THE TUMULTUOUS LIVES OF GALACTIC DWARFS AND THE MISSING SATELLITES PROBLEM

ANDREY V. KRAVTSOV, OLEG Y. GNEDIN, AND ANATOLY A. KLYPIN

ABSTRACT

Hierarchical cold dark matter (CDM) models predict that Milky Way–sized halos contain several hundred dense low-mass dark matter satellites (the substructure), an order of magnitude more than the number of observed satellites in the Local Group. If the CDM paradigm is correct, this prediction implies that the Milky Way and Andromeda are filled with numerous dark halos. To understand why these halos failed to form stars and become galaxies, we need to understand their history. We analyze the dynamical evolution of the substructure halos in a high-resolution cosmological simulation of Milky Way–sized halos in the CDM cosmology. We find that about 10% of the substructure halos with the present masses $\lesssim 10^8 - 10^9 M_\odot$ (circular velocities $V_{\text{vir}} \lesssim 30$ km s$^{-1}$) had considerably larger masses and circular velocities when they formed at redshifts $z \gtrsim 2$. After the initial period of mass accretion in isolation, these objects experience dramatic mass loss because of tidal stripping. Our analysis shows that strong tidal interaction is often caused by actively merging massive neighboring halos, even before the satellites are accreted by their host halo. These results can explain how the smallest dwarf spheroidal galaxies of the Local Group were able to build up a sizable stellar mass in their seemingly shallow potential wells. We propose a new model in which all the luminous dwarf spheroidal galaxies in the Local Group are descendants of the relatively massive ($\gtrsim 10^9 M_\odot$) high-redshift systems, in which the gas could cool efficiently by atomic line emission, and which were not significantly affected by the extragalactic ultraviolet radiation. We present a simple galaxy formation model based on the trajectories extracted from the simulation, which accounts for the bursts of star formation after strong tidal shocks and the inefficiency of gas cooling in halos with virial temperatures $T_{\text{vir}} \lesssim 10^4$ K. Our model reproduces the abundance, spatial distribution, and morphological segregation of the observed Galactic satellites. The results are insensitive to the redshift of reionization.

Subject headings: cosmology: theory — galaxies: dwarf — galaxies: evolution — galaxies: formation — galaxies: halos — methods: numerical

1. INTRODUCTION

Semianalytic models of galaxy formation (Kauffmann et al. 1993; Bullock et al. 2000; Somerville 2002; Benson et al. 2002) and numerical simulations (Klypin et al. 1999b; Moore et al. 1999a) have convincingly showed that the expected number of dark matter clumps around the galactic Milky Way (MW)–sized halos exceeds the observed number of satellites by an order of magnitude. The discrepancy may indicate that the amplitude of the small-scale primordial density fluctuations is considerably lower than expected in the cold dark matter (CDM) scenarios (e.g., Kamionkowski & Liddle 2000; Zentner & Bullock 2003) or that dark matter is self-interacting (Spergel & Steinhardt 2000). An alternative “astrophysical” interpretation is that the mismatch indicates that galaxy formation in dwarf halos is inefficient.

Several plausible physical processes may suppress gas accretion and star formation in dwarf dark matter (DM) halos. The cosmological UV background, which reionized the universe at $z \gtrsim 6$, heats the intergalactic gas and establishes a characteristic time-dependent minimum mass for halos that can accrete gas (e.g., Efstathiou 1992; Thoul & Weinberg 1996; Quinn et al. 1996; Navarro & Steinmetz 1997; Gnedin & Hui 1998; Kitayama & Ikeuchi 2000; Gnedin 2000; Dijkstra et al. 2004). The gas in the low-mass halos may be photoevaporated after reionization (Barkana & Loeb 1999; Shaviv & Dekel 2003; Shapiro et al. 2004). In particular, Shaviv & Dekel (2003) recently argued that halos with circular velocities of up to $\sim 30$ km s$^{-1}$ can be photoevaporated by the UV background. At the same time, the ionizing radiation may quickly dissociate molecular hydrogen, the only efficient coolant for low-metallicity gas in such halos, and prevent star formation before the gas is completely removed (e.g., Haiman et al. 1997).

The combined effect of these processes is likely to leave all DM halos with masses $\lesssim$ few times $10^9 M_\odot$ dark. This is consistent with current observational constraints that indicate that halos with $M < 10^{10} M_\odot$ are virtually devoid of galaxies (van den Bosch et al. 2003). It is thus remarkable that the dynamical masses of some of the Local Group (LG) dwarfs are only $\sim 10^7 M_\odot$ (Mateo 1998). How could such galaxies form stars despite the suppressing processes listed above?

One possibility is that they manage to accrete a certain amount of gas before the universe is reionized (Bullock et al. 2000), with the implicit assumption that this gas can be subsequently converted to stars. However, it is likely that gas cooling and star formation in such small systems is inefficient. For example, cosmological simulations with a self-consistent treatment of H$_2$ chemistry and radiative transfer indicate that star formation is strongly suppressed in halos with masses $M \lesssim 5 \times 10^8 M_\odot$ at all redshifts, even before reionization (Chiu et al. 2001). In addition, the galaxies may not be able to form sufficiently early to accrete the gas in the first place if the power spectrum normalization is low or the universe was

1 Department of Astronomy and Astrophysics, Center for Cosmological Physics, 5640 South Ellis Avenue, University of Chicago, Chicago, IL 60637; andrey@oddjob.uchicago.edu.
2 Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218; ognedin@stsci.edu.
3 Department of Astronomy, New Mexico State University, MSC 4500, P.O. Box 30001, Las Cruces, NM 88003-8801; aklypin@nmsu.edu.
DWARFS AND MISSING SATELLITES

reionized early, as indicated by the first-year *Wilkinson Microwave Anisotropy Probe* (WMAP) results (Spergel et al. 2003; Kogut et al. 2003). An alternative proposal was recently suggested by Stoehr et al. (2002, 2003) and corroborated by Hayashi et al. (2003), who argued that the host halos of the low-luminosity dwarf spheroidal galaxies may be considerably more massive than previously thought. In this case, the large halo mass could allow an object to resist the suppressing effects of UV background.

In this paper we study the dynamical evolution of dwarf satellite halos around the MW-sized hosts in self-consistent cosmological simulations. We show that the evolution of such objects is complex and often involves dramatic tidal stripping, interactions with other satellites, mass loss, and changes of internal structure. Most importantly, we find that some of the satellites that have small masses and circular velocities at the present were once considerably more massive and could have plausibly formed stars in the past. We argue that the evolution of these objects may explain how the smallest dwarfs in the LG managed to form their stellar populations.

The paper is organized as follows. In § 2 we describe the details of the numerical simulation used in our analysis. In § 3 and § 4 we discuss the algorithm used to identify halos and the method used to construct their evolutionary tracks. In § 5 we present the main results for the dynamical evolution, abundance, and radial distribution of the DM halos. In § 6 we present a model for star formation in these systems and compare results with the observed abundance and spatial distribution of the LG dwarfs. We discuss the implications of our results and compare our model with the previous studies in §§ 7 and 8. Finally, in § 9 we summarize our findings and conclusions.

2. SIMULATION

We used the Adaptive Refinement Tree N-body code (ART; Kravtsov et al. 1997; Kravtsov 1999) to follow the evolution of three MW-sized halos in the concordance ΛCDM cosmology: $\Omega_m, \Omega_\Lambda, h, \sigma_8 = (0.3, 0.7, 0.7, 0.9)$. The simulation starts with a uniform 256$^3$ grid covering the entire computational box. This grid defines the lowest (zeroth) level of resolution. Higher force resolution is achieved in the regions corresponding to collapsing structures by recursive refining of all such regions by using an adaptive refinement algorithm. Each cell can be refined or de-refined individually. The cells are refined if the particle mass contained within them exceeds a certain specified value. The grid is thus refined to follow the collapse objects in a quasi-Lagrangian fashion.

The galactic halos were simulated in the comoving box of 25 h$^{-1}$ Mpc; they were selected to reside in a well-defined filament at $z = 0$. Two halos are neighbors, located 425 h$^{-1}$ kpc (i.e., $\approx 610$ kpc $\sim 2R_{\text{vir}}$) from each other. The configuration of this pair thus resembles that of the LG. The third halo is isolated and is located $\sim 2$ Mpc away from the pair.

A multiple mass resolution technique was used to set up initial conditions. Namely, a Lagrangian region corresponding to a sphere of radius equal to 2 virial radii around each halo was resampled with the highest resolution particles of mass $m_p = 1.2 \times 10^6$ h$^{-1} M_\odot$, corresponding to 1024$^3$ particles in the box, at the initial redshift of the simulation ($z_i = 50$). The high mass resolution region was surrounded by layers of particles of increasing mass with a total of five particle species. Only regions containing the highest resolution particles were adaptively refined, and the threshold for refinement was set to correspond to the mass of the four highest resolution particles. The maximum of 10 refinement levels was reached in the simulations corresponding to the peak formal spatial resolution of 150 comoving pc. Each host halo is resolved with $\sim 10^5$ particles within its virial radius at $z = 0$.

From this point, we refer to the isolated halo as G$_1$ and the halos in the pair as G$_2$ and G$_3$. These halos are called B$_1$, C$_1$, and D$_1$, respectively, in Klypin et al. (2001), and we refer the reader to their paper for further details. The main properties of these three host halos, namely, the virial mass, radius, and maximum circular velocity, are given in Table 1. We choose to define the virial radius (and the corresponding virial mass) as the radius encompassing the density of 180 times the mean density of the universe. For the commonly used overdensity of 340, the virial radius and masses for G$_1$, G$_2$, and G$_3$ are $R_{\text{vir}} = 231, 212, \text{and } 213$ h$^{-1}$ kpc and $M_{\text{vir}} = 1.45 \times 10^{12}, 1.13 \times 10^{12}, \text{and } 1.14 \times 10^{12}$ h$^{-1} M_\odot$, respectively. The masses are in the range of possible MW halo masses (Klypin et al. 2002).

Figure 1 shows the mass aggregation history of the three host halos. They have similar masses at present but rather different evolutionary histories. In all cases, there is a period of very rapid mass assembly at $z \approx 2-3$, followed by a relatively quiescent accretion at $z \approx 1.5$, the behavior typical of particles. The maximum of 10 refinement levels was reached in the simulations corresponding to the peak formal spatial resolution of 150 comoving pc. Each host halo is resolved with $\sim 10^5$ particles within its virial radius at $z = 0$.

![FIG. 1.—Mass aggregation histories for the three MW-sized host halos analyzed in this study.](image-url)
hierarchically forming halos (Wechsler et al. 2002). Host G\textsubscript{1} undergoes a spectacular multiple major merger at \( z \approx 2 \), which results in a dramatic mass increase on a dynamical timescale. Halos G\textsubscript{2} and G\textsubscript{3} increase their mass in a series of somewhat less spectacular major mergers that could be seen as mass jumps at \( 5 < z < 1 \). All three systems accrete little mass and experience no major mergers at \( z \leq 1 \) (or look-back time of \( \approx 8 \) Gyr) and thus could host a disk galaxy. Note, however, that halos G\textsubscript{1} and G\textsubscript{3} experience minor mergers during this period.

3. HALO IDENTIFICATION

In this study we use a variant of the bound density maxima (BDM; Klypin et al. 1999a) halo-finding algorithm to identify halos both within (subhalos) and outside the host halos. Throughout this paper we use the terms “subhalo,” “substructure,” and “satellite” interchangeably to indicate the distinct gravitationally self-bound halos located within the virial radius of a larger halo, which we call the “host.” The division is illustrated in Figure 2.

The BDM algorithm first finds positions of local maxima in the density field smoothed on a certain scale. Starting with the highest overdensity particle, we surround each potential density maximum by a sphere of radius \( r_{\text{find}} = 10 \, h^{-1} \) kpc and exclude all particles within this sphere from further search. The search radius is defined by the size of the smallest systems we aim to identify. We verified that the results do not change if this radius is decreased by a factor of up to 4. After all potential halo centers are identified, we analyze the density distribution and velocities of surrounding particles to test whether the center corresponds to a gravitationally bound clump. Specifically, we construct the density, circular velocity, and velocity dispersion profiles around each center and iteratively remove unbound particles (see Klypin et al. 1999a for details). We then construct final profiles using only bound particles and use them to calculate such halo properties as the maximum circular velocity \( V_m \), mass \( M \), etc.

The virial radius is meaningless for the subhalos within a larger host as their outer layers are tidally stripped, and the extent of the halo is truncated. The definitions of the outer boundary of a subhalo and its mass are thus somewhat ambiguous. We adopt the truncation radius, \( r_t \), at which the logarithmic slope of the density profile constructed from the bound particles becomes larger than \( -0.5 \) as we do not expect the density profile of the CDM halos to be flatter than this slope. Empirically, this definition roughly corresponds to the radius at which the density of the gravitationally bound particles is equal to the background host halo density, albeit with a large scatter. For some halos \( r_t \) is larger than their virial radius.
radius. In this case, we set $r_t = R_{\text{vir}}$. Throughout this paper, we denote the minimum of the virial mass and mass within $r_t$ simply as $M$. For each halo we also construct the circular velocity profile $V_c(r) = \frac{GM(<r)}{r}^{1/2}$ and compute the maximum circular velocity profile $V_m$.

Figure 2 shows the particle distribution in the halo G1 at $z = 0$ along with the halos (circles) identified by the halo finder. The particles are color coded on a gray scale according to the logarithm of their density to enhance visibility of substructure clumps. The radius of the largest circle indicates the actual virial radius, $R_{\text{vir}}$, of the host halo ($R_{\text{vir}} = 298 \, h^{-1} \, \text{kpc}$); the radii of the other halos are the minimum of the truncation radius $r_t$ and $R_{\text{vir}}$. The figure demonstrates that the algorithm is efficient in identifying the substructure down to small masses.

4. CONSTRUCTING TRAJECTORIES

The halo finder described above was run at the 96 saved epochs between $z = 10$ and 0 with a typical spacing of $\sim(1-2) \times 10^8$ yr between outputs. For each epoch, the halo finder produced a halo catalog with positions, velocities, radii $r_h = \min(r_t, R_{\text{vir}})$, masses $m(<r_h)$, maximum of the circular velocity profile $V_m$, and the radius at which the maximum occurs, $r_{\text{max}}$. In addition, for each halo we save indices of all gravitationally bound DM particles located within $r_h$.

This information is used to identify the progenitors of halos at successive epochs. Specifically, for a current epoch $z_i$, starting at $z = 0$, we search progenitors for each halo at several previous epochs $z_{i-j}$ as follows. First, we select a given fraction, $f_{\text{bd}}$, of the most bound particles of the halos at the epochs of consideration. We then compare the fraction of these particles that is common to all pairs of halos at successive epochs and assume that the halo with the highest common fraction is the progenitor. The trajectories used in this study were constructed using $f_{\text{bd}} = 0.25$. As the halo catalogs may miss some halos, especially near the completeness limit of the simulation, if the progenitor is not found at the previous epoch, we need to search at the earlier epoch, etc. In particular, if a halo is located within the search radius $r_{\text{sr}}$ of some larger
system, it will not be identified by the halo finder. In constructing the trajectories we search for progenitors of a halo at \( z \) at epochs up to \( z_{\text{max}} - 4 \). In the dominant majority of cases, the progenitors are found at the previous epoch, \( z_{\text{max}} - 1 \). We experimented with other algorithms for progenitor identification and found the adopted prescription to be the most reliable and efficient.

5. RESULTS

5.1. Tidal Stripping and Dynamical Evolution of Satellite Halos

Figure 3 shows three examples of the evolution of satellite halos. In Figure 3 (middle row) we plot the tidal force experienced by each object. The force was calculated both directly from the gravitational potential field computed in the simulation and analytically from the neighbor halo catalogs, as described in Appendix A. The figure shows the trace of the tidal tensor, \( F_{i,d} = \sum \alpha \alpha \alpha \), which is a good measure of the overall tidal field.

In all cases the tidal force experienced by the satellite coincides approximately with the time when the object is closest to the host, as expected. The figure shows, for example, that at later epochs the tidal force calculated directly using the potential from the simulation can be well approximated by the analytical force from the host halo (see eq. [A4]). However, at earlier epochs (e.g., the highest peak in the left column) the force from the host underestimates the total tidal force. Thus, the overall tidal stripping is produced not only by the host halo but also by the massive neighbor halos, even before the host is formed (Gnedin 2003b).

The tidal heating by multiple halos is similar to “galaxy harassment” in clusters of galaxies (Moore et al. 1996, 1999b), except that it may occur when the halo is still isolated. As Figure 3 shows, the true force computed from the potential can be recovered if the analytical contributions of all neighboring halos are included. Their contribution is particularly important during major mergers of the host, when the centers of two or more massive halos are located in the close vicinity of each other and the satellite halos. Our analytic estimate describes the strong tidal peaks remarkably well but becomes inaccurate for low (a few Gyr \(^{-2} \)) values of \( F_{i,d} \).

The amount of energy imparted to the halo depends on the square of the tidal force (eq. [A6]). Thus, by far the strongest tidal heating experienced by an object is during the highest tidal peaks. For the object in the left column of Figure 3, for example, most of the stripping and disruption is due to the tidal peak at \( t \sim 4 \) Gyr \((z \sim 1.5)\). At this epoch, the host halo is not yet fully assembled and is undergoing a major merger with three other massive halos. It is at this epoch, however, that the satellite experiences the most dramatic tidal mass loss. Subsequent tidal peaks result in only a relatively mild stripping. The object in Figure 3 (middle) also suffers a dramatic mass loss at \( t \sim 4-5 \) Gyr. In this case, however, the efficient tidal stripping continues because of the later pericentric passages and associated peaks in the tidal force. Finally, the satellite shown in the figure experiences only a relatively mild tidal stripping. This satellite orbits in the outer regions of the host and never reaches the central \( \approx 60 \) kpc.

Note that the pericenter of the third satellite is larger at the late epochs compared with the pericenter at \( t \approx 4 \) Gyr. This is contrary to a naive expectation that the pericenter should stay constant or decrease with time if dynamical friction is efficient. The real situation is clearly more complicated. The satellite can lose as well as gain the orbital energy. The latter can occur via a three-body interaction. Indeed, in examining individual trajectories, we found cases in which a satellite gains orbital energy via the “slingshot” acceleration, a classic three-body interaction.

Figure 3 demonstrates that some satellites with small maximum circular velocity and mass at \( z = 0 \) were substantially more massive during the early stages of their evolution. The mass of the object in the middle column is \( \approx 10^{10} \) \( M_\odot \) and its circular velocity is more than 40 km s\(^{-1} \) at \( t = 4 \) Gyr. At the present epoch, they are only \( 2 \times 10^{8} \) \( M_\odot \) and 18 km s\(^{-1} \) respectively. In the extreme cases we find changes of mass and \( V_{\text{max}} \) by a factor of 200 and 8, respectively (see Fig. 6).

At the same time, the object in the right column of Figure 3 has a considerably larger pericenter and experiences weaker tidal force by more than an order of magnitude. Consequently, its mass and circular velocity change little during the evolution. What is the relative frequency of such cases compared with the cases of dramatic mass loss? We address this question in the next section.

5.2. Internal Structure Evolution

Strong tidal forces experienced by orbiting halos lead to a substantial mass loss, preferentially at the outer radii. The changes in the inner regions are more subtle and occur at a slower rate but can nevertheless be significant (e.g., Klypin et al. 1999a; Hayashi et al. 2003; Stoehr et al. 2003; Kazantzidis et al. 2004b). Figure 4 shows the maximum values of \( M \) and \( V_{\text{max}} \) reached by a satellite during its evolution versus their present values. Most of the surviving satellites experience only mild evolution, less than a factor of 2 in \( V_{\text{m}} \). Yet, there is a fair number of cases in which the evolution is significant. The average changes in \( M \) or \( V_{\text{m}} \) do not seem to depend on the halo mass.

Figure 5 shows the ratio of the mass at \( z = 0 \) to the maximum mass achieved by each satellite during its evolution versus the ratio of the maximum circular velocities at these two epochs. The figure shows a strong correlation between the two ratios:

\[
\frac{M^0}{M_{\text{max}}} = \left( \frac{V^0_m}{V_{\text{max}}^m} \right)^{\beta}, \quad \beta \approx 3-4,
\]

where \( V^0_m \) and \( M^0 \) are the values at the present and \( V_{\text{max}}^m \) is the maximum circular velocity at the epoch when the halo reached the maximum mass, \( M_{\text{max}} \). This correlation shows that the internal structure of satellites readjusts as they lose mass because of tidal stripping. Note that the decrease of \( V_{\text{m}} \) indicates the decrease in density within the inner radius of \( \approx 2.16 r_s \), where \( r_s \) is the NFW scale radius (Navarro et al. 1997). The adjustment is such that the virial correlation,

\[
M \propto V_{\text{m}}^\alpha, \quad \alpha \approx 3-4,
\]

is approximately maintained at all times.

This can be seen in Figure 6, which shows the tracks of individual satellites in the \( M-V_{\text{m}} \) plane. The satellites shown were selected from all three galactic hosts. We selected objects with large changes in mass to maximize the dynamic range. The figure shows that both during the periods of mass growth while evolving in isolation and the periods of mass decrease because of tidal stripping, halos approximately move up and down the power-law dependence of equation (2). For instance, the track shown in Figure 6 (top left) starts at \( M \approx 3 \times 10^8 h^{-1} \)
present-day value $V_m$ and $V_z$ are changing with redshift. The same zero point of the relation can be maintained only if both $M_{\text{vir}}$ and $V_m$ are changing with time in a certain way, specified above.

During the second stage of evolution, the subhalo experiences tidal forces from the host and other halos, and its mass and extent are tidally truncated. The average density $\rho_0$ within the truncation radius $R_t$ is approximately constant along the orbit and is proportional to the background density of the host halo at the pericenter of the subhalo’s orbit. The truncation radius scales with the truncated mass, $M_t$, as $R_t \propto (M_t/\rho_0)^{1/3}$, so $M_t \propto V_m^{3/2}$. The velocity $V_m$ is constant along the satellite trajectory, we obtain the following relation: $M_t \propto V_m^{3/2}$.

5.3. Evolution of Halos in the $M-V_m$ Plane

In the previous section we showed that individual halos maintain the $M \propto V_m^n$ relation during their evolution. This explains why the same correlation between the mass and $V_m$ holds for both subhalos and isolated halos (Avila-Reese et al. 1999; Bullock et al. 2001). We also find that the mass–circular velocity relations for the subhalos at $z = 0$ and for their progenitors at the epoch when the maximum mass was reached have similar amplitudes and slopes ($\approx 3.3$).

In this section we consider the mechanism behind such behavior in more detail. We can interpret the observed evolution of mass and $V_m$ for a given subhalo by dividing it into the following two stages: (1) mass growth while evolving in isolation and (2) mass decrease due to tidal stripping after the halo is accreted by its host. The transition between the two stages typically occurs at $z \approx 2$ for the mass range of subhalos and hosts considered here.

We fit the slope $\alpha$ for the trajectories in the $M-V_m$ plane for all satellites separately in the two regimes. We fit only tracks of halos with $V_m > 15$ km s$^{-1}$ at $z = 0$ and with at least 10 redshift outputs. In calculating the average slope, $\alpha_0$, and the dispersion of the sample, $\sigma_\alpha$, we weight $\alpha$ of each halo by the error of the fit. The isolated halos have the average slope $\alpha_0 = 4.7$, with the dispersion $\sigma_\alpha = 1.2$. The truncated halos have $\alpha_0 = 2.9$ and $\sigma_\alpha = 1.2$. Thus, the slopes in the two regimes seem to be somewhat different.

These different slopes can be linked to the different average densities of the dwarf halos in the two regimes. In the mass growth stage, when the average density of the universe is $\bar{\rho}(z) = \rho_0(1 + z)^3$, the mass and velocity are given by the virial scaling relation $M_{\text{vir}} \propto V_{\text{vir}}^3(1 + z)^{-3/2}$. Also, Bullock et al. (2001) showed that, as long as the NFW model is an adequate description of the halo profile, $V_m$ and $V_{\text{vir}}$ are related through the concentration parameter approximately as $V_m/V_{\text{vir}} \propto c_{\text{vir}}^{-1/4}$. The median concentration itself varies with the mass and redshift as $c_{\text{vir}} \propto (1 + z)^{-1}M_{\text{vir}}^{-0.13}$. Since all the variables scale as some power of $1 + z$, it is natural to approximate the evolution of the mass and maximum velocity as $M_{\text{vir}} \propto (1 + z)^{-q}$ and $V_m \propto (1 + z)^{-p}$, with the above relations leading to $(29/32)q = 3p + (3/4)$. The slope $\alpha_0 = q/p = 4.7$ is achieved for $q = 2.8$ and $p = 0.6$, although the scatter in the value of $\alpha$ implies a corresponding scatter in the exponents $q$ and $p$.

Note that the slope of the $M_{\text{vir}}-V_m$ relation is steeper than the virial $\alpha \approx 3$ because the virial parameters of isolated halos depend on the mean density of the universe and that density is changing with redshift. The same zero point of the relation can be maintained only if both $M_{\text{vir}}$ and $V_m$ are changing with time in a certain way, specified above.

During the second stage of evolution, the subhalo experiences tidal forces from the host and other halos, and its mass and extent are tidally truncated. The average density $\rho_0$ within the truncation radius $R_t$ is approximately constant along the orbit and is proportional to the background density of the host halo at the pericenter of the subhalo’s orbit. The truncation radius scales with the truncated mass, $M_t$, as $R_t \propto (M_t/\rho_0)^{1/3}$, so $M_t \propto V_m^{3/2}$. The velocity $V_m$ is constant along the satellite trajectory, we obtain the following relation: $M_t \propto V_m^{3/2}$.
in agreement with the average fit in this regime. Figure 6 shows that in this regime the power-law relation between $M_t$ and $V_m$ has a significant dispersion, which is due to the variation of $\rho_t$ along the trajectory. Nevertheless, as long as the distance of closest approach to the host halo remains the same, the average relation is well maintained.

Thus, we expect the mass-velocity relation to be constrained by the two limiting slopes, 3 and 5. The actual mass accretion and mass loss history may vary from halo to halo, but the same $M-V_m$ relation is maintained throughout the evolution, with a transition from the initial slope, $\nu = 4.7$, for the isolated halos to the later slope, $\nu = 3$, for the tidally truncated halos.

5.4. Abundance and Radial Distribution of Galactic Satellites

Figure 7 shows the cumulative velocity functions (CVFs), the number of satellites with maximum circular velocity $V_m$ larger than a given value, for the objects located within $200 \, h^{-1}$ kpc of their host halo. The figure compares the CVFs for the DM satellites and observed satellites of the MW and Andromeda and highlights the “missing satellite problem” (Kauffmann et al. 1993; Klypin et al. 1999b; Moore et al. 1999a): a large difference in the number of dwarf-sized DM satellites in simulations and the observed number of dwarfs in the LG.

Figure 8 shows the normalized cumulative radial distribution of the DM satellites compared with the radial distribution of satellites around the MW within the same radius. The LG data are from the compilation of Grebel et al. (2003). The figure clearly shows that the spatial distribution of dwarf galaxies around the MW is more compact than the distribution of the DM population. The median distance of observed satellites within $200 \, h^{-1}$ kpc is 60 and 85 $h^{-1}$ kpc for the MW and M31, respectively. For the DM satellites the corresponding median distances are 116, 121, and 120 $h^{-1}$ kpc. Although the median for M31 satellites is smaller than that of the DM satellites, their radial distributions are formally consistent. However, the comparison with the M31 satellites is difficult at present because typical distance errors are $\sim 20-50$ kpc ($\approx 70$ kpc for some galaxies), comparable to the distance to the host.

---

4 Note that uncertainty in the velocity anisotropy affects the conversion of the line-of-sight rms velocity of dSph galaxies to $V_m$. In the plot we assume an isotropic velocity distribution. Our reanalysis of numerical simulations of Gnedin (2003a) shows that tidal truncation and heating of galaxies leads to the preferential removal of radial orbits and the development of the tangentially biased dispersion in the outer parts. A similar result has been found by Kazantzidis et al. (2004a) and Moore et al. (2003). The solution of the Jeans equation for $V_m$ is sensitive to the exact value of the anisotropy parameter (Zentner & Bullock 2003; Kazantzidis et al. 2004b).

5 We use the circular velocities compiled by Klypin et al. (1999b) with updated values of circular velocity for the Large and Small Magellanic Clouds of $V_m = 50$ and 60 km s$^{-1}$, respectively (van der Marel et al. 2002).
For the MW satellites the typical distance errors are an order of magnitude smaller, and the comparison is considerably more meaningful. The Kolmogorov-Smirnov (K-S) test gives a probability of $(6.8) \times 10^{-4}$ that the MW satellites are drawn from the same radial distribution as the DM satellites. This has also been pointed out recently by Taylor et al. (2004), who compared the spatial distribution of the MW satellites with results of their semianalytic model of galaxy formation. Thus, in addition to the vastly different abundances of the observed and predicted satellites, there is a discrepancy in the radial distribution. Models that aim to reproduce the abundance of the LG satellites should therefore be able to reproduce the radial distribution as well.

6. A MODEL OF STAR FORMATION IN DWARF HALOS

6.1. Description of the Model

To gain insight into which halos might become luminous and which might not, we implement the following simple model of star formation. We use the standard assumption that the gas within the halos with the virial temperature $T_{\text{vir}} \gg 10^4$ K dissipates its energy via radiative cooling and forms a disk. We then apply the empirical Schmidt law to calculate the star formation rate (SFR) in radial shells within the disk. The novel features of our model include (1) the use of mass accretion and stripping history of the dwarf halos extracted from simulation, (2) the effects of photoionizing extragalactic background by using the filtering mass, (3) the effects of inefficient dissipation of the gas at $T_{\text{vir}} \approx 10^4$ K, and (4) bursts of star formation due to strong tidal shocks. The details of the model are as follows:

1. Using the mass assembly history (MAH) of a given satellite halo directly from the simulation instead of a semianalytic approach, we are able to trace the major merger events, as well as the quiescent accretion of material. The halo mass increases in both regimes, but the star SFRs are very different. The use of simulation MAHs allows us to determine the accretion epoch of a satellite and follow its mass loss due to tidal stripping.

2. At each time output we calculate the accreted mass since the last time step, $\Delta M$. We increase the total gas mass, $M_g$, in the satellite by the amount of cold gas in a single halo with the mass $\Delta M$ at that epoch: $\Delta M_g = f_g(M, z)\Delta M$. The fraction $f_g$ takes into account the photoevaporation of baryons by extragalactic UV flux, using the filtering scale parameterization of Gnedin (2000) and taking the redshift of reionization to be $z_r = 7$. See Appendix B and equation (B3) for details. After the satellite enters the host halo, the accretion of new gas is halted and the disk scale length is fixed, although stars may continue to form from the remaining reservoir of cold gas.
the luminous satellites for the three hosts ranges from luminous satellites in our model described in center of their host. The dashed lines show the velocity function for the comparable to the observed range. and simulated objects are selected within the radius of 200 h^{-1} kpc from the center of their host. The dashed lines show the velocity function for the luminous satellites in our model described in § 6. The minimum stellar mass of the luminous satellites for the three hosts ranges from ≈10^5 to ≈10^8 M☉, comparable to the observed range.

We distribute the gas on a spherically symmetric grid of 50 radial shells, according to the surface density of an exponential disk: \( \Sigma_g(r) = \Sigma_0 \exp(-r/r_d) \). We use the observed Schmidt law of star formation to estimate the SFR: \( \dot{\Sigma}_s = 2.5 \times 10^{-4} (\Sigma_g/1M_☉/pc^2)^4M_☉/kpc^2/yr^1 \). Only the shells above the threshold \( \Sigma_g > \Sigma_{th} \equiv 5M_☉/pc^2 \) form stars (Kennicutt 1998).

3. The scale length of the disk is determined by its angular momentum. For a rotationally supported disk it is approximately \( r_d = 2^{-1/2} \lambda r_{vir} \). The value of the angular momentum parameter is drawn randomly from the probability distribution,

\[
p(\lambda) d\lambda = \frac{1}{\sqrt{2\pi} \sigma_\lambda} \exp\left[-\frac{\ln \lambda/\bar{\lambda}}{2\sigma_\lambda^2}\right] d\lambda, \tag{3}
\]

with \( \bar{\lambda} = 0.045 \) and \( \sigma_\lambda = 0.56 \), according to the latest measurement by Vitvitska et al. (2002). This is a key assumption of the semianalytic models of galaxy formation.

However, small halos at high redshift could cool by atomic hydrogen only to about 10^4 K. If their virial temperature is only slightly above that equilibrium temperature, the gas would not be able to dissipate enough to reach a rotationally supported state. Instead, its distribution would be more extended, which can have important implications for the star formation with a density threshold \( \Sigma_{th} \). This effect is particularly important for dwarf halos.

We model the effect of inefficient dissipation by adopting the expansion factor that depends on the ratio of the virial temperature to the equilibrium temperature 10^4 K. The gas would reach a Boltzmann distribution with the density \( M/r^3 \propto \exp(-\Phi/kT) \), where \( \Phi \) is the potential energy. Using the maximum circular velocity instead of the temperature and ignoring the slow variation of the potential, we can express the scale length of the gas as \( r_d \propto \exp(c(V_4/V_m)^2) \), where \( c \) is a normalization factor and \( V_4 = 16.7 \text{ km s}^{-1} \) is the virial velocity corresponding to \( T_{vir} = 10^4 \text{ K} \). We find that \( c = 10 \) is a best fit to the abundance and radial distribution of the LG galaxies (see § 6.2). This scaling also provides a good description of the extent of the gas within halos in the cosmological galaxy formation simulation described in Kravtsov & Gnedin (2003). Thus, we set the size of the gaseous disk at each time step to be

\[
r_d = 2^{-1/2} \lambda r_{vir} e^{10(V_4/V_m)^2}. \tag{4}
\]

Of course, \( r_d \) is not allowed to exceed the tidal radius of the halo, \( r_t \). The gas in large halos with \( V_m \gg V_4 \) can cool efficiently and reach rotational support, but for small halos with \( V_m \approx V_4 \) the extended distribution reduces the central concentration of the gas and hinders star formation.

4. Strong tidal forces, such as in the interacting or merging galaxies, may lead to a burst of star formation throughout the dwarf galaxy. The association of starbursts with strong peaks of the tidal force is motivated by theoretical models (Mayer et al. 2001a) and, to a certain extent, by observations (Zaritsky & Harris 2004). The latter suggest that the tidally triggered star formation in the SMC can be accurately modeled as an instantaneous burst of star formation. Zaritsky & Harris (2004) find the best fit to their data when the SFR varies as \( r^{-4.6} \) with the distance to the Galaxy. The tidal interaction parameter, \( I_{tid} \) (see eq. [A7]), which reflects the integrated effect of a single tidal shock, is the most natural candidate for the parameterization of the tidally triggered SFR. Ignoring the adiabatic correction, it varies with the distance to the perturber approximately as \( I_{tid} \propto r^{-4} \) (but see the discussion in § 6.2).
We allow for the starburst mode of star formation, when the tidal interaction parameter exceeds a threshold value. After experimenting with different thresholds, we find that $I_{\text{tid, th}} = 4 \times 10^3 \text{ Gyr}^{-2}$ provides the best simultaneous fit to the velocity function and spatial distribution of the satellites. In all radial shells, a fraction $f_s = I_{\text{tid}}/4 \times 10^3 \text{ Gyr}^{-2}$ (with a maximum of $f_s = 0.5$) of the available gas is converted into stars instantaneously. The normalization of $f_s$ is somewhat arbitrary and can be adjusted to fit the stellar masses of the satellites. Since the starburst changes drastically the distribution of gas in the galaxy, new infalling gas may have a very different angular momentum. Therefore, after each starburst we recalculate the value of $f_s$ according to equation (3).

The external tidal force determines the truncation radius $R_t$ of the satellite, outside which all stars and gas are lost. In a static gravitational field, the radius of the Roche lobe is set by the condition that the average density of matter in the satellite equals twice the local ambient density (for the isothermal sphere potential). In a dynamic situation of the satellite on stars instantaneously. The normalization of $f_s$ is somewhat arbitrary and can be adjusted to fit the stellar masses of the satellites. Since the starburst changes drastically the distribution of gas in the galaxy, new infalling gas may have a very different angular momentum. Therefore, after each starburst we recalculate the value of $f_s$ according to equation (3).

The truncation occurs near the maximum of the tidal force along the orbit, usually at the perigalactic distance.

The knowledge of the external tidal force also allows us to estimate the tidal heating of stars in the satellite. After each tidal shock, typically one per orbit, the velocity dispersion of stars in each radial shell increases by the amount

$$\sigma^2 = \frac{1}{3} M_{\text{tid}} \left( \frac{r}{1 \text{ kpc}} \right)^2 \text{km}^2 \text{s}^{-2}$$

(Gnedin 2003b). The mass-weighted dispersion $\sigma$ may serve as an indicator of the morphological type of the satellite. In § 6.2, we adopt the ratio of the rotation velocity to the velocity dispersion, $v_{\text{rot}}/\sigma$, as a possible criterion. In practice, we compute $v_{\text{rot}}$ as the circular velocity of the NFW halo at the radius enclosing all bound stars.

6.2. Results

We show the predictions of our model for the CVF and radial distribution of luminous satellites by dashed lines in Figures 7 and 8. The model reproduces fairly well both observational statistics. The abundance of luminous satellites and the shape of the velocity function are in reasonable agreement with observations. The stellar masses of the satellites are in the range $10^5 M_\odot \lesssim M_* \lesssim 10^{10} M_\odot$, similar to the observed range of stellar masses of the LG galaxies (e.g., Dekel & Woo 2003). As can be seen in Figure 4, all the luminous systems in our model, including those that have masses and circular velocities of the smallest dwarfs, were relatively massive ($M \gtrsim 10^9 M_\odot$ and $V_m \gtrsim 30 \text{ km s}^{-1}$) at some point in their evolution.

In comparison with the observed velocity functions, it is worth noting that the conversion between the line-of-sight stellar velocity dispersions and maximum circular velocity is somewhat uncertain (Stoehr et al. 2002; Zentner & Bullock 2003; Kazantzidis et al. 2004b). Thus, at this point it is not worth trying to reproduce the observed function exactly.

The median distances of luminous satellites to their respective hosts are 59, 91, and 73 $h^{-1}$ kpc for halos G1, G2, and G3, respectively. This is close to the median values for the MW and M31 and lower than the median distance of the overall DM satellites ($\lesssim 120 h^{-1}$ kpc; see § 5.4). The $K$-S probability that the radial distribution of these luminous satellites is drawn from the same distribution as that of the MW are 97%, 1%, and 75% for the three hosts, respectively. Although there are apparent fluctuations due to the differences in the evolutionary histories of the three hosts, the luminous satellites in our model have a clear tendency to be more centrally concentrated than the overall DM satellite population.

On the other hand, tidally triggered bursts of star formation are not limited to the central parts of the host halo. In the isolated halo G1, where the sample of satellites is not contaminated by close proximity to another host, the tidal heating parameter $I_{\text{tid}}$ scales with the present distance as $I_{\text{tid}} \propto r^2$, $a = -3.7 \pm 0.2$, for $r < 1 h^{-1}$ Mpc, where the quoted error is the standard deviation of the whole sample, not the error of the mean. This is consistent with the expected slope $a = -4$ (see eq. [A6], ignoring the adiabatic correction). However, if we limit the sample to the halos of interest, i.e., only those capable of forming stars (with the maximum $V_m^2 > V_A = 16.7 \text{ km s}^{-1}$, the virial velocity corresponding to $T_{\text{vir}} = 10^4 \text{ K}$), then the slope is significantly shallower: $a = -1.8 \pm 0.2$. Furthermore, if we consider only the largest tidal parameters that might lead to starbursts ($I_{\text{tid}} > 10^3 \text{ Gyr}^{-2}$), then the distribution is almost independent of distance: $a = -0.3 \pm 0.1$. Thus, the current location of the satellite in the host galaxy gives very little indication of whether it had tidal starbursts in the past.

We find a large variety of star formation histories for the luminous satellites. Most systems have a single initial burst lasting up to 2 Gyr. For some this is the only star-forming activity, while others have a constant SFR at $1 M_\odot \text{ yr}^{-1}$ up to $z = 0.7$ or bursty star formation continuing until $z = 0.2$. There are also objects that have only a single tidally triggered burst at $z \sim 1$. The tendency is for more massive satellites to have more extended star formation. The median mass-weighted epoch of star formation in halos with $V_m < 30 \text{ km s}^{-1}$ is between $t_{\text{med}} = 1$ and 4 Gyr cosmic time (corresponding to redshifts between $z = 5$ and 1.7), while the more massive halos have $t_{\text{med}}$ up to 7 Gyr ($z = 0.7$).

Our simple model also predicts central stellar densities in reasonable agreement with observations: roughly constant, $\Sigma_* \sim 5-50 M_\odot \text{ pc}^{-2}$, for $M_* < 10^8 M_\odot$ and rising with the stellar mass as $\Sigma_* \sim M_*/(10^8 M_\odot)$ $M_\odot \text{ pc}^{-2}$ for systems with $M_* > 10^8 M_\odot$. The satellites located within 100 kpc of their host galaxy have typically higher central densities ($\gtrsim 50 M_\odot \text{ pc}^{-2}$) than the more distant satellites.

The results listed above are valid for all dwarf satellite galaxies regardless of their evolutionary history. In addition, as discussed in the previous section, our model tracks the tidal heating of stars formed in each halo. We can therefore attempt a crude morphological classification of galaxies based on the amount of heating they experienced. This is motivated by the observations that dSph galaxies have low values of the ratio of rotation velocity to the random velocity dispersion, $v_{\text{rot}}/\sigma \lesssim 1$. The galaxies of transition type dIrr/dSph have $v_{\text{rot}}/\sigma \lesssim 2$ (see, e.g., Grebel et al. 2003 and references therein). Theoretical models of Mayer et al. (2001a, 2001b) also indicate that the
tidally heated dSph-like remnants of low surface brightness spiral galaxies have small $v_{\text{rot}}/\sigma$.

We use the circular velocity at the radius enclosing all the stellar mass, $v_c^{\text{out}}$, as a proxy for $v_{\text{rot}}$. The rotation velocity will, in general, be smaller than the circular velocity because some of the kinetic energy is in the form of the random motions. Also, we account only for direct tidal heating and do not take into account tidally induced heating via bar and bending instabilities. The exact value of $v_{\text{rot}}/\sigma$ for our galaxies is thus somewhat uncertain as our $\sigma$ may be regarded as lower limit. We experimented with several values for the classification threshold in the range $1 < v_{\text{rot}}/\sigma < 3$, but the main trends are not sensitive to the specific choice in this interval. We chose the value of $v_{\text{rot}}/\sigma = 3$ for the classification shown in Figure 9. For the observed galaxies, we combined dwarf spheroidal, transition type, and dwarf elliptical galaxies in one broad class of spheroidal systems, using Table 1 of Grebel et al. (2003). The figure shows that our model is consistent with the observed trend of morphological segregation. Most spheroidal systems are located within $300 \, \text{h}^{-1} \, \text{kpc}$ of the host, while irregular systems are found at a wide range of radii.

7. DISCUSSION

In the previous sections we presented the results of the dynamical evolution of galactic satellites in self-consistent cosmological simulations. One of the main findings is that the internal structure of the satellites responds to the changes of mass in a remarkably regular way. Namely, during the periods of both mass growth and tidal mass loss, the maximum circular velocity of a halo changes as $V_m \propto M^{1/3}$. The slope $\alpha$ is $\sim 4$ when the mass grows and $\alpha \approx 3$ when the mass decreases because of tidal stripping. The latter result was also obtained by Hayashi et al. (2003) in their noncosmological simulations of satellite evolution.

The overall evolution of subhalo population is such that their $M-V_m$ relation is similar to that of isolated halos (with $\alpha \approx 3.3$). The fact that isolated halos and subhalos have similar mass–circular velocity relations may hint at why the fundamental planes of galaxies in clusters and in the field are similar (Dressler et al. 1987; Djorgovski & Davis 1987; Mobasher et al. 1999; Bernardi et al. 2003) and why the scatter in the Tully-Fisher relation is so small (Kannappan et al. 2002).

The fact that the circular velocity decreases with decreasing mass means that the systems experiencing dramatic mass loss experience a significant change in circular velocity. We find that about 10% of the subhalos with masses less than $10^8 - 10^9 \, M_\odot$ or $V_m < 30 \, \text{km s}^{-1}$ at $z = 0$ have considerably larger masses and circular velocities at earlier epochs. This may explain how such apparently small objects as UMi and Dra could have formed stars, given that the gas accretion is expected to be strongly suppressed by the UV background (e.g., Thoul & Weinberg 1996; Gnedin 2000). In our model, these systems were once sufficiently massive ($V_m \gtrsim 30 \, \text{km s}^{-1}$) to accrete gas and form stars, but the accretion was halted when they started to experience tidal mass loss.

After the accretion of new gas stops, these systems may continue to form stars in bursts as they are tidally stirred (e.g., Mayer et al. 2001b). Interestingly, we find that the strongest tidal interaction may occur even before the halo is accreted by the host. Some satellites experience the strongest tidal force from multiple halos at early epochs in major mergers during the assembly of their host (see Fig. 3). Such mergers are frequent at early epochs, and we find that in general all satellites forming stars experience multiple bursts in the first 2–3 Gyr of their evolution. We present a simple model for star formation in dwarf halos and apply it to the evolutionary tracks extracted from the simulations. As shown in Figure 7, the model is successful in reproducing the abundance of luminous satellites around M31 and the MW.

The spatial distribution of dwarfs around the MW offers another independent challenge to any model of satellite evolution. Figure 8 shows that our model reproduces the observed distribution reasonably well. The distribution of luminous satellites is more compact than the overall population of subhalos because stars form only in objects that were sufficiently massive at high redshifts. Because of the strong mass and redshift dependence of spatial bias, such objects are considerably more clustered around the host than smaller halos that form at a wide range of redshifts. Correspondingly, we find that luminous objects were accreted by the host systematically earlier (by $\Delta z \sim 0.5$–1) than smaller mass dark subhalos.

One of the remarkable features of our model is that the results are not sensitive to the details of the reionization history of the universe. For example, all the presented results are nearly intact if we change the assumed redshift of reionization from the fiducial value of $z_r = 7$ to $z_r = 15$ (see Appendix B). The physical reason behind this insensitivity to reionization is the inefficiency of gas cooling and star formation in small-mass ($T_{\text{vir}} \lesssim 10^4 \, \text{K}$) systems. This is because gas in such systems cannot cool via hydrogen line emission and must rely on the inefficient H$_2$ cooling. Such redshift-independent suppression of gas cooling is observed in cosmological simulations of Chiu et al. (2001) and Kravtsov & Gnedin (2003). The important implication is that properties of the population...
of galactic satellites are determined by the physics of galaxy formation rather than by the UV background and reionization.

Our results can qualitatively explain the morphological segregation of the LG galaxies (e.g., Grebel 2000). As shown in Figure 9, a simple division of model galaxies into irregular and spheroidal based on the amount of tidal heating they experienced during their evolution reproduces the main observed trend. Most spheroidal systems are located within 300 $h^{-1}$ kpc, while irregular galaxies are found almost uniformly at all distances. Therefore, our results support the scenario that spheroidal systems form via strong tidal heating (Mayer et al. 2001a, 2001b). Note, however, that tidal heating is not restricted to the host. It can occur early on, before the host is assembled, within merging subgroups.

Interestingly, this explains a puzzling presence of the Cetus and Tucana dSph galaxies at the outskirts of the LG some 700 and 1000 kpc from the nearest massive spiral galaxy (MW or M31). We also find $\sim 1$–2 galaxies with significant heating (the ratio of the rotational velocity to the velocity dispersion of $v_{\text{rot}}/\sigma < 1$) at distances $\sim 1000$ kpc from their hosts. The tidal heating of these systems occurred in small groups that are being accreted by the host at the current epoch (see also Gnedin 2003b for a similar effect in clusters of galaxies). As the tidal force unbinds satellites from such accreting groups, isolated dSph galaxies may be found at large distances from the primary host.

Also, early tidal interaction, experienced, for example, by the system shown in the left column of Figure 3, and subsequent interaction with other subhalos may lead to the increase of orbital energy and apocenter distance. This scenario would also explain presence of dSphs at large distances from the primary. The main points in both scenarios are that the primary is not the only source of tides and that the present-day environment is not necessarily indicative of a dwarf galaxy’s past.

One of the most interesting candidates for the “missing” dark halos is the population of compact high-velocity clouds (CHVCs) of neutral hydrogen ($H\,i$; e.g., Blitz et al. 1999; Braun & Burton 1999). This idea has recently been boosted by the detection of a concentration of CHVCs near M31 (Thilker et al. 2004). It is thus interesting to consider the amount of gas associated with the subhalos that remain dark in our model. The cumulative gas mass function associated with dark halos is remarkably consistent for all three host halos: $N(M_{\text{gas}} > 10^7 M_\odot)$ at $10^7 M_\odot < M_{\text{gas}} < 10^8 M_\odot$ within 200 $h^{-1}$ kpc. Most of the gas mass is thus in most massive subhalos. The total mass of gas associated with such halos within 200 $h^{-1}$ kpc is $M_{\text{gas}} \approx 2 \times 10^8 M_\odot$, the similar number for all three hosts. If we assume that on the average about 10% of gas is neutral (Maloney & Putman 2003; Thilker et al. 2004), the total mass in neutral hydrogen is $M_{\text{H}} \approx 2 \times 10^7 M_\odot$.

The predicted number of dark clouds with $M_\odot > 10^6 M_\odot$ is $\sim 50$–100. A fraction of the observed CHVCs can thus be associated with the small-mass DM halos. Within the central 50 kpc, however, the number of halos with such gas masses is only $\sim 2$–5. We cannot therefore explain 25 CHVCs observed by Thilker et al. (2004) within this radius around M31. It is possible that simulations underpredict the number of small-mass halos due to overmerging. To check this, we will require higher resolution simulations. On the other hand, we did not take into account processes such as ram pressure stripping, which would further reduce the number of halos with gas. Another possibility is that most of the observed M31 CHVCs are gas clouds in tidal streams, such as the Magellanic Stream, and are not associated with distinct dark matter halos (Putman et al. 2003).

8. COMPARISON WITH PREVIOUS WORK

Possible astrophysical solutions to the missing satellite problem have been considered in the last several years. Here we discuss the main differences between our model and the models proposed in previous studies.

Bullock et al. (2000), Somerville (2002), and Benson et al. (2002) discussed the formation and evolution of dwarf galactic satellites by using semianalytic models of different degrees of sophistication. The conclusion reached by all these studies is that the extragalactic UV background can greatly suppress the gas accretion and star formation in the majority of low-mass ($V_m < 30$ km s$^{-1}$) halos. A small fraction of the dwarf halos that harbor stellar systems were assumed to have formed (i.e., assembled a significant fraction of their mass) before reionization, when the level of UV radiation was low. This is because in all these studies the maximum circular velocity of subhalos was assumed to be constant as the mass is tidally stripped. There was thus a simple one-to-one mapping between the circular velocity observed at $z = 0$ and that at the time of accretion. Our results show that this assumption is incorrect (see also Hayashi et al. 2003; Kazantzidis et al. 2004b). Another key difference is that tidal mass loss in our model can occur before a halo is accreted by the host, as a result of interactions with other halos. These effects are not accounted for in any of the semianalytic models.

The implicit assumption in the above models is that the small systems would be able to retain the accreted gas and form stars after reionization. This assumption was justified at the time, as the first calculations of photoevaporation of gas indicated that halos with $V_m \geq 10$ km s$^{-1}$ might retain their gas (Barkana & Loeb 1999). More recent calculations, however, show that the gas could be gradually removed from halos of up to $V_m \approx 30$ km s$^{-1}$ (Shaviv & Dekel 2003). In light of this result, the previous models would not be able to explain the formation and properties of luminous dwarfs, as the star formation in small halos would be suppressed after reionization. It would thus be difficult to explain the more extended star formation histories derived for many dSph galaxies in the LG (Grebel 2000).

In our model, the small-mass dwarfs are identified with the halos that were relatively massive at high redshift and could retain the gas and form stars after reionization. The star formation histories of dwarfs are thus more extended and in better accord with observations. As noted in the previous section, our model is also insensitive to the epoch of reionization and can accommodate the early reionization suggested by results of the WMAP satellite (Spergel et al. 2003).

Our model and all the models discussed above are qualitatively different from the proposal of Stoehr et al. (2002, 2003). These authors argued that the maximum circular velocity of the LG dwarfs may be systematically underestimated because it is derived from the stellar velocity dispersion within radii considerably smaller than $r_{\text{max}}$, the radius at which the maximum halo velocity, $V_m$, is reached (see, however, Kazantzidis et al. 2004b). Stoehr et al. (2002; see also Hayashi et al. 2003) suggested that the luminous dwarfs may be harbored by the most massive satellites of the DM halos. This has an important physical implication: if the dwarfs indeed occupy 12 or so of the most massive halos, then there exists a certain mass scale below which galaxy formation is completely suppressed. If, on the other hand, the dwarf galaxies occupy satellites with a variety of masses ($\sim 10^7$–$10^9 M_\odot$), one has to explain why

$^6$ The solutions that invoke astrophysics of galaxy formation within the standard CDM framework rather than modifications of the properties of dark matter particles and/or the shape of the initial power spectrum.
some fraction of small halos managed to light up the stars, while most others did not.

If the idea of Stoehr et al. (2002) is correct, our results indicate that circular velocities of dwarf spheroidal halos should have been even larger (by a factor of 2 or more) than the values inferred from the current observations. This could make halos of some galaxies uncomfortably massive. For example, Stoehr et al. (2002) derive the maximum circular velocity for the Draco galaxy in the range ~35–55 km s\(^{-1}\). This implies the preaccretion values of \(V_{\text{max}} \gtrsim 70\) km s\(^{-1}\) and a preaccretion mass comparable to those of M32, NGC 205, and M33. The fact that the luminosity of the Draco galaxy is almost 4 orders of magnitude lower than luminosities of these galaxies would present a major puzzle.

In addition, the radial distribution of the most massive satellites should be consistent with the observed radial distribution of the MW satellites. We find that in our simulations the radial distribution of subhalos with the largest \(V_m\) is between that of the luminous satellites and all DM satellites shown in Figure 8. In a study of a larger sample of cluster halos, De Lucia et al. (2004) find that the radial distribution of the most massive halos is even more extended than that of the smaller mass objects. A similar point was made recently by Taylor et al. (2004), who used semianalytic models for a subhalo population to show that the radial distribution of the most massive halos is more extended than that of the MW satellites at \(\gtrsim 3\) \(\sigma\) level. A caveat to this argument is that the sample of MW satellites may be incomplete at large distances, and more faint dwarf galaxies will be discovered in the future (Willman et al. 2004).

9. CONCLUSIONS

We presented a study of the dynamical evolution of galactic satellites by using self-consistent high-resolution cosmological simulations of three MW-sized halos. Our main results and conclusions are as follows.

1. We find that \(\approx 10\%\) of the substructure halos that have masses of less than \(10^8–10^9\) \(M_\odot\) at the present epoch had considerably higher masses and circular velocities when they formed at \(z > 2\). After the initial period of mass accretion, while these objects evolve in isolation, they suffer dramatic mass loss due to tidal stripping by actively merging massive neighboring halos. Strong tidal interactions can occur even before the dwarfs are accreted by their primary host halos.

2. The decrease in mass due to tidal stripping is accompanied by the decrease in the maximum circular velocity, such that the objects evolve along a \(M-V_m^\alpha\) relation with \(\alpha \approx 3–4\).

3. These results indicate that some of the systems that have small masses and circular velocities at \(z = 0\) could have had masses comparable to those of the SMC and LMC in the past. This can explain how the smallest dwarf spheroidal galaxies observed in the LG were able to build up sizable stellar masses in such shallow potential wells.

4. We present a simple galaxy formation model based on the evolutionary tracks extracted from the simulations. The novel features of the model are the starburst mode of star formation after the strong peaks of the tidal force and the accounting for the inefficient dissipation of gas in halos with \(T_{\text{vir}} \lesssim 10^4\) K. The model can successfully reproduce the circular velocity function, radial distribution, morphological segregation of the observed MW satellites, and the basic properties of galactic dwarfs such as stellar masses and densities.

We would like to thank Nick Gnedin for providing us with the results of his filtering-mass calculation in numerical form and for his hospitality at the University of Colorado at Boulder, where this paper was completed. We are also grateful to Andrew Zentner for useful comments and Stelios Kazantzidis for communicating results of his calculations prior to publication. The simulations presented here were performed on the Origin2000 at the National Center for Supercomputing Applications. This work was supported by the National Science Foundation under grant AST 02-06216 and AST 02-39759 to the University of Chicago. O. Y. G. is supported by the STScI Fellowship.

APPENDIX A

CALCULATION OF THE TIDAL FORCE

In our analysis we use the external tidal force experienced by each satellite halo to estimate the strength of tidal interaction. We calculate the force both directly from the gravitational potential computed in the simulation and by using an analytical approximation for the neighbor halos.

To compute the tidal force numerically from the local potential \(\Phi\), we estimate its second spatial derivative at the center of mass of the satellite,

\[
F_\alpha \equiv -\left(\frac{d^2 \Phi}{dR_\alpha dR_\beta} \right)_{0} r_\beta \equiv F_{\alpha \beta} r_\beta, \tag{A1}
\]

where \(r\) is the radius vector in the satellite reference frame and \(R\) is the radius vector in the perturber reference frame. The potential \(\Phi\) is calculated on the original refinement grid by using the ART gravity solver. In the calculation of the potential, we subtract the self-contribution of the halo and consider only the external tidal potential.

In a study of galaxy interactions in clusters of galaxies, Gnedin (2003b) used the Savitzky-Golay smoothing filter to interpolate the potential on a plane and calculate its derivatives from a smooth polynomial function. We employ a similar scheme but with the adaptive fourth-order interpolating polynomials in each of the three orthogonal planes around the satellite center of mass,

\[
P_d(x, y) = \sum_{k, l=0}^{4} c_{k\ell}x^k y^l, \tag{A2}
\]
and the same for the $x$-$z$ and $y$-$z$ planes. The fourth-order expansion ensures a smooth second derivative of the potential. In each of the planes we extract an $n \times n$ subgrid centered on the original grid point, nearest to the satellite center. To obtain a uniform accuracy of the tidal force for satellites of different sizes, we choose the size of the subgrid cells to be closest to $\frac{1}{4}$ of the satellite’s tidal radius. The coefficients $c_{ij}$ are calculated by minimizing the $\chi^2$ deviation,

$$\chi^2 = \sum_{i,j=1}^{n} \left[ P_d(x_i, y_j) - \Phi(x_i, y_j) \right]^2,$$

by using the CERN Program Library routine MINUIT.\footnote{See http://wwwasdoc.web.cern.ch/wwwasdoc/minuit.} We have experimented with $n = 16$, 32, and 64 and found that $n = 64$ provides the most accurate derivatives, as tested on the analytical NFW models. The tidal tensor components $F_{\alpha\beta}$ are then calculated by analytical differentiation of equation (A2).

We compare the real tidal force due to the overall mass distribution in the simulation with the contributions of all neighboring halos, including the host halo. We model the halos with an NFW density profile and take their mass $M_{\text{vir}}$ and virial radius $r_{\text{vir}}$ directly from the halo catalogs generated by the halo finder (see § 3). We determine the scale radius of the NFW model for the satellite halos from the position of the maximum circular velocity, $r_s = r_{\text{max}}/2.16$. For the host halo, we use the parameterization

$$c_{\text{NFW}} \equiv r_{\text{vir}}/r_s = 16a^{3/2},$$

which is a best fit to the density profile of the analyzed host halos. The analytical tidal force in the reference frame of the satellite is then readily calculated using equation (5) of Gnedin et al. (1999):

$$F(r) = \frac{GM(R)}{R^3} \left(3 - \mu\right) (n \cdot r)n - r,$$

where $r$ is the radius vector within the satellite, $R$ is the distance to the perturber, $n \equiv R/R$, $\mu \equiv d \ln M/d \ln R$, and $M(R)$ is the enclosed mass of the NFW model:

$$M(R) = M_{\text{vir}} \frac{\ln (1 + R/r_s) - 1 + (1 + R/r_s)^{-1}}{\ln (1 + c_{\text{NFW}}) - 1 + (1 + c_{\text{NFW}})^{-1}}.$$

Figure 3 shows that the approximate tidal force calculated in this manner is quite accurate, especially near the maximum of the tidal force.

Although the tidal force along the satellite trajectory varies rapidly with time, most of the tidal heating of stars and dark matter particles occurs near the strong peaks of the tidal force. Each of these tidal peaks can be considered an independent tidal shock (Gnedin & Ostriker 1999; Gnedin 2003b). The amount of tidal heating, such as the increase of the velocity dispersion, is proportional to the integral over the peak of tidal force:

$$I_{\text{tid}}(t_n) \equiv \sum_{\alpha, \beta} \left( \int F_{\alpha\beta} dt \right)^2 \left(1 + \frac{\tau_n^2}{t_{\text{dyn}}^2} \right)^{-3/2},$$

where the sum extends over all components of the tidal tensor, $\alpha, \beta = \{x, y, z\}$. The last factor is the correction for the conservation of adiabatic invariants of stellar orbits during the tidal shock (see Gnedin & Ostriker 1999). Here $\tau_n$ is the effective duration of peak $n$ at time $t_n$, and $t_{\text{dyn}}$ is the dynamical time of the satellite. We take $t_{\text{dyn}} = 2\pi r_{1/2}/v_{\text{rot}}$, where $r_{1/2}$ is the half-mass radius of the stellar disk and $v_{\text{rot}}$ is the circular velocity of the appropriate NFW model at $r_{1/2}$. The cumulative tidal heating parameter is the sum over all tidal peaks:

$$I_{\text{tid}} = \sum_n I_{\text{tid}}(t_n).$$

This parameter determines the increase of the velocity dispersion of stars (eq. [6]) in our model of dwarf galaxy formation (§ 6).

APPENDIX B

FILTERING-MASS SCALE

We estimate the suppression of gas accretion due to the extragalactic UV background by using the filtering-mass scale derived by Gnedin (2000). He defined $M_f$ as the mass of the halo that would lose half its baryons, compared with the universal baryon fraction. This filtering mass relates to the Jeans mass of the intergalactic gas integrated over the cosmic history:

$$M_f(a) = M_{f,0} f(a)^{3/2}, \quad f(a) = \frac{3}{a} \int_0^a x T_4(x) \left(1 - \left(\frac{x}{a}\right)^{1/2}\right) dx,$$

where $T_4(x)$ is the temperature function. The filtering mass is given by

$$M_f(a) = \frac{9}{2\pi^2} \frac{1}{\nu a} 3^{3/2} a^{3/2} f(a)^{3/2},$$

where $\nu = 0.05$. This parameter is used in our model of dwarf galaxy formation (§ 6).
where $M_{f0} = 2.5 \times 10^{11} \, h^{-1} \, \Omega_{\text{b}}^{-1/2} \mu^{-3/2} \, M_{\odot}$, $\mu \approx 0.59$ is the mean molecular weight of the fully ionized gas, and the integration extends over the expansion factor, $a$ (Gnedin 2000, eq. [6]). The temperature of the cosmic gas $T_4$ is expressed in units of $10^4$ K for convenience.

Here we propose an analytical fit to the results of Gnedin (2000), assuming a simple dependence of the temperature on the expansion factor: $T_4(a) = (a/a_0)^{\alpha}$ for $a \leq a_0$, $T_4(a) = 1$ for $a_0 < a < a_r$, and $T_4(a) = (a/a_r)^{-1}$ for $a \geq a_r$.

These three distinct stages can be clearly seen on Figure 1 of Gnedin (2000). They correspond to (1) the epoch before the first $\text{H} \, \text{II}$ regions form, $z > z_0$; (2) the epoch of the overlap of multiple $\text{H} \, \text{II}$ regions, $z < z < z_0$; and (3) the epoch of complete reionization, $z < z_r$. In the first stage, before redshift $z_0 \equiv 1/a_0 - 1 \approx 8$, the temperature is rising as the newly formed stars ionize their neighboring regions. The parameter $\alpha$ controls the rate of growth of the extragalactic UV flux; we find $\alpha = 6$ to be the best fit. During the overlap stage, between redshifts $z_0$ and $z_r \equiv 1/a_r - 1 \approx 7$, the temperature is kept constant at roughly $10^4$ K as the cosmic $\text{H} \, \text{II}$ regions overlap. After the universe is fully ionized, at redshifts below $z_r$, the temperature falls adiabatically with the cosmic expansion.

With these analytical expressions for $T_4(a)$, we integrate equation (B1) analytically:

$$f(a) = \begin{cases} \frac{3}{(2 + \alpha)(5 + 2\alpha)} \left( \frac{a}{a_0} \right)^{\alpha}, & a \leq a_0, \\ \frac{3}{a} \left[ \frac{1}{2 + \alpha} - \frac{2(a/a_0)^{-1/2}}{5 + 2\alpha} \right] + \frac{a^2}{a_0^2} \left( \frac{5 - 4(a/a_0)^{-1/2}}{10} \right), & a_0 < a < a_r, \\ \frac{3}{a} \left[ \frac{1}{2 + \alpha} - \frac{2(a/a_0)^{-1/2}}{5 + 2\alpha} \right] + \frac{a^2}{a_0^2} \left( \frac{5 - 4(a/a_0)^{-1/2}}{10} \right) + \frac{a a_0}{3} - \frac{a_0^2}{3} \left[ 3 - 2(a/a_r)^{-1/2} \right], & a \geq a_r. \end{cases}$$

(B2)

The virial circular velocity of the halo is $V^2 = GMH(z)(\Delta_{\text{vir}}/2)^{1/2}$, where $H(z) = H(0)\left[\Omega_{\text{b}}(1 + z)^3 + \Omega_{\Lambda}\right]^{1/2}$ is the Hubble constant and $\Delta_{\text{vir}}$ is the virial overdensity with respect to the critical density, parameterized by Bryan & Norman (1998) as $\Delta_{\text{vir}}(z) = 18\pi^2 + 82x - 39x^2$ and $x \equiv \Omega(z) - 1$. The virial temperature is $T_{\text{vir}} = 36(V_c/1 \text{ km s}^{-1})^2$ K.

Our analytical fit is convenient for accurate modeling of the photoheating effect in semianalytical models of galaxy formation. Its versatile form, with the two parameters $z_0$ and $z_r$, allows a simple recalculation of the filtering mass for a different redshift of reionization than was assumed in the simulation of Gnedin (2000). It can also be easily adapted to describe two epochs of reionization or the early extended reionization suggested by WMAP (Spergel et al. 2003). For illustration, we show in Figure 10 (right) the filtering mass as a function of redshift for two choices of the reionization redshift. Figure 10 (left) shows the filtering circular velocity and the corresponding values of the maximum velocity $V_m$ and virial temperature $T_{\text{vir}}$.

Gnedin (2000) provided the following expression for the amount of cold gas left in the halo of mass $M$:

$$f_6(M, z) = \frac{f_6}{[1 + 0.26M/36]^7},$$

(B3)
where \( f_b \approx 0.14 \) is the universal baryon fraction. In § 6 we use this fraction of cold gas to model the star formation history of the satellite galaxies. To account for the inefficiency of atomic gas cooling at \( T < 10^4 \) K, we apply equation (B3), substituting for \( M_f \) the maximum of \( M_f(z) \) or \( M_4 \), the halo mass corresponding to \( T_{\text{vir}} = 10^4 \) K.

REFERENCES

Avila-Reese, V., Firmani, C., Klypin, A., & Kravtsov, A. V. 1999, MNRAS, 310, 527

Barkana, R., & Loeb, A. 1999, ApJ, 523, 54

Benenson, A. J., Frenk, C. S., Lacey, C. G., Baugh, C. M., & Cole, S. 2002, MNRAS, 333, 177

Bernardi, M., et al. 2003, AJ, 125, 1866

Blitz, L., Spergel, D. N., Teuben, P. J., Hartmann, D., & Burton, W. B. 1999, ApJ, 514, 818

Braun, R., & Burton, W. B. 1999, A&A, 341, 437

Bryan, G. L., & Norman, M. L. 1998, ApJ, 495, 80

Bullock, J. S., Kolatt, T. S., Sigad, Y., Somerville, R. S., Kravtsov, A. V., Klypin, A. A., Primack, J. R., & Dekel, A. 2001, MNRAS, 321, 559

Bullock, J. S., Kravtsov, A. V., & Weinberg, D. H. 2000, ApJ, 539, 517

Chiu, W. A., Gnedin, O. Y., & Ostriker, J. P. 1997, ApJ, 514, 818

Dekel, A., & Woo, J. 2003, MNRAS, 344, 1131

De Lucia, G., Kauffmann, G., Springel, V., & White, S. D. M., Lanzoni, B., Stoeber, F., Tormen, G., & Yoshida, N. 2004, MNRAS, 348, 333

Dijkstra, M., Haiman, Z., Rees, M. J., & Weinberg, D. H. 2004, ApJ, 601, 666

Djorgovski, S., & Davis, M. 1987, ApJ, 313, 59

Dressler, A., Lynden-Bell, D., Burstein, D., Davies, R. L., Faber, S. M., Terlevich, R., & Wegner, G. 1987, ApJ, 313, 42

Efthathiou, G. 1992, MNRAS, 256, 43P

Gnedin, O. Y. 2000, in Proc. 33rd ESLAB Symp., Star Formation from the Small to the Large Scale, ed. F. Favata, A. A. Kaas, & A. Wilson (ESA SP-445; Noordwijk: ESA) 87

Gnedin, N. Y. 2003, in Dark Matter in Galaxies, ed. S. Ryder, D. J. Pisano, M. Walker, & K. Freeman (San Francisco: ASP), in press

Kogut, A., et al. 2003, ApJS, 148, 161

Krabbe, A. V. 1999, Ph.D. thesis, New Mexico State Univ.

Krabbe, A. V., & Gnedin, O. Y. 2003, ApJ, submitted (astro-ph/0305199)

Kravtsov, A. V., Klypin, A. A., & Khokhlov, A. M. 1997, ApJS, 111, 73

Maloney, P. R., & Putman, M. E. 2003, ApJ, 589, 270

Mateo, M. L. 1998, ARA&A, 36, 435

Mayer, L., Governato, F., Colpi, M., Moore, B., Quinn, T., Wadsley, J., Stadel, J., & Lake, G. 2001a, ApJ, 559, 754

Mebascher, B., Guzman, R., Aragon-Salamanca, A., & Zepf, S. 1999, MNRAS, 304, 225

Moore, B., Staveley-Smith, L., Freeman, K. C., Gibson, B. K., & Barnes, D. G. 2003, ApJ, 586, 170

Quinn, T., Katz, N., Lake, G., Dressler, A., & Oemler, A. 1996, Nature, 379, 613

Moore, B., Kazantzidis, S., Diemand, J., & Stadel, J. 2003, MNRAS, submitted (astro-ph/0301066)

Moore, B., Lake, G., Quinn, T., & Stadel, J. 1999b, MNRAS, 304, 465

Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493

Navarro, J. F., & Steinmetz, M. 1997, ApJ, 478, 13

Putman, M. E., Staveley-Smith, L., Freeman, K. C., Gibson, B. K., & Barnes, D. G. 2003, ApJ, 559, 719

Quinn, T., Katz, N., & Efthathiou, G. 1996, MNRAS, 278, L49

Shapiro, P. R., Iliev, I. T., & Raga, A. C. 2004, MNRAS, 348, 753

Shaviv, N. J., & Dekel, A. 2003, MNRAS, submitted (astro-ph/0305257)

Somerville, R. S. 2002, ApJ, 572, L23

Spergel, D. N., & Steinhardt, P. J. 2000, Phys. Rev. Lett., 84, 3760

Spergel, D. N., et al. 2003, ApJS, 148, 175

Stoehr, F., White, S. D. M., Springel, V., Tormen, G., & Yoshida, N. 2003, MNRAS, 345, 1133

Stoehr, F., White, S. D. M., Tormen, G., & Springel, V. 2002, MNRAS, 335, L84

Taylor, J. E., Silk, J., & Babul, A. 2004, in IAU Symp. 220, Dark Matter in Galaxies, ed. S. Ryder, D. J. Pisano, M. Walker, & K. Freeman (San Francisco: ASP), in press

Thilker, D., Braun, R., Walterbos, R., Corbelli, E., Lockman, F., Murphy, E., & Maddalena, R. 2004, ApJ, 601, L39

Thoul, A. A., & Weinberg, D. H. 1996, ApJ, 465, 608

van den Bosch, F. C., Yang, X., & Mo, H. J. 2003, MNRAS, 340, 771

van der Marel, R. P., Alves, D. R., Hardy, E., & Suntzeff, N. B. 2002, AJ, 124, 2639

Vitvitska, M., Klypin, A. A., Kravtsov, A. V., Wechsler, R. H., Primack, J. R., & Bullock, J. S. 2002, ApJ, 581, 799

Wechsler, R. H., Bullock, J. S., Primack, J. R., Kravtsov, A. V., & Dekel, A. 2002, ApJ, 568, 52

Willman, B., Governato, F., Dalcanton, J. J., Reed, D., & Quinn, T. 2004, MNRAS, submitted (astro-ph/0403001)

Zaritsky, D., & Harris, J. 2004, ApJ, 604, 167

Zentner, A. R., & Bullock, J. S. 2003, ApJ, 598, 49