THE SURFACE DENSITY PROFILE OF THE GALACTIC DISK FROM THE TERMINAL VELOCITY CURVE

STACY S. MCGAUGH

Department of Astronomy, Case Western Reserve University, Cleveland, OH 44106, USA; stacy.mcgaugh@case.edu

Received 2015 September 2; accepted 2015 November 18; published 2015 December 30

ABSTRACT

The mass distribution of the Galactic disk is constructed from the terminal velocity curve and the mass discrepancy–acceleration relation. Mass models numerically quantifying the detailed surface density profiles are tabulated. For \( R_0 = 8 \) kpc, the models have stellar mass \( 5 < M_\odot < 6 \times 10^{10} M_\odot \), scale length \( 2.0 < R_s < 2.9 \) kpc, LSR circular velocity \( 222 < V_0 < 233 \) km s\(^{-1}\), and solar circle stellar surface density \( 34 < \Sigma_d(R_0) < 61 \) \( M_\odot \) pc\(^{-2}\). The present interarm location of the solar neighborhood may have a somewhat lower stellar surface density than average for the solar circle. The Milky Way appears to be a normal spiral galaxy that obeys scaling relations like the Tully–Fisher relation, the size–mass relation, and the disk maximality–surface brightness relation. The stellar disk is maximal, and the spiral arms are massive. The bumps and wiggles in the terminal velocity curve correspond to known spiral features (e.g., the Centaurus arm is a ~50% overdensity). The rotation curve switches between positive and negative over scales of hundreds of parsecs. The rms amplitude \( \langle (dV/dr)^2 \rangle^{1/2} \approx 14 \) km s\(^{-1}\) kpc\(^{-1}\), implying that commonly neglected terms in the Jeans equations may be nonnegligible. The spherically averaged local dark matter density is \( \rho_{\odot,DM} \approx 0.009 M_\odot \) pc\(^{-3}\) (0.34 GeV cm\(^{-3}\)).

Adiabatic compression of the dark matter halo may help reconcile the Milky Way with the \( c-V_{200} \) relation expected in \( \Lambda \)CDM while also helping to mitigate the too-big-to-fail problem, but it remains difficult to reconcile the inner bulge/bar-dominated region with a cuspy halo. We note that NGC 3521 is a near twin to the Milky Way, having a similar luminosity, scale length, and rotation curve.

Key words: Galaxy: fundamental parameters – Galaxy: kinematics and dynamics – Galaxy: structure

Supporting material: machine-readable table

1. INTRODUCTION

The structure of our Galaxy is notoriously difficult to discern given our location within it. The traditional picture of a disk plus bulge has progressed to include both thick and thin disks and a prominent bar component. The thickened portion of the central bar may account for much or perhaps even all of what was traditionally considered the bulge (e.g., Shen et al. 2010). Detailed models of the nonaxisymmetric bar have been constructed (e.g., Martinez-Valpuesta & Gerhard 2015; Portail et al. 2015), there has been enormous progress in mapping the vertical structure of the disk (e.g., Bienaymé et al. 2014; Binney et al. 2014), and the stellar halo is now known to contain considerable substructure (e.g., Helmi 2008).

Despite these advances, we persist in parameterizing the radial surface brightness profile of the primary stellar component of the Galaxy as an exponential disk. This is a crude approximation that ignores variations due to spiral structure, the kinematic effects of which have been detected (Siebert et al. 2012; Williams et al. 2013; Faure et al. 2014). Even with the simple exponential disk approximation, estimates of basic parameters like the scale length of the disk range from \( R_d \approx 2 \) kpc (e.g., Gerhard 2002) to 4 kpc (e.g., Benjamin et al. 2005).

It would be good to move beyond the exponential disk approximation. Here we seek to supplement traditional photometric constraints on the stellar mass distribution with a different technique based on kinematic information. The result is a numerical estimate of the stellar surface density profile \( \Sigma_d(R) \).

A basic result from the mass modeling of external spiral galaxies is that features in the azimuthally averaged light profile have corresponding “bumps and wiggles” in the rotation curve. This can be phrased as Sancisi’s law: “For any feature in the luminosity profile there is a corresponding feature in the rotation curve and vice versa” (Sancisi 2004). This is quantified by the mass discrepancy–acceleration relation (MDAR; McGaugh 2004, 2014), which empirically relates the baryonic mass distribution to the rotation curve. We utilize this correspondence to infer features in the stellar surface density profile from those observed in the terminal velocity curve of the Milky Way.

2. GALACTIC MASS MODELS

The ideal map of the Galaxy would include complete 6D phase-space information for every star. Within such a map one can imagine perceiving not just bulge and disk, or even thick disk and stellar halo, but a distinct stellar population for each and every star-forming event. Chemical tagging (e.g., De Silva et al. 2009; Bland-Hawthorn et al. 2010; Quillen et al. 2015) in the era of large surveys like GALAH (De Silva et al. 2015) and Gaia (Perryman et al. 2001) should help to move us closer to this ideal.

A desirable subset of the ideal map would be a 2D image \( \Sigma(R_\odot,\phi) \) of the Milky Way as seen face-on by an external observer (de Vaucouleurs & Pence 1978; Churchwell et al. 2009). At this juncture it is clear that such an observer would witness a strong bar in the central regions of the Milky Way. Presumably, spiral arms would be perceptible as well. While there has been a great deal of recent work on the Galactic bar, the mass contained in the spiral arms remains uncertain. Such features are certainly known to exist, both from star counts and in the distribution of tracers in the \( l-v \) diagram (Binney & Merrifield 1998).
Here we attempt to take one small step forward by applying what we have learned from external galaxies to the Milky Way. The product is a numerical, nonparametric representation of the azimuthally averaged surface density profile $\Sigma_d(R)$. This is not simply a kinematic estimation of the disk scale length, as it includes the bumps and wiggles presumably induced by spiral arms.

2.1. Assumed Galactic Parameters

Our chief interest here is to find the relative variation in stellar surface density as a function of radius. It is beyond the scope of this paper to address many of the outstanding problems in Galactic structure, so we make some specific assumptions in order to move forward. The absolute values of the surface densities will likely need to be tweaked as more precise values of the Galactic constants are nailed down, but we expect that the relative variations—the bumps and wiggles of interest here—will persist.

Specifically, we assume $R_0 = 8$ kpc and $\Theta_0 = 220$ km s$^{-1}$. These set the scale of the rotation curve to which we fit. We keep $R_0$ fixed but let $\Theta_0$ vary. The inferred variation is within the uncertainty in the solar motion.

Recent work indicates a slightly larger Milky Way (e.g., Chatzopoulos et al. 2015). The relative variation of the bumps and wiggles is not strongly affected, and the absolute normalization of the surface densities is only affected to a small degree. The total mass of the Galaxy varies with $R_0$, but not enough to alter any of the conclusions drawn here. Perhaps the strongest effect of $R_0$ is on the shape of the rotation curve, which rises unnaturally if $R_0$ becomes too large. Consistency with the measured rotation curve shapes of external galaxies prefers $R_0 \leq 8$ kpc (Olling & Merrifield 1998).

2.2. Method

The procedure is that described in Section 5 of McGaugh (2008). We start with a purely exponential stellar disk,

$$\Sigma_d(R) = \Sigma_d(R_0)e^{-(R-R_0)/R_d}. \tag{1}$$

McGaugh (2008) found that the kinematic data preferred short disk scale lengths, so for an initial guess we adopt $R_d = 2$ kpc and a surface density at the solar ring $\Sigma_d(R_0) = 35 M_\odot$ pc$^{-2}$ (Flynn et al. 2006). We compute the rotation curve of the stellar disk $V_d(R)$ using the GIPSY (van der Helst et al. 1992) task ROTMOD, which numerically solves the Poisson equation for the stipulated mass distribution. An exponential vertical profile with scale height $h_v = 300$ pc (Siegel et al. 2002) is assumed.

Next, we compute the corresponding baryonic rotation curve $V_b$ including the bulge ($V_b$) and gas ($V_g$):

$$V_b^2(R) = V_d^2(R) + V_b^2(R) + V_g^2(R). \tag{2}$$

The treatment of the bulge is described in more detail below. The gas distribution is adopted from Olling & Merrifield (2001), including both atomic and molecular gas corrected for helium, as in McGaugh (2008).

The baryonic rotation curve provides an estimate of the full rotation curve $V_c$ by way of the MDAR (McGaugh 2004, 2014):

$$V_c^2 = DV_b^2. \tag{3}$$

The MDAR is an empirical relation between the amplitude of the mass discrepancy $D$ and the force per unit mass $g_b = V_d^2/R$ generated by the baryons. In effect, it quantifies Sancisi’s law. To represent the MDAR, we use

$$D = \left(1 - e^{-\sqrt{g_b}/a_1}\right)^{-1} \tag{4}$$

(see Famaey & Binney 2005; McGaugh 2008; Famaey & McGaugh 2012), with $a_1 = 3700$ km$^2$ s$^{-2}$ kpc$^{-1}$ (Begeman et al. 1991; McGaugh 2004, 2011, 2012, 2014). This is the same functional form adopted by McGaugh (2008).

The rotation curve computed in this way is compared to the observed rotation curve from the terminal velocities observed interior to the solar circle. The input stellar surface density of the disk $\Sigma_d(R)$ is adjusted by changing the surface density in each ring by hand to match the features in the terminal velocity curve (see Section 5 of McGaugh 2008). This procedure is repeated as illustrated in Figure 1 until an adequate fit is obtained.

2.3. Terminal Velocity Data

We apply the method described above separately to the first- and fourth-quadrant terminal velocities. For a given choice of $(R_0, \Theta_0)$, the terminal velocities in the two quadrants are in effect separate realizations of the Galactic rotation curve. Differences between the quadrants may reflect real differences in the gravitational potential stemming from asymmetry in the mass distribution of the Galactic disk. We take the terminal velocity at face value and derive corresponding $\Sigma_d(R)$ independently in each quadrant.

First-quadrant data are adopted from the CO observations of Clemens (1985), as these seem to underpin many published estimates of the Galactic rotation curve. Data for the fourth quadrant are taken from the CO observations of Luna et al. (2006) and the HI observations of McClure-Griffiths & Dickey (2007). The latter are the same data utilized in McGaugh (2008).

The terminal velocity data provide an excellent tracer of the rotation interior to the solar radius through

$$V = v_t + \Theta_0 \sin \ell \tag{5}$$

(Binney & Merrifield 1998), provided that the gaseous tracers are in circular motion at the tangent points. This is a good approximation at larger radii, where the velocity dispersion is much less than the circular speed in both stars and gas. It breaks down as we approach the center of the galaxy and material becomes entrained in the Galactic bar. We fit the data over the range1 $3$ kpc $< R < 8$ kpc $(22^\circ < |\ell| < 90^\circ)$, with the understanding that the inner portion of this range may be affected by the bar. We do not fit the data within 3 kpc on the presumption that it certainly is.

Our procedure requires other judgement calls beyond the decision of where the eccentricities of orbits due to the bar become too great. We cannot hope to fit every tiny bump and wiggle. Nor should we do so, as some may be due to errors or nongravitational (e.g., gas) physics. Examination of the terminal velocity data reveals several qualitatively distinct features. There are broad features extending over several

---

1 The practical upper limit from which useful constraints come is $|\ell| \approx 70^\circ$ ($R \approx 7.5$ kpc).
Imagine that these features are shocks or strong deviations in velocity that appear as sudden spikes in the HI data. These we do not fit to the data. Iterative steps are shown from top to bottom, ending when a satisfactory fit is achieved (bottom; CO data marked in red). The correct y-axis values are shown for the final fit; previous iterations are offset by a factor of two in density (left) and by 50 km s\(^{-1}\) in velocity (right) for clarity.

**Figure 1.** Sequence of steps illustrating the terminal velocity fitting procedure. The disk surface density is shown in the left panel; the resulting velocities are shown in the right panel. We start with a smooth exponential disk model (straight line at top left) that provides what is usually considered an adequate description of the terminal velocity data (fourth-quadrant HI data from McClure-Griffiths &Dickey [2007; gray dots] and CO data from Luna et al. [2006; circles]). We then redistribute mass in rings (McGaugh 2008), recompute the mass model, and compare it to the data. Iterative steps are shown from top to bottom, ending when a satisfactory fit is achieved (bottom; CO data marked in red). The correct y-axis values are shown for the final fit; previous iterations are offset by a factor of two in density (left) and by 50 km s\(^{-1}\) in velocity (right) for clarity.

degrees of Galactic latitude, the most prominent of which is that extending from \(\ell \approx -40^\circ\) to \(-50^\circ\) in the fourth quadrant (Figure 1). These we fit. There are also sudden, sharp deviations in velocity that appear as sudden spikes in the HI data of McClure-Griffiths & Dickey (2007, e.g., at \(\ell = -48^\circ\) and \(-54^\circ\)). These we do not fit. Indeed, no plausible mass distribution can explain such sudden changes in velocity. We imagine that these features are shocks or strong flows of gas where our line of sight crosses a spiral arm fragment. Finally, there are intermediate features, small and sometimes sharp but distinct from the sudden spikes. These we fit as we can, without obsessing over differences of a few kilometers per second that are smaller than turbulence in the gas (Kalberla &Kerp 2009).

The formal uncertainties in the terminal velocity measurements are typically only a few kilometers per second. These are small compared to systematic uncertainties, particularly the degree to which the assumption of circular motion holds. Since these are not quantified, we make no attempt at a formal fit that minimizes \(\chi^2\). Rather, we consider a fit to have converged if the rotation curve passes through the bulk of the data and captures the observed pattern of bumps and wiggles. Though we make no claim to have obtained a formal best fit, we are not aware of any other results that fit the terminal velocity data in such detail.\(^2\)

2 The nearest comparable work is that of Sofue et al. (2009), who fit a possible dip in the rotation curve outside the solar circle with a ring of mass farther out.

**2.4. Disk Thickness**

Considerable effort has been made to understand the vertical structure of the disk (e.g., Bienaymé et al. 2014; Binney et al. 2014; Piffl et al. 2014a). Separate thin and thick disk components can be perceived, and these may have different radial scale lengths (Juric et al. 2008). The exact vertical structure of the disk need not be purely exponential any more than its radial structure, and this is essential to the determination of the vertical restoring force to the disk.

Here we are interested in the radial rather than the vertical force. The vertical structure plays a relatively minor role in the computation of \(V_2^2/R\). This is illustrated by Figure 2, which shows the terminal velocities and corresponding rotation curve for three models of differing scale height. The models are otherwise identical, sharing the same radial mass distribution \(\Sigma_d(R)\). In addition to the nominal assumed scale height of 300 pc, a razor-thin disk and a thicker disk with \(h_z = 600\) pc are shown.

As expected (Binney & Tremaine 1987), the thinner disk rotates slightly faster and the thicker one more slowly, all other things being equal. In addition, thinner disks respond more dramatically to variations in the surface density profile, also as expected. However, the absolute difference between plausible models is not great. We therefore fix the disk thickness to 300 pc (Siegel et al. 2002) and do not distinguish between thick and thin disks. For this particular problem, this distinction is small, with differences that are smaller than those caused by turbulent motion in the gas.

**2.5. The Bulge-bar**

We are interested here in the structure of the Galactic disk, and in particular the detailed radial variation of its stellar surface density. We make no attempt to infer this outside the solar radius, where the tangent point method cannot be applied, nor inside a radius of 3 kpc \(|\ell| < 22^\circ\), where noncircular motions are important. For the present purpose, the nature of the central
component of the galaxy—whether it is a bulge or a bar or some combination thereof—is not terribly important. However, it is necessary to account for the integrated interior mass. To this end, we approximate the central “bulge” component with a numerical model based on the COBE light distribution (Binney et al. 1997), as described by McGaugh (2008).

Here we vary the normalization of the bulge component to check its effect on the inferred disk surface densities. As one might expect, the bulge plays only a minor role, and only at small radii. We build models with different bulge fractions to explicitly quantify its effect.

3. MASS MODELS

We construct mass models with three components: a stellar disk, a central bulge, and a gas disk. The gas disk is based on the work of Olling & Merrifield (2001) and is identical to that used in McGaugh (2008). The bulge model is also that used in McGaugh (2008), but here we vary the bulge fraction, adopting three cases: zero bulge, a nominal bulge fraction close to 20% of the total light, and a heavy bulge that could be taken to represent a bulge light fraction of \( \sim 1/3 \), or equivalently, a bulge with a smaller light fraction but with a mass-to-light ratio heavier than that of the disk. These cases presumably bracket reality.

A model is built for the stellar disk for each of the three choices of bulge fraction. This is done separately in the first and fourth quadrants, treating the terminal velocity curve from each as an independent estimate of the rotation curve. This produces a total of six models. The fourth-quadrant model with zero bulge is basically identical to that in Table 3 of McGaugh (2008), with the exception that the subtle outward force of the gas component at small radii, ignored before, is treated rigorously here.

Fits to the terminal velocities are shown in Figure 3. The models are very similar over the range fit. The effect of the different bulge fractions is apparent only at small radii (\( |\ell| < 22 \)).

The bulk properties of the models are given in Table 1. All models assume a galactocentric distance \( R_0 = 8 \) kpc. Scaling to other values of \( R_0 \) is not straightforward, but the basic pattern of bumps and wiggles would persist.

The first column of Table 1 provides a label for each model composed of the quadrant for which the terminal velocity data have been fit and the bulge size (Zero, Moderate, or Big). Q1

Figure 2. Effect of disk thickness on the terminal velocities (left) and corresponding rotation curve (right). Fits (central line) have been made assuming an exponential vertical profile with \( h_z = 300 \) pc (Siegel et al. 2002). The other lines show models with identical radial surface density profiles but different vertical scale heights. A razor-thin disk is shown (line having more negative terminal velocities and higher rotation velocities) together with a thick disk with \( h_z = 600 \) pc (line with less negative terminal velocities and lower rotation velocities). All other things being equal, thinner disks rotate faster and respond more strongly to bumps and wiggles in the surface density profile. A disk of any plausible thickness could be made to fit the data by changing the mean surface density (which needs to be slightly higher for a thicker disk) and modulating the relative amplitude of the bumps and wiggles. Though details may change for different assumptions, the pattern of bumps and wiggles would remain the same.

Figure 3. Fits to the terminal velocity data for models with varying bulge fraction in the first (positive velocities) and fourth (negative velocities) quadrants. The CO data of Clemens (1985) are shown in the first quadrant. In the fourth quadrant, the CO data of Luna et al. (2006) are shown as large circles and the HI data of McClure-Griffiths & Dickey (2007) as small dots. For reference, the smooth model of McClure-Griffiths & Dickey (2007) is shown as the dashed curve in the fourth quadrant, as is a simple linear fit to the first-quadrant data (\( v_t = 167.3 - 2.325 \ell \)). The solid lines show our detailed fits to the data. The models differ significantly only at small radii (\( R < 3 \) kpc; \( |\ell| < 22^\circ \)). No attempt is made to fit the data in this inner region where noncircular motions become important. Lines representing larger bulge fractions can be distinguished by their higher \( |v_t| \).
The distance to the Galactic center is assumed to be $R_0 = 8.0$ kpc. The total gas mass in all models is $M_g = 11.8 \times 10^9 M_\odot$. This includes both atomic and molecular gas and has been corrected to include helium and metals. The mass-to-light ratios assume that the total luminosity of the Milky Way is $L_V = 37.3$, $L_I = 42.9$, and $L_K = 86.5 \times 10^9 L_\odot$ (Drimmel & Spergel 2001; Flynn et al. 2006; Just et al. 2015).

The point of these models is to move beyond the usual approximation of an exponential disk. Nevertheless, it is useful to fit an exponential disk to the inferred surface densities as a reference. Equation (1) is fit over the range $3 \text{ kpc} < R < 8 \text{ kpc}$ where the terminal velocities have been fit. The fifth and sixth columns give the surface density at the solar radius and the scale length that result from this fit.

It is interesting to see how the parameters of the fitted exponential disk vary. In the first quadrant, the bumps and wiggles average out and return a fitted exponential indistinguishable from the smooth initial guess. In the fourth quadrant, a higher surface density is inferred at larger radii, leading to fits with longer scale lengths and higher $\Sigma_d(R_0)$. It is well known that such fits depend on the range over which the fit is made. That we obtain somewhat different results from the first and fourth quadrants may go some way to explaining the range of results found in the literature, which may themselves be fit over different radial and azimuthal ranges.

Another item to note is that the fitted surface density of the solar circle need not be identical to that of the solar neighborhood. That is to say, our local patch extending over a small range of azimuths need not be identical to the ring centered at $R_0$ averaged over all azimuths. As it happens, the exponential disk fits to the bumps and wiggles in the first quadrant return a stellar surface density at the solar radius very near to that measured locally (e.g., Flynn et al. 2006). In contrast, the fits to the fourth-quadrant data indicate a rather higher surface density. This difference may just be fluke, but it may also indicate a real variation from one side of the Galaxy to the other. Taken literally, it appears that the surface density averaged around all azimuths may exceed that of the solar neighborhood. That is, the Sun might reside in a patch that is a bit underdense for its radius, consistent with its current interarm location.

Figure 4 shows the surface density profiles of the models over the radial range to which exponential fits are made. Models with different bulge fractions look similar, while those from different quadrants are notably different. Also shown for reference is the stellar surface density profile found by Bovy & Rix (2013) [$\Sigma_{d,R}(R_0) = 38 M_\odot \text{ pc}^{-2}$, $R_d = 2.15$ kpc] from their analysis of the vertical force. This is in good agreement with that found here from the radial force.

The seventh column of Table 1 reports the radius at which the rotation curve of the baryonic mass model peaks. This would be 2.2 scale lengths for a purely exponential disk, but can be different for general mass distributions. For these models, the radius where $V_b$ peaks is a bit larger than $2.2 R_d$. The eighth column gives the total circular velocity at this radius, which is useful for making a fair comparison to external galaxies. The ninth column gives the rotation attributable to the baryons (stars and gas) at $R_p$. All models work out to be essentially maximal with $V_b/V_p \approx 0.85$. This reflects the fact that within the solar radius, the Galaxy resides in a portion of the MDAR where the mass discrepancy is modest.

Column 10 gives the rotation velocity that an external observer would measure at large radii. This is taken to be the velocity at the edge of the H I distribution at 20 kpc, where the H I surface density drops to the typical sensitivity limit of $1 M_\odot \text{ pc}^{-2}$. This outer velocity is useful for comparing the Milky Way to other galaxies in the Tully–Fisher relation. Note that this $V_f$ is not equal to the circular velocity of the LSR, which is reported for each model in column 11. All models have slowly declining rotation curves at large radii. One consequence of this is that we cannot measure $\Theta_0$, assume that the rotation curve is flat, and expect this to be an adequate measure for comparison to external galaxies.

Columns 12 and 13 of Table 1 give the Oort constants $A$ and $B$. These are determined from the local gradient in the...
rotation curve just inside and outside of the solar radius. The models are not particularly well constrained at the solar radius, as the tangent point method is only effective interior to the solar circle. The radial run of the Oort constants is discussed in Section 4.2.

The final three columns of Table 1 give the stellar mass-to-light ratio of each model Milky Way in the V, I, and K bands. Integration of the models fit to the terminal velocities provides the stellar mass. For the total luminosity of the Milky Way we adopt the K-band disk luminosity of $6.92 \times 10^{10} L_\odot$ of Drimmel & Spergel (2001) and follow their lead in correcting this upward to include a 20% bulge fraction, resulting in a total luminosity of $L_K = 8.65 \times 10^{10} L_\odot$. The other luminosities assume $V - I = 0.90$ and $V - K = 2.46$, which are local colors from Flynn et al. (2006) and Just et al. (2015). These yield $L_V = 3.73 \times 10^{10} L_\odot$ and $L_I = 4.29 \times 10^{10} L_\odot$.

Note that the total K-band luminosity adopted here, after correction for the bulge, is very similar to the disk-only luminosity found by extrapolation of the local surface brightness by Just et al. (2015) before inclusion of the bulge. They adopt a longer scale length than found here, which may account for part of this discrepancy. However, it is not obvious that the results of Drimmel & Spergel (2001) and Just et al. (2015) can entirely be reconciled. Indeed, Just et al. (2015) note that the single star Arcturus makes a substantial contribution to the local surface brightness measurement, so we remain cautious about how accurately these quantities are known.

The K-band mass-to-light ratio of all models is very nearly $T^K_\odot = 0.6 M_\odot/L_\odot$, consistent with the expectations of population synthesis models (McGaugh & Schombert 2014). Indeed, this value was adopted in the calibration of the MDAR (McGaugh 2014). However, obtaining the same mass-to-light ratio for the Milky Way is not guaranteed, as the luminosity estimate is independent of the stellar mass estimate. That the models return a value consistent with the MDAR calibrated by external galaxies provides some hope that the luminosity estimates are not too far off. The mass-to-light ratios in V and I follow from the adopted colors and are also reasonable from the perspective of stellar populations (compare to Table 7 of McGaugh & Schombert 2014). Licquia et al. (2015) quote a very similar $I$-band mass-to-light ratio to what we find here. They find a redder global color than assumed here, leading to a correspondingly higher V-band mass-to-light ratio.

In our models, the disk mass declines as the bulge fraction increases. This trade-off is necessary to keep the total mass in the right ballpark. Indeed, to accommodate an increasing bulge fraction, it is necessary to reduce the mass of the disk in the inner regions. This is usually accomplished by letting the scale length of the disk grow (e.g., Flynn et al. 2006). However, it is no longer possible to fit the bumps and wiggles if we stretch out the disk too much. To address this issue, we limit the disk mass reduction to the region of the bulge/bar by adopting a Freeman (1970) Type II profile\(^3\) with a constant surface density region interior to a radius that depends on the bulge fraction. The zero models increase exponentially all the way to the center. For $B/T > 0$, $\Sigma_f = \Sigma_c = \text{constant for } R < R_c$. These values are tabulated in the detailed mass models given in Table 2.

Type II profiles are a common morphology for the surface brightness profiles of barred spiral galaxies. One might imagine that the constant-density region of the disk represents the azimuthally averaged bar, while the bulge is a thicker component. However, no attempt has been made to match the details of either kinematics or photometry in the inner region where this trade-off is made. Indeed, the choice of inner profile is rather arbitrary. The parameters $\Sigma_c$ and $R_c$ are highly degenerate, and similar results could be obtained with different combinations. Given this and the uncertainty in the bulge itself, we simply choose an $R_c - \Sigma_c$ pair that accommodates the target bulge fraction. To fit the remainder of the rotation curve, we fit the pattern of bumps and wiggles, allowing the exact bulge mass to vary slightly in order to also fit the amplitude of the rotation curve. For this reason, the final bulge fraction need not be exactly the target fraction.

Table 2 gives the complete mass model for each case. The first column specifies the model by quadrant and bulge fraction, as in Table 1. The second column is the radius in kpc. The third and fourth columns are the surface density of the stellar disk and the circular speed of the gravitational potential it generates. Similarly, the fifth and sixth columns are the surface density and rotation curve of the bulge component, and the seventh and eighth columns those of the gas disk. The ninth column gives the total rotation curve of the model determined from Equation (3). The model is extrapolated with an exponential disk well beyond the limits of the data.

The mass models tabulated in Table 2 are shown in detail in Figure 5 (first quadrant) and Figure 6 (fourth quadrant). Bumps in the stellar surface density cause corresponding wiggles in the rotation curve. The pattern of bumps and wiggles is similar regardless of bulge fraction, but differs between the two quadrants. Close comparison of the rotation curves in the two quadrants reveals that $V(R)$ is a bit higher at middle radii in the first quadrant relative to the fourth quadrant, with the reverse being true at the edges of the data. Blinking between the two resembles the flapping of a bird’s wings. Since the quadrant midpoints are offset by 90°, this is consistent with the effects of an $m = 2$ mode perturbation, i.e., spiral arms.

4. DISCUSSION

The results discussed above amplify the initial work of McGaugh (2008). The pattern of features in the terminal velocity curve can be fit by a corresponding pattern of features in the surface density profile of the stellar disk. Here we discuss some of the implications of these models.

4.1. Spiral Structure

We have applied the MDAR to infer the azimuthally averaged stellar surface density implied by the pattern of bumps and wiggles observed in the terminal velocity curve. If these kinematically inferred features are real, then they should correspond to physical structures. These are presumably spiral arms.

There is a rich history to the study of spiral structure in the Milky Way (Binney & Merrifield 1998). Spiral arms are known features. We can thus check whether the structures we infer correspond to known spiral arms.

Figure 7 shows the positions of known giant molecular clouds (GMCs) and H II regions from the compilation of Hou

---

3 More generally, allowing a break in the exponential disk profile may help alleviate the tension between the mass of the bulge-bar and the scale length of the outer disk.
et al. (2009). These objects are good tracers of spiral arms. We also show lines of sight to known spiral arms: the Carina, Centaurus, and Norma arms in the fourth quadrant, and in the first quadrant the Sagittarius arm and the molecular ring in Aquilla. The latter feature may simply be a tightly wound spiral arm (Dobbs & Burkert 2012).

There is a good correspondence between known spiral arms and features in the terminal velocity curve. The most prominent is the Centaurus arm. This manifests as the wide dip in the fourth-quadrant rotation curve around $R \approx 5$ kpc (Figure 6). This can be seen in the surface density profile as a broad ($\sim 1$ kpc wide) overdensity extending a little beyond 6 kpc.

The Centaurus arm must represent a prominent overdensity of stellar mass to have the observed effect. Bear in mind that the mass models are axially symmetric, but fit to a single quadrant’s data. So, on the one hand, a bump must be large to have any effect on the azimuthally averaged surface density profile. On the other hand, the terminal velocities only probe

### Table 2
Detailed Milky Way Mass Models

| Model | $R$ (kpc) | $\Sigma_d$ ($M_\odot$ pc$^{-2}$) | $V_d$ (km s$^{-1}$) | $\Sigma_b$ ($M_\odot$ pc$^{-2}$) | $V_b$ (km s$^{-1}$) | $\Sigma_g$ ($M_\odot$ pc$^{-2}$) | $V_g$ (km s$^{-1}$) | $V_c$ (km s$^{-1}$) |
|-------|-----------|-------------------------------|-------------------|-------------------------------|-------------------|-------------------------------|-------------------|------------------|
| Q1ZB  | 0.1       | 2029                          | 15.1              | 0                             | 0                 | 0                             | $-0.7$            | 20.5             |
|       | 0.2       | 1930                          | 29.3              | ...                           | ...               | ...                           | $-1.5$            | 36.1             |
|       | 0.3       | 1836                          | 42.3              | ...                           | ...               | ...                           | $-2.2$            | 49.9             |
|       | 0.4       | 1746                          | 54.4              | ...                           | ...               | ...                           | $-3.0$            | 62.5             |
|       | 0.5       | 1661                          | 65.5              | ...                           | ...               | ...                           | $-3.8$            | 74.0             |

(This table is available in its entirety in machine-readable form.)

---

**Figure 5.** Mass profiles (left) and rotation curves (right) for first-quadrant models designated by bulge fraction. Symbols as per Figure 3. The stellar disk is represented by dashed lines, the bulge-bar by dotted lines, and the gas disk by dot-dashed lines. Their sum is given by the solid line. The bumps and wiggles in the mass profiles at left can cause the corresponding features in the rotation curve (upper solid lines in right panels), as fit to the terminal velocity data of Clemens (1985, points).
the vicinity around the tangent point at each radius, so features there may have an exaggerated effect. Nevertheless, it is striking that spiral structure does appear to have the expected effect on the observed rotation curve.

Taking the data at face value, the Centaurus arm represents a 40% enhancement over the smooth exponential fit of the corresponding models in Table 1. These fits include the overdensity. Simply taking the “background” disk density as the line between the edges of the Centaurus spiral feature raises the inferred enhancement to 60%. This is an overdensity of mass associated with the Centaurus arm, not just light.

In the first quadrant, both the Sagittarius arm and the molecular ring/Aquila arm leave distinctive features in the terminal velocities and corresponding surface densities. Neither are as broad or massive as the Centaurus arm, but both have large density contrasts that cause abrupt changes in the rotation curve. These bumps and wiggles appear to be real features owing to the expected effects of spiral structure.

The correspondence between features in the mass distribution inferred from the terminal velocity curves and known spiral arms gives some confidence that the method employed here is on the right track. Just as the Milky Way has a central bar, so too it has spiral arms. These have the expected effect on the velocity field and the surface density profile of the stellar disk. An obvious next step would be the construction of nonaxisymmetric models $\Sigma(R, \phi)$. These entail further degeneracies that are beyond the scope of this work.

4.2. The Derivative of the Rotation Curve and the Oort Parameters

There are clear bumps and wiggles in the terminal velocity data. These appear to correspond to real variations in the surface density of stars caused by spiral arms. An interesting consequence is that there are nonnegligible variations in the derivative of the rotation curve over scales of hundreds of parsecs.

To a first approximation, the rotation curve is roughly flat over many kpc. Extrapolation of the models to larger radii predicts a slowly declining rotation curve. However, on scales of hundreds of parsecs, $V(R)$ can switch between rising and falling rather suddenly. These local changes in $dV/dR$ can have
important effects on the determination of quantities that depend on it (Figure 8).

Examples of quantities that depend on the derivative of the rotation curve include the Oort constants and the vertical restoring force of the disk. The definition of the Oort constants depends explicitly on \( \frac{dV}{dR} \), so their inferred value will differ between surveys that cover different radii and azimuths if these happen to encounter different bumps and wiggles. This...
The derivative of the rotation curve swings between positive and negative several times over the range depicted in Figure 8. The frequency with which this happens depends somewhat on how we choose to fit the observed bumps and wiggles in the terminal velocities. There is more variation than we have chosen to fit, though as discussed in Section 2.3, the smaller-scale fluctuations are less likely to be dynamical structures. In the current models, the sign of \( dV/dR \) changes over scales of hundreds of parsecs, with the actual zero crossing being even more sudden.

To quantify the amount of variation in the derivative, we compute its rms over the range depicted in Figure 8. In the first-quadrant model, \( \langle (dV/dR)^2 \rangle^{1/2} = 22 \text{ km s}^{-1} \text{kpc}^{-1} \). In the fourth quadrant, \( \langle (dV/dR)^2 \rangle^{1/2} = 14 \text{ km s}^{-1} \text{kpc}^{-1} \). For comparison, \( V/R \approx 30 \text{ km s}^{-1} \text{kpc}^{-1} \) at \( R = 7.5 \text{ kpc} \), so the variation in the derivative is not much smaller than the orbital frequency just interior to the solar neighborhood.

If the fine-grained rotation curve is not smooth, analyses that simplify the Jeans equations by assuming \( dV/dR = 0 \) may be incorrect, or at least run the risk of introducing systematic errors. For example, a nonzero rotation curve gradient on hundreds of parsec scales might help to explain mild inconsistencies in the vertical force estimated over several kiloparsecs (Bovy & Rix 2013). If this effect is important, one might expect it to manifest as apparent discrepancies within different subsets of the Gaia data. That is, as these data become available, analyses that assume \( dV/dR = 0 \) may give different results when applied to subsets of the data representing distinct regions with different local gradients. Imposing a uniform assumption about the gradient may lead to perplexing results.

4.3. The Milky Way in the Context of External Spirals

An obvious question is how the Milky Way compares to other spiral galaxies. From the perspective of the Copernican principle, one would expect it to be a normal spiral galaxy. Occasionally, one finds cause to think it peculiar, which must also be true at some level of detail since all objects are individuals. Here we use the models constructed above to place the Milky Way in context.

Figure 9 shows the baryonic Tully–Fisher relation with the Milky Way highlighted. All six Milky Way models from Table 1 are plotted together with data for other galaxies from McGaugh (2005a) and McGaugh & Schombert (2015). The sum of the baryonic mass components is plotted against two measures of the rotation velocity: that measured at the peak of the baryonic contribution to the rotation curve \( V_p = V(R_p) \) being the generalized version of \( V_{2,2} \), and that in the outer, more nearly flat portion of the rotation curve \( V_f \).

The Milky Way, as modeled here, falls within the scatter of the Tully–Fisher relation for either measure of the rotation speed. Indeed, all six models lie comfortably within the scatter and are hardly distinguishable in the Tully–Fisher plane. There is perhaps a hint that the Milky Way lies on the lower right side of the very small scatter, but this is well within the uncertainties. Similarly, adopting a different value of \( R_0 \) will vary the rotation velocity and mass, but not beyond the scatter for plausible values of \( R_0 \). By this standard, the Milky Way is a normal spiral.

Figure 10 shows the disk size–mass and disk thickness–mass relation for spiral galaxies. In both cases, the rotation speed is used as a proxy for mass. This is a good proxy (see Figure 9), but slightly different measures of the velocity are available: \( V_p \) in the left panel, and the maximum observed velocity \( V_{\text{max}} \) in the right panel. There is a strong correlation between \( V_p \) and \( V_{\text{max}} \), so these suffice here. The exponential scale length is used...
to characterize the sizes of disk galaxies in the left panel of Figure 10, which uses the same data as in Figure 9. For disk thickness in the right panel, data for edge-on galaxies are taken from Kregel & van der Kruit (2005) and Kregel et al. (2005).

As with the baryonic Tully–Fisher relation, the Milky Way appears to be a normal spiral galaxy. It is perhaps a bit compact for its rotation speed, but this is within the ample scatter in the size–mass relation. Segregation between the models of Table 1 is more apparent here. This is also true in terms of disk thickness, where $h = 300 \text{ pc}$ is assumed for models of all radial scale lengths. Nonetheless, the Milky Way sits comfortably within the scatter shown by spiral galaxies generally.

The global properties inferred for the Milky Way by the modeling of the terminal velocities made here suggest that we live in a fairly ordinary spiral galaxy.

### 4.4. NGC 3521: A Milky Way Twin

In comparing the Milky Way to other galaxies, we noticed not only that it follows the same correlations as other spirals, but that it consistently falls in the same spot as one other galaxy in the comparison sample. This object is NGC 3521.

NGC 3521 has a baryonic mass, rotation velocity, and scale length that are very similar to those of the Milky Way. Indeed, these quantities are identical within the uncertainties. While it is common for galaxies to fall in the same spot along the Tully–Fisher relation, such Tully–Fisher pairs of galaxies often have different disk scale lengths (de Blok & McGaugh 1996; Tully & Verheijen 1997; Courteau & Rix 1999; McGaugh 2005b). The intrinsic scatter in the Tully–Fisher relation is small, while that in the size–mass relation is large (see Figures 9 and 10).

In the case of NGC 3521, the scale length is close to that of the Milky Way, as well as its position on the Tully–Fisher relation. Indeed, the similarity persists in even greater detail. Figure 11 shows the rotation curve of NGC 3521 from the THINGS survey (de Blok et al. 2008; Walter et al. 2008). Also plotted are the zero bulge models of the Milky Way from both the first and fourth quadrants (Table 2). The rotation curves of the two galaxies are practically indistinguishable.

There is a general lesson here. Galaxies that share a similar baryonic mass distribution share a similar rotation curve (Persic & Salucci 1991; McGaugh 2014). Nature builds galaxies from a very strict recipe (Disney et al. 2008).

Galaxies with structural similarities to the Milky Way have been noted before: de Vaucouleurs & Pence (1978) highlighted NGC 1073, NGC 4303, NGC 5921, and NGC 6744. These galaxies have similar global properties and morphological classifications. Apparently, NGC 3521 can be added to this list. De Vaucouleurs & Pence (1978) classified the Milky Way as SAB(rs)bc, NGC 3521 is classified as SABbc (de Vaucouleurs et al. 1991). Apparently, both galaxies contain bars (see also Zeilinger et al. 2001), though the similarity of the inner regions is less pronounced.

### 4.5. Maximal Disks

The disk inferred here for the Milky Way is maximal. The baryons dominate the gravitational potential interior to the Sun and contribute the bulk of the rotational support at $R_p$.

One working definition of a maximal disk is one in which the ratio of the rotation due to the baryonic component to the total rotation at $R_p$ is $V_b/V_p \approx 85\%$ (Sackett 1997). The Milky Way models constructed here all fall in the range $0.85 < V_b/V_p < 0.87$. It therefore appears that the Milky Way has a maximal disk. This is consistent with the findings of Sackett (1997) and Bovy & Rix (2013), and with the expectation for a galaxy of the inferred surface density (McGaugh 2005b, Figure 12).

Indeed, it is necessary for the disk to dominate the gravitational potential in order to generate the bumps and wiggles in the terminal velocity curve. Such features are the natural consequence of disk self-gravity (e.g., Sellwood 2014). They cannot be supported by the dynamically hot, quasi-spherical dark matter halo (Binney & Tremaine 1987). That the procedure applied here works suggests that the disk is massive.

The importance of disk self-gravity means that spiral arms are overdensities in mass, not just light. If spiral arms were not enhancements in mass, then there would be no reason to expect an association between them and features in the terminal velocity curve. Since the expected imprint of massive spiral arms on the velocity field is observed, it follows that they are indeed massive.

It also follows that the disk in which spiral arms are embedded is also massive. If we seek to reduce the disk mass, we are obliged to increase the contrast of the overdensity that the arms represent. This is not small to start, and it becomes greater as the surface density of the disk declines. In effect, the

Figure 10. Scale lengths (left) and thicknesses (right) of disks as a function of rotation speed. The data in the left panel are the same as those in Figure 9. The data in the right panel are from Kregel & van der Kruit (2005) and Kregel et al. (2005) and show the ratio of exponential disk scale length to scale height. Differences between the models are more apparent here than in Figure 9. The Milky Way does not stand out from either distribution, though it is somewhat compact for its rotation velocity.
spiral arms must become more massive as the disk becomes less massive. This places an effective lower limit on the maximality of the disk that depends only on how unreasonable a contrast we are willing to tolerate. The estimate of the contrast here is already relatively large (see Drimmel & Spergel 2001).

4.6. The Implied Dark Matter Distribution

The implied distribution of dark matter follows trivially once the baryon distribution is specified. The portion of the velocity attributable to the dark matter halo can be found from the data in Table 2 through

$$V_{\text{DM}}^2 = V_e^2 - V_b^2. \quad (6)$$

More generally, this can be expressed analytically as

$$V_{\text{DM}}^2 = V_b^2 (D - 1). \quad (7)$$

(McGaugh 2004, 2014). The amplitude of the mass discrepancy increases outward as the centripetal acceleration decreases. In the solar neighborhood the mass discrepancy $|\Theta_b/V_b(R_0)|^2$ is $D_0 \approx 1.5$. At $R_p$, $D_p = [V_p/V_b(R_p)]^2 \approx 1.35$. At smaller radii, $D \to 1$ so that it becomes difficult to perceive the role of dark matter in the inner few kiloparsecs (Portail et al. 2015).

Once the functional form of the MDAR is specified, the rotation curve due to the dark matter halo follows. This is only as uncertain as the empirical calibration of the MDAR and the mass-to-light ratio of a given galaxy. The density of dark matter can be inferred from $V_{\text{DM}}$ through solution of the Poisson equation. Assuming a spherical halo,

$$4\pi G \rho_{\text{DM}} = 2 \left( \frac{V_{\text{DM}}}{R} \right) \left( \frac{\partial V_{\text{DM}}}{\partial R} \right) + \left( \frac{V_{\text{DM}}}{R} \right)^2. \quad (8)$$

Applying this to the solar neighborhood, we infer a spherically averaged dark matter density of

$$\rho_{0,\text{DM}} \approx 0.009 \ M_\odot \ pc^{-3} = 0.34 \ GeV \ cm^{-3}. \quad (9)$$

This quantity is of obvious interest to laboratory searches for dark matter and is very similar to other estimates (e.g., Holmberg & Flynn 2000; Salucci et al. 2010; McMillan 2011; Bovy & Tremaine 2012; Strigari 2013; Piffl et al. 2014b, 2014a, 2015; Read 2014). Indeed, it is so consistent with previous results that it did not warrant mention in McGaugh (2008), though the same result can be derived from the information provided there. It is equally trivial to derive the

![Figure 11](image1.png)

**Figure 11.** Rotation curve of NGC 3521 (de Blok et al. 2008) compared to the $B/T = 0$ Milky Way models Q1ZB (red line) and Q4ZB (blue line). The two galaxies are near twins, having similar masses, disk scale lengths, and rotation curves.

![Figure 12](image2.png)

**Figure 12.** Maximality of disks as a function of rotation speed (left) and characteristic baryonic surface density (defined as $\Sigma_b = (3M_b)/(4R_p^2)$ by McGaugh 2005b, right). The comparison data are the same as those in Figure 9. The Milky Way has a maximal disk, as expected for a high surface brightness spiral (see also Figure 23 of Bovy & Rix 2013).
dark matter density at any other point in the Galaxy by combining Equations (7) and (8).

The dark matter profile dictated by Equation (7) does not, in general, follow any of the traditional analytic prescriptions for dark matter halos. We can nevertheless fit such halo models, which give a tolerable description of the data over the modest range of radii probed. For example, a pseudo-isothermal halo characterized as

$$V_{ISO}(R) = V_\infty \left[ 1 - \left( \frac{R_C}{R} \right) \arctan \left( \frac{R}{R_C} \right) \right]^{1/2}$$

fits the data with a core radius $R_C = 3$ kpc and an asymptotic velocity $V_\infty = 177$ km s$^{-1}$ (Table 3).

Similarly, the Navarro–Frenk–White (NFW) halo (Navarro et al. 1997) is characterized by a concentration $c$ and a characteristic velocity $V_{200}$. This is the orbital velocity of a test particle on a circular orbit at the quasi-virial radius $R_{200}$. This quasi-virial radius contains a mass density 200 times the critical density of the universe and is typically far beyond the reach of observation. We can nevertheless use this notionally radial variable to define a radial variable $x = R/R_{200}$ and the concentration $c = R_{200}/R_s$, where $R_s$ is the scale radius where the density profile rolls over (Navarro et al. 1997). The rotation curve of an NFW halo is

$$V_{NFW}(R) = V_{200} \left[ \frac{\ln(1 + cx) - cx/(1 + cx)}{x \ln(1 + c) - c/(1 + c)} \right]^{1/2}. \hspace{1cm} (11)$$

Fitting this to the dark matter distribution indicated by the MDAR over the radial range $3 \text{ kpc} < R < 8 \text{ kpc}$ gives $c = 5.2$ and $V_{200} = 264$ km s$^{-1}$ (Table 3).

The NFW halo fit directly to the dark matter distribution given by the MDAR implies a rather large mass for the Milky Way of $M_{200} \approx 6 \times 10^{12} M_\odot$. However, our fit over the radial range $3 \text{ kpc} < R < 8 \text{ kpc}$ has little power to constrain the total mass of the halo. Indeed, NFW halos are highly self-degenerate, so that a correlated series of $c$–$V_{200}$ values yield nearly indistinguishable results over finite ranges of radii (hence the banana-shaped contours in Figure 4 of de Blok et al. 2001). Consequently, an adequate description of the data is also provided by an NFW halo with $c = 7.5$ and $V_{200} = 180$ km s$^{-1}$. Though not the formal best fit, this is well within the uncertainties. The total mass of this halo is $M_{200} = 1.9 \times 10^{12} M_\odot$, more in line with observational determinations from tracers at large radii (Sakamoto et al. 2003).

As a consequence of the degeneracy between $c$ and $V_{200}$, it is possible to fit still lower mass halos ($\lesssim 10^{12} M_\odot$). A frequently noted advantage (e.g., Gibbons et al. 2014) of a low-mass Milky Way halo is that it eases the so-called too-big-to-fail problem (Boylan-Kolchin et al. 2011). This comes at the price of excessively high concentrations ($c \approx 20$). These are not consistent with the predictions of ΛCDM, which provides a well-defined mass–concentration relation (Macciò et al. 2008). Simply lowering the mass of the Milky Way halo does not provide a satisfactory solution as the amount of substructure depends on density, not just mass (McGaugh et al. 2003; Zentner & Bullock 2003): a high concentration is just as bad as a high mass in this context.

The mass–concentration relation is predicted by dark-matter-only simulations (Navarro et al. 1997). One inevitable effect of forming a luminous galaxy within a dark matter halo is adiabatic compression. This has the effect of raising the effective concentration of the resulting halo above that of the primordial initial condition to which the mass–concentration relation applies.

We use the procedure of Sellwood & McGaugh (2005) to make an approximate fit to model 4QZB. Both the primordial and compressed halo are illustrated in Figure 13. The compressed halo is the one subject to observational constraint and no longer has a purely NFW form. We can nevertheless fit it as if it were, with the resulting parameters given in Table 3. The result is a rather low mass halo ($M_{200} = 4 \times 10^{11} M_\odot$) with a concentration ($c = 14$) that is too high for ΛCDM (Macciò et al. 2008). However, this is not the right result to compare to the prediction of simulations, which provide the mass–concentration relation prior to compression. The precompressed, primordial halo is found to have a mass $M_{200} = 6.1 \times 10^{11} M_\odot$ with a concentration $c = 7.1$. This is nicely consistent with the predicted mass–concentration relation.

This result appears quite favorable. Not only is the Milky Way consistent with the mass–concentration relation once we have accounted for adiabatic compression, but the mass is low enough with a reasonable concentration to help with the too-big-to-fail problem. However, it would be premature to call this a complete solution, as the problem extends beyond the virial radius of the Milky Way (Garrison-Kimmel et al. 2014).

There is another problem with the model in Figure 13. Close examination reveals that the rotation velocity is slightly overfit.
at small radii and underfit at large radii. This is a consequence of the cuspy center of the NFW halo, which predicts more mass at small radii and less at large radii than indicated by the data (McGaugh et al. 2007). While the fit illustrated in Figure 13 is within the uncertainties, this problem rapidly exacerbates as we allow for a nonzero bulge component. As the mass distribution of the stars becomes more concentrated than the inward extrapolation of a pure exponential disk, the compression of the halo becomes more severe. The inner rotation curve becomes unrealistic at fairly modest bulge fractions. This problem comes as no surprise, as it has been seen before (Dubinski 1994; Abadi et al. 2003). The compression of an initially cuspy dark matter halo by a dense stellar bulge predicts much more dark matter at small radii than tolerated by observations (Portail et al. 2015). We only escape this problem in Figure 13 because we have made no attempt to model the inner 3 kpc. Piffl et al. (2015) similarly conclude that compression in the inner regions needs to be counteracted in some way.

Note that it does not matter whether the inner mass concentration is a bulge or a bar: the compression occurs in either case (Sellwood & McGaugh 2005). While bulge formation may be chaotic, the adiabatic assumption remains fairly good (Choi et al. 2006). The secular growth of a bar within the disk is the poster child for an adiabatic process. Dynamical friction of the bar against the halo (Debattista & Sellwood 2000; Athanassoula 2003) may transfer angular momentum to the halo, but this is unlikely to counteract the compression (Sellwood 2008). We can, as always, invoke feedback, but it is far from obvious that this will have the desired effect. We therefore urge caution in the interpretation of Figure 13 and Table 3, which appears promising but leaves important questions unaddressed.

4.7. The Predicted Vertical Force

The surface density of baryons can be constrained by the vertical restoring force to the disk (Kuijken & Gilmore 1989; Bovy & Rix 2013). This provides an independent check on the surface densities derived here from the radial force. More generally, the dark matter density inferred from the vertical force may differ from the spherical average given in Section 4.6 if the halo is oblate or if there is a distinct “dark disk” in addition to the quasi-spherical halo (Read 2014; Silverwood et al. 2015). It is therefore interesting to compare the results derived from vertical and radial force analyses.

The vertical force is given by the baryonic surface density and a term containing \( \partial V^2 / \partial R \) that emerges from the Poisson equation as in Equation (8). This “tilt” term can be cast in terms of the Oort parameters so that the vertical force \( K_z \) is given by

\[
K_z = 2\pi G \left( \Sigma_d + \Sigma_b + \Sigma_g \right) + 2Z (A^2 - B^2). \tag{12}
\]

We can use the information from Table 2 to predict \( K_z \).

The predicted vertical force of model 4QMB is shown in Figure 14. Also shown are the data of Bovy & Rix (2013). Other fourth-quadrant models give similar results. First-quadrant models have more variation, but follow the same trend.

From examination of Figure 14, it is clear that the models constructed here from the radial force and the MDAR are compatible with the vertical force measured by Bovy & Rix (2013). Indeed, the fourth-quadrant models provide very nearly as good a fit to these data as does the exponential fit made by Bovy & Rix (2013). This occurs despite the fact that the two quantities are completely independent. The force predicted by our models follows from Equation (12) with no adjustment. Indeed, this is a true prediction, as the essential information already appears in Table 3 of McGaugh (2008).

Looking in detail at Figure 14, the only place where the model does not provide a good description of the data is over the 1 kpc range \( 5.5 < R < 6.5 \) kpc. This is where the contribution of \( (A^2 - B^2) \) is greatest (Figure 8) owing to the Centaurus spiral arm. Elsewhere, the match is impeccable.

The difference between model and data may be real, as the two probe different regions of the Galaxy. The data used by Bovy & Rix (2013) are located primarily along \( \ell \approx 0 \) (J. Bovy 2013, private communication). The fit to the terminal velocities follows the locus of tangent points. This locus deviates most from \( \ell = 0 \) precisely at the midlatitudes where the model and data disagree. If the feature in the terminal velocities is indeed due to the Centaurus spiral arm, one would expect it to shift in radius with azimuth and perhaps change in amplitude as well. Thus, the one apparent failing of the model may actually contain further information about Galactic structure.

The vertical force depicted in the top panel of Figure 14 is evaluated at \( Z = 1.1 \) kpc. Equation (12) depends linearly in \( Z \) on the \( A^2 - B^2 \) term. This tilt term can become very important when the derivative of the rotation curve is large. To illustrate this sensitivity, in the bottom panel we illustrate the relative force evaluated at \( Z = 600 \) pc and \( Z = 2 \) kpc. While these retain the same basic shape as the \( Z = 1.1 \) kpc value at most radii, their amplitude changes quite a bit where the tilt is large.

Variation in the tilt has important implications for attempts to evaluate the vertical force in order to determine the local dark matter density (e.g., Bienaymé et al. 2014; Silverwood et al. 2015). Normally, one assumes that \( \partial V^2 / \partial R = 0 \), or at least that it varies slowly and continuously. This does not appear to be the case (Figure 8). Instead, the tilt term can vary substantially from place to place and change suddenly. This might go some way to reconciling contrary analyses (e.g., Bovy & Tremaine 2012 and Moni Bidin et al. 2012), but it is dreadfully inconvenient. If this effect is significant, it may manifest as apparent inconsistencies within the Gaia data when one assumes no fluctuations in the derivative of the rotation curve. It also appears that the odds are good that many tracers could be affected: there are fluctuations everywhere (Faure et al. 2014; Antoja et al. 2015; Xu et al. 2015).

4.8. The Relation to MOND and the Nature of Dark Matter

We have used the MDAR (McGaugh 2004) to connect the terminal velocity curve to the surface density of the Galactic stellar disk. With the calibration of the MDAR with external galaxies (McGaugh 2014), this is a purely empirical exercise. This approach is valid irrespective of the physics underlying it. What works for other spirals works for the Milky Way.

How the MDAR comes to be remains a mystery. In the context of \( \Lambda \)CDM, we are obliged to imagine that this very uniform scaling relation somehow emerges from the chaotic process of feedback during galaxy formation. Alternatively, it could be that the appearance of a universal effective force law in galaxy data (the MDAR) is an indication of an actual modification of the force law (MOND; Milgrom 1983).
In the conventional dark matter context, the baryons are embedded in a dark matter halo. There is no way to attempt the exercise successfully performed here because there is too much freedom in the dark halo model. If we start by assuming an exponential disk, we never get to the point of fitting the bumps and wiggles, as we must first fix the scale length. This is effectively impossible, as there is degeneracy between the dark matter halo and baryon distribution: one can shorten or lengthen the disk scale length to accommodate one halo model or another. If, for example, we start by assuming an NFW halo—a very reasonable starting point in the context of $\Lambda$CDM—then we are immediately pushed toward adopting longer scale lengths for the reasons discussed in Section 4.6. As a consequence of the central cusp in NFW halos, we never approach the maximum disk limit and never suspect that the bumps and wiggles might be connected to the subdominant baryon mass. If instead we start with a maximum disk, then we are driven to overestimate the stellar surface density at large radii to explain the bumps and wiggles near the solar circle. This leaves little room for dark matter at small radii, so we may be inclined to adopt a halo model with a low-density core. This can fit the data, but we have no cosmological context for the halo parameters and remain afflicted with considerable degeneracy between them and the baryons. In short, this exercise has not been done before because it is not possible without the MDAR.

The MDAR was uniquely anticipated by Milgrom (1983). The mapping between terminal velocities and features in the baryon distribution is very natural in MOND (McGaugh 2008). It is not natural in the context of dark matter.

That said, the results for the vertical force may pose a problem for MOND. The vertical force is computed conventionally in Section 4.7. The tilt term in Equation (12) is derived from the normal Poisson equation, not a modified version thereof (Bekenstein & Milgrom 1984). The derivative of the rotation curve thus implicitly assumes that dark matter is the reason why the total rotation exceeds that predicted by Newton for the baryons. As can be seen in Figure 14, this works quite well.

In MOND, the Newtonian prediction for the radial force is amplified as the acceleration decreases below the critical value, $a_0$ (Milgrom 1983). One would naively expect the vertical force to be enhanced by the same factor as the radial force. The result is to make the dynamical scale length longer in MOND than it is conventionally (by a factor of $\sim 1.25$; Bienaymé et al. 2009). Nevertheless, the conventional computation of the vertical force is consistent with a purely Newtonian maximal disk plus the dark matter halo specified by the MDAR: it has not been stretched by the factor predicted by MOND (but see below).

The match to the vertical force using the conventional formula is very good (Figure 14). It would seem like an extraordinary coincidence that this should occur by accident. By the same token, this can also be said for MOND fits to rotation curves in general: it is hard to imagine that this is a coincidence devoid of physical meaning.

It is worth noting that the Milky Way is not unique in this apparent mismatch between radial and vertical forces in MOND. Angus et al. (2015) analyzed the vertical velocity dispersions of the face-on galaxies of the DiskMass survey (Bershady et al. 2010) in the context of MOND. In general, it is possible to obtain a fit to the rotation curve or to the vertical velocity dispersion profile, but not to both simultaneously. In the case of the DiskMass survey, the thickness of each disk is
not observed directly, so it is possible\(^4\) to obtain a simultaneous fit, but this only comes at the expense of making the disks much thinner than indicated by edge-on samples of similar morphological types. Intriguingly, the shapes of the profiles for both the radial and vertical forces are well predicted, but the amplitude is offset. It is as if MOND is more active in the radial direction than in the vertical direction: \(D(Z|R) < D(R)\).

One interpretation is that the MDAR is simply an empirical scaling law and does not embody new physics. This is tempting, but it leaves unanswered why it occurs in the first place. It is also tempting to dismiss MOND entirely for this discrepancy, and that might be the correct thing to do. We should bear in mind, however, that we did not get this far without it: this apparent failing of MOND is not a success of LCDM. Indeed, it might be considered a success if a LCDM galaxy formation model came within a factor of 1.25 of matching the scale lengths inferred independently from the radial and vertical forces, if such a test could ever be made. Moreover, MOND does correctly predict (Bienaymé et al. 2009) the tilt angle of the velocity ellipsoid (Siebert et al. 2008).

MOND makes a number of definitive predictions (Milgrom 2015). A precise mapping between radial and vertical force is not one of them. For this, we need a specific version of the theory: one can construct MOND theories by modifying either gravity (e.g., Bekenstein & Milgrom 1984; Milgrom 2010) or inertia (Milgrom 1994, 2006). The increase of the scale length by a factor of 1.25 over the Newtonian expectation is specific to the formulation of Bekenstein & Milgrom (1984). In the case of modified inertia, the theory is inevitably nonlocal (Milgrom 1994), with the consequence that the dynamics is trajectory dependent. So it is conceivable that the modest apparent mismatch between radial and vertical forces in MOND is telling us something further about the correct underlying theory.

Another possibility to consider is something else entirely. For example, Blanchet & Le Tiec (2009) have proposed a type of dipolar dark matter that reproduces MOND’s successes in galaxies while preserving those of LCDM on larger scales. Similarly, a dark matter superfluid (Khoury 2015) may possibly explain galaxy dynamics while behaving differently in clusters of galaxies (Berezhiani & Khoury 2015). It is unclear at present what these theories predict for the vertical force of disks.

What is clear is that it is important to explore new ideas, with an open mind as to whether the mass discrepancy problem is caused by some form of (not necessarily cold) dark matter or a modification of dynamical laws. A minimum requirement for a successful theory is to explain the observed coupling between baryons and dynamics. It is far from obvious that one can reasonably hope that the unique effective radial force law observed in galaxies embodied by the MDAR will somehow emerge from the chaotic feedback processes widely invoked in the context of cold dark matter.

\(^4\) A further complication is that the conventional analysis of the DiskMass data (Martinsson et al. 2013) obtains submaximal disks, contrary to the results here. The mean DiskMass mass-to-light ratio is a factor of \(~2\) lower than that anticipated by stellar population models (McGaugh & Schombert 2014, 2015). One possibility is that the stellar population that dominates the lines from which the vertical velocity dispersion is measured has a different scale height than the bulk of the stellar mass (e.g., the fractional contribution to the spectra by red supergiants can be larger than their contribution to the mass). Another possibility is that we are not yet in a position to rigorously explain the vertical force in disks, even conventionally.

4.9. Hobgoblins of Inconsistency

The models presented here are intended to provide a first step forward from the simplistic assumption of an exponential stellar disk. As discussed above, they have a number of virtues, such as a realistic radial mass distribution that correlates with observed spiral structure and is consistent with independent data. However, they are not a complete solution, and we should be aware of a number of minor respects in which they are not self-consistent.

The models constructed here are azimuthally symmetric. Though the radial mass distribution \(\Sigma(R)\) is not smooth as in the usual exponential approximation, we have made no attempt to construct two-dimensional models \(\Sigma(R, \phi)\). Yet the bumps and wiggles that we identify in the surface density profile appear to correspond to spiral arms, which certainly vary in azimuth. Indeed, this is apparent in the difference between the first- and fourth-quadrant terminal velocity curves, which imply bumps and wiggles at different radii. This is expected as the stellar mass follows the pitch angles of spiral arms to different radii as the azimuth varies. In effect, the models provide a snapshot of \(\Sigma(R)\) along the azimuths probed by the terminal velocities.

In external galaxies, the full velocity field is observed, and a true azimuthal average is made. This is what goes into the MDAR. In the Milky Way we only sample the rotation curve in the first and fourth quadrants, and not at all on the side opposite the Galactic center. Consequently, local variations in the surface density can have a larger impact on the rotation curves deduced from the terminal velocities than they might over a complete azimuthal average. This probably results in the features inferred here being stronger than they would be after azimuthal averaging, emphasizing the importance of local features.

The likelihood of azimuthal as well as radial variations further complicates the prospects for Jeans analyses. In addition to the radial variation in \(\langle dV/dR \rangle^2\) discussed in Section 4.2, one should also worry about \(dV/d\phi\) and how disk structure varies around the disk (Olling & Dehnen 2003). The Milky Way disk could be grand design or a patchwork of flocculent spiral structure, rendering the usual azimuthally symmetric, radially smooth exponential disk approximation inadequate for the analysis of complex data like those provided by Gaia.

The terminal velocity curves in the first and fourth quadrants are different, and this difference appears to reflect a real difference in the structure of the Galaxy. This difference leads to differences in the first- and fourth-quadrant models. These differences lead to rather different inferences for the local disk surface density \(\Sigma(R_0)\) and LSR velocity \(\Theta_0\). The models fit to the first-quadrant terminal velocities are quite consistent with the assumed (solar neighborhood) inputs for these values. The fourth-quadrant models prefer larger values of both \(\Sigma(R_0)\) and \(\Theta_0\). This may indicate that the solar neighborhood is a bit underdense relative to the azimuthal average for \(R_0\). This is quite reasonable: there is no reason for the local patch around the Sun to be exactly average, and the Sun is known to currently reside in a low-density interarm region.

A greater difficulty arises from the difference in \(\Theta_0\). A particular value of \(\Theta_0 = 220\ km\ s^{-1}\) has been assumed (Section 2.1) in order to derive the rotation curve. The fits to the rotation curve then imply a slightly larger value for \(\Theta_0\). We should therefore iterate the solution, changing the rotation curve and refitting the surface densities. In practice, this makes little difference to the result and falls within the uncertainties of
solar motion and gas turbulence, and indeed of the variation in structure that we are attempting to ascertain.

First-quadrant models give a circular velocity for the LSR of $\Theta_0 \approx 224 \text{ km s}^{-1}$, similar to the assumed value. For the observed proper motion of Sgr A* of $30.24 \text{ km s}^{-1} \text{kpc}^{-1}$ (Reid & Brunthaler 2004), this implies a rather high solar motion of $V_0 \approx 18 \text{ km s}^{-1}$. While not inconceivable (Bovy et al. 2015), this is not consistent with the low solar motion found from the terminal velocity data themselves by Clemens (1985, $V_0 = 7 \text{ km s}^{-1}$). Matters become slightly worse if we adopt $(\Theta_0 + V_0)/R_0 = 30.57 \text{ km s}^{-1} \text{kpc}^{-1}$ estimated from star-forming regions by Reid et al. (2014), as this implies $V_0 \approx 21 \text{ km s}^{-1}$.

In contrast, the fourth-quadrant fit yields a higher LSR velocity of $\Theta_0 \approx 232 \text{ km s}^{-1}$. For the adopted $R_0$, this is consistent with the proper motion of Sgr A* if the solar motion is $V_0 \approx 10 \text{ km s}^{-1}$. Similarly, the star-forming regions imply $V_0 \approx 13 \text{ km s}^{-1}$. These are plausible values consistent with many independent determinations (Binney 2010; McGaugh & Binney 2010; Schönrich et al. 2010; Sharma et al. 2014).

At present, the true value of $\Theta_0$ seems systematically uncertain at the $\pm 10 \text{ km s}^{-1}$ level. Indeed, this particular issue is greatly complicated by local gradients in the surface density and rotation curve in the immediate vicinity of the Sun (Olling & Merrifield 1998). The former should vary as we encounter the Perseus arm slightly outside the solar radius (Siebert et al. 2012). This can, in turn, cause a sudden change in $V(R)$. Consequently, it does not seem possible to improve much on the present models without a great deal more information than currently available.

5. CONCLUSIONS

We have applied the MDAR calibrated by external galaxies (McGaugh 2004, 2014) to the terminal velocity curves observed in the first (Clemens 1985) and fourth (Luna et al. 2006; McClure-Griffiths & Dickey 2007) quadrants to construct mass models for the Milky Way. These models provide a nonparametric numerical estimate of the surface density profile of the stellar disk $\Sigma_d(R)$. This provides a first step beyond the simple assumption of an exponential disk.

The Galactic disk inferred here is maximal and favors a short scale length. The precise value of the scale length depends on the portion of the disk that is fit, potentially relieving the tension between apparently discrepant determinations. The scale length itself, while useful, is less fundamental than the pattern of structure to which it is fit.

The bumps and wiggles observed in the terminal velocities correspond well to known spiral arms. The arms are massive and have the expected effect on the observed kinematics. These conclusions are independent of the bulge fraction or assumptions about the disk thickness, though these do have a small effect on the details of individual models.

The Milky Way appears to be a normal spiral galaxy. It obeys scaling relations like the Tully–Fisher relation, the size–mass relation, and the disk maximality–surface brightness relation. It is somewhat compact for its stellar mass, but resides well within the large intrinsic scatter of the size–mass relation.

In comparing the Milky Way to other galaxies, it is important to compare equivalent measures. It is sometimes claimed that the Milky Way deviates from the Tully–Fisher relation, but these claims seem to stem largely from using $\Theta_0$ as a proxy for other quantities. The LSR velocity $\Theta_0$ is neither the peak of the rotation curve ($V_p > \Theta_0$) nor the outer, quasi-flat velocity ($V_f < \Theta_0$). Though the difference between these quantities seems small, the scatter in the Tully–Fisher relation is very tight (McGaugh & Schombert 2015), so any difference is readily apparent.

The distribution of stellar mass in the Galactic disk inferred here from the radial force is consistent with that inferred independently from the vertical force (Bovy & Rix 2013). Indeed, the vertical force is correctly predicted from our models with no adjustment. The only point of disagreement between the two is where they probe different parts of the Galaxy, emphasizing the importance of local structures like spiral arms.

One consequence of the bumps and wiggles in the terminal velocity curves and their correspondence to variations in the stellar surface density is that the gradient of the rotation curve $dV/dr$ fluctuates on the scale of hundreds of parsecs. This fluctuation is not subtle, having rms amplitude $\langle (dV/dr)^2 \rangle^{1/2} \approx 14 \text{ km s}^{-1} \text{kpc}^{-1}$ in the fourth quadrant. This has important implications for Jeans analyses, which frequently invoke the flatness of the rotation curve to ignore terms involving its gradient. Such approximations are unlikely to be adequate, especially to the analysis of the upcoming Gaia data. Indeed, it is not even adequate to assume a finite slope: $dV/dr$ switches signs repeatedly on kiloparsec scales.

The distribution of dark matter can be inferred from our models. The amplitude of the mass discrepancy locally is $D(R_0) \approx 1.5$, leading to a spherically averaged density of dark matter in the solar neighborhood $\rho_{\text{DM}} \approx 0.009 \text{ M}_\odot \text{ pc}^{-3}$ (0.34 GeV cm$^{-3}$). These values are likely accurate to 20%, though we caution that systematic errors outweigh random ones.

More generally, the detailed radial distribution of dark matter can be empirically inferred from the models. This is not precisely equivalent to any of the common halo forms, but can be tolerably approximated by an adiabatically compressed NFW halo. The compression is important to reconciling the Milky Way halo with the mass–concentration relation expected in $\Lambda$CDM and may help alleviate (though not solve) the too-big-to-fail problem. We do not, however, attempt to fit the inner 3 kpc, where the cusp-core problem persists.

As the only theory that predicted the MDAR, the successes of these models can be interpreted as successes of MOND. However, we identify a modest but apparently real tension between the radial and vertical forces in MOND: when the radial force is well fit, the vertical force is overpredicted. This may be a genuine problem for MOND, but it may also be a hint about the deeper theory underlying the observed phenomena. We should keep an open mind about the underlying cause of the MDAR, which may be a hint about the nature of dark matter as well as modified dynamical laws.

Irrespective of the underlying cause of the MDAR, the models presented here are based on an empirical calibration thereof. As such, they are empirically valid. While the MDAR phenomenon deserves a better explanation than is currently available, there is nothing to preclude us from using this information in our exploration of the Galaxy.

I thank Benoit Famaey, James Binney, Heather Morrison, Joss Bland-Hawthorn, and Paul Harding for conversations on Galactic structure, and Jerry Sellwood and Harley Katz for work on adiabatic compression. This publication was made possible through the support of a grant from the John Templeton Foundation. The opinions expressed in this
