



# A Lower Limit on the Mass of Our Galaxy from the H3 Survey

Dennis Zaritsky<sup>1</sup> , Charlie Conroy<sup>2</sup> , Huanian Zhang<sup>1</sup> , Rohan P. Naidu<sup>2</sup> , Ana Bonaca<sup>2</sup> , Nelson Caldwell<sup>2</sup> ,  
Phillip A. Cargile<sup>2</sup> , and Benjamin D. Johnson<sup>2</sup>

<sup>1</sup> Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721-0065, USA; [dennis.zaritsky@gmail.com](mailto:dennis.zaritsky@gmail.com)

<sup>2</sup> Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA

Received 2019 September 4; revised 2019 November 22; accepted 2019 November 24; published 2020 January 15

## Abstract

The timing argument provides a lower limit on the mass of the Milky Way. Using a sample of 32 stars at  $R > 60$  kpc drawn from the H3 Spectroscopic Survey and mock catalogs created from published numerical simulations, we find that  $M_{200} > 0.91 \times 10^{12} M_{\odot}$  with 90% confidence. We recommend using this limit to refine the allowed prior mass range in more complex and sophisticated statistical treatments of Milky Way dynamics. The use of such a prior would have significantly reduced many previously published uncertainty ranges. Our analysis suggests that the most likely value of  $M_{200}$  is  $\approx 1.5 \times 10^{12} M_{\odot}$ , but establishing this as the Milky Way mass requires a larger sample of outer halo stars and a more complete analysis of the inner halo stars in H3. The imminent growth in the sample of outer halo stars due to ongoing and planned surveys will make this possible.

*Unified Astronomy Thesaurus concepts:* Milky Way Galaxy (1054); Milky Way dark matter halo (1049); Milky Way stellar halo (1060)

## 1. Introduction

The mass of our own galaxy would seem to be one of a modest set of fundamental parameters that is both necessary for a complete understanding of our universe and relatively straightforward to measure. However, convergence on a precise final answer has proven difficult. Zaritsky et al. (1989) presented three mass estimates, two of which are based on a statistical treatment of the radial velocities of satellite galaxies and one of which is a lower limit based on the properties of a single, extreme satellite (Leo I):  $(0.93^{+0.41}_{-0.12}, 1.25^{+0.84}_{-0.32}, 1.3^{+0.33}_{-0.33}) \times 10^{12} M_{\odot}$ . Exemplar, state-of-the-art studies conclude that the Galaxy’s mass is  $(0.85^{+0.23}_{-0.26}, 0.96^{+0.29}_{-0.28}, 1.29^{+0.37}_{-0.37}) \times 10^{12} M_{\odot}$  based on satellite dynamics and the stellar escape velocity (Patel et al. 2018; Grand et al. 2019). The casual observer may be somewhat underwhelmed by 30 years of “progress.”

This cursory comparison is misleading. Studies in the intervening years identified areas ripe for improvement in earlier analyses (e.g., Li & White 2008; Patel et al. 2018), utilized the ever-increasing amount of available data (e.g., Watkins et al. 2010), most importantly by adding proper motions (e.g., Boylan-Kolchin et al. 2013), and developed the cosmological framework in which we construct our models (e.g., Barber et al. 2014; Carlesi et al. 2017; Patel et al. 2017). The surprise is that better data in combination with more sophisticated models did not manifestly result in evident quantitative improvements. Derived parameter uncertainties grew or remained at best constant even as the quantity and quality of the data improved, because the increasingly sophisticated models contain more degrees of freedom. The apparent stagnation is therefore partly an illusion. Improvements over the years have implicitly tested for systematic biases in previous, oversimplified and underconstrained models. However, as happens sometimes, there were no such biases of a magnitude above and beyond the quoted precision of the earlier measurements. This result should be both disappointing and reassuring to practitioners in the field.

How to proceed? There is a tension between using increasingly realistic models and the larger parameter space these models access. One way to mitigate this tension is to develop more constraining parameter priors. Lower mass limits, such as provided by a measurement of the escape velocity or the timing argument, could possibly help, but have historically been discounted when in conflict with other results (e.g., Xue et al. 2008; Gibbons et al. 2014). The rejection of this additional information was not unreasonable given that the largest (most constraining) lower mass limit came from the study of the single satellite galaxy Leo I (Zaritsky et al. 1989; Li & White 2008), which might not be on a simple orbit (Sales et al. 2007; Mateo et al. 2008).<sup>3</sup> Subsequent work, including new proper motion measurements and a careful theoretical treatment, concluded that Leo I is indeed bound to the Galaxy and likely to be a strong constraint (Boylan-Kolchin et al. 2013). However, any mass limit from Leo I is still susceptible to the usual concern of relying too heavily on the properties of a single object.

Our aim is to reinforce a meaningful lower mass limit. In doing this we have two key points of emphasis. First, we aim to use the most distant dynamical tracers available to measure the total Galactic mass without having to extrapolate too much in radius. Meeting this goal can be challenging because tracer particles are generally available in quantity only to distances  $< 50$  kpc (e.g., Williams et al. 2017), and the Milky Way’s virial radius is likely to be 250 kpc or larger. Second, we want to obtain a conservative, but stringent inference using a straightforward and simple analysis.

The two principal methods for determining a Galactic lower mass limit are based on the escape velocity and the timing argument.<sup>4</sup> The escape velocity argument is highly intuitive:

<sup>3</sup> An alternative lower limit that is also highly constraining comes from the Local Group timing argument, but that measures the sum of the Milky Way and M31 masses. In this case uncertainty is introduced by the partitioning of the calculated mass.

<sup>4</sup> A non-dynamical lower limit can be formulated using the cosmological baryon fraction (Zaritsky & Courtois 2017), but the uncertainties are large until the circumgalactic gas mass is measured more precisely.

the Galaxy needs to have sufficient mass to retain the fastest moving stars that are still found near it. Several treatments of the measurement of escape velocity have been published recently (Piffl et al. 2014; Williams et al. 2017; Deason et al. 2019; Grand et al. 2019). While limits can be obtained from the fastest moving stars in any sample, the most meaningful limits come from large samples of stars where the high-velocity tail truly includes stars moving at nearly the escape velocity. Such studies are necessarily carried out with samples of stars at modest Galactic radii of, at best, a few tens of kiloparsecs. Because of this limitation, evaluating the total Galactic mass requires a model of the Galactic potential and the phase space distribution of the stars. The uncertainties arising from the model choice can then be dominant (for example, see discussion by Deason et al. (2019) comparing their inference to that of Piffl et al. (2014), which differs despite similar direct measurements).

The second method that can provide a lower mass limit is the timing argument (Kahn & Woltjer 1959; Sandage 1986). The timing argument posits that all matter initially expands away from the future galaxy due to universal expansion and that the subsequent gravitational pull of the galaxy eventually recalls some of that matter. In the simplest version of this scenario, the tracer particles—galaxies or stars—follow a radial orbit around a point mass. In a slightly more complex version, orbit crossing leads to the behavior seen in secondary infall models without angular momentum (Fillmore & Goldreich 1984; Bertschinger 1985) or with angular momentum (White & Zaritsky 1992). More complete treatments are currently being investigated to define the splashback radius and study the growth of structure (e.g., Diemer et al. 2017).

In the simplest version, the system mass can be solved for analytically if the current separation and relative velocity between the Galaxy and a tracer particle are known. One does not need to know the phase space distribution of the population of tracer particles. The advantage provided by the timing argument comes from using one additional piece of information, namely the constrained orbital time (which is the age of the universe). Because this analysis does not require us to assume either a velocity distribution function or a radial tracer density profile, we do not need detailed knowledge of the accretion history of the tracer particles or the assumption of dynamical relaxation.

The timing argument is an exceedingly elegant and robust treatment, particularly given the now well-measured age of the universe. There are a few, limited caveats. First, the standard solution is for a radial orbit. If the tracer particle has non-zero tangential velocity then the actual mass can be significantly larger than that inferred (see Boylan-Kolchin et al. 2013 for how the additional tangential velocity component of Leo I as measured with the *Hubble Space Telescope* affected the inferred Galactic mass). Second, the Galaxy is modeled as a point mass. Neglecting both its extent and its growth over time results in the classical timing argument underestimating the current mass within the orbit of the tracer particle. Finally, there is the possibility of an invalid result if the tracer particle has a large peculiar velocity resulting from interactions between it and other mass concentrations. This, for example, was one of the principal concerns in the application of the timing argument to Leo I (Sales et al. 2007; Mateo et al. 2008). If Leo I has been flung toward the Milky Way by an encounter with a third galaxy, then its current position and velocity do not reflect

those of the assumed cosmologically motivated, radial orbit about the Milky Way. All of these concerns are addressed when using numerical simulations both to calibrate the estimator and to determine the likelihood of catastrophic failures. Catastrophic failures can be mitigated by repeating the analysis using multiple, independent tracer particles.

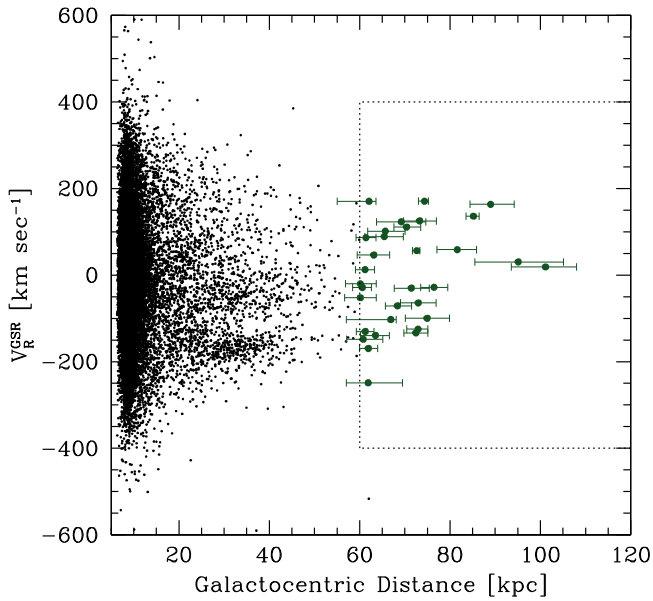
Numerical simulations confirm that the simple timing argument systematically underestimates the system mass and that statistical corrections are necessary to accurately recover the true mass (Kroeker & Carlberg 1991; Li & White 2008). For the Local Group timing argument, using the orbit of M31 and the Milky Way, that statistical correction requires increasing the result obtained from the application of the simple timing argument by about a factor of 1.7. Those studies also confirm that the scatter between the estimated and true masses can be large when using only M31 and the Galaxy, or Leo I and the Galaxy. Further studies have expanded on the use of the basic principles of this approach (Phelps et al. 2013).

We present the application of the timing argument to a sample of outer halo ( $R > 60$  kpc) stars from the H3 Spectroscopic Survey (Conroy et al. 2019). We discuss the sample of stars in Section 2, discuss the application and results from the timing argument to both the models and data in Section 3, discuss the results in Section 4, including an illustration of how constraining a robust lower mass limit can be on previous results, and summarize our conclusions in Section 5.

## 2. The Sample

H3 is a sparsely sampled survey of  $\sim 15,000$  deg<sup>2</sup> of sky that will provide stellar parameters and spectrophotometric distances for  $\approx 200,000$  stars to  $r = 18$  (Conroy et al. 2019). The procedure used for the parameter determinations and distance estimation is developed and presented by Cargile et al. (2019). From the currently available set of observed and analyzed stars, about 80,000, we select those with fitting flag set to 0 (denoting fitting was problem-free), spectral signal-to-noise ratio per pixel  $> 3$ , a stellar rotational velocity  $< 5$  km s<sup>-1</sup>, an effective temperature  $< 7000$  K, and a tangential velocity,  $v_T$ , that is  $< 1000$  km s<sup>-1</sup>. These criteria are all meant to remove stars with problematic and unphysical parameter determinations and were arrived at after a careful examination of numerous fits. From that sample, we then select those stars with Galactocentric distance  $> 60$  kpc (Figure 1) to focus on stars that are likely to have completed only a few Galactic orbits at most.

In Figure 1 we highlight the 32 outer halo stars within a selection box that adds the criterion that the radial velocity in the Galactic Standard of Rest,  $v_R^{\text{GSR}}$ , satisfies  $|v_R^{\text{GSR}}| < 400$  km s<sup>-1</sup>. This selection removes the one star at  $R > 60$  kpc that clearly does not belong in our sample (with  $v_R^{\text{GSR}} < -500$  km s<sup>-1</sup>). Whether this star represents a particularly catastrophic parameter determination or a physically compelling outlier, is unknown and merits follow-up, but in neither case is this star suitable for our mass estimation. A preliminary investigation of the model fit for this one star suggests that this system has a large residual flux in the IR and is therefore perhaps an unresolved binary star for which the parameter estimation failed. The radial velocity errors in general are of the order of a few km s<sup>-1</sup>, so the distance errors dominate in our calculations and are plotted in the figure. The distribution of stars follows the apparent caustics that can be visually defined at  $R < 60$  kpc. Substructure is clearly evident



**Figure 1.** The H3 sample and outer halo stars. From the full sample we select a sample of 32 outer halo stars ( $R > 60$  kpc,  $-400 \text{ km s}^{-1} \leq v_R^{\text{GSR}} \leq 400 \text{ km s}^{-1}$ ) for the timing argument analysis, as shown highlighted within the dotted selection box.

**Table 1**  
Auriga Model Parameters

Auriga Run	$M_{200}$ ( $10^{12} M_{\odot}$ )	$R_{200}$ (kpc)
6	1.04	214
23	1.58	245
24	1.49	241
27	1.75	254

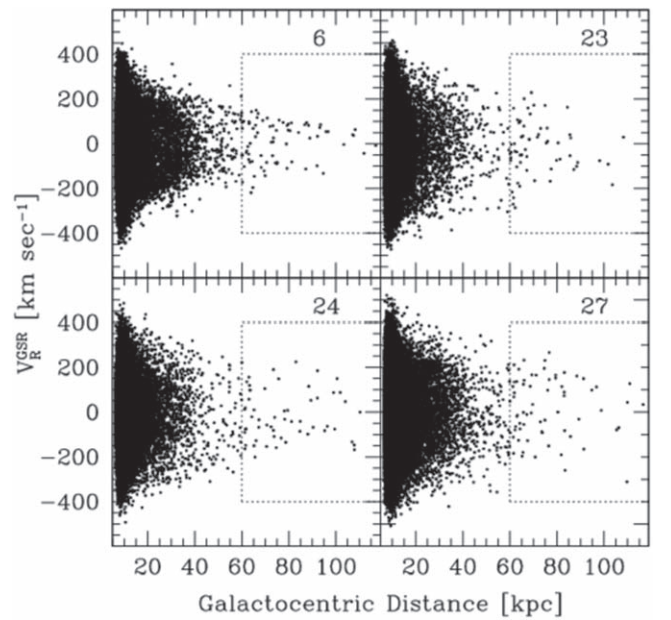
at  $R < 50$  kpc and illustrates the difficulty faced by any statistical analysis of the velocities where one needs to model the phase space distribution.

### 3. The Timing Argument

The basic timing argument is simple in its application. Using the equations provided by Sandage (1986), we find the smallest orbital phase,  $\theta$ , that solves the equations relating distance, velocity, and orbital time for a selected star. The orbital time is taken to be the age of the universe, 13.75 Gyr (Hinshaw et al. 2013). Once we solve for the orbital phase, we calculate the mass.

#### 3.1. Calibration Using Simulations

We construct H3 mock catalogs from the Auriga suite of simulations (Grand et al. 2017). Specifically, we have catalogs that mock the H3 selection, including a 10% distance uncertainty, for their model numbers 6, 16, 21, 23, 24, and 27. We construct the plots corresponding to our Figure 1, applying all of the same selection criteria except those aimed at removing stars with poor spectral fits. We find that models 16 and 21 have dominant substructures visible at large radii ( $R > 60$  kpc) that are qualitatively different from what we find in H3 (but perhaps similar to what is seen in H3 at smaller radii, between 25 and 40 kpc). The remainder of this discussion uses only the four models that do not show such structure (see



**Figure 2.** Mock H3 catalogs created from Auriga models (6, 23, 24, and 27) from Grand et al. (2017). The selection box of outer halo stars from Figure 1 is also shown. The model number is given in the upper right of each panel. Each model contains roughly twice as many stars in the selection box as H3 because we apply the additional data quality and model fitting criteria to the actual data.

Table 1 and Figures 2 and 3) and are, at least visually, plausible matches to the data.

We are interested in using timing argument calculations for stars at  $60 \text{ kpc} < R < 120 \text{ kpc}$  to recover  $M_{200}$ , where  $M_{200}$  is the mass contained inside a radius within which the mean density equals 200 times the critical density. When we apply the timing argument calculation to all of the star particles in each of these four models and then scale the recovered timing mass and the calculated apocenter<sup>5</sup> of each star particle by the corresponding  $M_{200}$  and  $R_{200}$  of each model, we find the relation between derived mass and apocenter shown in the upper panel of Figure 4.

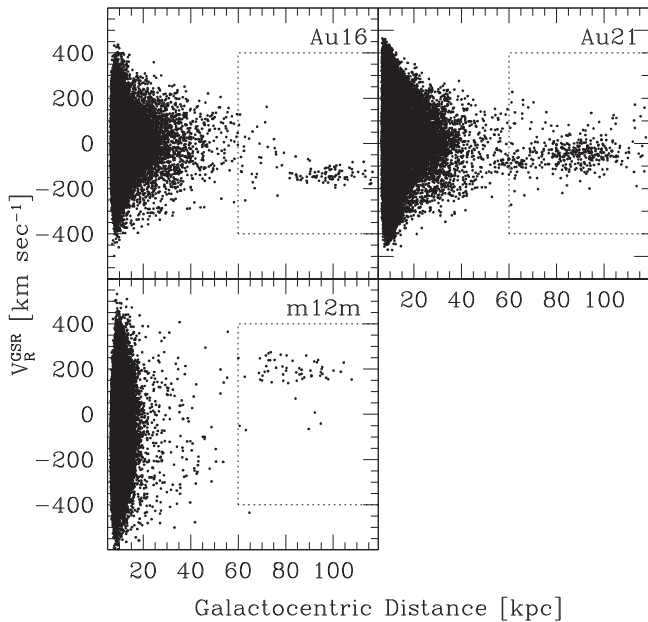
As expected, the timing mass significantly underestimates  $M_{200}$  in all cases. The closest any estimate gets to the true value in any of the four models is less than half of it. The models show very little scatter in the estimates along the sequence, demonstrating that strong orbital perturbations are not significantly affecting the estimates. Because the four models together have roughly eight times as many stars as we have in the H3 sample (257 versus 32), the likelihood of a star in H3 resulting in an estimate closer to  $M_{200}$  than what we find in the models is small ( $\sim 12\%$ ). Therefore, the results from analyzing the models suggest that one can take the star with the largest timing mass estimate in H3, multiply that result by 2.3 ( $1/0.43$ ) to estimate  $M_{200}$ , and still have a conservative estimate of  $M_{200}$ .

#### 3.2. Application to the H3 Outer Halo Sample

Among the H3 sample, the largest resulting timing mass comes from H3 star 117408280 ( $\alpha = 355^\circ 479908$ ,  $\delta = 15^\circ 82972$ ) and is  $0.49 \times 10^{12} M_{\odot}$ . Using the calibration

<sup>5</sup> We utilize the calculated apocenter distances using the timing argument equations rather than the apocenter distances presented in the H3 catalog, which are calculated in a completely independent manner. We do this to avoid introducing different model assumptions into our current analysis and discussion.



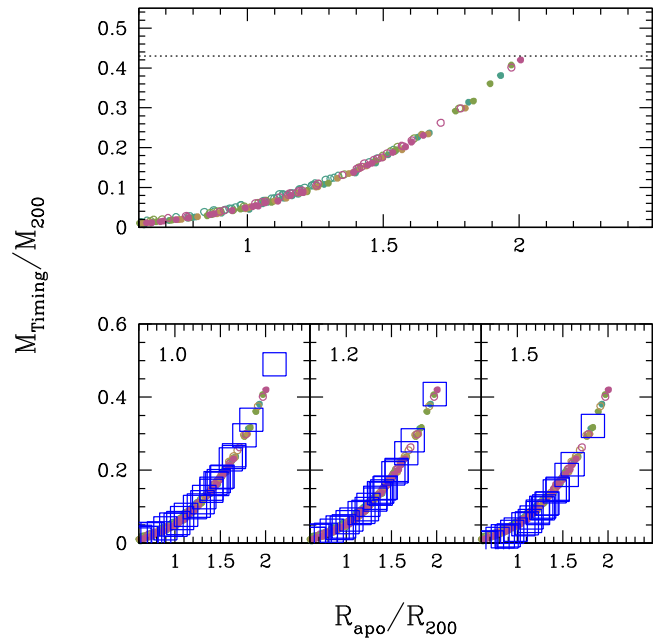


**Figure 3.** The phase space diagrams for the three rejected models (Auriga models 16 and 23, and Latte m12m). Note the clear substructure seen in these relative to the analogous diagrams for either the data (Figure 1) or the accepted models (such as shown in Figure 2).

factor of 2.3 yields a timing mass estimate for  $M_{200}$  of  $1.1 \times 10^{12} M_{\odot}$ . Simulating the effect of the distance uncertainty by drawing 1000 realizations of the distance and re-evaluating the resulting mass estimate, we find that with 90% confidence  $M_{200} > 0.97 \times 10^{12} M_{\odot}$ . This star has a negative radial velocity, eliminating any concern that it happens to be an unbound, hypervelocity star (Brown et al. 2014).

The weakness of this argument is that it relies on identifying the largest, most constraining timing mass estimate. Again, as with Leo I, we are at the mercy of a single object. Therefore, we now approach the problem in a complementary manner by looking at the distribution of timing mass estimates. We compare the distribution of  $M_{\text{Timing}}/M_{200}$  for the H3 data to that drawn from the four simulations. Of course, we do not yet know  $M_{200}$  for the Galaxy. However, if we choose an acceptable value of  $M_{200}$  then the resulting H3 distribution should be statistically consistent with those produced from the mock catalogs. We visualize this approach in the bottom panels of Figure 4, where we have rescaled the H3 results using three hypothetical values of  $M_{200}$  for the Milky Way ( $10^{12}$ ,  $1.2 \times 10^{12}$ , and  $1.5 \times 10^{12} M_{\odot}$ ). To calculate the radial rescaling, we use the relationship between  $M_{200}$  and  $R_{200}$ ,  $M_{200} = (800/3)\pi\rho_c R_{200}^3$ , where  $\rho_c$  is the critical density and we have confirmed the appropriateness of this relation to the values given with the Auriga simulations. The differences are most evident at the largest values of  $M_{\text{Timing}}/M_{200}$ , but the entire distribution of points slides along the locus and this behavior is what we use to obtain our mass estimate.

We apply the Kolmogorov–Smirnov (KS) test to assess which rescaled  $M_{\text{Timing}}/M_{200}$  distributions of H3 can be discriminated from that of the mock catalogs. If we can reject with  $> 90\%$  confidence the hypothesis that the resulting  $M_{\text{Timing}}/M_{200}$  distribution is drawn from the same parent sample as the distribution obtained from the mock catalogs, then we reject the adopted  $M_{200}$  as valid for the Milky Way. In



**Figure 4.** Distribution of scaled timing argument results. We scale the derived mass by  $M_{200}$  for each of the outer halo star particles in each of the four simulations, and scale the calculated apocenter distances by  $R_{200}$  in the upper panel. The different models are represented by different colors, and inbound and outbound stars by closed and open symbols, respectively. All estimates are less than half of the true value regardless of the model and whether the inward or outward moving stars are used. The dotted horizontal line represents  $M_{\text{Timing}}/M_{200} = 0.43$  and lies above all estimates. In the lower panels, we reprise the simulation results in all panels and present the scaled values for the H3 outer halo stars (large open squares) for three different adopted values of  $M_{200}$  ( $10^{12}$ ,  $1.2 \times 10^{12}$ , and  $1.5 \times 10^{12} M_{\odot}$ ) in each of the three panels, as labeled. There are roughly a factor of eight more tracer particles in the combined simulations, so we do not expect the H3 sample to sample the extreme upper end of the distribution better than the simulations. This consideration and visual examination of the figure suggest that  $M_{200}$  for the Milky Way is  $> 10^{12} M_{\odot}$ .

this way, we conclude that  $M_{200} > 0.91 \times 10^{12} M_{\odot}$  (and that  $M_{200} < 2.13 \times 10^{12} M_{\odot}$ ). The KS test, by construction, tends to focus on deviations in the middle of the distribution rather than at the extremes, making this a valued complement to our prior estimate that used only the most extreme star. Indeed, this general behavior is also the case here, where the maximum deviation in the cumulative distributions between data and mocks for our lower limit of  $M_{200} = 0.91 \times 10^{12} M_{\odot}$  occurs halfway through the ranked values of  $M_{\text{Timing}}/M_{200}$  for the H3 stars.

The weakness of this argument is that because we are relying on the distribution of a sample of tracers, such as is done in the estimation of escape velocity, we are assuming that the model is a fair representation of both the tracer distribution in phase space and the underlying potential. We have already seen that substructure can vary significantly from model to model.

We mitigate this concern by combining four different models, hoping to average over any undetected substructure at  $R > 60$  kpc, but even so, an underlying systematic difference in the modeling can skew the result. We address this potential overdependence on one set of simulations by re-deriving our mass limits using the Latte suite of models from the FIRE project (Wetzel et al. 2016; Hopkins et al. 2018) as examined using the tools provided by Sanderson et al. (2018). We use the three different high-resolution models (m12f, m12i, and m12m), which we view in three different orientations to

increase our sample. The models have virial masses of between  $1.2$  and  $1.6 \times 10^{12} M_{\odot}$ . We reject one of the models (m12m) because it exhibits strong substructure at the relevant radii that is qualitatively different than what we observe in H3. Again, using the KS test and the distribution of normalized mass estimates, we conclude with 90% confidence that  $M_{200} > 0.98 \times 10^{12} M_{\odot}$  (and that  $M_{200} < 2.56 \times 10^{12} M_{\odot}$ , with a preferred value of  $1.63 \times 10^{12} M_{\odot}$ ).

The use of different simulations, provided one rejects models with grossly different properties at the relevant radii (see Figure 3), has a minimal impact on our results. We expect the H3 outer halo sample of stars to represent various substructures, but models that have far more substructure than what is evident in the data are an inappropriate comparison sample.

It is highly reassuring that multiple approaches produce nearly identical limits. We adopt the smaller of the limits as the limit that we quote ( $M_{200} > 0.91 \times 10^{12} M_{\odot}$ ) but note that the preferred values of  $1.38 \times 10^{12} M_{\odot}$  and  $1.63 \times 10^{12} M_{\odot}$  (from the most likely scaling values) provide further guidance that the Milky Way mass might actually be  $\approx 1.5 \times 10^{12} M_{\odot}$ .

#### 4. Discussion

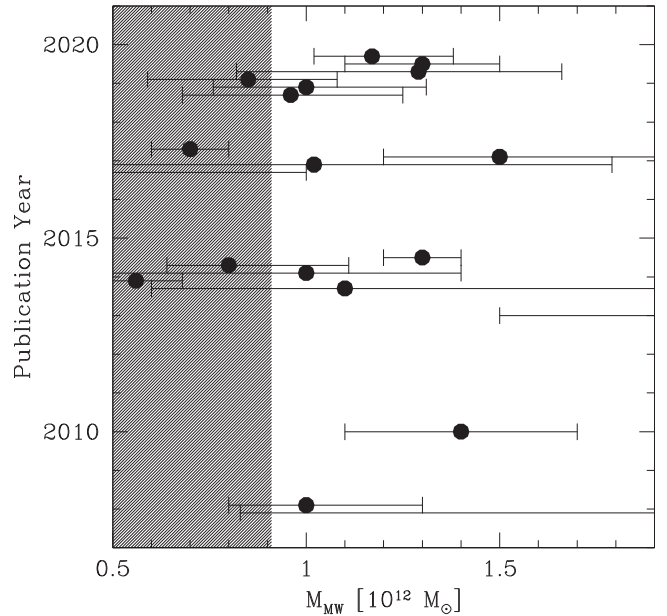
These results will strengthen as the number of independent probes increases. Our mass limit is consistent with that drawn previously from Leo I (Zaritsky et al. 1989; Boylan-Kolchin et al. 2013), lending further credence to both our current limit and that from Leo I, and more data are imminent. First, H3 itself when complete will be three times the size of the current sample. Second, other large surveys of halo stars are ongoing or about to start (Majewski et al. 2017; Martell et al. 2017; Dalton et al. 2012; Deng et al. 2012; Guiglion et al. 2019), from which we expect to gain more examples of such stars. It will be increasingly more difficult to dismiss a lower mass limit of  $\sim 10^{12} M_{\odot}$ .

##### 4.1. How Constraining Is Our Lower Mass Limit?

In Figure 5, we compare a selection of recent measurements (publication since 2008) of the Milky Way mass to our 90% confidence lower mass limit of  $0.91 \times 10^{12} M_{\odot}$ . Most of the published masses refer to either  $M_{200}$  or the virial mass. Only in one of these cases is the mass quoted within a fixed physical radius of 300 kpc. All of these are thought to be roughly interchangeable given the other outstanding uncertainties.

Each of the published analyses, even if it incorporates data from previous work, is independent. We argue that once provided with a robust, non-trivial lower limit each study could significantly benefit by limiting the mass range over which it considers models. In many cases in the figure, about half of the internally allowed Milky Way mass range is inconsistent with the lower limit we provide.

Two results whose uncertainty ranges lie entirely or nearly entirely within our excluded zone in Figure 5 are those of Gibbons et al. (2014) and Dierickx & Loeb (2017). Both of these are based on an analysis of the Sagittarius dwarf, suggesting that perhaps there is something that is not well understood about this system. A third recent analysis of this system concludes that the mass enclosed within 100 kpc is  $\sim 7 \times 10^{11} M_{\odot}$  (Fardal et al. 2019), which is in tension with the previous estimates and more in line with our estimates.

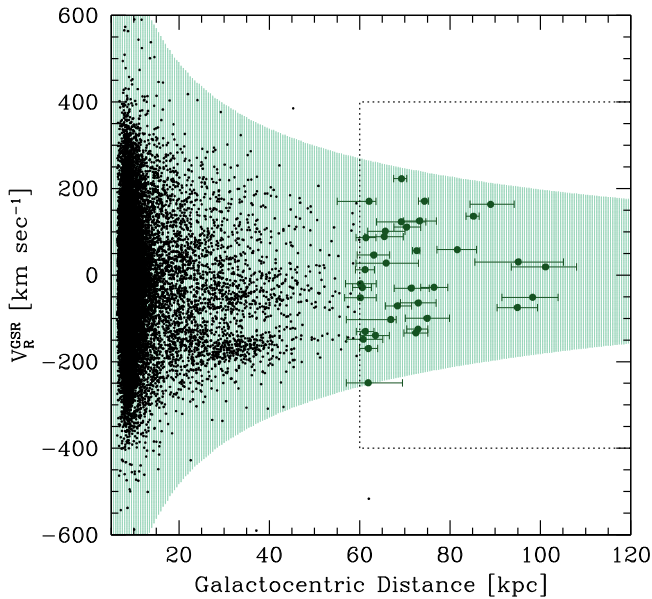


**Figure 5.** A sample of literature measurements of the Milky Way mass. The shaded region illustrates the region excluded by our 90% confidence lower limit. Only two previous measurements are strongly argued against, but for most of the measurement our lower limit reduces the allowed range of values and so provides a useful prior. The data shown represent 19 results (from oldest to most recent): Li & White (2008), Xue et al. (2008), Watkins et al. (2010), Phelps et al. (2013), Barber et al. (2014), Gibbons et al. (2014), Cautun et al. (2014), Kafle et al. (2014), Piffl et al. (2014), Dierickx & Loeb (2017), Patel et al. (2017), Fragione & Loeb (2017), Carlesi et al. (2017), Patel et al. (2018), Deason et al. (2019), Patel et al. (2018), Grand et al. (2019), Posti & Helmi (2019), Callingham et al. (2019). Note that Patel et al. (2018) is represented twice, as the sixth to last and fourth to last.

##### 4.2. Could Our Limit Be a Significant Underestimate?

We associated the star that has the largest timing mass estimate with the upper end of the modeled timing mass estimates to derive a mass limit. It is possible that this star does not represent the upper end of the distribution of Milky Way stars and results in a significant underestimate of the total mass. The complementary limit obtained by using the full distribution of timing mass estimates argues against such a premise, but we noted that this analysis has its own potential problems, namely its dependence on models that might not be producing an accurate match to the phase space distribution of the dynamical tracers, even if they are accurately representing the gravitational potential.

We previously rejected one star at very large negative velocities from the sample and we have not yet considered stars with  $R < 60$  kpc. Are there similar stars at  $R < 60$  kpc that support an argument for a significantly larger mass estimate? In Figure 6 we reprise our presentation of the distribution of H3 stars and highlight, in the shaded region, where stars need to be to result in timing mass estimates that are less than or equal to that of our most constraining star at  $R > 60$  kpc. We find only three stars in the figure that are outside the shaded region (there are a few more beyond the boundaries of the plot but those are very likely to be spurious). We conclude that there is no indication, even when considering stars at  $R < 60$  kpc, of a convincing population of stars that would support a mass that is significantly larger than that resulting from the most extreme star in our sample at  $R > 60$  kpc. Instead, we find a number of stars near the boundary of the shaded region at both positive



**Figure 6.** The timing argument and  $M_{200} = 1.1 \times 10^{12} M_{\odot}$ . We compare the distribution of stars in this phase space, reproducing Figure 1, to the region that produces timing mass estimates equal to or smaller than that produced by our most constraining H3 star at  $R > 60$  kpc (shaded region). The shaded region traces faithfully the distribution of stars at  $R < 60$  kpc and leaves little doubt that the few stars outside it are some type of interloper rather than indicative of a much larger Milky Way mass.

and negative  $v_R^{\text{GSR}}$ , providing support for the timing mass estimate derived previously using the one star.

This figure raises the question of whether we should be fitting to data at all radii, rather than just to outer halo stars. There are, however, a few complications in working with stars at smaller radii. First, the steepness of the caustic curve at small radius means that the mass estimation is much more sensitive to distance errors than it is at the larger distances we have worked at. Additionally, the radial velocities are also more dependent on the proper motion measurements than for stars at large distance, where the line-of-sight velocity and the radial velocity in the Galactic frame are closer to being parallel. Lastly, working at small radii, there are greater concerns regarding the impact of perturbations. All of this does not lead us to conclude that these data are not useful, but rather that they require a more complete treatment, including perhaps a closer examination of numerical simulations of halo stars, which is beyond the current study.

## 5. Summary

We propose that investigators exploring increasingly more complex and sophisticated mass modeling of the Galaxy utilize simple and robust mass estimates to constrain the range of mass priors that they explore. In this particular study, we use the timing argument and the H3 spectroscopic catalog to present such a limit. We use mock catalogs constructed from published, independent simulations to test our methodology and validate our conclusions. Our derived, conservative, 90% confidence lower limit on the mass of the Milky Way,  $0.91 \times 10^{12} M_{\odot}$ , already helps eliminate much of the mass range that previous models were unable to exclude. Our analysis suggests a Milky Way mass,  $M_{200}$ , of  $\approx 1.5 \times 10^{12} M_{\odot}$ . The properties of stars at  $R < 60$  kpc qualitatively support this result, but

those stars become increasingly more difficult to analyze as  $R$  becomes smaller. Instead, we advocate for larger samples that may identify more stars at large  $R$ . Outer halo surveys are growing prodigiously. We expect a lower mass limit, such as that produced here, to continue to gain in confidence and be used generally to refine complex dynamical modeling.

We thank the Hectochelle operators Chun Ly, ShiAnne Kattner, Perry Berlind, and Mike Calkins, and the CfA and U. Arizona TACs for their continued support of this long-term program. This research was supported in part by the National Science Foundation under grant No. NSF PHY-1748958 and AST-1812461. D.Z. acknowledges the hospitality of DARK, University of Copenhagen where the majority of this paper was written.

*Facility:* MMT (Hectochelle).

## ORCID iDs

Dennis Zaritsky <https://orcid.org/0000-0002-5177-727X>  
 Charlie Conroy <https://orcid.org/0000-0002-1590-8551>  
 Huanian Zhang <https://orcid.org/0000-0002-0123-9246>  
 Rohan P. Naidu <https://orcid.org/0000-0003-3997-5705>  
 Ana Bonaca <https://orcid.org/0000-0002-7846-9787>  
 Nelson Caldwell <https://orcid.org/0000-0003-2352-3202>  
 Phillip A. Cargile <https://orcid.org/0000-0002-1617-8917>

## References

- Barber, C., Starkenburg, E., Navarro, J. F., McConnachie, A. W., & Fattahi, A. 2014, *MNRAS*, **437**, 959
- Bertschinger, E. 1985, *ApJS*, **58**, 39
- Boylan-Kolchin, M., Bullock, J. S., Sohn, S. T., Besla, G., & van der Marel, R. P. 2013, *ApJ*, **768**, 140
- Brown, W. R., Geller, M. J., & Kenyon, S. J. 2014, *ApJ*, **787**, 89
- Callingham, T. M., Cautun, M., Deason, A. J., et al. 2019, *MNRAS*, **484**, 5453
- Cargile, P. A., Conroy, C., Johnson, B. D., et al. 2019, arXiv:1907.07690
- Carlesi, E., Hoffman, Y., Sorce, J. G., & Gottlöber, S. 2017, *MNRAS*, **465**, 4886
- Cautun, M., Frenk, C. S., van de Weygaert, R., Hellwing, W. A., & Jones, B. J. T. 2014, *MNRAS*, **445**, 2049
- Conroy, C., Bonaca, A., Cargile, P., et al. 2019, *ApJ*, **883**, 107
- Dalton, G., Trager, S. C., Abrams, D. C., et al. 2012, *Proc. SPIE*, **8446**, 84460P
- Deason, A. J., Fattahi, A., Belokurov, V., et al. 2019, *MNRAS*, **485**, 3514
- Deng, L.-C., Newberg, H. J., Liu, C., et al. 2012, *RAA*, **12**, 735
- Diemer, B., Mansfield, P., Kravtsov, A. V., & More, S. 2017, *ApJ*, **843**, 140
- Dierickx, M. I. P., & Loeb, A. 2017, *ApJ*, **847**, 42
- Fardal, M. A., van der Marel, R. P., Law, D. R., et al. 2019, *MNRAS*, **483**, 4724
- Fillmore, J. A., & Goldreich, P. 1984, *ApJ*, **281**, 1
- Fragione, G., & Loeb, A. 2017, *NewA*, **55**, 32
- Gibbons, S. L. J., Belokurov, V., & Evans, N. W. 2014, *MNRAS*, **445**, 3788
- Grand, R. J. J., Deason, A. J., White, S. D. M., et al. 2019, *MNRAS*, **487**, L72
- Grand, R. J. J., Gómez, F. A., Marinacci, F., et al. 2017, *MNRAS*, **467**, 179
- Guiglion, G., Battistini, C., Bell, C. P. M., et al. 2019, *Msngr*, **175**, 17
- Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, *ApJS*, **208**, 19
- Hopkins, P. F., Wetzel, A., Kereš, D., et al. 2018, *MNRAS*, **480**, 800
- Kaffe, P. R., Sharma, S., Lewis, G. F., & Bland-Hawthorn, J. 2014, *ApJ*, **794**, 59
- Kahn, F. D., & Woltjer, L. 1959, *ApJ*, **130**, 705
- Kroeker, T. L., & Carlberg, R. G. 1991, *ApJ*, **376**, 1
- Li, Y.-S., & White, S. D. M. 2008, *MNRAS*, **384**, 1459
- Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., et al. 2017, *AJ*, **154**, 94
- Martell, S. L., Sharma, S., Buder, S., et al. 2017, *MNRAS*, **465**, 3203
- Mateo, M., Olszewski, E. W., & Walker, M. G. 2008, *ApJ*, **675**, 201
- Patel, E., Besla, G., & Mandel, K. 2017, *MNRAS*, **468**, 3428
- Patel, E., Besla, G., Mandel, K., & Sohn, S. T. 2018, *ApJ*, **857**, 78
- Phelps, S., Nusser, A., & Desjacques, V. 2013, *ApJ*, **775**, 102
- Piffl, T., Scannapieco, C., Binney, J., et al. 2014, *A&A*, **562**, A91

- Posti, L., & Helmi, A. 2019, [A&A](#), **621**, [A56](#)
- Sales, L. V., Navarro, J. F., Abadi, M. G., & Steinmetz, M. 2007, [MNRAS](#), **379**, [1475](#)
- Sandage, A. 1986, [ApJ](#), **307**, [1](#)
- Sanderson, R. E., Wetzel, A., Loebman, S., et al. 2018, [arXiv:1806.10564](#)
- Watkins, L. L., Evans, N. W., & An, J. H. 2010, [MNRAS](#), **406**, [264](#)
- Wetzel, A. R., Hopkins, P. F., Kim, J.-h., et al. 2016, [ApJL](#), **827**, [L23](#)
- White, S. D. M., & Zaritsky, D. 1992, [ApJ](#), **394**, [1](#)
- Williams, A. A., Belokurov, V., Casey, A. R., & Evans, N. W. 2017, [MNRAS](#), **468**, [2359](#)
- Xue, X. X., Rix, H. W., Zhao, G., et al. 2008, [ApJ](#), **684**, [1143](#)
- Zaritsky, D., & Courtois, H. 2017, [MNRAS](#), **465**, [3724](#)
- Zaritsky, D., Olszewski, E. W., Schommer, R. A., Peterson, R. C., & Aaronson, M. 1989, [ApJ](#), **345**, [759](#)