A SEMIANALYTIC MODEL FOR COSMOLOGICAL REHEATING AND REIONIZATION DUE TO THE GRAVITATIONAL COLLAPSE OF STRUCTURE

WEISUEH A. CHIU
Joseph Henry Laboratories, Department of Physics, Princeton University, Princeton, NJ 08544; chiu@astro.princeton.edu

AND

JEREMIAH P. OSTRIKER
Princeton University Observatory, Princeton, NJ 08544; jpo@astro.princeton.edu

Received 1999 July 16; accepted 1999 December 10

ABSTRACT

We present a semianalytic model for the thermal and ionization history of the universe at 1000 ≳ z ≳ 3. This model incorporates much of the essential physics included in full-scale hydrodynamical simulations, such as (1) gravitational collapse and virialization; (2) star/quasar formation and subsequent ionizing radiation; (3) heating and cooling; (4) atomic and molecular physics of hydrogen; and (5) the feedback relationships between these processes. In addition, we model the process of reheating and reionization using two separate phases, self-consistently calculating the filling factor of each phase. Thus, radiative transfer is treated more accurately than simulations published to date have done: we allow to lowest order for the inhomogeneity of the sources and the sinks of radiation. After calibrating and checking the results of this model against a hydrodynamical simulation, we apply our model to a variety of Gaussian (adiabatic power spectra) and non-Gaussian (texture and isocurvature) cold dark matter (CDM)–dominated cosmologies normalized to cluster abundances. Our model is also normalized to observations of the ionizing UV intensity J_{21} ≃ 1 at redshift z = 4. Our major conclusions include: (1) the epoch of reheating (starting late at z ≳ 30 or early at z ≳ 80) depends most strongly on the power spectrum (late: adiabatic; early: texture or isocurvature); (2) because of the effects of gas clumping, full reionization occurs at z ≳ 10 in all models; (3) the cosmic microwave background radiation (CMBR) polarization anisotropy will be a strong discriminant between late and early reheating models; (4) the fraction of baryons sequestered in stars and quasars in early reheating models appears to be greater than the observational limit, while the fraction in late reheating models is well below it; (5) the average degree of nonlinearity for collapsing baryons remains roughly constant during reheating, a possible explanation of which is feedback, which regulates the pace of reheating through the Jeans criterion; and (6) the evolution of the bias of luminous objects can potentially discriminate strongly between Gaussian and non-Gaussian probability density functions.

Subject headings: cosmology: theory — galaxies: formation — intergalactic medium — methods: analytical

1. INTRODUCTION

In this paper we semianalytically model the thermal and ionization history of the universe and its effects on the formation of luminous objects. Many parts of this complex early evolution of the universe have been studied previously, beginning with the papers of Binney (1977), Silk (1977), and Rees & Ostriker (1977), the last of whom used models of the thermal history of the universe to investigate the formation of galaxies. Tegmark et al. (1997) used a simplified model of dark matter halo profiles to derive the minimum mass of objects that are able to cool in a Hubble time. The evolution of the intergalactic medium (IGM) has been explored semianalytically by numerous authors, including Shapiro, Giroux, & Babul (1994), Haiman & Loeb (1998), and Valageas & Silk (1999). The work of Haiman & Loeb (1998) was a hybrid scheme utilizing hydrodynamical simulations of spherically symmetric collapse (Haiman & Loeb 1997). Recently, Ciardi et al. (1999) used a semianalytic model for galaxy and star formation in conjunction with high-resolution N-body simulations. The detailed evolution of the IGM was addressed in a full hydrodynamical simulation by Ostriker & Gnedin (1996) and Gnedin & Ostriker (1997; hereafter GO97). These studies explicitly allowed for clumping of the IGM, but were only able to simulate a small volume (a cube of side 2 h^{-1} Mpc). Ostriker & Gnedin (1996) found that a mass resolution of about 10^4 M_☉ was required to resolve early epochs of reheating and reionization. A more recent simulation (Gnedin 2000) incorporates an inhomogeneous radiation field for the first time.

The more general problem of galaxy (as opposed to dark matter halo) formation has also been addressed in full hydrodynamical simulations by many investigators, including Katz, Weinberg, & Hernquist (1996), GO97, and recently by Pearce et al. (1999). Pearce et al. (1999) simulated a large volume, but with much coarser mass resolution than GO97, which thus was not meant to address the details of reheating and reionization. Kauffmann, Nusser, & Steinmetz (1997) developed a hybrid scheme, using a semi-
analytic model of gas dynamics and star formation in conjunction with $N$-body simulations of the formation of dark matter halos. In most cases, including all the hydrodynamical studies cited here, only Gaussian cold dark matter (CDM) has been considered.

Here we attempt to incorporate much of the essential physics modeled by GO97 into a semianalytic model for the reheating and reionization of the universe. The work is not explicitly three-dimensional. Rather, we treat the multiphase medium in a statistical fashion and treat collapse in the approximation of the Press-Schechter method (Press & Schechter 1974). We consider the universe as an evolving two-phase medium, one phase (I) being neutral and the other (II) ionized. Of course, we still must make many assumptions as to which physical processes are important and which can be neglected. The main assumption we make is that UV photons from an early generation of stars and quasars are the main source of energy for the reheating and reionization of the universe.

Our model also includes the following physical processes: (1) gravitational collapse and virialization, (2) nonequilibrium atomic and molecular physics of hydrogen, (3) heating and ionization by UV photons, (4) cooling, (5) simulated star/quasar formation, (6) clumping, and (7) feedback between star formation and the thermal and ionization evolution of the universe. All of these processes are calculated within the context of the standard evolution of the background cosmological model. In determining the evolution of structure formation, we explore several models for the probability density function (PDF) and power spectrum of the initial cosmological density field. Once the background cosmology and normalization of the power spectrum are fixed, our semianalytic model has only two free parameters, related to the efficiency of star formation and UV photon generation. These two parameters we fix by calibrating our results with GO97 and requiring each cosmological scenario to reproduce approximately the observed ionizing photon intensity at $z = 4$. The model we develop neglects the following processes: (1) spatial variations and correlations beyond the two phases, (2) nonvirialization shock heating from gravitational collapse and supernova explosions, (3) local optical depths to UV photons and the frequency evolution of the photon energy spectrum, and (4) helium, dust, and heavy elements in the IGM. Because these effects are likely to be important at late times, we only consider our models a reasonable approximation for $z \geq 3$.

A word on motivation may be useful here. In almost all respects, the full numerical hydrodynamical simulations provide a better physical treatment than we are implementing here; for example, the neglected processes 1–4 not treated by the current work are included in GO97. Our treatment, however, is preferable in two related ways: it is not limited by numerical resolution, and it allows, to lowest order, for a multiphase (ionized/neutral) IGM that follows from allowing for the spatial distribution of radiation sources. Given the overall relative strengths of the full numerical treatment, one can ask why bother to do anything less? The answer is threefold:

1. First, residual doubts about numerical resolution always remain, whatever tests are made, so it is important to see how robust the conclusions of the numerical work are when resolution limits are removed.

2. The full numerical approach is so expensive that only a very small number of models can be explored. The result is that one is again concerned with questions of the robustness and generality of the results to variations of the adopted model. How sensitive are conclusions to the assumed nature of the perturbations (Gaussian/non-Gaussian), amplitude of the power spectrum, etc.? Semianalytic methods allow one to explore the parameter space.

3. The full numerical results, paradoxically, in so far as they are better and increasingly comprehensive in physical treatment, tell one less and less. Understanding depends on knowing which physical processes and initial conditions are essential to the results, and which merely modify the details of the overall picture. A semianalytic treatment allows one to add and subtract from the modeling to determine what matters and what does not.

The organization of this paper is as follows. In § 2, we present the elements of our adopted physical model for reheating and reionization. In § 3, we apply it to several Gaussian and non-Gaussian cosmological models. In § 4, we discuss some potential observational constraints on these models, as well as the nonlinearity and clustering bias of collapsing baryons. We summarize our results in § 5. In the appendices, we delineate our treatment of the standard physics of the universe, as well as the specific calculational implementation (e.g., finite-difference equations) of our semianalytic model.

2. A MODEL FOR THE PHYSICS OF REIONIZATION

In this section, we describe the semianalytic model we use for the physics of reionization. We concentrate on the treatment of a two-phase medium with sources and sinks of radiation. In particular, we define our basic differential equations and describe how we calculate the formation rate of bound objects, the UV production rate, and the clumping of cosmic gas.

2.1. Definitions and Conservation Equations

Since we are considering a two-phase medium, we will need to understand the time evolution of the filling factor of ionized regions, $f_\text{II}$. The two other independent quantities required are the ionization fraction in the ionized region, $x_\text{II}$, and the mean (volume-averaged) UV ionizing ($\geq 13.6$ eV $\equiv \epsilon_0$) intensity, $J$. The corresponding intensity in the ionized region is $J_\text{II} = J/f_\text{II}$. In our notation, $n_\text{HI}$ is the total mean hydrogen density.

Considering only hydrogen ionization, local particle conservation in the ionized region gives

$$\frac{d}{dt} (n_\text{HI} x_\text{II}) = -3 \frac{\dot{a}}{a} n_\text{HI} x_\text{II} + n_\text{HI} (1 - x_\text{II}) \frac{4\pi J_\text{II} \sigma_\text{P}}{\epsilon_0}$$

$$+ \langle \epsilon \rangle n_\text{HI}^2 (1 - x_\text{II}) k_{ci} - x_\text{II}^2 \langle \epsilon \rangle R \rangle$$

$$= -3 \frac{\dot{a}}{a} n_\text{HI} x_\text{II} + n_\text{HI} (1 - x_\text{II}) \frac{\langle \epsilon \rangle \sigma_\text{P}}{J_\text{II} \epsilon_0}$$

$$+ \langle \epsilon \rangle n_\text{HI}^2 (1 - x_\text{II}) k_{ci} - x_\text{II}^2 \langle \epsilon \rangle R \rangle,$$ (1)

where the first term on the right-hand side is from the expansion of the universe, the second term is from photoinionization, and the third term is from collisional ionizations (rate coefficient $k_{ci}$) and radiative recombinations (rate coef-
The photoionization rate per hydrogen atom is given (in ergs s\(^{-1}\)) by

\[ \frac{dE}{dt} = S - [4 + (\gamma - 1)] \frac{\dot{a}}{a} E - n_0 n_H (1 - x_H) 4\pi J_H \sigma_x - \epsilon_0 n_0 x_H \frac{df_H}{dt} \]

(3)

where we have used the relationship between \( E, E_H, \) and \( J_H \).

We use a value of \( \gamma = 1.5 \) throughout. The integrated source function is given by \( S \), the Hubble expansion is accounted for in the second term, energy lost to maintaining the ionization of the ionized region is given by the third term, and the last term is due to energy lost to ionizing new (i.e., previously neutral) material. The effective UV energy loss cross section is \( \sigma_x \).

The definitions of \( \sigma_x \) and \( \epsilon_x \) are as follows. If the ionizing specific intensity is \( J_x \), then the ionizing intensity is

\[ J = \int_{v_0}^{v_x} J_x dv, \]

(5)

where \( v_x \) is a high-frequency cutoff (we use 7.4 keV).\(^2\) For a power law

\[ J_x = J_0 v_0^{-\gamma} \left( \frac{v}{v_0} \right)^\gamma, \]

(6)

we have the relation

\[ J = \frac{J_0}{\gamma - 1} \left[ 1 - \left( \frac{v_0}{v_x} \right)^{(\gamma - 1)} \right]. \]

(7)

The photoionization rate per hydrogen atom is given (in s\(^{-1}\)) by

\[ \Gamma_x \equiv \int_{v_0}^{v_x} 4\pi J_x \sigma_x \frac{dv}{h_p v} \]

(8)

\[ = 4\pi J_0 \int_{v_0}^{v_x} \left( \frac{v_0}{v} \right)^\gamma \sigma_x \frac{dv}{v}, \]

(9)

where \( h_p \) is Planck’s constant. Equating this with \( 4\pi J_x \sigma_x/\epsilon_0 \) gives

\[ \sigma_x \equiv \frac{\gamma - 1}{1 - (v_0/v_x)^{(\gamma - 1)}} \int_{v_0}^{v_x} \left( \frac{v_0}{v} \right)^\gamma \sigma_x \frac{dv}{v}. \]

(10)

Similarly, assuming that all the excess energy from each photoionization goes into heat and produces no UV photons, the rate of UV energy loss per hydrogen atom (ergs s\(^{-1}\)) is

\[ \Gamma_y \equiv \int_{v_0}^{v_y} 4\pi J_y \sigma_y \frac{dv}{h_p v} \]

(11)

\[ = 4\pi J_0 \int_{v_0}^{v_y} \left( \frac{v_0}{v} \right)^\gamma \sigma_y \frac{dv}{v}. \]

(12)

Equating this with \( 4\pi J_x \sigma_x \) gives

\[ \sigma_y \equiv \frac{\gamma - 1}{1 - (v_0/v_y)^{(\gamma - 1)}} \int_{v_0}^{v_y} \left( \frac{v_0}{v} \right)^\gamma \sigma_y \frac{dv}{v}. \]

(13)

The difference between these cross sections gives the cross section for photoheating, or, equivalently, a photoheating rate (ergs s\(^{-1}\)) of

\[ \Gamma_{ph} \equiv 4\pi J_x (\sigma_x - \sigma_y). \]

(14)

In order to fully specify the model, we need to be able to determine \( S \) as well as have a relationship between \( E \) and \( f_H \). These turn out to be intimately related.

2.2. Formation Rate of Bound Objects

We assume that stars/quasars only form in virialized collapsed objects. Hence, the first input into the star formation rate is the formation rate of bound objects. In § 2.3, we relate these bound objects to UV energy output in order to derive expressions for the source function \( S \) and the relationship between the mean radiation energy density, \( \langle E \rangle \), and the filling factor, \( f_H \).

The generalized Press-Schechter formalism for a PDF \( P(v) \) and normalization factor \( f^{-1} \equiv \int_0^\infty P(v) dv \) gives the comoving number density of collapsed objects with initial comoving radius between \( R \) and \( R + dR \) as

\[ N_R dR = f P[v(R, t)] \frac{3}{4\pi R^3} \frac{\delta \ln v(R, t)}{\delta \ln R} v(R, t) \frac{dR}{R}, \]

(15)

\[ = \frac{\delta_c}{\sigma(R) D(t)}, \]

(16)

\[ \frac{\delta \ln v(R, t)}{\delta \ln R} = \frac{d \ln \sigma^{-1}(R)}{d \ln R} = \sigma(R), \]

(17)

where \( \sigma(R) \) is the rms fluctuation on a scale \( R \) linearly extrapolated to the present, \( \delta_c \approx 1.69 \) is the standard linear overdensity of collapse, and \( D(t) \) is the linear growth factor normalized to \( D(t_0) = 1 \). The original Press-Schechter (Press & Schechter 1974) method was derived for a Gaussian PDF, and has been checked in that case with \( N \)-body simulations (Efstathiou et al. 1988). These expressions in the non-Gaussian case were derived by Chiu, Ostriker, & Strauss (1998) and recently tested through \( N \)-body simulations for several non-Gaussian models (Robinson & Baker 1999).

For each model, we filter the power spectrum using a top-hat window to obtain the rms fluctuation, \( \sigma \). The power spectra we use are from Bardeen et al. (1986) and Sugiyama (1995) in the adiabatic cases; Pen, Spergel, & Turok (1994) and Pen & Spergel (1995) in the textures case; and Peebles (1999a, 1999b) in the isocurvature CDM (ICDM) case.

In general, we cannot use only the Press-Schechter number density to determine the formation rate of bound objects. This is because the net time rate of change of the number density of objects at a particular scale consists of both a formation rate \( (\dot{N}_{\text{form}}) \) and a destruction rate...
Thus, the actual formation rate may be higher than the rate inferred from Press-Schechter alone. Of course, we must constrain these rates to be consistent with the Press-Schechter formalism,

$$\dot{N}_R = \dot{N}_{R, \text{form}} - \dot{N}_{R, \text{dest}} \tag{18}$$

where \(\dot{N}_R\) is given by differentiating equations (15)–(17) with respect to time. We seek a semianalytic methodology applicable to an arbitrary PDF that gives these two rates.

Sasaki (1994) models the formation and destruction rates of these objects by assuming that the destruction efficiency rate, \(\phi \equiv \dot{N}_{R, \text{dest}}/\dot{N}_R\), has no characteristic scale; “destruction” by incorporation into larger objects is hierarchical and scale-invariant. In this section, we generalize his derivation to non-Gaussian PDFs.

The assumption of no characteristic scale of destruction efficiency, in addition to the requirement that the net rate of formation minus destruction be consistent with Press-Schechter, leads to the conclusion that \(\phi\) must be a function of only time. Sasaki (1994) derives the following relation for the destruction probability per unit time:

$$\phi(t) = \frac{D}{D}, \tag{19}$$

which we find to be independent of the PDF. It follows that the probability that an object created at \(t_1\) exists at \(t_2\) without having been incorporated into a larger object is given by

$$p_{\text{surv}}(t_2 | t_1) = \exp \left[ -\int_{t_1}^{t_2} \phi(t) dt \right] = \frac{D(t_1)}{D(t_2)}. \tag{20}$$

Combining \(\phi\) and the Press-Schechter consistency requirement, equation (18) leads to a formation rate of the form

$$\dot{N}_{R, \text{form}} = N_{R, v}(R, t) \frac{D}{D} \left( -\frac{P'}{P} \right), \tag{21}$$

where \(P' = dP/dv\). The quantity \(-P'/P\) for a Gaussian and an exponential \(P(v)\) takes on the forms

$$\frac{-P'}{P} = v, \quad \text{if } P(v) \propto e^{-v^2/2}; \tag{22}$$

$$\frac{-P'}{P} = \omega, \quad \text{if } P(v) \propto e^{-av} \tag{23}$$

Textures and non-Gaussian ICDM both have PDFs that are well approximated by exponentials (Gooding et al. 1992; Park, Spergel, & Turok 1991; Peebles 1999a, 1999b). For textures, \(\omega = 1.45\) (using Park et al. 1991 as described in Chiu et al. 1998), and for non-Gaussian ICDM, \(\omega = 0.67\) (Peebles 1999b). Recalling that

$$\frac{D}{D} = f(\Omega) \frac{\dot{a}}{a}, \tag{24}$$

where the velocity factor \(f(\Omega) \approx \Omega^0.6\), the formation rate in closed form is

$$\dot{N}_{R, \text{form}} dR = f[-P'(v)] \frac{3}{4\pi R^3} f(\Omega) \frac{\dot{a}}{a} \alpha(R) v^2(R, t) \frac{dR}{R} \tag{25}$$

This is the formation rate per comoving volume we use in our model. For a power law \(\sigma(R) \propto R^{-\alpha}\), the logarithmic slope \(\alpha \ln v/\ln R = \alpha\) is a constant.

This formalism for merging and formation rates is consistent with the work of Blain & Longair (1993), who used a numerical extension of Press-Schechter with simple forms for the merging probability. All other work of this type is based on Lacey & Cole (1993, 1994), who construct conditional probabilities by extending the work of Bond et al. (1991). Unfortunately, this method is heavily dependent upon the PDF being Gaussian, and is sufficiently complicated that extending it to non-Gaussian PDFs is not straightforward. Furthermore, since objects are “constantly growing” in the Lacey & Cole (1993) formalism, rather than in “punctuated equilibrium” as implied by equation (20), it is difficult to directly compare the two methods. However, one sign that they might be roughly consistent is that Viana & Liddle (1996) obtain similar results for cluster abundances whether they use the Lacey & Cole (1993) or the Sasaki (1994) methodology.

We further note that during reheating and reionization, structure formation is largely hierarchical in most of the models we consider. This is because only a small fraction of the universe needs to collapse in order to release enough energy to reheat and reionize the universe. Collapsed objects are thus still “rare” \((v > 1)\), and their net formation rates are thus dominated by their creation rates. The use of the Sasaki results therefore provide only a small correction. In the limit where the “destruction” rates are negligible, all semianalytic merging formalisms based on Press-Schechter must give the same answer in order to be consistent with equation (18).

### 2.3. UV Photon Production and the Filling Factor of the Ionized Universe

Here we discuss our model of UV photon production and how the UV energy density is related to the fraction of the universe that is ionized. A summary of our model is as follows. We assume that star formation in collapsed halos is the source of UV photons. The number density of these sources is determined using the Press-Schechter formalism, with the requirement that the baryonic halo mass is greater than the Jeans mass; the luminosity of each source is determined by the fraction of cooling gas and several efficiency factors, as well as the time since its collapse. These sources are assumed to be scattered randomly (i.e., a Poisson point process), and each surrounded by a Strömgren sphere whose radius depends on the mass, time of formation, and time since formation of the object, and the ionization state of the ionized region as a whole. The volume emissivity of these objects is assumed to be uniform, and they are assumed to radiate radially into the surrounding IGM. Assuming that at a fixed time, the volume of each Strömgren sphere per unit luminosity is the same as that for the universe as a whole (and therefore is a function only of time), we require consistency among the radius of each sphere, the filling factor, the mean intensity, the number density of sources, and the total UV luminosity. This consistency determines the relationship between the filling factor and the UV energy density of the universe.

#### 2.3.1. Determining the UV Production Rate

Our basic model for UV production is that individual halos that can undergo star formation will emit UV radi-
ionized radially, and thus be surrounded by a sphere of material, a “cosmological Strömgren sphere.” Consider halos of baryonic mass in the range $M_b$ to $M_b + dM_b$, corresponding to a comoving length scale between $R$ and $R + dR$, formed in the time interval $\tilde{t}$ to $\tilde{t} + d\tilde{t}$, and “observed” at time $t$. The comoving number density of such objects is given by

$$N(R, \tilde{t}, t) dR d\tilde{t} = \tilde{N}_{\text{form}}(R, \tilde{t}) p_{\text{surv}}(t | \tilde{t}) dR d\tilde{t}, \quad (26)$$

where the notation is the same as in the previous section. The second factor accounts for the destruction of objects of scale $R$ through merging into larger objects. Although we are using Sasaki’s method to determine $\tilde{N}_{\text{form}}$ and $p_{\text{surv}}$ as described above, the form of equation (26) is completely general. Note that $M_b$ and $R$ are related by $M_b = 4\pi R^3 \rho_{\text{b,0}}/3$. Following the prescriptions of Cen (1992) and Gnedin (1996), we assume that star formation is spread over a dynamical timescale, so that the UV luminosity of such an object is given by

$$L(M_b, \tilde{t}, t) = (1 - f^*_b) M_b c^2 \epsilon_{\text{cool}} \epsilon_b \epsilon_{\text{hm}} \epsilon_{\text{UV}} \left(\frac{t - \tilde{t}}{t_{\text{dyn}}}\right),$$

where $f^*_b$ is the fraction of the baryonic mass that has already been processed into stars/quasars and $t_{\text{dyn}}$ is the dynamical time associated with the object at its time of formation (see eq. [C1]). We are also assuming here that the stars/quasars are uniformly distributed throughout the collapsed halo.

The efficiencies, $\epsilon^*$, are as follows: $\epsilon_{\text{cool}}$ is the mass fraction of the halo that is cooling, $\epsilon_b$ is the mass fraction of cooling gas that physically ends up in collapsed objects, $\epsilon_{\text{hm}}$ is the mass fraction of collapsed objects that are quasars or high-mass stars (i.e., with substantial UV flux), and $\epsilon_{\text{UV}}$ is the mass-to-UV efficiency of the high-mass stars/quasars. The last two have been previously discussed in GO97 and Gnedin (1996). The cooling fraction, $\epsilon_{\text{cool}}$, we derive below in Appendix C; it is a function of the halo virial mass and the redshift of collapse. We cannot use $\epsilon_{\text{UV}}$ directly because we are not resolving the mass elements within individual cooling halos. The resolution factor, $\epsilon^*$, parameterizes this uncertainty. We use a net efficiency

$$\epsilon_{\text{eff}} = \epsilon_b \epsilon_{\text{hm}} \epsilon_{\text{UV}}, \quad (28)$$

normalized by requiring, for each model, that the ionizing intensity $J_{21} = 1$ at $z = 4$. We separately estimate $f^*_b$ (which is necessary in order to derive $f^*$) using the results of GO97 (who use $\epsilon_{\text{UV}} = 6 \times 10^{-5}$ and $\epsilon_{\text{hm}} = 0.16$), and obtain a value of $f^*_b \approx 0.5$. We use this same resolution factor in all our models. With this specified, there is, effectively, only one free parameter here, $\epsilon_{\text{eff}}$, and that is fixed by normalizing to observations.

Given $f^*_b$, we calculate the stellar fraction, $f^*$, by integrating over the UV production rate $S$, given in equation (31), and dividing by the appropriate efficiencies:

$$f^*(t) = \int_0^t \frac{S(\tilde{t}) d\tilde{t}}{\epsilon_b \epsilon_{\text{hm}} \epsilon_{\text{UV}} n_{\text{H}} m_{\text{H}} c^2 / X},$$

and

$$f^*_b = \frac{\epsilon^* \int_0^t S(\tilde{t}) d\tilde{t}}{\epsilon_{\text{eff}} n_{\text{H}} m_{\text{H}} c^2 / X}. \quad (30)$$

Here $X$ is the hydrogen mass fraction, and we have expressed the result in terms of the $J_{21}$-normalized effective efficiency, $\epsilon_{\text{eff}}$, and the resolution factor, $\epsilon^*$. As a check on our value of $f^*_b$, we compare $f^*_b$ at $z = 4$ between our model and the GO97 simulation, given the same cosmological parameters. We find the values to be consistent, with $f^*_b \approx 2\%$ at this redshift (see also Fig. 1).

The global mean UV production rate per unit (physical) volume is given by

$$S(t) = \dot{a}(t)^{-3} \int_0^\infty dR \int_0^{|\tilde{t}|} d\tilde{t} \psi(R, \tilde{t}) N(R, \tilde{t}, t) L(M_b, \tilde{t}, t), \quad (31)$$

where the function $\psi(R, \tilde{t})$ is a selection function for objects that are able to collapse and form stars/quasars. In order to account for the fact that the Jeans mass will be different in the two phases, we define $\psi$ to be

$$\psi(R, \tilde{t}) = \{(1 - \tilde{f}) \theta[R - R_{\text{Jeans}}(\tilde{t})] + \tilde{f}_\text{th} \theta[R - R_{\text{Jeans}}(\tilde{t})]\}, \quad (32)$$

where $\theta$ is the step function (zero if the argument is $<0$ and unity if the argument is $>0$), $\tilde{f}$ is the filling factor of the ionized region at time $\tilde{t}$, $R_{\text{Jeans}}$ is the Jeans length in the neutral (cold) region, and $R_{\text{Jeans}}$ is the Jeans length in the ionized (hot) region. Note that we are assuming that an object, once formed, can only be destroyed by merging and forming a larger object. We are thus assuming that virialized gas halos are “self-shielded” from ionizing radiation due to their high density. That is, they cannot be heated to the point of dynamical instability, nor can they be “evaporated” away.

2.3.2. Determining the Relationship Between the Filling Factor and the UV Energy Density

Because the differential equation (4) includes derivatives of both the filling factor, $f^*$, and the UV energy density, $E$, we require an additional relationship between the two to close our system of equations. To do this, we consider defi-
nitions of the filling factor and energy density in terms of UV sources and the ionized volumes surrounding them.

If we make the standard assumption that the UV sources are randomly distributed, then the filling factor can be related to the average number of “overlapping” ionized regions, \( \lambda(t) \), seen by a random volume element:

\[
\lambda(t) = \int_0^\infty dR \int_0^t \! \! \! d\bar{t} \psi(R, \bar{t}) N(R, \bar{t}, t) \frac{4\pi}{3} r_{II}(M_b, \bar{t}, t). \tag{33}
\]

The quantity \( r_{II} \) is the comoving radius of the ionized region around the collapsed object (assumed to be much smaller than the local Hubble radius). For Poisson probabilities, the filling factor is related to \( \lambda(t) \) by

\[
f_{II}(t) = 1 - e^{-\lambda(t)} . \tag{34}
\]

Similarly, considering the contribution of each source to the mean intensity \( \bar{J}(t) \) gives

\[
4\pi \bar{J}(t) = \int_0^\infty dR \int_0^t \! \! \! d\bar{t} N(R, \bar{t}, t) \int_0^{a(t)R} L(M_b, \bar{t}, t) \frac{4\pi}{3} [a(t)R]^3 \tag{35}
\]

\[
= \int_0^\infty dR \int_0^t \! \! \! d\bar{t} N(R, \bar{t}, t) L(M_b, \bar{t}, t) \times r_{II}(M_b, \bar{t}, t)a^{-2}(t). \tag{36}
\]

Here we have neglected the optical depth within an isolated ionized region, and assume that no flux escapes beyond \( r_{II}(M_b, \bar{t}, t) \). In addition, we have assumed that each central object emits radiation radially. Therefore, the mean energy density, \( \bar{E}(t) = 4\pi \bar{J}(t)/c \), at time \( t \) is given by

\[
\bar{E}(t) = \frac{1}{c} \int_0^\infty dR \int_0^t \! \! \! d\bar{t} N(R, \bar{t}, t)L(M_b, \bar{t}, t) \times r_{II}(M_b, \bar{t}, t)a^{-2}(t). \tag{37}
\]

To relate equations (33) and (37), we need a relationship between the luminosity of an object and the radius of the ionized sphere around it. If we consider the ionized phase as a whole, with its ionization history governed by equation (2), the effective total ionized volume per unit luminosity is \( \lambda/(a^3 S) \) (for the actual volume, we must use \( f_{II} \) instead of \( \lambda \)). This is because \( \lambda \) gives the number of times to count each volume element (expected number of overlaps if \( \lambda \gtrsim 1 \)), and \( a^3 S \) is the average comoving luminosity density. Because of the aggregate nature of our model, we make the lowest order self-consistent assumption that the effective volume per unit luminosity for each source is the same that for the total:

\[
\frac{4\pi r_{II}^3}{3L(M_b, \bar{t}, t)} = \frac{\lambda(t)}{a^3 S(t)}. \tag{38}
\]

![Graph](image-url)

**Fig. 2.—** Results of reionization calculation with GO97 parameters: \( f_{II} \) is the filling factor; \( \langle J_{21} \rangle \) is the global average of the ionizing intensity \( J_{21} \); \( \langle 1 - x \rangle \) is the neutral fraction (thick line, global average; thin line, ionized region); and \( \langle T \rangle \) is the temperature. Results from the GO97 simulation are denoted by asterisks.
Using equation (38) in equation (37) gives

\[ E(t) = \frac{1}{c} \left[ \frac{3\lambda(t)}{4\pi S(t)} \right]^{1/3} a^{-3} \int_0^\infty dR 
\times \left[ \int_{\hat{t}}^t d\tilde{t} \psi(R, \tilde{t}) N(R, \tilde{t}, t) L^{2/3}(M_b, \tilde{t}, t) \right]. \] (39)

This defines a relationship between \( E \) and \( \lambda \) in terms of independently determined quantities.

Note that equation (38) is equivalent to assuming that the ionized volume around an isolated object is proportional to its UV luminosity with a coefficient that depends only on time \( t \), and requiring self-consistency. This is obviously a simplification, but should be a reasonable approximation, since we are considering the effects of luminous objects on the IGM, where the ionization physics is a function of the temperature and density at \( t \). In addition, significant parts of the “history” of the luminous object (i.e., its star formation history as well as its merging history) are already accounted for by equations (26)–(27). As long as objects live longer than the recombination time, this assumption is consistent with the work of Shapiro & Giroux (1987) on cosmological \( \text{H II} \) regions.

We also have neglected spatial correlations among these sources. This approximation should be less important for the texture model, which is well approximated by uncorrelated seeds, than for the Gaussian models. As long as the filling factor is small, however, these correlations are likely not to be important, since most of the universe is neutral. We find below that in the Gaussian cases we consider, reionization occurs quite rapidly, and this rapid percolation likely overwhelms any spatial correlations, which tend to slow down the percolation.

Finally, we have neglected the optical depths within the collapsed object, as well as in the gas in the Strömgren sphere surrounding each object. We assume both that the covering factor of other halos within the sphere is small, and that there is a sharp transition between the ionized and neutral regions. This approximation is somewhat alleviated by assuming a clumping factor for the gas (see § 2.4). With our clumping factor, our model is self-consistent in the volume average, accounting for all photoionizations and recombinations in the ionized region.

2.4. Nonuniformity of Cosmic Gas

The importance of clumping in determining the ionization history of the universe was stressed by GO97. In this section, we use Press-Schechter to estimate the clumping factor, dividing the baryons into clumped (virialized) and diffuse components. A similar procedure was described in Valageas & Silk (1999), who added a term to account for Lyα forest clouds (see their eq. [25]).

We assume that the clumped component consists of baryons that have collapsed into virialized halos. Using equation (26), the fraction of the baryonic mass of the universe that has collapsed, \( df_m \), in the time interval \( \tilde{t} \) to \( \tilde{t} + d\tilde{t} \),

![Figure 3](image-url)

Fig. 3.—Results of reionization calculation for flat Gaussian models, \( G + \Lambda \text{CDM} \): \( f_0 \) is the filling factor; \( \langle J_{21} \rangle \) is the global average of the ionizing intensity \( J_{21} \); \( \langle 1 - x \rangle \) is the neutral fraction (thick line, global average; thin line, ionized region); and \( \langle T \rangle \) is the temperature.
and survives without being incorporated into a larger object to time \( t \), is given by
\[
d f_m(t) = \int dR \frac{D(t)}{D(t)} N_{form}(R, i) \frac{4\pi R^3}{3} \psi(R, i) \frac{d\tilde{t}}{\tilde{t}},
\]
(40)
where \( \psi(R, i) \) is the selection function that filters out objects below the Jeans scale. The fraction of baryonic mass in collapsed objects at time \( t \) is given by integrating over \( \tilde{t} \):
\[
f_m(t) = \int_{\tilde{t}^c}^{\tilde{t}} \frac{d f_m(\tilde{t})}{d\tilde{t}} d\tilde{t}.
\]
(41)
Note that we do not take the collapsed fraction \( f_m(t) \) to be equal to the fraction of mass above the Jeans mass at time \( t \). This is because during reheating, the Jeans mass increases, but objects below the Jeans mass that formed at earlier times (when the Jeans mass was lower) can still survive, as long as they are not merged into larger objects. If all of these objects are in virialized halos, then their overdensity relative to the mean density will be
\[
\rho_{\text{vir}} = \frac{\Delta_{\text{vir}}}{\Omega(t)},
\]
(42)
where \( \Delta_{\text{vir}} \) is the overdensity relative to the critical density. Thus, the fraction of the volume in the universe that they occupy will be
\[
f_v(t) = \frac{f_m}{\delta_{\text{vir}}},
\]
(43)
Self-consistency requires that the diffuse component be underdense, so that \( \rho_{\text{diff}} < \rho \). Taking the volume average gives
\[
\rho_{\text{diff}} = f_v \rho_{\text{vir}} + (1 - f_v) \rho_{\text{diff}},
\]
(44)
which implies that
\[
\frac{\rho_{\text{diff}}}{\rho} = \frac{1 - f_m}{1 - f_v}.
\]
(45)
The volume-averaged square density will be
\[
\langle \rho^2 \rangle = f_v \rho_{\text{vir}}^2 + (1 - f_v) \rho_{\text{diff}}^2
\]
\[
= \frac{1}{\rho^2} \left[ f_m \delta_{\text{vir}}^2 + \frac{(1 - f_m)^2}{1 - f_v} \right].
\]
(47)
The average clumping factor, \( q \), is thus
\[
q \equiv \frac{\langle \rho^2 \rangle}{\rho^2}
\]
\[
= f_m \delta_{\text{vir}}^2 + \frac{(1 - f_m)^2}{1 - f_v}.
\]
(49)
This is the clumping factor we use in our model. As shown in Figure 1, the results of this semianalytic calculation are in fair agreement with the results of the numerical simulation given in GO97.

We keep track of the ionization and temperature history of the purely diffuse phase of the ionized region. We do this...
in order to properly calculate the Gunn-Peterson optical depth, which probes the diffuse IGM. This phase is by definition unclumped, and is underdense relative to the mean (eq. [45]). Note, however, that this diffuse phase plays no feedback role; it simply responds to the other physics going on.

Our model assumes that UV sources are uniformly distributed throughout collapsed halos and that all halos carry the same overdensity. These two assumptions will tend to reduce clumping, so equation (49) probably gives a lower bound to the clumping factor. However, as noted in the semianalytic model of Miralda-Escudé, Haehnelt, & Rees (1999) as well as in the numerical simulation of Gnedin (2000), the recombination rate may be reduced in highly clumped regions if such regions are shielded from outside radiation. The simulation of Gnedin (2000) shows that only those clumped regions far from UV sources will have this property, since clumped regions with UV sources will be ionized by the sources they contain. A comparison of the Miralda-Escudé et al. (1999) results with our model indicates that their effective clumping factor (that is, the ratio of the global recombination rate to the rate for a homogeneous universe) as normalized to observations is a factor of 2–3 smaller than the clumping factor we calculate around the epoch of reionization. On the other hand, as shown in Figure 1, the results of GO97 are somewhat higher than our model. These comparisons only apply to Gaussian cosmologies, since Miralda-Escudé et al. (1999) and GO97 only test those cases, and the former depends strongly on the PDF. Miralda-Escudé et al. (1999) note, however, that the late reionization of highly clumped gas does not necessarily increase the number of photons required to complete reionization. This was validated in the recent simulation of Gnedin (2000), in which the “overlap” phase of reionization, which is our primary concern in this paper, occurs in a relatively short period of time (10% of a Hubble time), and precedes the slow “postoverlap” stage better described by the model of Miralda-Escudé et al. (1999).

3. RESULTS

In this section we present the results of our calculations for a range of cosmological models. We begin with a “calibration” model, which uses the same cosmological parameters as were used by GO97 ($\Omega_0 = 0.35$, $\Omega_\Lambda = 0.65$, $h = 0.70$, and $\Omega_b = 0.03$). We then consider adiabatic Gaussian CDM ($G + \Lambda$CDM), textures ($T + \Lambda$CDM), and non-Gaussian isocurvature CDM (ICDM) as proposed by Peebles (1999a, 1999b). To help assess the sensitivity of the results to the PDF and the power spectrum, we also consider an “artificial” case of a texture PDF with an adiabatic CDM power spectrum ($T + \Lambda$CDM). We use a Hubble parameter $h = 0.65$, a baryon density parameter $\Omega_b h^2 = 0.0125$ throughout, and consider flat ($\Lambda$-dominated) and open models with matter density parameters of $\Omega_0 = 1$, 0.45, 0.35, and 0.25. All models are normalized to cluster abundances (using the method of Chiu et al. 1998 for all models except ICDM, which uses Peebles 1999a, 1999b). The cosmological parameters associated with the models we consider are listed in the first three columns of Table 1.

![Fig. 5.—Results of reionization calculation for flat texture models with adiabatic CDM power spectrum, $T + \Lambda$CDM](image_url)
As noted above, the UV efficiency, $\varepsilon_{\text{eff}}$, in each run is adjusted to reproduce $J_{21} = 1$ at $z = 4$. There are no other adjustable parameters in the model.

### Table 1

**Cosmological Models Considered and Calculated Results**

| Cosmological Parameters | Reionization: $z$ for $\langle 1 - x \rangle$ of | CMBR | IGM (at $z = 3$) |
|-------------------------|-----------------------------------------------|------|-----------------|
|                         | $10^{-1}$                                     | $10^{-3}$ | $\tau_{\text{rec}}$ | $\tau_{\text{GP}}$ | $f_*$ |
| Gaussian Adiabatic CDM  |                                              |        |                 |                 |           |
| $\Omega_b$ | $0.1$ | $0.0$ | $0.60$ | late | $13.1$ | $10.3$ | $1.4$ | $0.069$ | $0.14$ | $0.047$ |
| $\Omega_{\Lambda}$ | $0.45$ | $0.55$ | $0.71$ | late | $10.6$ | $7.6$ | $1.3$ | $0.065$ | $0.17$ | $0.027$ |
| $\sigma_8$ | $0.35$ | $0.65$ | $0.76$ | late | $10.2$ | $7.3$ | $1.4$ | $0.068$ | $0.17$ | $0.025$ |
| $\Omega_b$ | $0.25$ | $0.75$ | $0.85$ | late | $9.8$ | $7.2$ | $1.6$ | $0.077$ | $0.19$ | $0.025$ |
| $\Omega_{\Lambda}$ | $0.45$ | $0.55$ | $0.68$ | late | $11.9$ | $9.7$ | $1.8$ | $0.087$ | $0.17$ | $0.045$ |
| $\sigma_8$ | $0.35$ | $0.65$ | $0.71$ | late | $11.8$ | $9.8$ | $2.1$ | $0.099$ | $0.17$ | $0.049$ |
| $\Omega_b$ | $0.25$ | $0.75$ | $0.76$ | late | $11.7$ | $10.0$ | $2.6$ | $0.12$ | $0.21$ | $0.058$ |

Texture + CDM

| $\Omega_b$ | $0.1$ | $0.0$ | $0.37$ | early | $17.2$ | $14.4$ | $5.8$ | $0.28$ | $0.08$ | $0.35$ |
| $\Omega_{\Lambda}$ | $0.45$ | $0.55$ | $0.50$ | early | $12.8$ | $10.6$ | $6.5$ | $0.32$ | $0.13$ | $0.37$ |
| $\sigma_8$ | $0.35$ | $0.65$ | $0.56$ | early | $12.0$ | $10.0$ | $6.6$ | $0.35$ | $0.14$ | $0.38$ |
| $\Omega_b$ | $0.25$ | $0.75$ | $0.66$ | early | $11.0$ | $9.1$ | $3.9$ | $0.38$ | $0.16$ | $0.42$ |
| $\Omega_{\Lambda}$ | $0.45$ | $0.55$ | $0.47$ | early | $13.6$ | $11.3$ | $5.9$ | $0.33$ | $0.11$ | $0.42$ |
| $\sigma_8$ | $0.35$ | $0.65$ | $0.50$ | early | $12.7$ | $10.6$ | $4.6$ | $0.43$ | $0.11$ | $0.45$ |
| $\Omega_b$ | $0.25$ | $0.75$ | $0.66$ | early | $12.3$ | $10.1$ | $1.8$ | $0.46$ | $0.12$ | $0.48$ |

Texture + Adiabatic CDM Power Spectrum

| $\Omega_b$ | $0.1$ | $0.0$ | $0.35$ | early | $13.6$ | $10.8$ | $1.0$ | $0.065$ | $0.15$ | $0.021$ |
| $\Omega_{\Lambda}$ | $0.45$ | $0.55$ | $0.47$ | early | $11.8$ | $9.7$ | $1.2$ | $0.073$ | $0.18$ | $0.016$ |
| $\sigma_8$ | $0.35$ | $0.65$ | $0.53$ | late | $11.6$ | $9.6$ | $1.3$ | $0.074$ | $0.19$ | $0.016$ |
| $\Omega_b$ | $0.25$ | $0.75$ | $0.62$ | late | $11.2$ | $9.5$ | $1.7$ | $0.097$ | $0.22$ | $0.018$ |

Non-Gaussian Isocurvature CDM

| $\Omega_b$ | $0.20$ | $0.80$ | $0.90$ | early | $18.9$ | $16.5$ | $11$ | $0.53$ | $0.44$ | $0.20$ |

#### 3.1. Comparison with GO97 Simulation

In Figure 2 we show the filling factor, $f_{\text{II}}$, the ionizing intensity, $J_{21}$, the neutral fraction, $\langle 1 - x \rangle$, and the gas temperature, $T$, in the case where we mimic the GO97 calculation. While the ionizing intensity was adjusted to be consistent with the GO97 value at $z = 4$, the evolution of $J_{21}$ is similar in the two calculations. As in the GO97 calculation, reionization to $1 - x = 10^{-3}$ occurs suddenly at $z \sim 10$, while reheating occurs more gradually. The main difference between this work and GO97 is that reionization and reheating begin somewhat earlier in our model. This may in part be due to the limited resolution and single-phase treatment of GO97.

In Figure 1, we show the stellar fraction and clumping factor for this reionization calculation. Again, the values are similar to those in GO97. The clumping factor shows the greatest discrepancy, but the impact of this on the epoch of reionization appears to be minimal. The earlier onset of star formation shown in the top panel by our semianalytic model may reflect the limited resolution of GO97, which while adequate for average objects is insufficient for early, dense, small ones.

#### 3.2. Reheating and Reionization in Various Cosmologies

On the whole, the results are quite similar. It is thus encouraging that two completely different methods of calculating the reheating and reionization of the universe give comparable results.

The results of our calculations for the models we consider are presented in Table 1 and Figures 3–8. The general picture of the evolution in these models is as follows. Qualitatively, the evolution of reheating and reionization can be divided into two broad categories, characterized by late or early reheating, corresponding to late or early evolution of the filling factor, $f_{\text{II}}$. In models with late reheating, which include the Gaussian models and the $T + \Lambda CDM$ model, the filling factor rises from $f_{\text{II}} < 10^{-4}$ at $z = 30$–50 to $f_{\text{II}} \sim 1$ at $z = 10$–20. For models with early reheating, which include the pure texture $T + CDM$ models and the ICDM model, the filling factor is already $f_{\text{II}} > 0.1$ at $z \sim 80$, but rises slowly to $f_{\text{II}} = 1$ at $z = 10$–20. A caveat to these results (as well as that of GO97) is that they are based on a similar framework for translating gravitational collapse into UV photons. Major changes to this framework would undoubtedly change the results.

For an adiabatic CDM-type power spectrum, the fluctuation spectrum, $\sigma_R$, rises logarithmically, or nearly logarithmically, as $R \to 0$. Hence, at the Jeans scale of $\sim 0.01 h^{-1}$
Mpc, there is a logarithmic cutoff in power for $z > \sigma_R \sim 10^{-20}$. Furthermore, cooling becomes less efficient at the Jeans scale at later times. These two factors lead to a suppression of reheating and reionization until late times. Full reionization occurs after the filling factor approaches unity, at redshifts $z \lesssim 10$. Note that by "reionization" we mean $f_{HI}$ approaching unity, but by "full reionization" we mean the epoch at which the individual ionized regions start overlapping substantially, so that $\lambda = -\ln(1 - f_{HI}) \gg 1$.

These results for Gaussian CDM models are consistent with the work of GO97, Ciardi et al. (1999), and Haiman & Loeb (1998), but vary somewhat from the results of Valageas & Silk (1999). The latter find a slightly lower value for the redshift of reionization ($z \sim 6-7$). This difference is not likely to be due to gas clumping, as proposed by Ciardi et al. (1999), since both this work and GO97 incorporate clumping. In fact, our work uses a prescription for clumping very similar to that of Valageas & Silk (1999). We do, however, use a different prescription for calculating the radius of the cosmological Strömgren spheres, but the rapidity with which reionization occurs may mitigate the effect of this differences. While the exact value of the epoch of reionization remains uncertain, various methods, including ours, give the range $z = 8 \pm 2$ for Gaussian CDM models.

For T + CDM and ICDM power spectra, the fluctuation spectrum rises as a power law for small $R$, and the Jeans scale becoming nonlinear with $\delta M/M \sim 1$ at $z \sim 50$. At this redshift, cooling at the Jeans scale is still efficient, and UV photons are created in abundance. At $z \sim 100$, the Jeans scale is already mildly nonlinear, with $\delta M/M \sim 0.5$, causing the universe to already be substantially reheated and reionized ($f_{HI} > 10\%$) at $z \sim 80$. But because of the early formation of nonlinear structures, the clumping factor is higher, and hence the recombination rate is greater. Full reionization is thus delayed until $z = 10-20$. Detailed calculations for reheating and reionization have not been done for these models, so comparisons with other work cannot be made.

4. DISCUSSION

In this section, we begin by discussing several possible observational tests to constrain different cosmological models. We then discuss the nonlinearity of collapsing baryons in relation to the idea that reheating can be regulated through the Jeans criterion. Finally, we discuss the issue of physical bias—the bias of luminous objects—by considering bias at the Jeans mass limit. For this, we make use of a generalized form of statistical bias based on the peak-background split.

4.1. Observational Constraints on Cosmological Models

4.1.1. CMBR: The Compton $y$ and Optical Depth to Recombination

The $y$-parameter measures the cosmic microwave background radiation (CMBR) spectral distortion away from a blackbody spectrum caused by Compton-Thomson scattering. It is defined by

$$\dot{y} = \frac{k_B(T_e - T)}{m_e c^2} \sigma_T n_e c , \tag{50}$$
where $T_e$ is the electron temperature, $T$ is the CMBR temperature, and the other symbols have their usual meanings. When structure formation begins in earnest and the gas temperature rises, low-energy CMBR photons will be scattered to higher energies.

The $y$ we calculate in our models using equation (50) are shown in Table 1. Again, the $T + CDM$ and ICDM models behave similarly, with $y = 2 - 11 \times 10^{-7}$, and the Gaussian and $T + ACDM$ models have $y = 1 - 3 \times 10^{-7}$. Thus, they differ by a factor of $\sim 3$. The COBE satellite (Fixsen et al. 1996) put a bound on the spectral distortion, $y \leq 1.5 \times 10^{-5}$. (51)

Unfortunately, the prospects for future experimental bounds on $y$ that are significantly better than equation (51) are not good (D. Wilkinson 1998, private communication).

A potentially stronger constraint on reionization comes from the CMBR anisotropy spectrum. The optical depth to recombination is given by

$$\tau_{rec} = \int n_e c \sigma_T dt.$$ (52)

Because we stop our calculation at $z = 3$, we can only give a lower bound to the recombination optical depth. On small angular scales, reionization results in the damping of the temperature anisotropy spectrum by a factor of $e^{-\tau_{rec}}$. However, reionization can lead to a significant polarization anisotropy.

In the models we consider, those with an adiabatic CDM power spectrum give a small optical depth of $\tau_{rec} \lesssim 0.1$, while the other models have $\tau_{rec} \gtrsim 0.3$ (see Table 1). According to Zaldarriaga (1998), a negative detection for polarization anisotropy on the $1^\circ$ scale with sensitivity $\sim 1 \mu K$ would be able to rule out these moderate optical depth models. The upcoming MAP and Planck satellites may be able to distinguish between the low- and high-reionization optical depths. As long as polarization foregrounds do not limit the measurements, the MAP and Planck satellites should be able to detect reionization optical depths of $\sim 0.02$ and 0.004, respectively (D. Spergel 1999, private communication).

4.1.2. Constraints on the IGM: The Gunn-Peterson Optical Depth and the Baryon Fraction in Stars

The Gunn-Peterson optical depth as a function of the scale factor $a$ is given by (Peebles 1993)

$$\tau_{GP} = \frac{3\Lambda_{21} \lambda_{21} n_{HI}(a)c}{4H_0 a\sqrt{\Omega_\Lambda a^{-3} + \Omega_\Lambda + (1 - \Omega_0 - \Omega_\Lambda)a^{-2}}}.$$ (53)

where $\Lambda_{21} = 6.25 \times 10^8$ s$^{-1}$ is the Ly$\alpha$ decay rate, and $\lambda_{21} = 1216$ Å is the wavelength. The Gunn-Peterson optical depth at $z = 3$ is shown in Table 1. The results for all models are near the upper limits of $0.1 - 0.35$ at $z = 4$ established by observations (Jenkins & Ostriker 1991; Giallongo et al. 1994). Because we have neglected the effect of large-scale shock heating (which would increase the temperature)
as well as metals and other coolants (which would increase the star formation rate), it is plausible that our final neutral fractions are overestimated, leading to an overestimate of $\tau_{\text{GP}}$.

A stronger discriminant between models may be their predicted star formation histories. Figure 9 shows the baryon fraction in stars/quasars predicted by the various models we have considered. Clearly, the early-reheating models require that a significantly greater fraction of the baryons be tied up in stars/quasars or their remnants. The observational limit that we consider is from the analysis by Rauch et al. (1997). Their work indicates that Ly$\alpha$ clouds account for the majority of the baryons in the universe, and implies a strong upper bound of 0.29 on the fraction of baryons in the universe that can have been converted into stars/quasars by a redshift $z < 3$. The models with adiabatic power spectra are well within this constraint, with stellar fractions $f_\star < 0.06$. The T + ACDM models gives lower values than the Gaussian models because of the lower value of cluster normalization required. The ICDM model is within the limit of $f_\star = 0.2$. The pure texture T + CDM gives stellar fractions that are at or above the limit set by Rauch et al. (1997). While we have used $\Omega_b h^2 = 0.0125$ in these models rather than the value $\Omega_b h^2 = 0.024$ given in Rauch et al. (1997), we expect the fraction of baryons converted to stars/quasars to be higher for models with higher $\Omega_b$. In particular, because the cooling time is inversely proportional to the baryon density, to lowest order, the fraction $f_\star \sim t_{\text{cool}}^{-1} \sim \Omega_b$. Using this scaling behavior, the higher value of $\Omega_b h^2 = 0.024$ yields $f_\star \gtrsim 0.38$ for all texture and ICDM models, well above the observational limit.

4.2. Nonlinearity of Collapsing Baryons

The work of Ostriker & Gnedin (1996) suggested that reheating and reionization may be regulated by the Jeans criterion. In particular, Ostriker & Gnedin (1996) stated that the Jeans mass and the nonlinear mass tracked each other during reheating, and suggested a feedback mechanism to explain this feature. In our model, however, because there are two distinct phases, the “average” Jeans mass is not the most physically meaningful concept. However, we can still consider the degree to which collapsing baryons—those halos that satisfy the Jeans criterion in their respective phases—are nonlinear. We can therefore explore whether the possibility of a Jeans-feedback mechanism may be generic to all the cosmologies we consider.

In Figure 10, we show the value of the degree of nonlinearity, $\nu$, in the ionized and neutral phases, as well as the value of $\nu$, volume-averaged between the two phases. As defined in equation (16), the degree of nonlinearity essentially characterizes “how many $\sigma$ fluctuations” are necessary for collapse, and is related to the fraction of the universe that is collapsing. We show the average degree of nonlinearity, $\nu$, to show the transition that occurs during
reheating. This will approximate, to lowest order, the average \( v \) of observed objects. The early evolution of \( v \) is determined primarily by the power spectrum, and is only weakly dependent on the PDF. This is, of course, related to the fact that the Jeans scale depends primarily on the thermal evolution of the universe, which we found in the previous section to be determined mainly by the power spectrum on small scales.

At epochs \( z \lesssim 30 \), all the models look quite similar to those found by Ostriker & Gnedin (1996): during reheating, the average degree of nonlinearity remains about constant. This may be because of feedback between star formation and the temperature of the universe once reheating has begun, keeping the volume-averaged nonlinear scale of \( v = 2-3.5 \). A possible physical basis for this feedback is as follows. As structure collapses, star formation accelerates, increasing the ionized fraction of the universe. However, because the temperature rises in these regions, the Jeans mass in these regions becomes higher, baryonic collapse slows, and further star formation there is quenched. The result is a slowing of the global star formation rate. However, this feedback stops when ionized regions have overlapped substantially, since there is no more additional neutral gas to ionize.
4.3. Physical Bias from Statistical Bias: Objects at the Jeans Scale

As far as we know, a necessary condition for matter to be luminous is that the baryons collapse to high density. Thus, the first-order criterion we use to explore physical bias is that objects have a mass greater than or equal to the Jeans mass. We begin, however, by reviewing statistical bias in the peak-background split formalism and its extension to non-Gaussian PDFs.

4.3.1. Generalized Statistical Bias

The peak-background split was first suggested by Kaiser (1984) and recently revisited in the case of non-Gaussian models by Robinson, Gawiser, & Silk (1998). In this model, the density contrast field, $\delta(x)$, is split into a short-wavelength “signal,” $\delta_s$, and a long-wavelength “background,” $\delta_b$. The split is made in such a manner that $\delta_s$ and $\delta_b$ are largely uncorrelated, $\langle \delta_s \delta_b \rangle \approx 0$. The wavelengths of $\delta_s$ corresponds to the scales of the objects whose clustering properties are of interest, and the wavelengths of $\delta_b$ corresponds to the scales at which the clustering properties (and bias) are to be measured. In this prescription, the local threshold for collapse, $\delta_c$, is modified locally to $\delta_c - \delta_b(x)$, and the (observed) Eulerian density is modified from the Lagrangian one by a factor $1 + \delta_b(x)$. Let the average comoving number density of objects of scale $R$ to
R + dR be given by \( \bar{N}_R(\delta_c) \, dR \), where we have explicitly shown the dependence on the critical threshold of collapse \( \delta_c \). Then the local comoving number density per unit scale \( dR \) is given by

\[
N_R(x) = [1 + \delta_b(x)] \bar{N}_R(\delta_c - \delta_b(x))
\]

\[
= [1 + \delta_b(x)] \bar{N}_R(\delta_c) - \delta_b(x) \frac{\partial \bar{N}_R(\delta_c)}{\partial \delta_c} + O[\delta_b^2(x)],
\]

where in the second line, we have used a Taylor expansion to first order. The long-wavelength “bias at birth,” \( b_{*}(R) \), can be defined by the ratio of the correlation functions at the time when the objects of scale \( R \) are identified,

\[
b^2_{*}(R) \equiv \frac{\xi(R)}{\xi_c},
\]

where we have defined

\[1 + \xi(R) \equiv \frac{\langle N_R(x) N_R(0) \rangle}{\bar{N}_R^2},\]

and \( \xi \) is the usual mass correlation function. The bias at birth is thus given by

\[b_{*}(R) = 1 - \frac{1}{\bar{N}_R} \frac{\partial \bar{N}_R}{\partial \delta_c}.
\]

The value of \( \bar{N}_R \) is given from Press-Schechter by equations (15)-(17). Differentiating with respect to the critical threshold \( \delta_c \) and rearranging gives the following general form for \( b_{*}(R) \):

\[b_{*}(R) = 1 + \frac{1}{\delta_c} \left[ \left( \frac{-P'}{P} \right) v(R, t) - 1 \right],
\]

where \( P' \equiv dP/dv \). Examples of the ratio \( -P'/P \) are given in equations (22)-(22). In the cases we consider, the bias at birth will be given by

\[
b_{*}(R) = 1 + \frac{1}{\delta_c} \begin{cases}
0.67 v^2 - 1 & \text{Gaussian PDF}, \\
1.45 v - 1 & \text{texture PDF}, \\
0.67 v - 1 & \text{non-Gaussian ICDM PDF}.
\end{cases}
\]

The Gaussian case has been derived previously by Mo & White (1996) and others. The non-Gaussian cases are similar to the derivation for \( v \gg 1 \) by Kaiser (1984). It is important to note that for small \( v \), an exponential PDF such as that from textures or from the ICDM model gives a (slightly) larger bias than a Gaussian PDF; but for extremely rare events with \( v \gg 1 \), a Gaussian PDF leads to a substantially larger bias than exponential PDFs. This was previously noted by Gooding et al. (1992).

A similar calculation yields the bias of objects of a scale \( \geq R \) to be

\[
b^*_{*}(\geq R) = \frac{\int_{R}^{\infty} dR \bar{N}_R b_{*}(R)}{\int_{R}^{\infty} dR \bar{N}_R}.
\]

As noted by Mo & White (1996), the total bias is simply the average of the individual biases, weighted by the mean number density of objects at each scale. An equivalent form for \( b^*_{*}(\geq R) \) was derived by Robinson et al. (1998).

### 4.3.2. Bias of Objects at the Jeans Scale

We now consider this bias at birth in the cosmological models we have analyzed. This is equivalent to considering models that are dominated by constant merging at the epochs we consider, so that objects do not “survive” long enough for their clustering bias to dynamically decrease.

The Jeans birth bias, \( b_{*,*} \), shown in Figure 11, is dependent on the PDF and the degree of nonlinearity at the Jeans scale, \( v_j \). At high redshift, well before reionization, the Gaussian models exhibit the strongest bias, with \( b_{*,*} \geq 10 \) at redshifts \( z > 30 - 40 \). After a pause during reheating, during which \( b_{*,*} \) remains in the range of 5-7, the bias drops quickly after full reionization to a level of \( b_{*,*} = 1.5 - 2 \) at \( z = 3 \). The \( T + \Lambda \) CDM model has an intermediate bias, falling to \( b_{*,*} \sim 10 \) at redshift \( z \sim 80 \), pausing during reheating at \( b_{*,*} \sim 3 \), and then falling to \( b_{*,*} = 1.5 - 2 \) at \( z = 3 \). The pure texture \( T + CDM \) and ICDM models have a low bias, which always remains \( b_{*,*} \leq 3 \), with the \( T + CDM \) bias declining slowly to from 2.5 to 1.5 after reionization and the ICDM model bias remaining near 1.

In contrast to many of the observational constraints considered previously, the evolution of \( b_{*,*} \) depends on both the PDF and the fluctuation spectrum. While the value of \( v_j \) depends primarily on the power spectrum, the relationship between \( v_j \) and \( b_{*,*} \) is quite different for different PDFs.

### 4.3.3. Comparison with Other Bias Calculations

Many numerical simulation–based studies of large-scale (\( \sim 8 \) h⁻¹ Mpc) galaxy bias have been done recently, although all of them concentrate on Gaussian adiabatic CDM scenarios. Each of these studies uses a different prescription to identify “galaxies.” For instance, Katz, Hernquist, & Weinberg (1999), who identify galaxies with dense, gravitationally bound star-forming regions, determine (for \( \Lambda \)-dominated models) that bias decreases from \( b \sim 3.5 \) at \( z = 5 \) to \( b \sim 2 \) at \( z = 3 \). This is close to or slightly above the Jeans birth bias (\( b \sim 3 \) and 2, respectively) in our model for comparable cosmologies. On the other hand, Kauffmann et al. (1999) combine numerical simulation with semianalytic galaxy formation, and find for \( L_\ast \) galaxies that bias falls from \( b \sim 4.5 \) to \( b \sim 2.5 \) in the same redshift range. This is simply due to the fact that \( L_\ast \) galaxies are rarer than “typical” galaxies at these redshifts. Finally, Blanton et al. (2000) investigate the bias of “recently formed” galaxies, finding that for these objects, \( b \sim 4.5 \) at \( z = 3 \). Again, the higher bias is due to the fact that the Jeans mass is significantly less than the masses of the galaxies under consideration.

### 5. SUMMARY AND CONCLUSIONS

We have developed a semianalytic model for the reheating and reionization of the universe. Our primary assumption is that UV photons from early generations of star formation provide the energy to heat and ionize the primeval gas. We have explored the thermal and ionization evolution of the universe in a variety of cosmological models.

Interestingly, we have found that many of the results found in a set of specific hydrodynamic simulations of the \( \Lambda \) CDM model (G097) appear to be generic: (1) reheating and reionization are distinct processes, and clumping is important for understanding both of them; (2) full reionization is a sudden process of percolation and occurs near
redshift $z = 12 \pm 5$; (3) the pace of reheating appears to be regulated by feedback through the Jeans criterion; and (4) hierarchical clustering models are broadly consistent with the measured Gunn-Peterson opacity of the IGM. However, models do appear to have significant differences, between which observations in the near future might be able to distinguish.

Our major conclusions are as follows:

1. The most important determinant of the thermal history of the universe for $z \leq 100$ is the power spectrum at the Jeans scale. The shape of the PDF is only of secondary importance to the overall thermal history. For both Gaussian and texture PDFs, models with adiabatic CDM-type power spectra, which have a logarithmically diverging fluctuation spectrum $\sigma(R)$ at small scales, lead to reheating beginning at $z \sim 30$. These we refer to as “late-reheating models.” For models with a power-law divergence at small scales, such as textures or ICDM, reheating has already begun at $z \sim 80$. These we refer to as “early-reheating models.”

2. In all models, full reionization (overlap of ionized regions) occurs as $z \sim 10$. In models with late reheating, reionization proceeds after reheating has plateaued. In the early reheating models, the increased clumping of gas caused by earlier structure formation delays full reionization until $z \sim 10$. This conclusion is subject to modification,
however, if there are major deficiencies in our prescription for star formation, which remains a poorly understood process. This limitation is ameliorated by the normalization of our models to require that it reproduce the observed UV intensity $J_{21} \approx 1$ at $z = 4$.

3. With regards to observational constraints from the CMBR, the polarization anisotropy caused by Compton scattering off free electrons will be a strong discriminant between models. The late- and early-reheating models give different Compton optical depths to recombination, which may be testable with experiments with sensitivity of $\sim 1 \mu$K.

4. With regards to observational constraints from the IGM, a strong test is the baryon fraction in stars/quasars, which appears to disfavor the early-reheating models. Observations suggest that this fraction is $f_\text{b} < 0.29$ for $\Omega_b h^2 = 0.024$. Scaling our results for early-reheating models of $f_\text{b} = 0.20$ for ICDM and $f_\text{b} = 0.35$ for texture (obtained with $\Omega_b h^2 = 0.0125$) gives this fraction as, respectively, $f_\text{b} \sim 0.38$ and $f_\text{b} \sim 0.67$. Late-reheating models, on the other hand, are well within the observational limit.

5. The degree of nonlinearity on the Jeans scale averaged between the neutral and ionized phases, $\langle v_J \rangle \sim \langle 1/\sigma(M_d) \rangle$, decreases before and after reheating, but remains about constant during reheating. A possible explanation for this involves a feedback mechanism that regulates the pace of reheating through the Jeans criterion.

6. Since $v_J$ can be related to galaxy bias, we find that the average bias at the Jeans scale drops before and after reheating, but remains roughly constant during reheating. Gaussian models generically give much higher biases than the texture and non-Gaussian ICDM models, and is a potential discriminant between models.

W. A. C. and J. P. O. thank David Spergel and David Wilkinson for helpful comments, Nick Gnedin for providing results from his simulations, and an anonymous referee who helped improve the focus and presentation of our results. W. A. C. thanks David Spergel, David Wilkinson, and Peter Meyers, who were all on his dissertation defense committee. This work was supported in part by NSF grant AST 93-18185.

### APPENDIX A

#### THERMAL HISTORY OF NEUTRAL AND IONIZED PHASES

The temperature of the neutral (cold) phase will be determined primarily by adiabatic expansion and Compton scattering of electrons off the CMBR. We take the electron fraction of the neutral phase, $x_\text{e}$, to be equal to the residual electron fraction, $x_0$, from Peebles (1993), of

$$x_\text{e} = x_0 \equiv \frac{1.2 \times 10^{-5} \Omega_0^{1/2}}{h \Omega_b},$$  \hspace{1cm} (A1)

where $\Omega_0$ is the present-day value of the total mass density parameter, $h$ is the Hubble parameter, and $\Omega_b$ is the present-day baryon density parameter. We define the gas internal energy and particle number density as

$$u \equiv \frac{3}{2} n k_B T,$$  \hspace{1cm} (A2)

$$n = \frac{n_\text{b}}{\mu} \equiv \frac{n_\text{H}}{\mu X},$$  \hspace{1cm} (A3)

where $k_B$ is Boltzmann’s constant, $T$ is the temperature, $n$ is the total particle density, $n_\text{b}$ is the baryon density, $\mu^{-1}$ is the particles per baryon, and $X$ is the hydrogen mass fraction. We have assumed an adiabatic index $\gamma_a = 5/3$.

For simplicity, we consider only cooling from hydrogen and electrons, using the compilation of Anninos et al. (1998). In erg cm$^{-3}$ s$^{-1}$, these rates, followed by their reference and formulae (if not too complicated) are:

*Compton cooling (heating).*—Peebles (1993): $\Gamma_{cc} n_e = 5.65 \times 10^{-36} a^{-4} [T - (2.73/a)] n_e$.

*Collisional excitation cooling.*—Black (1981), Cen (1992): $\Gamma_{ce} n_e = 7.50 \times 10^{-19} (1 + T^{3/4})^{-1} \exp (-118348/T) n_e n_{\text{H}0}$

*Collisional ionization cooling.*—Shapiro & Kang (1987), Cen (1992): $\Gamma_{ic} n_e = 2.18 \times 10^{-11} k_{cc} n_e n_{\text{H0}}$

*Recombination cooling (case A).*—Black (1981), Spitzer (1978): $\Gamma_{cc} n_e = 8.70 \times 10^{-27} T^{1/2} T_\text{c}^{-0.4} (1 + T_6^{0.7})^{-1} n_e n_{\text{H}1+}$

*Molecular hydrogen cooling.*—Formula from Lepp & Shull (1983): $\Gamma_{bc} n_e$ $\propto$ $n_e$

*Bremsstrahlung cooling.*—Black (1981), Katz, Weinberg, & Hernquist (1996): $\Gamma_{bc} n_e \equiv 1.43 \times 10^{-27} T^{1/2} [1.1 + 0.34 \exp \left[ - (5.5 - \log_{10} T)^3 / 3 \right]] n_e n_{\text{H}1+}$

Here $k_{cc}$ is the $H^0 + e^-$ collisional ionization rate coefficient taken from the Janev et al. (1987) compilation, and $T_n$ is the temperature in units of $10^n$ K.

We define the total cooling rate to be

$$\Gamma_{\text{tot}, e} \equiv \ell (\Gamma_{cc} + \Gamma_{ic} + \Gamma_{rc} + \Gamma_{H2, e} + \Gamma_{bc}) + \Gamma_{cc},$$  \hspace{1cm} (A4)

where the clumping factor $\ell$ is required for the particle-particle cooling rates, but not the Compton cooling rate. The
equations for thermal evolution in the H\textsc{i} and H\textsc{ii} regions are thus

\[ \dot{u}_I = -5 \frac{\dot{a}}{a} u_I - \Gamma_{\text{tot}, c} x_I n_H, \]  
(A5)

\[ \dot{u}_{\text{II}} = -5 \frac{\dot{a}}{a} u_{\text{II}} + \Gamma_{\text{ph}} (1 - x_{\text{II}}) n_H - \Gamma_{\text{tot}, c} x_{\text{II}} n_H, \]  
(A6)

where the photoheating rate, $\Gamma_{\text{ph}}$, is given by equation (14), and we have assumed that the electron density, $n_e$, is equal to the ionized hydrogen density, $n_{H^+}$. The equations for the temperature evolution are thus

\[ k_B T_I = -2 \frac{\dot{a}}{a} k_B T_I - \frac{2X}{3\mu} \Gamma_{\text{tot}, c} x_I, \]  
(A7)

\[ k_B T_{\text{II}} = -2 \frac{\dot{a}}{a} k_B T_{\text{II}} + \frac{2X}{3\mu} \left[ \Gamma_{\text{ph}} (1 - x_{\text{II}}) - \Gamma_{\text{tot}, c} x_{\text{II}} \right], \]  
(A8)

where $X$ is the hydrogen mass fraction.

These temperatures are used to determine the comoving Jeans length in each region:

\[ R_{J,I,\text{II}} = \frac{\pi}{k_{J,I,\text{II}}} = \sqrt{\frac{5\pi k_B T_{I,\text{II}}}{12\bar{\rho} m_H a^2}}, \]  
(A9)

where $k_j$ is the comoving Jeans wavenumber and $\bar{\rho}$ is the total mean density (including dark matter). The Jeans mass in baryons is thus

\[ M_{J,I,\text{II}} = \frac{4\pi}{3} \bar{\rho}_{0,b} R_{J,I,\text{II}}^3, \]  
(A10)

where $\bar{\rho}_{0,b}$ is the current mean density in baryons.

**APPENDIX B**

**MOLECULAR HYDROGEN FORMATION**

We follow the procedure described in Anninos et al. (1998) to solve for the evolution of molecular hydrogen. The reactions involving molecular hydrogen are given in Table 2. The formation of H\textsubscript{2} is primarily through the intermediaries H$^-$ and H\textsubscript{2}$^+$, which Anninos et al. (1998) note to be nearly in their (small) equilibrium abundances at all times. Neglecting the reaction

| Reaction Number | Reaction |
|-----------------|----------|
| 7               | H$^0 + e^- \rightarrow H^- + \gamma$ |
| 8               | H$^- + H^0 \rightarrow H_2 + e^-$ |
| 9               | H$^0 + H^+ \rightarrow H_2^+ + \gamma$ |
| 10              | H$_2^+ + H^0 \rightarrow H_2 + H^+$ |
| 11              | H$_2 + H^+ \rightarrow H_2^+ + H^0$ |
| 12              | H$_2 + e^- \rightarrow 2H^0 + e^-$ |
| 13              | H$_2 + H^0 \rightarrow 3H^0$ |
| 14              | H$^- + e^- \rightarrow H^0 + 2e^-$ |
| 15              | H$^- + H^0 \rightarrow 2H^0 + e^-$ |
| 16              | H$^- + H^+ \rightarrow 2H^0$ |
| 17              | H$^- + H^+ \rightarrow H_2^+ + e^-$ |
| 18              | H$_2^+ + e^- \rightarrow 2H^0$ |
| 19              | H$_2^+ + H^- \rightarrow H_2 + H^0$ |
| 23              | H$^- + \gamma \rightarrow H^0 + e^-$ |
| 24              | H$_2 + \gamma \rightarrow H_2^+ + e^-$ |
| 25              | H$_2^+ + \gamma \rightarrow H^0 + H^+$ |
| 26              | H$_2^+ + \gamma \rightarrow 2H^+ + e^-$ |
| 27              | H$_2 + \gamma \rightarrow H_2^+ + 2H^0$ |
| 28              | H$_2 + \gamma \rightarrow 2H^0$ |
$H_2^+ + H^- \rightarrow H_2 + H^0$, which is only a small-order correction, the equilibrium abundance of $H^-$, $x_{H^-} \equiv n_{H^-}/n_H$, can be written independently of $H_2^+$:

$$x_{H^-} = \frac{k_7 n_H x_e}{k_8 n_H + k_{14} n_e + k_{15} n_H + (k_{16} + k_{17}) n_{H^+} + k_{23}},$$

where the variables $k_i$ are the rate coefficients with subscripts referring to the reaction number in Table 2. Then, given $n_{H^-}$, the equilibrium abundance of $H_2^+$, $x_{H_2^+} \equiv n_{H_2^+}/n_H$ can be written with no additional assumptions as

$$x_{H_2^+} = \frac{k_9 n_H x_{H^-} + k_{11} n_{H^+} x_{H^-} + k_{17} n_{H^-} x_{H^+} + k_{24} x_{H_2}}{k_{10} n_H + k_{18} n_e + k_{19} n_{H^-} + k_{25} + k_{26}}.$$

The evolution equation for $H_2$ is thus

$$\dot{x}_{H_2} = C - D x_{H_2},$$

where the creation and destruction rates are

$$C \equiv k_8 n_{H^-} x_{H^0} + k_{10} n_{H_2^+} x_{H_2^0} + k_{19} n_{H^+} x_{H^-},$$

$$D \equiv k_{11} n_{H^+} + k_{12} n_e + k_{13} n_{H^0} + k_{24} + k_{25} + k_{26}.$$

### APPENDIX C

**GAS COOLING IN VIRALIZED HALOS**

In Press-Schechter theory, the dynamical time for halos collapsing at $z_{\text{col}}$ is the same regardless of its mass, since the virialized density of the halo depends only on the collapse time. Specifically, the dynamical time we use is

$$t_{\text{dyn}}(z_{\text{col}}) = \frac{3\pi}{32G\rho_{\text{vir}}(z_{\text{col}})},$$

$$\rho_{\text{vir}}(z_{\text{col}}) = \Delta_{\text{vir}}[\Omega(z_{\text{col}})]\rho_c(z_{\text{col}}),$$

where $\rho_c$ is the critical density, and the overdensity factor $\Delta_{\text{vir}}$ depends only on $\Omega$ at the time of collapse.

It is a much more involved process, however, to determine the cooling time, $t_{\text{cool}}$. Although the mass density is the same for all halos collapsing at a given epoch, different mass halos have different virial temperatures, as well as different temperature histories. Because the primordial chemistry of the expanding universe not in equilibrium, we must follow the thermal and ionization history of a halo in order to determine its cooling time at virialization.

Some previous semianalytic models (e.g., Cole, Fisher, & Weinberg 1994) have simply assumed a collisional equilibrium cooling function, $\Lambda_c(T) \equiv \Gamma_{\text{tot},e}/n_H$, which only depends on temperature. Tegmark et al. (1997) considered the minimum mass that could cool within a Hubble time. Their results, as well as those of Ostriker & Gnedin (1996) and GO97, showed that molecular hydrogen plays a major role in the cooling of early halos (as envisaged by Kashlinsky & Rees 1983 and Couchman & Rees 1986). The formation of $H_2$, however, is both nonequilibrium and density dependent. Thus, we calculate a cooling rate for each halo, assuming a uniform spherical collapse of gas that ends up at the virial density. Our cooling function depends on the temperature of the halo and the collapse redshift. We do make the simplifying assumption that the same cooling function, $\Lambda_c(T, z)$, applies for the entire halo.

To calculate $\Lambda_c(T, z)$, we extensively modified the nonequilibrium “T0D” ionization code of Abel (1998)\(^3\) so that it modeled a spherical collapse. This code incorporates all of the physics of Anninos et al. (1998) and Abel et al. (1997). We assume a uniform sphere of gas and dark matter, which collapses to one-half its radius of maximum expansion and which heats the gas to the virial temperature. Before turnaround, we assume the standard parametric form for spherical collapse, given by

$$r = A(1 - \cos \theta)$$

$$t = B(\theta - \sin \theta)$$

$$A^3 = GMB^2$$

$$t_a = \pi B$$

$$r_{\text{max}} = 2A,$$

\(^3\) T0D (Tom 0D code; T. Abel, 1998) is available at: http://zeus.ncsa.uiuc.edu/~abel/PGas/codes.html.
where $t_{\text{ta}}$ is the time at turnaround, $r_{\text{max}}$ is the maximum radius of expansion, and $M$ is the total mass of the halo. After turnaround, we adopt the following simple functional form for the radius as a function of time:

$$r = A \left[ 2 - \frac{\tau^2}{\tau^2 + 8 \text{sech}(\tau/2)} \right],$$  \hspace{1cm} (C8)

$$\tau \equiv \frac{t - t_{\text{ta}}}{B}. \hspace{1cm} (C9)$$

For the velocity dispersion, $\sigma^2$, of the gas, we assume that shock heating leads to a rise from its value at turnaround, $\sigma^2_{\text{ta}}$, to the virial velocity dispersion, $\sigma^2_{\text{vir}}$, via the equation

$$\sigma^2 = \sigma^2_{\text{ta}} + (\sigma^2_{\text{vir}} - \sigma^2_{\text{ta}}) \left[ 1 - \text{sech}(\tau^2/2) \right]. \hspace{1cm} (C10)$$

Plots of these functions are shown in Figure 12.

The calculated value of the ratio $t_{\text{dyn}}/t_{\text{cool}}$ for gas at the virial density is given in Figures 13 and 14 for the cases of $\Omega_0 = 1.0$ and $\Omega = 0.35$, respectively. Also shown is the comoving Jeans length for gas in the neutral phase. The narrow ridge in which the cooling time is longer than the dynamical time extending from $R \approx 0.02$, $z \approx 25$ to $R \approx 0.01$, $z \approx 100$ for the $\Omega_0 = 1$ case, and from $R \approx 0.04$, $z \approx 20$ to $R \approx 0.02$, $z \approx 100$ for $\Omega_0 = 0.35$, corresponds to the period in which $H_2$ has been destroyed, but atomic line cooling is not yet efficient. This "pause" in the cooling efficiency was first identified by Ostriker & Gnedin (1996). Note that $H_2$ cooling remains efficient later for $\Omega_0 = 0.35$ than for $\Omega_0 = 1$ (the tip of the banana-shaped region to the left of the ridge extends to lower redshift).

This calculation is essentially a simplified version of the Tegmark et al. (1997) calculation, but we use the information in the entire "cooling grid" in our work. This is particularly important at later times, since the cooling efficiency actually begins to decrease for large masses (this is why we have clusters of galaxies rather than gigantic $\sim 10^{15} M_\odot$ megagalaxies). These cooling grids can be precalculated for a given background cosmology, and are assumed to be the same for any PDF.
Even if the average cooling time of a halo is greater than or equal to the dynamical time, it is possible that the central, densest parts of a halo can still cool. To see why, consider a constant cooling function $\Lambda_\rho$. Because the cooling time is inversely proportional to the density, $t_{\text{cool}} \propto \rho^{-1}$, and the dynamical time, $t_{\text{dyn}} \propto \rho^{-1/2}$, the ratio of the dynamical to the cooling times is proportional to $\rho^{1/2}$:

$$\varphi \equiv \frac{t_{\text{dyn}}}{t_{\text{cool}}} \propto \rho^{1/2}.$$  \hfill (C11)

If we make the Ansatz that only regions cooling on a timescale shorter than the dynamical timescale, so that $\varphi \geq 1$, can undergo star formation, then there will be a maximum radius, possibly equal to the virial radius, outside of which star formation will not occur. For a singular isothermal sphere, the density is proportional to $r^{-2}$, and the mass within a radius $r$ is proportional to $r$. Therefore, the fraction of the baryonic mass of the halo undergoing star formation is simply proportional to $\varphi$:

$$\frac{M_\star[r(\varphi = 1)]}{M_{\star,\text{vir}}} = r(\varphi = 1) = \frac{\rho_{\text{vir}}}{\sqrt{3} \rho(r)} \propto \varphi.$$  \hfill (C12)

If we define $\varphi_0$ to be the value of $\varphi$ for gas at the virial density of the halo, then since the mean density of the halo $\rho_{\text{vir}}$ equals the local density when $r = r_{\text{vir}}/\sqrt{3}$, the fraction of a halo that is cooling faster than the dynamical timescale, $\epsilon_{\text{cool}}(M_\star)$, will be...
Fig. 14.—Ratio of dynamical time to cooling time for Ω₀ = 0.35, as a function of collapse redshift and comoving radius. Notation is as in Fig. 13.

given by

$$\epsilon_{\text{cool}}(M) = \min \left[ \frac{\varphi_0}{\sqrt{3}}, 1 \right].$$  \hfill (C13)

This is, of course, an upper limit, since halos will have cores, so that the density no longer scales as $r^{-2}$. To account for core radii, we set $\epsilon_{\text{cool}} = 0$ if $\varphi_0/\sqrt{3} < 0.01$, corresponding to a core radius equal to about 1/30 of the virial radius.

APPENDIX D
CALCULATIONAL IMPLEMENTATION

D1. SCALED EQUATIONS

In order to simplify our differential equations (2) and (4), we convert to dimensionless energy densities by dividing $E$ and $\dot{S}$ by $\epsilon_0 n_H$:

$$\mathcal{E} \equiv \frac{E}{\epsilon_0 n_H},$$  \hfill (D1)

$$\mathcal{S} \equiv \frac{\dot{S}}{\epsilon_0 n_H}.$$  \hfill (D2)
We scale the luminosity, \(L(M_b, \tilde{t}, t)\) (eq. [27]), by \(\epsilon_0\):

\[
\mathcal{L}(M_b, \tilde{t}, t) = \frac{L(M_b, \tilde{t}, t)}{\epsilon_0} \\
= [1 - f_\star(\tilde{t})] \frac{M_b \epsilon_{\text{cool}} \epsilon_{\text{eff}} c^2}{\epsilon_0 \dot{t}_{\text{dyn}}} \left( \frac{t - \tilde{t}}{\dot{t}_{\text{dyn}}} \right) \exp \left[ - \left( \frac{t - \tilde{t}}{\dot{t}_{\text{dyn}}} \right) \right].
\]

(D3)

We also make the following definitions:

\[
n_0 \equiv n_H a^3 \\
\zeta \equiv \delta \lambda^{-1/3} \\
\beta \equiv n_H c \sigma_p \\
\xi \equiv n_H c \sigma_c \\
\Gamma \equiv \phi n_H \chi \\
K_{ci} \equiv \phi n_H k_{ci} \\
H \equiv \frac{\dot{a}}{a}
\]

(D4) \quad (D5) \quad (D6) \quad (D7) \quad (D8) \quad (D9) \quad (D10)

\eta^{-1} \equiv x_{hi}(1 - f_\star) = x_{hi} e^{-\lambda}.

(D11)

The expression for \(\mathcal{S}\) in closed form from equation (31) is thus

\[
\mathcal{S} = n_0^{-1} \int_0^\infty dM_b \int_0^t d\tilde{t} N(M_b, \tilde{t}) \mathcal{L}(M_b, \tilde{t}, t) ;
\]

(D12)

the expression for \(\zeta\), using equation (39), is

\[
\zeta = a \frac{3}{4 \pi \mathcal{S}} \left( \frac{3}{4 \pi \mathcal{S}} \right)^{1/3} n_0^{-4/3} \int_0^\infty dM_b \int_0^t d\tilde{t} N(M_b, \tilde{t}) \mathcal{S}(M_b, \tilde{t}, t) \mathcal{S}^{4/3}(M_b, \tilde{t}, t).
\]

(D13)

Note that although the integrands that determine \(\mathcal{S}\) and \(\zeta\) depend on \(f_\star\) through the selection function \(\psi\), this dependence is only for \(f_\star\) evaluated at times \(\tilde{t} < t\), since \(\mathcal{L}(M_b, \tilde{t} = t, t) = 0\).

These integrals can be simplified by making the definition

\[
\tilde{\psi}(\tilde{t}) \equiv \int dR \psi(R, \tilde{t}) \tilde{N}_{\text{form}}(R, \tilde{t}) [1 - f_\star(\tilde{t})] \epsilon_{\text{cool}}(R, \tilde{t}) \frac{4 \pi R^3}{3} \left( m_H c^2 \epsilon_{\text{eff}} \right) X \frac{\epsilon_0}{X},
\]

(D14)

which is the cumulative source function for star formation, with energy in Rydbergs. The scaled source function \(\mathcal{S}\) is given by smoothing \(\tilde{\psi}\) over a dynamical time and accounting for the destruction of halos through mergers. The smoothing function \(\phi(\tilde{t})\), from the definition of \(L\), is given by

\[
\phi(\tilde{t}) \equiv \frac{1}{\dot{t}_{\text{dyn}}} \left( \frac{t - \tilde{t}}{\dot{t}_{\text{dyn}}} \right) \exp \left[ - \left( \frac{t - \tilde{t}}{\dot{t}_{\text{dyn}}} \right) \right] \frac{D(\tilde{t})}{D(t)};
\]

(D15)

so the scaled source is given by

\[
\mathcal{S}(t) = \int_0^t d\tilde{t} \phi(\tilde{t}) \phi(\tilde{t}) .
\]

(D16)

Similarly, defining

\[
\tilde{\psi}_{4/3}(\tilde{t}) \equiv \int dR \psi(R, \tilde{t}) \tilde{N}_{\text{form}}(R, \tilde{t}) \left[ 1 - f_\star(\tilde{t}) \right] \epsilon_{\text{cool}}(R, \tilde{t}) \frac{4 \pi R^3}{3} \left( m_H c^2 \epsilon_{\text{eff}} \right) X \frac{\epsilon_0}{X} \right)^{4/3},
\]

(D17)

we obtain from equation (D13)

\[
\zeta = a \frac{3}{4 \pi \mathcal{S}} \left( \frac{3}{4 \pi \mathcal{S}} \right)^{1/3} \int_0^t d\tilde{t} \phi(\tilde{t}) \frac{\phi(\tilde{t})}{\phi(\tilde{t})} .
\]

(D18)

In terms of the scaled source function \(\mathcal{S}\), the differential equation for \(f_\star\), the fraction of baryons in stars/quasars, is

\[
f_\star = \epsilon_\star \frac{\mathcal{S}(t) \epsilon_0 X \epsilon_{\text{eff}} m_H c^2}{X}.
\]

(D19)
obtained by differentiating equation (30). The ionization fraction and UV energy evolution equations (2) and (4) thus become

$$\dot{x}_i = x_i[(1 - x_i)K_{ei} - x_i \Gamma] + (1 - x_i) \frac{\beta E}{f_i},$$

(D20)

$$\dot{\varepsilon} = \varepsilon - [\gamma H + (1 - x_i)\xi] \varepsilon - \eta^{-1} \lambda,$$

(D21)

having replaced $f_i$ by $(1 - f_i)\lambda$. Our constraint equation, derived from equation (39), is

$$\lambda = \left(\frac{\varepsilon}{\xi}\right)^3,$$

(D22)

and $\varepsilon$ and $\xi$ are defined above.

## D2. Finite-Difference Equations

The differential equations (D20) and (D21) are nonlinear and “stiff,” so forward differencing such as by Runge-Kutta is not stable. Two methods often used in such stiff systems are “implicit” and “semisimilicic” differencing (Press et al. 1992). Implicit differencing evaluates the right-hand side derivatives at the new location, but is only guaranteed to have a stable solution in closed form for linear systems. In our case, the nonlinearity of the equations precludes using purely implicit differencing. Semisimilicic differencing linearizes the right-hand side derivatives, inverting a matrix at each step, but is not guaranteed to be stable. Furthermore, because of the complicated nature of the right-hand side derivatives, the Jacobian matrix required for these methods is extremely cumbersome to evaluate. Below, we use a simple “mostly backward” differencing scheme.

At each time step, we first evaluate $\varepsilon$ and $\xi$ using equations (D12) and (D13) and update the temperature of each phase using equations (A7) and (A8). At very high redshift, $\varepsilon$ and $\xi$ will be zero; we leave $x_i$ at the postrecombination residual ionization, and $\varepsilon$ and $\lambda$ as zero. When $\varepsilon$ and $\xi$ become nonzero, we solve for the new values of $\lambda$ and $\varepsilon$ simultaneously from equations (D21) and (D22), using the previous time step value for $x_i$. We then solve the ionization equation (D20) to update the value of $x_i$. Finally, we update the abundance of $H_2$. The finite-differencing implementation of this procedure follows below.

In differencing equation (D21), we take the values of $(1 - x_i)$ and $\xi$ at the current time step, $t_i$. The other values are evaluated at the advanced time step $t_{i+1}$. The finite-difference equation for $dt = t_{i+1} - t_i$ is thus

$$\varepsilon_{i+1} - \varepsilon_i = dt \varepsilon_i + dt [\gamma H_{i+1} + (1 - x_i)\xi_{i+1}] \varepsilon_{i+1} - \eta^{-1} (\lambda_{i+1} - \lambda_i).$$

(D23)

Replacing $\lambda_{i+1}$ with $(\varepsilon_{i+1}/\xi_{i+1})^3$, equation (D22), and collecting terms gives the equation

$$\left(\frac{\varepsilon_{i+1}}{\xi_{i+1}}\right)^3 \eta^{-1} + \varepsilon_{i+1} U - V = 0,$$

(D24)

with

$$U \equiv 1 + dt \gamma H_{i+1} + dt (1 - x_i)\xi_{i+1},$$

(D25)

$$V \equiv \lambda_i \eta^{-1} + \varepsilon_i + \varepsilon_{i+1} dt.$$

(D26)

While this cubic equation has an algebraic solution, it is faster to solve for $\varepsilon_{i+1}$ iteratively using Newton’s method. Equation (D24) is of the form

$$f(y) \equiv Cy^3 + Uy - V = 0,$$

(D27)

where $y = \varepsilon_{i+1}$. Newton’s method iterates to a solution using

$$y_{j+1} = y_j - \frac{f(y_j)}{f'(y_j)} = y_j - \frac{Cy_j^3 + Uy_j - V}{3Cy_j^2 + U}.$$

(D28)

We use the subscript $j$ so as not to confuse this iterative root finding with our time-step labels $i$. The key to Newton’s method is a good “first guess” for $y$. Since $U$, $V$, $C > 0$, we know that the solution $\bar{y}$ satisfies $\bar{y} < (V/C)^{1/3}$ and $\bar{y} < V/U$. Therefore, our first guess is $y_0 = \min \{ (V/C)^{1/3}, V/U \}$. In cases where $U \ll C^{1/3}V^{2/3}$ or $U \gg C^{1/3}V^{2/3}$, the iteration will converge in only one or two steps. Once a solution for $\varepsilon_{i+1}$ is found, we then use equation (D22) to find $\lambda_{i+1}$.

We use the updated values for $\varepsilon$ and $\lambda$ in the ionization equation (D20). We evaluate all values at the advanced time step. The finite-difference equation is thus

$$x_{II,i+1} - x_{II,i} = dt \left[x_{II,i+1}(1 - x_{II,i+1})K_{ei,i+1} - x_{II,i+1} \Gamma_{i+1} + (1 - x_{II,i+1})\beta_{i+1} \frac{\varepsilon_{i+1}}{f_{II,i+1}}\right].$$

(D29)

Collecting terms gives

$$x_{II,i+1}^2 dt(\Gamma_{i+1} + K_{ei,i+1}) + x_{II,i+1} \left[1 + dt \left(\beta_{i+1} \frac{\varepsilon_{i+1}}{f_{II,i+1}} - K_{ei,i+1}\right)\right] = x_{II,i} + dt \beta_{i+1} \frac{\varepsilon_{i+1}}{f_{II,i+1}},$$

(D30)
or equivalently,

\[ x_{\text{II},i+1} + Ax_{\text{II},i+1} - B = 0, \tag{D31} \]

where

\[ A \equiv \frac{1 + dt(\beta_{i+1} E_{i+1}/f_{\text{II},i+1} - K_{\text{II},i+1})}{dt(\Gamma_{i+1} + K_{\text{II},i+1})}, \tag{D32} \]

\[ B \equiv \frac{\chi_{\text{II}} + dt \beta_{i+1} E_{i+1} f_{\text{II},i+1}}{dt(\Gamma_{i+1} + K_{\text{II},i+1})}. \tag{D33} \]

This can be solved for \( x_{\text{II},i+1} \) either algebraically or iteratively.

The new values of the electron fraction are used to update the molecular hydrogen abundance. Converting the differential equation for the evolution of \( H_2 \) (eq. [B3]) to a backward difference equation gives

\[ x_{\text{H}_2,i+1} = \frac{x_{\text{H}_2,i} + C_{i+1} dt}{1 + D_{i+1} dt}, \tag{D34} \]

where the creation, \( C \), and destruction, \( D \), rates are given by equations (B4) and (B5). We use the equilibrium abundances of \( H^- \) and \( H^+_2 \), as given by equations (B1) and (B2).

REFERENCES

Abel, T., Anninos, P., Zhang, Y., & Norman, M. L. 1997, NewA, 2, 181
Anninos, P., Zhang, Y., Abel, T., & Norman, M. L. 1998, NewA, 2, 209
Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986, ApJ, 304, 15
Binney, J. J. 1977, ApJ, 215, 483
Black, J. H. 1981, MNRAS, 197, 553
Blain, A. W., & Longair, M. S. 1993, MNRAS, 264, 509
Blanton, M., Cen, R., Ostriker, J. P., Strauss, M. A., & Teegmark, M. 2000, ApJ, 531, 1
Bond, J. R., Cole, S., Efstathiou, G., & Kaiser, N. 1991, ApJ, 379, 440
Cen, R. 1992, ApJS, 78, 341
Chiu, W., Ostriker, J. P., & Strauss, M. A. 1998, ApJ, 494, 479
Ciardi, B., Ferrara, A., Governato, F., & Jenkins, A. 1999, MNRAS, submitted (preprint astro-ph/9907189)
Cole, S., Fisher, K. B., & Weinberg, D. H. 1994, MNRAS, 267, 785
Couchman, H. M. P., & Rees, M. J. 1986, MNRAS, 221, 53
Efstathiou, G., Frenk, C. S., White, S. D. M., & Davis, M. 1988, MNRAS, 235, 715
Fixin, D. J., Cheng, E. S., Gales, J. M., Mather, J. C., Shafer, R. A., & Wright, E. L. 1996, ApJ, 473, 576
Giallongo, E., D'Odorico, S., Fontana, A., McMahon, H. G., Savaglio, S., Cristiani, S., Molaro, P., & Trevese, D. 1994, ApJ, 425, L1
Gnedin, N. Y. 1996, ApJ, 450, 1
———. 2000, ApJ, 535, in press (preprint astro-ph/9909383)
Gnedin, N. Y., & Ostriker, J. P. 1997, ApJ, 488, 581
Gooding, A. K., Park, C., Spergel, D. N., Turok, N., & Gott, J. R., III. 1992, ApJ, 393, 42
Haiman, Z., & Loeb, A. 1997, ApJ, 483, 23
———. 1998, ApJ, 503, 505
Janey, R. K., Langer, W. D., Evans, K. J., & Post, D. E. J. 1987, Elementary Processes in Hydrogen-Helium Plasmas (New York: Springer)
Jenkins, E. B., & Ostriker, J. P. 1991, ApJ, 376, 33
Kaiser, N. 1984, ApJ, 284, L9
Kashlinsky, A., & Rees, M. J. 1983, MNRAS, 205, 955
Katz, N., Hernquist, L., & Weinberg, D. H. 1999, ApJ, 523, 463
Katz, N., Weinberg, D. H., & Hernquist, L. 1996, ApJS, 105, 19
Kauffmann, G., Colberg, J. M., Diaferio, A., & White, S. D. M. 1999, MNRAS, 307, 529
Kauffmann, G., Nusser, A., & Steinmetz, M. 1997, MNRAS, 286, 795
Lacey, C. G., & Cole, S. 1993, MNRAS, 262, 627
Lepp, S., & Shull, J. M. 1983, ApJ, 270, 578
Miralda-Escudé, J., Haehnelt, M., & Rees, M. J. 1999, A&A, in press
Mo, H. J., & White, S. D. M. 1996, MNRAS, 282, 347
Ostriker, J. P., & Gnedin, N. Y. 1996, ApJ, 472, L63
Park, C., Spergel, D. N., & Turok, N. 1991, ApJ, 372, L53
Pearce, F. R., et al. 1999, ApJ, 521, L99
Peebles, P. J. E. 1993, Principles of Physical Cosmology (Princeton: Princeton Univ. Press)
———. 1999a, ApJ, 510, 523
———. 1999b, ApJ, 510, 531
Pen, U., & Spergel, D. N. 1995, Phys. Rev. D, 51, 4099
Press, W. H., & Schechter, P. 1974, ApJ, 187, 425
Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, Numerical Recipes in C (2d ed.; Cambridge: Cambridge Univ. Press)
Rauch, M., et al. 1997, ApJ, 489, 7
Rees, M. J., & Ostriker, J. P. 1977, MNRAS, 179, 541
Robinson, J., & Baker, J. E. 1999, preprint (astro-ph/9905098)
Robinson, J., Gawiser, E., & Silk, J. 1998, preprint (astro-ph/9805181)
Sasaki, S. 1994, PASJ, 46, 427
Shapiro, P. R., & Giroux, M. L. 1987, ApJ, 321, L107
Shapiro, P. R., Giroux, M. L., & Babul, A. 1994, ApJ, 427, 25
Shapiro, P. R., & Kang, H. 1987, ApJ, 318, 638
Silk, J. 1977, ApJ, 211, 638
Spitzer, L. 1978, Physical Processes in the Interstellar Medium (New York: Wiley)
Sugiyama, N. 1995, ApJS, 100, 281
Tegmark, M., Silk, J., Rees, M. J., Blanchard, A., Abel, T., & Palla, F. 1997, ApJ, 474, 1
Valageas, P., & Silk, J. 1999, A&A, 347, 1
Viana, P. T. P., & Liddle, A. R. 1996, MNRAS, 281, 323
Zaldarriaga, M. 1998, Ph.D. thesis, MIT