The Astrophysical Journal, 573:597–613, 2002 July 10

ΛCDM-BASED MODELS FOR THE MILKY WAY AND M31. I. DYNAMICAL MODELS

Anatoly Klypin
Astronomy Department, New Mexico State University, Box 30001, Department 4500, Las Cruces, NM 88003-0001
and
HongSheng Zhao [1] and Rachel S. Somerville
Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, UK
Received 2002 July 10; accepted 2002 March 12

ABSTRACT

We apply standard disk formation theory with adiabatic contraction within cuspy halo models predicted by the standard cold dark matter (ΛCDM) cosmology. The resulting models are confronted with the broad range of observational data available for the Milky Way and M31 galaxies. We find that there is a narrow range of parameters that can satisfy the observational constraints, but within this range, the models score remarkably well. Our favored models have virial masses of 10^{12} and 1.6×10^{12} $M_\odot$ for the Galaxy and for M31, respectively, average spin parameters $\lambda \approx 0.03$–0.05, and concentrations $C_{\rm vir} = 10$–17, typical for halos of this mass in the standard ΛCDM cosmology. The models require neither dark matter modifications nor flat cores to fit the observational data. We explore two types of models, with and without the exchange of angular momentum between the dark matter and the baryons. The models without exchange give reasonable rotation curves, fulfill constraints in the solar neighborhood, and satisfy constraints at larger radii, but they may be problematic for fast rotating central bars. We explore models in which the baryons experience additional contraction due to loss of angular momentum to the surrounding dark matter. These models produce similar global properties, but the dark matter is only a 25% of the total mass in the central 3 kpc region, allowing a fast rotating bar to persist. According to preliminary calculations, our model galaxies probably have sufficient baryonic mass in the central ~3.5 kpc to reproduce recent observational values of the optical depth to microlensing events toward the Galactic center. Our dynamical models unequivocally require that about 50% of all the gas inside the virial radius must not be in the disk or in the bulge, a result that is obtained naturally in standard semianalytic models. Assuming that the Milky Way is "typical," we investigate whether the range of virial masses allowed by our dynamical models is compatible with constraints from the galaxy luminosity function. We find that if the Milky Way has a luminosity $M_K = -24.0$, then these constraints are satisfied, but if it is more luminous (as expected if it lies on the Tully-Fisher relation), then the predicted space density is larger than the observed space density of galaxies of the corresponding luminosity by a factor of 1.5–2. We conclude that observed rotation curves and dynamical properties of "normal" spiral galaxies appear to be consistent with standard ΛCDM.

Subject headings: cosmology: theory — galaxies: individual (M31) — Galaxy: structure

1. INTRODUCTION

Modeling the mass distribution within the Milky Way (MW) and M31 galaxies is a classical problem that has seen many past iterations (Einasto 1972, 1979; Innanen 1973; Schmidt 1975). Because of our unique ability to observe these systems in exquisite detail, this problem has continued to be a fertile testing ground for theories of galactic structure and galaxy formation (Kent 1989; Dehnen & Binney 1998; Wilkinson & Evans 1999; Evans & Wilkinson 2000; Olling & Merrifield 2001). In the absence of a coherent halo formation theory, early models were plagued by the arbitrary structure parameters of the dark halos. Recent developments in the theory of dark matter halo formation make this sort of modeling much more meaningful because the halo parameters (such as total mass, peak rotation speed, and concentration parameter) may no longer be considered free parameters but are correlated, and these correlations are specified within the context of a given ΛCDM cosmology (cold dark matter plus a cosmological constant Λ) (e.g., Bullock et al. 2001 and references therein).

While very successful on large scales, the standard cosmological theory is experiencing mounting difficulties on the scales of galaxies and dwarf galaxies. The so-called substructure problem (the predicted overabundance of subhalos relative to observed dwarf satellites of the MW and the M31 galaxies [Klypin et al. 1999; Moore et al. 1999a]) may not be a problem. The substructures can remain dark clumps if we invoke the presence of a photoionizing background that "squelches" star formation in small halos (Bullock, Kravtsov, & Weinberg 2000; Somerville 2001). And indeed, known radio/optical lenses show significant variance of image flux ratios from that of smooth lens models, which has been suggested by several groups to be evidence for dark/faint clumps in the line of sight (Chiba 2002; Metcalf & Zhao 2002; Dalal & Kochanek 2002; Keeton 2002; Bradač et al. 2002). However, the so-called cusp and concentration problems remain thorny. The most acute manifestations of these problems are encountered for objects that are dark matter–dominated—namely, the central parts of low surface brightness galaxies and dwarf galaxies (Moore 1994; Flores & Primack 1994; Moore et al. 1999b; van den Bosch & Swaters 2000). For example, the high central density of ΛCDM dark matter halos appears to be inconsistent with the observed rotation curves of dwarf and low surface brightness galaxies (Blais-Ouellette et al. 1999; Côté, Carignan, & Freeman 2000; Blais-Ouellette, Amram, & Carignan 2001; de Blok et al. 2001).

[1] Visiting Professor at Institute of Theoretical Physics, National Astronomical Observatory, Beijing 100012, China.

It is interesting to ask how ΛCDM models fare in their predictions for luminous, high surface brightness galaxies, such as the Milky Way and M31. These systems probe the dark matter properties on mass scales of $10^{12}\ M_\odot$, which are significantly larger than the 10^8–$10^9\ M_\odot$ scales tested by inner rotation curves of dwarf galaxies. The Milky Way is particularly interesting for constraining the central cusp of the dark matter because the inner 3–10 kpc region of the Milky Way is well studied. Apart from the rotation curve and integrated light distribution, the Milky Way offers many more opportunities to study very detailed data that are not generally available for extragalactic systems, e.g., individual stellar radial velocities and proper motions in the solar neighborhood and, recently, Galactic bulge microlensing events.

It is likely that the halo of our Galaxy consists mostly of weakly interacting elementary particles (WIMPs) rather than black holes or failed stars (MACHOs), a necessary condition to apply the standard cosmological approach. The very low microlensing counts by machos in the inner 50 kpc halo of the MW imply that elementary particles make up $\geq$80%–90% of the halo (Lasserre et al. 2000; Zhao & Evans 2000). Therefore it should be valid to apply halo formation models to our Galaxy, which predict an outer dark matter profile of the form $\rho \propto r^{-3}$. In this regard, earlier halo models using cored isothermal halos with a $\rho \propto r^{-2}$ outer profile and a divergent mass (e.g., Olling & Merrifield 2001), or very steep outer profiles $\rho \propto r^{-5}$ (e.g., Evans & Wilkinson 2000), are unphysical and should be updated.

Navarro & Steinmetz (2000) and Eke, Navarro, & Steinmetz (2001) were the first to use mass modeling of the Milky Way to constrain cosmological models. The dark mass inside the solar radius was compared with observational limits derived from previous mass modeling. This could give misleading results because of inconsistent modeling of the halo profiles. This happens because unlike in galaxies with low baryon content such as dwarf or low surface brightness (LSB) galaxies, the density of the inner halo in high-surface brightness galaxies can deviate significantly from the original Navarro, Frenk, & White (1996, hereafter NFW) profile owing to a strong compression during the baryonic collapse (Mo, Mao, & White 1998). It might seem that this can only exacerbate the problem, because the amount of dark matter in the central region will always be increased. In fact, this is not the case because the shape of the entire density profile is modified, which allows and requires a change (decrease) in the virial mass of the dark matter halo. In turn, this reduces the amount of dark matter in the central region. This complicated interplay between different effects makes it necessary to repeat the mass modeling of the Milky Way and M31 including a self-consistent treatment of the baryons and dark matter component.

The amount of dark matter inside the central ~3 kpc of the Milky Way and M31 introduces another layer of problems. It is very likely that both galaxies have fast rotating bars. A rotating bar experiences dynamical friction induced by the dark matter. If the mass of the dark matter surrounding the bar is too large, the dynamical friction could slow down the bar rotation over a short timescale (Weinberg 1985; Debattista & Sellwood 2000), in which case we should not expect to see fast rotating bars. This argument sometimes is presented as the most severe problem the ΛCDM cosmological models encounter on small scales (Sellwood 2000; Sellwood & Kosowski 2001; Binney & Evans 2001). It is clear that some effect of this kind must exist, but a realistic estimate of its magnitude is still difficult for current analytic works (Weinberg 1985) and *N*-body simulations (Debattista & Sellwood 2000).

The *COBE* map of the Galaxy and the high rate of microlensing events toward the Galactic bulge have been used as arguments that most of the inner Galaxy's mass is in a heavy stellar bar (Zhao 1996b; Zhao & Mao 1996). This leaves very little room for a massive dark matter halo (Binney, Bissantz, & Gerhard 2000). We address these issues in this paper.

Modeling the Milky Way and M31 turns out to be challenging because of their complex formation processes. The present-day dark matter halos can have density profiles very different from the original dark matter prediction (e.g., the NFW profile). They are squeezed and deformed by the gravitational force of the baryons, which collapse out of the NFW halo to form heavy disks and dense nuclear bulges. The amount of squeezing also depends on the angular momentum transfer between the baryons and the dark matter halo. The dark matter cusp is elusive to detection because the gravitational force at the center is dominated by the baryons. Apart from detailed data, the Milky Way provides a good laboratory for studying the transfer of angular momentum from the rapidly rotating baryons to the dark matter halo through dynamical friction. Such a transfer must have taken place in the MW and M31 since both have a prominent triaxial bulge/bar and spiral arms in their disks.

Studying the MW halo can set general limits on dark matter halos only if the MW is typical among galaxies that form in a dark matter halo of a given mass. This "typicality" hypothesis can be tested to some degree with our parallel study of the nearby galaxy M31. It can also be tested by producing many realizations of a Milky Way-like galaxy using a Monte Carlo–based semianalytic model of galaxy formation. The semianalytic approach allows us to make statistical predictions of the formation histories of ensembles of galaxies hosted by halos of a given mass and to incorporate the combined stochastic effects of distributions of mass accretion and merging history, halo spin parameter, and gas accretion and star formation history. This approach has the unique ability to allow us to estimate the chances of finding a certain combination of structural parameters that simultaneously produce the observed stellar populations and gravitational potential in galaxies like the MW and M31. We pursue this question in a companion paper (R. Somerville, A. Klypin, & H. Zhao 2002, in preparation, hereafter Paper II).

In summary, the goal of this series of papers is to test whether the current standard theory of galaxy formation, which seems to be very successful on large scales and very problematic on dwarf galaxy scales, can produce L_* galaxies such as the MW and M31. In this paper, this is done by combining classical dynamical modeling of detailed data for the MW and M31 with current cosmologically based models of dark halos and disk formation.

2. MODEL INGREDIENTS

2.1. *Mass Components*

In traditional mass models of the MW one decomposes the baryons into several descriptive components: the

nucleus, the bulge, the bar, the spheroid, the thin disk, the thick disk, and the cold interstellar medium in the disk. It is somewhat problematic to determine where a component starts and ends, and this invariably adds to confusion of terminology. For example, the bar sits in the overlap region between the bulge and the disk, and most stars in this region rotate in the same sense with similar angular speed. So is the bar a distinct identity on top of an oblate bulge, is it merely the triaxial bulge, or is it the nonaxisymmetric part of the disk inside corotation? The lack of a three-dimensional map of the Milky Way makes it uncertain whether or not the disk is truncated inside the corotation of the bar. These distinctions are subjective at some level and sometimes unnecessary when the important quantity is the total mass distribution of the baryons.

We take a different approach. We divide up the mass of the Galaxy into only two components: a dark halo component and a summed-up baryonic component. We describe this baryonic sum by a function $M_b(< r)$, which is the combined mass of baryons inside the radius r. It is understood that the mass distribution is far from spherical. In fact, it is highly flattened, and triaxial inside $\sim$3.5 kpc. We incorporate some minimal modeling of the dynamical effects of flattening and triaxiality, but without repeating earlier more rigorous three-dimensional triaxial dynamical models (Zhao 1996a, 1996b; Häfner et al. 2000).

Our models for the bulge/bar are motivated by the Zhao (1996b) model of the Galactic bulge/bar. That model used *COBE* DIRBE dereddened infrared maps (Weiland et al. 1994) and a collection of stellar kinematics data in the direction of the central parts of the Galactic bulge. The fit to the data indicated that the bar has an elongated boxy triaxial shape. Zhao (1996a, 1996b) added a steep oblate nucleus to the best-fit so-called G2 model of Dwek et al. (1995). The nucleus and a massive central black hole are important for modeling the compressed density of the dark matter within 100 pc. The total mass of the bulge, the bar, and the nucleus was $(2.2 \pm 0.2) \times 10^{10}\ M_\odot$. No dark matter was explicitly included in the model of Zhao (1996b). A massive Miyamo-Nagai analytic disk potential was used to make the rotation curve flat within 8 kpc. Here, instead of the Miyamo-Nagai disk, we use a double-exponential disk as in Kent, Dame, & Fazio (1991). The combined density of nucleus, bulge/bar, and disk are modeled by the following three components:

$$\rho_b = \rho_1 + \rho_2 + \rho_3 \,, \tag{1}$$

$$\rho_1 = \rho_{1,0}(s_1)^{-1.85}\exp(-s_1) \,, \tag{2}$$

$$\rho_2 = \rho_{2,0}\exp\left(-s_2^2/2\right) \,, \tag{3}$$

$$\rho_3 = \rho_{3,0}\exp(-s_3) \,, \tag{4}$$

where $\rho_{1,0}$, $\rho_{2,0}$, and $\rho_{3,0}$ are characteristic densities, determined by the corresponding total masses M_1, M_2, and M_3 of the components ρ_1, ρ_2, and ρ_3. The dimensionless radii s_1, s_2, and s_3 are given by

$$s_1^2 = \frac{0.6^2(x^2+y^2)+z^2}{z_0^2} \,, \tag{5}$$

$$s_2^4 = \frac{\left[(0.26x)^2+(0.42y)^2\right]^2+z^4}{z_0^4} \,, \tag{6}$$

$$s_3 = \frac{\sqrt{x^2+y^2}+12|z|}{r_d} \,. \tag{7}$$

The above aspect ratios in the components $\rho_1(s1)$, $\rho_2(s2)$, and $\rho_3(s3)$ are taken from Zhao (1996b) to reproduce the shape of the *COBE* bar and the disk of Kent et al. (1991). We take the vertical scale height $z_0 = 400$ pc from Zhao (1996b), and we take the disk scale length to be $r_d = 3000$ or 3500 pc and the distance to the Galactic center to be $R_0 = 8000$ or 8500 pc. These parameter combinations produce acceptable fits to the light distribution of the Milky Way after proper normalization of the three components. Zhao (1996b) normalizes the ρ_1 and ρ_2 components so that they have equal strength at about a distance $z = 0.75z_0$ on the minor axis. This fixes the ratio $\rho_{1,0} : \rho_{2,0}$ so that $\rho_1 + \rho_2$ joins smoothly from an observed power-law nucleus to the observed *COBE* bar; this corresponds to $M_1 : (M_1 + M_2) \sim 0.15$. The ratio $\rho_{2,0} : \rho_{3,0}$ is fixed so that $(M_1 + M_2) : M_3 = 1 : 5$, approximately matching the bulge-to-disk K-band luminosity ratio $1.1 \times 10^{10} : 4.9 \times 10^{10}\ L_\odot$ of Kent et al. (1991). Kent et al. in fact introduced a linear tapering of their disk scale length from 250 pc at the solar radius to 167 pc at the center for better matching of the observed surface brightness map. This makes their disk less massive than our double-exponential disk if both are normalized to the surface density and the volume density at the solar radius. The same is true of the triaxial models of Freudenreich (1998), where the axisymmetric disk is tailored into a triaxial hole inside 3.5 kpc.

We do not introduce this fine tailoring of the disk, instead favoring the mathematical simplicity of a ρ_2 bar and a ρ_3 disk, rather than a true bar or disk. The extra mass contained in our ρ_2 disk is at the price of reducing the mass of our ρ_3 bar given the same budget of mass or light within the central 3.5 kpc.

We have deliberately chosen not to label the three components ρ_1, ρ_2, and ρ_3 as nucleus, bar and disk given the uncertain truncations of these latter components. Instead they should be treated as convenient functions to describe the overall distribution of baryons. There is no separate thick disk component. When computing the rotation curve we approximate ρ_3 as a razor-thin disk and ρ_1 and ρ_2 as spherical. The enclosed baryonic mass is computed by spherically averaging ρ_b:

$$M_b(< r) = m_{\rm BH} + 4\pi\int_0^r \langle\rho_b\rangle r^2\, dr \,, \tag{8}$$

where

$$\langle\rho_b\rangle \equiv \int \rho_b \frac{d\Omega}{4\pi} \,, \tag{9}$$

and we include a central black hole of mass $m_{\rm BH} = 2.6 \times 10^6\ M_\odot$ (Genzel et al. 2000) in the Milky Way models and $m_{\rm BH} = 3.5 \times 10^7\ M_\odot$ in the M31 models (Kormendy 1988; Kormendy & Bender 1999). The total mass of the Galaxy inside a radius r is computed as

$$M(r) = M_b(r) + M_{\rm dm}(r) \,, \tag{10}$$

where $M_{\rm dm}(r)$ is the mass profile of the dark matter. We use equation (8) in our calculations for the baryon mass profile,

but a more user-friendly analytic approximation would be

$$M_b(<r) = m_{\rm BH} + 0.025 M_b^{\rm vir}\left[1-\exp(-2.64 r^{1.15})\right] + 0.142 M_b^{\rm vir}\left[1-(1+r^{1.5})\exp(-r^{1.5})\right] + 0.833 M_b^{\rm vir}\left[1-\left(1+\frac{r}{r_d}\right)\exp\left(-\frac{r}{r_d}\right)\right], \quad (11)$$

where r is the radius in kpc, $M_b^{\rm vir} = M_1 + M_2 + M_3$ is the total mass in cooled baryons and r_d is the scale length of the disk.

In our Milky Way models the parameters of the bulge are kept close to those of the original Zhao (1996b) bulge. This is done to preserve the agreement with the *COBE* data and with estimates of stellar radial velocities. Another reason is the microlensing counts, which may be a problem for models of the Milky Way (Zhao & Mao 1996). The rate of microlensing events was estimated for the Zhao (1996b) bulge and was found to be consistent with the constraints on the counts (Peale 1998).

2.2. *Dark Matter Halo Properties*

A sparse set of observational data can generally be fitted by a large range of models, and sometimes unphysical models give excellent fits. Dehnen & Binney (1998) give plenty of examples of this kind. To obtain physically interesting models, it is important to impose constraints based on a physical theory of halo formation and to select models that are not only compatible with data but are also physically well motivated. In this paper we use three results from halo formation theory: the shape of the mass density profile, the halo concentration-mass relationship, and the distribution of halo spin parameter (angular momentum).

Throughout this paper, we assume that the Hubble constant H_0 is equal to 70 km s^{-1} Mpc^{-1}, that the universe is flat, and that the contribution of matter (dark and baryonic) to the critical density is equal to $\Omega_0 = 0.3$. We assume that initially the dark matter density profile is described by an NFW profile (Navarro et al. 1997):

$$\rho_{\rm halo}(r) = \frac{\rho_s}{x(1+x)^2}, \quad x = r/r_s , \quad (12)$$

$$M_{\rm halo}(r) = 4\pi\rho_s r_s^3 f(x) \quad (13)$$

$$= M_{\rm vir} f(x)/f(C) , \quad (14)$$

$$f(x) = \ln(1+x) - \frac{x}{1+x} , \quad (15)$$

$$C = r_{\rm vir}/r_s , \quad (16)$$

$$M_{\rm vir} = \frac{4\pi}{3}\rho_{cr}\Omega_0\delta_{\rm th} r_{\rm vir}^3 , \quad (17)$$

where C and $M_{\rm vir}$ are the halo concentration and virial mass, and $r_{\rm vir}$ is the virial radius. In the above equations, the parameter $\rho_{\rm cr}$ is the critical density of the universe and $\delta_{\rm th}$ is the overdensity of a collapsed object in the "top-hat" collapse model ($\delta_{\rm th} \approx 340$ for our cosmological model). Two independent parameters—C and $M_{\rm vir}$—completely define all relevant halo properties.

The more general density profiles of dark matter halos should be a double power law of the form $\rho \propto x^{-\gamma}(1+x^{\alpha})^{(\gamma-\beta)/\alpha}$, where x is the rescaled radius, while γ, β, and α prescribe the steepness of the inner cusp, the outer slope, and the sharpness of the transition, respectively (Zhao 1996a). For CDM-type structure formation models the dark matter has an inner cusp $\rho_{\rm halo} \propto r^{-\gamma}$ with γ in the range 1–1.5 (e.g., Moore et al. 1998; Ghigna et al. 2000; Klypin et al. 2001). However, for radii larger than 0.5%–1% of the virial radius (1–3 kpc for our galaxies) the difference between these profiles and equation (12) is relatively small (Klypin et al. 2001). Baryonic infall has a tendency to reduce this difference even more, as we will discuss in § 4.2. Because the baryonic compression is very large and still quite uncertain, and because of the relatively small differences in the density profiles, we will not consider profiles with a steeper $1 < \gamma \leq 1.5$ inner slope here.

While we have introduced the halo concentration C as an independent parameter, in practice it is found to be strongly correlated with the virial mass (Navarro et al. 1997; Bullock et al. 2001). For the ΛCDM model, Bullock et al. (2001) give

$$C = 15\text{–}3.3\log(M_{\rm vir}/10^{12}\ h^{-1}\ M_{\odot}) . \quad (18)$$

Statistical ensembles of halos at fixed mass show 40% rms fluctuations of the concentration around this mean. For $10^{12}\ M_{\odot}$ halos this gives a range of $C = 11$–21. Halos of this mass may host galaxies of all types, not only the Sb galaxies considered in this paper. One may argue that perhaps elliptical, lenticular, and Sa galaxies form in high-concentration halos and LSB galaxies are hosted by low-concentration halos. Thus, the range for "normal" Sb galaxies may be somewhat narrower than the full spread for all halos. Tentatively, we assume that the concentration should be in the range $C = 10$–17.

The angular momentum of dark matter halos is also theoretically constrained. Halos in cosmological simulations have small amounts of angular momentum, characterized by the dimensionless spin parameter

$$\lambda = J|E|^{1/2}/GM_{\rm vir}^{5/2} , \quad (19)$$

where J is the angular momentum and E is the total energy of the halo. The spin parameter has a lognormal distribution, which does not depend (or has very little dependence) on cosmological parameters, halo mass, or redshift (Barnes & Efstathiou 1987; Lemson & Kauffmann 1999; Vitvitska et al. 2001). We use the parameters given by Vitvitska et al. (2001). The maximum of the distribution of λ is at $\lambda = 0.035$ and λ has a 90% probability of being in the range

$$\lambda = 0.02\text{–}0.10 . \quad (20)$$

Just as in the case of the concentration, it is likely that there is a correlation of morphological type and spin, with Sb galaxies avoiding halos with both extremely low and high values of λ.

For each set of model assumptions, we can estimate the spin parameter of its precollapse halo (halo before baryonic infall). After the fitting is done and we have the rotation curve, we find the angular momentum of the disk: $J_{\rm disk} = 2\pi\int V_{\rm rot}\Sigma(r)r^2 dr$, where $\Sigma(r)$ is the surface density of the disk. We then assume that the specific angular momentum of the whole system $j = J/M_{\rm vir}$ is equal to the specific angular momentum of the present-day disk $j_{\rm disk} = J_{\rm disk}/M_{\rm disk}$. This gives us the total angular momentum J. Using the parameters of the precollapse halo, we estimate the potential energy of the system and, assuming virial

equilibrium, find the spin parameter. The assumption of virial equilibrium should be quite reasonable because our systems are not interacting or merging.

The assumption that $j = j_{\rm disk}$ is a sensible and conventional starting point and should require little further discussion. We neglect the angular momentum of the bulge in our calculations. This does not give large errors because the angular momentum of bulges in our models is much smaller that the angular momentum of disks. Because we restrict our modeling to late-type galaxies, bulge masses are also small (<20% of the disk mass), so in the worst case we are making 20% error. The original angular momentum of the dark matter halo comes from gravitational (tidal) interactions with its environment. Thus, the dark matter and the gas experience the same torque in the process of halo assembly and should initially have (almost) the same specific angular momentum.

The main uncertainty is what happens to the gas and its angular momentum from the moment it crosses the virial radius until it finally settles into the disk. As we shall find, a significant fraction of the gas must not end up in the disk. In our models, about half of the gas is not found in the disk or bulge. The models presented here do not address the question of where that gas is—whether it is still inside the virial radius, was expelled from the halo, or was never in the halo. For simplicity, we assume that the specific angular momentum was not affected by whatever processes (presumably some form of feedback) were responsible for removing the gas from the disk or preventing its collapse. Another issue is the exchange of angular momentum between the baryons and the dark matter. Some redistribution of the angular momentum may have happened, but we do not know how significant it was. We consider two types of models, with and without exchange of angular momentum. In both cases, within a specific set of assumptions about the angular momentum exchange (or lack thereof), we can thus find the initial angular momentum of the gas that is presently in the disk.

2.3. *Baryonic Compression*

We start with models with no exchange of angular momentum between different components. The effects of baryonic infall are treated following same approach as that of Blumenthal et al. (1986), Flores et al. (1993), and Mo, Mao, & White (1998). It is assumed that the compression occurs adiabatically and the angular momentum of each component is preserved as the baryons settle into the exponential disk. Assuming additionally that the velocity anisotropy is preserved, one obtains the following relation:

$$G[M_b(r) + M_{\rm dm}(r)]r = GM_{\rm halo}(r_i)r_i \ , \qquad (21)$$

where j is the specific angular momentum, r_i is the average radius of a dark matter particle before the baryonic compression, $M_b(r)$ and $M_{\rm dm}$ are the baryonic mass and dark matter mass inside the final radius r, and the $M_{\rm halo}(r_i)$ is the halo mass before contraction, given by

$$M_{\rm halo}(r_i) = \frac{M_{\rm dm}(r)(\Omega_b + \Omega_{\rm dm})}{\Omega_{\rm dm}} \ . \qquad (22)$$

Equation (21) can be considered as an equation for the initial radius r_i corresponding to a given final radius r. This equation is solved numerically.

2.4. *Exchange of Angular Momentum between Baryons and DM*

Models with exchange of angular momentum between the baryons and dark matter are more complicated. The exchange probably happens at late stages of the baryonic infall when the baryon density becomes large and a nonaxisymmetric component may develop as a result of the excitation of spiral waves and/or barlike modes. Giant molecular clouds may play some role in this process. Dynamical friction can then result in a transfer of angular momentum from the baryons to the dark matter. Because the dark matter gains angular momentum, it moves farther from the Galactic center. Thus, the density of the dark matter in the central region decreases. In the early stages of galaxy formation, when most of the baryons were still in gaseous form, we might expect that those nonaxisymmetric features were more prevalent and more dynamically important than at the present time, when most of the baryons are locked in stars.

It is difficult to estimate the exact amount of angular momentum that would be lost by the baryons in such a situation, but we know that it cannot be very large. It is constrained by two factors: the angular momentum of the disk at present and the condition that the initial spin parameter should be close to the typical (median) value for halos. Models without the exchange of angular momentum require $\lambda \approx 0.02$–0.03, which is already close to the most probable spin $\lambda = 0.035$ of a dark matter halo. If the disk loses some angular momentum during its assembly, the estimate of the initial spin parameter increases almost linearly with the lost angular momentum. If too much angular momentum is lost, the spin parameter becomes unacceptably large and we would be forced to conclude that our Galaxy formed from the extreme tail of the distribution. Still, because the distribution of spin parameter is quite broad, there is some room for angular momentum exchange. The spin parameter can be 2 times larger than for the no-exchange models and still be considered "typical." Thus for our model with angular momentum exchange, we assume that the disk loses a factor of 1.5–2 of its angular momentum during the collapse. For simplicity we assume that the formation of the disk happens in two stages. During the first stage, when the baryons experience most of the collapse, they preserve their angular momentum. During the second stage the disk shrinks further, losing some of its angular momentum to the dark matter.

We use the approach outlined in the previous section to compute the state of the system at the end of the first stage of adiabatic compression (in which angular momentum is conserved). We then consider a spherical shell of dark matter with radius r, thickness dr, density $\rho_{\rm dm}$, and specific angular momentum

$$j = rV_c = \sqrt{G[M_b(r) + M_{\rm dm}(r)]r} \ . \qquad (23)$$

It has a total mass $dM_{\rm dm}$, and total angular momentum of

$$dJ = j\, dM_{\rm dm} \ , \qquad (24)$$

$$dM_{\rm dm} = 4\pi\rho_{\rm dm} r^2\, dr \ . \qquad (25)$$

We estimate the amount of angular momentum lost by the baryons (disk and bulge) when a mass element of baryons ΔM_b moves from radius $r + dr$ to radius r. The baryons

lose angular momentum

$$dJ_b = dM_b\left[\left(V_c + \frac{dV_c}{dr}dr\right)(r + dr) - V_c r\right]. \quad (26)$$

This angular momentum is deposited into the dark matter, which produced the dynamical friction. Thus, the final angular momentum of the DM shell is

$$dJ_f = dJ + dJ_b\ . \quad (27)$$

Shells at different distances should be affected because the dynamical friction is not a local process. Nevertheless, for simplicity, we deposit all the angular momentum into the shell through which the disk or bulge element moves. This is a reasonable approximation because (1) a significant fraction of the dynamical friction is due to the elements of dark matter closest to the sinking baryonic material and (2) the volume affected by the deposition of the angular momentum for all shells is very extended—up to 20 kpc.

A shell of dark matter at radius r, which acquired specific angular momentum,

$$j_f - j = dJ_b/dM_{\rm dm}\ , \quad (28)$$

then moves to a new radius r_f. Here j_f is the final specific angular momentum and is related to the mass model by

$$j_f^2 = GM(r_f)r_f\ . \quad (29)$$

Equating initial and final angular momenta for each shell, we get an implicit equation for the final radius r_f:

$$j_f = j\left[1 + \frac{A\Delta M}{4\pi\rho_{\rm dm}r^3}\right], \quad (30)$$

$$A = 1 + \frac{r}{V_c}\frac{dV_c}{dr}\ , \quad (31)$$

$$\Delta M = M_{b,f} - M_b\ . \quad (32)$$

Here $M = M_{\rm dm} + M_b$ is the total mass inside a radius r. Equation (30) is solved numerically. The solution also gives the mass inside a final radius r_f. Equation (30) has the same structure as equation (21). The only difference is the term on the right-hand side, which is the correction due to angular momentum deposition.

We can get a rough estimate of the effect if we neglect the term A (which is close to unity) and introduce the average density excess produced by infalling baryons $\Delta\rho = \Delta M/(4\pi r_i^3/3)$. If we further assume that $M \propto r$ (a good approximation for most radii), then equation (30) takes the form

$$r_f \approx r_i\left(1 + \frac{\Delta\rho}{3\rho_{\rm dm}}\right). \quad (33)$$

It is clear that during the initial stages of the collapse, when the density of the baryons was about 10 times smaller than the density of the dark matter, the exchange of angular momentum had little impact on the dark matter: $r_f \approx r_i$. The effect peaks at around $\Delta\rho = 3\rho_{\rm dm}$. At even larger values of the density ratio, the approximation fails because a small amount of dark matter cannot exert significant dynamical friction on a large mass of baryons. Nevertheless, at the peak of its importance, the effect is potentially quite large, with $r_f \approx 2r_i$, resulting in a decrease of the dark matter density by a factor of 10.

3. OBSERVATIONAL CONSTRAINTS

3.1. *Constraints for the Milky Way*

In this section we describe the observational data that we use to confront our models.

Satellite dynamics and modeling of the Magellanic Clouds (e.g., Zaritsky et al. 1989; Fich & Tremaine 1991; Lin, Jones, & Klemola 1995; Kochanek 1996) provide constraints on the mass of our Galaxy on large scales ($\simeq$50–100 kpc). The mass of the Milky Way inside 100 kpc is estimated to be $(5.5 \pm 1) \times 10^{11}\ M_\odot$ from dynamics of the Magellanic Clouds (Lin et al. 1995). Using constraints from the escape velocity and motions of satellite galaxies, Kochanek (1996) estimates the mass of the Galaxy inside 100 kpc to be $(5–8) \times 10^{11}\ M_\odot$. We adopt the following constraints on the mass inside 100 kpc (Dehnen & Binney 1998):

$$M_{R<100\,{\rm kpc}} = (7 \pm 2.5) \times 10^{11}\ M_\odot\ . \quad (34)$$

For the outer part of the rotation curve of our Galaxy we use the data summarized in Dehnen & Binney (1998). The radial circular velocity v_r of an object at Galactocentric coordinates l and b is related to the circular velocity at radius r, $v_c(r)$, by

$$W(r/R_0) = \frac{v_r}{\sin l\cos b} = \frac{R_0}{r}v_c(r) - v_c(R_0)\ , \quad (35)$$

where R_0 is the Sun's distance to the Galactic center. Data for H II regions (Brand & Blitz 1993) and for classical Cepheids (Pont et al. 1997) are used to derive $W(r/R_0)$ for up to twice the solar distance.

The surface density of gas and stellar components at the solar radius is estimated by Kuijken & Gilmore (1989) to be

$$\Sigma_{\rm stars+gas} = 48 \pm 8\ M_\odot\ {\rm Mpc}^{-2}\ . \quad (36)$$

We also use another local constraint: the vertical force K_z at 1.1 kpc above the Galactic plane. Using data on the kinematics of K dwarfs, (Kuijken & Gilmore 1989, 1991) give the limit on the total density of matter inside 1.1 kpc:

$$|K_z(R_0, z = 1.1\ {\rm kpc})|/2\pi G = (71 \pm 6)\ M_\odot\ {\rm Mpc}^{-2}\ . \quad (37)$$

Oort's constants,

$$A = (v_c/R - dv_c/dR)/2\ , \quad (38)$$

$$B = -(v_c/R + dv_c/dR)/2\ , \quad (39)$$

provide constraints on the circular velocity curve at the solar radius. We adopt the same values as Dehnen & Binney (1998):

$$A = 14.5 \pm 1.5\ {\rm km\ s^{-1}\ kpc^{-1}}\ , \quad (40)$$

$$B = -12.5 \pm 2\ {\rm km\ s^{-1}\ kpc^{-1}}\ , \quad (41)$$

$$A - B = 27 \pm 1.5\ {\rm km\ s^{-1}\ kpc^{-1}}\ . \quad (42)$$

The corresponding $v_c = (A - B)R_0$ at the solar neighborhood is 200–240 km s^{-1}.

We use observations of terminal velocities $v_{\rm terminal}$ to constrain the mass distribution and the rotational velocities inside the solar radius. If l is the Galactic longitude, then for an axisymmetric model the terminal velocity is related to the circular velocity v_c by

$$v_{\rm terminal} = v_c(R_0\sin l) - v_c(R_0)\sin l\ , \quad (43)$$

where R_0 is the distance to the Galactic center. We make use of data from the H I surveys of Knapp, Stark, & Wilson (1985) and Kerr et al. (1986). The terminal velocities are compatible with the results of Malhotra (1995). Dehnen & Binney (1998) present a more comprehensive comparison of different surveys.

To constrain the mass in the central parts of the MW, we use data on stellar motions. We make use of the results of Genzel et al. (2000), who found $M = 3 \times 10^7\ M_\odot$ for the mass inside 10 pc and $M = 10^7\ M_\odot$ at 4 pc. Kinematics of OH/IR stars (Lindqvist, Habing, & Winnberg 1992) constrain the mass on the $\approx$100 pc scale.

More indirect constraints on the mass of baryons and dark matter in the central part of the galaxy can be obtained from the existence of persistent rapidly rotating bars and from the observed optical depth to microlensing events in the direction of the Galactic center. Based on the former, the work of Debattista & Sellwood (2000) suggests that dark matter must make up less than a quarter of the dynamical mass within one disk scale length. However, the halo models used in that work were rather unphysical $n = 3$–5 polytropic cored models, and so it is difficult to know whether these results apply to cuspy CDM halos. From the microlensing constraints, Zhao & Mao (1996) concluded that all of the dynamical mass within 3.5 kpc must be in an optimal stellar bar to account for observed optical depth of $(3\text{–}10) \times 10^{-6}$. More recent data suggest that the optical depth may be much lower than this. Both of these pieces of evidence suggest that models with a minimal dark matter contribution in the inner part are to be preferred. We discuss how our models fare with respect to both of these constraints in §§ 5.2 and 6; however, we do not impose either of these constraints a priori.

We use the K-band luminosity of our Galaxy derived by Drimmel & Spergel (2001):

$$L_K = 8.9 \times 10^{10}\ L_\odot\ , \tag{44}$$

corresponding to a magnitude of $M_K = -24.0$. This value is consistent with the estimate of Malhotra et al. (1996), but it is larger than the value $6 \times 10^{10}\ L_\odot$ adopted by Kent et al. (1991). It is about 0.4–0.6 mag fainter than the value implied by the K-band Tully-Fisher relation of Tully & Pierce (2000), assuming a rotation velocity of 220 km s^{-1} for the Galaxy (see the discussion in § 4.3).

3.2. *M31: Light and Kinematics*

Throughout this paper, we assume that the distance to M31 is 770 kpc, based on Cepheids and red clump stars (Freedman & Madore 1990; Kennicutt et al. 1998; Stanek & Garnavich 1998). In some earlier papers a smaller distance of 690 kpc was used. Results used below were rescaled to our adopted distance.

The M31 galaxy has a more extended disk than our galaxy. Walterbos & Kennicutt (1988) studied the surface brightness distribution of M31 in four colors, U, B, V, and R, tracing the disk to 20 kpc. The exponential length in the R band was found to be 5.7 ± 0.3 kpc, as compared with 2.5–3.5 kpc for our Galaxy. The scale length was found to be larger in bluer bands (6.4 kpc in B). We use the R-band value in our modeling because it is less affected by dust extinction and more likely to trace the overall stellar mass distribution. A significant fraction of the light ($\frac{1}{3}$–$\frac{1}{2}$, depending on the band) is from the bulge. The bulge of M31 is triaxial, as evidenced by a significant twist in isophotes and complex gas velocity patterns. Like the Milky Way, M31 has a nucleus as well. In fact, the nucleus can be resolved into two central peaks separated by about 1 pc. A black hole of mass $\sim 3.5 \times 10^7\ M_\odot$ is harbored in the less luminous peak.

We model the bulge of M31 with a scaled-up version of the Milky Way bulge (eqs. [1]–[4]) and disk, using a disk scale radius of $r_d = 5.7$ kpc. We include a black hole with mass $3.5 \times 10^7\ M_\odot$—more than 10 times more massive than the Milky Way black hole (Kormendy & Bender 1999). R-band photometry along the major axis of M31 is taken from Walterbos & Kennicutt (1987). We also use r-band photometry from Kent (1987). The latter results are shifted up by 0.45 mag to match the R-band surface brightness. We apply a correction of 0.25 mag to both the bulge and disk to account for extinction in our Galaxy and apply an additional correction of 0.74 mag to the disk for internal extinction (Kent 1989).

We use the CO observations of Loinard, Allen, & Lequeux (1995) to determine the rotational velocity inside the central 10 kpc. H I measurements extend the rotation curve to 30 kpc (Brinks & Burton 1984). The CO and H I velocities agree in the overlapping central region of the galaxy. The central 2 kpc region of M31 shows anomalously high velocities and a misalignment of the major axes of the bulge and disk. Those anomalies are attributed to the presence of a triaxial bar/bulge in M31 (e.g., Stark 1977; Stark & Binney 1994; Berman 2001). Just as in the case of the Milky Way, we use a simplified spherical model of the central bulge. The model can not reproduce the details of the circular velocity function in the central 2–5 kpc, but it gives useful constraints on the mass models.

Additional constraints are obtained from stellar kinematics in the central region of M31. We use the estimates of the rotation velocity $v_{\rm rot}$ and the line-of-sight velocity dispersion σ from Table 2 of Kormendy & Bender (1999). We then estimate the mass, assuming an isothermal distribution with isotropic velocities $M = 2(\sigma^2 + v_{\rm rot}^2)r/G$. At distances 25″–50″ ($\approx$100–200 pc) from the center of M31 the rotational velocity is small, the velocity dispersion is almost constant (150–160 km s^{-1}) and the assumed density profile of the nucleus is close to r^{-2}. Thus, our approximation seems to be quite reasonable for these radii.

4. MODELS OF THE MILKY WAY AND M31

4.1. *Comparison with Observations*

Our goal is to find whether a set of models, which we consider plausible or theoretically well motivated, are compatible with the observational constraints. We do not attempt to find best-fit models, but present a range of models and discuss whether they satisfy the constraints we have adopted. Once we find by trial and error that a model is compatible with the constraints, we do not attempt to further improve the quality of the fit by fine-tuning the parameters of the model.

Tables 1–3 list the parameters of our models. Most of the parameters and values have already been introduced and discussed in §§ 2 and 3. The last row in each table gives the ratio of the sum of disk and bulge masses to the mass of baryons expected inside the virial radius ($f_b M_{\rm vir}$). We assume the universal ratio of baryons to dark matter is

TABLE 1
THE MILKY WAY GALAXY: NO EXCHANGE OF ANGULAR MOMENTUM

Parameter	Constraints	Model A_1 (Favored)	Model A_2 (Maximum Disk)	Model A_3 (Maximum Concentration)	Model A_4 (Maximum Halo)
Virial mass $M_{\rm vir}$ ($M_\odot$)	...	1.0×10^{12}	0.71×10^{12}	1.0×10^{12}	2.0×10^{12}
Virial radius $r_{\rm vir}$ (kpc)	...	258	230	258	325
Halo concentration C	10–17	12	5	17	10
Disk mass M_3 (eq. [3]; $M_\odot$)	...	4×10^{10}	6×10^{10}	3.5×10^{10}	4.0×10^{10}
Mass $M_1 + M_2$ (eq. [3]; $M_\odot$)	...	0.8×10^{10}	1.2×10^{10}	0.7×10^{10}	0.8×10^{10}
Stellar mass $r < 3.5$ kpc ($M_\odot$)	...	1.9×10^{10}	3.1×10^{10}	1.5×10^{10}	1.9×10^{10}
Disk exponential scale length r_d (kpc)	2.5–3.5	3.5	3.0	3.5	3.5
Maximum circular velocity $V_{c,\max}$ (km s^{-1})	...	228	246	235	237
Maximun halo circular velocity with no compression (km s^{-1})	...	163	123	178	197
Solar distance, $R_\odot$ (kpc)	7–8.5	8.0	8.5	8.5	8.5
Baryon surface density at $R_\odot$ ($M_\odot$ pc^{-2}	48 ± 8	53	62	40	46
Total surf. density within 1.1 kpc at $R_\odot$ ($M_\odot$ pc^{-2})	71 ± 6	75	74	63	69
Oort's constants $A - B$ (km s^{-1} kpc^{-1})	27 ± 1.5	26.8	28.1	27.6	27.9
Mass inside 100 kpc (10^{11} $M_\odot$)	7.5 ± 2.5	5.8	3.8	6.1	9.0
Disk + bulge mass-to-light ratio $(M/L)_K$ ($M_\odot/L_\odot$)	...	0.54	0.81	0.47	0.54
Spin parameter λ	0.02–0.10	0.031	0.022	0.037	0.018
$M_{\rm dm}/M_{\rm disk+bulge}$ ($r < 3$ kpc)	...	0.95	0.40	1.31	1.02
Fraction of " galactic " baryons	...	0.48	1.00	0.42	0.24

$f_b \equiv \Omega_b/(\Omega_{\rm dm} + \Omega_b) = 0.1$. Our results can be easily adjusted to any other ratio.

With the exception of models A_2 and B_2, all the models pass the constraints we have adopted. Typically the fitting was done by first assuming a mass and concentration for the halo and then by tuning the disk parameters.

In models A_2 and B_2 we used a different approach. Here we started by assuming a disk mass of $M_3 = 6 \times 10^{10}$ $M_\odot$. This mass is only 20%–50% more massive than the value used in our favored models, A_1 and B_1. It is often quoted as the fiducial mass of the disk (Binney & Tremaine 1987). We then tried to find an acceptable fit to the data by changing parameters of the halo. Even assuming an extremely low halo mass and unrealistically low concentration, both models fail. In model A_2 the maximum circular velocity and the surface density at the solar radius are both too large. Model B_2 is a bit better, but still the local surface density is too large, the angular momentum is uncomfortably large, and the mass inside 100 kpc is too small. We can decrease the local surface density by reducing the exponential length of the disk. But that has the side effect of increasing the disk mass inside the central 5 kpc region. As a result, the rotation velocity gets unacceptably large and the model fails again.

Some cautions should be given regarding the bulge mass and the mass distribution inside the central $\approx$3 kpc. The values M_1, M_2, and M_3 are the masses of three interpenetrating components, which are the dominant components in the regions of the nucleus, the bulge/bar, and the disk, respec-

TABLE 2
THE MILKY WAY GALAXY: MODELS WITH THE EXCHANGE OF ANGULAR MOMENTUM

Parameter	Constraints	Model B_1 (Favored)	Model B_2 (Maximum Disk)	Model B_3 (Large Mass)	Model B_4 (Small Exchange)
Virial mass $M_{\rm vir}$ ($M_\odot$)	...	1.0×10^{12}	0.71×10^{12}	1.5×10^{12}	1.0×10^{12}
Virial radius $r_{\rm vir}$ (kpc)	...	258	230	295	258
Halo concentration C	10–17	12	10	10	12
Disk mass M_3 ($M_\odot$)	...	5×10^{10}	6×10^{10}	5×10^{10}	5×10^{10}
Mass M_1 + M_2 (M$_\odot$)	...	1×10^{10}	1.2×10^{10}	1×10^{10}	1×10^{10}
Stellar mass $r < 3.5$ kpc ($M_\odot$)	...	2.7×10^{10}	2.8×10^{10}	2.7×10^{10}	2.6×10^{10}
Disk exponential scale length r_d (kpc)	2.5–3.5	3.0	3.5	3.0	3.0
Intermediate exponential scale r_d (kpc)	...	6.0	7.0	6.0	4.5
Maximum circular velocity $V_{c,\max}$ (km s^{-1})	...	223	216	225	234
Maximun halo circular velocity, no infall (km s^{-1})	...	163	139	178	163
Solar distance $R_\odot$ (kpc)	7–8.5	8.5	8.5	8.5	8.5
Baryon surface density at $R_\odot$ ($M_\odot$ pc^{-2})	48 ± 8	52	69	52	52
Total surface density within 1.1 kpc at $R_\odot$ ($M_\odot$ pc^{-2})	71 ± 6	68	79	70	72
Oort's constants $A-B$ (km s^{-1} kpc^{-1})	27 ± 1.5	26.5	25.4	26.5	28.0
Mass inside 100 kpc (10^{11} $M_\odot$)	7.5 ± 2.5	6.0	4.6	7.7	6.0
Bulge + disk mass-to-light ratio $(M/L)_K$ ($M_\odot/L_\odot$)	...	0.67	0.81	0.67	0.67
Spin parameter λ	0.02–0.10	0.058	0.076	0.044	0.050
$M_{\rm dm}/M_{\rm disk+bulge}$ ($r < 3$ kpc)	...	0.24	0.14	0.24	0.31
Fraction of " galactic " baryons	...	0.6	1.0	0.4	0.6

TABLE 3
PARAMETERS OF MODELS FOR THE M31 GALAXY

Parameter	Model C_1 (With No Exchange)	Model C_2 (With Exchange)
Virial mass M_{vir} ($M_\odot$)	1.60×10^{12}	1.43×10^{12}
Virial radius r_{vir} (kpc)	300	290
Halo concentration C	12	12
Disk mass M_d ($M_\odot$)	7.0×10^{10}	9.0×10^{10}
Bulge mass M_b ($M_\odot$)	1.9×10^{10}	2.4×10^{10}
Disk exponential scale r_d (kpc)	5.7	5.7 (8.55)
Stellar mass $r < 3.5$ kpc ($M_\odot$)	2.8×10^{10}	3.6×10^{10}
DM mass $r < 3.5$ kpc ($M_\odot$)	2.5×10^{10}	1.8×10^{10}
Bulge $(M/L)_R$ ($M_\odot/L_\odot$)	3.0	3.8
Disk $(M/L)_R$ ($M_\odot/L_\odot$)	0.93	1.2
Maximum circular velocity $V_{c,max}$ (km s^{-1})	269	262
Max. halo circular velocity $V_{halo,max}$ without compression (km s^{-1})	184	176
Mass inside 100 kpc radius ($M_\odot$)	8.5×10^{11}	8.1×10^{11}
Spin parameter λ	0.036	0.057
Fraction of "galactic" baryons	0.56	0.80

tively. The ρ_3 component is a constant-height thin exponential disk, which remains as such even inside the central $\approx$3 kpc region occupied by the Galactic bar. This is unphysical and is just a simplification for the convenience of the fitting. In reality the disk should be mixed with the bar-bulge in the central region and should constitute a single component with the distribution of the central nonaxisymmetric bar. These complexities are only important for a transitional region between the bulge and the disk at radii 2–4 kpc, where the disk mass is comparable to the bulge mass. Inside 2 kpc the disk is much smaller than the bulge and it does not matter how we treat the disk.

Some important quantities for the best models of the Milky Way and M31 are presented in Figures 1–5. Figure 1 compares the observed terminal velocities in our Galaxy with the predictions of model A_1. The quality of the fit

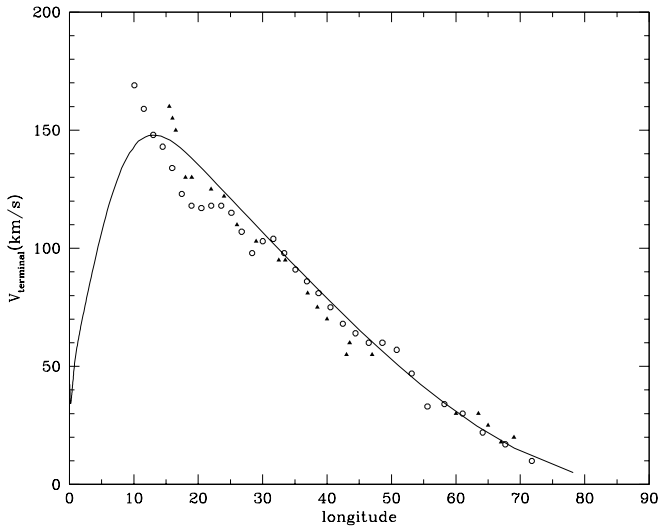

FIG. 1.—Dependence of terminal velocities on galactic longitude l. The full curve is for model A_1. Symbols show observational data from the H I measurements of Knapp et al. (1985; *circles*) and Kerr et al. (1986; *triangles*). At small angles the deviations from circular velocities are expected to be large owing to the central bar. This is clearly seen at $l < 20^o$.

decreases at small angles because of noncircular motions produced by the central bar. Even larger distortions are found in the rotation curve of M31, which has a large bar. The usual way to deal with the deviations is to ignore the data at small angles. For example, Dehnen & Binney (1998) ignore data for angles smaller than 17°, which corresponds to $\approx$2.5 kpc. Olling & Merrifield (1998) used the data of Malhotra (1995), which also start at $\approx$2.5 kpc. Without a more realistic (nonaxisymmetric) model of the bar one cannot do better than this. The problem is that we can not exclude the possibility that small 5–10 km s^{-1} disturbances related to the bar can be felt even at larger radii. This could bias our parameter estimates, which are sensitive to what happens at those radii. For this reason, we still consider the fit produced by model A_1 to be acceptable in spite of the fact that for longitudes 20°–30° the model values exceed the observational points by 5–10 km s^{-1}.

Figure 2 shows the rotation curve of our Galaxy. The observational data are compared with models A_1 and B_1. We also show the contributions of different components. For both models the dark matter dominates in the outer parts of the Galaxy, but the radius at which the contribution of the dark matter equals that of the baryonic component are very different. For model A_1 the masses contributed by baryons and dark matter are equal at 3.5 kpc. For model B_1 this radius is 11 kpc. The difference is due to a combination of two factors: expulsion of dark matter by dynamical friction and an increase in the assumed disk+bulge mass. In the central 5 kpc of the Galaxy the contribution of the dark matter relative to the disk and bulge is very different for models A_1 and B_1. In the case of model A_1 and in all other models with no exchange of angular momentum, the dark matter is never a strongly dominant component, but it always contributes 40%–60% of the total mass. In other words, these models have submaximal disks (Bottema 1997; Courteau & Rix 1999). In the case of models with exchange of the angular momentum, there is little dark matter in the central region of our Galaxy.

Figure 3 shows the distribution of mass in the Milky Way galaxy for a very large range of scales, from 5 pc to 200 kpc. The following observational constraints are used. The first two points at 5–10 pc are from studies of stellar radial veloc-

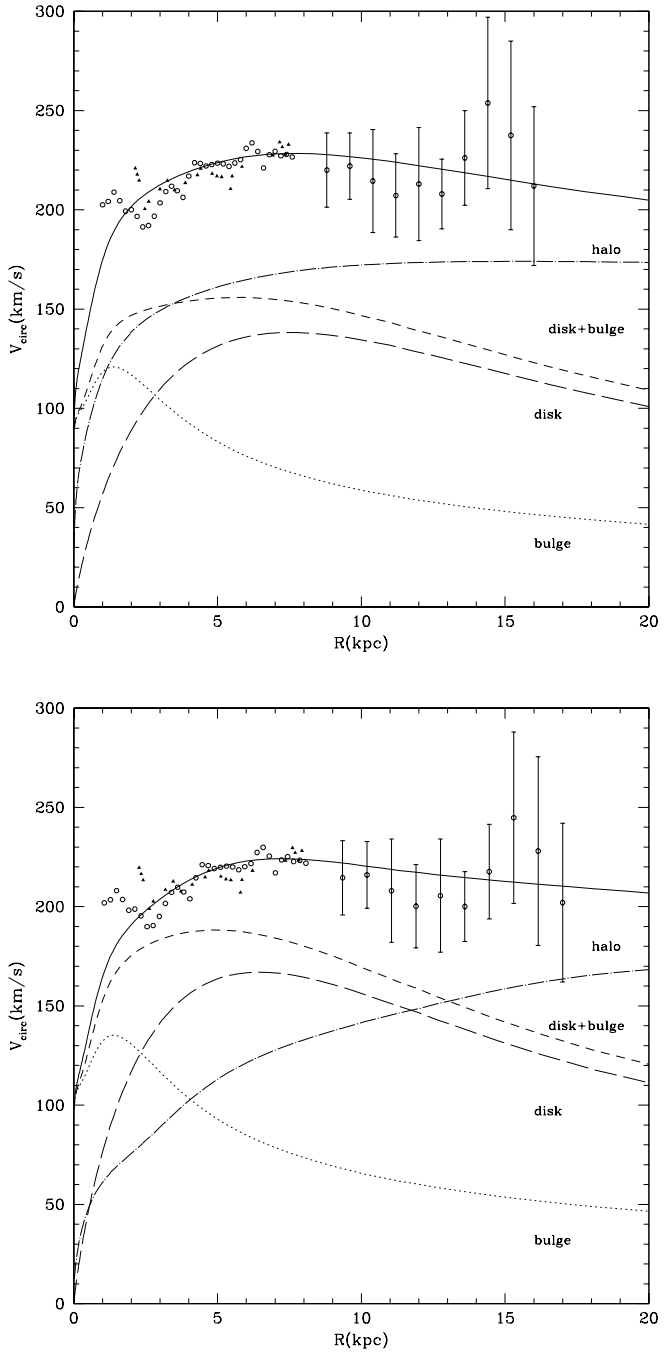

FIG. 2.—Rotation curve for our favorite models A_1 (no exchange of angular momentum) and B_1 (with the exchange). Note that the dark matter dominates only in the outer part of the Milky Way. Symbols show observational data from H I measurements of Knapp et al. (1985; *circles*) and Kerr et al. (1986; *triangles*).

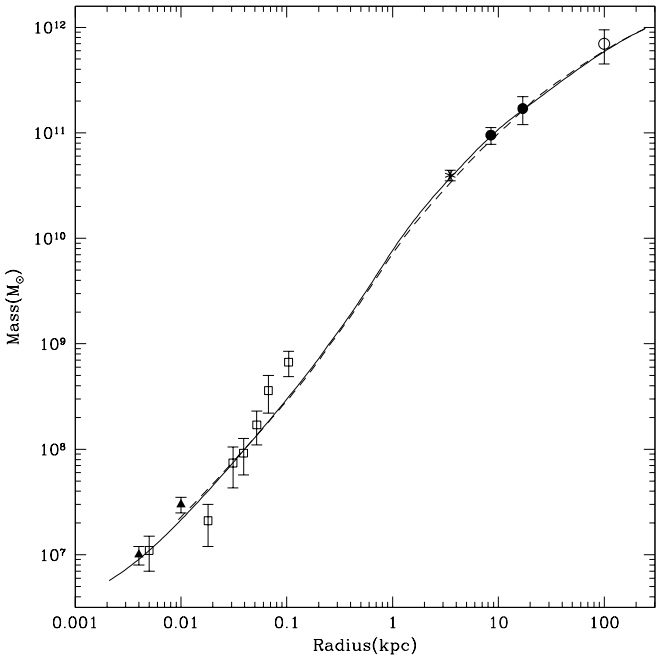

FIG. 3.—Mass distribution of the MW galaxy for model A_1 (*full curve*) and model B_1 (*dashed curve*). The large dots with error bars are observational constraints. From small to large radii the constraints are based on the following: stellar radial velocities and proper motions in the Galactic center; radial velocities of OH/IR stars; modeling of the bar using DIRBE and stellar velocities; rotational velocity at the solar radius; and dynamics of satellites.

ities and proper motions in the Galactic center. Squares are based on kinematics of OH/IR stars (Lindqvist et al. 1992). The point at 3.5 kpc is based on the Zhao (1996a, 1996b) model of the bar. Because the model was compared with the data on stellar kinematics (inner rotation curve and radial velocity dispersion), it gives a constraint on the total mass: $4 \times 10^{10}\ M_\odot$, with an uncertainty of about 20%. For the next data point at 8.5 kpc we simply assume that the circular velocity is 220 ± 20 km s^{-1}, which covers the whole range of reasonable values. We then estimate the mass as $M = v^2 r/G$. The last observational point is the constraint from the motions of satellite galaxies discussed in § 3. The central data points were not used either in our fitting or in the analysis of the bulge (Zhao 1996b). Nevertheless, they come fairly close to the extrapolation of our model into the very center of our Galaxy. The theoretical curves for our favored models A_1 and B_1 are very close to each other, which is not surprising because they fit the same data and have the same global dark matter content. The largest deviation of the models from the data is for the mass inside 100 pc, where the observational estimate is twice larger than the prediction of the models. Even at this point the disagreement is not alarming because the observational data are likely more uncertain than the formal error.

What is remarkable about Figure 3 is that it spans more than 5 orders of magnitude in radius and mass. It is encouraging that, without fine-tuning, our models are consistent with observations of the dynamical mass of the MW over this huge range.

Finding an acceptable model for M31 was relatively easy because there are much less data. In particular, we do not have kinematic constraints for the disk, which would be equivalent to constraints at the solar position in our Galaxy. Our model seems to reproduce reasonably well the dynamical mass of M31 from 100 pc to ≈100 kpc. Our model does not produce the very large wiggles exhibited by the observed rotation curve. The wiggles at 5 and 9 kpc are likely due to noncircular motions induced by the bar and, thus, as discussed before, cannot be reproduced by any axisymmetric model. The bulge of M31 is almost twice as massive as the bulge of our Galaxy. It is also slightly (30%) more compact. The disk of M31 is also more massive, but it is more extended. As a result, in the central 5 kpc of the M31 the

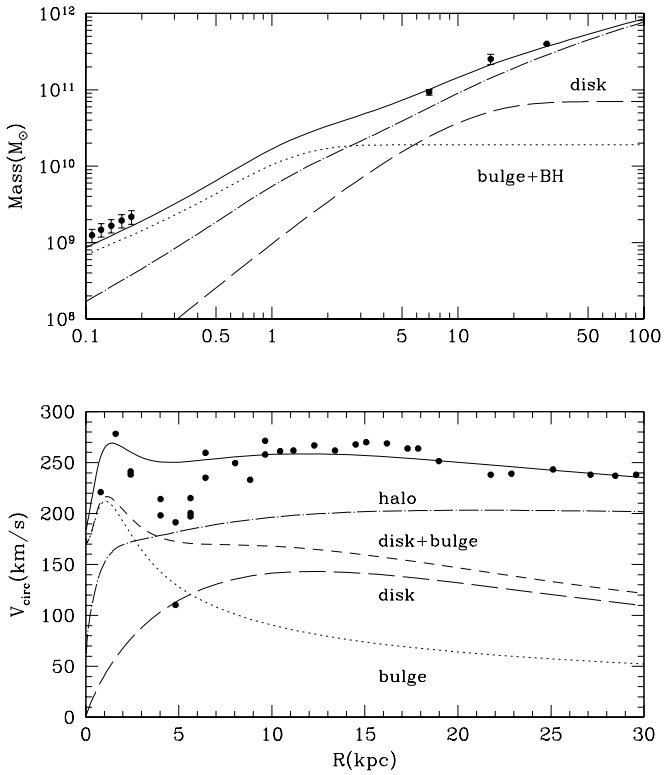

FIG. 4.—Mass distribution (*top*) and rotation velocity (*bottom*) of the M31 galaxy for model C_1. The large dots with error bars are observational constraints. In the bottom panel the circles show results of CO ($r < 10$ kpc) and H I ($r > 10$ kpc) observations. The large dip at around 5 kpc is likely due to noncircular motions induced by the bar. Observational data points in the top panel at small radii are from stellar motions in the nucleus. Vertical bars correspond to 20% errors in mass. Data points at 7, 15, and 30 kpc correspond to circular velocities of 240 ± 10, 270 ± 20, and 240 ± 10 km s^{-1}.

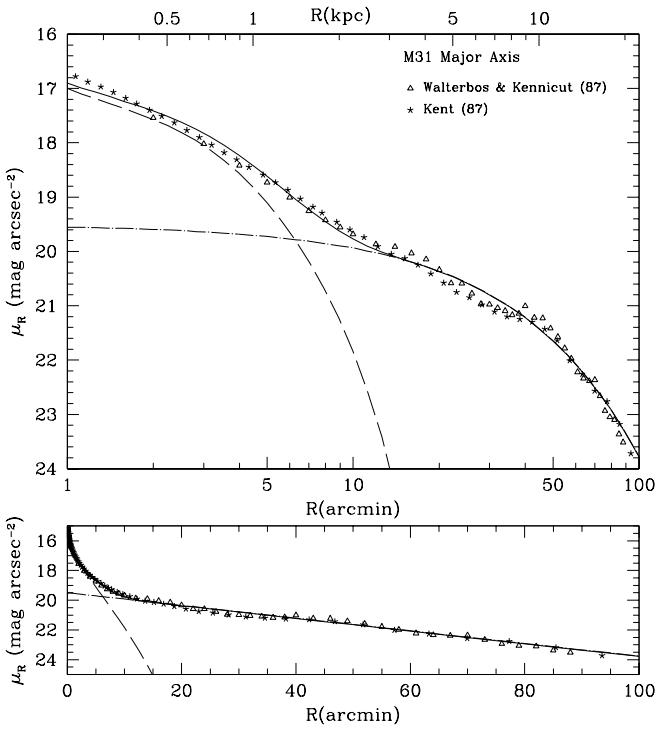

FIG. 5.—Surface brightness of M31 in the *R* band on linear (*bottom*) and logarithmic (*top*) scales. Deviations from the observational results are less than 0.2 mag.

bulge is a much more dominant component as compared with the bulge of our Galaxy.

The surface brightness profile in the *R* band, shown in Figure 5, is used as an additional constraint. An accurate fit (the same as for the mass modeling) is obtained for stellar mass-to-light ratios of $M/L = 0.93\ M_\odot/L_\odot$ and $M/L = 3\ M_\odot/L_\odot$ for the disk and the bulge, respectively. These results are quite consistent with the expectations from stellar population synthesis models for a galaxy with the $B-R$ color of M31, which imply an overall stellar mass-to-light ratio $(M/L)_R = 0.85\ M_\odot/L_\odot$ (Bell & de Jong 2001). The bulge M/L is likely overestimated, and it would be reduced by a correction for internal dust absorption. For example, internal absorption of 0.25 mag inside the bulge would reduce M/L to 2.4. Our estimates of the M/L ratios are significantly different from those of Kent (1989), who give $M/L = 5\ M_\odot/L_\odot$ and $M/L = 10.4\ M_\odot/L_\odot$ for the bulge and the disk. Most of the differences come from differences in bands (*r* instead of *R*) and from the fact that Kent's luminosities were not corrected for absorption. If we include those corrections (see § 3.2) and also scale the results to the distance of 770 pc, Kent's estimates become $M/L = 2.5\ M_\odot/L_\odot$ for the disk ($M/L \approx 6.7$ in *B*) and $M/L = 2.4\ M_\odot/L_\odot$ for the bulge. Such a large M/L ratio for the disk, more typical of an old bulge population, seems a bit problematic. The main reason that Kent (1989) obtains such a large M/L lies in the assumed profile of the dark matter: constant density through the whole galaxy. This results in the extreme case of a maximal disk and very little dark matter inside 20–25 kpc.

4.2. *Effects of Compression by Baryonic Infall*

The sinking of baryons into the central part of a galaxy increases the depth of the gravitational potential. This leads to an increase of the dark matter density. To some degree, this process provides coupling of the baryons and the dark matter. As a result, in the case of adiabatic compression without the exchange of the angular momentum, it appears difficult to find a case where one of the components (dark matter or baryons) is significantly larger than the other. Figure 2 (*top panel*) gives an example of this effect. From 100 pc to about 6 kpc the contributions of baryons and dark matter to the circular velocity are about equal.

We use two simple models to study the effect of adiabatic compression on the ratio of dark matter and baryons. In the first case we take model A_1 and change the mass of baryons while keeping the disk scale length and the bulge-to-disk ratio constant. In the second case, we vary the scale length of the disk, keeping the mass of the disk the same as in model A_1; we ignore the bulge to make the case more clear. Figure 6 shows the ratio of the circular velocity produced by the dark matter to that of the baryons.

Figure 6 illustrates that for a realistic range of parameters the contribution of baryons can not be significantly larger that of the dark matter. Even in the case where the disk is 4 times more massive or 4 times more compact than in model A_1, the contribution of the dark matter to the rotation curve is at least $\frac{1}{3}$. In other words, in models with cuspy halos, with adiabatic compression and no exchange between components it is difficult to have a maximum disk.

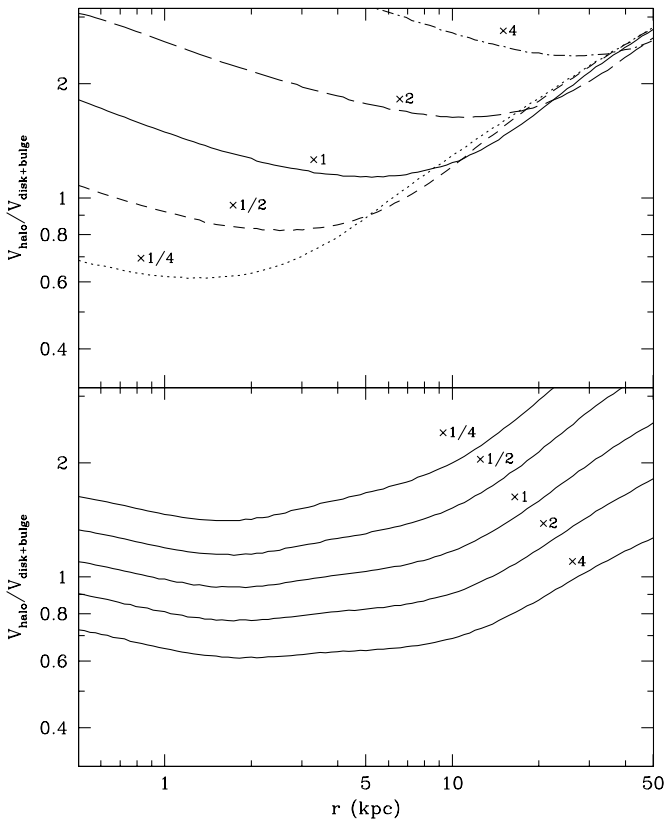

FIG. 6.—Ratio of the circular velocity produced by the dark matter to that of the baryons for two models of adiabatic compression. For the top panel the exponential disk scale length r_d was changed by a factor indicated in the plot. The factor of unity (×1) is for the assumption of a disk mass $4 \times 10^{10}\ M_\odot$ and scale length $r_d = 3.5$ kpc. This model has no bulge. Curves in the bottom panel show the same quantity when the disk scale length is held fixed at $r_d = 3.5$ kpc the total baryonic mass is varied. Parameters of the ×1 curve are the same as for model A_1.

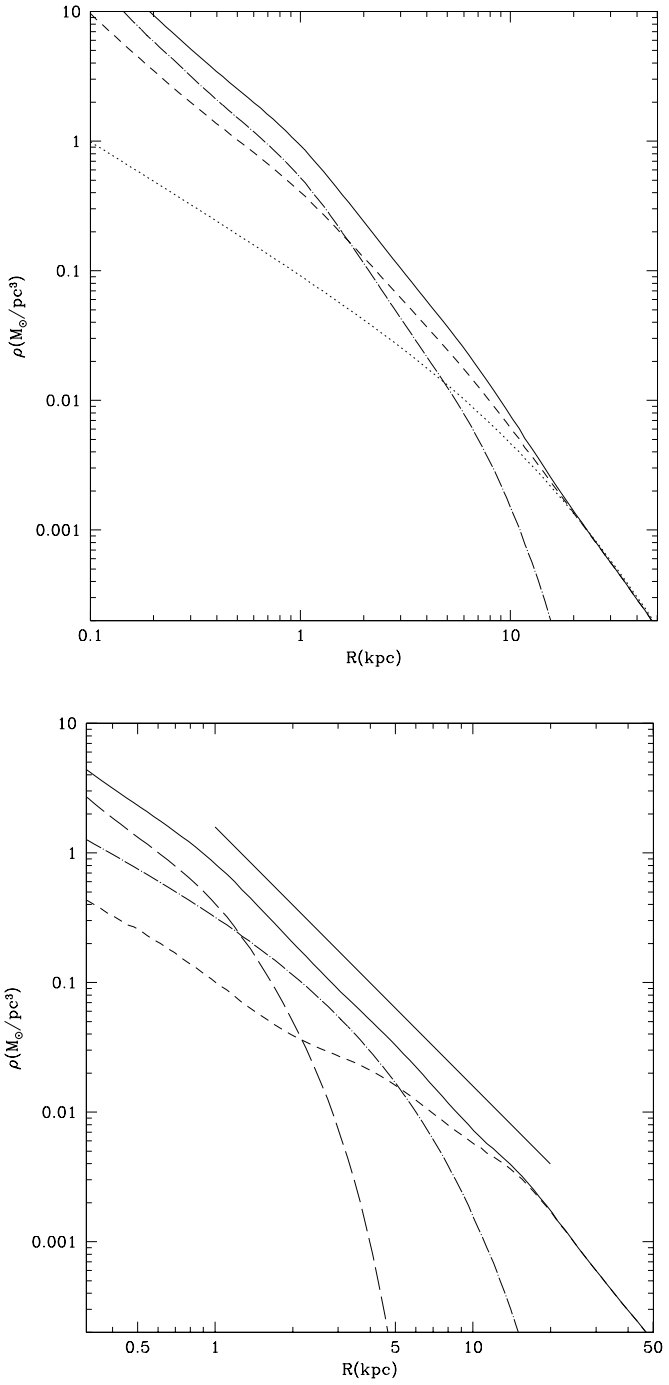

FIG. 7.—Density distribution for different components in models A_1 (*top*) and B_1 (*bottom*). Different curves show the total density (*solid curve*) and dark matter (*short-dashed curve*). *Top:* The total bulge+disk density is shown by the dot-dashed curve. Inside ≈3 kpc the baryons and the dark matter closely follow one another. This illustrates the self-adjusting nature of the compression. The dotted curve shows the dark matter before the compression by baryons. The compression is clearly a dominant effect inside the central 10 kpc. *Bottom:* The bulge is shown by the long-dashed curve; the disk is shown by dot-dashed curve. The components have very different profiles, but their sum is very close to a power law with slope −2, shown by the straight line. This "disk-halo conspiracy" is the natural result of the adiabatic dark matter compression.

It is interesting to note that the opposite is also true: in the central part of a galaxy it is difficult to have a model where the mass of the dark matter is much larger than that of the baryons. The model with a 4 times more extended disk in the top panel (×4) is an example of what should represent a low surface brightness galaxy. Its surface density is 16 times smaller than for our Galaxy, and its mass in stars is only 1% of the virial mass, yet the baryons at the solar radius contribute about $\frac{1}{4}$ of the rotational velocity.

In view of this self-adjusting behavior of the dark matter, it is easier to understand the profiles of different components in realistic models. The top panel in Figure 7 shows the dark matter and the combined disk and bulge density profiles. For comparison we also show the dark matter profile before the baryonic infall. Note that the change in the dark matter profile is quite large, yet the final profile manages to follow the baryonic density very closely in the central region. The bottom panel shows different components for model B_1. Each component has a complicated shape. Different components become prominent at different radii. The net effect is rather remarkable—the sum of all components is very close to a power law with slope −2. The deviations from this power law are less than 10% on scales of 1–20 kpc. One may argue that this occurs because we are fitting a rotation curve that is nearly flat, and so we get a r^{-2} density profile. This is not exactly true. The observed deviations are quite substantial. The final fit is much closer to a pure power law than what is required by the observations. This "disk-halo conspiracy" is the natural result of the coupling of baryons and dark matter via the adiabatic compression.

4.3. *Constraints from the Global Luminosity Function*

If the luminosity of our Galaxy is typical of that of other disk-type galaxies that form in halos of the same mass, this also has implications for the global luminosity function. We now investigate the number density of galaxies like the Milky Way, using the revised Press-Schechter approximation of Sheth & Tormen (1999). The Sheth-Tormen approximation provides us with the number density of dark matter halos as a function of their mass, per unit mass interval, $dn/dM_h(M_h)$. Once we have pinned down the luminosity of the Milky Way, we can calculate the mass-to-light ratio within the virial radius of the dark matter halo hosting the galaxy in a given model. We can then convert the halo number density per mass interval to a galaxy number density per magnitude interval as a function of magnitude $[dn/d(\mathrm{mag})]$, which may be compared with the observed luminosity function.[2] We neglect the contribution from galaxies that are "satellites" in much larger mass halos (i.e., cluster galaxies) and assume that the adopted mass-to-light ratio holds for disk galaxies only. We then compare this prediction with the *K*-band luminosity function for all galaxies and for late-type galaxies from the 2MASS survey (Kochanek et al. 2001).

We use two different approaches to estimate the *K*-band luminosity of the Galaxy. The total luminosity of our Galaxy is notoriously difficult to measure directly because we live in the middle of it, and dust extinction is important even in the *K* band. Drimmel & Spergel (2001) recently performed a detailed analysis in which they estimated the *K*-band luminosity of the Milky Way by fitting a model to the *COBE* DIRBE photometric data. They found $M_K = -24.0$ for the total luminosity of the Galaxy.

The other approach is to assume that the Galaxy is "typical" and should lie on the Tully-Fisher relation (TFR). Unfortunately, this approach is also uncertain because the Tully-Fisher data in the *K* band are sparse and the results in the literature are somewhat inconsistent. If we use the TFR of Malhotra et al. (1996), assuming that the velocity width of the Milky Way $\Delta V = 2V_c = 440$ km s^{-1}, then we obtain $M_K = -23.37$, about 0.7 mag fainter than the direct DIRBE measurement. If we use instead $\Delta V = 480$ km s^{-1}, which Malhotra et al. obtain by scaling from M31 [$\Delta V(\mathrm{MW}) = (220/250)\Delta V(\mathrm{M31})$, with $\log_{10}[\Delta V(\mathrm{M31})] = 2.737$], then we get -23.69, only 0.3 mag fainter than the direct measurement. This would place the MW about 1 σ brightward of the TFR, which seems not unreasonable. However, the K'-band TFR recently published by Tully & Pierce (2000) and Rothberg et al. (2000) gives $M_K = -24.4$ for $\Delta V = 440$ km s^{-1}. The slope is very similar to the relation of Malhotra et al., but the zero-point is almost 1 full magnitude brighter. This problem is only exacerbated if we consider that the Cepheid calibration data used by Malhotra et al. is consistent with a Hubble parameter of $H_0 = 71$ km s^{-1} Mpc^{-1}, while the ground-based K' data imply $H_0 = 81$ km s^{-1} Mpc^{-1}. If we scale the zero-point of the ground-based data to the lower value of $H_0 \simeq 70$ km s^{-1} Mpc^{-1}, all the galaxies are further away and therefore are even brighter, making the discrepancy worse (the Milky Way would then have $M_K = -24.7$).

[2] The relevant transformation is $dn/d(\mathrm{mag}) = (\Upsilon_h L)/[-2.5\log_{10}(e)]dn/dM_h(M_h) = -0.921\ M_h\ dn/dM_h(M_h)$, where $\Upsilon_h = M_h/L$ is the mass-to-light ratio within the virial radius of the halo, and L is the assumed luminosity of the Milky Way.

Malhotra et al. (1996) base their TFR on photometry from DIRBE and have a total of seven galaxies in their sample (including the Milky Way). The ground-based K'-band data considered by Rothberg et al. (2000) and Tully & Pierce (2000) are from the Ursa Major sample (see also Verheijen 2001) and additional galaxies from the Pisces filament, comprising a total of 69 galaxies with four galaxies having Cepheid distances. All of these galaxies have multiband photometry in *B*, *R*, *I*, and K', and the TF relations in the optical bands agree well with those from other samples, including the very well-studied *I* band. Therefore it seems that the zero point for the ground-based sample should be more reliable. This suggests that perhaps there is some sort of calibration offset in the DIRBE photometry, and perhaps calls into question the validity of using even the Drimmel & Spergel (2001) result. We shall consider the implications of using three different choices for the total *K*-band luminosity of the Galaxy: the direct DIRBE result, the raw ground-based TFR result, and the ground-based TFR result scaled to $H_0 = 70$ km s^{-1} Mpc^{-1}. The three sets of squares plotted in Figure 8 correspond to these three values.

In each case the higher of the two connected squares shows the number density that we would obtain if every halo hosted a disk galaxy. The lower point shows the value obtained if only half of the halos host disk-type galaxies

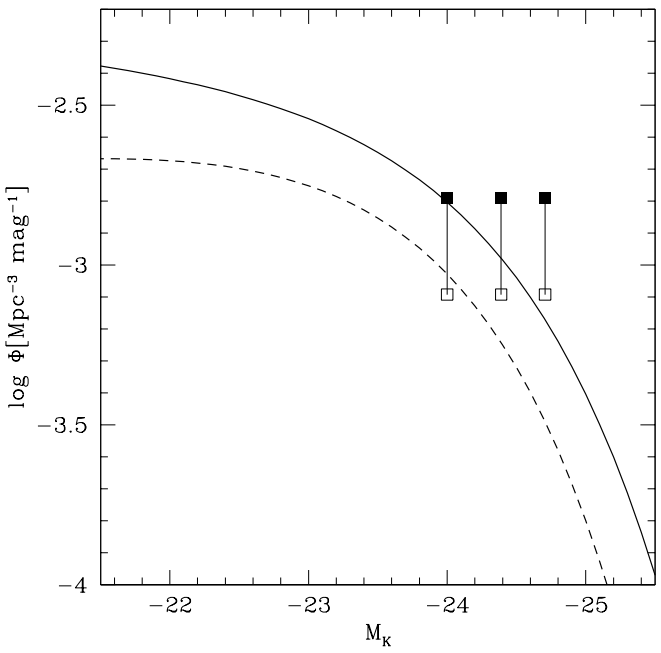

FIG. 8.—*K*-band galaxy luminosity function. The curved lines show the observed *K*-band luminosity function from the 2MASS survey (Kochanek et al. 2001) scaled to $H_0 = 70$ km s^{-1} Mpc^{-1}; the solid line shows the results for all types of galaxies and the dashed line shows the results for late-type galaxies only. Points show the prediction of the number density of halos hosting "Milky Way" galaxies, using the halo virial mass from model A_1 or B_1 and an assumed Milky Way luminosity. Solid symbols show the value obtained if all halos are assumed to host disk galaxies, and they are to be compared with the solid lines. Open symbols show the result obtained if only half of the halos host disk galaxies and are to be compared with the dashed lines. The three sets of connected points are for three different assumed values for the total *K*-band luminosity of the Milky Way (see text). The "bright" and "intermediate" Milky Way luminosities imply a contradiction between the number density of dark matter halos predicted by CDM and the observed luminosity function, while the "faint" luminosity is marginally consistent.

(probably a reasonable lower limit, based on semianalytic modeling). The points can easily be scaled up or down for different assumptions about the fraction of disk-hosting halos, but for reasonable assumptions the result should lie between the lower and upper points. Unfortunately, the interpretation of the luminosity function constraint is quite sensitive to the uncertain value of the MW's total luminosity. For the bright normalization (from the TFR of Rothberg et al. (2000) scaled to $H_0 = 70$ km s^{-1} Mpc^{-1}), the number density of halos harboring MW galaxies is about a factor of 2 higher than the observations. For the faint normalization (the direct value from Drimmel & Spergel 2001), the number density of MW halos is marginally consistent with the observed luminosity function.

It is of some concern that for the Milky Way normalization that we regard as the most reliable (the ground-based TFR scaled to the standard value of H_0), the luminosity function constraint is violated. If this normalization turns out to be correct, then this may suggest that the Milky Way is harbored by a halo with a larger virial mass. The number density of dark matter halos drops fairly rapidly with virial mass so this relaxes the constraint somewhat. Figure 9 shows the number density of dark matter halos as a function of their virial mass, normalized by the observed number density of galaxies at the assumed luminosity of the Milky Way from Kochanek et al. (2001). We see that the faint and even the intermediate Milky Way luminosity can be accommodated fairly comfortably within the mass range allowed by our dynamical modeling ($10^{12}\ M_\odot < M_{\rm vir} < 2\times 10^{12}\ M_\odot$, with brighter Milky Ways favoring larger virial masses), however, the bright Milky Way normalization is too high even at the upper virial mass limit of $2 \times 10^{12}\ M_\odot$.

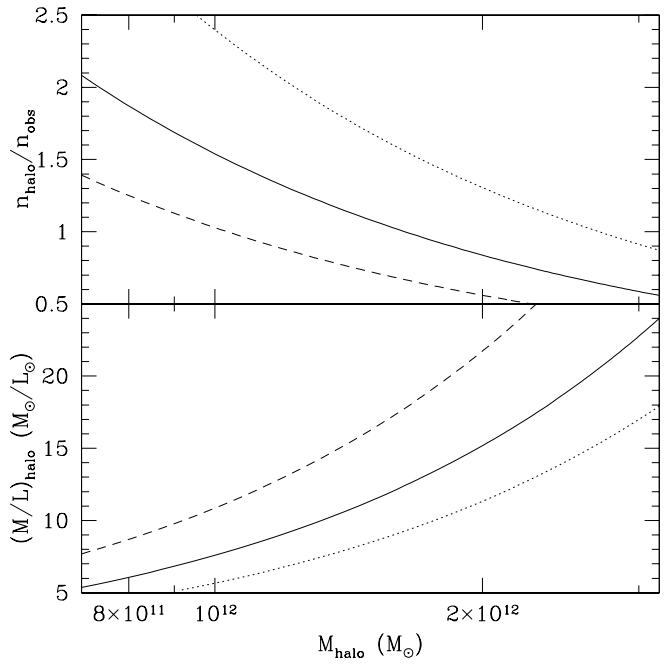

FIG. 9.—Constraints on the virial mass of the Milky Way from the observed luminosity function. The top panel shows the ratio of the predicted to observed number densities of Milky Way galaxies as a function of the assumed average virial mass of dark matter halos hosting galaxies with the luminosity of the Milky Way. The three lines correspond to the three values of the Milky Way luminosity discussed in the text, where the dotted line is the bright normalization, the solid line is the intermediate normalization, and the dashed line is the faint normalization. The plotted line should be less than unity in order for the model to satisfy the luminosity function constraint. Larger masses have lower number densities and therefore more easily satisfy the constraint. The bottom panel shows the halo mass-to-light ratio (i.e., the halo virial mass divided by the total galaxy luminosity) as a function of halo mass. Again, the three lines correspond to the three different Milky Way luminosities considered in the text.

5. DISCUSSION

There are numerous aspects of these models of the MW and the M31 galaxies that are important to consider. We start with very small scales and proceed to larger ones.

5.1. *Central Region and the Black Hole*

It is now quite well established that black holes of mass $\sim 10^6$–$10^8\ M_\odot$ reside in the centers of most galaxies. The case for a $2.6 \times 10^6\ M_\odot$ black hole at the center of the Milky Way is demanded by rising radial velocity and proper motion dispersions within a few parsecs of the center (Ghez et al. 2000 and references therein). Recently, Gondolo & Silk (1999) showed that if the black hole (BH) is grown adiabatically, the central dark matter will be drawn into a dense spike, with density exceeding $10^8\ M_\odot$ pc^{-3}. This would make the Galactic center very luminous in terms of neutrino flux if the dark matter is made of annihilating neutralinos. Ullio, Zhao, & Kamionkowski (2001), however, argue that it is unlikely for the BH to grow all its mass while at rest in the center because of galaxy merging in CDM models. A spiraling-in BH would actually "heat" and reduce the central density of the dark matter. The seed is likely to be massive ($\geq 10^5\ M_\odot$), because otherwise it would not have enough time to spiral into the center by dynamical friction. Realistic models of loss-cone capture show that the rate of stellar capture is around $10^6\ M_\odot$ per Hubble time (Magorrian & Tremaine 1999). This implies that the BH could at most double its mass during the adiabatic phase.

Following Ullio et al. (2001) we take into account the effect of adiabatic growth from a seed BH of half the final value. The flux is very likely to be small because in our models the final peak of the dark matter density outside the event horizon of the BH is well below $10^8\ M_\odot$ pc^{-3}, which is the amount needed for significant annihilation.

5.2. *Microlensing Constraints*

Microlensing events toward the Galactic center provide a lower limit on the mass of baryons within the inner part of the Galaxy. Combined with the observed terminal velocities, this places upper limits on the mass of dark matter that can be present and implies that the inner Galaxy must be mostly baryonic with a massive bulge and disk. The microlensing events toward the Galactic bulge are still not completely understood observationally and theoretically, in particular the nature of two long-duration events near the $l = 3°$, $b = -3°$ field. Several values for the spatially averaged optical depth are reported with large discrepancies in the mean values. The earlier values are too high for axisymmetric bulge and disk models, and they are barely consistent even with a massive bar. A $2.2 \times 10^{10}\ M_\odot$ bar pointing about 13°–20° away from the Galactic center–Sun line plus a full disk with a mass of $1.7 \times 10^{10}\ M_\odot$ inside 3.5 kpc are required by detailed dynamical models of the *COBE* map and by the kinematics of the bulge stars (Zhao 1996b). Zhao, Spergel, & Rich (1995) found that the bulk of the lenses are on the near side of this elongated bar. This

bar+disk model can produce an optical depth consistent with the lower 1 σ error bar of the observed microlensing optical depth of $(3.3 \pm 1.2) \times 10^{-6}$ for all stars observed by the Optical Gravitational Lensing Experiment (OGLE; Udalski et al. 1994) toward Baade's window ($l = 1°$, $b = -4°$). Zhao, Rich, & Spergel (1996) found that this model with a flat initial mass function with very few brown dwarfs can fit the microlensing event duration distribution. Peale (1998) studied the optical depth and the distribution of durations of lensing events for the bulge with the shape used in our paper. Both statistics were found to be compatible with observational data.

Zhao & Mao (1996) pointed out that except for very contrived models, the higher reported optical depth values are in conflict with the total lensable material in the inner Galaxy. The optical depth for the low-latitude red clump subsample from MACHO (Alcock et al. 1996) is in the range $(3\text{–}10) \times 10^{-6}$ at the 95% confidence level. This would require contrived models involving the Sun–Galaxy center line being the diagonal line of a homogeneous rectangular bar. This discrepancy has been highlighted by Kuijken (1999) and by Binney et al. (2000) for axisymmetric models. Both confirmed that the need of extremely contrived bar and disk models if the optical depth is as large as $(3\text{–}10) \times 10^{-6}$.

The most recent analysis of bulge microlensing, however, suggests a lower optical depth in general. Statistics have improved, as we now have over 500 events. In particular, the value for red clump giants is the range $(1.4\text{–}2) \times 10^{-6}$ depending on whether the long duration events are included (P. Popowski 2001, private communications). Sevenster & Kalnajs (2001) suggested that the 3 kpc arm might correspond to a stellar ring and also account for a large fraction of the microlenses.

Without making detailed models of the nonaxisymmetric bar and the ring, we can do a quick check of whether our models are consistent with the microlensing data by scaling the results of Zhao et al. (1996b) and Peale (1998) for the masses of baryonic components in our models. If we use only the ρ_2 bar component (mass $10^{10}\ M_\odot$), then the optical depth is $\sim 10^{-6}(M_2/10^{10}\ M_\odot) \sim 10^{-6}$ (Zhao et al. 1995). In our model A_1 there is another $\sim 10^{10}\ M_\odot$ of baryons inside 3.5 kpc, which is formally in the ρ_3 axisymmetric disk. As we have discussed, this thin unperturbed disk inside the nonaxisymmetric bar is only a mathematical simplification. It is unlikely that it can stay thin or even can survive inside the bar. The most probable scenario is that the bar was formed from the disk at some epoch, when the disk became unstable. In this case the disk becomes the bar and we should expect that the inner disk has the same distribution as the bar. In this case, the microlensing counts should include all the baryonic mass inside 3.5 kpc, which doubles the estimate of the optical depth. Then our best models A_1 and B_1 are consistent with an optical depth of $(1.4\text{–}2) \times 10^{-6}$.

5.3. *Comparison with Other Work*

Our estimate for the virial mass of the M31 halo $1.4 \times 10^{12}\ M_\odot$ is consistent with earlier estimates that used kinematics of satellites. Courteau & van den Bergh (1999) found a total mass of $(1.33 \pm 0.18) \times 10^{12}\ M_\odot$ inside 260 kpc, and Evans et al. (2000) and Evans & Wilkinson (2000) found $1.23^{+1.8}_{-0.6} \times 10^{12}\ M_\odot$. The main difference between our results and Evans et al. (2000) is that they assumed a halo density profile very different from an NFW profile.

For the Milky Way halo, our models A_1 and B_1 predict a mass $M_{R<100\,\mathrm{kpc}} = (5.8\text{–}6) \times 10^{11}\ M_\odot$, consistent with what Dehnen & Binney (1998) find in their "standard models 1–4," $M_{R<100\mathrm{kpc}} = (6.0\text{–}6.6) \times 10^{11}\ M_\odot$. Our disk also has a mass of $(4\text{–}5) \times 10^{10}\ M_\odot$, consistent with their finding of $M_d = (4.2\text{–}5.1) \times 10^{10}\ M_\odot$.

Our results and conclusions differ from those of Hernandez, Avila-Reese, & Firmani (2001), who model the Milky Way in the framework of a cosmological scenario. Hernandez et al. (2001) find that the standard halo profiles are not compatible with the observed properties of the Milky Way and that "the rotation curve for our Galaxy implies the presence of a constant density core in its dark matter halo." The gross features of our approach and that of Hernandez et al. (2001) are similar. The parameters of the cosmological models are only slightly different. Adiabatic infall corrections are also used, and the dark matter profiles are also similar. The main difference was in the mass of the halo assumed for the Galaxy. The average virial mass in the models of Hernandez et al. (2001) was assumed to be $M_{\mathrm{vir}} = 2.8 \times 10^{12}\ M_\odot$, which is almost 3 times larger than the virial mass in our favored models A_1 and B_1. If we ran our models for this large virial mass, we would also reject the model (our model A_4 has the largest virial mass, $M_{\mathrm{vir}} = 2.0 \times 10^{12}\ M_\odot$, and even this model is marginal, with a rather large rotation velocity and an abnormally low concentration). The reason that Hernandez et al. (2001) adopted such a large virial mass is their perhaps overly stringent interpretation of the Kochanek (1996) estimate of the mass inside 50 kpc. Hernandez et al. (2001) assume that the mass was $M_{50} = (4.9 \pm 0.5) \times 10^{11}\ M_\odot$, while Kochanek (1996) gives $M_{50} = (3.2\text{–}5.5) \times 10^{11}\ M_\odot$ at the 90% confidence level if Leo I is excluded from the analysis. For comparison, our models A_1 and B_1 have $M_{50} = (3.7\text{–}3.8) \times 10^{11}\ M_\odot$, which is compatible with Kochanek (1996). Wilkinson & Evans (1999) give even larger uncertainties on the mass: $M_{50} = 5.4^{+0.2}_{-3.6} \times 10^{11}\ M_\odot$.

There is no doubt that estimates of the mass at large radii, 50–100 kpc, are crucial for constraining the halo parameters. Values of M_{50} larger than $5 \times 10^{11}\ M_\odot$ are very likely to be difficult to reconcile with the standard cosmological models. The problem is that the observational constraints are still very much uncertain.

It is interesting to compare our results with the results of Binney & Evans (2001), who came to the conclusion that "cuspy halos favored by the cold dark matter cosmology (and its variants) are inconsistent with the observational data." Their main argument was a combination of three constraints. (1) The mass of baryonic matter in the disk and bulge inside the solar radius should be larger than $3.9 \times 10^{10}\ M_\odot$ to produce enough microlensing events. (2) The surface density of the dark matter inside 1.1 kpc from the plane at the solar radius should be $30 \pm 15\ M_\odot\ \mathrm{pc}^{-2}$ to satisfy the constraints from local stellar kinematics and stellar and gas content. (3) The sum of all components should reproduce the circular velocity curve at 2–4 kpc. Our models A_1 and B_1 easily satisfy the first two constraints. For example, there is a mass of $4.5 \times 10^{10}\ M_\odot$ of baryons inside the solar radius in our model A_1, and the dark matter surface density is $22\ M_\odot\ \mathrm{pc}^{-2}$. In our modeling we do not use the circular velocity at 2–4 kpc because corrections due to noncircular motions are large and uncertain. Still, even if we accept the Binney &

Evans results for the circular velocity, the differences appear to be small—10 km s^{-1} for 2–4 kpc, which is within the observational uncertainty. We also note that the treatment of the dark matter and the bar by Binney & Evans are not realistic. The profile of the dark matter was assumed to be an unmodified NFW, which as we have discussed is not appropriate for the inner part of our Galaxy, where adiabatic compression will have substantially modified the dark matter profile. The bulge was modeled as an elliptical exponential disk with scale-length 1 kpc. We have used more realistic models for the dark matter and a more detailed bar model that matches numerous observational constraints. We also rely on the more detailed treatment of lensing events made by Zhao et al. (1995) and by Peale (1998) for the same bar used in our analysis. These differences explain why we come to different conclusions.

6. CONCLUSIONS

We study models of the Milky Way and the Andromeda galaxies based on the standard paradigm of disk and dark matter halo formation within a ΛCDM cosmology. The models produced acceptable fits for numerous observational data. Here we itemize our main conclusions.

1. Cuspy NFW profiles may be compatible with observational data on MW and M31.

2. The compression of the dark matter by baryons is a very significant effect in the inner parts of galaxies and must be taken into account. It provides a coupling between the baryons and the dark matter. The result of this coupling is that the total density distribution does not show any wiggles or features corresponding to the transition from one component to another. Thus, the disk-halo "conspiracy" is a natural prediction of our models.

3. The rate of microlensing events is a strong constraining factor for models with standard cuspy halo density profiles. These models are compatible with the recent low estimates of the optical depth if all the stellar material inside 3.5 kpc is optimally distributed in a $\approx 2 \times 10^{10}\ M_\odot$ bar.

4. In none of our models does the dark matter dominate the central parts (≈3 kpc) of the galaxies. Since both the MW and M31 appear to be typical high surface brightness galaxies (HSB), we expect that this is true for other HSB galaxies. If there has been no transfer of angular momentum between the baryons and the dark matter, HSB galaxies should have submaximal disks.

5. Including the effects of a modest transfer of angular momentum from baryons to dark matter (expected during the nonaxisymmetric process of bar formation) can produce models with 2–3 times less dark matter in the central 3 kpc region of the Galaxy. This is likely to sustain fast rotating bars such as those observed in the Milky Way and M31. These models also allow a slightly heavier disk without producing too high a rotation curve.

6. Our dynamical models suggest that the virial mass of the dark matter halo hosting the Milky Way and M31 is in the range $10^{12}\ M_\odot < M_{\rm vir} < 2 \times 10^{12}\ M_\odot$. We use the predicted halo mass function to estimate the number density of such halos and compare this with the observed number density of galaxies. We find that if the *K*-band luminosity of the Milky Way is $M_K = -24$, as found by Drimmel & Spergel (2001) based on direct fits to DIRBE data, then the predicted number density of dark matter halos hosting Milky Way galaxies is not inconsistent with the observed *K*-band luminosity function of Kochanek et al. (2001). If the Milky Way has a larger luminosity of $M_K = -24.4$ to -24.7, as would be the case if it lies on the Tully Fisher relation found by Tully & Pierce (2000) and Rothberg et al. (2000), then the number density of "Milky Way" halos is more than a factor of 2 higher than the observed number density of galaxies at this luminosity. Increasing the virial mass of the Milky Way's halo to the upper limit allowed by the dynamical modeling ($M_{\rm vir} < 2 \times 10^{12}\ M_\odot$) allows the fainter end of the Tully-Fisher range to be accommodated, but not the brightest values.

7. A significant fraction of the baryons within the virial radius of the halo must not be in the disk or bulge of the Milky Way and M31. All acceptable models required that the disk and bulge should contain not more than $\approx \frac{1}{4}$–$\frac{1}{2}$ of the baryons expected to be inside the virial radius of the halo in the absence of feedback. It is beyond the scope of this paper to address the state and the location of the "lost" baryons, but this is in keeping with the usual assumptions of semianalytic models, and we address this issue in Paper II. We note that this conclusion is based on two independent arguments. Models that have more than half of the baryons in the central luminous part fail to produce acceptable fits for either the rotation curve or for the local surface density. When compared with the observed luminosity function, these models also produce too large a number density of bright galaxies (see item 6 above). This is related to the usual "overcooling" problem (see Balogh et al. 2001 for a recent discussion).

We acknowledge support from the grants NAG 5-3842 and NST 98-02787. H. S. Z. acknowledges hospitality while visiting New Mexico State University. A. K. acknowledges hospitality and support from the Institute of Astronomy, Cambridge. We thank James Binney and Rene Walterbos for inspiring discussions.

REFERENCES

Alcock, C., et al. 1996, ApJ, 461, 84
Balogh, M. L., Pearce, F. R., Bower, R. G., & Kay, S. T. 2001, MNRAS, 326, 1228
Barnes, J., & Efstathiou, G. 1987, ApJ, 319, 575
Bell, E. F., & de Jong, R. S. 2001, ApJ, 550, 212
Berman, S. 2001, A&A, 371, 476
Binney, J., Bissantz, N., & Gerhard, O. 2000, ApJ, 537, L99
Binney, J. J., & Evans, N. W. 2001, MNRAS, 327, L27
Binney, J., & Tremaine, S. 1987, Galactic Dynamics (Princeton: Princeton Univ. Press)
Blais-Ouellette, S., Amram, P., & Carignan, C. 2001, AJ, 121, 1952
Blais-Ouellette, S., Carignan, C., Amram, P., & Côté, S. 1999, AJ, 118, 2123
Blumenthal, G., Faber, S. M., Flores, R., & Primack, J. R. 1986, ApJ, 301, 27
Bottema, R. 1997, A&A, 328, 517
Bradač, M., Schneider, P., Steinmetz, M., Lombardi, M., King, L. J., & Porcas, R. 2002, A&A, submitted (astro-ph/0112038)
Brand, J., & Blitz, L. 1993, A&A, 275, 67
Brinks, E., & Burton, W. B. 1984, A&A, 141, 195
Bullock, J. S., Kolatt, T. S., Sigad, Y., Somerville, R. S., Kravtsov, A. V., Klypin, A., Primack, J. P., & Dekel, A. 2001, MNRAS, 321, 559
Bullock, J. S., Kravtsov, A. V., & Weinberg, D. H. 2000, ApJ, 539, 517
Chiba, M. 2002, ApJ, 565, 17
Côté, S., Carignan, C., & Freeman, K. C. 2000, AJ, 120, 3027
Courteau, S., & Rix, H. 1999, ApJ, 513, 561

Courteau, S., & van den Bergh, S. 1999, AJ, 118, 337
Dalal, N., & Kochanek, C. 2002, ApJ, submitted (astro-ph/0111456)
Debattista, V. P., & Sellwood, J. A. 2000, ApJ, 543, 704
de Blok, W. J. G., McGaugh, S. S., Bosma, A., & Rubin, V. C. 2001, ApJ, 552, L23
Dehnen, W., & Binney, J. 1998, MNRAS, 294, 429
Drimmel, R., & Spergel, D. 2001, ApJ, 556, 181
Dwek, E., et al. 1995, ApJ, 445, 716
Einasto, J. 1972, in IAU Symp. 44, External Galaxies and Quasi-Stellar Objects, ed. D. E. Evans (Dordrecht: Reidel), 44, 37
———. 1979, in IAU Symp. 84, The Large-Scale Characteristics of the Galaxy, ed. W. B. Burton (Dordrecht: Reidel), 84, 451
Eke, V. R., Navarro, J. F., & Steinmetz, M. 2001, ApJ, 554, 114
Evans, N. W., & Wilkinson, M. I. 2000, MNRAS, 316, 929
Evans, N. W., Wilkinson, M. I., Guhathakurta, P., Grebel, E. K., & Vogt, S. S. 2000, ApJ, 540, L9
Fich, M., & Tremaine, S. 1991, ARA&A, 29, 409
Flores, R. A., & Primack, J. R. 1994, ApJ, 427, L1
Flores, R., Primack, J. R., Blumenthal, G. R., & Faber, S. M. 1993, ApJ, 412, 443
Freedman, W. L., & Madore, B. F. 1990, ApJ, 365, 186
Freudenreich, H. T. 1998, ApJ, 492, 495
Genzel, R., Pichon, C., Eckart, A., Gerhard, O. E., & Ott, T. 2000, MNRAS, 317, 348
Ghez, A. M., Morris, M., Becklin, E. E., Tanner, A., & Kremenek, T. 2000, Nature, 407, 349
Ghigna, S., Moore, B., Governato, F., Lake, G., Quinn, T., & Stadel, J. 2000, ApJ, 544, 616
Gondolo, P., & Silk, J. 1999, Phys. Rev. Lett., 83, 1719
Häfner, R., Evans, N., Dehnen, W., & Binney, J. 2000, MNRAS, 314, 433
Hernandez, X., Avila-Reese, V., & Firmani, C. 2001, MNRAS, 327, 329
Innanen, K. A. 1973, Ap&SS, 22, 393
Keeton, C. R. 2002, ApJ, submitted (astro-ph/0111595)
Kennicutt, R. C., et al. 1998, ApJ, 498, 181
Kent, S. M. 1987, AJ, 94, 306
———. 1989, AJ, 97, 1614
Kent, S. M., Dame, T. M., & Fazio, G. 1991, ApJ, 378, 131
Kerr, F. J., Bowers, P. F., Kerr, M., & Jackson, P. D. 1986, A&AS, 66, 373
Klypin, A., Kravtsov, A. V., Bullock, J. S., & Primack, J. R. 2001, ApJ, 554, 903
Klypin, A., Kravtsov, A. V., Valenzuela, O., & Prada, F. 1999, ApJ, 522, 82
Knapp, G. R., Stark, A. A., & Wilson, R. W. 1985, AJ, 90, 254
Kochanek, C. S. 1996, ApJ, 457, 228
Kochanek, C. S., et al. 2001, ApJ, 560, 566
Kormendy, J. 1988, ApJ, 325, 128
Kormendy, J., & Bender, R. 1999, ApJ, 522, 772
Kuijken, K. 1999, Ap&SS, 267, 217
Kuijken, K., & Gilmore, G. 1989, MNRAS, 239, 605
———. 1991, ApJ, 367, L9
Lasserre, T., et al. 2000, A&A, 355, L39
Lemson, G., & Kauffmann, G. 1999, MNRAS, 302, 111
Lin, D. N. C., Jones, B. F., & Klemola, A. R. 1995, ApJ, 439, 652
Lindqvist, M., Habing, H. J., & Winnberg, A. 1992, A&A, 259, 118
Loinard, L., Allen, R. J., & Lequeux, J. 1995, A&A, 301, 68
Magorrian, J., & Tremaine, S. 1999, MNRAS, 309, 447
Malhotra, S. 1995, ApJ, 448, 138
Malhotra, S., Spergel, D. N., Rhoads, J. E., & Li, J. 1996, ApJ, 473, 687
Metcalf, R. B., & Zhao, H. S. 2002, ApJ, 567, L5
Mo, H. J., Mao, S., & White, S. D. M. 1998, MNRAS, 295, 319
Moore, B. 1994, Nature, 370, 629
Moore, B., Ghigna, S., Governato, F., Lake, G., Quinn, T., Stadel, J., & Tozzi, P. 1999a, ApJ, 524, L19
Moore, B., Governato, F., Quinn, T., Stadel, J., & Lake, G. 1998, ApJ, 499, L5
Moore, B., Quinn, T., Governato, F., Stadel, J., & Lake, G. 1999b, MNRAS, 310, 1147
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, ApJ, 462, 563
———. 1997, ApJ, 490, 493
Navarro, J. F., & Steinmetz, M. 2000, ApJ, 528, 607
Olling, R. P., & Merrifield, M. R. 1998, MNRAS, 297, 943
———. 2001, MNRAS, 326, 164
Peale, S. J. 1998, ApJ, 509, 177
Pont, F., Queloz, D., Bratschi, P., & Mayor, M. 1997, A&A, 318, 416
Rothberg, B., Saunders, W., Tully, R. B., & Witchalls, P. L. 2000, ApJ, 533, 781
Sellwood, J. A. 2000, ApJ, 540, L1
Sellwood, J. A., & Kosowski, A. 2001, in ASP Conf. Ser. 240, Gas and Galaxy Evolution, ed. J. E. Hibbard, M. P. Rupen, & J. H. van Gorkom (San Francisco: ASP)
Sevenster, M., & Kalnajs, A. 2001, AJ, 122, 885
Schmidt, M. 1975, ApJ, 202, 22
Sheth, R. K., & Tormen, G. 1999, MNRAS, 308, 119
Somerville, R. 2001, ApJL, in press (astro-ph/0107507)
Stanek, K. Z., & Garnavich, P. M. 1998, ApJ, 503, L131
Stark, A. A. 1977, ApJ, 213, 368
Stark, A. A., & Binney, J. 1994, ApJ, 426, L31
Tully, R. B., & Pierce, M. J. 2000, ApJ, 533, 744
Udalski, A., et al. 1994, Acta Astron., 44, 165
Ullio, P., Zhao H. S., & Kamionkowski, M. 2001, Phys. Rev. D, 54, 43504
van den Bosch, F. C., & Swaters, R. A. 2001, MNRAS, 325, 1017
Verheijen, M. 2001, ApJ, 563, 694
Vitvitska, M., Klypin, A., Kravtsov, A. V., Bullock, J. S. Wechsler, R. H., J. R., & Primack, J. R. 2001, ApJ, submitted (astro-ph/0105349)
Walterbos, R. A. M., & Kennicutt, R. C. 1987, A&AS, 69, 311
———. 1988, A&A, 198, 61
Weinberg, M. D. 1985, MNRAS, 213, 451
Weiland, J. L., et al. 1994, ApJ, 425, L81
Wilkinson, M. I., & Evans, N. W. 1999, MNRAS, 310, 645
Zaritsky, D., Olszewski, E. W., Schommer, R. A., Peterson, R. C., & Aaronson, M. 1989, ApJ, 345, 759
Zhao, H. S. 1996a, MNRAS, 278, 488
———. 1996b, MNRAS, 283, 149
Zhao, H., & Evans, N. W. 2000, ApJ, 545, L35
Zhao, H., & Mao, S. 1996, MNRAS, 283, 1197
Zhao, H., Rich, R. M., & Spergel, D. N. 1996, MNRAS, 282, 175
Zhao, H., Spergel, D. N., & Rich, R. M. 1995, ApJ, 440, L13