WHAT SHAPES THE LUMINOSITY FUNCTION OF GALAXIES?

A. J. BENSON

California Institute of Technology, MS 105-24, 1200 East California Boulevard, Pasadena, CA 91125;
abenson@astro.caltech.edu

AND

R. G. BOWER, C. S. FRENK, C. G. LACEY, C. M. BAUGH, AND S. COLE,

Institute for Computational Cosmology, Physics Department, University of Durham,
South Road, Durham DH1 3LE, England, UK; c.s.frenk@durham.ac.uk,
c.g.lacey@durham.ac.uk, c.m.baugh@durham.ac.uk,
shaun.cole@durham.ac.uk

ABSTRACT

We investigate the physical mechanisms that shape the luminosity function of galaxies in hierarchical clustering models. Beginning with the mass function of dark matter halos in the ΛCDM (Λ cold dark matter) cosmology, we show, in incremental steps, how gas cooling, photoionization at high redshift, feedback processes, galaxy merging, and thermal conduction affect the shape of the luminosity function. We consider three processes whereby supernovae and stellar wind energy can affect the forming galaxy: (1) the reheating of cold disk gas to the halo temperature; (2) expansion of the hot, diffuse halo gas; and (3) complete expulsion of cold disk gas from the halo. We demonstrate that while feedback of form 1 is able to flatten the faint end of the galaxy luminosity function, this process alone does not produce the sharp cutoff observed at large luminosities. Feedback of form 2 is also unable to solve the problem at the bright end of the luminosity function. The relative paucity of very bright galaxies can only be explained if cooling in massive halos is strongly suppressed. This might happen if thermal conduction near the centers of halos is very efficient, or if a substantial amount of gas is expelled from halos by process 3 above. Conduction is a promising mechanism, but an uncomfortably high efficiency is required to suppress cooling to the desired level. If, instead, superwinds are responsible for the lack of bright galaxies, then the total energy budget required to obtain a good match to the galaxy luminosity function greatly exceeds the energy available from supernova explosions. The mechanism is only viable if the formation of central supermassive black holes and the associated energy generation play a crucial role in limiting the amount of stars that form in the host galaxy. The models that best reproduce the galaxy luminosity function also give reasonable approximations to the Tully-Fisher relation and the galaxy autocorrelation function.

Subject headings: conduction — cooling flows — galaxies: evolution — galaxies: formation —
galaxies: luminosity function, mass function

1. INTRODUCTION

The luminosity function of galaxies is one of the most basic properties of the galaxy population, yet it contains many valuable clues to the process of galaxy formation. The basic physical mechanisms that determine the form of the luminosity function were first described by Rees & Ostriker (1977) and White & Rees (1978). In this picture, galaxy formation is regulated by the rate at which gas is able to cool in the parent dark matter halos. These authors suggested that the sharp cutoff in the galaxy luminosity function arose from the long cooling times of gas in high-mass halos (or high-mass protogalaxies, in the case of Rees & Ostriker). The model has been developed by many authors to follow in great detail the formation of galaxies in a hierarchical universe. Key improvements are the inclusion of galaxy merging and the evolution of stellar populations (White & Frenk 1991; Cole 1991; Kauffmann, White, & Guiderdoni 1993; Lacey et al. 1993; Cole et al. 1994, 2000; Kauffmann et al. 1999; Somerville & Primack 1999; Benson et al. 2002). Such models are now being strongly tested by high-precision measurements of the galaxy luminosity function from large redshift surveys, such as the Two-Degree Field Galaxy Red-shift Survey (2dFGRS) and Two Micron All-Sky Survey (2MASS) (Cole et al. 2001; Kochanek et al. 2001).

While the key physics of gas cooling and merging is now thought to be modeled with reasonable accuracy (Benson et al. 2001, 2002; Yoshida et al. 2002; Helly et al. 2003; Voit et al. 2002), other physics crucial to establishing the shape of the luminosity function remains poorly understood. The first uncertainty is the " feedback " needed to regulate the formation of dwarf galaxies, and hence reconcile the rather shallow slope of the faint end of the observed luminosity function with the relatively steep mass function of dark matter halos. While outflows of gas from galaxies have been observed at both low and high redshift (Martin 1999; Pettini et al. 2002), the complex physics at work has not yet been understood in detail, and most models of galaxy formation simply adopt phenomenological rules to describe their effects. Previous work has typically assumed that the dominant feedback mechanism is the reheating of cold gas in the disk to the temperature of the diffuse gas halo (White & Frenk 1991; Cole et al. 1994; Kauffmann et al. 1999; Efstathiou 2000), although complete expulsion of disk gas from the halo was considered by Somerville & Primack (1999). However, although this form of feedback solves the

problem with the faint-end slope of the luminosity function (as it was originally designed to do), it creates a second problem of matching the very sharp (exponential with luminosity) cutoff seen at the bright end. In addition, the effectiveness with which the gas needs to be reheated seems excessive compared to observations of galactic outflows (Martin 1999) and to what is seen in simulations of this process in individual galaxies (Strickland & Stevens 2000). These papers suggest that the mass in the outflow is comparable to the gas mass that is turned into stars.

All current models of galaxy formation, calculated using either gasdynamical simulations (Pearce et al. 2001; Kay et al. 2002; Murali et al. 2002; Weinberg et al. 2003) or semianalytic techniques (e.g., Kauffmann et al. 1999; Somerville & Primack 1999; Cole et al. 2000), exhibit strong gas cooling in the central regions of groups and clusters. This leads to the formation of extremely bright galaxies (which are never seen in reality) unless some additional suppression of the gas cooling is assumed. The suppression mechanisms that have been considered in semianalytical models are (1) simply switching off cooling and/or star formation in the most massive halos (e.g., Kauffmann et al. 1999); (2) redistributing the gas within the halo so that it becomes so rarefied that its cooling time is longer than the age of the universe (e.g., Cole et al. 2000); or (3) suppressing cooling until the halo is completely virialized, e.g., van Kampen, Jimenez, & Peacock (1999). The model presented in Cole et al. (2000) adopted a low value for the cosmic baryon fraction ($\Omega_b = 0.02$) and a model for the halo gas distribution in which the core radius was a function of the amount of gas that had cooled in previous generations of halos. In combination, these assumptions were able to produce a good match to the galaxy luminosity function; however, both are now disfavored by current observational data (e.g., O'Meara et al. 2001; Allen et al. 2001; Ettori, De Grandi, & Molendi 2002). If the more recent, higher value of the baryon abundance ($\Omega_b \approx 0.04$) is adopted, the growing core radius has little effect on the cooling rates (since the cooling radius is then significantly beyond the typical core radii).

In this paper, we investigate physical processes that may be responsible for shaping the bright end of the luminosity function. In addition to the conventional feedback process, in which cold gas in the disk is reheated to the temperature of the diffuse gas halo, the feedback energy may be used to regulate the formation of the galaxy in two further ways. While the wind flowing from the gas disk may contain a relatively small mass, the energy in the disk outflow may be transferred to the existing hot gas halo, causing it to expand within the halo potential. This makes the central gas more diffuse, lengthens its cooling time, and so reduces the rate at which gas is supplied to the cold disk. This type of process is seen in simulations of the effect of injection of relativistic radio-emitting plasma in the centers of clusters (Quilis, Bower, & Balogh 2001; Brüggen et al. 2002). An alternative is to assume that the gas expelled from the disk does not mix with the virialized diffuse gas halo and that the energy per particle is sufficiently high that the material escapes the confining potential and is never recaptured.

A further means by which the supply of cold gas to the galaxy disk may be reduced is suggested by recent *Chandra* and *XMM* observations of galaxy clusters (Peterson et al. 2001; Tamura et al. 2001; Fabian et al. 2001; Johnstone et al. 2002; Böhringer et al. 2002; McNamara et al. 2001; Nulsen et al. 2002). These have led to a revival of the idea that thermal conduction may be an important source of

heating in the central regions of clusters (Narayan & Medvedev 2001; Gruzinov 2002; Voigt et al. 2002). Indeed, the conductivity has been inferred to be close to the Spitzer rate expected for a fully ionized plasma (Fabian, Voigt, & Morris 2002). Heating due to conduction could plausibly counteract the excessive cooling predicted to occur in the most massive halos and thereby explain the dearth of highly luminous galaxies in the universe.

In the remainder of this paper, we demonstrate quantitatively the importance of each physical mechanism and its role in setting the form of the luminosity function. We demonstrate that the milder forms of feedback are unable to account simultaneously for the sharp break in the luminosity function and the flat faint-end slope. The two processes discussed above, thermal conduction and superwinds, may provide the answer to this problem. Thermal conduction is capable of suppressing the formation of the brightest objects if the conductive efficiency is high. Alternatively, a model that includes energetic gas expulsion is also able to produce reasonable fits to the observed luminosity function, but the energy required to power this "superwind" is larger than that available from supernova explosions and stellar winds.

2. THE MODEL

We compute the luminosity function of galaxies in a cold dark matter (CDM) universe using a development of the semianalytic model, GALFORM, described by Cole et al. (2000). Our changes to the model concern the treatment of the gas distribution in halos and the treatment of feedback.

2.1. *Feedback*

In Cole et al. (2000), the star formation rate was affected by feedback generated when cold gas was reheated by energy injected by supernova explosions and stellar winds. This process transfers gas from the star-forming disk of the galaxy to a diffuse corona or halo. The rate at which cold gas is reheated is assumed to be related to the star formation rate, $\dot{M}_*$, by

$$\dot{M}_{\rm reheat} = \left(\frac{V_{\rm disk}}{V_{\rm hot}}\right)^{\alpha} \dot{M}_* , \qquad (1)$$

where $V_{\rm disk}$ is the circular speed at the half-mass radius of the gas disk and $V_{\rm hot}$ and α are adjustable parameters. Here we include this feedback mechanism, but we also consider two further schemes, which correspond to different processes by which the energy from supernova explosions and stellar winds can regulate the gas cooling and star formation rates. A further fraction of the total energy will be radiated away and so is not available for feedback. We assume that the way the energy is apportioned among the three processes is fixed by the local supernova environment and does not depend on the mass of the galaxy or the instantaneous star formation rate.[1] Thus, feedback in our model is specified by three parameters: $\epsilon_{\rm reheat}$, $\epsilon_{\rm halo}$, and $\epsilon_{\rm sw}$. It is useful to note that the total energy available from supernovae is on the

---

[1] In principle the supernova's local environment—the interstellar medium—could depend on global properties of the galaxy, such as its mass, but we neglect any such dependence here.

order of 10^{49} ergs for each solar mass of stars that is formed, for a standard initial mass function (IMF).

1. *Disk reheating.*—This is the feedback scheme explored by Cole et al. (2000). Gas is removed from the disk of the galaxy at a rate proportional to the star formation rate. This gas can be thought of either as being ejected at close to, but below, the escape velocity of the halo or as being heated to the halo virial temperature. In the first case, the gas will subsequently be heated back to the halo virial temperature, through mixing with the halo gas or through further shocks when a new halo forms. In either case, the gas does not leave the halo and so is available for cooling in the future. We parameterize the energy invested in reheating the cold disk gas as $\epsilon_{\text{reheat}} 10^{49}$ ergs per solar mass of stars formed. If we approximate the gravitational potential wells of galaxies as being self-similar, then the energy required to eject the gas from the disk should scale as V_{disk}^2, and one obtains equation (1) with $\alpha = -2$. Specifically, if we assume that gas is ejected with energy per unit mass equal to V_{disk}^2, then the rate at which gas is reheated is given by

$$\dot{M}_{\text{reheat}} = \frac{5.6\epsilon_{\text{reheat}}}{(V_{\text{disk}}/300 \text{ km s}^{-1})^2} \dot{M}_* . \quad (2)$$

We therefore have the following relation between ϵ_{reheat} and V_{hot}:

$$\epsilon_{\text{reheat}} = 0.18 \left(\frac{V_{\text{hot}}}{300 \text{ km s}^{-1}} \right)^2 . \quad (3)$$

If the circular speed of the disk is equal to V_{hot}, the mass of gas reheated is equal to the mass formed into stars. Feedback of this form is motivated by the study of Efstathiou (2000), although the efficiency assumed here is somewhat higher.

2. *Energy injection into the gaseous halo.*—The feedback energy is assumed to redistribute the hot, diffuse gas so that the central density of the gas is reduced and the central cooling time is increased. The energy injected per unit mass of stars formed is parameterized as $\epsilon_{\text{halo}} 10^{49}$ ergs $M_\odot^{-1}$. The nature of this type of feedback process can be understood as follows: The total energy injected into the halo grows with time until it becomes comparable to the gravitational binding energy of the gaseous halo. At this point, the cooling rate drops dramatically. Since the binding energy of a halo is proportional to $M_{\text{halo}}^{5/3}$, the fraction of the total gaseous mass that must be turned into stars to achieve this balance is strongly dependent on halo mass. Larger halos are therefore able to cool a larger fraction of their baryon reservoir to form galaxies. This mechanism is described in more detail in Bower et al. (2001) and, as noted above, is motivated by simulations of the injection of radio plasma in cluster centers (Quilis et al. 2001; Brüggen et al. 2002).

3. *Gas expulsion from the halo.*—In this scheme, an energy of $\epsilon_{\text{sw}} 10^{49}$ ergs $M_\odot^{-1}$ is invested in heating a small mass of the cold gas disk to an energy much greater than the binding energy of the halo. If the energy is contained within this superheated phase (and is not shared with the diffuse gas halo, as assumed in mechanism 2 above), the wind may be able to escape completely from the halo. This mechanism is motivated by the strong, high-velocity winds inferred to exist around Lyman break galaxies (Pettini et al. 2002; Adelberger et al. 2003) and has recently been studied theoretically by Shu, Mo, & Mao (2003). We assume that the energy is

injected with a thermal distribution of particle energies, i.e., so that the fraction of mass ejected with energy in the range E–$(E + dE)$ is proportional to $\exp(-E/E_{\text{av}})dE$, where E_{av} is the mean ejection energy per particle. The fraction of the superwind gas with sufficient energy to escape the halo is computed using this distribution. Some of the expelled gas will be recaptured when a deeper potential well forms. To estimate this, we calculate the fraction of the expelled gas with energy less than the depth of the new potential well and allow this fraction to be recaptured. This process continues as deeper potential wells form until either all of the gas is recaptured or $z = 0$ is reached.

We assume that the mass flux of material in the wind as it leaves the disk, $\dot{M}_{\text{sw}}$, is proportional to the star formation rate, with coefficient of proportionality β_{sw}. The energy in the superwind as it leaves the disk is given by

$$\dot{E}_{\text{sw}} = 10^{49} \epsilon_{\text{sw}} \dot{M}_* \text{ergs} , \quad (4)$$

where $10^{49}\epsilon_{\text{sw}}$ ergs is the energy per unit mass of star formation. The characteristic specific energy of the wind is simply $E_{\text{av}} = \dot{E}_{\text{sw}}/\dot{M}_{\text{sw}} \equiv \frac{1}{2} V_{\text{sw}}^2$, and so the characteristic velocity of the wind is

$$V_{\text{sw}} = 1002\sqrt{\epsilon_{\text{sw}}/\beta_{\text{sw}}} \text{ km s}^{-1} . \quad (5)$$

We assume that material ejected from the disk requires an energy per unit mass $\lambda_{\text{sw}} V_{\text{disk}}^2$ in order to escape also from the halo. Therefore, in the case in which all of the wind material has the same energy, superwinds will be driven from halos with $V_{\text{disk}} < 709(\epsilon_{\text{sw}}/\beta_{\text{sw}}\lambda_{\text{sw}})^{1/2}$ km s^{-1}, and not from deeper potential wells. For typical halos in our model, $\lambda_{\text{sw}} = 2.9$ gives a good estimate of the energy required to escape the combined gravitational pull of the galaxy and its dark matter halo, and so we use this value throughout this work. If we assume a distribution of energies in the wind, then there will be a smooth transition, as V_{disk} increases, between the regime in which the wind escapes and the regime in which it is retained. Thus, the actual mass flux in the superwind escaping from the halo is given by

$$\dot{M}_{\text{sw}} = f_{\text{eject}} (E_{\text{av}}/\lambda_{\text{sw}} V_{\text{disk}}^2) \beta_{\text{sw}} \dot{M}_* , \quad (6)$$

where, for our chosen energy distribution, $f_{\text{eject}}(x) = \exp(-x)$. The material that does escape may be recaptured by larger halos forming at later times. Our estimate of the fraction of the wind that escapes and the fraction that can be recaptured by larger halos is described in the Appendix. We regard this mechanism as being more uncertain than mechanisms 1 and 2, because the expelled gas must leave the halo without sharing its energy with the diffuse hot component. It must therefore punch a well-collimated hole through the halo or remain contained within a buoyant bubble that is convected out of the halo (Quilis et al. 2001; Springel & Hernquist 2003; Kay et al. 2002).

In general, it seems plausible that some fraction of the total feedback energy will be processed into each one of these forms, and we consider models that include a combination of these processes on an equal footing with models that include only one.

2.2. *Halo Gas Distribution*

The hot gaseous component in dark matter halos is assumed initially to have a density profile, $\rho_g(r)$, given by

the β-model; i.e., at radius r in the halo (whose dark matter density profile is assumed to have the NFW profile; Navarro, Frenk, & White 1996, 1997),

$$\rho_g(r) = \frac{\rho_0}{\left[1 + (r/r_c)^2\right]^{3\beta/2}} , \qquad (7)$$

where ρ_0 is the density at the center of the halo, r_c is the radius of the " core," and β is a parameter that sets the slope of the profile on scales larger than r_c.

Departing from the prescription of Cole et al. (2000), we adopt a gas density profile in the absence of energy injection with fixed $r_c = 0.07 r_{\rm vir}$ (and $\beta = \frac{2}{3}$) in all halos. This provides a reasonable match to gasdynamic simulations of non-radiative gas in clusters (e.g., Eke, Navarro, & Frenk 1998) and to the observed X-ray profiles of relaxed clusters (e.g., Allen et al. 2001). The normalization of the profile is determined by the total diffuse gas mass remaining in the halo, and the temperature of the gas is set assuming hydrostatic equilibrium. As a boundary condition, we set the temperature at the virial radius equal to the virial temperature (eq. [4.1] in Cole et al. 2000). This " default " profile is modified if the diffuse gas gains further energy (" excess energy ") as a result of energy injection (process 2 in § 2.1). We use the prescription described in Bower et al. (2001), in which the excess energy first causes the slope of the gas profile to decrease down to a minimum value of $\beta_{\rm min} = 0.2$, after which it increases the boundary temperature of the gas halo. The effect of the excess energy is to decrease the central density of the gas, lengthening its cooling time. Mass is conserved by pushing some of the diffuse gas outside the halo; however, this gas can be recaptured as the total halo mass (and thus the gravitational binding energy) increases. We assume that the excess energy is conserved during mergers between halos, although the results are not qualitatively affected by a small dilution or amplification of energy during halo mergers.

We define the effective cooling time at radius r, $t'_{\rm cool}(r)$, as the maximum true cooling time (i.e., that defined by Cole et al. 2000) occurring at smaller radii plus the free-fall time from r to the halo center. This ensures that $dt'_{\rm cool}/dr > 0$, so the cooling radius is always a smoothly increasing function of radius. Experiments with different approaches show that the results we present here are not sensitive to the details of this prescription.

2.3. *Conduction*

Conduction in the ionized gas can transport energy into the inner regions of the halo, effectively increasing the cooling time of the gas there. The rate at which energy is deposited into the shell between radii r and $r + dr$ is given by

$$\Sigma \, dr = 4\pi \frac{d}{dr}\left(\kappa r^2 \frac{dT}{dr}\right) dr , \qquad (8)$$

where the conductivity, κ, may depend on radius through its temperature dependence. We approximate Σ as

$$\Sigma = \alpha_{\rm cond} 4\pi \kappa_{\rm S} T , \qquad (9)$$

where $\kappa_{\rm S}$ is the Spitzer conductivity (Spitzer 1962) and $\alpha_{\rm cond}$ is a parameter that absorbs the dependence on the shape of the temperature profile, as well as any difference between the actual conductivity and the Spitzer rate. For a power-law temperature profile, $T \propto r^a$, and conductivity of the Spitzer form, $\kappa_{\rm S} \propto T^{5/2}$, $\alpha_{\rm cond} = f_{\rm sp} a(1 + 7a/2)$, where $f_{\rm sp}$ is the ratio of the conductivity to the Spitzer value. Adopting a temperature profile with $a = 0.4$, as suggested by recent X-ray observations (Voigt et al. 2002), gives $\alpha_{\rm cond} = 0.96 f_{\rm sp}$. Adopting a linear temperature gradient gives $\alpha_{\rm cond} = 4.5 f_{\rm sp}$.

The heating rate due to conduction is subtracted from the radiative cooling rate to give a net cooling rate for the gas. This net cooling rate is used to compute the cooling time. Conduction causes the cooling radius to become smaller than in the standard model. The result is a suppression of cooling in hot halos, and the mass at which this effect becomes important is determined by the parameter $\alpha_{\rm cond}$.

2.4. *Photoionization and Merging*

While we employ the detailed model of galaxy merging developed by Benson et al. (2002), we choose to use a much simpler model of the effects of reionization than used in that paper, in order to incorporate them easily into our calculation of the galaxy luminosity function. We simply assume that galaxy formation is completely suppressed by reionization in dark matter halos with circular velocities below $V_{\rm reion}$ after redshift $z_{\rm reion}$. Unless otherwise stated, we adopt $V_{\rm reion} = 50$ km s^{-1} and $z_{\rm reion} = 6$. With this choice, this simple model matches the results of the full calculation of Benson et al. (2002) quite well.

3. RESULTS

Throughout this paper, we use the same parameter values adopted by Benson et al. (2002), with the exception of a larger baryon fraction corresponding to $\Omega_b = 0.045$, consistent with constraints from big bang nucleosynthesis (O'Meara et al. 2001). We assume a ΛCDM universe with mean matter density $\Omega_0 = 0.3$, cosmological constant term $\Omega_\Lambda = 0.7$, Hubble constant[2] $H_0 = 70$ km s^{-1} Mpc^{-1}, and linear fluctuation amplitude on spheres of radius $8 \, h^{-1}$ Mpc $\sigma_8 = 0.9$. Of these parameters, the uncertainty in the value of σ_8 has the greatest effect on the model results. While several studies support a value of $\sigma_8 \sim 0.9$ (e.g., Bacon et al. 2003, Hoekstra, Yee, & Gladders 2002 [gravitational lensing]; Eke, Cole, & Frenk 1996, Viana, Nichol, & Liddle 2002 [cluster abundance]; Spergel et al. 2003 [cosmic microwave background (CMB)]; Sievers et al. 2003 [Sunyaev-Zeldovich effect]), other recent analyses have suggested lower values, $\sigma_8 \sim 0.7$ (e.g., Peacock 2003 [large-scale structure]; Melchiori et al. 2002 [CMB]; Jarvis et al. 2003 [gravitational lensing]; Allen et al. 2003, Smith et al. 2003 [cluster abundance]). A discussion of recent results may be found in Wang et al. (2002). Except where specified, we show models based on $\sigma_8 = 0.9$; however (as we shall see), taking the lower value, $\sigma_8 = 0.7$, considerably eases the energy budget and/or reduces the conduction efficiency required to match the galaxy luminosity function. Throughout, we use the halo mass function derived from N-body simulations by Jenkins et al. (2001),[3] instead of the Press-Schechter mass function used by Cole et al. (2000).

---

[2] Here and below h denotes the Hubble constant in units of 100 km s^{-1} Mpc^{-1}.

[3] Note that we use the Jenkins et al. (2001) mass function to compute the abundances of halos at $z = 0$ but retain the Press-Schechter approximation when computing halo merger trees. An improved calculation would use a self-consistent computation of halo merging histories.

We compare our model with recent determinations of the *K*-band luminosity function assuming a Kennicutt stellar IMF (Kennicutt 1983). In order to facilitate comparison between models, we have kept the IMF fixed and assumed a negligible fraction of brown dwarf stars.[4] We choose the *K* band in order to minimize the sensitivity of our results to recent star formation and to dust obscuration. The model of Cole et al. (2000) includes a detailed and fully self-consistent calculation of dust extinction, which is used in this work. We find, however, that dust obscuration has a negligible effect on our results for the *K*-band luminosity function (typically shifting the bright end faintward by less than 0.1 mag). For the observational comparison, we use the local *K*-band luminosity functions measured by Cole et al. (2001) and Kochanek et al. (2001) (both based on the 2MASS survey) and the local ($z < 0.1$) luminosity function derived from the much deeper *K*-band survey of Huang et al. (2003). The analysis by Huang et al. suggests a faint-end slope, $\alpha_K = -1.37 \pm 0.10$, steeper than the values found by Cole et al. (2001; $\alpha_K = -0.93$) and by Kochanek et al. (2001; $\alpha_K = -1.09$). The latter two are also in good agreement with the faint-end slope of the *z*-band luminosity function measured by Blanton et al. (2003; $\alpha_z = -1.08$) from the SDSS survey. The *z*-band data should also be little affected by residual star formation and dust extinction but have a deeper surface brightness limit. These discrepancies indicate that there remain significant systematic uncertainties in current measurements of the faint end of the *K*-band luminosity function.

3.1. *A Model with a Constant Mass-to-Light Ratio*

In Figure 1, we show the simplest possible model of the luminosity function, which we call model 1 (shown as the dashed line). In this model, the mass function of dark matter halos (Jenkins et al. 2001) has been converted into a luminosity function simply by assuming a fixed mass-to-light ratio ($M/L_K = 11 \ M_\odot/L_{K,\odot}$), chosen so as to match the knee of the observed luminosity function. As is well known, this produces a luminosity function that is much steeper at the faint end than is observed and also fails to cut off at bright magnitudes (the halo mass function does possess a cutoff, but it occurs at much higher mass and lower abundance than shown in the plot).

3.2. *Model Including Cooling Only*

White & Rees (1978) argued that the difference between the halo mass and the galaxy luminosity functions is due to the dependence of the gas cooling time on halo mass and to feedback processes. We use the GALFORM semianalytic model to follow gas cooling and star formation in a merging hierarchy of dark matter halos in the ΛCDM cosmology. In order to illustrate the simplest possible model first, we do not include photoionization suppression, feedback, galaxy merging, or conduction. The result is model 2 (shown as a dotted line in Fig. 1). It clearly displays the "overcooling problem": gas has cooled into the smallest halos resolved in the calculation, producing an overabundance of faint galaxies. As a result, the faint-end slope of the luminosity function is much too steep. It should be noted that the results for this cooling-only model are sensitive to the mass resolution
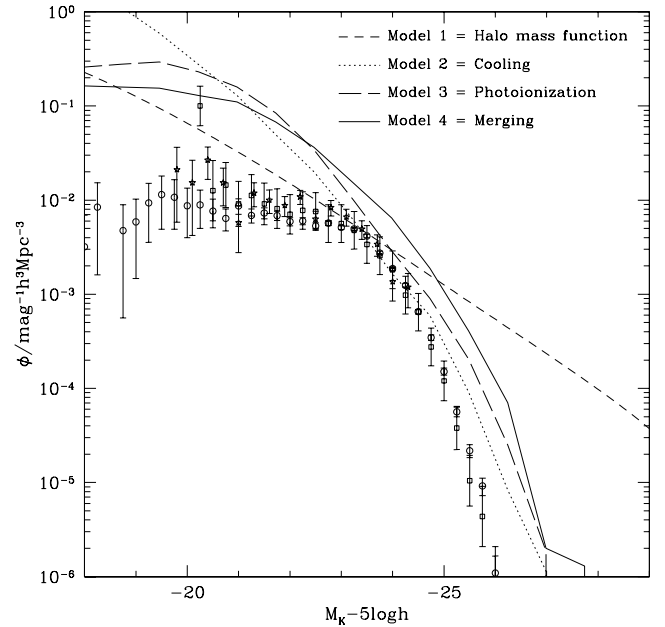
---

[4] We set $\Upsilon = 1$, in the notation of Cole et al. (2000).

FIG. 1.—*K*-band luminosity function of galaxies. The points show the observational determinations of Cole et al. (2001; *circles*), Kochanek et al. (2001; *squares*), and Huang et al. (2003; $z < 0.1$, *stars*). Lines show model results. Model 1 (*dashed line*) shows the result of converting the dark matter halo mass function into a galaxy luminosity function by assuming a fixed mass-to-light ratio chosen to match the knee of the luminosity function. Model 2 (*dotted line*) shows the result from GALFORM when no feedback, photoionization suppression, galaxy merging, or conduction is included. Models 3 and 4 (*long-dashed and solid lines, respectively*) show the effects of adding photoionization and then galaxy merging.

of the GALFORM merger trees. The cooling time for halo gas decreases with decreasing halo mass down to halo virial temperatures around 10^4 K. Below this scale, cooling becomes ineffective unless molecular hydrogen is abundant. This means that for the luminosity function to be fully converged at all galaxy luminosities, we would need to resolve halos with virial temperatures as low as 10^4 K at all redshifts at which there is significant cooling. To obtain the results shown for model 2, we ran GALFORM at the highest mass resolution that was computationally feasible, which comes close to resolving 10^4 K halos at $z = 0$. We believe that our results are substantially converged, but we cannot be certain that they would not change if the mass resolution were increased further.

3.3. *Model with Photoionization and Merging*

The problem of mass resolution encountered in model 2 is effectively removed once we include feedback processes that are effective in low-mass halos. Model 3 (long-dashed line in Fig. 1) shows the effects of including photoionization suppression. As described by Benson et al. (2002), the formation of low-mass galaxies is suppressed after the universe is reionized, and as we noted in § 2, we adopt a simplified model of this process, which nevertheless is a good approximation to a full calculation. The resolution of the GALFORM merger trees is then sufficient to follow the lowest mass halos that are able to form galaxies, resulting in a converged solution. However, photoionization suppression alone is unable to produce a sufficiently flat faint end in the luminosity function. In addition, model 3 also produces more bright galaxies, since less gas has been locked into

stars in the smallest halos. When both galaxy merging (which was artificially switched off in models 1 and 2) and photoionization are included, as in model 4 (solid line in Fig. 1), the faint end remains too steep, and even more bright galaxies are produced. In the models that follow, we investigate the effects of including feedback processes in addition to photoionization and merging.

3.4. Model with Feedback—Disk Reheating

A solution to the faint-end problem is illustrated by model 5.3 in Figure 2. Here feedback is included through the reheating of disk gas in star-forming galaxies. We use the standard prescription of Cole et al. (2000), but with a larger value of $\epsilon_{reheat} = 0.41$ (equivalent to $V_{hot} = 450$ km s^{-1}), which is required in order to obtain a faint-end slope similar to that in Cole et al. for the larger value of Ω_b assumed in this work. This form of feedback flattens the luminosity function considerably, resulting in reasonably good agreement with the observed faint end. While the slope is not as flat as that measured by Cole et al. (2001) or Blanton et al. (2003), it is in good agreement with the steeper slope reported by Huang et al. (2003). This achievement carries a price, however—the overabundance of bright galaxies (formed through excessive cooling in massive halos) is exacerbated, as there is now a much greater mass of diffuse hot gas remaining in the larger halos. This gas is sufficiently dense that the central regions are able to cool; consequently, model 5.3 produces far too many bright galaxies. This result depends little on the choice of σ_8. Adopting $\sigma_8 = 0.7$ makes the brightest galaxies only 0.5 mag fainter. This clearly demonstrates a long-standing problem in semianalytic models: previous calculations have either assumed low values of Ω_b or have invoked rather artificial ways to prevent the cooling that forms these overluminous objects.

The energy requirements of model 5.3 are substantial but not excessive. The reheating energy of 0.41×10^{49} ergs $M_\odot^{-1}$

should be compared to the total energy available from supernova explosions, which is approximately 0.7×10^{49} ergs $M_\odot^{-1}$ for a Salpeter IMF or 0.9×10^{49} ergs $M_\odot^{-1}$ for the Kennicutt IMF adopted in GALFORM. In a halo of circular velocity 250 km s^{-1}, the mass of gas reheated is more than 3 times the mass of gas formed into stars. If the level of reheating is reduced, as illustrated by models 5.1 ($\epsilon_{reheat} = 0.03$) and 5.2 ($\epsilon_{reheat} = 0.13$), then the formation of small galaxies is not suppressed sufficiently to match the observed luminosity function.

3.5. Models with Energy Injection

In models 6.1–6.3 (Fig. 3), we investigate the effect of assuming that a fraction of the supernova and stellar wind energy heats the diffuse gas halo, causing it to expand. This form of feedback suppresses the formation of both bright and faint galaxies, but it does not produce a sharp break in the luminosity function. Models 6.1 and 6.2 illustrate how as the energy spent in heating the diffuse halo increases, the break in the luminosity function becomes less pronounced. If the heating is made even stronger, the resulting luminosity function approaches a power law (model 6.3). This form of feedback clearly cannot solve the problem of overproduction of bright galaxies.

3.6. Model with Conduction

We now consider two possible schemes that are capable of producing a good match to the observed luminosity function. The first involves balancing radiative cooling with thermal conduction. The second involves expelling gas from dark matter halos at such high energies that it is subsequently unable to cool.

Thermal conduction is expected to imprint a special scale on the galaxy population because of the strong temperature
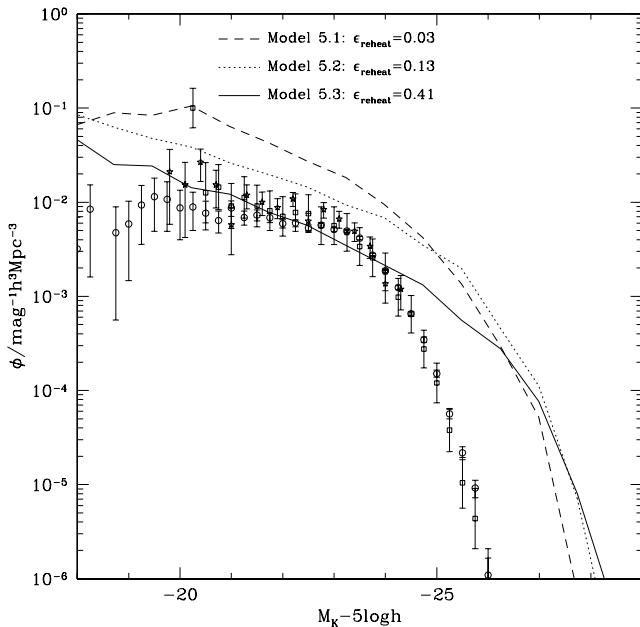

FIG. 2.—Model 5. Starting from model 4, disk reheating is added in order to suppress the formation of small galaxies. Results are shown for three levels of energy input ($\epsilon_{reheat} = 0.03, 0.13$, and 0.41). The data points are the same as in Fig. 1.
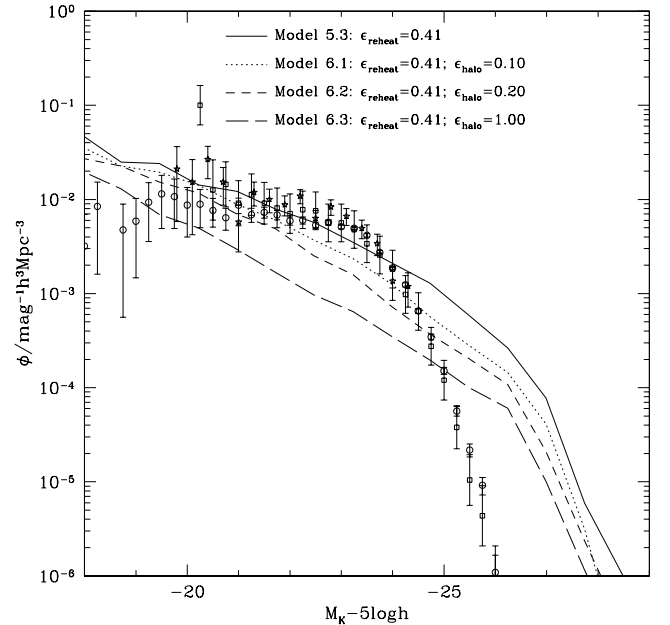

FIG. 3.—Model 6. These models illustrate the effect of energy injection. Starting from model 5.3, the effect of heating the diffuse gas halo is included. Models 6.1–6.3 have $\epsilon_{halo} = 0.1, 0.2$, and 1.0, respectively, in addition to $\epsilon_{reheat} = 0.41$. Increasing the energy available for this form of feedback suppresses the formation of both bright and faint galaxies but does not produce a sharp break in the luminosity function.
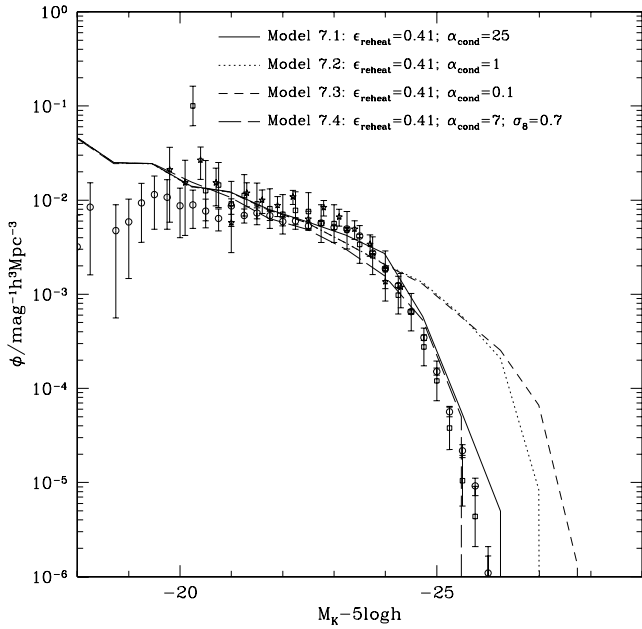
FIG. 4.—Model 7. These models illustrate the effect of thermal conduction. In model 7.3 ($\alpha_{\rm cond} = 25$), conduction is assumed to be highly efficient (it is unlikely that such a high efficiency is physically plausible). More realistic conduction efficiencies are illustrated in models 7.2 ($\alpha_{\rm cond} = 1$) and 7.1 ($\alpha_{\rm cond} = 0.1$). For model 7.4, we adopt a lower value for σ_8; a conduction efficiency of $\alpha_{\rm cond} = 7$ then gives a reasonable match to the observed luminosity function. In all cases, the energy feedback parameters are set to $\epsilon_{\rm halo} = 0.0$ and $\epsilon_{\rm reheat} = 0.41$.

dependence of the Spitzer conductivity rate. Model 7.3 (solid line in Fig. 4) shows the result of including conduction with $\alpha_{\rm cond} = 25$ in model 5.3 ($\epsilon_{\rm reheat} = 0.41$, $\epsilon_{\rm halo} = 0$). Such a high value of the conductivity is indeed effective in suppressing the formation of the most massive galaxies, since it prevents efficient gas cooling in group and cluster-sized halos. The brightest galaxies in the luminosity function are instead built through mergers. The result is a rather good match to the observed galaxy luminosity function. However, the conduction efficiency that we have assumed is extremely high. For the model to operate at the required level, we must assume both that the conductivity is not suppressed below the Spitzer value and that the effective temperature gradient is steeper than $T \propto r^{2.5}$ in the region of the cooling radius. Note that the Spitzer formula for thermal conductivity in an ionized plasma breaks down if the conductivity becomes too high, i.e., if the conduction "saturates," as described by Cowie & McKee (1977). The models shown in this paper do not take account of this saturation limit. However, using the estimate of the saturated heat flux from Cowie & McKee (1977), we have checked that our results are not significantly affected by saturation (for model 7.3, which has the most extreme conduction of all our models, there is only a small increase in the number of the very brightest galaxies).

In models 7.1 and 7.2, we show the effect of assuming a more modest conduction efficiency ($\alpha_{\rm cond} = 1.0$ and 0.1, respectively). In these models, conduction is not sufficient to suppress cooling in the larger halos adequately.

If we adopt a lower value for σ_8, however, a lower conduction efficiency gives a reasonable match to the observed luminosity function. Model 7.4 shows the luminosity function for the case $\sigma_8 = 0.7$ and $\alpha_{\rm cond} = 7$. This conduction

efficiency could be achieved if the temperature gradient was $T \propto r^{1.3}$ and the conduction was only slightly suppressed below the Spitzer value. Although this is still a high rate of conduction, it offers a promising route for explaining the bright end of the galaxy luminosity function.

3.7. *Model with Superwinds*

The expulsion of gas from halos at high energy can, in principle, strongly suppress the formation of later generations of galaxies, hence affecting the shape of the luminosity function. Starting from model 5.2, we add further feedback energy that expels cold gas completely, not only from the disk but also from the halo. The superwind must have high energy in order that the expelled material not be recaptured by more massive halos. The effect of a low-power superwind is illustrated by model 8.1 (dashed line in Fig. 5). This model, with $\epsilon_{\rm sw} = 0.27$ and $\beta_{\rm sw} = 3$, has a relatively weak superwind. This gas expulsion is in addition to the reheating of cold disk gas ($\epsilon_{\rm reheat} = 0.13$); we have assumed that there is no heating of the diffuse halo ($\epsilon_{\rm halo} = 0.0$). Although the winds eject a large amount of gas, most of the material is recaptured as larger halos collapse, and the luminosity function differs little from that of model 5.2.

In model 8.2 (dotted line in Fig. 5), we have set $\epsilon_{\rm sw} = 5.0$ and $\beta_{\rm sw} = 1$, corresponding to a mean energy per superwind particle of $E_{\rm av} = 15$ keV. Such an energetic wind is required to ensure that very little material is recaptured by group halos. In this model, the superwind dominates the feedback energy budget; indeed, the total energy required (5.13×10^{49} ergs $M_\odot^{-1}$) significantly exceeds that available from supernovae alone. The model comes much closer to matching the luminosity function but still overproduces
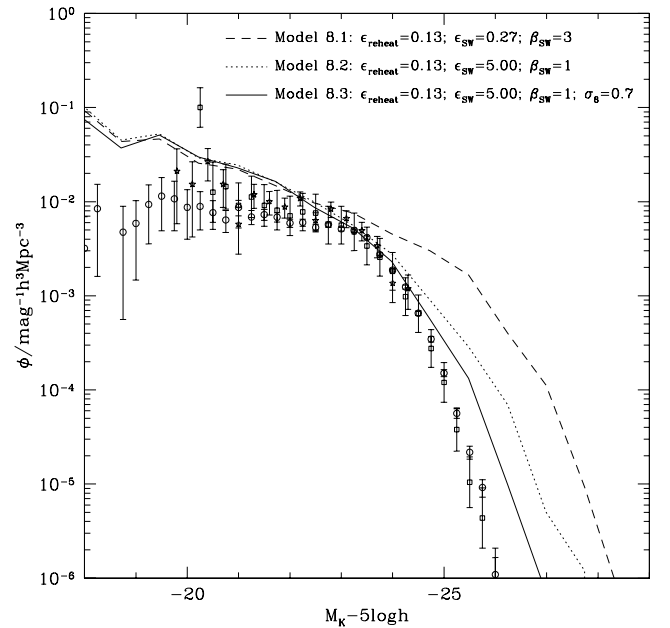

FIG. 5.—Model 8. These models illustrate the effect of superwinds. In model 8.1, an energy of $\epsilon_{\rm sw} = 0.27$ drives a weak superwind (with $\beta_{\rm sw} = 3$); disk reheating has efficiency $\epsilon_{\rm reheat} = 0.13$, and there is no heating of the diffuse halo ($\epsilon_{\rm halo} = 0$). A much more powerful wind is needed to create a break in the luminosity function. Model 8.2 ($\epsilon_{\rm sw} = 5.0$; $\beta_{\rm sw} = 1$) illustrates the effect of increasing the superwind power. An improved match to the luminosity function can be achieved with the same power if σ_8 is lower. Model 8.3 shows a model with $\epsilon_{\rm sw} = 5.0$, $\beta_{\rm sw} = 1$, and $\sigma_8 = 0.7$.

bright galaxies. We can increase the superwind energy further, but a factor of 2 increase only results in a small improvement in the luminosity function. If we include conduction as well as superwinds, it is, of course, possible to improve the match, but a high conduction efficiency ($\alpha_{cond} \gg 1$) is still needed. Increasing the mass loading of the wind substantially (by increasing β_{sw}) results in too few galaxies around the knee of the luminosity function.

As we found in the case of conduction, the luminosity function can be matched more easily if we adopt a lower value for σ_8. The case $\epsilon_{sw} = 5.0$, $\beta_{sw} = 1$, and $\sigma_8 = 0.7$ is illustrated by model 8.3 (solid line in Fig. 5). Given the uncertainties of our recapture prescription, this model gives a reasonable match to the luminosity function; it has a strong break at the correct luminosity and overproduces bright objects only marginally. There are a variety of ways to further improve the match to observations: we could increase the energy injection (but $\epsilon_{sw} > 10.0$ is required) or increase the mass loading of the wind so that the curve is shifted faintward. An alternative strategy is to allow for a low level of conduction: $\alpha_{cond} \sim 1$ is sufficient to produce a significant improvement in the match to the luminosity function when $\sigma_8 = 0.7$ and superwinds are present.

The parameters α_{cond} and ϵ_{sw} are highly degenerate in their effects on the luminosity function (i.e., increasing either of them suppresses the bright end). While current computational limitations make it prohibitively expensive to perform an accurate χ^2 fit of the model parameters to the data, a crude estimate of the χ^2 surface for these two parameters shows that the data prefer models with strong superwinds ($\epsilon_{sw} \approx 6$) and high conductivity ($\alpha_{cond} \approx 30$) for a model with $\sigma_8 = 0.93$. Lowering σ_8 to 0.7 reduces the requirements to $\epsilon_{sw} \approx 3$ and $\alpha_{cond} \approx 25$. Further investigation of the χ^2 surface would require a Bayesian prior to specify formally a physically allowed range for this parameter.

3.8. *Other Considerations*

While our primary goal in this paper is to examine the luminosity function of galaxies, it is prudent to check whether our models are in reasonable agreement with other basic properties of the galaxy population, such as the Tully-Fisher relation and the galaxy autocorrelation function. We compare the models that best fit the galaxy luminosity function to these observables. We retain the model parameters found earlier, and we do not attempt to adjust these or any other parameters in this comparison. We defer a more detailed comparison of our models with a wider range of observational constraints to a future paper.

3.8.1. *Tully-Fisher Relation*

Simultaneously matching the galaxy luminosity function and the Tully-Fisher relation has been a long-standing problem for CDM-based semianalytic models (see, e.g., White & Frenk 1991; Kauffmann et al. 1993; Cole et al. 2000). In particular, Cole et al. (2000) found that while their best model reproduced the observed slope and scatter in the Tully-Fisher relation, the predicted circular velocities for galaxy disks were about 30% larger than the values measured for the data.

Figure 6 shows the Tully-Fisher relation predicted by four of our models (including those that most successfully match the observed luminosity function). It is readily
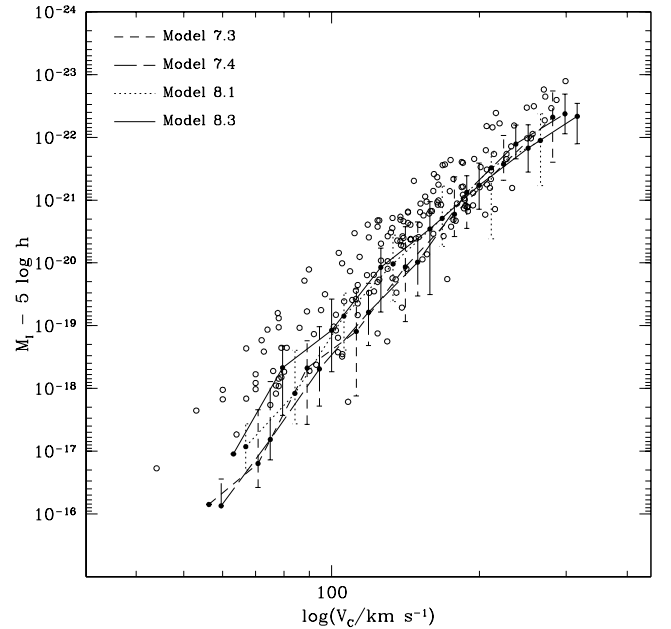

Fig. 6.—*I*-band Tully-Fisher relation of galaxies. Circles show the observational sample of Mathewson, Ford, & Buchhorn (1992). Lines show the median relation for model galaxies. We follow Cole et al. (2000) and select only galaxies with a bulge-to-total ratio in dust-extinguished *I*-band light between 0.02 and 0.24 and a disk gas fraction of at least 10%. Circular velocities for model galaxies are computed at the disk half-mass radius. Error bars enclose the inner 80% of the distribution of model galaxies.

apparent that our models all perform reasonably well, in that they reproduce the slope and scatter of the observed relation, although they do not match the zero point exactly. Model 8.3 does best, being offset by only approximately 10% in circular velocity over most of the range shown. This model, in fact, performs rather better than that of Cole et al. (2000), although an even better match would be desirable. This may involve revising the assumption of adiabatic invariance made in calculating the effects of the disk on the inner structure of the dark matter halo (see Cole et al. 2000) or considering the effects of the angular momentum of the superwind material or the details of our star formation prescription.

3.8.2. *Galaxy Correlation Function*

One of the most remarkable successes of semianalytic models of galaxy formation in the ΛCDM cosmology is their ability to match the observed galaxy autocorrelation function (Kauffmann et al. 1999; Benson et al. 2000). As pointed out by Benson et al. (2000), the clustering of galaxies may be understood physically by reference to the halo occupation distribution (see, e.g., Peacock & Smith 2000; Cooray & Sheth 2002), which specifies the number of galaxies that occupy dark matter halos of a given mass. We show, in Figure 7, the mean number of galaxies brighter than $M_B - 5\log h = -19.5$ per halo as a function of halo mass for four of our models. For reference, we also plot the fits obtained by Berlind et al. (2003) to the semianalytic model of Cole et al. (2000) and to a smoothed particle hydrodynamics calculation of galaxy formation as solid and dotted lines, respectively. The two models considered by Berlind et al. (2003) both gave reasonable fits to the galaxy correlation function.
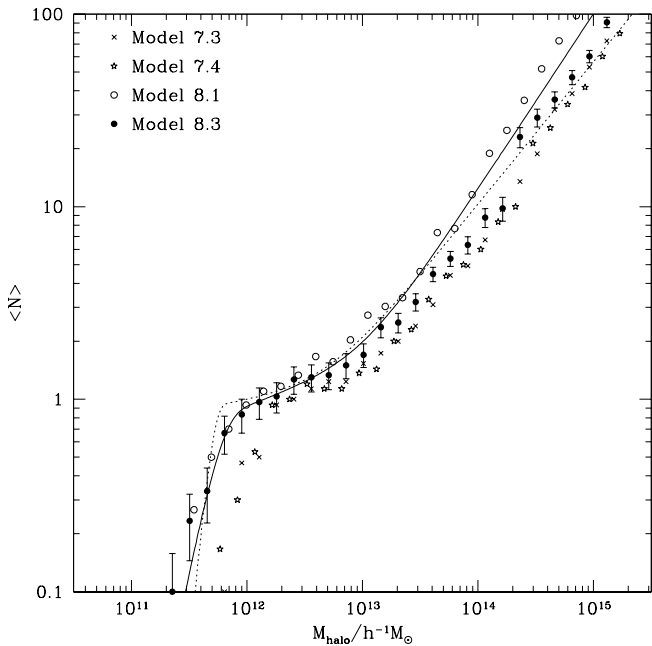
FIG. 7.—Mean number of galaxies brighter than $M_B - 5\log h = -19.5$ as a function of halo mass for some of our models. For reference, lines show the fitting functions of Berlind et al. (2003). The solid line shows their fit to the results of the Cole et al. (2000) semianalytic model, while the dotted line shows their fit to a smoothed particle hydrodynamics calculation.
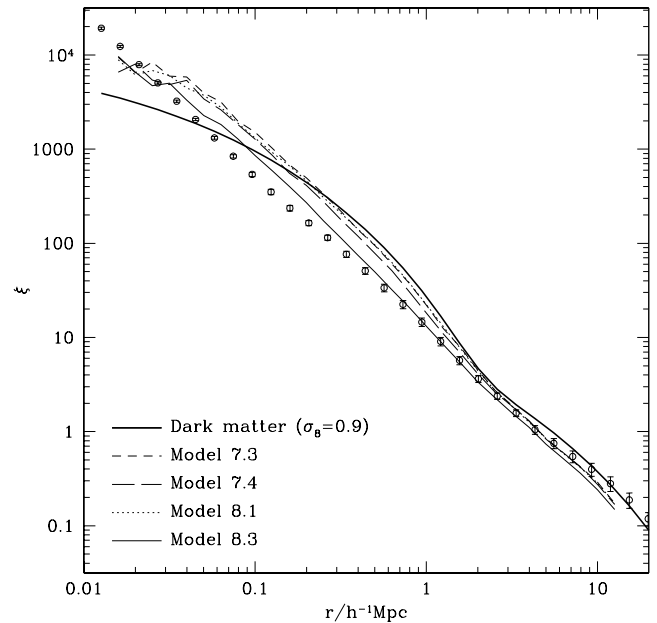

FIG. 8.—Galaxy real-space two-point correlation function. Circles show the observational determination of Baugh (1996). The heavy solid line shows the correlation function of dark matter for our adopted ΛCDM cosmology (with $\sigma_8 = 0.9$). The remaining lines show results for model galaxies brighter than $M_B - 5\log h = -19.5$, as indicated in the key.

Our new models give rise to a mean halo occupancy that is qualitatively similar to those found by Berlind et al. (2003)—a power law at high masses, which flattens before cutting off sharply at some critical mass. Quantitatively, the models with conduction (7.3 and 7.4) produce significantly fewer galaxies in low- and high-mass halos, while agreeing with the Berlind et al. (2003) fits at a halo mass of a few times 10^{12} h^{-1} $M_\odot$. The models with superwinds (8.1 and 8.3) agree more closely with the Berlind et al. (2003) fits, although model 8.3 produces somewhat fewer galaxies in clusters. Since none of our models differ greatly from those of Berlind et al. (2003), we may expect them also to produce correlation functions in reasonable agreement with the observational data, at least on intermediate and large scales. On smaller scales, the mean number of galaxies per halo is not sufficient to specify the correlation function; this requires knowledge of the distribution of galaxy pairs per halo.

Figure 8 shows correlation functions (computed using the techniques of Benson et al. 2000) for the same four models, compared to both the predicted dark matter correlation function and the observed galaxy correlation function. As expected, all models do well at matching the data on scales larger than about 1 h^{-1} Mpc. On smaller scales, all models are "antibiased" with respect to the dark matter, but they produce somewhat too much clustering. Benson et al. (2000) found similar results for their semianalytic model. Model 8.3 gives, in fact, clustering results very similar to theirs (for the same ΛCDM cosmology). The other models produce even stronger correlations on small scales, providing a worse match to the data. Model 8.3 does better than model 8.2 because it places relatively fewer galaxies in clusters, thereby having many fewer pairs on small scales. We speculate that a model that performs even better at

matching the observed luminosity function (e.g., producing a sharper cutoff at bright magnitudes) may also do even better at matching the observed correlation function.

4. DISCUSSION

We have examined the key physics thought to be necessary to explain the shape of the galaxy luminosity function in a CDM universe. While the cooling and condensation of gas in a merging hierarchy of dark matter halos remains the fundamental process through which galaxies form, we have demonstrated that at least two other processes must act to shape the luminosity function. A *minimal* model requires the inclusion of feedback mechanism(s) (beyond the heating resulting from the photoionization of the pregalactic gas) to flatten the faint end of the luminosity function and to suppress cooling at the centers of the massive halos of groups and clusters. If a fraction of the energy liberated by supernovae and winds goes into reheating disk gas and/or heating the diffuse gas halo, then the formation of faint galaxies is suppressed, resulting in a flattened faint-end slope that matches the available observational data adequately. These mechanisms on their own, however, are unable to produce a sharp cutoff at the bright end of the luminosity function. We have shown that there are two possible (but quite extreme) processes that can achieve this: (1) thermal conduction at about or above the Spitzer rate and (2) expulsion of gas from halos in superwinds at temperatures high enough to prevent its subsequent recapture.

The high value of α_{cond} required to suppress the bright end of the luminosity function is discouraging, since it implies both that the conductivity must be close to the Spitzer value (despite the presence of microgauss magnetic fields in clusters; Narayan & Medvedev 2001; Taylor, Fabian, & Allen 2002) and that the effective temperature

gradient must be somewhat steeper than the gradients observed in galaxy clusters (Allen et al. 2001; Fabian et al. 2002).

However, our method for calculating the effect of conduction is highly simplified. In particular, there are two issues that are not well addressed. First, our calculation is based on the heat flux through a shell. We have not considered the total extent of the region in which conduction must be effective. We can define a radius r_{out}, such that the initial thermal energy in the region $r_{cool} < r < r_{out}$ is equal to the total energy radiated from the region $r < r_{cool}$, where r_{cool} is the cooling radius in the presence of conduction. Thermal energy needs to be conducted from a radius at least as large as r_{out} in order to balance the radiative losses from smaller radii. For an isothermal halo with an r^{-2} density profile, we find $r_{cool} r_{out} = r'^2_{cool}$, where r'_{cool} is the cooling radius calculated in the absence of conduction. The radius r_{out} can be large, but the temperature gradient that is required to transport heat effectively across this region is a strongly declining function of radius. Thus, so long as r_{out} lies within the virialized region of the cluster, this is unlikely to modify our solution significantly. Second, our method assumes that the logarithmic temperature gradient within r_{cool} is constant. If, instead, the gradient is peaked in this region, our method would underestimate the heat deposited in the shell at r_{cool}. These two simplifications (which affect the required conduction coefficient in opposite senses) are difficult to model accurately and without ad hoc assumptions. Clearly, this is a problem that needs to be addressed by numerical simulations of cooling in conductive plasmas. These would allow us to calibrate our simple model of conductive heating and to infer what conductivity is required, by removing the degeneracy between the shape of the halo temperature profile and the suppression factor f_{sp} for the conductivity.

The alternative model invokes highly energetic superwind outflows. The expulsion of gas from the halo reduces the reservoir of baryons available for cooling at later times and reduces the number of stars formed by each parcel of cold gas. However, the fraction of baryons required to be in this "superheated" phase at $z = 0$ is quite high, and the energy budget is exceptional. We found the energy budget of a promising superwind model to be in the region of 5×10^{49} ergs per solar mass of stars formed. It is quite possible that our model may not treat the recapture process adequately and that we may therefore be overestimating the total energy requirement. It is difficult to model the process better, however, because the expelled gas may cool adiabatically and then be accreted by a completely different halo; alternatively, some material may escape into void regions and evade recapture altogether. A further problem is the need to assume a particular distribution for the wind particle energies. Our assumption that the distribution is thermal may lead us to overestimate the energy requirement.

On average, the mass of baryons ejected from halos amounts to 7% of the mean cosmic value. This corresponds to about 1.3 times the fraction of baryons turned into stars. A compilation of observations by Martin (1999) suggests that the total mass-loss rate from disk galaxies could certainly be as high as twice the star formation rate. If all of this mass loss were energetic enough to produce a superwind, these observations would provide some justification for the high value of ϵ_{sw} required in our model. The total energy budget, however, exceeds that available from supernovae ($\sim 0.9 \times 10^{49}$ ergs $M_\odot^{-1}$ for the Kennicutt IMF), even though we have not allowed for radiative losses. Although the exact energy output from supernovae is uncertain, and the conditions we have assumed for superwind escape are also uncertain, this suggests that the supernova energy must need to be supplemented by energy released during black hole formation (Ensslin et al. 1998; Bower et al. 2001; Cavaliere, Lapi, & Menci 2002).

Finally, we have briefly examined the ability of some of our most successful models to match the observed Tully-Fisher relation and galaxy correlation function. Without adjusting any parameters, we find, as in many previous semianalytic calculations, that our models adequately match the slope and scatter of the Tully-Fisher relation but tend to overpredict the circular velocity at a given luminosity (or, equivalently, underpredict the luminosity at given circular velocity). In most cases, the agreement is at least as good as in previous studies, but for one of our superwind models the discrepancy in the circular velocity zero point is reduced to only 10%. At least one of our models (8.3) matches the galaxy autocorrelation function reasonably well (as well as the semianalytic model of Benson et al. 2000). Importantly, this success reflects the fact that clusters in this model contain relatively fewer galaxies, and thus many fewer pairs on small scales, than clusters in other models. This demonstrates how measurements of galaxy clustering place important constraints on halo mass–dependent feedback processes, such as those considered in this study. Our results encourage a more detailed investigation of how our models fare when compared with a wider range of observables, a program that we intend to pursue in future work.

5. CONCLUSIONS

One of the simplest properties of the galaxy population, the luminosity function, is now reasonably well determined observationally (although, as we have noted, some significant disagreements about the faint-end slope remain). We have shown that our theoretical understanding of this fundamental property has progressed enormously in recent years, to the extent that it is now possible to explain the form of the luminosity function in quantitative detail. Nevertheless, two crucial ingredients—feedback and conduction—are understood only poorly at best. Further study of these physical processes is required before we can truly claim a satisfactory understanding of the luminosity function of galaxies.

APPENDIX

EJECTION AND RECAPTURE OF SUPERWINDS

The wind is ejected at speed V_{sw} and so has a mean energy per unit mass of $E_{av} = \frac{1}{2} V_{sw}^2$. We assume that to escape the halo it needs energy per unit mass $\lambda_{sw} V_{disk}^2$. Assuming a thermal distribution of energies in the wind [i.e., $f(E) \propto \exp(-E/E_{av})$], the fraction of the mass flowing out of the disk that also escapes from the halo is given by

$$f_{eject} = \int_{x_{sw}}^{\infty} \exp(-x)\, dx = \exp(-x_{sw}) , \tag{A1}$$

where $x = E/E_{av}$ and $x_{sw} = \lambda_{sw} V_{disk}^2 / E_{av}$. For the material that escapes the halo, the mean energy excess over that needed to eject it from the halo is

$$E_{esc} = \frac{E_{av} \int_{x_{sw}}^{\infty} (x - x_{sw}) \exp(-x)\, dx}{\int_{x_{sw}}^{\infty} \exp(-x)\, dx} = E_{av} , \tag{A2}$$

i.e., the mean (kinetic+thermal) energy per unit mass for the gas escaping from the halo is the same as the mean (kinetic+thermal) energy per unit mass for the gas flowing out of the disk (since only the highest energy particles escape). Using, therefore, the same energy distribution, we assume that the fraction of particles recaptured in a halo of virial velocity V_{halo} is

$$f_{cap} = \int_{0}^{x_{cap}} \exp(-x)\, dx = 1 - \exp(-x_{cap}) , \tag{A3}$$

where $x_{cap} = C_{cap} V_{halo}^2 / E_{av}$ and C_{cap} is a numerical coefficient of order unity. We have used $C_{cap} = 1$.

REFERENCES

Adelberger, K. L., Steidel, C. C., Shapley, A. E., & Pettini, M. 2003, ApJ, 584, 45
Allen, S. W., Schmidt, R. W., Fabian, A. C., & Ebeling, H. 2003, MNRAS, 342, 287
Allen, S. W., et al. 2001, MNRAS, 324, 842
Bacon, D. J., Massey, R. J., Refrigier, A. R., & Ellis, R. S. 2003, MNRAS, 344, 673
Baugh, C. M. 1996, MNRAS, 280, 267
Benson, A. J., Cole, S., Frenk, C. S., Baugh, C. M., & Lacey, C. G. 2000, MNRAS, 311, 793
Benson, A. J., Lacey, C. G., Baugh, C. M., Cole, S., & Frenk, C. S. 2002, MNRAS, 333, 156
Benson, A. J., Pearce, F. R., Frenk, C. S., Baugh, C. M., & Jenkins, A. 2001, MNRAS, 320, 261
Berlind, A. A., et al. 2003, ApJ, 593, 1
Blanton, M. R., et al. 2003, ApJ, 592, 819
Böhringer, H., Matsushita, K., Churazov, E., Ikebe, Y., & Chen, Y. 2002, A&A, 382, 804
Bower, R. G., Benson, A. J., Lacey, C. G., Baugh, C. M., Cole, S., & Frenk, C. S. 2001, MNRAS, 325, 497
Brüggen, M., Kaiser, C. R., Churazov, E., & Ensslin, T. A. 2002, MNRAS, 331, 545
Cavaliere, A., Lapi, A., & Menci, N. 2002, ApJ, 581, L1
Cole, S. 1991, ApJ, 367, 45
Cole, S., Aragón-Salamanca, A., Frenk, C. S., Navarro, J. F., & Zepf, S. E. 1994, MNRAS, 271, 781
Cole, S., Lacey, C. G., Baugh, C. M., & Frenk, C. S. 2000, MNRAS, 319, 168
Cole S., et al. 2001, MNRAS, 326, 255
Cooray, A., & Sheth, R. 2002, Phys. Rep., 372, 1
Cowie, L. L., & McKee, C. F. 1977, ApJ, 211, 135
Efstathiou, G. 2000, MNRAS, 317, 697
Eke, V. R., Cole, S. M., & Frenk, C. S. 1996, MNRAS, 282, 263
Eke, V. R., Navarro, J. F., & Frenk, C. S. 1998, ApJ, 503, 569
Ensslin, T. A., Wang, Y., Nath, B. B., & Biermann, P. L. 1998, A&A, 333, L47
Ettori, S., De Grandi, S., & Molendi, S. 2002, A&A, 391, 841
Fabian, A. C., Mushotzky, R. F., Nulsen, P. E. J., & Peterson, J. R. 2001, MNRAS, 321, L20
Fabian, A. C., Voigt, L. M., & Morris, R. G. 2002, MNRAS, 335, L71
Gruzinov, A. 2002, preprint (astro-ph/0203031)
Helly, J. C., Cole, S., Frenk, C. S., Baugh, C. M., Benson, A., Lacey, C., & Pearce, F. R. 2003, MNRAS, 338, 913
Hoekstra, H., Yee, H. K. C., & Gladders, M. D. 2002, ApJ, 577, 595
Huang, J.-S., Glazebrook, K., Cowie, L. L., & Tinney, C. 2003, ApJ, 584, 203
Jarvis, M., Bernstein, G. M., Fischer, P., Smith, D., Jain, B., Tyson, J. A., & Wittman, D. 2003, AJ, 125, 1014
Jenkins, A., Frenk, C. S., White, S. D. M., Colberg, J. M., Cole, S., Evrard, A. E., Couchman, H. M. P., & Yoshida, N. 2001, MNRAS, 321, 372

Johnstone, R. M., Allen, S. W., Fabian, A. C., & Sanders, J. S. 2002, MNRAS, 336, 299
Kauffmann, G., Colberg, J. M., Diaferio, A., & White, S. D. M. 1999, MNRAS, 303, 188
Kauffmann, G., White, S. D. M., & Guiderdoni, B. 1993, MNRAS, 264, 201
Kay, S. T., Pearce, F. R., Frenk, C. S., & Jenkins, A. 2002, MNRAS, 330, 113
Kennicutt, R. C., Jr. 1983, ApJ, 272, 54
Kochanek, C. S., et al. 2001, ApJ, 560, 566
Lacey, C., Guiderdoni, B., Rocca-Volmerange, B., & Silk, J. 1993, ApJ, 402, 15
Martin, C. L. 1999, ApJ, 513, 156
Mathewson, D. S., Ford, V. L., & Buchhorn, M. 1992, ApJS, 81, 413
McNamara, B. R., et al. 2001, ApJ, 562, L149
Melchiori, A., Bode, P., Bahcall, N. A., & Silk, J. 2003, ApJ, 586, L1
Murali, C., Katz, N., Hernquist, L., Weinberg, D. H., & Davé, R. 2002, ApJ, 571, 1
Narayan, R., & Medvedev, M. V. 2001, ApJ, 562, L129
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, ApJ, 462, 563
———. 1997, ApJ, 490, 493
Nulsen, P. E. J., David, L. P., McNamara, B. R., Jones, C., Forman, W. R., & Wise, M. 2002, ApJ, 568, 163
O'Meara, J. M., Tytler, D., Kirkman, D., Suzuki, N., Prochaska, J. X., Lubin, D., & Wolfe, A. M. 2001, ApJ, 552, 718
Peacock, J. A. 2003, in AIP Conf. Proc. 666, The Emergence of Cosmic Structure, ed. S. S. Holt & C. S. Reynolds (New York: AIP), 275
Peacock, J. A., & Smith, R. 2000, MNRAS, 318, 1144
Pearce, F. R., Jenkins, A., Frenk, C. S., White, S. D. M., Thomas, P. A., Couchman, H. M. P., Peacock, J. A., & Efstathiou G. 2001, MNRAS, 326, 649
Peterson, J. R., et al. 2001, A&A, 365, L104
Pettini, M., Rix, S. A., Steidel, C. C., Adelberger, K. L., Hunt, M. P., & Shapley, A. E. 2002, ApJ, 569, 742
Quilis, V., Bower, R. G., & Balogh M. L. 2001, MNRAS, 328, 1091
Rees, M. J., & Ostriker, J. P. 1977, MNRAS, 179, 541
Shu, C., Mo, H. J., & Mao, S. 2003, MNRAS, submitted (astro-ph/0301035)
Sievers, J. L., et al. 2003, ApJ, 591, 599
Smith, G. P., Edge, A. C., Eke, V. R., Nichol, R. C., Smail, I., & Kneib, J.-P. 2003, ApJ, 590, L79
Somerville, R. S., & Primack, J. R. 1999, MNRAS, 310, 1087
Spergel, D., et al. 2003, ApJS, 148, 175
Spitzer, L. 1962, Physics of Fully Ionized Gases (New York: Wiley)
Springel, V., & Hernquist, L. 2003, MNRAS, 339, 289
Strickland, D. K., & Stevens, I. R. 2000, MNRAS, 314, 511
Tamura, T., et al. 2001, A&A, 365, L87
Taylor, G. B., Fabian, A. C., & Allen, S. W. 2002, MNRAS, 334, 769
van Kampen, E., Jimenez, R., & Peacock, J. A. 1999, MNRAS, 310, 43
Viana, P. T. P., Nichol, R. C., & Liddle A. R. 2002, ApJ, 569, L75

Voigt, L. M., Schmidt, R. W., Fabian, A. C., Allen, S. W., & Johnstone, R. M. 2002, MNRAS, 335, L7

Voit, G. M., Bryan, G. L., Balogh, M. L., & Bower, R. G. 2002, ApJ, 576, 601

Wang, X., Tegmark, M., Jain, B., & Zaldarriaga, M. 2002, preprint (astro-ph/0212417)

Weinberg, D. H., Davé, R., Katz, N., & Hernquist, L. 2003, ApJ, in press

White, S. D. M., & Frenk, C. S. 1991, ApJ, 379, 52

White, S. D. M., & Rees, M. J. 1978, MNRAS, 183, 341

Yoshida, N., Stoehr, F., Springel, V., & White, S. D. M. 2002, MNRAS, 335, 762