Draft version November 22, 2015
Preprint typeset using LaTeX style emulateapj v. 12/16/11

# THE SIGNIFICANCE OF THE LMC/SMC FOR UNDERSTANDING THE MILKY WAY SATELLITES

Yu Lu[1], et al.

*Draft version November 22, 2015*

## ABSTRACT

The Large and Small Magellan Clouds (LMC/SMC) are significant companions of the Milky Way galaxy and provide useful information about the formation of the Milky Way. By analyzing a suite of high-resolution $N$-body simulations of Milky Way-sized host halos and semi-analytic model predictions, we find that the existence of the LMC/SMC-like companions strongly correlates with the formation history and the concentration of the host halo. We explore the parameter space of a flexible semi-analytic galaxy formation model with Markov-Chain Monte-Carlo and find that ignoring the prior knowledge about the host properties that are correlated with the existence of the LMC/SMC can lead to a biased inference for galaxy formation physics. In particular, when the galaxy formation model is calibrated to reproduce the observed stellar mass for Milky Way satellite galaxies, different host halos are found to require different levels of feedback strength, leaving imprints on the mass-metallicity relation of dwarf galaxies. Moreover, using MCMC, we rigorously show that the general wisdom of strong outflow with evenly fixed metals cannot explain the stellar mass function and the mass-metallicity relation of Milky Way satellites. The result suggests that two different physical mechanisms are needed to explain the mass function and the metallicity relation.

## 1. INTRODUCTION

Our own galaxy, the Milky Way galaxy, provides an excellent laboratory for constraining cosmology and galaxy formation physics. These studies require accurate model for the formation of galaxies and the host halo to make meaningful comparisons between theory and observation. Conventionally, a Milky Way host halo is defined by a given present-day mass, but the importance of properties other than mass when selecting Milky Way (MW) analogues has been realized in recent studies (e.g Busha et al. 2011b). For instance, it has been shown that the concentration of the MW halo can change the statistical significance when matching simulated MW and the observed (Mao et al. 2015). We investigate in this paper the implications of the existence of high-mass satellite galaxies, the Large and Small Magellanic Clouds (LMC/SMC), for understand the formation of the MW system.

The LMC and SMC are significant members of the satellite galaxies of the MW. Both of them are measured to have maximum circular velocities $V_{\rm max} \sim 60{\rm kms}^{-1}$ with magnitudes $M_{\rm v} = -18.5$ and -17.1, respectively (van der Marel et al. 2002; Stanimirović et al. 2004; van den Bergh 2000). The next brightest satellite is Sagittarius, about 4 magnitudes dimmer, with $V_{\rm max} \sim 20{\rm kms}^{-1}$ (Strigari et al. 2007). In terms of stellar mass, the LMC and SMC have significantly higher stellar masses than the rest of the MW satellite galaxy population. The stellar mass of the LMC is about $10^{8.9}{\rm M}_{\odot}$, and the stellar mass of the SMC is about $10^{8.2}{\rm M}_{\odot}$, which are 1.4 and 0.7 dex higher than the stellar mass of Sagittarius. This difference is often referred as MW satellite galaxy mass gap (e.g. Jiang & van den Bosch 2015). For the rest of the satellite population, the differences between two sequential galaxies are no larger than 0.5 dex (the 5th and the 6th). These distinct features make the LMC/SMC special in the MW satellite population. The significance of the the LMC and SMC as major satellites of the MM and their implications have been discussed in the literature from both theoretical and observational point of view by many authors.

Busha et al. (2011b) have studied the likelihood that MW-like systems host massive satellite galaxies. They assumed a certain host halo mass for the MW and applied a sub halo abundance matching model (SHAM) to study the statistics for the MW to host LMC/SMC. They found that MW-like objects in their selection had a 5%11% chance to host two subhalos as large or as luminous as the SMC with relevant properties, including their circular velocity, distance from the center of the MW, and velocity within the MW halo, in agreement with previous simulation results (e.g. Boylan-Kolchin et al. 2010). In Busha et al. (2011a), the authors assumed a satellite population and asked what the implications were for the properties of the host halo, including its mass. They constrained the MW virial mass, $1.2^{+0.7}_{-0.4}$ (stat.)$\pm$0.3 (sys.) $\times 10^{12}{\rm M}_{\odot}$ (68% confidence). In addition, they also calculated the probability distribution for the density profile of the MW and its satellite accretion history, and found that although typical satellites of $10^{12}{\rm M}_{\odot}$ halos are accreted over a wide range of epochs over the last 10 Gyr, a $\sim$72% probability that the LMC and SMC were accreted within the last Gyr, and a 50% probability that they were accreted together. In recent studies, Mao et al. (2015) and Wang et al. (2015) demonstrated the need to consider the dependence of satellite occupation on concentration when inferring the mass or other properties of the MW halo from satellites. These studies, however, are limited to the dark matter skeleton of the MW system.

The baryonic processes of galaxy formation are known to have significant impact on the satellite galaxy population. A number of authors have adopted various approaches to model the formation of MW satellite galaxies. Koposov et al. (2009) used a number of toy models

[1] The Observatories, The Carnegie Institution for Science, 813 Santa Barbara Street, Pasadena, CA 91101, USA

to add galaxy properties to dark matter halos generated using a combination of Press-Schechter theory with semi-analytic models for tracking subhalo orbits (**?**). They found that it was very difficult to model objects as bright as LMC and SMC without allowing for an extremely high star formation efficiency. This finding extended the underprediction of high-mass subhalos to an underprediction of luminous satellites, implying that the MW system is unusual owing to the existence of LMC/SMC. Similarly, Okamoto et al. (2010) explored a range of feedback models to add galaxies to some of the high-resolution Aquarius halos (Springel et al. 2008), and found again difficult to readily reproduce halos with luminosities as bright as the MCs when the luminosities of lower mass satellite galaxies are matched.

While this difficulty could stem from wrong physics being considered in current galaxy formation model, it is possible that influence of the underlying dark matter structure causes the issue. As the backbone of galaxy formation, the dark matter halo and its associated subhalos impose important constraints on the properties of galaxies that form in them. It has widely acknowledged that the formation history, number of substrctures, environment and so on can affect galaxy formation. To accurately model the formation of a particular galaxy, for instance the Milky Way, requires prior knowledge about the host dark matter halo. This paper is dedicated to investigating the host halo prior. We try to gain insight on what halo properties are of interests when modeling the formation of MW satellite galaxies. To archive this, study a set of $N$-body simulations of MW-sized halos, and apply a semi-analytic model on the merger trees extracted from those simulations to study the galaxy properties.

Moreover, one has to admit that current galaxy formation models are premature. Even for a well defined galaxy formation model, large uncertainties still exist, and need to be constrained using observational data. One needs to exhaustively explore the parameter space to identify "successful" models to interpret the physics of galaxy formation in a meaningful way. We archive our goals to understand how halo prior influence inferences of galaxy formation physics, we employ a MCMC machinery that is joint with our SAM to explore the parameter space of galaxy formation. We investigate the effect of halo prior and gain insight into what aspects of modeling and observational work are needed to further tighten the constraints on the formation of MW satellite galaxies.

Last but not the least, metallicities of galaxies are expected to put interesting constraints on feedback models (Peeples & Shankar 2011; Lu et al. 2015c,a; Ma et al. 2015). Low-mass galaxies, for instance MW satellite galaxies, are best places to test feedback models, because feedback are thought to be most effective in these system for their shallow gravitational potential. In this paper, we also explore the parameter space of the semi-analytic model to understand the constraining power of the metallicity relation of MW satellites.

In this paper, we describe the simulations and the modification of the SAM we use in §2. The results we obtain in this study are presented in §3. We summarize the conclusions of the study and discuss the implications in §4.

## 2. METHODOLOGY

### 2.1. *The simulations*

In this study, we use two sets of $N$-body simulations, a cosmological simulation of c125-2048 (Matthew R. Becker in prep) and a suite of high-resolution zoom-in simulations of halos with similar final masses (Mao et al. 2015). The c125-2048 simulation is a dark matter-only cosmological simulation run with $2048^3$ particles and a side length of $125\mathrm{Mpc}h^{-1}$, particle mass of $1.8 \times 10^7\mathrm{M}_\odot h^{-1}$ . The softening length is $0.5\mathrm{Kpc}h^{-1}$, constant in comoving length. The cosmological parameters are $\Omega_\mathrm{M} = 0.286$, $\Omega_\Lambda = 0.714$, $h = 0.7$, $\sigma_8 = 0.82$, and $n_s = 0.96$. We select all the halos that have virial mass $M_\mathrm{vir} = 10^{12.1} \pm 0.03\mathrm{M}_\odot$ at $z = 0$, which yields a large sample $\sim 1300$ halos. The zoom-in simulations consist of 46 halos with final mass in the same mass range selected from a lower resolution sister simulation c125-1024. The parameters and initial conditions of these two boxes are identical, but c125-1024 contains $1024^3$ particles and starts at a different redshift. The mass of the high-resolution particles in the zoom-in simulations is $3.0 \times 10^5\mathrm{M}_\odot h^{-1}$. The softening length in the highest-resolution region is $170\mathrm{pc}h^{-1}$ comoving. Out of the 46 halos, we adopt 38 of them for the studies in this paper and discard 8 others because they include very large lagrangian volume making the SAM calculation practically intractable. For detailed description of these simulations, including halo identification and merger tree construction, readers are referred to Mao et al. (2015).

With these simulated dark matter halos, we first study the mass accretion histories (MAHs) of MW halos. The MAH is defined as the virial mass of the main-branch halo as a function of redshift or time (Springel et al. 2005). Figure 1 shows the MAH of three randomly selected host halos from the simulations. For each halo in the zoom-in simulation, we identify the same halo from the c125-2048 simulation by matching the locations. The figure shows that the host MAHs from the two simulations with different mass resolutions are very similar within the mass resolution limit of the c125-2048 simulation. Tasitsiomi et al. (2004) proposed a 2-parameter model, which combines the exponential model (Wechsler et al. 2002) and the power-law model van den Bosch (2002), to describe a halo MAH. The model is

$$M(z) = M_0 \left(\frac{1+z}{1+z_0}\right)^{-\beta} \exp\left[-\gamma(z-z_0)\right]. \qquad (1)$$

Using this model, McBride et al. (2009) fit the MAHs of halos in the Millennium simulation in a large halo mass range, and found that the 2-parameter model remarkably well captures the halo MAHs. Following the previous work, we also fit the MAHs of the MW halos in our simulations using this model. We overplot the best fit model in Figure 1 for the three examples. As one can see, the overall shape of the MAHs is well captured by the best fit model. We also show how the best-fit parameters for the MAHs from high-res and low-res simulations compare in Figure2. The fitting parameters are in general agreement. We point out that a merit of this fitting model is that the model parameters characterize the accretion rate of a halo at early and late epochs. In

this model, we find that

$$\frac{\mathrm{d}\log M}{\mathrm{d}z} = -\left(\frac{\beta}{1+z}+\gamma\right), \quad (2)$$

and

$$\frac{\mathrm{d}\log M}{\mathrm{d}\log a} = \beta+\gamma(1+z). \quad (3)$$

The derivation indicates that at the present time, when $z$=0 and $a=1$, the accretion rate $\mathrm{d}\log M/\mathrm{d}\log a$ is $\beta+\gamma$, and at early times, the accretion rate is more characterized by parameter $\gamma$. Motivated by these indications, we will use $\beta+\gamma$ and $\gamma$ extracted from the best-fit model to characterize halo accretion rate at late and early times for each halo, and to study the correlation between the MAH and the satellite population.

We also measure the concentration parameter of the simulated host halos. The median concentration of the large low-resolution sample is $c_{\rm median}=11.5$. The concentrations predicted for the counterparts in the high-resolution and the low-resolution simulations do not always agree, especially the low-concentration ones. Halos with low concentration in the low-resolution run tend to show large deviation in their concentration parameter. This is different from the MAH fitting parameters, which seem to agree much better between the two sets of the simulations. We will use both observations in this study, taking advantage of the large number of halos from the low-res simulation for better statistics and the high mass resolution from the zoom-in simulation.

### 2.2. *The SAM and an approximation scheme for predicting the satellite population*

To study the baryonic component of MW satellites, we adopt the SAM developed by Lu et al. (2011, 2014b) and apply the model to the merger trees extracted from the high-res zoom-in simulations. Because the high-resolution simulations have very high mass and time resolution, the merger trees have a large number of branches that end up with a satellite galaxy below mass range interested in this study. Therefore, following the full tree in a SAM calculation is inefficient for this study and prohibits efficient parameter space exploration. To allow exhaustive parameter space exploration, we adopt a scheme in the SAM to speed up the SAM calculation by ignoring satellite galaxies that are significantly below the lower mass limit of interests. In the scheme, we focus the computation only on the subhalos that are still present at $z=0$, but ignore the sub halos that are accreted into the final host too early to survive until the present day. The high mass resolution of the simulation guarantees that the stellar mass of the satellite galaxies produced in the ignored branches would be much lower than our interests in a sensible model. We note that subhalos that are accreted into other branches are not ignored in this scheme even though they have been disrupted by the present time, because they can affect the evolution of the satellite galaxies we are interested in at $z=0$ In Figure 3, we compare the stellar masses of the satellite galaxies predicted using the new scheme and the full SAM. We find that only 4% of the satellites down to $M_*=10^4\mathrm{M}_\odot$ are predicted with a deviation in the stellar mass larger than 2%. We conclude that the new scheme accurately reproduces the stellar masses of all satellites in the mass range we care about. We therefore adopt the approximate scheme in this study. This scheme typically reduces the computation time by a factor of 4 for each MW merger tree. The benefit in speeding up the calculation allows us to use MCMC to sample the posterior distribution of the model under data constraints of the MW satellite stellar mass function.

### 2.3. *Likelihood function for Milky Way satellite galaxy stellar mass function*

To conduct a Bayesian inference, we need to write down the likelihood function of the model. The data we use in this paper is the stellar mass function of the MW satellite galaxies, down to the $11^{th}$ most massive satellite galaxies ($M_*=2.9\times10^5\mathrm{M}_\odot$). The incompleteness of MW satellite counts down to this mass is negligible (REF). The theoretical prediction for the mass function is straightforward, but the likelihood function of the satellite mass function is unknown. Poisson distribution is a model one considers, but Boylan-Kolchin et al. (2010) found that Poisson distribution does not accurately describe the distribution function of subhalo mass function predicted by $N$-body simulation. Similar findings have also been reported by other authors (e.g. Busha et al. 2011a; Cautun et al. 2014). In a recent paper, Mao et al. (2015) argued that the deviation from Poisson can be largely attributed to the effect of varying concentration for a fixed halo mass. To choose an accurate model for the distribution function of the satellite galaxies predicted by a model, we apply a fiducial SAM to the merger tree set. Taking the advantage of large sample of the low-res MW set, we extract the distribution of the predicted mass function for modeling the likelihood function.

The fiducial model we use here was tuned to match the field galaxy stellar mass function of the local Universe (Lu et al. 2014b). We apply the fiducial model to the low-resolution merger trees to make a prediction for the MW satellite population. For each merger tree, we extract the subhalo mass function from the $N$-body simulation itself, and predict the satellite stellar mass function using the fiducial SAM. In Figure 4, we show the subhalo mass function on the left panel and the SAM predicted satellite stellar mass function on the right panel. We show the mass functions down to the $10^{th}$ most massive subhalo (satellite galaxy in right panel). The blue box extends horizontally from the lower to upper quartile values of the mass of $i^{\rm th}$ subhalo (satellite), with a middle line indicating the median of the mass distribution. The whiskers extend from the box to show the range of the mass distribution. The gray shaded region covers the 1-$\sigma$ range of the accumulative distribution function $N(>M)$ as a function of $\log M$. In comparison, we also show the expected 1-$\sigma$ scatter of the mass function assuming it is Poisson distributed with red vertical lines with error bars. The mismatch between the expected Poisson error bar and the simulation predicted 1-$\sigma$ region (grey region) indicates that the real distribution of the mass functions is deviated from Poisson. Moreover, we also overplot the observed MW satellites in the right panel. Without further tuning the model against the data, the fiducial model prediction is in a good agreement with the data, with all the data points covered by the $1-\sigma$ region

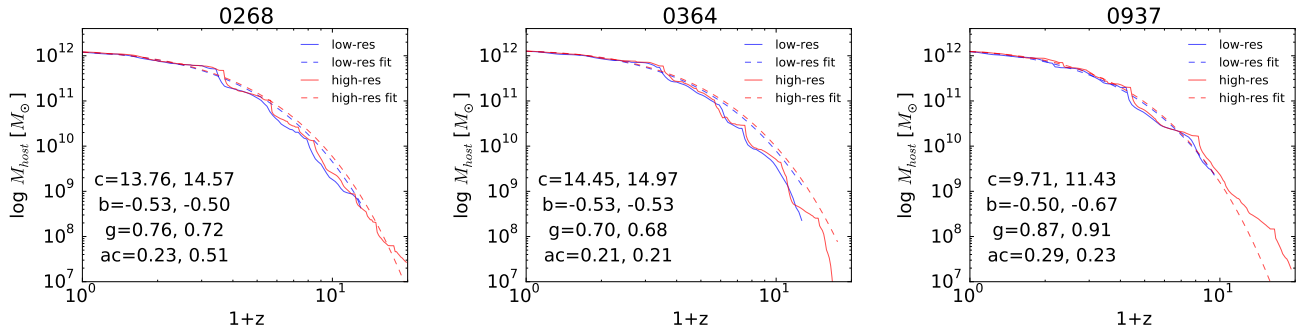


FIG. 1.— The halo mass accretion histories of three typical host halos. The blue solid line denotes the MAH predicted by the low-res c125-2048 simulation, and the solid red line denotes the MAH of the same halo predicted by the high-res zoom-in simulation. The dashed lines with same color denote the corresponding fitting model using Eq.1. The halo concentration and fitting parameters are listed in each panel. The first column is for the low-res halo, and the second column is for the high resolution counterpart.

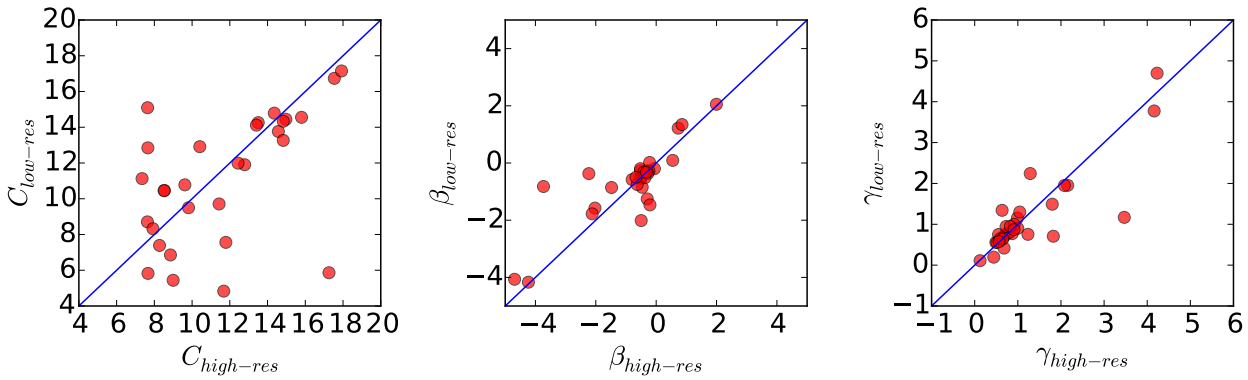


FIG. 2.— We compare the parameters low-res and high-res host halos. Left panel: the concentration parameter predicted in the high-res simulation and in the low-res simulation for same halos. Middle panel and right panel shows the comparisons of the fitting parameters for the host halo mass accretion history.

of the predicted distribution.

The probability distribution function of the number count of subhalo is found to be better described by the Negative Binomial Distribution (NBD) (Boylan-Kolchin et al. 2010), which has also been used to describe the number of satellite galaxies in HOD models (e.g. Berlind & Weinberg 2002),

$$P(N|r,p) = \frac{\Gamma(N+r)}{\Gamma(r)\Gamma(N+1)} p^r (1-p)^N \,, \qquad (4)$$

where $N$ is the number of subhalos per host, $\Gamma(x) = (x-1)!$ is the Gamma function, and $r$ and $p$ are two parameters in the NBD. Boylan-Kolchin et al. (2010) found that the distribute deviates from Poisson because of a small fraction of intrinsic scatter, and defined a parameter $s_{\rm I} = \frac{\sigma_{\rm I}}{\mu}$, the fractional scatter from the intrinsic scatter, $\sigma_{\rm I}$, with respect to the Poisson scatter, $\mu$. The two parameters in the NBD are then determined as

$$p = \frac{1}{1+s_{\rm I}^2\mu} \,, r = \frac{1}{s_{\rm I}^2}. \qquad (5)$$

In Figure 5, we show the distribution of subhalo count of MW mass hosts in three different subhalo mass bins. The range covered by the mass bin is noted in each panel. In agreement with previous studies, we also find that the NBD describes the subhalo number count distribution much better than the Poisson distribution. Boylan-Kolchin et al. (2010) found that $s_{\rm I} \approx 0.18$ yields a good fit to the subhalo mass function distribution predicted by the Millennium-II Simulation. We find, however, a larger value for $s_{\rm I}$ is needed to describe the distribution. In the figure, we use $S_{\rm I} = 0.26$. When the expected number counts becomes smaller ($\bar{N} < 4$), the NBD approaches to the Poisson distribution, making it hard to distinguish between the two models. We also show the satellite galaxy number count distribution at three different stellar masses. Same to the subhalo mass function distribution, the number count distribution for given stellar mass bin can also be accurately described by the NBD. Busha et al. (2011a) found that adding an exponential tail to the NBD can better capture the distribution at very high $N$ for SHAM galaxies. Because the tail only covers a small fraction of the probability distribution, we ignore this part for keeping the model simple without losing accuracy above the level our MCMC can capture.

Using the fiducial SAM, we make prediction for the satellite galaxy mass function for each MW host. We show the distribution of the number of satellite galaxies in three stellar mass bins in the lower panel of Figure5. Again, we use $S_{\rm I} = 0.26$ for the NBD plotted in the figure

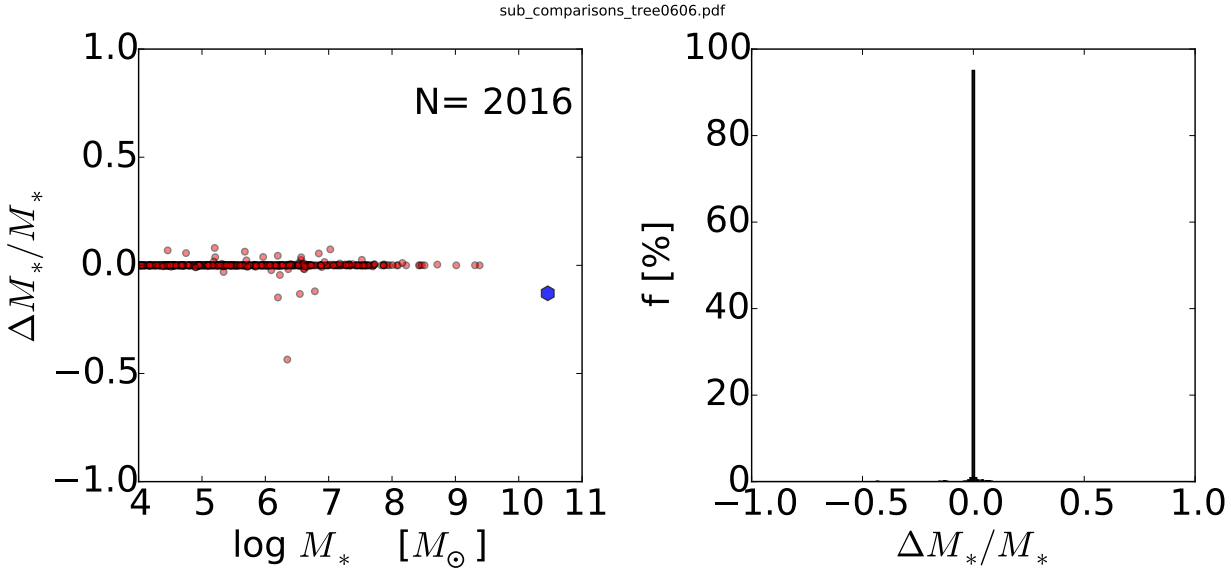


FIG. 3.— The error in predicted stellar mass of satellite galaxies introduced by the approximate treatment as a function of stellar mass for a typical high-res halo. Left: the x-axis is the stellar mass predicted by the approximate model, and y-axis denotes the fractional difference between the stellar masses predicted by the approximate model and the full SAM. The blue polygon denotes the central galaxy, and the red circles denote the satellites in the host halo. Right: the distribution of the fractional error of the predicted stellar mass. More than 95% satellite with stellar mass larger than $10^4 M_\odot$ have stellar mass deviated from the full SAM predicted value by less than 2%.

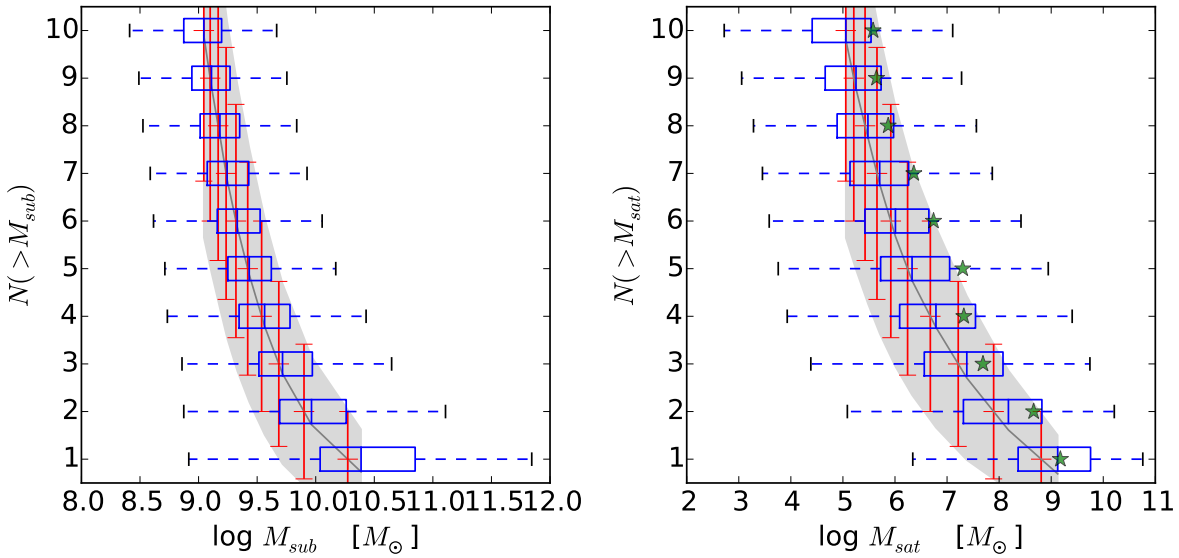


FIG. 4.— The left shows the sub-halo mass function of the MW host halos in the c125-2048 simulation. The right shows the satellite stellar mass function predicted by the fiducial SAM using the merger trees of the MW host halos in the c125-2048 simulation. In each panel, the horizontal whiskers show the distribution of the mass of the subhalo mass (or stellar mass) of the $i$th most massive subhalo (or satellite). The size of the box shows the quartiles of the distribution. The vertical bar in the middle of the box marks the median of the distribution in mass. The outer bars connected by dashed lines extend to the most extreme values. The red error bars show the standard deviation in the mass functions assuming the distribution of the mass functions follows a Poisson distribution. The gray shaded region shows the 1-$\sigma$ range of the simulation (SAM) predicted mass function function. The dark gray line in the middle of the shaded region denotes the median of the distribution. In the satellite mass function panel (right), the green stars show the observational data of the stellar masses of the most massive MW satellite in stellar mass.

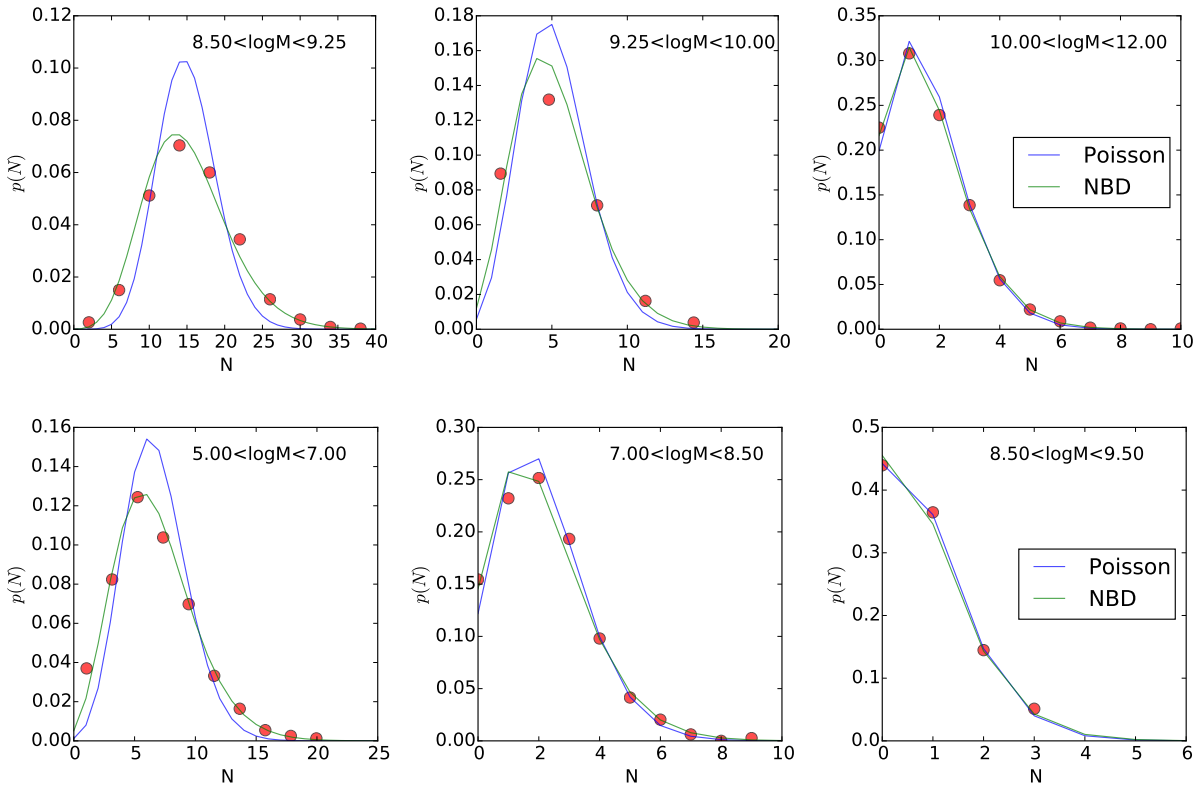


FIG. 5.— Upper: the distribution of the number of subhalos in three mass bins. Lower: the distribution of the number of satellites in three stellar mass bins. In each panel, the red circles denote the distribution of subhalo mass function (satellite stellar mass function) predicted by the low-res MW host simulations (and SAM applied to the merger trees). The blue line and the green lines show the Poisson distribution and negative binomial distribution with the same mean as the simulation predicted, respectively. For the negative binomial distribution, we have assumed $s_{\rm I} = 0.26$.

to describe the distribution of the stellar mass number counts as well. Nevertheless, we can imagine that the value of the intrinsic scatter can depend on galaxy formation model. In principle, the distribution can be simulated each time if a large number of merger trees are adopted. Limited by computation, we take $s_{\rm I}$ as a nuisance parameter to be sampled by MCMC in our study. The posterior distribution of $s_{\rm I}$ will be marginalized in subsequent analysis.

We have studied the overall shape of the number count distribution of subhalos and satellite masses. It is interesting to test the distribution functions by looking at the higher order moments of the predicted distribution quantitatively. Because the likelihood function will be apply to the MCMC posterior sampler based on no more than 38 halos, the tests we do here are based on a sample size of 38 halos. Using the mass functions predicted using the 38 halos, we compute the 2-4th order moments of each halo mass (and stellar mass) bin of the predicted mass functions. We compare the SAM predicted moments with the Poisson moments and the NBD moments. The goal is to test if the theoretical distribution is inconsistent with what the model predicts based on finite number of samples at 2-4th order moments.

The way we test it is the following. We generated 10000 Monte Carlo mass function sets based on an assumed distribution function, Poisson or NBD. To mimic the largest sample we can model in this paper, each Monte Carl sample contains 38 mass functions. We use the mean values predicted by the 38 simulated halos and SAM to assign the expectation values for the Monte Carlo mass functions. For each 38 Monte Carlo sample, we compute the 2nd, 3rd, 4th moments, yielding 10000 realizations of the moments. We used these 10000 values to construct a reference distribution for each moment of a simple size of 38 hosts. In Figure 6 we show the distribution of the Monte Carlo subhalo mass function and satellite stellar mass function of sample size of 38 hosts assuming the samples are Poisson distributed. The upper row shows the distribution of the 2nd, 3rd, and 4th moments of the subhalo mass function, while the lower row shows the distribution of the moments of the satellite stellar mass function. In each panel, we subtract the theoretical value from the Monte Carlo values and show the distribution of the residual. For a given mass bin, the distribution is shown by the violin symbol with a red error bar encompasses the Monte Carlo sample values. The blue line denotes the $N$-body simulation prediction (or SAM prediction) of each moment with respect to the theoretical moments. Due to limited sample size (38 hosts in this case), the simulated mass functions are expected to have some deviation from the theoretical expectations, even if the underlying distribution is Poisson, but the deviation should not be at odd with respect to the reference distribution defined by the Monte Carlo samples. As one can see, however, the deviation is too large to be consistent with a Poisson distribution because it is not comfortably bounded the Monte Carlo samples, especially the low-mass bins. The Monte Carlo simulations suggests that the while the simulation predicted distribution is still consistent with Poisson at high-mass bins, it is clearly inconsistent with Poisson for low mass bins revealed by higher order moments. As we have discussed previously, when the expected number counts is low, the difference between the simulated distribution and a Poisson distribution is very small. The deviation is only visible for low-mass bins, as we visualized in Figure 5.

In comparison, we carry out the same test for the NBD. Figure 7 similarly shows the distribution of the moments but assuming that the distribution follows the NBD. As one can see, the higher order moments of the simulated mass functions are generally consistent with the assumed NBD for the adopted sample size. Supported by these tests, we adopt the NBD as the model for the likelihood function of the satellite stellar mass function. Note that we also sample the parameter $S_{\rm I}$ as a nuisance parameter using the MCMC but do not try to predict it using the small sample size of host halos.

## 3. RESULTS

### 3.1. *LMC/SMC and halo mass assembly history*

We now study what makes a MW mass host to have satellite galaxies as massive as the LMC and SMC. In Figure 8, we show the correlation between the mass of a satellite galaxy and its host subhalo property of 38 MW systems. We choose to look at the first massive satellite and the second massive satellite in the first and second columns. The x-axis denotes the SAM predicted stellar mass for the satellites. For a given galaxy formation model, the most massive satellite in each host halo has very different stellar mass. In some systems, the most massive satellite can have a stellar mass higher than $10^{10}{\rm M}_\odot$, while in some other systems, the most massive satellite has a very low mass, $\sim 10^5{\rm M}_\odot$. The question is then what causes the big difference.

In the first row of Figure 8, we show the correlation between the stellar mass of the satellites and their instantaneous subhalo mass. It is clear that the galaxy mass is correlated with the host subhalo mass with a substantial scatter. Higher mass subhalos tend to host a high mass satellite, but for a given stellar mass, the scatter of the host subhalo mass is about a dex. We track the merger history of the subhalos and find the halo mass at the time when it is first accreted into another halo. In the middle panel, we show the halo mass at accretion as a function of the final stellar mass of those satellites. The correlation between the two quantities is even stronger, with a smaller scatter. This suggests that the halo mass at accretion is a better predictor of the final stellar mass of massive satellites. The third row shows the effect of model on the stellar mass of massive satellites. We vary the model parameters by a large amount from the fiducial model, and call this test model M1. The parameter set is off from the fiducial model and is out side the 95% posterior of the model constrained by the field stellar mass function (Lu et al. 2014b). While the model predicts a different stellar mass for each subhalo, the rank order of the systems based on their massive satellite stellar mass does not generally change. What this means is that a system is predicted to host a high mass satellite galaxy by an arbitrary model will also host a relatively high mass satellite if the galaxy formation physics is changed systematically for all galaxies.

Given that the final stellar mass of massive satellite galaxies are correlated with instantaneous subhalo mass, we further explore what host halo properties are correlated with suhhalo mass. In Figure9, we show the distri-

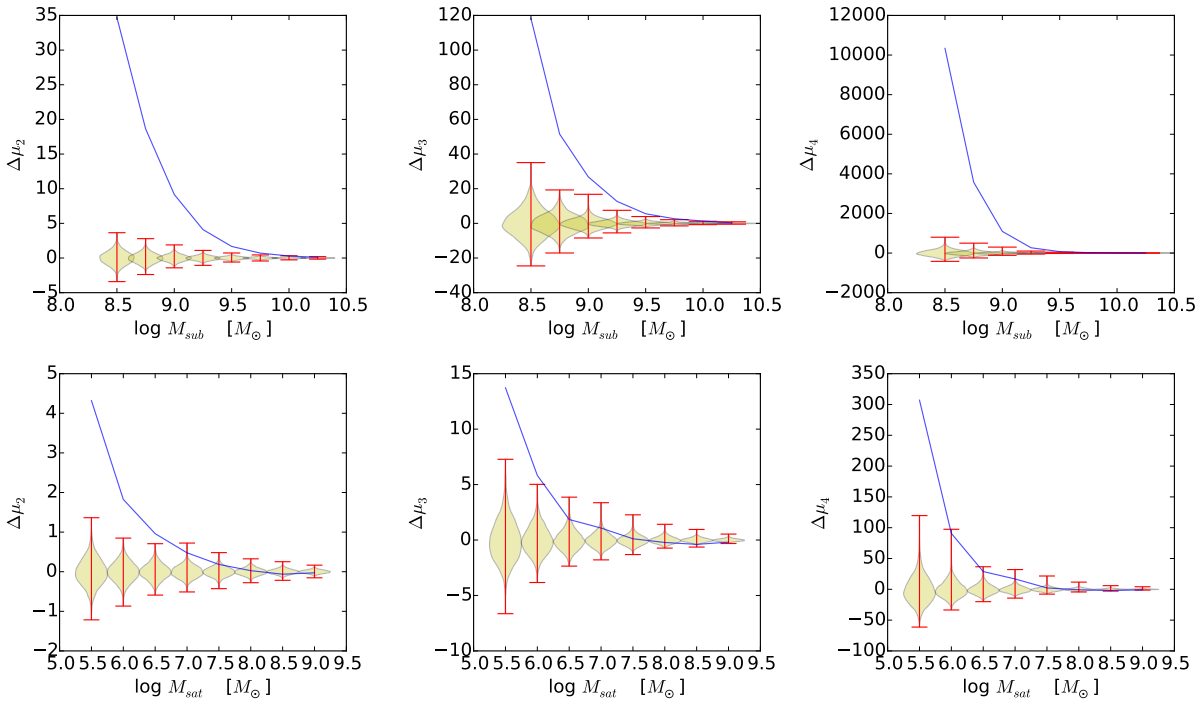


FIG. 6.— The 2nd, 3rd, and 4th moments of the distribution of subhalo and satellite mass functions as a function of mass relative to the Poisson moments. The blue line is the moments predicted out of the MW host predicted by the low-res simulation samples, subtract the expected moments assuming the distributions are Poisson. The yellow distributions are the reference distributions predicted by Monte Carlo simulation with the same number of samples. The simulation predicted mass functions have moments that are largely deviated from Poisson for low-mass bins. The 2nd, 3rd, and 4th moments of the distribution of subhalo and satellite mass functions as a function of mass relative to the negative binomial distribution moments. The blue line is the moments predicted out of the MW host predicted by the low-res simulation samples, subtract the expected moments assuming the distributions are negative binomial distribution. The yellow distributions are the reference distributions predicted by Monte Carlo simulation with the same number of samples. The simulation predicted mass functions have moments that are consistent with the negative binomial distribution for all masses.

bution of all the simulated MW host halos in the parameter space defined by the MAH fitting parameters, $\beta+\gamma$ and $\gamma$. All the halos populate a particular region of the parameter space with a concentration around $\beta+\gamma = 0.3$ and $\gamma = 0.7$. This distribution is in agreement with what is found in other simulations (McBride et al. 2009; Taylor 2011). Different from other authors, we choose to use $\beta+\gamma$ as one of the axises for its physical meaning. In the diagram, halos in the upper left branch have rapid late time accretion, while those in the lower right branch have rapid early time accretion. We also note that when $\beta = 0$, the fitting model is exactly an exponential form and reduces to the Wechsler et al. (2002) model. The model that is occupied by a large population of halos is not too far from an exponential form. These halos do not belong to the populations that have extremely high rate of accretion at late or early times among the entire population.

We fit the mass profile with a NFW model (Navarro et al. 1997) and use the size of the symbol to denote the characteristic core size of the halo. The size of the circle is linearly proportional to $r_s$ and inversely proportional to the concentration parameter $c = R_{\rm vir}/r_s$. One can find that the $r_s$ and concentration are closely correlated with the fitting parameters of the MAH. In general, halos have faster late-time accretion tend to have lower concentration (and larger $r_s$ for given $R_{\rm vir}$). This result is in agreement with findings of Wechsler et al. (2002); Bullock et al. (2001); Zhao et al. (2003) and explained by models (e.g. Lu et al. 2006).

We also color code each halo according to the mass of the most massive subhalo hosted. Interestingly, the mass of the most massive subhalo is also correlated with the MAH parameters and the concentration of the host. Hosts that have higher accretion rate at late times and lower concentration (larger $r_s$ for given $R_{\rm vir}$) tend to host high mass subhalos. Motivated by this finding, we plot the distribution of host halos in a diagram defined by halo concentration and $\beta+\gamma$ in the right panel of Figure9. Using the same color coding, we find a clear sequence that halos with higher $\beta+\gamma$, indicating a higher recent accretion rate, have lower concentration and tend to host a higher mass subhalo. We define a division line in the $c-\beta+\gamma$ diagram to split simulated hosts into two groups. Halos in the lower-right corner have higher concentration and slower recent accretion, and their most massive subhalo typically have higher masses. Halos in the upper-left have lower concentration and faster recent accretion, and do not tend to host a high mass subhalo. We use a division line, which is shown by a dashed in in the right-panel of Figure9, to split the halo samples into two groups. We name halos in the upper-left part of the diagram Group 1 halos, which are less concentrated, have more rapid recent accretion, and tend to host high-mass subhalos. We name halos in the lower-right part of the diagram Group 2 halos, which are more concentrated,

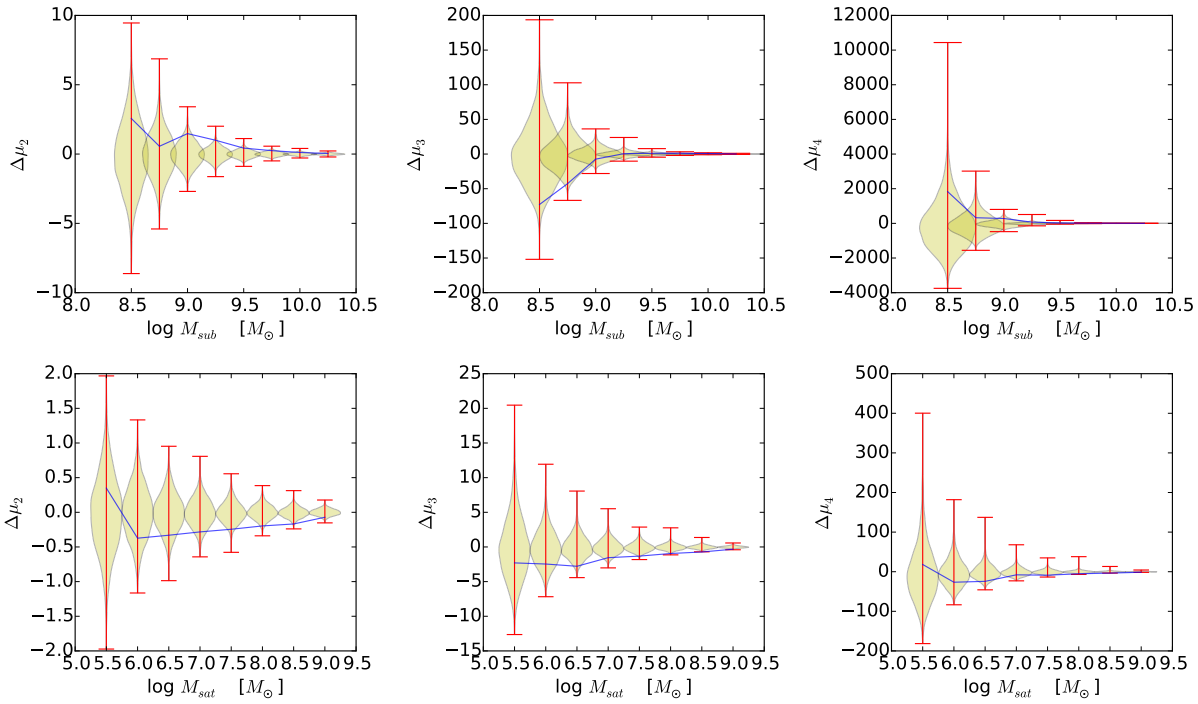


FIG. 7.— Same as Fig.6 but assuming the underlying distribution follows the NBD.

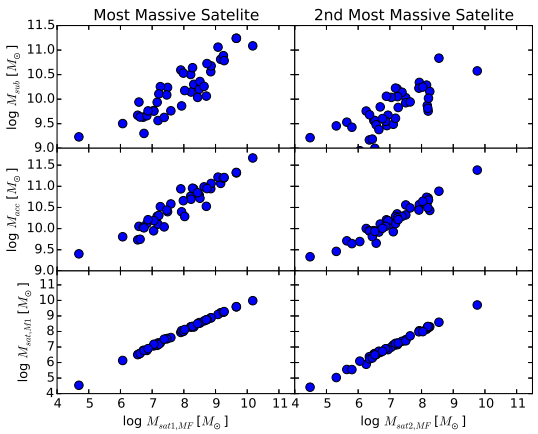


FIG. 8.—

have slower recent accretion, and do not tend to host high-mass subhalos. The division defined here, roughly split the entire high-resolution host halos into two subsamples with equal size, likewise the low-resolution host halos. For the high-resolution hosts, we have 20 halos in Group 1, and 18 halos in Group 2.

### 3.2. *Satellite galaxy stellar mass function*

We have classified simulated MW halos into two categories based on their concentration and formation history. While the division is arbitrarily defined, the whole high-resolution sample is split into two subsamples with roughly equal sizes. We stress that we do not have clear prior knowledge about which type of underlying dark matter halo the MW host halo belongs to, even though one can argue that the high stellar masses of LMC and SMC indicate high-subhalo masses, so the MW host is presumably Group 1 type. We do not have solid observational evidence to accurately constrain the subhalo mass of LMC and SMC. It is possible that the MW host halo is Group 2 type, and the dark matter mass of the LMC and SMC subhalos are relatively low. Theoretically, it is also uncertain to make prediction for the ratio between stellar mass and subhalo mass. Nevertheless, we expect that host halos with different subhalo population would require different star formation efficiency or feedback strengths to produce certain amount of stellar mass as observed. To test how the halo prior affects the inference of galaxy formation physics, we use the two halo groups as two explicit prior constraints to perform Bayesian inference from the MW satellite stellar mass function.

We adopt the stellar masses and memberships of MW satellite galaxies compiled in McConnachie (2012). We utilize the likelihood function we defined based on the NBD in Sec.2.3, and only include 11 most massive satellite galaxies into likelihood evaluation. We use MCMC to sample the posterior probability density distribution. Two separate runs are performed with Group 1 host halos and Group 2 host halos, respectively. In both runs, we allow six model parameters and one nuisance parameter to vary within their prior range. These model parameters are $\Sigma$, which is threshold gas surface density for star formation in units of $M_\odot pc^{-2}$, $\alpha_{LD}$ the normalization of mass-loading factor for feedback outflow, $\beta_{LD}$ the power index for the circular velocity dependence of the mass-loading factor, $V_{out}$ a characteristic halo circular velocity in km/s, below which all outflow mass leaves the host halo, $\beta_{out}$ the steepness of the transition from total outflow for halos with $V_c < V_{out}$ to no outflow for halos with $V_c > V_{out}$, and $\gamma_{RI}$ the rate for the outflow

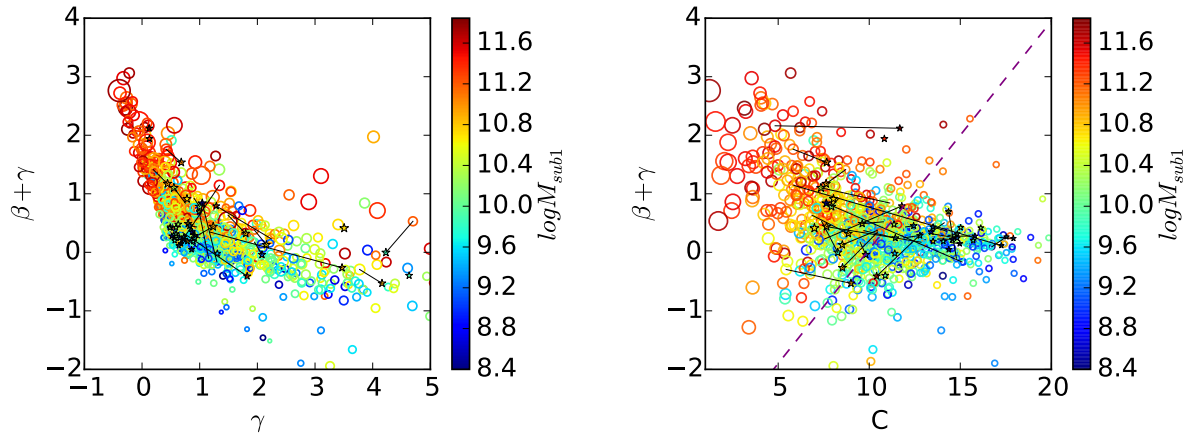


FIG. 9.— The distributions of the simulated MW hosts in parameter space. The left panel show the distribution of the host halo mass accretion history fitting parameters. The parameter $\gamma$ indicates the early accretion rate, and $\beta+\gamma$ indicates the late accretion rate. The distribution has a mode peaks at $\gamma = 0.8$, $\beta+\gamma = 0.3$. Each circle represents a MW host from the low-resolution c125-2048 simulation and each star represents a high-res MW host. The right panel shows the concentration parameter versus $\beta+\gamma$. The color represents the mass of the most massive subhalo hosted by each MW host halo, indicated in the color bar. The size of each symbol is proportional to the size of NFW core radius $r_s$ (or inversely proportional to the halo concentration parameter, $C$). The black solid lines in each panel connect the counterpart halos in the low-res and the high-res simulations. The purple dashed line in the right panel splits the distribution into two parts. Halos in the upper-left corner have lower concentration, and higher recent accretion rate, and their most massive subhalo typically has higher mass.

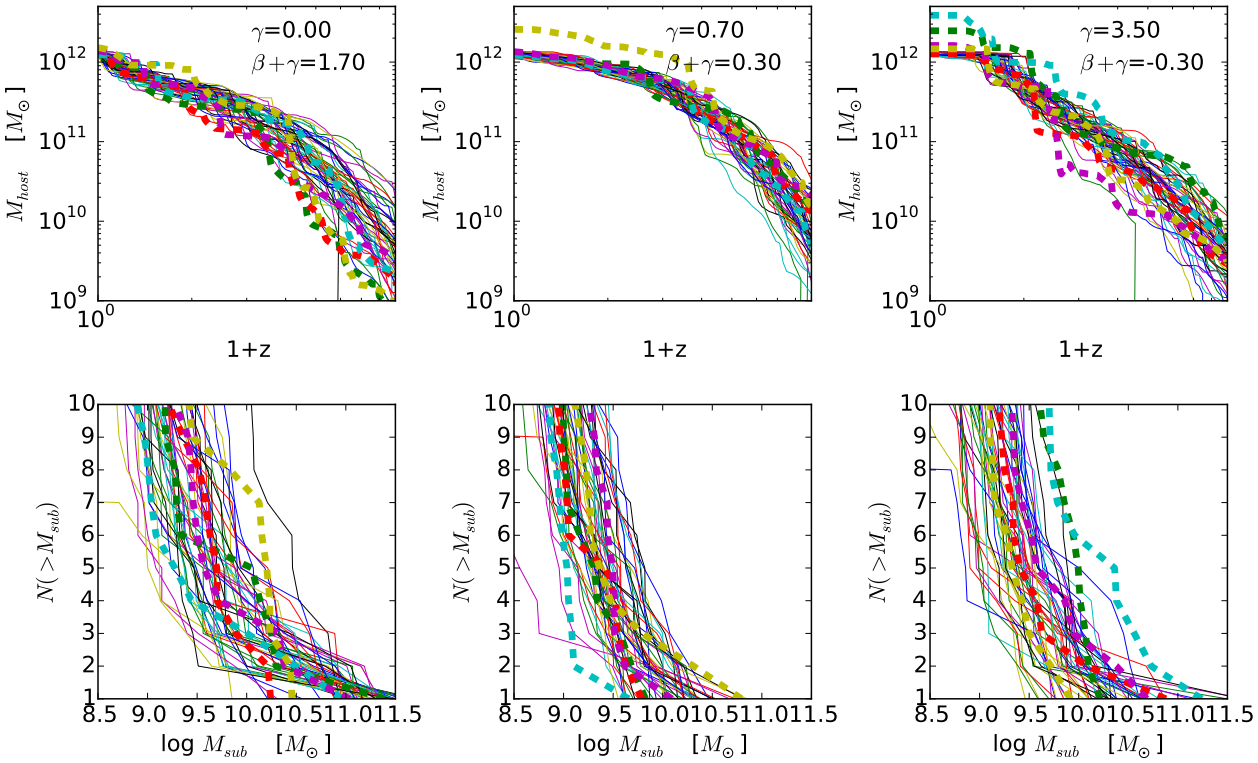


FIG. 10.— The host MAHs (upper) and the subhalo mass function (lower) predicted by the low-res simulation and the high-res resolution. The thin lines denote the mass functions of the low-res sample, and the thick dashed lines denote the high-res sample. We group host halos based on their fitting parameters for the mass accretion history and show each group in a column. From left to right $\{\gamma, \beta+\gamma\} = \{0, 1.7\}, \{0.7, 0.3\}$, and $\{3.5, -0.3\}$.

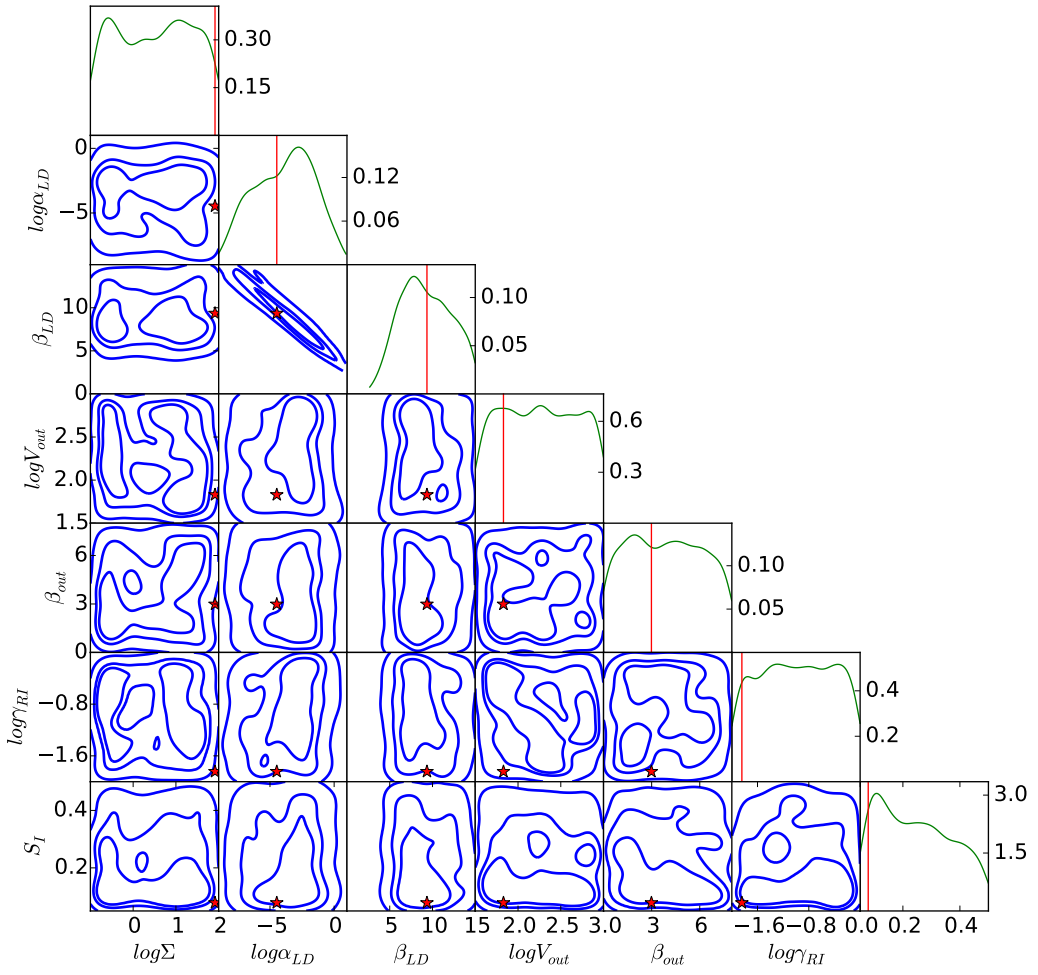


FIG. 11.— Posterior distribution of the SAM parameters constrained with MW satellite stellar mass function. Group 1 merger trees are used. The stars mark the position where the best model is.

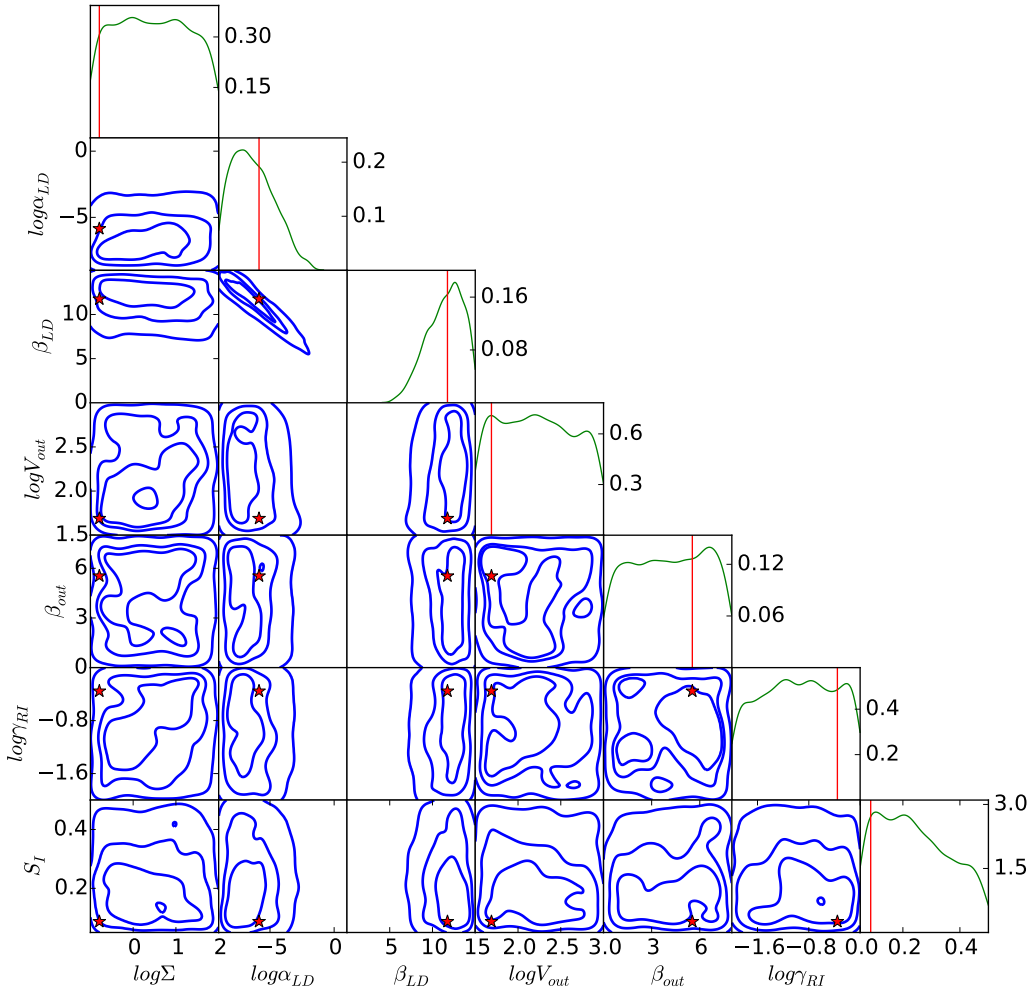


FIG. 12.— Posterior distribution of the SAM parameters constrained with MW satellite stellar mass function. Group 2 merger trees are used. The stars mark the position where the best model is.

mass to reincorporate back into the halo. For detailed explanation of these parameters, readers are referred to Lu et al. (2014b). We also include a nuisance parameter, $S_{\rm I}$, characterizing the intrinsic scatter of the predicted stellar mass function. The parameters, $\beta_{\rm LD}$, $\beta_{\rm out}$, and $S_{\rm I}$ are given uniform priors in linear scale, and $\Sigma$, $\alpha_{\rm LD}$, $\gamma_{\rm RI}$ are assigned uniform priors in logarithmic scale. The ranges of the priors are shown in Figure11. For each run, we run the MCMC for 20,000 iterations with 144 parallel chains using the differential evolution algorithm (Ter Braak 2006). The convergence test is done with the Gelman-Rubin test (Gelman & Rubin 1992), requiring $\hat{R} < 1.2$. After removing outliers and pre-turn-in states, we obtain 2,000,000 posterior samples from the MCMC. For these runs, the "best fit" model (maximum likelihood) is close to the median model of the full posterior.

We show the 2-D marginalized posterior distribution for the Group 1 run in Figure 11, and the posterior for the Group 2 run in Figure 12. As one can see, under the same data constraints, using different halo prior results in different the posterior in galaxy formation model parameters. An obvious change is that the normalization and the power index for the parameterization of the the mass-loading factor of outflow. In the model, the outflow mass-loading factor is parameterized as

$$\eta = \alpha_{\rm LD} \left( \frac{V_{\rm c}}{220{\rm km/s}} \right)^{-\beta_{\rm LD}}, \qquad (6)$$

where $V_{\rm c}$ is the circular velocity of a halo, $\alpha_{\rm LD}$ and $\beta_{\rm LD}$ are model parameters to be constrained. The normalization parameter $\alpha_{\rm LD}$ for Group 1 is higher than that for Group 2. Correspondingly, the power index for Group 1 is lower than that for Group 2. The normalization for the mass-loading factor is defined as the mass-loading factor for halos with a circular velocity $V_c = 220$km/s. This parameter directly represents the strength of SN feedback driven outflow. A higher $\alpha$ requires a higher fraction of

SN energy to power outflows to keep the baryon fraction of a halo low. The Group 1 halos are those tend to contain massive subhalos. The high-mass subhalos typically have higher baryon mass to start with before being accreted into the host. To keep the baryon mass fraction low in those high-mass subhalos, the model is required to have stronger feedback to suppress star formation. The Group 2 halos, which do not tend to have massive subhalos, typically require lower feedback to allow relatively higher stellar mass fraction to fit the satellite mass function.

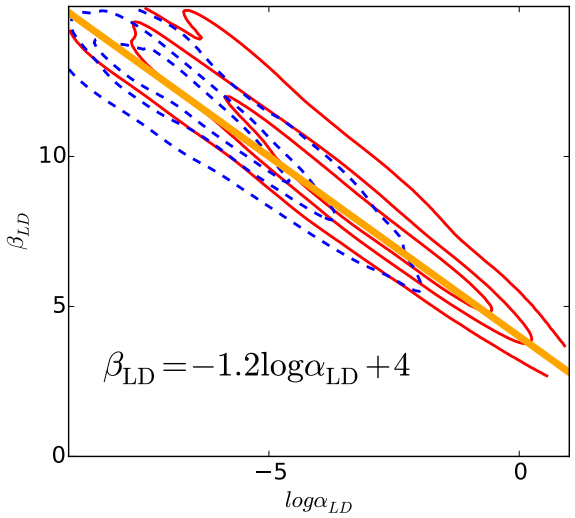


FIG. 13.— The red contours denote the posterior for the feedback outflow mass-loading parameters using the Group 1 halo prior. The blue contours denote the posterior for the same parameters of the model using the Group 2 halo prior. The solid linear line, $\beta_{\rm LD} = -1.2 \log \alpha_{\rm LD} + 4$ captures the degeneracy between the two model parameters in both runs.

It is clear that the two parameters for outflow mass-loading factor are strongly degenerate, regardless which halo prior is adopted. The degeneracy can be described by a linear function as $\beta_{\rm LD} = A \log \alpha_{\rm LD} + B$, where $A$ and $B$ are the slope and the intercept of the linear function. The straight line in Figure 13 captures the slope and intercept, with $A = -1.2$ and $B = 4$. Recall that the outflow mass-loading factor is parameterized as

$$\eta = \alpha_{\rm LD} \left( \frac{V_{\rm c}}{220 {\rm km/s}} \right)^{-\beta_{\rm LD}}. \quad (7)$$

One finds that

$$\log \eta = \log \alpha_{\rm LD} - \beta \log V_{220}, \quad (8)$$

where $V_{220}$ is halo circular velocity in unit of 220km/s. The linear function capturing the degeneracy can be rewritten as

$$-\frac{B}{A} = \log \alpha - \beta \frac{1}{A}. \quad (9)$$

Comparing Eq.8 and Eq.9, we find that in order to have any combination of $\alpha_{\rm LD}$ and $\beta_{\rm LD}$ to exactly follow a linear function defined by fixed A and B, the following condition needs to be fulfilled. The condition is when $\log V_{220} = \frac{1}{A}$, $\log \eta = -\frac{B}{A}$. For degeneracy found in this work, we find that the outflow mass-loading factor models have $\eta \approx 2100$, when $V_{\rm c} = 32$km/s. This is the generic feature for all models that are on the ridge of the degeneracy. This velocity scale is interesting because it is between the circular velocity of SMC and the third massive MW satellite, Sagittarius. The degenerate models collectively require extremely high outflow mass-loading factor for subhalos with circular velocity as low as (or lower than ) 32 km/s. It would be interesting to investigate the significance of this velocity scale using other data sets of the galaxy population.

We marginalize the posterior to show how the model, together with its predetermined halo prior, reproduces the MW satellite stellar mass function. The reproduced stellar mass function using the two underlying dark matter halo priors are shown in Figure 14. The red bands denote the posterior predictive distribution using the Group 1 halo prior. The solid line in the middle of the bands denotes the median of the predictive distribution, and the bands from darker to lighter color enclose 20%, 50%, and 80% of the predictive distribution. Similarly, the blue bands show the predictive distribution using the Group 2 halo prior with three levels of color darkness. In addition, we also show the observed satellite mass function of the MW in the same figure. We find that when the model applied to Group 1 host, it can achieve very good match to the observed stellar mass function for the most massive 11 satellite galaxies. When the model is applied to Group 2 hosts, while the model matches the mass function equally well as using Group 1 host for satellites with mass lower than SMC, it still tends to predict lower stellar masses for the 2 most massive satellite galaxies. It is because the hosts do not host high-mass subhalos, precluding the solutions for making high enough mass to match the high-mass satellite galaxies. We compute the marginalized likelihood for each of the inference using different halo priors. The Bayes Factor, which is defined as the ratio between the marginalized likelihood of the model based on the Group 1 halo prior and the marginalized likelihood of the model based on the Group 2 halo prior. We find that the Bayes Factor $\mathcal{B} = \frac{\mathcal{M}_1}{\mathcal{M}_2} = 1.42$, which the halo prior of Group 1 type MW host is only weakly preferred by the data over the halo prior of Group 2 type host. Due to large uncertainties, the model is not able to distinguish between the two types of underlying dark matter halo using the observed MW satellite galaxy stellar mass function. Further inferences involving more data are needed.

### 3.3. *Metallicity relation of MW satellites*

We make predictions for the stellar-phase metallicity as a function of stellar mass for the MW satellite galaxies based on the two halo priors and the corresponding models specifically tuned for the halo prior. Instead of choosing one single model (the best or any arbitrary one), we use the posterior samples obtained with MCMC to produce the posterior predictive distribution Lu et al. (2012). The predictive distribution marginalizes over the model uncertainties that are allowed by the uncertainties of the constraining data, and quantifies the confidence levels of the predictions for the given model and prior. The posterior predictive distributions for both host halo

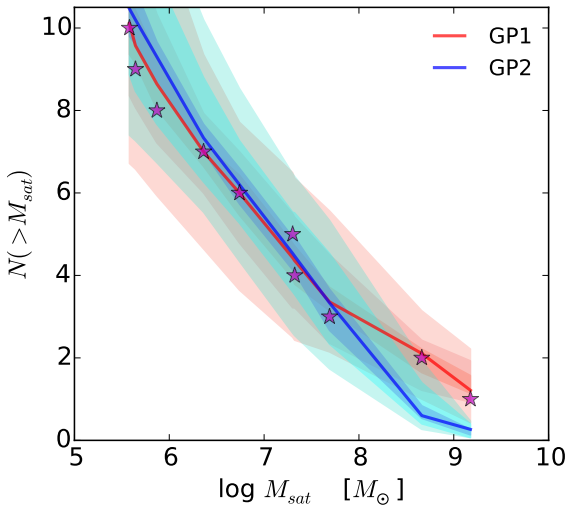


FIG. 14.— The posterior predictive distribution of the MW satellite stellar mass function. The red is predicted by the model constrained using host halo group 1, and the blue is predicted by the model constrained using host halo group 2. The bands with decreasing intensity are the 20%, 50%, and 80% predictive distribution for each halo group. Observational data from (McConnachie 2012) are shown by star symbols.

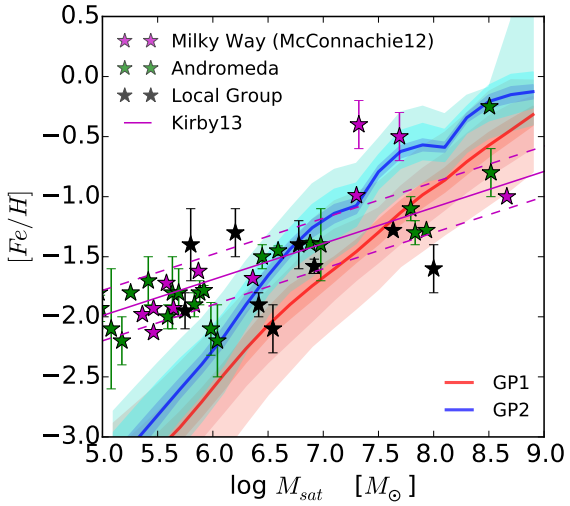


FIG. 15.— The posterior predictive distribution of the stellar-phase metallicity as a function of stellar mass. The red is predicted by the model constrained using host halo group 1, and the blue is predicted by the model constrained using host halo group 2. Observational data from (McConnachie 2012) for different populations are shown by star symbols.

priors are shown in Figure 15. In the figure, the red bands are predicted using the posterior constrained for the Group 1 halo prior, and the blue bands are predictions using the posterior constrained with the Group 2 halo prior. The bands with different intensity of each color show the 20%, 50%, and 80% predictive distribution. We find that while the model can be tuned to reproduce the satellite stellar mass function using either halo prior, they make rather different predictions for the stellar-phase metallicity-stellar mass relation. For the Group 2 prior, because the hosts typically do not have high-mass subhalos, the model needs weaker feedback and outflow to yield a relatively higher stellar mass for given subhalo mass. The consequence of the weaker outflow is to leave more metals in the galaxies, resulting in a higher metallicity for given stellar mass. In contrast, the Group 1 prior imposes host halos that typically host high-mass subhalos. The high-mass subhalos require higher outflow, which take away metals with hydrogen masses from the galaxy. The result suggests that for given galaxy formation scenario, when it is calibrated to the stellar mass function, the properties of the underlying dark matter halo, including the formation history, concentration and subhalo masses, can leave imprints in the satellite galaxy metallicity relation. Because the high-mass end of the subhalo mass function is correlated with the host formation history and concentration, the metallicity may also reflect the formation and structure properties of the host halo of the MW and its analogues.

An outstanding problem of the model we used is that the model predicts significant steeper slopes of the mass-metallicity relation, regardless of which halo prior is adopted. As shown in Figure15, the models that are constrained to the stellar mass function predict significantly lower metallicity for low-mass satellite galaxies, while they match the observation at the high-mass end. For $M_* \sim 10^{5.5}\mathrm{M}_\odot$, the models underpredict the metallicity by at least a factor of few. This is a generic problem of the model family as pointed out in Lu et al. (2014a). To test if the model family is able to reproduce the mass-metallicity relation and what determines the slope of the metallicity relation, we carry out another set of MCMC runs. In these runs, we use the slope and intercept of the mean metallicity relation of observed galaxies as constraints. We adopt the mean relation of Kirby et al. (2013), which shows that

$$\log\left(\frac{Z_*}{Z_\odot}\right) = 0.3 \pm 0.02 \log\left(\frac{M_*}{10^6 \mathrm{M}_\odot}\right) + (-1.69 \pm 0.04). \quad (10)$$

We take the best fit slope, $a_{\rm obs} = 0.3$ and intercept $b_{\rm obs} = 1.69$ and their $1-\sigma$ errors $\sigma_a = 0.02$ and $\sigma_b = 0.04$ as data constraints. For each model, we predict the stellar-phase metallicity for MW satellite galaxies using a group of MW host merger trees. We use a linear model to fit the stellar-phase metallicity v.s. stellar mass for the predicted samples, and yield the best fit parameters $a$ and $b$. We define a $\chi^2$-like objective function as

$$\chi^2 = \left(\frac{a - a_{\rm obs}}{\sigma_a}\right)^2 + \left(\frac{b - b_{\rm obs}}{\sigma_b}\right)^2. \quad (11)$$

We then use MCMC to sample the discrenpancy function defined as

$$\mathcal{T} = \exp\left(-\frac{\chi^2}{2}\right), \quad (12)$$

to explore the parameter space. We stress that this is not a formal Bayesian inference because we do not utilize a likelihood function that calculates the likelihood of data given model. We use this rather loosely defined discrepancy function to only test which parameter space can result in a mass-metallicity relation that is consistent

with observed slope and intercept.

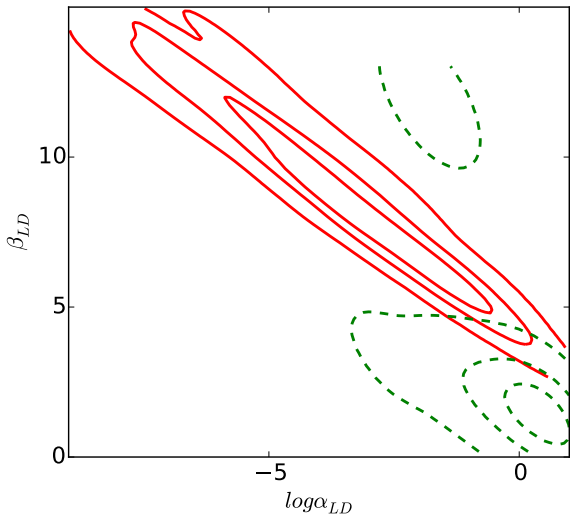


FIG. 16.— The posterior distributions of the model constrained to two different data. The red contours denote the posterior for the feedback outflow mass-loading parameters constrained by the MW satellite stellar mass function. The green contours denote the posterior for the same parameters of the model constrained by the stellar-phase metallicity relation (NOT YET CONVERGED!!!). The same merger tree set (group 1) is adopted.

Figure 16 shows the contours of the target distribution constrained by the metallicity relation. We compare the contours with the posterior distribution constrained by the stellar mass function using the Group 1 halo prior. It is obvious that the two constraints are completely shifted off from from each other. The contours suggest that the stellar mass function requires a steep power-law for the mass-loading parameter as a function of halo circular velocity, while the metallicity relation requires a shallow velocity dependence for the mass-loading parameter. This is reflected in the modes of the contours. The stellar mass function constrained contours have a mode at $\beta_{\rm LD} \sim 7$, but the metallicity relation constrained contours have a model at $\beta_{LD} \sim 2$. Even extending to higher confidence range, the 90% confidence range contour (the outer one) constrained by the two different data sets are still incompatible.

Marginalizing over the MCMC states, we predict the mean metallicity relation and the corresponding satellite stellar mass function. Those are shown in Figure 17 in comparison with those predictions with model constrained to the stellar mass function. First of all, the constrained model recovers the slope of the observed metallicity relation remarkably well. The reproduce metallicity relation agrees with observational mean relation with a much shallower slope than that constrained by the mass function. On the other hand, however, the model produces too abundant MW satellite galaxies with stellar mass below $10^{8.5}{\rm M}_{\odot}$, which is in a strong contradiction with the observed MW satellite stellar mass function.

This suggests that the parameterization for feedback that affects the star formation and metallicity has problem in explaining both data set simultaneously. In our model, and also many other models, feedback is captured in a form of strong outflow. The outflow expels cold baryonic matter out of the virial radius of the halo. Metals produced by star formation are assumed to mixed with the ISM and blown away by the outflow at the same time. The low stellar mass fraction of dwarf galaxies requires a large mass-loading factor for low-mass halos, resulting a too low metallicity. On the other hand, to match the metallicity relation, the mass-loading factor is required to be a much weaker function of halo circular velocity, so the outflow will not expel too much metal mass, but it results in too high baryon mass fraction in dwarf galaxies. To fix the problem in the ejective feedback scenario, presumably one has to assume a differential scaling relation for the metal-loading with respect to the mass loading. If we keep the mass-loading factor scaling relation inferred from the stellar mass function, the metal loading needs to be reduced significantly for low-mass galaxies, meaning that outflow from low-mass galaxies needs to be metal poor with respect to the ISM. This seems to be odd because outflow is powered by SN explosion, where metals are produced. It is hard for stronger outflow to carry less metals. On the other hand, if the process that suppresses star formation is different from the process that affects the metal budget of galaxies, we may be able to solve the problem. As Lu et al. (2015b) argued, feedback affecting star formation history may be preventive in nature. Feedback prevents baryon from collapsing into dark matter halos, and outflow that has a weak halo mass (circular velocity dependence) mainly affects the metallicity of low-mass galaxies. Such a model seems to be able to mach the baryon mass fraction and metallicity relation much better. Systematic parameter space exploration on such a model under well understood data constraints is needed to drawn conclusive answer.

## 4. DISCUSSION

In this paper, we analyze a suite of $N$-body cosmological simulations of Milky Way-sized halos. The analysis shows that the existence of high-mass subhalo is correlated with the formation history of the host halo and its concentration. For fixed halo mass, halos that have rapid recent accretion characterized by the fitting parameter $\beta + \gamma$ of the mass accretion history tend to host high-mass subhalo that are still exist at the present day. Those halos also tend to have lower concentration. The result suggests that the existence of LMC/SMC in the Milky Way halo indicates that the Milky Way is hosted by a dark matter halo that has lower concentration than average halos at the same mass. The formation history of the Milky Way halo may also have a particular shape, with significant accretion since $z = 0.5$. However, due to large uncertainties of galaxy formation model, we still cannot conclusively determine the properties of the MW host halo in terms of its concentration and formation history based on our inferences from solely the MW satellite galaxy stellar mass function.

The tends we find in the dark matter halo properties are still interesting. It has been known that the halo formation time is highly correlated with the concentration of the halo (Wechsler et al. 2002). Halos of same mass but different formation history can have very different characteristics or reside in different environments (e.g. Bullock et al. 2001; Allgood et al. 2006; Macciò et al. 2007). The differences in environment may give rise to

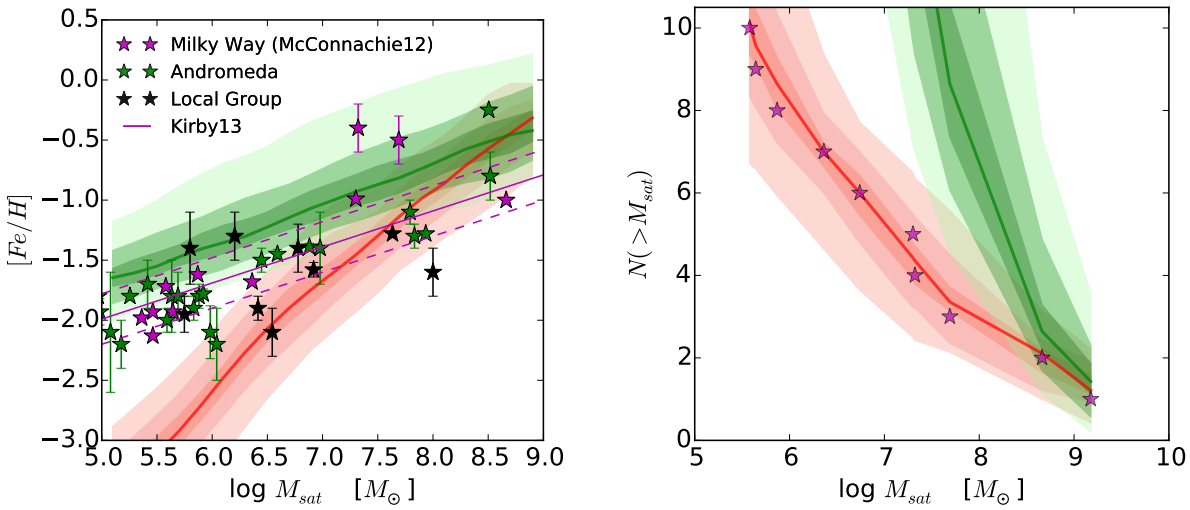


FIG. 17.— Left: the posterior predictive distribution of the stellar-phase metallicity as a function of stellar mass. Right: the posterior predictive distribution of the MW satellite stellar mass function. We only use the Group 1 halo prior, but constrained to different data. For both panels, the red bands show the predictive distribution for the model constrained to the satellite stellar mass function, and the green bands show the distribution for the model constrained to the stellar-phase metallicity relation. For each set, the color bands from dark to light are the 20%, 50%, and 80% predictive distribution.

different subhalo population that are recently accreted from the environment. For this reason, the abundance of subhalos within a dark matter halo does not only correlate with the mass of the halo (e.g. Kravtsov et al. 2004) but also the formation history of the halo (e.g. Zentner et al. 2005; Zhu et al. 2006). For this reason, it is important to understand the properties other than the mass when we study the formation of the MW, because the structure of the host halo, formation history, and its substructures may contain critical information to tell the MW from other galaxies with similar mass.

In this paper, we explore the effect of the host halo prior, which is characterized by the recent MAH and concentration, on the inferences of the formation physics of dwarf galaxies in the Milky Way. We find that when applying different host halo priors the observed stellar mass function requires different strength and circular velocity dependence for the feedback process. Although galaxy formation models are uncertain in accurately capturing how feedback works in detail, when the uncertainties of model parameters are marginalized, we find that host halos with rapid recent accretion require much stronger feedback than those without strong recent accretion. It is simple to understand because halos with significant recent accretion tend to host high-mass subhalos, which needs lower stellar mass to halo mass ratio to match the mass function.

We highlight that the different feedback strengths, according to the underlying dark matter halo, leave observable imprint in the mass-metallicity relation of dwarf galaxies. This is because outflow does not only suppress star formation but also expels produced metals out of galaxies. When the model requires stronger feedback in the model, it inevitably results in lower metallicity in dwarf galaxies. The result indicates that if we understand how feedback works in galaxy formation, we can use the model to learn about the underlying dark matter halo formation history of the Milky Way by using the observed metallicity relation. We stress, however, this requires accurate models of galaxy formation.

Our current understanding about feedback is still poor. We have shown that the model, even with very large flexibility, still cannot consistently reconcile the stellar mass function and the metallicity relation of MW dwarfs simultaneously. Based on the ejective feedback scenario, the model that can reproduce the mass function is needed to have very large mass-loading factor for halos with circular velocity $\sim 30$km/s, but the model predicts too low metallicity for low-mass galaxies because much of the metal mass is lost in outflow. To reproduce the observed metallicity relation, the outflow needs to have a much shallower halo circular velocity dependence. The inconsistency between the inferred feedback models seems to suggest that two different feedback mechanisms are needed to explain the mass function and the metallicity relation.

In future, we will perform investigations to understand the implication of star formation histories of Milky Way dwarf galaxies. The variations in the host halo formation history (reflected in the halo concentration and feedback strength) and their impact on star formation histories will be further studies. The star formation history of Milky Way dwarfs may also carry information about the formation history of the host halo.

## ACKNOWLEDGEMENTS

## REFERENCES

Allgood, B., Flores, R. A., Primack, J. R., et al. 2006, MNRAS, 367, 1781

Berlind, A. A., & Weinberg, D. H. 2002, ApJ, 575, 587

Boylan-Kolchin, M., Besla, G., & Hernquist, L. 2010, ArXiv e-prints, arXiv:1010.4797
Bullock, J. S., Kolatt, T. S., Sigad, Y., et al. 2001, MNRAS, 321, 559
Busha, M. T., Marshall, P. J., Wechsler, R. H., Klypin, A., & Primack, J. 2011a, ApJ, 743, 40
Busha, M. T., Wechsler, R. H., Behroozi, P. S., et al. 2011b, ApJ, 743, 117
Cautun, M., Hellwing, W. A., van de Weygaert, R., et al. 2014, MNRAS, 445, 1820
Gelman, A., & Rubin, D. 1992, Statistical Science, 7, 457
Jiang, F., & van den Bosch, F. C. 2015, MNRAS, 453, 3575
Kirby, E. N., Cohen, J. G., Guhathakurta, P., et al. 2013, ApJ, 779, 102
Koposov, S. E., Yoo, J., Rix, H.-W., et al. 2009, ApJ, 696, 2179
Kravtsov, A. V., Berlind, A. A., Wechsler, R. H., et al. 2004, ApJ, 609, 35
Lu, Y., Blanc, G. A., & Benson, A. 2015a, ApJ, 808, 129
Lu, Y., Mo, H. J., Katz, N., & Weinberg, M. D. 2006, MNRAS, 368, 1931
—. 2012, MNRAS, 421, 1779
Lu, Y., Mo, H. J., Lu, Z., Katz, N., & Weinberg, M. D. 2014a, MNRAS, 443, 1252
Lu, Y., Mo, H. J., & Wechsler, R. H. 2015b, MNRAS, 446, 1907
Lu, Y., Mo, H. J., Weinberg, M. D., & Katz, N. 2011, MNRAS, 416, 1949
Lu, Y., Wechsler, R. H., Somerville, R. S., et al. 2014b, ApJ, 795, 123
Lu, Z., Mo, H. J., & Lu, Y. 2015c, MNRAS, 450, 606
Ma, X., Hopkins, P. F., Faucher-Giguere, C.-A., et al. 2015, ArXiv e-prints 1504.02097, arXiv:1504.02097
Macciò, A. V., Dutton, A. A., van den Bosch, F. C., et al. 2007, MNRAS, 378, 55
Mao, Y.-Y., Williamson, M., & Wechsler, R. H. 2015, ApJ, 810, 21
McBride, J., Fakhouri, O., & Ma, C.-P. 2009, MNRAS, 398, 1858
McConnachie, A. W. 2012, AJ, 144, 4
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493
Okamoto, T., Frenk, C. S., Jenkins, A., & Theuns, T. 2010, MNRAS, 406, 208
Peeples, M. S., & Shankar, F. 2011, MNRAS, 417, 2962
Springel, V., White, S. D. M., Jenkins, A., et al. 2005, Nature, 435, 629
Springel, V., Wang, J., Vogelsberger, M., et al. 2008, MNRAS, 391, 1685
Stanimirović, S., Staveley-Smith, L., & Jones, P. A. 2004, ApJ, 604, 176
Strigari, L. E., Bullock, J. S., Kaplinghat, M., et al. 2007, ApJ, 669, 676
Tasitsiomi, A., Kravtsov, A. V., Gottlöber, S., & Klypin, A. A. 2004, ApJ, 607, 125
Taylor, J. E. 2011, Advances in Astronomy, 2011, 6
Ter Braak, C. J. F. 2006, Stat. Comput., 16, 239
van den Bergh, S. 2000, The Galaxies of the Local Group (Cambridge)
van den Bosch, F. C. 2002, MNRAS, 331, 98
van der Marel, R. P., Alves, D. R., Hardy, E., & Suntzeff, N. B. 2002, AJ, 124, 2639
Wang, W., Han, J., Cooper, A. P., et al. 2015, MNRAS, 453, 377
Wechsler, R. H., Bullock, J. S., Primack, J. R., Kravtsov, A. V., & Dekel, A. 2002, ApJ, 568, 52
Zentner, A. R., Berlind, A. A., Bullock, J. S., Kravtsov, A. V., & Wechsler, R. H. 2005, ApJ, 624, 505
Zhao, D. H., Jing, Y. P., Mo, H. J., & Börner, G. 2003, ApJ, 597, L9
Zhu, G., Zheng, Z., Lin, W. P., et al. 2006, ApJ, 639, L5