/2}$]. Both the stronger dependence on $T_{\text{vir}}$ and the inverse scaling on redshift result from the dependence of the cold fraction on $T_{\text{vir}}$. Note that for still higher virial temperatures, the cold fraction would decrease to negligibly small values and the surface brightness would drop sharply, but this happens only for exceedingly large halos.

In order to illustrate the robustness of the above conclusions, we have computed the surface brightness in two variants of our model. First, to facilitate comparison with earlier work, we adopt $t_{\text{cool}} = t_{\text{Hub}}$ rather than $t_{\text{cool}} = 0.2 t_{\text{Hub}}$, as the condition used to compute the cold fraction $f$. The effect of this change on the surface brightness at $z = 5$ is shown by the dotted curve in Figure 2. As expected, the cold fraction remains unity for halos up to a higher virial temperature, which reduces the surface brightness by an order of magnitude at $T_{\text{vir}} \geq 10^6$ K. Second, we note that our simplified model would conflict with observations if we applied it to local galaxy clusters (see, e.g., Fabian 1994 for a review). Our model would predict a nonnegligible amount of cold gas at the center of at least the lowest mass clusters (e.g., $M \sim 10^{14} M_\odot$ at redshift $z \sim 0$), where observations show little evidence for cold gas. This could mean that, in clusters, the cool gas converts quickly and efficiently into (low-mass) stars. Alternatively, there could be heat input (e.g., from higher mass star formation, SNe, etc.) that effectively slows down the cooling. To mimic this latter scenario, we have increased the cooling time by a factor of 10, equivalent to requiring $t_{\text{cool}} = 0.2 t_{\text{Hub}}$. This condition ensures that $f \sim 0$ is predicted for all $z \sim 0$ galaxy clusters. The effect of this change (at $z = 5$) is shown by the dashed curve in Figure 2. The cold fraction is decreased, which enhances the surface brightness by up to an order of magnitude for halo with $T_{\text{vir}} \geq 10^5$ K.

4. OBSERVATIONAL PROSPECTS: CONSTRAINTS ON GALAXY AND QUASAR FORMATION

The main results of the previous two sections are (1) in galaxy-sized halos ($T_{\text{vir}} \geq 10^6$ K) at redshifts $z \geq 3$, a significant fraction of the gas should be cold, and (2) if this gas is illuminated by ionizing radiation from a central quasar, it should be kept photoionized and result in a detectable Lyz fuzz of characteristic size of $\gtrsim 2''$ and surface brightness $\gtrsim 10^{-17}$ ergs s$^{-1}$ cm$^{-2}$ arcsec$^{-2}$. A handful of existing observations in various other contexts have already reached these relevant levels for the Lyz surface brightness. Examples are the imaging of a protocluster region of Lyman break galaxies at redshift $z \approx 3$ in a narrowband Lyz filter, reaching a sensitivity of $\sim 10^{-16}$ ergs s$^{-1}$ cm$^{-2}$ arcsec$^{-2}$ in a $\sim 16$ hr observation with the Palomar 200 inch telescope (Steidel et al. 2000). Similar sensitivities have been reached in $\sim 5$ hr with the Keck telescope in blank field searches for high-redshift Lyz galaxies (Cowie & Hu 1998) and should also be achievable by Subaru. Extended Lyz emission has also been detected around high-redshift radio galaxies (e.g., De Breuck et al. 2000) and radio-loud quasars (possibly related to outflows; e.g., Heckman et al. 1991a, 1991b; Bremer et al. 1992). Nevertheless, observations have not yet targeted bright quasars to search for extended Lyz emission at similar depths. Less sensitive observations (reaching a few times $10^{-16}$ ergs s$^{-1}$ cm$^{-2}$ arcsec$^{-2}$) did target bright quasars. Bremer et al. (1992) finds extended emission around two of the ten quasars, but a search of another twelve quasars at similar redshifts showed emission only around one source (Hu & Cowie 1987). Yet another search around a $z = 4.7$ quasar (Hu et al. 1996) has uncovered discrete Lyz-emitting companions, rather than an extended continuous “fuzz.”

In summary, the few existing observations, typically utilizing $\sim 1$ hr integrations on 4 m telescopes, probe surface brightnesses above $\sim 10^{-16}$ ergs s$^{-1}$ cm$^{-2}$ arcsec$^{-2}$ and reveal that luminous quasars are not generally enveloped by Lyz fuzz at this sensitivity. This puts a mild constraint on the models, although emission at this level is expected only in the largest halos (see Fig. 2). However, observations that probe bright quasars at about an order of magnitude deeper would either detect Lyz fuzz around most sources, or else lead to strong constraints on the type of models discussed here. Provided that the Lyz fuzz is detected, its surface brightness, shape and extent, together with the line profiles, would constitute an invaluable direct probe of galaxy formation.

It is interesting to consider constraints that would arise on galaxy and quasar formation should the current trend of not detecting fuzz persist at fainter fluxes. One possibility is that quasars turn on only during the later stages of galaxy formation, i.e., at a time when most of the cold gas has already settled to a thin disk, and/or turned into stars. A lack of Lyz fuzz would then be naturally explained by the absence of significant amounts of spatially extended cold gas during the luminous quasar phase. To avoid detectability at the surface brightness threshold of $\sim 10^{-17}$ ergs s$^{-1}$ cm$^{-2}$ arcsec$^{-2}$, Figure 2 suggests that in halos with $T_{\text{vir}} \geq 10^6$ K, the flux has to be reduced by a factor of $\gtrsim 10$ relative to the predictions of our simple models. This implies, in turn, that $\gtrsim 90\%$ of the cold gas must already have settled to a disk (or disappeared). It has been argued that the presence of a disk is indeed a prerequisite for the central BH to grow (Sellwood & Moore 1999); furthermore, cold gas in a thin disk is locally Toomre-unstable and might rapidly turn into stars (e.g., Mo et al. 1998). Independent “evidence” favoring this scenario is that the heavy element abundances in the broad-line region of quasars always appear to be high (Hamann 1999), even at high redshifts. This requires one or two generations of massive stars to precede the activation of the central quasar.

2 See http://www.subarutelescope.org.
There are other, less attractive scenarios without Ly\textsubscript{a} emission. If the cold gas is quickly turned into stars and/or collapses to the central regions or to a thin disk, this could eliminate Ly\textsubscript{a} reprocessing, at least until cold gas is replenished by further accretion or merger with another halo (i.e., resulting in a short Ly\textsubscript{a} duty cycle). However, this would also imply that a fraction \( \sim t_{\text{dyn}}/t_Q \gtrsim 10\% \) of all quasars should still show Ly\textsubscript{a} fuzz, unless the cold gas disappears on an exceedingly short timescale \( (\ll t_{\text{dyn}}) \) (see Haiman \& Hui 2000 for constraints on the lifetime \( t_Q \) of the luminous quasar phase). Alternatively, one might envision that the cold gas resides in clumps with a small covering factor, allowing most of the ionizing photons to leak out along line of sights traversing only hot (collisionally ionized) medium. However, this explanation requires a minimum cold clump size that exceeds the Jeans mass in the cold phase (Rees 1988). Hence, the postulated large clumps are unstable and would fragment to smaller pieces, increasing the covering factor to approximately unity. Yet another scenario with no Ly\textsubscript{a} fuzz is if the quasar (or its associated wind) has blown out most of the gas from the halo. However, in this case, one would still expect to see a fluorescent Ly\textsubscript{a} outflow around a fraction of quasars “caught in the act” of removing the gas.

Dust absorption might strongly suppress the Ly\textsubscript{a} flux escaping from a medium, even if the medium is optically thin in the Lyman continuum, and has been thought to cause the lack of detections of protogalaxies in early Ly\textsubscript{a} surveys. A strong quenching of Ly\textsubscript{a} by dust, however, does not necessarily occur (see, e.g., Neufeld 1991, or Pritchet 1994 for a general discussion), especially if most of the Ly\textsubscript{a} photons originate in an ionized layer with relatively low dust opacity. Furthermore, Ly\textsubscript{a}-emitting galaxies have been found at high redshift (e.g., Hu et al. 1996; Hu, Cowie, \& McMahon 1998), as expected in models with lower galactic dust abundance, and inhomogeneous dust distribution (Haiman \& Spaans 1999). Based on these observations, it would appear unlikely that dust can suppress the Ly\textsubscript{a} fuzz from around all high-redshift quasars. Indeed, the dust abundance in the early, spatially extended, collapsing phase of the high-redshift halos is likely to be significantly lower than in star-forming galaxies.

Even if dust does not suppress the Ly\textsubscript{a} emission itself, however, an important question is whether the nucleus can be obscured by dust and rendered undetectable at optical wavelengths. If this is typical, then Ly\textsubscript{a} fuzz should be expected around submillimeter sources, rather than around optical quasars. The extended Ly\textsubscript{a}-emitting blob of Steidel et al. (2000) has been found to be an exceptionally bright submillimeter source (Chapman et al. 2001), with no visible continuum source. The Ly\textsubscript{a} line luminosity for this object is \( \sim 2 \times 10^{10} L_\odot \), while the bolometric luminosity inferred from the submillimeter detections is greater than \( 10^{13} L_\odot \). This implies that less than 1% of the UV produced by the source is available to power the extended Ly\textsubscript{a} emission and that most of this UV is used up for this purpose (to explain the lack of any continuum source). It would be surprising if extended Ly\textsubscript{a} fuzz was typically produced in a similar way, since it requires a fine-tuning of the unobscured fraction to match the amount of surrounding cold gas. Nevertheless, it would be invaluable to target bright submillimeter sources (i.e., those with redshift estimates) in deep Ly\textsubscript{a} searches, to clarify what fraction of them do produce Ly\textsubscript{a} emission. The primary driver (stellar vs. quasar UV light) of both the submillimeter and Ly\textsubscript{a} emissions remain unclear in the Chapman et al. (2001) source (as well as in other bright submillimeter sources). Nevertheless, the lack of strong submillimeter emission in a second Ly\textsubscript{a} blob in the same data set suggests that the line emission is not necessarily powered by a dust-obscured source.

Our models imply a second interesting, although somewhat less stringent constraint on the amount of extended cold gas around quasars. As discussed in \S 3.1 above, we find that it is sufficient to convert \( \sim 1\% \) of the total gas mass into a central BH in order to keep most of the cold gas ionized (cf. eq. [3]). We emphasize that this must indeed happen if any ionizing radiation is to escape from the halo. Observations typically indicate that a large fraction of the ionizing radiation from quasars does escape, even for the quasar with the highest known redshift at \( z = 5.8 \) (Fan et al. 2000b). The abundance of this object implies a halo mass \( M_{\text{halo}} \sim 10^{13} M_\odot \), while its luminosity, under the assumption that it equals the Eddington limit, implies a BH mass of \( M_{\text{bh}} \sim 4 \times 10^9 M_\odot \) (Haiman \& Loeb 2001). For this halo, under the assumption of a constant cold fraction \( f = 0.5 \), equation (3) would imply that a BH as large as \( M_{\text{bh}} \sim 4 \times 10^{10} M_\odot \) is needed to ionize the cold gas and allow the ionizing radiation to escape. Based on the profile \( f(r) \) derived in our models, the requisite BH mass in smaller, \( M_{\text{bh}} \sim 1.3 \times 10^{10} M_\odot \). Nevertheless, this mass is close to 1% of the total gas mass and is a factor of \( \sim 2 \) higher than the BH mass inferred directly from the luminosity. We conclude that a large escape fraction of the ionizing continuum from bright, high-redshift quasars requires either (1) massive central BHs whose masses are a significant fraction (\( \sim 1\% \)) of the gas mass or (2) that some of the cold gas has settled to a disk (or disappeared in a blowout).

Finally, we note that any significant continued gas infall at large radius, or a major merger spreading cold gas over an extended region, could still reprocess much of the quasar’s radiation to Ly\textsubscript{a}, although the surface brightness at large radii might drop below detectable levels.

5. CONCLUSIONS

In this paper, we have studied what would happen to a protogalaxy if a bright quasar were to turn on during the early stages of its assembly. Using a simple spherical model for the distribution of a two-phase gas, we find that such a system should have a substantial amount of cold, photoionized gas with a spatially extended distribution. We expect this conclusion to be generic, owing to efficient radiative cooling at high redshifts. The cold, photoionized gas is detectable as an extended Ly\textsubscript{a} fuzz enshrouding the quasar, with a characteristic angular size of a few arcseconds and a surface brightness of \( \sim 10^{-18} \) to \( 10^{-16} \) ergs s\(^{-1}\) cm\(^{-2}\) arcsec\(^{-2}\). Existing observations have not yet targeted luminous quasars at these sensitivities, although at \( \sim 10^{-16} \) erg s\(^{-1}\) cm\(^{-2}\) arcsec\(^{-2}\), most quasars do not appear to be enveloped by extended emission. Future, deep Ly\textsubscript{a} imaging of few arcsecond regions around luminous quasars (as well as bright submillimeter sources) should be possible with current instruments. While a detection of Ly\textsubscript{a} fuzz would provide a direct probe of galaxy formation; nondetections at the level of \( 10^{-17} \) erg s\(^{-1}\) cm\(^{-2}\) arcsec\(^{-2}\) would already imply strong constraints.

We thank Eliot Quataert for useful discussions and an anonymous referee for comments that improved this paper.
This work began when M. J. R. was visiting Princeton University as Scribner Lecturer; he acknowledges support from this source and from the Royal Society. Z. H. was supported by NASA through the Hubble Fellowship grant HF-01119.01-99A, awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA under contract NAS 5-26555.

REFERENCES

Baker, A. C., & Clements, D. L. 1997, in Extragalactic Astronomy in the Infrared, ed. G. A. Mamon, T. X. Thuan, & J. Tran Thanh Van (Paris: Editions Frontieres), 295
Bremer, M. N., Fabian, A. C., Sargent, W. L., Steidel, C. C., Boksenberg, A., & Johnstone, R. M. 1992, MNRAS, 258, 23
Bohringer, H., & Hensler, G. 1989, A&A, 215, 147
Bunker, A. J., Marleau, F. R., & Graham, J. R. 1998, ApJ, 116, 2086
Cavaliere, A., & Vittorini, V. 2000, ApJ, 543, 599
Cen, R., & Haiman, Z. 2000, ApJ, 542, L75
Chapman, S. C., Lewis, G. F., Scott, D., Richards, E., Borys, C., Steidel, C. C., Adelberger, K. L., & Shapley, A. E. 2001, ApJ, 548, L17
Cowie, L. L., & Hu, E. M. 1998, AJ, 115, 1319
De Breuck, C., Röttgering, H., Miley, G., van Breugel, W., & Best, P. 2000, A&A, 362, 519
Fabian, A. C. 1994, ARA&A, 32, 277
Fall, S. M., & Efstathiou, G. 1980, MNRAS, 193, 189
Fall, S. M., & Rees, M. J. 1985, ApJ, 298, 18
Fan, X., et al. 2000a, AJ, 119, 1
Fan, X., et al. 2000b, ApJ, 540, 1167
Fardal, M. A., et al. 2000, ApJ, submitted, astro-ph/0007205
Ferrarese, L., & Merritt, D. 2000, ApJ, 539, L9
Gebhardt, K., et al. 2000, ApJ, 539, L13
Gould, A., & Weinberg, D. 1996, ApJ, 468, 462
Haiman, Z., & Hui, L. 2000, ApJ, 547, 27
Haiman, Z., & Loeb, A. 2001, ApJ, 552, 459
Haiman, Z., & Spaans, M. 1999, ApJ, 518, 138
Haiman, Z., Spaans, M., & Quataert, E. 2000, ApJ, 537, L5
Hamann, F. 1999, ARA&A, 37, 487
Heckman, T. M., Lehner, M. D., Miley, G. K., & van Breugel, W. 1999, ApJ, 370, 78
Heckman, T. M., Lehner, M. D., Miley, G. K., & van Breugel, W. 1991a, ApJ, 381, 373
Hu, E. M., & Cowie, L. L. 1987, ApJ, 317, L7
Hu, E. M., Cowie, L. L., & McMahon, R. G. 1998, ApJ, 502, L99
Hu, E. M., McMahon, R. G., & Egami, E. 1996, ApJ, 459, L53
Hu, E. M., Songaila, A., Cowie, L. L., & Stockton, A. 1991, ApJ, 368, 28
Katz, N. 1991, ApJ, 368, 325
Kauffmann, G., & Haehnelt, M. 2000, MNRAS, 311, 576
Lacey, C., & Cole, S. 1993, MNRAS, 262, 627
Madau, P., & Pozzetti, L. 2000, MNRAS, 312, 9
Magnor, J., et al. 1998, AJ, 115, 2285
McKee, N., Sasaki, S., & Suto, Y. 1998, ApJ, 495, 555
Menou, K., Haiman, Z., & Narayan, V. 2000, ApJ, submitted
Mo, H. J., Mao, S., & White, S. D. M. 1998, MNRAS, 295, 319
Mo, H. J., & Miralda-Escudé, J. 1996, ApJ, 469, 589
Moore, B., Ghigna, S., Governato, F., Lake, G., Quinn, T., Stadel, J., & Tozzi, P. 1999, ApJ, 524, L19
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493 (NFW)
Navarro, J. F., & Steinmetz, M. 1997, ApJ, 478, 13
Navarro, J. F., & White, S. D. M. 1994, MNRAS, 267, 401
Neufeld, D. A. 1991, ApJ, 370, L85
Osterbrock, D. E. 1989, Astrophysics of Gaseous Nebulae and Active Galactic Nuclei (Sausalito: University Science Books)
Pritchet, C. J. 1994, PASP, 106, 1052
Rees, M. J. 1988, MNRAS, 231, 91
Sellwood, J. A., & Moore, E. M. 1999, ApJ, 511, 660
Steidel, C. C., Adelberger, K. L., Giavalisco, M., Dickinson, M., & Pettini, M. 1999, ApJ, 519, 1
Steidel, C. C., Adelberger, K. L., Shapley, A. E., Pettini, M., Giavalisco, M., & Dickinson, M. 2000, ApJ, 532, 170
van den Bosch, F. C., & Dalcanton, J. J. 2000, ApJ, 534, 146
White, S. D. M., & Rees, M. J. 1978, MNRAS, 183, 341