THE ASTROPHYSICAL JOURNAL, 599:759–772, 2003 December 20


# COMBINING *WILKINSON MICROWAVE ANISOTROPY PROBE* AND SLOAN DIGITAL SKY SURVEY QUASAR DATA ON REIONIZATION CONSTRAINS COSMOLOGICAL PARAMETERS AND STAR FORMATION EFFICIENCY

WEIHSUEH A. CHIU
US Environmental Protection Agency, 808 17th Street NW, Washington, DC 20460; chiu@astro.princeton.edu

XIAOHUI FAN
Steward Observatory, University of Arizona, Tucson, AZ 85721; fan@as.arizona.edu

AND

JEREMIAH P. OSTRIKER
Princeton University Observatory, Princeton, NJ 08544;
and Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK;
jpo@astro.princeton.edu
*Received 2003 April 12; accepted 2003 August 26*

## ABSTRACT

We present constraints on cosmological and star formation parameters based on combining observations from the *Wilkinson Microwave Anisotropy Probe* (*WMAP*) and high-redshift quasars from the Sloan Digital Sky Survey (SDSS). We use a semianalytic model for reionization that takes into account a number of important physical processes both within collapsing halos (e.g., $H_2$ cooling) and in the intergalactic medium (e.g., $H_2$ cooling, Compton cooling, and photoionization heating). We find that the Gunn-Peterson absorption data provide tight constraints on the power spectrum at small scales in a manner analogous to that derived from the cluster mass function. Assuming that the efficiency of producing UV photons per baryon is constant, the constraint takes on the form $\sigma_8\Omega_0^{0.5} \approx 0.33$ in a flat, Λ-dominated universe with $h = 0.72$, $n = 0.99$, and $\Omega_b h^2 = 0.024$. However, the calculated optical depth to electron scattering of $\tau_{\rm es} \sim 0.06$ is well below the value found by *WMAP* of $0.17 \pm (0.04 \sim 0.07)$. Since the *WMAP* constraints on $\tau_{\rm es}$ are somewhat degenerate with the value of the spectral index $n$, we then permit the primordial spectral index $n$ to float and consider the 1 $\sigma$ *WMAP*-only determination of $\Omega_0 h^2 = 0.14 \pm 0.02$ (implying $\Omega_0 = 0.27 \pm 0.04$), while normalizing the power spectrum using *WMAP*. In addition, we allow the UV efficiency to be greater in the past. Combining the *WMAP* constraints with the quasar transmission data, our analysis then favors a model with $\tau_{\rm es} = 0.11^{+0.02}_{-0.03}(\Omega_0/0.27)^{-1}$, $n = 0.96^{+0.02}_{-0.03}(\Omega_0/0.27)^{-0.57}$, implying a *WMAP* normalization of $\sigma_8 = 0.83^{+0.03}_{-0.05}(\Omega_0/0.27)^{0.53}$ (all at 95% confidence) and an effective UV efficiency that was at least ∼10 times greater at $z \gg 6$. The implied UV efficiency is not unreasonable for stars, spanning the range from $10^{-5.5}$ to $10^{-4}$. These results indicate that the quasar and *WMAP* observations are consistent. If future observations confirm an optical depth to electron scattering $\tau_{\rm es} \sim 0.1$, then it would appear that no more "exotic" sources of UV photons, such as miniquasars or active galactic nuclei, are necessary. However, unless one considers more radical sources of UV photons or alternative forms for the power spectrum of density fluctuations, one cannot achieve a value of $\tau_{\rm es} \gtrsim 0.17$ without violating some combination of constraints from quasar transmission data from $z = 4$ to 6 and *WMAP* measurements at large scales.

*Subject headings:* cosmic microwave background — cosmology: theory — galaxies: formation — intergalactic medium — quasars: general

## 1. INTRODUCTION

The primeval spectrum of cosmic density fluctuations on large scales is determinable with great precision from cosmic microwave background (CMB) radiation measurements (Page et al. 2003). In order to determine the total relevant spectrum, information on small scales is also necessary; these become nonlinear at early times, so information concerning reionization in the epochs $20 \lesssim z \lesssim 6$ provides vital clues.

Recent observations of high-redshift quasars have provided the first observational signatures of the epoch of reionization. Spectra of quasars at redshift $z \lesssim 6$ indicate that the universe was almost fully ionized up to $z \sim 6$, since even a small neutral fraction in the intergalactic medium (IGM) would have led to complete absorption of a quasar's continuum radiation. However, the first absorption spectra of quasars at higher redshift indicate that the abundance of neutral hydrogen increases significantly for $z \gtrsim 6$.

Reionization has generally been assumed to be caused by ionizing photons created in early generations of stars and/or quasars. Given this premise, hydrodynamic simulations as well as semianalytic calculations seem to indicate that the process of reionization should occur in several distinct stages. First, cosmological gas falls into deep enough potential wells (caused by dark matter halos) that they can cool and collapse to high enough densities to produce stars and/or quasars. The UV photons produced in this process ionizes the local surroundings, first within the halo itself and then outside the halo, creating cosmological H II regions. This is generally referred to as the "pre-overlap" stage. As the abundance of these H II regions increases (because of additional "galaxy" formation), they eventually start to "overlap" so that gas in the IGM becomes exposed to multiple sources of ionizing radiation. After this "overlap" stage, the IGM becomes optically thin except for inside self-shielded, high-density clouds (those without ionizing sources).

In this picture, the reionization history of the universe depends on both the growth of density perturbations and the efficiency of star/quasar formation. The former is a complex function of the standard cosmological parameters: the density of the universe $\Omega_0$, the baryon abundance $\Omega_b$, the expansion rate $H_0$, and the mass power spectral index $n$ and normalization $\sigma_8$. The latter can be calculated with some precision based on atomic physics (e.g., cooling rates) but ultimately depends on some unknown parameters relating to the efficiency of turning mass into UV photons that can escape into the IGM.

To this overall picture has been added the recent observations by the *Wilkinson Microwave Anisotropy Probe* (*WMAP*) satellite (Bennett et al. 2003). In particular, the high values of the electron optical depth to last scattering $\tau_{\rm es} = 0.17 \pm 0.04$ (Kogut et al. 2003) and $\tau_{\rm es} = 0.17 \pm 0.07$ (Spergel et al. 2003) seem to indicate a much earlier epoch or reionization of $z_{\rm rei} = 17 \pm 5$ (Spergel et al. 2003). How can these measurements be reconciled with the quasar Gunn-Peterson observations? Do they indicate a source of ionizing photons that cannot be accounted for through standard star formation?

In this paper, we first examine how quasar transmission measurements constrain cosmological parameters. In particular, we use the fact that the first generation of UV-generating objects are in fact the tail of the distribution—the rare events that have collapsed to high enough density to produce stars and/or quasars. They thus provide a unique probe of the small-scale power spectrum (at $\lesssim 1$ Mpc scales) in a way analogous to the way the X-ray cluster mass function probes $\sim$10 Mpc scales. Our constraints are derived from quasars' absorption measurements, using theoretical predictions based on the detailed semianalytic model developed by Chiu & Ostriker (2000), with some improvements derived from recent hydrodynamic and semianalytical work on reionization.

Second, we address the question of the consistency of the *WMAP* and quasar measurements, paying particular attention to the degeneracy in the *WMAP* data between the optical depth $\tau_{\rm es}$ and the spectral index $n$. Combining the *WMAP* constraints with the quasar transmission data, our analysis favors a model with somewhat lower values of $\tau_{\rm es}$ and $n$ than implied by the *WMAP* data alone but that are still consistent at the 1 $\sigma$ level.

The organization of this paper is as follows. In § 2 we describe the modeling of reionization and Gunn-Peterson absorption. In § 3 we compare observations from both quasars and *WMAP* to model predictions in order to constrain cosmological parameters. In § 4 we discuss our results and discuss their implications. We summarize and conclude in § 5.

## 2. MODELING REIONIZATION AND GUNN-PETERSON ABSORPTION

### 2.1. *Summary of the Model*

The details of the semianalytic model are described in Chiu & Ostriker (2000); the basic principles are summarized here. It is based on a two-phase model of the universe in which a statistical filling factor for ionized gas is self-consistently calculated. It is assumed that the cold, neutral phase has no sources and evolves passively with the expansion of the universe. The hot, ionized phase contains the ionizing sources and evolves in line with local particle and energy conservation averaged over the phase. The temperature of each phase is calculated using standard physics, with photoheating as the source of heat in the ionized phase and cooling via multiple mechanisms, including $H_2$, atomic lines, and Compton scattering. The time evolution is determined by global particle and energy conservation. In each case, we consider only the regions outside of collapsed gas halos, which of course will rise to the virial temperature. Such collapsed halos are calculated separately and considered potential sources of ionizing radiation. We relate the global ionizing energy density to the filling factor by self-consistently calculating the effective ionized volume surrounding each ionizing source, thus ensuring through energy conservation that all ionizing photons are accounted for.

The abundance and properties of these potential ionizing sources are calculated on the basis of the Press-Schechter formalism (Press & Schechter 1974), constrained by the Jeans criterion (which utilizes the calculated gas temperatures) and by a cooling criterion (the cooling time must be less than the dynamical time). Cooling in the halos includes the important contributions from $H_2$ cooling. Halos that satisfy the Jeans criterion and that can cool efficiently are sources of ionizing radiation. We calculate the luminosity of each ionizing source using the Schmidt law

$$L(M_b) = (1 - f_*) M_b c^2 \epsilon_{\rm esc} \epsilon_* \epsilon_{\rm UV} t_{\rm dyn}^{-1} , \qquad (1)$$

where $f_*$ is the fraction of baryons already turned into stars (a small correction), $M_b$ is the baryonic mass of a halo, $\epsilon_{\rm esc}$ is the escape fraction from the halo, $\epsilon_*$ is a resolution factor (related to the fraction of gas that can form stars) determined through calibration (see Chiu & Ostriker 2000, eq. [28], and below), $\epsilon_{\rm UV}$ is the mass-to-UV efficiency (where we have absorbed the notation in Chiu & Ostriker 2000 of $\epsilon_{\rm hm}\epsilon_{\rm UV}$ into a single efficiency), and $t_{\rm dyn}$ is the dynamical time of the halo. Note that this halo luminosity is somewhat simpler than that used in Chiu & Ostriker (2000) but is consistent with that used in hydrodynamic simulations.

Calibration was done by comparing the redshift of overlap, the ionizing intensity, and the fraction of baryonic mass in stars between the semianalytic results and the hydrodynamic simulation of Gnedin (2000a, 2000b). This simulation was chosen because it is the only published simulation that simulates a statistically "average" universe (as opposed to an individual halo collapse) with sufficient resolution to follow the overlap process and continues at least to redshift 4. The cosmological parameters were $\Omega_0 = 0.3$, $\Omega_\Lambda = 0.7$, $h = 0.7$, $\Omega_b = 0.04$, $n = 1$, and $\sigma_8 = 0.9$. The data used for calibration from Gnedin (2000a, 2000b) are the redshift of overlap $z_* \approx 7$, the ionizing intensity $J_{21} \approx 0.3$ at $z = 4$, and the fraction of baryonic mass in stars $f_* \approx 0.04$ at $z = 4$. The values of $\epsilon_*$ and $\epsilon_{\rm UV}$ in the semianalytic model were adjusted to best match these three values. The results were $\epsilon_* = 0.03$ and $\epsilon_{\rm UV} = 1.2 \times 10^{-5}$, which gave $z_* = 7$, $J_{21} = 0.6$, and $f_* = 0.05$. Note that Gnedin (2000a, 2000b) used $\epsilon_* = 0.05$ and $\epsilon_{\rm UV} = 4 \times 10^{-5}$, which are remarkably similar to the values in the semianalytic model given the differences in approach. In fact, because we take each halo as a whole, it should not be surprising that $\epsilon_*$ is smaller for the semianalytic model than for the hydrodynamic model.

The determination of the escape fraction $\epsilon_{\rm esc}$ bears some additional discussion. We use a simple Strömgren sphere approximation (see Appendix) and derive

$$\epsilon_{\rm esc} \approx (1-\eta)^2 , \tag{2}$$

where the quantity $\eta$ is given by

$$\eta = \min\left(1,\ \sqrt{\frac{\Delta_v \bar{\rho}_b \alpha_R}{3\epsilon_* \epsilon_{\rm UV} m_{\rm H}^2 c^2 E_0^{-1} t_{\rm dyn}^{-1}}}\right) , \tag{3}$$

where $\Delta_v \approx 178$ is the virial overdensity, $\bar{\rho}_b$ is the mean baryonic density, $\alpha_R$ is the recombination rate, $m_{\rm H}$ is the hydrogen mass, and $E_0$ is 13.6 eV. In this approximation, we assume that the halo baryons have a $r^{-2}$ density run, the UV output per baryon is constant in the halo, the halo is in ionization equilibrium, and recombinations determine the amount of photons absorbed locally.

The ionizing source luminosity defined above, then, is the luminosity seen by the IGM, i.e., outside of the halos. A further assumption is made in the pre-overlap stage: that at any particular time, the ionized volume around an isolated source is linearly related to its UV luminosity. The constant of proportionality is determined through the global ionization and energy balance, but the size of the H II regions as a function of luminosity is a linear relation. This assumption is consistent with previous work on cosmological H II regions (e.g., Shapiro & Giroux 1987). The filling factor is calculated by assuming the ionized regions are randomly distributed. When ionizing sources are rare, the "effective volume" surrounding each ionized source is the same as that for cosmological Strömgren spheres. However, when overlap begins to occur, this effective volume is adjusted self-consistently to ensure that no ionizing photons are "lost." The self-consistency is ensured through the ionizing energy conservation equation (Chiu & Ostriker 2000, eq. [4]).

The outputs of the model that we use below are the filling factor $Q$, the temperature $T_4$, and the ionizing intensity $J_{21}$.

### 2.2. *Density Distribution of Cosmic Gas*

One of the most important determinants of how reionization evolves is the degree of gas clumping. The clumping factor, defined as

$$\mathscr{C} \equiv \frac{\langle \rho_b^2 \rangle}{\langle \rho_b \rangle^2} , \tag{4}$$

is important not only for determining the ionization balance (and hence the filling factors), but also for determining the *thermal* balance. Note that this clumping factor is used only in the ionized region.

In order to determine the clumping factor, the probability density function (pdf) for the gas density must be known. Miralda-Escudé, Haehnelt, & Rees (2000) found that a good fit to the volume-weighted pdf as seen in hydrodynamic simulations is

$$P_V(\Delta)\, d\Delta = A \exp\left[-\frac{(\Delta^{-2/3} - C_0)^2}{2(\delta_0/3)^2}\right] \Delta^{-\beta}\, d\Delta , \tag{5}$$

where the overdensity $\Delta = \rho_b/\bar{\rho}_b$, $\delta_0 = 7.61/(1+z)$ is related to the linear rms gas density fluctuation and $\beta$ is related to the density run at high densities. The parameters $A$ and $C_0$ are found by requiring the mass and volume to be normalized to unity.

This formula, with the parameterizations of $\beta$ and $\delta_0$ found in Miralda-Escudé et al. (2000), has been used by many others in calculating characteristics of reionization (e.g., Songaila & Cowie 2002; Fan et al. 2002). However, it is not often noted that the parameterization is based on a single simulation with a particular set of cosmological parameters. In particular, Miralda-Escudé et al. (2000) used the simulation reported in Miralda-Escudé et al. (1996), which was a $\Lambda$ cold dark matter model with $\Omega_0 = 0.4$, $\Omega_\Lambda = 0.6$, $\Omega_b h^2 = 0.015$, $h = 0.65$, $n = 1$, and $\sigma_8 = 0.79$. Since in this paper we are considering variants in the background cosmology, it is certainly not sufficient to use the Miralda-Escudé et al. (2000) parameterization for the gas pdf. In particular, the amount of power at small scales will depend significantly on all the cosmological parameters (especially $\sigma_8$). We therefore analyzed a second simulation (which uses different cosmological parameters) and generated a second set of parameter fits. The second simulation is one by R. Cen (2002, private communication) in which $\Omega_0 = 0.3$, $\Omega_\Lambda = 0.7$, $\Omega_b h^2 = 0.017$, $h = 0.67$, $n = 1$, and $\sigma_8 = 0.9$. The values of $\beta$ and $\delta_0$ are given for both Miralda-Escudé et al. (2000) and Cen in Table 1.

As is clear from Table 1, the value of $\delta_0$ and to a lesser extent $\beta$ depend on cosmology. This is not surprising given that $\delta_0$ *must* depend on the power spectrum. The cosmological dependence on $\beta$ is less clear, especially since the best-fit value of $\beta$ depends significantly on the density run in collapsed halos. Because the Cen simulation is at much higher resolution and includes more realistic physics, we simply fit $\beta$ to the Cen results as a function of redshift, with a maximum of $\beta_{\rm max} = 2.5$, corresponding to an isothermal sphere:

$$\beta \approx \min\left(2.5,\ 3.2 - \frac{4.73}{1+z}\right) , \quad z \geq 4 , \tag{6}$$

where we are considering only redshifts $z \geq 4$.

As for a prescription for finding $\delta_0$, we note that the gas pdf also predicts the fraction of mass in collapsed virialized halos:

$$f_{\rm gas\,pdf}({\rm collapsed}) \approx \int_{6\pi^2}^{\infty} \Delta P_V(\Delta)\, d\Delta . \tag{7}$$

The integration point $6\pi^2$ is derived from the fact that for a singular isothermal sphere, the local overdensity density at

TABLE 1

Fit Parameters for Gas Probability Density Function Derived from Simulations

| | MHR[a] | | Cen[b] | |
|---|---|---|---|---|
| Redshift | $\beta$ | $\delta_0$ | $\beta$ | $\delta_0$ |
| 6 | 2.50 | 1.09 | 2.52 | 2.6 |
| 5 | ... | ... | 2.41 | 3.0 |
| 4 | 2.48 | 1.53 | 2.25 | 3.6 |
| 3 | 2.35 | 1.89 | ... | ... |
| 2 | 2.23 | 2.54 | ... | ... |

[a] Miralda-Escudé et al. 2000.
[b] R. Cen 2002, private communication.

the virial radius is $6\pi^2$. Therefore, we are approximately taking into account all gas within virialized halos. Similarly, from linear theory, we can use the Press-Schechter formalism (Press & Schechter 1974) to find the same quantity if we know the correct filtering radius $R_f$:

$$f_{\rm PS}(\text{collapsed}) \approx \sqrt{\frac{2}{\pi\sigma_{R_f}^2}} \int_{\delta_c}^{\infty} \exp\left(-\frac{\delta^2}{2\sigma_{R_f}}\right) d\delta \,, \quad (8)$$

where $\delta_c \approx 1.69$ and $\sigma_{R_f}$ is the linear rms mass fluctuation filtered with a top hat of radius $R_f$. However, the mass fraction should be equal by the two calculations:

$$f_{\rm gas\,PDF}(\text{collapsed}) = f_{\rm PS}(\text{collapsed}) \,. \quad (9)$$

By analyzing both simulations, we find the following relation leads to satisfactory results:

$$R_f \approx \frac{R_{\rm J}}{4} \,, \quad (10)$$

where $R_{\rm J}$ is the Jeans length defined by

$$R_{\rm J} = \sqrt{\frac{5\pi k_{\rm B} T}{12 G\bar{\rho}\mu m_{\rm H} a^2}} \,, \quad (11)$$

$T$ is the gas temperature, $\bar{\rho}$ is the mean total density, and $\mu$ is the baryons per particle. With this prescription for finding $\delta_0$ and $\beta$, we need only use the normalization constraints to find $A$ and $C_0$ for an arbitrary background cosmology. While by no means perfect, this procedure should account to first order for the dependence of the gas pdf on background cosmology and is certainly more accurate than treating the evolution of $\delta_0$ as independent of cosmology. We note here that the "filtering" scale is smaller than the "current" Jeans scale, consistent with the fact that the IGM retains a "thermal memory" of the Jeans scale in the past.

Given the gas pdf, one can determine the clumping factor by simple integration over all densities. However, in reality there is a high-density cutoff due to the the very high density regions not participating in the ionization balance because of self-shielding. However, the determination of the cutoff in our prescription is not difficult, since we have explicitly made a separation between the "in halo" and "out of halo" calculations. Thus, we calculate the clumping factor by integrating

$$\mathscr{C} = \frac{\int_0^{6\pi^2} \Delta^2 P_V(\Delta)\, d\Delta}{\int_0^{6\pi^2} P_V(\Delta)\, d\Delta} \,. \quad (12)$$

As above, we assume that gas with overdensity greater than $6\pi^2$ is within collapsed halos.

We should note that this treatment produces smaller clumping factors ($\sim$5) than typically used in semianalytic treatments. This is because we are essentially using a "hybrid" between the reionization treatment of Miralda-Escudé et al. (2000) and the usual clumping factor methods. Conceptually, Miralda-Escudé et al. (2000) considered a density-dependent ionization fraction: they assumed that the universe was ionized up to a critical density, above which it was neutral and self-shielded. Here we are assuming that the collapsed halos are self-shielded and then treating the rest of the universe using the clumping factor approach. As described in the Appendix, at redshifts before full reionization, halos are optically thick if they do not contain sources themselves. This treatment does not include screening and evaporation of minihalos (e.g., Haiman, Abel, & Madau 2001; Barkana & Loeb 2002), but we discuss the potential effects of minihalos in § 4 below.

To summarize, the semianalytic model actually has *three* phases: the "interior of halos," the hot ionized H II regions, and the cold neutral regions. The "interior of halos" phase includes all the gas in the universe that is in halos with baryonic mass greater than the Jeans mass at their time of formation. Because we account for halo survival after their time of formation, the masses of ionizing sources at any particular time retain a thermal memory of the Jeans mass in the past (i.e., the effective "filtering mass," as described by Gnedin & Hui 1998). We assume that these regions have virialized with isothermal sphere density runs and therefore include all gas with overdensity greater than $6\pi^2$. In this phase, the cooling criterion and the Schmidt law are used to calculate the ionizing source function, and the photoionizations and recombinations are treated through the escape fraction calculation (a Strömgren sphere approximation). Those photons that "escape" this phase then enter the "second phase," which is the ionized H II regions. In this phase, a clumping factor approach is used, with a high-density cutoff at $6\pi^2$. The last phase is the cold neutral part of the universe, which uses up ionizing photons only when "converted" into the ionized phase.

### 2.3. *Modeling Gunn-Peterson Absorption*

The statistics of Gunn-Peterson absorption have been reviewed by numerous others (Miralda-Escudé et al. 2000; McDonald & Miralda-Escudé 2001; Fan et al. 2002; Songaila & Cowie 2002). To summarize, we begin with the standard optical depth at resonance for a gas with neutral density $n_{\rm H\,I}$ at redshift $z$:

$$\tau = \frac{3\Lambda_{2p\to 1s}\lambda_\alpha^3 n_{\rm H\,I}(z)}{8\pi H_0\sqrt{\Omega_0(1+z)^3 + \Omega_\Lambda}} \,, \quad (13)$$

where $\Lambda_{2p\to 1s}$ is the decay rate ($6.25 \times 10^8$ s$^{-1}$), $\lambda_\alpha$ is the Ly$\alpha$ wavelength ($1.216 \times 10^{-5}$ cm), and the other symbols have their usual meaning in a cosmological context.

Now we make the standard assumptions of ionization equilibrium and a uniform ionizing intensity in the ionized regions to obtain the neutral fraction. Considering only photoionization and radiative recombination,

$$n_{\rm H\,I}\Gamma_{21}J_{21} = \frac{X(X+Y/4)}{m_p^2 R_4 T_4^{-0.7}\bar{\rho}_b^2\Delta^2} \,, \quad (14)$$

where $Y = 0.24$ is the helium mass fraction (we assume helium is singly ionized), $R_4 \approx 4.2 \times 10^{-13}$ cm$^3$ s$^{-1}$ is the recombination rate at a temperature $T = 10^4$ K, $T_4 = T/10^4$ K is the temperature at mean density in units of $10^4$ K, and $\Gamma_{21}$ is the photoionization rate for $J_{21} = 1$. We use the method of Hui & Gnedin (1997) to calculate the temperature at mean density after reionization (see also Hui & Haiman 2003). Plugging in values for the various other constants [$\Gamma_{21} = 4.35 \times 10^{-12}$ s$^{-1}$, $m_p = 1.67 \times 10^{-24}$ g, $\bar{\rho}_b = 1.88 \times 10^{-29}(1+z)^3\Omega_b h^2$ g cm$^{-3}$] gives

$$n_{\rm H\,I} = 7.63 \times 10^{-12}\frac{(\Omega_b h^2)^2 T_4^{-0.7}\Delta^2(1+z)^6}{J_{21}} \text{ cm}^{-3} \quad (15)$$

or

$$\tau = \frac{0.316\Omega_b^2 h^3 (1+z)^6}{\sqrt{\Omega_0 (1+z)^3 + \Omega_\Lambda}} \frac{T_4^{-0.7}}{J_{21}} \Delta^2 , \quad (16)$$

where $T_4$, $\Delta$, and $J_{21}$ depend on the redshift $z$ but only $\Delta$ also depends on spatial location. It is helpful to reformulate this for $\Omega_0(1+z)^3 \gg \Omega_\Lambda$ and scaled from $\Omega_b h^2 = 0.02$ and $\Omega_0 h^2 = 0.14$:

$$\tau = 1.536 \left( \frac{1+z}{1+5.5} \right)^{4.5} \sqrt{\frac{0.14}{\Omega_0 h^2}} \left( \frac{\Omega_b h^2}{0.02} \right)^2 \frac{T_4^{-0.7}}{J_{21}} \Delta^2 . \quad (17)$$

Defining the optical depth for a uniform medium ($\Delta = 1$) $\tau_u$ as

$$\tau_u \equiv 1.536 \left( \frac{1+z}{1+5.5} \right)^{4.5} \sqrt{\frac{0.14}{\Omega_0 h^2}} \left( \frac{\Omega_b h^2}{0.02} \right)^2 \frac{T_4^{-0.7}}{J_{21}} , \quad (18)$$

we can define the mean transmitted flux ratio at a given redshift

$$\mathscr{T}_z = \langle \exp(-\tau_u \Delta^2) \rangle = Q \int_0^\infty P_V(\Delta) \exp(-\tau_u \Delta^2) d\Delta , \quad (19)$$

where $Q$ is the volume filling factor for the ionized regions (we have now adopted the standard notation; Chiu & Ostriker 2000 used "$f$" as the volume filling factor). Note that strictly speaking, there should be a high-density cutoff. However, since the exponential has a power of $\Delta^2$, for large values of $\Delta$ the optical depth is very large. The high-density end of the pdf contributes very little to the integral, and the results are thus not sensitive to this cutoff. In fact, at the redshifts of interest, the integral is usually dominated by values of $\Delta < 1$, as was noted by Barkana (2002) and others.

Our semianalytic model provides the values of $Q$, $T_4$, and $J_{21}$, and from the previous section, we have a prescription for determining $P_V(\Delta)$. Thus, we have a complete model from which to calculate the expected Gunn-Peterson absorption.

## 3. COMPARING MODEL PREDICTIONS AND OBSERVATIONS

Having calibrated our model using various hydrodynamic simulations, we are left with one free parameter, $\epsilon_{\rm UV}$, in addition to the cosmological parameters. Below we first consider the case where $\epsilon_{\rm UV}$ is constant and derive a constraint on $\sigma_8 \Omega_0^{0.5}$, similar to those from rich clusters, and compare the results and predictions to those found by *WMAP*. We then consider joint constraints from *WMAP* and the Sloan Digital Sky Survey (SDSS) quasars on the spectral index $n$ as well as the time dependence of $\epsilon_{\rm UV}$.

### 3.1. *Combining SDSS Quasar Data*

The quasar data we use combine the compilation at $z_{\rm abs} \lesssim 5.5$ from Songaila & Cowie (2002) (see our Table 2) and measurements of six SDSS quasars at $z > 5.7$ from Becker et al. (2001), Fan et al. (2003), and White et al. (2003) (our Table 3). Note that we did not include the J1044−0125 ($z = 5.74$) data in Table 3 because it is already incorporated into the compilation of Songaila & Cowie (2002).

TABLE 2

COMPILATION OF TRANSMISSION DATA FROM SONGAILA & COWIE (2002)

| Redshift | Mean Transmission $\mathscr{T}$ | $\sigma_{\rm scatter}$ | $\sigma_{\rm mean} = \sigma_{\rm scatter}/\sqrt{N}$ | $N$ |
|---|---|---|---|---|
| 4.09 | 0.352 | 0.352 | 0.027 | 15 |
| 4.34 | 0.334 | 0.334 | 0.020 | 20 |
| 4.61 | 0.260 | 0.260 | 0.017 | 15 |
| 4.93 | 0.162 | 0.162 | 0.022 | 5 |
| 5.20 | 0.107 | 0.107 | 0.022 | 8 |
| 5.51 | 0.074 | 0.074 | 0.011 | 7 |

In order to combine these data, we must take into account the fact that at redshifts $z \lesssim 5.6$, the uncertainties in the transmission data are dominated by intrinsic scatter, while at higher redshifts, the uncertainties are dominated by measurement errors. Since our model predicts the *mean* transmission, our likelihood function must properly account for both measurement error as well as the scatter. For a compilation of transmissions, such as that by Songaila & Cowie (2002), the contribution of the data $D_c = \{(z_j, \mathscr{T}_j)\}$ to the likelihood is simply

$$P(\mathscr{T}_z | D_c) = \exp\left[ -\frac{1}{2} \sum_j \frac{(\mathscr{T}_{z_j} - \mathscr{T}_j)^2}{\sigma_{j,\rm mean}^2} \right] , \quad (20)$$

$$\sigma_{j,\rm mean}^2 = \frac{\sigma_{j,\rm scatter}^2}{\sqrt{N}} . \quad (21)$$

The contribution to the likelihood from the compiled data is estimated this way.

For the individual transmission data, especially in the case where both measurement error and scatter are important, the simplest way to do this is to estimate scatter $\sigma_{\rm scatter}$ and add in quadrature with the measurement error $\sigma_{\rm meas}$ to obtain the total uncertainty in the mean for each data point. In particular, for individual transmission data $D_i = \{(z_j, T_j)\}$with known (Gaussian) errors $\sigma_{\rm meas}$ and $\sigma_{\rm scatter}$, the probability of a mean transmission as a function of redshift $\mathscr{T}_z$ is

$$P(\mathscr{T}_z | D_i) = \exp\left[ -\frac{1}{2} \sum_j \frac{(\mathscr{T}_{z_j} - T_j)^2}{\sigma_{j,\rm meas}^2 + \sigma_{j,\rm scatter}^2} \right] . \quad (22)$$

Note that if the measurement error is negligible, then equation (22) reduces to equation (21). To estimate $\sigma_{\rm scatter}$, we use the Songaila & Cowie (2002) compilation for $z \lesssim 5.6$ (interpolated) and use the data themselves (binned) to determine $\sigma_{\rm scatter}$ at higher redshifts. These results are also tabulated in Table 3. The total likelihood function $\mathscr{L}$ is then given by the product of the two separate likelihoods.

### 3.2. *SDSS Quasar Constraints with a Constant UV Efficiency Compared with WMAP*

For this initial comparison, we fix the other cosmological parameters to their *WMAP*-only best-fit values: $n = 0.99$, $\Omega_b h^2 = 0.024$, and $h = 0.72$. We use the *WMAP*-only results from Spergel et al. (2003), Table 1, to ensure clarity in the

TABLE 3

TRANSMISSION DATA

| Redshift Range | Transmission $T \pm \sigma_{\rm meas}$ | $\sigma_{\rm scatter}$ | Reference | SDSS Quasar | Comments |
|---|---|---|---|---|---|
| 4.84–5.00 | $0.1334 \pm 0.0011$ | 0.046 | 1 | J0836+0054 | |
| 5.00–5.17 | $0.0809 \pm 0.0011$ | 0.055 | 1 | J0836+0054 | |
| 5.17–5.33 | $0.0523 \pm 0.0008$ | 0.055 | 1 | J0836+0054 | |
| 5.33–5.50 | $0.0692 \pm 0.0010$ | 0.029 | 1 | J0836+0054 | |
| 5.50–5.66 | $0.1185 \pm 0.0011$ | 0.029 | 1 | J0836+0054 | |
| 5.25–5.41 | $0.1324 \pm 0.0036$ | 0.029 | 1 | J1030+0524 | |
| 5.41–5.58 | $0.0996 \pm 0.0033$ | 0.029 | 1 | J1030+0524 | |
| 5.58–5.74 | $0.0418 \pm 0.0033$ | 0.020 | 1 | J1030+0524 | |
| 5.74–5.95 | $0.0242 \pm 0.0038$ | 0.020 | 1 | J1030+0524 | |
| 6.0–6.17 | $0.0010 \pm 0.0009$ | 0.0005 | 2 | J1030+0524 | |
| 6.0–6.17 | $0.0043 \pm 0.0088$ | ... | 2 | J1030+0524 | Ly$\beta$ |
| 5.95–6.15 | $0.0031 \pm 0.0149$ | 0.0005 | 3 | J1048+4637 | |
| 5.50–5.70 | $0.0541 \pm 0.0014$ | 0.029 | 1 | J1148+5251 | |
| 5.70–5.90 | $0.0139 \pm 0.0019$ | 0.020 | 1 | J1148+5251 | |
| 6.0–6.10 | $0.0 \pm 0.0063$[a] | ... | 2 | J1148+5251 | |
| 6.0–6.10 | $0.0 \pm 0.335$[a] | ... | 2 | J1148+5251 | Ly$\beta$ |
| 6.10–6.32 | $0.0 \pm 0.0021$[a] | ... | 2 | J1148+5251 | |
| 6.10–6.32 | $0.0 \pm 0.051$[a] | ... | 2 | J1148+5251 | Ly$\beta$ |
| 5.00–5.17 | $0.1429 \pm 0.0027$ | 0.055 | 1 | J1306+0356 | |
| 5.17–5.33 | $0.0922 \pm 0.0025$ | 0.055 | 1 | J1306+0356 | |
| 5.33–5.50 | $0.0936 \pm 0.0025$ | 0.029 | 1 | J1306+0356 | |
| 5.50–5.66 | $0.0679 \pm 0.0027$ | 0.029 | 1 | J1306+0356 | |
| 5.66–5.83 | $0.0699 \pm 0.0033$ | 0.020 | 1 | J1306+0356 | |
| 5.55–5.75 | $0.0383 \pm 0.0258$ | 0.020 | 3 | J1630+4027 | |

NOTE.—Ly$\alpha$ unless otherwise noted.

[a] These are conservative upper limits based on the hypothesis of Ly$\alpha$ emission from an intervening galaxy.

REFERENCES.—(1) Becker et al. 2001; (2) White et al. 2003; (3) Fan et al. 2003.

data underlying our analysis. Note, however, that these last two values have additional independent lines of support, as noted in Spergel et al. (2003).

We use a Bayesian method to determine the cosmological constraints. The priors we use are $\Omega_0 \in [0.15, 0.40]$ and $\sigma_8 \in [0.5, 1.0]$. This region essentially bounds the 99% region reported for the combined CMB and cluster analysis of Melchiorri et al. (2003).

In order to constrain $\Omega_0$ and $\sigma_8$ separately from $\epsilon_{\rm UV}$, we marginalize over the latter parameter to yield the marginalized likelihood distribution:

$$\mathscr{L}(\Omega_0, \sigma_8) = \int \mathscr{L}(\Omega_0, \sigma_8, \epsilon_{\rm UV})\, d\epsilon_{\rm UV} \; . \tag{23}$$

The central values and percentile limits are found by integrating over $\mathscr{L}(\Omega_0, \sigma_8)$.

The result of this analysis is presented in Figure 1, where we plot the likelihood contours in the plane $\Omega_0$-$\sigma_8$. As can be seem by the figure, there is a considerable degeneracy in this plane similar to that derived from rich clusters. Thus, the cosmological constraint from Gunn-Peterson absorption can be summarized by

$$\sigma_8 \Omega_0^{0.5} = 0.33 \pm 0.01 \; , \tag{24}$$

where the error term is statistical error only (68%).

Given the complex physics of the reionization model, it is difficult to determine a meaningful systematic error bar. However, we did investigate the effects of changing $h$ and $n$. The net result was to shift the right-hand side of equation

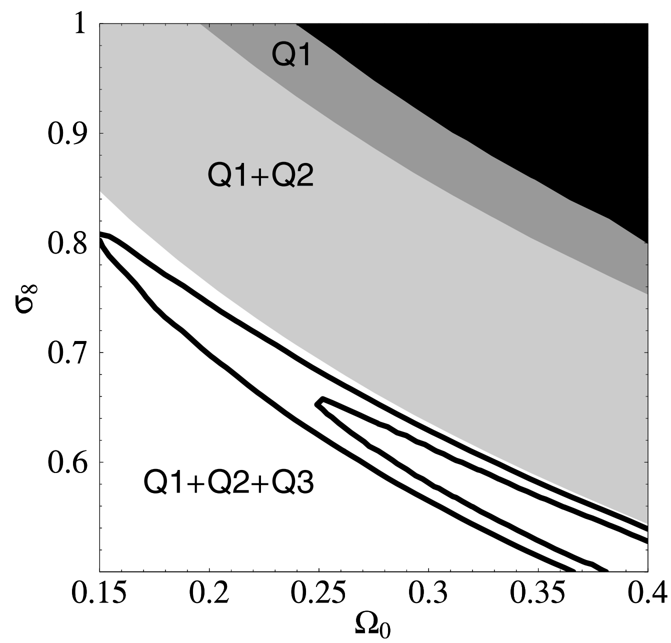


FIG. 1.—Summary of cosmological constraints in the $\Omega_0$-$\sigma_8$ plane for a constant UV efficiency. The constraints Q1, Q2, and Q3, as described in the text, are labeled in the shaded regions (the plus sign in this context denotes "intersection"). The thick solid lines are the formal 68% and 95% contours using all the quasar data.

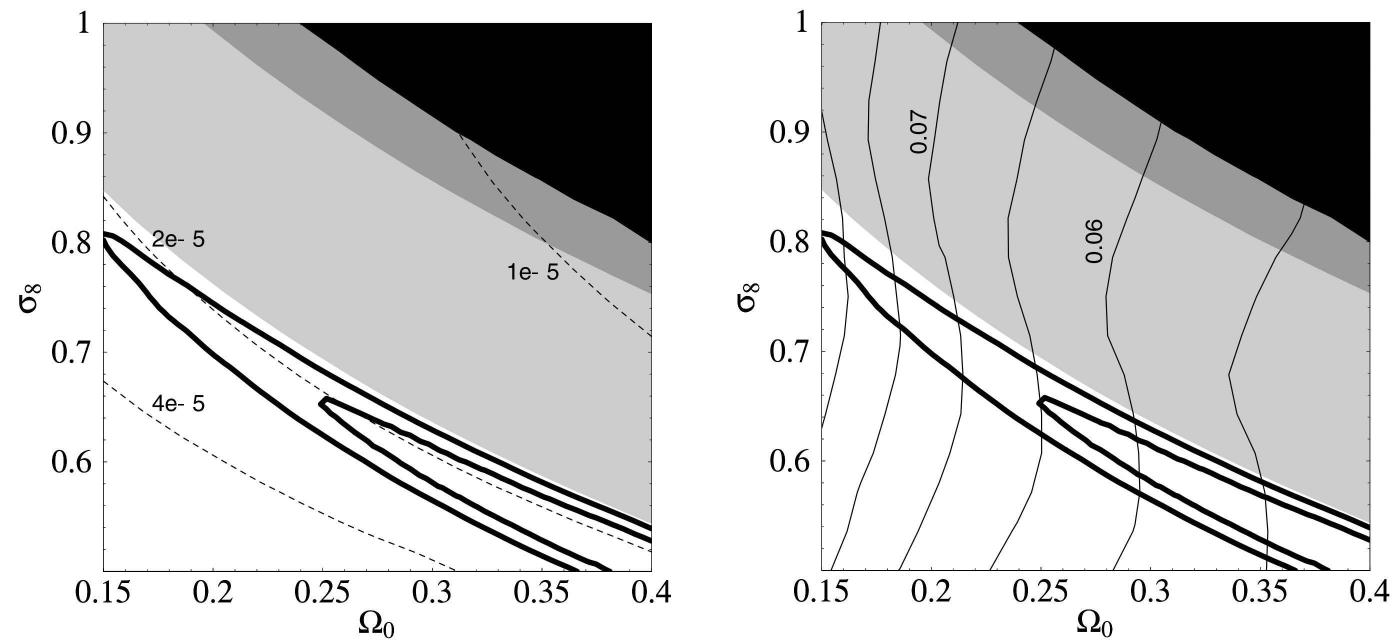


FIG. 2.—Values of the UV efficiency $\epsilon_{\rm UV}$ (*left*) and the optical depth to electron scattering $\tau_{\rm es}$ (*right*), overlaid on the quasar data constraints (Fig. 1), for a constant UV efficiency.

(24) so that

$$\sigma_8\Omega_0^{0.5} \approx (0.33 \pm 0.01)\left(\frac{0.72}{h}\right)^{0.9+(n-1)/2}\left(\frac{n}{0.99}\right)^{-0.85}, \quad (25)$$

so that lower levels of $h$ or $n$ increase the right-hand side. Thus, lowering $h$ or $n$ slightly would make the Gunn-Peterson constrain more consistent.

These results for a constant UV efficiency are somewhat discordant with *WMAP*'s marginalized value of $\sigma_8\Omega_0^{0.5} = 0.48 \pm 0.12$, although only at the 1.25 $\sigma$ level. They are consistent with the cluster determinations of Bahcall et al. (2003) of $\sigma_8\Omega_0^{0.6} = 0.33 \pm 0.03$ (note different exponent), although again at about the 1 $\sigma$ level.

The actual *value* of $\epsilon_{\rm UV}$ bears some discussion. Although we have left this parameter as completely free, there are certainly astrophysical constraints on its value. In our 95% region, we find that $\epsilon_{\rm UV} \sim 2 \times 10^{-5}$ (Fig. 2, *left*). For a Scalo mass function with metal-enriched stars (1/20 solar metallicity), the value calculated from population synthesis is $\sim 5 \times 10^{-5}$ (Wyithe & Loeb 2003a). Given the uncertainties and the relative simplicity of our model, the correspondence is quite remarkable.

The more striking inconsistency is that in all cases, the optical depth to electron scattering,

$$\tau_{\rm es} \equiv \int_0^{1000} dz\, \frac{dt}{dz} c\sigma_{\rm T} n_e\,, \quad (26)$$

where $\sigma_{\rm T} = 6.652 \times 10^{-25}$ cm$^2$ is the Thomson cross section and $n_e$ is the electron density, is too low. Our model finds that $\tau_{\rm es} \sim 0.06$ (Fig. 2, *right*), whereas Spergel et al. (2003) and Kogut et al. (2003) report *WMAP*-only best-fit values of $\tau_{\rm es} = 0.17 \pm 0.07$ and $0.17 \pm 0.04$, respectively. Several articles written in the wake of these results (e.g., Haiman & Holder 2003; Ciardi, Ferrara, & White 2003; Cen 2003b) have discussed this inconsistency and indicated that star formation must have begun much earlier than previously thought. We consider this issue in the following analysis.

### 3.3. *Combined WMAP and SDSS Constraints with a Time-varying UV Efficiency*

There are good astrophysical reasons to believe that $\epsilon_{\rm UV}$ may effectively increase with redshift. For instance, the first generation of stars would have been metal-free, and stellar models predict them to have a significantly higher UV output per baryon (about 4 times higher was reported by Wyithe & Loeb 2003a). In addition, the creation of metals and henceforth dust, which would obscure ionizing sources, may also lead to lower *effective* efficiencies at lower redshift than at higher redshift.

Thus we now consider constraints on $n$ and the time dependence of $\epsilon_{\rm UV}$ from *WMAP* and SDSS quasars jointly. To limit the dimensionality of the parameter space, we keep the baryon abundance and the Hubble constant at their *WMAP* best-fit values and in addition fix $\Omega_0 h^2 = 0.14$ to its *WMAP* best-fit value (below we also consider the sensitivity to this parameter). For each value of $n$ and $\tau_{\rm es}$, *WMAP* predicts a unique best-fit value of the normalization $\sigma_8$ (based on the code provided by Verde et al. 2003 and the accompanying data files from Hinshaw et al. 2003 and Kogut et al. 2003). The normalization is given approximately by (for our fixed values of $h$, $\Omega_0 h^2$, and $\Omega_b h^2$)

$$\sigma_8 \exp(-\tau_{\rm es}) = 0.765 + 0.6(n-1)\,, \quad (27)$$

where formula (27) is good to better than 0.5%. Note that at the *WMAP* best-fit values of $n = 0.99$ and $\tau_{\rm es} = 0.17$, this formula gives the *WMAP* best-fit value of $\sigma_8 = 0.9$.

We assume the following heuristic form for the time dependence of the UV efficiency:

$$\epsilon_{\rm UV} = \epsilon_{\rm UV,0}\left[1 + A\exp\left(\frac{-f_*}{f_{*,\rm crit}}\right)\right], \quad (28)$$

TABLE 4

$z_{\rm trans}$ AS A FUNCTION OF $n$ AND $f_{*,\rm crit}$ FOR SEMIANALYTIC RUNS

| | $z_{\rm trans}$ for $n =$ | | | | |
|---|---|---|---|---|---|
| $f_{*,\rm crit}$ | 0.94 | 0.97 | 1.00 | 1.03 | 1.06 |
| $3\times10^{-5}$ ............. | 16.5 | 20 | 24.5 | 30 | 35 |
| $1\times10^{-4}$ ............. | 14.5 | 18 | 22 | 27 | 32 |
| $3\times10^{-4}$ ............. | 12.5 | 15.5 | 19 | 23.5 | 28 |
| $1\times10^{-3}$ ............. | 10.5 | 13 | 16 | 20 | 24 |

where $f_*$ is the fraction of baryons in stars. Thus, the luminosity of each ionizing source is given by

$$L(M_b) = (1 - f_*)M_b c^2 \epsilon_{\rm esc}\epsilon_*\epsilon_{\rm UV,0} t_{\rm dyn}^{-1}\left[1 + A\exp\left(\frac{-f_*}{f_{*,\rm crit}}\right)\right] . \tag{29}$$

The new parameters $A$ and $f_{*,\rm crit}$ define the time dependence of the efficiency. The essential feature is that at very high redshift, the total efficiency is $(1 + A)$ times greater than at lower redshift, with the transition occurring at $f_* \sim f_{*,\rm crit}$. This form is motivated by the notion that the first stars, being metal-free, had a higher effective efficiency, but that as metals were produced by this generation of stars and reinjected back into the IGM, the efficiency decreased. However, we emphasize that this model was simply selected heuristically and was not based on any detailed model of star formation feedback on the UV efficiency. For instance, it should be noted that the " enhancement " factor may be a combination of changes in the intrinsic efficiency $\epsilon_{\rm UV}$ and the resolution factor $\epsilon_*$ related to the fraction of cooling gas that forms stars. To give a sense of the meaning of $f_{*,\rm crit}$, Table 4 presents the transition redshifts $z_{\rm trans}$, where $f_* \approx (\ln 2) f_{*,\rm crit}$ so that $\epsilon_{\rm UV}(z_{\rm trans}) \approx \epsilon_{\rm UV,0}(1 + A/2)$.

We fix the value of $f_{*,\rm crit} = 10^{-4}$ (we discuss the sensitivity to this parameter below). Thus, for each point in the $n$-$\tau_{\rm es}$ plane, our procedure for searching parameter space is as follows: (1) pick a value of $A$; (2) adjust $\epsilon_{\rm UV,0}$ so that $\tau_{\rm es,model}$ (calculated by model) is consistent with $\tau_{\rm es}$; (3) stop if $\chi^2$ is minimized, otherwise go back to step 1. For each point in the $n$-$\tau_{\rm es}$ plane, we have a best-fit value of $\tilde{\chi}^2(n, \tau_{\rm es})$. We treat $\tilde{\chi}^2$ as approximately $-2\ln\mathscr{L}$, where $\mathscr{L}(n, \tau_{\rm es}|f_{*,\rm crit})$ is the likelihood function, and integrate (with uniform priors) to find the confidence regions (the results are insensitive to the priors).

The effective optical depth for several models is shown in Figure 3. The three illustrative " good-fitting " models are in the 95% region, and the " bad-fitting " model is outside the region. Clearly, high values of $n$ and $\tau_{\rm es}$ are ruled out by the data.

Considering all the constraints together, the results are shown in Figure 4. Our analysis strongly favors a narrow range (95% errors):

$$\tau_{\rm es} = 0.11^{+0.02}_{-0.03} , \tag{30}$$

$$n = 0.96^{+0.02}_{-0.03} \tag{31}$$

for $f_{*,\rm crit} = 10^{-4}$ and $\Omega_0 = 0.27$. This range is consistent with the *WMAP* results, also shown in the figure. The effect of changing $f_{*,\rm crit}$ is only to shift the $\tau_{\rm es}$ constraint up or down. For instance, the constraint for $f_{*,\rm crit} = 3\times10^{-5}$ is $\tau_{\rm es} = 0.12^{+0.02}_{-0.03}$, with no change in the $n$ constraint. The results for increasing $f_{*,\rm crit}$ are in the opposite direction, decreasing the constraint on $\tau_{\rm es}$ by $\sim$0.02 for a factor of 10 increase in $f_{*,\rm crit}$. We also investigated the sensitivity to the matter density $\Omega_0 h^2$ in the 1 $\sigma$ *WMAP*-only range of

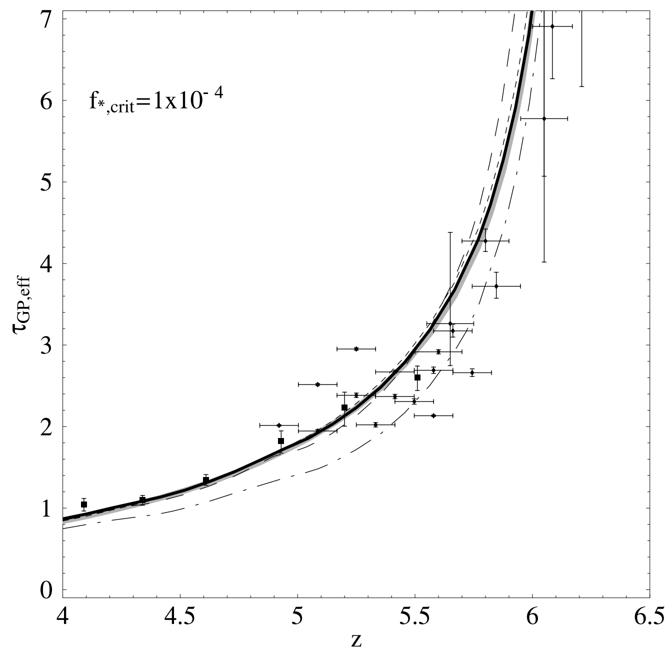


FIG. 3.—Plot of effective Ly$\alpha$ optical depth $\tau_{\rm eff} = -\ln(\mathscr{T})$ from the Songaila & Cowie (2002) compilation (*boxes*; error bars are $\sigma_{\rm mean}$), SDSS quasars (*diamonds*; errors do not include $\sigma_{\rm scatter}$), and model predictions for three model runs within the 95% region of the $n$-$\tau_{\rm es}$ plane for $f_{*,\rm crit} = 1\times10^{-4}$ and one model run outside of the region (i.e., ruled out by the data). The values of $(n, \tau_{\rm es})$, in decreasing likelihood, are as follows: *thick solid line*, (0.96, 0.11); *short-dashed line*, (0.93, 0.08); *long-dashed line*, (0.98, 0.13); *dot-dashed line*, (0.99, 0.14). For comparison, the thick gray line is the best-fit model for a constant efficiency at $\Omega_0 = 0.27$ and $n = 0.99$ ($\sigma_8 = 0.64$).

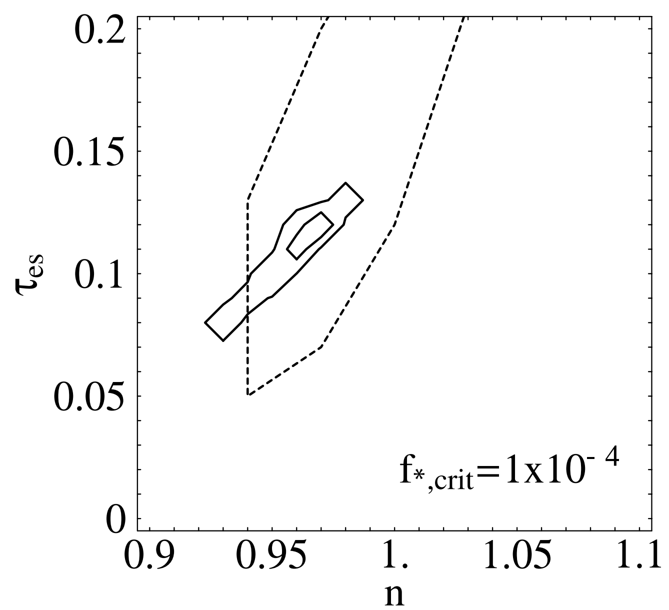


FIG. 4.—Constraints (68% and 95% contours) from quasar observations for a time-varying UV efficiency in the $n$-$\tau_{\rm es}$ plane for *WMAP*-normalized models with $f_{*,\rm crit} = 1\times10^{-4}$. Also shown (*short-dashed line*) is approximately the 68% constraint from *WMAP* (Spergel et al. 2003, Fig. 5).

$\Omega_0 h^2 = 0.14 \pm 0.02$ and found an approximate scaling relation

$$\tau_{\rm es} \approx 0.11 \left(\frac{\Omega_0}{0.27}\right)^{-1} , \quad (32)$$

$$n \approx 0.96 \left(\frac{\Omega_0}{0.27}\right)^{-0.57} . \quad (33)$$

Therefore, decreasing the matter density leads to higher values for the best-fit $\tau_{\rm es}$ and $n$.

## 4. DISCUSSION

In order to understand the quasar constraints, let us return to the constant efficiency case. Consider the following "subset" of the quasar data:

Q1. Requiring $z_{\rm overlap} \geq 6$ and using only the Ly$\alpha$ constraint from quasar J1148+5251 ($z_{\rm em} = 6.43$).

Q2. Using the Ly$\beta$ constraint from SDSS 1030+0524 ($z_{\rm em} = 6.28$).

Q3. Using the Songaila & Cowie (2002) constraints at $z \lesssim 5$.

Figure 1 also shows the effects of combining these three constraints. The intersection (Q1+Q2+Q3) subset of data appears to adequately depict the upper limit of the combined data set. In particular, these three sets of data contain much of the "information" in the full $\chi^2$ treatment. The lower limit of the combined data set is actually also contained within the constraint data in Q3 but changes as a function of $z_{\rm overlap}$. Since Q3 is the union of all these constraints for all values of $z_{\rm overlap}$, the lower limit does not appear.

This illustration shows that given the 3 degrees of freedom $\Omega_0$, $\sigma_8$, and $\epsilon_{\rm UV}$, the quasar data constrains to essentially a line in this space. This is because for each $\Omega_0$, the data (Q1+Q2) at $z \sim 6$ and the data (Q3) at $z \sim 4.5$ provide strong, essentially independent constraints on the remaining parameters $\sigma_8$ and $\epsilon_{\rm UV}$. Of course, the "length" and "width" of the line (i.e., the exact shape, as illustrated in Fig. 1) depend on the likelihood function in detail.

Now let us return to the time-varying efficiency case. For each value of $f_{*,\rm crit}$, we have four parameters $n$, $\tau_{\rm es}$, $\epsilon_{\rm UV,0}$, and $A$. Heuristically, we now have three constraints on the output: the data (Q1+Q2) at $z \sim 6$, the data (Q3) at $z \sim 4.5$, and the self-consistency of $\tau_{\rm es}$ (value "in" = value "out"). Therefore, once again, we expect a "line" in the four-dimensional parameter space (for fixed $f_{*,\rm crit}$). As before, the "length" and "width" of the line will depend on the likelihood function in detail.

We note also that the implied values for the normalization in our 95% region is $\sigma_8 \sim 0.83^{+0.03}_{-0.05}$ (based on *WMAP* best-fit), which for our assumed value of $\Omega_0$ gives $\sigma_8 \Omega_0^{0.6} = 0.38^{+0.015}_{-0.025}$, just overlapping with the cluster measurements $\sigma_8 \Omega^{0.6} = 0.33 \pm 0.03$. Incorporating the range $\Omega_0 h^2 = 0.12$–$0.16$ in our results gives $\sigma_8 \Omega_0^{0.6} \approx 0.32$–$0.44$, which is entirely consistent with the cluster measurements.

The next question is, what are the implications as to the implied values of the UV efficiency? Do the efficiencies make sense? Table 5 shows a number of parameter combinations from the 95% region. The implied efficiency at high redshift is on the order $\sim 10^{-4}$; a transition occurs at $z = 15 \sim 20$, and the efficiency at low redshift is $\epsilon_{\rm UV,0} = 10^{-5.5}$ to $\sim 10^{-5}$.

TABLE 5

Parameter Combinations in the 95% Confidence Region

| | $f_{*,\rm crit} = 10^{-4}$ | | | |
|---|---|---|---|---|
| $\tau_{\rm es}$ | $n$ | $\epsilon_{\rm UV,0}$ ($\times 10^{-4}$) | $A$ | $\epsilon_{{\rm UV,high}\,z}$ ($\times 10^{-4}$) |
| 0.09 | 0.94 | 0.12 | 7.8 | 0.92 |
| 0.10 | 0.95 | 0.096 | 12 | 1.1 |
| 0.11 | 0.96 | 0.077 | 17 | 1.3 |
| 0.12 | 0.97 | 0.058 | 26 | 1.5 |

The span of these efficiencies encompasses those typically calculated through population synthesis of $10^{-5}$ to $\sim 10^{-4.5}$. Given uncertainties of factors of a few in the gas collapse fraction (our resolution factor $\epsilon_*$), dust absorption (at lower redshift), etc., the values of the efficiencies do not seem unreasonable. They certainly do not approach the upper bound for conversion from nuclear reactions of $\sim 10^{-3}$. Sokasian et al. (2003) and Wyithe & Loeb (2003b) discuss further the star formation efficiency as relating to Population II and Population III stars.

Our results are consistent with the results of Cen (2003b) in that we find that to reach $\tau_{\rm es} \geq 0.17$ requires that the spectral index is positively tilted with $n \gtrsim 1.02$. In this case the effective UV efficiency was at least 100 times greater at $z \gg 6$. However, our calculations indicate that models with such high $n$ (and hence high power spectrum normalization $\sigma_8$) are inconsistent with quasar transmission measurements at redshift $z \lesssim 5$. This latter constraint is less stringent for lower values of $\Omega_0$ but would require $\Omega_0 \sim 0.21$ for the 95% interval to overlap with $\tau_{\rm es} = 0.17$. But our analysis indicates that this would require a heavily tilted spectral index of about $n \approx 1.1$, a combination of $\tau_{\rm es}$ and $n$ that would be inconsistent with Spergel et al. (2003) (as shown in Fig. 4).

Therefore, like other authors (e.g., Haiman & Holder 2003; Cen 2003b), we find that a value of $\tau_{\rm es} = 0.17$ is inconsistent with constraints at $z \lesssim 6$ for simple models of reionization and may require more exotic (although not necessarily implausible) methods for creating ionizing photons, such as miniquasars or an X-ray background. Alternative forms of the power spectrum of density fluctuations may also be able to relax these constraints. However, we conclude that *WMAP* data taken as a whole and quasar observations at $z \lesssim 6$ in fact are *entirely* consistent for reasonable values and time dependence for the UV efficiency and a power-law (nearly scale-invariant) initial power spectrum of density fluctuations. Our analysis shows the importance of taking into account the significant degeneracy between $\tau_{\rm es}$ and the primordial spectrum index $n$ determined by *WMAP*.

Having found a consistent set of evolutionary models, let us describe their properties in slightly greater detail. We consider the following two models: a model with constant UV efficiency with $h = 0.72$, $\Omega_0 = 0.27$, $n = 0.99$, $\sigma_8 = 0.64$, $\epsilon_{\rm UV} = 2.2 \times 10^{-5}$, and $\tau_{\rm es} = 0.06$ and a model with variable efficiency with $h = 0.72$, $\Omega_0 = 0.27$, $n = 0.96$, $\sigma_8 = 0.827$, $\epsilon_{\rm UV} = 7.7 \times 10^{-6}[1 + 17 \exp(-f_*/10^{-4})]$, and $\tau_{\rm es} = 0.11$. These are both near the peak of the likelihood distribution for $\Omega_0 = 0.27$.

The Gunn-Peterson optical depths are already shown in Figure 3. The best-fit models do not show much

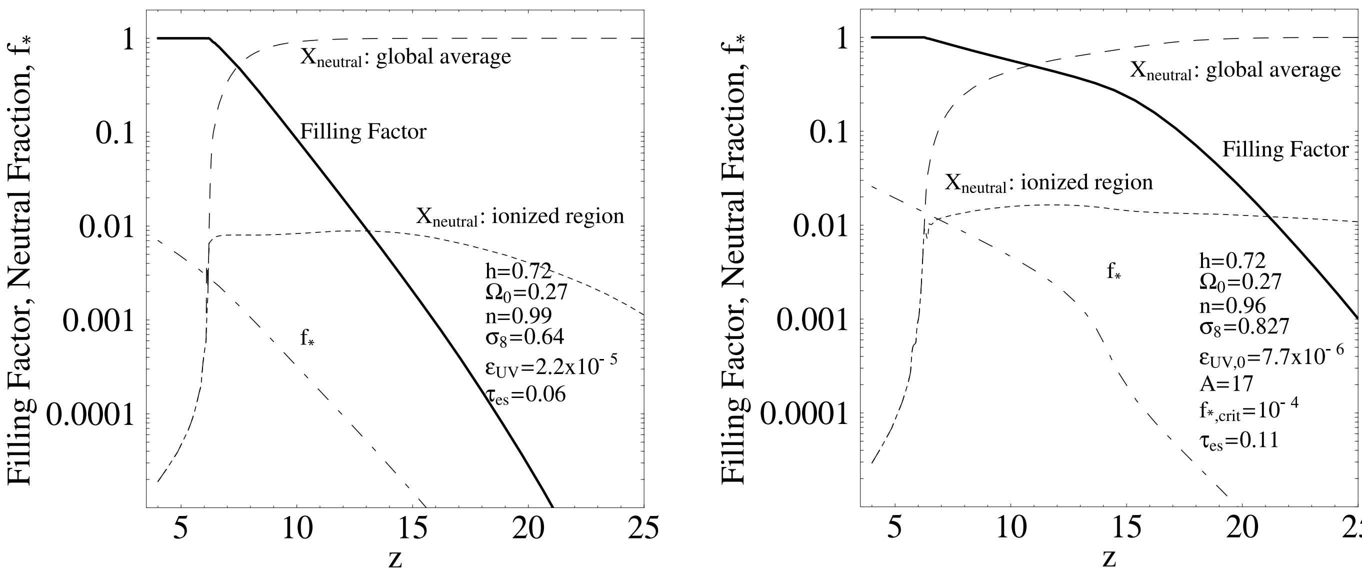


FIG. 5.—Some basic reionization properties of the best-fit models with a constant UV efficiency (*left*; $\tau_{\rm es} = 0.06$) and with a time-dependent UV efficiency (*right*; $\tau_{\rm es} = 0.11$) as a function of redshift $z$. Shown are the filling factor (*solid line*), the neutral hydrogen fraction in the ionized region (*short-dashed line*), the global average neutral hydrogen faction (*long-dashed line*), and the fraction of baryons in stars, $f_*$ (*long-short–dashed line*).

difference, as expected since the parameters were fitted to these data. Figure 5 shows the reionization properties for the two models. Although the last phase of reionization occurs rapidly in both models, the phase in which the average neutral fraction drops from unity to $\sim$0.1 takes a significantly longer time in the model with a variable UV efficiency. This is a "necessary" part of the model in order to achieve a higher electron scattering optical depth. The transition from the higher to the lower UV efficiency is clearly seen in the stellar baryon fraction $f_*$ at redshift $z \sim 15$. We note that even in the case of $\tau_{\rm es} = 0.11$, there are not two distinct epochs of ionization, in contrast to the calculations of Cen (2003a). Rather, there is an extended period during which the ionization increases from 1% to 10% before the rapid phase change at $z \sim 6$.

The thermal properties are show in Figure 6. Again, the rise in the global volume-averaged temperature is much more gradual with the variable UV efficiency. In addition, in this case, the final temperature is higher, but this is due to the higher power spectrum normalization ($\sigma_8$). These temperatures are somewhat lower that the peak temperatures derived by Hui & Haiman (2003), but this is probably due to the "sudden" reionization model used in those calculations. A more gradual reionization transition would smooth out the peak and would appear consistent with our

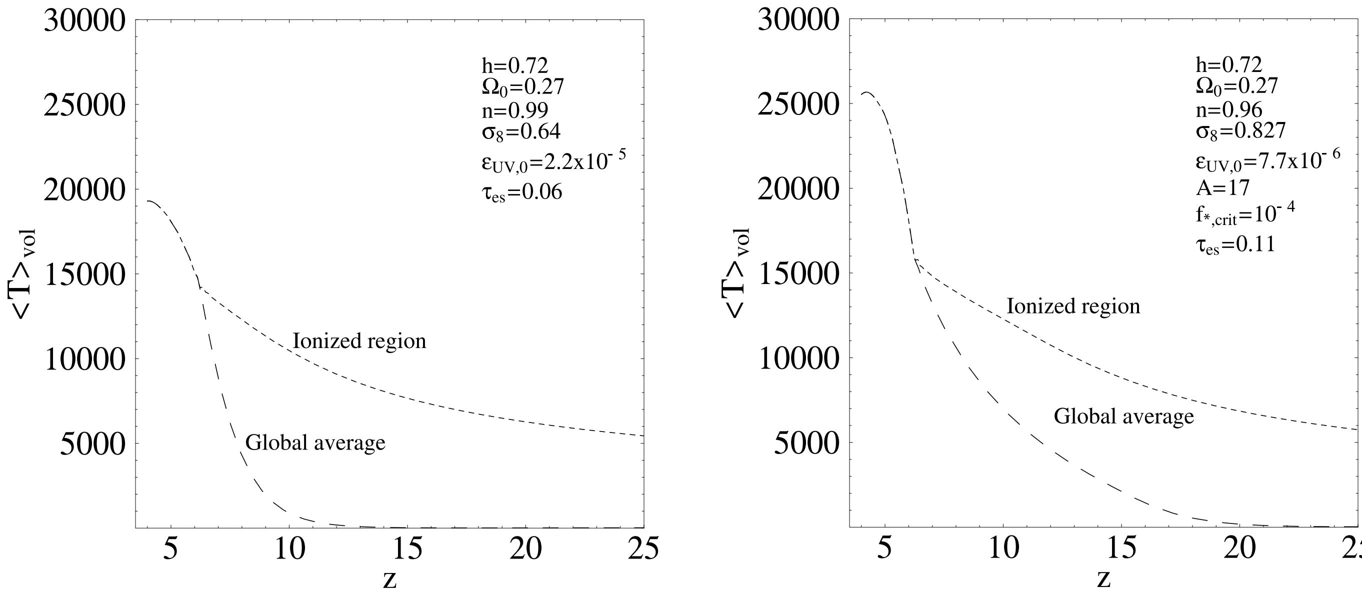


FIG. 6.—Thermal properties of the best-fit models with a constant UV efficiency (*left*; $\tau_{\rm es} = 0.06$) and with a time-dependent UV efficiency (*right*; $\tau_{\rm es} = 0.11$) as a function of redshift $z$. Shown are the volume-averaged temperature in the ionized region (*short-dashed line*) and the global volume average (*long-dashed line*).

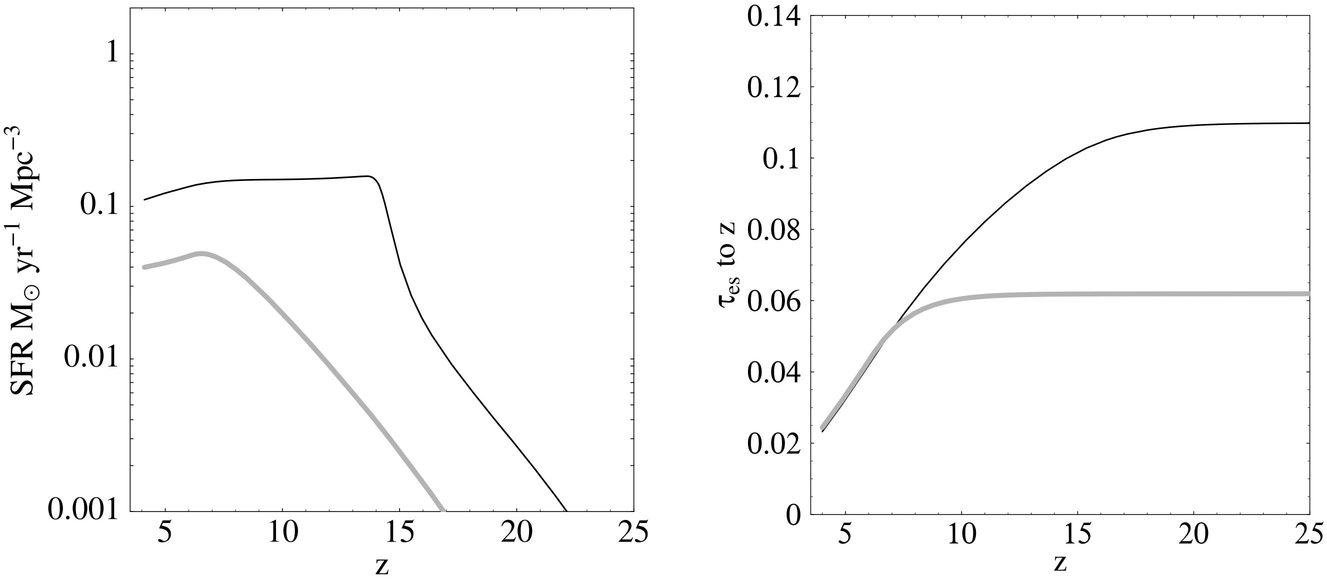


FIG. 7.—Star formation rate (*left*) and electron scattering optical depth to redshift $z$ (*right*) for the constant UV efficiency (*thick gray line*; $\tau_{es} = 0.06$) and the variable UV efficiency (*thin solid line*; $\tau_{es} = 0.11$) best-fit models.

calculations. For instance, Hui & Haiman (2003) consider a "stochastic" reheating process, and their derived temperatures at $z \sim 4$ are similar to ours.

In Figure 7 we show the star formation rate and the electron scattering optical depth to redshift $z$. Star formation rates at $z \sim 4$ from field galaxy measurements have been reported at about $10^{-2}$ to $10^{-3}$ $M_\odot$ yr$^{-1}$ Mpc$^{-1}$ (Steidel et al. 1999), while a recent lower limit of $\sim 10^{-3.1}$ $M_\odot$ yr$^{-1}$ Mpc$^{-1}$ has been reported at $z \sim 6$ (Standway, Bunker, & McMahon 2003). Our model results imply that there is exists a large population of unobserved sources at $z \sim 6$. We note that quasar observations alone require that $\tau_{es}$ has reached value of $\sim 0.05$ at $z \sim 7$. Finally, the electron scattering optical depth shows that in the constant efficiency case, the full optical depth is reached by $z \sim 10$, while in the variable efficiency case, the full optical depth is not reached until $z \sim 20$. This is an observational signature that could be detected by future CMB experiments.

We also note here that the very gradual increase of the filling factor in the models here, evolving from 0.01 to 1 over the redshift range $\sim 20$ to $\sim 6$, limits the impact of minihalos (halos with virial temperature $\lesssim 10^4$ K and that are not self-illuminating) on the overall evolution of reionization. As was discussed in Haiman et al. (2001), a rapid reionization can lead to minihalos dominating gas clumping after "overlap." This is because for rapid reionization all the minihalos that were present at the time of reionization are available for screening. However, for a more gradual reionization—one with a timescale much greater than the photoevaporation timescale—the impact is much smaller because only a few "new" minihalos would enter the ionized phase at any one time. One can crudely approximate the *relative* effects for a gradual reionization by $\sim Qt_{pe}$, where the photoevaporation time $t_{pe} \sim r_{virial}/10$ km s$^{-1}$ is from Haiman et al. (2001). We found that $Qt_{pe} < 0.05$ around the time of overlap for the best-fit models considered here, so the impact of minihalos should be a small correction to the post-overlap clumping.

A second potential effect of minihalos is that if they cluster around ionizing sources, then they can prevent ionizing photons from reaching the IGM until they are photoevaporated. This is the more likely situation here because of the more gradual nature of reionization. It is a much more complicated situation to model but could be accounted for by a more complicated calculation for the escape fraction (e.g., Barkana & Loeb 2002). However, for our analysis it would be important to calculate this minihalo escape fraction in a manner that includes dependence on background cosmology, a dependence that could be significant because the minihalo distribution depends strongly on the small-scale power spectrum. We therefore leave this problem for future work. However, we note that the net effect could accounted for by changing the time dependence of the *effective* UV efficiency. Therefore, one can interpret the "efficiencies" quoted here, which are freely parameterized, as "net" efficiencies after accounting for minihalos.

## 5. SUMMARY AND CONCLUSIONS

We derive constraints on several cosmological parameters based on observations of Gunn-Peterson absorption in high-redshift quasars and *WMAP* observations (Bennett et al. 2003). We use a semianalytic model for reionization (Chiu & Ostriker 2000) that takes into account a number of important physical processes both within collapsing halos (e.g., $H_2$ cooling) and in the intergalactic medium (e.g., $H_2$ cooling, Compton cooling, and photoionization heating). The model is also calibrated to hydrodynamic simulations. We also develop a method for estimating the gas pdf, which is important for properly calculating the mean absorption in the IGM, as a function of cosmological parameters.

We find that the Gunn-Peterson absorption data provide constraints on the power spectrum at small scales in a manner similar to that derived from the cluster mass function. Assuming that the efficiency of producing UV photons per baryon is constant, the constraint takes on the form

$\sigma_8\Omega_0^{0.5} = 0.33 \pm 0.01$ assuming a flat, $\Lambda$-dominated universe with $h = 0.72$, $n = 0.99$, and $\Omega_b h^2 = 0.024$. The best fit for the *WMAP* data (marginalized over all parameters) is reported as $\sigma_8\Omega_0^{0.5} = 0.48 \pm 0.12$, which differs by slightly more than 1 $\sigma$. However, the derived value for the optical depth to last-scattering $\tau_{\rm es} \approx 0.06$ differs significantly from the *WMAP*-determined value of $\tau_{\rm es} = 0.17 \pm 0.04$.

Since the *WMAP* constraints on $\tau_{\rm es}$ are somewhat degenerate with the value of the spectral index $n$ (Spergel et al. 2003), we then let the primordial spectral index $n$ float while fixing the best-fit *WMAP* determination of $\Omega_0 h^2 = 0.14$ and normalizing the power spectrum using *WMAP*. In addition, we allow the UV efficiency to have time dependence. Combining the *WMAP* constraints with the quasar transmission data, our analysis favors then a model with $\tau_{\rm es} = 0.11^{+0.02}_{-0.03}$, $n = 0.96^{+0.02}_{-0.03}$, implying $\sigma_8 = 0.83^{+0.03}_{-0.05}$ (all at 95% confidence), and an effective UV efficiency that was at least $\sim$10 times greater at $z \gg 6$. Including the dependence on $\Omega_0$ gives a scaling $\tau_{\rm es} \approx 0.11(\Omega_0/0.27)^{-1}$, $n \approx 0.96(\Omega_0/0.27)^{-0.57}$, implying a *WMAP* normalization $\sigma_8 \approx 0.83(\Omega_0/0.27)^{0.53}$. If future observations confirm this range for the optical depth to electron scattering $\tau_{\rm es} \sim 0.1$, then it would appear that no more "exotic" sources of UV photons are necessary. We are unable to find a model that is consistent with all observational and physical constraints that has an electron scattering optical depth $\tau_{\rm es} \gtrsim 0.17$. Thus, if additional data (for example, from the *WMAP* EE spectrum) finally require a value $\tau_{\rm es} \gtrsim 0.17$, then more exotic sources of early ionizing photons or alternative forms for the power spectrum will be required than those considered in this paper.

The authors would like to thank Renyue Cen kindly providing his simulation outputs, Martin Rees, Massimo Ricotti, and David Spergel for useful discussions, and referee Peiro Madau for his valuable comments. X. F. is grateful for support from the University of Arizona and a Sloan Research Fellowship.

## APPENDIX

## APPROXIMATE MODEL FOR ESCAPE FRACTION

In this appendix we consider a simple Strömgren sphere approximation for the escape fraction. The basic conditions for star forming halos in our semianalytic model are as follows:

1. In halos greater than the Jeans mass, the baryonic mass will collapse to a singular isothermal sphere with $\rho_b = (1/3)\Delta_v \bar{\rho}_b (r_v/r)^2$, where $\Delta_v \sim 178$ is the virial overdensity, $\bar{\rho}_b$ is the mean baryonic density, and $r_v$ is the virial radius. Here the total baryonic mass is $M_b = (4\pi/3)\Delta_v \bar{\rho}_b r_v^3$.
2. In halos where the cooling time $t_c$ is less than the dynamical time $t_d$, there will be star formation.
3. The rate of UV photon production per unit baryonic mass is $\theta = c^2 \epsilon_* \epsilon_{\rm UV} E_0^{-1} t_d^{-1}$, where $\epsilon_{\rm UV}$ is the mass to UV energy efficiency and $E_0 = 13.6$ eV. Here $\epsilon_*$ is the fraction of cooling baryons in the halo that are forming stars (the "resolution" factor) and $\epsilon_{\rm UV}$ is the fundamental conversion efficiency from stellar baryons to UV photons.

In regions that are optically thick, such that all photons are absorbed locally, then local ionization equilibrium would imply that $\theta\rho_b = n_{\rm H}^2 x^2 \alpha$, where $n_{\rm H}$ is the total hydrogen density, $x$ is the ionization fraction, and $\alpha$ is the recombination coefficient. This implies that the ionization fraction is $x = (\theta\rho_b/\alpha)^{1/2}/n_{\rm H} \propto r$. Therefore, as $r \to 0$, the ionized fraction also tends to 0.

To check for self-consistency, consider the optical depth over a central region. The average density over the sphere $r$ is $\langle \rho_b(r) \rangle = \Delta_v \bar{\rho}_b (r_v/r)^2$. Optical depth over this region is approximately $\langle \tau \rangle \sim \langle n_{\rm H}(1-x) \rangle \sigma r$, where $\sigma$ is an effective cross section. Since $\langle n_{\rm H} \rangle \sim r^{-2}$, this implies that $\langle \tau \rangle \sim (1-x) r^{-1}$. Therefore, as long as $(1-x)$ does not vanish, then the optical depth will become very large as $r \to 0$. Above, we established that $x \to 0$ as $r \to 0$ for an optically thick region. We therefore have a self-consistent picture for the central part of the halo.

Now, what is the condition so that the entire halo is optically thick? For a completely neutral halo, $\tau \sim \Delta_v \bar{\rho}_b \sigma r_v / m_{\rm eff}$. Plugging in numbers gives $\tau \sim r_v (1+z)^3/10$ kpc (in physical units). In comoving units, $\tau \propto (1+z)^2$. Therefore, at any redshift before full reionization ($z \gtrsim 6$), all halos of concern are optically thick if fully neutral. They become optically thin only if the neutral fraction is small.

Now consider a model based on a central source approximation. Let $S(r)$ be the number of ionizing photons emitted by a central source that pass through a sphere of radius $r$. The standard equation for $S(r)$ is given by

$$\frac{\partial S(r)}{\partial r} = -4\pi r^2 n_{\rm H}^2 x^2 \alpha \ . \tag{A1}$$

Using the definitions above, we obtain

$$\frac{\partial S(r)}{\partial r} = -4\pi \left( \frac{\Delta_v \bar{\rho}_b r_v^2}{3 m_{\rm eff}} \right)^2 \frac{x^2 \alpha}{r^2} \ . \tag{A2}$$

Now consider the number of photons produced within $r$, $\mathscr{S}(r) = \theta M(< r)$. Differentiating this with respect to $r$ gives

$$\frac{\partial \mathscr{S}(r)}{\partial r} = \frac{4\pi}{3} \theta \Delta_v \bar{\rho}_b r_v^2 \ . \tag{A3}$$

Now let us make the *Ansatz* that all these photons can be considered to be radially emitted. Then these two equations can be *combined* to yield

$$\frac{\partial S(r)}{\partial r} = \frac{4\pi}{3}\Delta_v \bar{\rho}_b r_v^2 \left(1 - \frac{\Delta_v \bar{\rho}_b r_v^2}{3\theta m_{\rm eff}^2}\frac{x^2\alpha}{r^2}\right) . \tag{A4}$$

Because we know that the halo is optically thick if $x < 1$, let us consider the following prescription: as long as $x < 1$, all photons are absorbed locally, so $\partial S/\partial r = 0$. This leads to the expression for the ionized fraction as a function of radius

$$x(r) = \min\left(\frac{r}{r_v}\sqrt{\frac{3\theta m_{\rm eff}^2}{\Delta_v \bar{\rho}_b \alpha}},\ 1\right) . \tag{A5}$$

If we define quenching fraction

$$\eta \equiv \sqrt{\frac{\Delta_v \bar{\rho}_b \alpha}{3\theta m_{\rm eff}^2}} , \tag{A6}$$

then $\eta \geq 1$ means that all of the photons are absorbed within the halo. If $\eta < 1$, then the ionized fraction has reached its maximum value of 1 at the quenching radius $r_q = \eta r_v$. It is important to note that $\eta$ depends only on redshift through the background density $\bar{\rho}_b$ and the dynamical time in $\theta$ and is essentially *the same for all halos*. (Actually, there is a weak dependence on the halo temperature through the recombination rate $\alpha$.)

If $\eta < 1$, then we use equation (A4) with $x = 1$ and the boundary condition $S(r_q) = 0$ to solve for $S(r)$ from $r_q$ to $r_v$. The result is

$$S(r_v) = \frac{4\pi}{3}\theta \Delta_v \bar{\rho}_b r_v^3 (1-\eta)^2 = M_b \theta (1-\eta)^2 . \tag{A7}$$

Thus, the factor $(1-\eta)^2$ is the fraction of UV photons created in the halo that actually escape—the escape fraction $\epsilon_{\rm esc}$. At high redshift, when $\eta \geq 1$, this factor is 0: all the photons are consumed within the halo.

To illustrate, consider the best-fit model with time-variable UV efficiency described in the main text. Using a temperature of $10^4$ K, $\Omega_0 h^2 = 0.14$, and $\Omega_b h^2 = 0.024$, the quenching fraction is

$$\eta \sim \frac{10^{-4}(1+z)^{0.75}}{\sqrt{\epsilon_* \epsilon_{\rm UV}}} . \tag{A8}$$

In our best-fit model, we fixed $\epsilon_* = 0.03$, while $\epsilon_{\rm UV}$ ranged from $\sim 0.8 \times 10^{-5}$ at low redshift to $1.3 \times 10^{-4}$ at high redshift. At very high redshift, we have

$$\eta \approx \left(\frac{1+z}{57}\right)^{0.75} . \tag{A9}$$

This implies that at very high redshift, when $\eta \geq 1$, all the photons are absorbed locally so that the escape fraction is essentially 0. This is to be expected since densities are much higher at high redshift. At redshifts near transition $\sim$15, when $\epsilon_{\rm UV} \sim 6 \times 10^{-5}$, we have

$$\eta \approx \left(\frac{1+z}{37}\right)^{0.75} , \tag{A10}$$

implying an escape fraction $\sim$0.2, comparable to values used in other semianalytic models. At $z \sim 6$, we have

$$\eta \approx \left(\frac{1+z}{9}\right)^{0.75} , \tag{A11}$$

implying an escape fraction $\epsilon_{\rm esc} \sim 0.03$. These values are all within the range used by others in semianalytic models (e.g., Cen 2003b; Wyithe & Loeb 2003a; Haiman & Holder 2003).

## REFERENCES

Bahcall, N. A., et al. 2003, ApJ, 585, 182
Barkana, R. 2002, NewA, 7, 85
Barkana, R., & Loeb, A. 2002, ApJ, 578, 1
Becker, R. H., et al. 2001, AJ, 122, 2850
Bennett, C. L., et al. 2003, ApJS, 148, 1
Cen, R. 2003a, ApJ, 591, 12
———. 2003b, ApJ, 591, L5
Chiu, W. A., & Ostriker, J. P. 2000, ApJ, 534, 507
Ciardi, B., Ferrara, A., & White, S. D. M. 2003, MNRAS, 344, L7
Fan, X., Narayanan, V. K., Strauss, M. A., White, R. L., Becker, R. H., Pentericci, L., & Rix, H.-W. 2002, AJ, 123, 1247
Fan, X., et al. 2003, AJ, 125, 1649
Gnedin, N. Y. 2000a, ApJ, 535, 530
———. 2000b, ApJ, 542, 535
Gnedin, N. Y., & Hui, L. 1998, MNRAS, 296, 44
Haiman, Z., Abel, T., & Madau, P. 2001, ApJ, 551, 599
Haiman, Z., & Holder, G. P. 2003, ApJ, 595, 1
Hinshaw, G., et al. 2003, ApJS, 148, 135
Hui, L., & Gnedin, N. Y. 1997, MNRAS, 292, 27
Hui, L., & Haiman, Z. 2003, ApJ, 596, 9
Kogut, A., et al. 2003, ApJS, 148, 161
McDonald, P., & Miralda-Escudé, J. 2001, ApJ, 549, L11
Melchiorri, A., Bode, P., Bahcall, N. A., & Silk, J. 2003, ApJ, 586, L1

Miralda-Escudé, J., Cen, R., Ostriker, J. P., & Rauch, M. 1996, ApJ, 471, 582
Miralda-Escudé, J., Haehnelt, M., & Rees, M. J. 2000, ApJ, 530, 1
Page, L., et al. 2003, ApJS, 148, 233
Press, W. H., & Schechter, P. 1974, ApJ, 187, 425
Shapiro, P. R., & Giroux, M. L. 1987, ApJ, 321, L107
Sokasian, A., Abel, T., Hernquist, L., & Volker, S. 2003, MNRAS, 344, 607
Songaila, A., & Cowie, L. L. 2002, AJ, 123, 2183
Spergel, D., et al. 2003, ApJS, 148, 175
Stanway, E. R., Bunker, A. J., & McMahon, R. G. 2003, MNRAS, 342, 439
Steidel, C. C., Adelberger, K. L., Giavalisco, M., Dickinson, M. E., & Pettini, M. 1999, ApJ, 519, 1
Verde, L. 2003, ApJS, 148, 195
White, R. L., Becker, R. H., Fan, X., & Strauss, M. A. 2003, AJ, 126, 1
Wyithe, J. S. B., & Loeb, A. 2003a, ApJ, 586, 693
———. 2003b, ApJ, 588, L69