Beyond Ultra-diffuse Galaxies. II. Environmental Quenching of Mass–Size Outliers among the Satellites of Milky Way Analogs

Jiaxuan Li (李嘉轩)[1] , Jenny E. Greene[1] , Johnny Greco[1,2] , Rachael Beaton[1] , Shany Danieli[1,5] , Andy Goulding[1] ,
Song Huang (黄崧)[1,3] , and Erin Kado-Fong[1,4]
[1] Department of Astrophysical Sciences, 4 Ivy Lane, Princeton University, Princeton, NJ 08544, USA; jiaxuanl@princeton.edu
[2] Center for Cosmology and AstroParticle Physics (CCAPP), The Ohio State University, Columbus, OH 43210, USA
[3] Department of Astronomy and Tsinghua Center for Astrophysics, Tsinghua University, Beijing 100084, People's Republic of China
[4] Physics Department, Yale Center for Astronomy & Astrophysics, PO Box 208120, New Haven, CT 06520, USA
Received 2023 February 27; revised 2023 July 4; accepted 2023 July 4; published 2023 September 11

Abstract

Recent observations have reignited interest in a population of dwarf galaxies that are large and diffuse for their mass, often called ultra-diffuse galaxies (UDGs). However, the origin and evolution of these mass–size outliers and the role of the environment are still unclear. Using the exquisitely deep and wide Hyper Suprime-Cam Strategic Survey images, we search for ultra-puffy galaxies (UPGs), defined as being 1.5σ larger than the average size for their mass, around Milky Way–like galaxies. We present the sizes and radial distributions of mass–size outliers and derive their quenched fraction to explore the impact of the environment. Surprisingly, despite being outliers in size, the UPGs have a similar quenched fraction as normal-sized satellites of Milky Way analogs in both observations and simulations, suggesting that quenching is not tied to being a mass–size outlier. The quenched fraction is higher for the UPGs associated with redder hosts, as well as those that are closer to the host in projection. In contrast, the UDGs are overall redder and more quiescent compared with normal satellites. We show that the classic definition of UDGs is heavily weighted toward quenched galaxies and thus cannot be used for a study of quenching of mass–size outliers.

Unified Astronomy Thesaurus concepts: Low surface brightness galaxies (940); Galaxy properties (615); Dwarf galaxies (416); Galaxy quenching (2040); Galaxy evolution (594)

1. Introduction

Dwarf galaxies, given their overwhelming abundance, shallow gravitational potential wells, diverse evolutionary histories, and high dark matter content, are ideal probes for testing the ΛCDM paradigm and understanding dark matter (Bullock & Boylan-Kolchin 2017) and baryonic processes (Sales et al. 2022). In particular, dwarf galaxies and their star formation history are sensitive to the environments where they reside. Nearly all isolated dwarf galaxies in the field are star-forming (Geha et al. 2012), but satellite dwarf galaxies in the Local Group and Local Volume hosted by Milky Way (MW)–mass galaxies are much more quiescent in star formation, with a near-unity quenched fraction at $M_\star \lesssim 10^8\,M_\odot$ (e.g., Grcevich & Putman 2009; Spekkens et al. 2014; Wetzel et al. 2015b; Baxter et al. 2021; Mao et al. 2021; Putman et al. 2021; Carlsten et al. 2022a; Greene et al. 2023; Karunakaran et al. 2022). Such a contrast indicates that environment plays an important role in regulating star formation in dwarf galaxies.

Several physical processes are believed to be responsible for the cessation of star formation (also known as "quenching") in dwarf galaxies. Isolated dwarf galaxies can be quenched by bursty stellar feedback (e.g., El-Badry et al. 2018), reionization (for ultra-faint dwarfs; Bullock et al. 2000; Benson et al. 2002; Somerville 2002; Tollerud & Peek 2018; Applebaum et al. 2021), or previously passing a more massive halo, as in the

case of so-called backsplash galaxies (Simpson et al. 2018; Benavides et al. 2021). For the satellite galaxy of an MW analog (a galaxy having a similar stellar or halo mass as the MW), the ram pressure experienced by the satellite as it moves through the hot circumgalactic medium (CGM) of the host galaxy could remove the gas and cause the cessation of star formation (e.g., Gunn & Gott 1972; McCarthy et al. 2008; Grcevich & Putman 2009; Emerick et al. 2016; Simpson et al. 2018; Tremmel et al. 2020; Samuel et al. 2022). It has been shown that ram pressure stripping can efficiently quench galaxies at $M_\star < 10^8\,M_\odot$ (Fillingham et al. 2016; Simpson et al. 2018; Buck et al. 2019). Starvation, where cold gas supplies are halted following infall to a more massive hot halo, is believed to be dominant for higher-mass satellites (Fillingham et al. 2016). Tidal stripping can also remove the gas from the satellite and cause quenching (Simpson et al. 2018), but strong tidal interactions are needed because the dark matter needs to be stripped before a significant gas loss happens. A satellite dwarf galaxy might also be preprocessed in a lower-mass group prior to falling into the current host galaxy, and such preprocessing would increase the quenched fraction (Wetzel et al. 2015a; Jahn et al. 2022; Samuel et al. 2022). Extensive studies have been done on the quenching of satellites in MW analogs using simulations (e.g., Simpson et al. 2018; Buck et al. 2019; Garrison-Kimmel et al. 2019; Simons et al. 2020; Akins et al. 2021; Joshi et al. 2021; Karunakaran et al. 2021; Font et al. 2022; Pan et al. 2023; Samuel et al. 2022), and many of them are able to reproduce the observed quenched fraction in the Local Group and Local Volume. It is now time to understand the relative importance of these mechanisms as a function of satellite and host properties.

---

[5] NASA Hubble Fellow.

Interestingly, several of these physical processes might be able to puff up satellite galaxies and turn them into outliers with respect to the average mass–size relation. An example class of mass–size outlier is the ultra-diffuse galaxies (UDGs; van Dokkum et al. 2015), which have a large size ($r_e > 1.5$ kpc) and low surface brightness ($\mu_0(g) > 24$ mag arcsec^{-2}). It is therefore valuable to ask if the same mechanism that dominates quenching also produces the observed population of mass–size outliers. This will provide a unique angle to peek into the formation and evolution of satellite galaxies in MW analogs and help us understand whether the mass–size outliers are a distinct class of satellite galaxies or just the large-size tail of the satellite population.

Large samples of mass–size outliers already exist in galaxy cluster environments, but lower-density environments such as groups are relatively unexplored (e.g., Karunakaran & Zaritsky 2023). In Li et al. (2022a, hereafter L22), we presented a sample of dwarf galaxies that are mass–size outliers among the satellites of MW analogs at $0.01 < z < 0.04$. As a follow-up study, in this paper, we ask whether the mass–size outliers are more quiescent compared to the "normal" satellite population of MW analogs. We take the mass–size outliers from L22 (described in Section 2) and present their characteristics, including their size and radial distributions, in Section 3. We then derive the quenched fractions as a function of satellite stellar mass, host color, and projected distance to the host (Section 4). We compare the quenched fractions of mass–size outliers with the bulk of the satellite population of nearby MW analogs and simulations. In Section 5, we discuss the implications of these results on the formation and quenching of mass–size outliers. We adopt a flat ΛCDM cosmology from Planck Collaboration et al. (2016) with $\Omega_m = 0.307$ and $H_0 = 67.7$ km s^{-1} Mpc^{-1}. We use the AB system (Oke & Gunn 1983) for magnitudes. The stellar masses used in this work are based on the Chabrier (2003) initial mass function.

2. Data and Sample Selection

In this section, we briefly summarize the data and mass–size outlier samples presented in L22. We refer interested readers to L22 for a detailed description of the samples and the search algorithms.

2.1. Mass–Size Outlier Sample

In L22, we performed a systematic search for low surface brightness galaxies (LSBGs) in Subaru's Hyper Suprime-Cam (HSC) Strategic Survey Program (Aihara et al. 2018; hereafter the HSC survey). The HSC survey is an imaging survey using the 8.2 m Subaru telescope and the Hyper Suprime Camera (Miyazaki et al. 2012, 2018), covering ~1000 deg^2 in five broad bands (grizy) and reaching a depth of $g = 26.6$, $r = 26.2$, and $i = 26.2$ mag (5σ point-source detection). In L22, we used the data release PDR2, which covers ~300 deg^2 and has a global sky subtraction to preserve low surface brightness features (Aihara et al. 2018; Li et al. 2022b).

We search for LSBGs following the method in Greco et al. (2018) with several modifications to suit PDR2 and improve the completeness and purity. Briefly, we run SExtractor (Bertin & Arnouts 1996) on the coadd images after removing bright extended sources and associated diffuse light. After applying an initial size and color cut to the output catalog, we model each source by running scarlet (Melchior et al. 2018)

and extract a series of structural and morphological parameters from the nonparametric model. We design a metric based on the color, size (r_e; circularized half-light radius), surface brightness ($\overline{\mu}_{\rm eff}(g)$; average surface brightness within r_e in the g band), and morphology and use it to remove false positives that are not likely to be real LSBGs. For the remaining objects, we fit a parametric model to estimate their sizes, magnitudes, and surface brightnesses. The completeness of the search is characterized by injecting mock galaxies into coadd images and recovering them with our algorithm. In the size range of $3'' < r_e < 14''$, our search is >70% complete to $\overline{\mu}_{\rm eff}(g) < 26.5$ mag arcsec^{-2} and >50% complete to $\overline{\mu}_{\rm eff}(g) \leqslant 27.0$ mag arcsec^{-2}. For comparison, Zaritsky et al. (2021) searched for LSBGs in the Dark Energy Camera Legacy Survey and reached ~25% completeness at $\mu_0(g) \approx 25.5$ mag arcsec^{-2}. Tanoglidis et al. (2021) reported the LSBG search using Dark Energy Survey data and showed a completeness of ~30% at $\overline{\mu}_{\rm eff}(g) \approx 26$ mag arcsec^{-2}. Therefore, our completeness is high compared with other LSBG searches, which ensures high fidelity in the scientific results. We remove LSBGs with a completeness of less than 10% for cleaner statistics.

Then we cross-match our LSBG sample with MW analogs selected from the NASA-Sloan Atlas[6] (Blanton et al. 2005, 2011) catalog by requiring $0.01 < z < 0.04$, $10.2 < \log M_\star / M_\odot < 11.2$, and being in the HSC PDR2 footprint. In this way, we select 922 MW analogs. For context, we compare our selection with two surveys that also search satellites of MW analogs. The Satellites Around Galactic Analogs (SAGA; Geha et al. 2017; Mao et al. 2021) survey is an ongoing spectroscopic survey of classical satellites of 100 MW analogs at 20 Mpc $< D <$ 40 Mpc. Their MW analogs are in a stellar mass range of $10.2 < \log M_\star / M_\odot < 11.0$. The Exploration of Local VolumE Satellites (ELVES; Carlsten et al. 2020, 2021, 2022a, 2022b) survey maps the satellites of 30 host galaxies with $M_\star > 10^{9.9} M_\odot$ and $D < 12$ Mpc. Most of the satellites have direct distances from either the tip of the red giant branch or surface brightness fluctuations, with the latter dominating for the faint satellites. The ELVES sample already includes MW and M31 satellites. To sum up, our MW analog selection is similar to that of SAGA. Because of the volume-limited criterion, ELVES includes several groups more massive than the SAGA hosts and ours.

We calculate the virial mass of the MW analogs based on the stellar-to-halo mass relation in Behroozi et al. (2010) using halotools[7] (Hearin et al. 2017). Then we convert virial mass to virial radius $R_{\rm vir}$ assuming a flat ΛCDM cosmology, where the virial overdensity only depends on cosmological parameters and redshift (Bryan & Norman 1998). For a given MW analog, we associate any LSBG that falls inside its projected virial radius with the MW analog as the host galaxy. If one LSBG is matched to more than one host, we assign it to the nearest host based on the angular separation normalized by the host virial radius. In total, we surveyed the 922 MW analogs, and we found 2510 LSBG candidates associated with 689 MW analogs. We assume that the LSBGs are at the same redshifts as their hosts.

After cross-matching, we select two mass–size outlier samples, namely, UDGs and *ultra-puffy galaxies* (UPGs).

---

[6] http://nsatlas.org
[7] https://halotools.readthedocs.io/en/latest/index.html

The UDGs are defined to have $r_e + \sigma(r_e) > 1.5\,\mathrm{kpc}$ and $\overline{\mu}_{\mathrm{eff}}(g) + \sigma(\overline{\mu}_{\mathrm{eff}}(g)) > 25\,\mathrm{mag\,arcsec}^{-2}$, which takes the 1σ measurement error into account. This definition is consistent with the original definition in van Dokkum et al. (2015) for galaxies with Sérsic indices $n \sim 1$. There are 412 LSBG candidates that satisfy the UDG definition, and they are defined as our UDG sample. As we argued in L22, the UDG definition does not consider the dependence of galaxy size on galaxy mass. A UDG with $r_e = 1.5\,\mathrm{kpc}$ is an outlier in size if its stellar mass is $10^7\,M_\odot$, but it will be normal-sized if its stellar mass is $10^{8.5}\,M_\odot$. Therefore, UDGs are not necessarily mass–size outliers. We propose the concept of UPGs, which are defined to lie 1.5σ above the average mass–size relation of the satellite galaxies. Because we are interested in mass–size outliers associated with MW analogs, we take the mass–size relation and its scatter from Carlsten et al. (2021), which is derived from the satellites of MW analogs in the Local Volume. The mass–size relation in Carlsten et al. (2021) is measured for $10^{5.5} < M_\star/M_\odot < 10^{8.5}$, and we linearly extrapolate this relation to $M_\star \sim 10^9\,M_\odot$ to select UPGs in this work. Among our LSBG candidates, we have 337 galaxies as our sample of UPGs that fall 1.5σ above the average mass–size relation. The UPGs are associated with 239 MW analogs.

2.2. Contamination Correction

We note that the LSBGs are associated with the MW-like hosts only in projection, so a certain fraction of them will be foreground or background galaxies that happen to be close to the MW analogs in the sky. Therefore, we apply a statistical interloper correction as follows. We randomly select a continuous patch of the sky of $24\,\mathrm{deg}^2$ in HSC PDR2 regardless of whether it contains MW analogs. Then we repeat the LSBG search in this area and apply the same cuts as we described above to remove false positives. There are 480 LSBG candidates in this area that represent possible contaminants for the UDG and UPG samples. Because distances are needed to define the UDG and UPG classes, we randomly match these 480 LSBGs with the 922 MW analogs that we have surveyed. With such "artificial" associations, we calculate the corresponding physical sizes and stellar masses and select UDGs and UPGs accordingly. Such matching was repeated 200 times, and we obtain 7625 artificial UDGs and 8267 artificial UPGs. These UDGs and UPGs are "artificial" only in the sense of being associated with random MW-like hosts. Therefore, the number densities of artificial UDGs and UPGs are $S_{\mathrm{UDG}} = 1.60 \pm 0.25$ and $S_{\mathrm{UPG}} = 1.72 \pm 0.23\,\mathrm{deg}^{-2}$, respectively. Equivalently, the contamination fractions for both samples are $f_{\mathrm{contam}}^{\mathrm{UDG}} \approx 35\% \pm 5\%$ and $f_{\mathrm{contam}}^{\mathrm{UPG}} \approx 45\% \pm 6\%$. We have checked that these fractions are not sensitive to the specific sky regions used to construct the artificial UDG and UPG samples (Li et al. 2022a). We will use these values to correct the contribution from the foreground and background interlopers when deriving the size and radial distributions in Sections 3.1 and 3.2.

Furthermore, the contaminants might not be homogeneous in color and could bias the color distribution, thereby changing the quenched fraction of the mass–size outliers. If the artificial UDGs (UPGs) have different properties from the real ones, we can use this information to derive a probability that an object is a real UDG (UPG) at the proposed distance. Indeed, we find that the $g - i$ color distribution of the "artificial" UDG (UPG) sample is bluer than that of the observed UDG (UPG) sample.

Specifically, the 25th, 50th, and 75th quantiles of the $g - i$ color are (0.50, 0.71, 0.82) for the real UPG sample and (0.42, 0.54, 0.72) for the artificial UPG sample. The color distribution of the observed UDGs (UPGs) will thus be biased to be bluer, probably because interlopers such as field dwarfs and spiral LSBGs tend to be blue. This motivates us to assign importance weights based on the $g - i$ color distribution and the contamination fraction f_{contam}.

Taking UDGs as an example, we first compute the normalized histograms of $g - i$ colors for both observed and artificial UDG samples (denoted as λ_k^{obs} and $\lambda_k^{\mathrm{artificial}}$ at the $(g - i)_k$ bin, where $\sum_k \lambda_k = 1$). The weight assigned to UDGs in color bin $(g - i)_k$ is then estimated to be $w_k = \max(1 - f_{\mathrm{contam}}^{\mathrm{UDG}}\,\lambda_k^{\mathrm{artificial}}/\lambda_k^{\mathrm{obs}}, 0)$. This weight stands for the possibility of not being a contaminant. The weights of UPGs are assigned following the same procedure. Such weights are applied when we calculate the quenched fraction in Section 4. We test this method by generating a population of mock UDGs (UPGs), adding a bluer population of contaminants, and recovering the underlying UDG (UPG) color distribution and quenched fraction. Such a mock test verifies that this color-based contamination subtraction is sufficient for this work, and our main results are robust against contamination subtraction.

3. Sample Characteristics

In this section, we present statistical analyses of the mass–size outliers to gain a comprehensive understanding of the sample. We derive the size (Section 3.1) and radial (Section 3.2) distributions, paving the way for the discussion of quenching and environmental effects.

3.1. Size Distribution

The size distribution of UDGs is an important topic, since it has been used to test the formation scenarios of UDGs (e.g., Amorisco & Loeb 2016; van der Burg et al. 2017). In this section, we calculate the size distributions of our UDG and UPG samples and compare them with the literature. We bin the sizes on a logarithmic scale ($\log r_e$) and calculate the size distribution $dn/d\log r_e$ as follows. First of all, we take the artificial UDGs in the random fields (described in Section 2) and calculate their size distribution per square degree as a proxy for contamination. For each host, we calculate the size distribution of the associated UDGs, and we multiply the contaminant size distribution by the virial area of the host and subtract it from the UDG size distribution. We combine the UDG size distributions for all hosts by taking the average size distribution, shown as green pentagons in Figure 1. For reference, we plot the contaminant size distribution in an average virial area as gray dots. The completeness is calculated for each size bin, and the correction is applied. The final size distribution is shown as red squares. Although the number of mass–size outliers scales with stellar mass, we do not weight each host according to stellar mass, since the stellar masses of the hosts in our sample are quite similar. The error shown in Figure 1 takes both the Poisson error and the measurement error in size into account.

We find that the UDG size distribution roughly follows a power law. We then fit a power law to our UDG size distribution and get $dn/d\log r_e \propto r_e^{-1.45 \pm 0.20}$. This is shallower than the power laws presented in van der Burg et al. (2016; $dn/d\log r_e \propto r_e^{-3.40 \pm 0.19}$) and van der Burg et al. (2017;
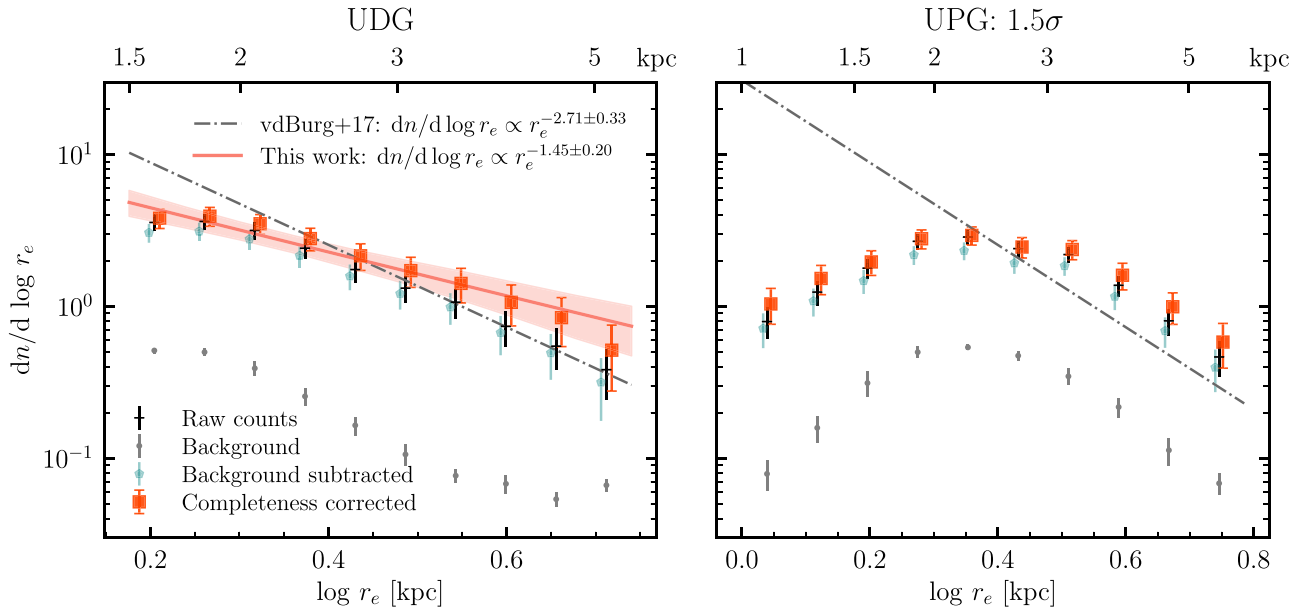
Figure 1. Size distribution of UDGs (left) and UPGs (right) hosted by MW analogs. The contribution from the foreground and background contaminants (gray dots) is subtracted, and the completeness effect has been corrected. We fit a power law to the size distribution of UDGs and get a power law $dn/d\log r_e \propto r_e^{-1.45\pm0.20}$. By comparing with the size distributions of UDGs in van der Burg et al. (2016, 2017), we find that the power law becomes steeper as the average halo mass increases. It is possible that denser environments provide stronger tidal forces and thus produce more small UDGs. The UPG size distribution is suppressed at the smaller-size end because small galaxies are not 1.5σ above the mass–size relation and thereby not included as UPGs.

$dn/d\log r_e \propto r_e^{-2.71\pm0.33}$, shown as gray dashed–dotted lines in Figure 1), where they also corrected for background contribution and completeness. Among others, van der Burg et al. (2016) focused on large galaxy clusters and found a quite steep power law; van der Burg et al. (2017) included less-massive groups, and the resulting power law is less steep. Combined with our findings, it might be possible that a denser environment produces more small UDGs, probably due to stronger tidal interactions.

The distribution of $\log r_e$ for UPGs is shown in the right panel of Figure 1. The size distribution is no longer monotonic because the UPG sample is constructed not by cutting at a certain size but by cutting along the mass–size relation. Consequently, for a given mass, many small UDGs are not in the UPG sample, since their sizes are not extreme on the mass–size plane. The size distribution of the UPG sample is thus suppressed at the smaller-size end. We also split the UDG and UPG samples based on their host stellar masses and distance to hosts. For UDGs, we consistently find shallow power laws with indices of ~-1.5, not depending on host stellar mass. Interestingly, we do not find a significant change in the size distribution as UDGs (UPGs) get closer to their hosts.

3.2. Radial Distribution

The spatial distribution of satellite galaxies around their host is proposed as a good tracer of the host dark matter halo profile and thus is a probe of environmental effects. Many studies have found that the radial distribution of satellites can be well fit by a Navarro–Frenk–White (NFW; Navarro et al. 1997) profile, although the radial distribution of satellites is a function of host mass, satellite mass, and satellite color (e.g., van den Bosch et al. 2005; Sales et al. 2007; Budzynski et al. 2012; Tal et al. 2012; Wang et al. 2014; Carlsten et al. 2020; McDonough & Brainerd 2022). On the other hand, the radial distribution of satellites is a sensitive probe of many physical processes,

including reionization (Kravtsov et al. 2004) and tidal disruption (Samuel et al. 2020). Subhalos can also be artificially disrupted due to numerical issues in simulations (van den Bosch et al. 2018; Carlsten et al. 2020), thus changing the radial distribution of satellites. Furthermore, it is unclear whether the mass–size outliers follow the same radial distribution as normal satellites (e.g., Tremmel et al. 2020). In cluster environments, the number of UDGs is found to be depleted near the centers of clusters, indicating that they are more likely to be disrupted near the cluster center (e.g., van Dokkum et al. 2015; van der Burg et al. 2016; Mancera Piña et al. 2018). The radial distribution of mass–size outliers in groups is not as well explored. In this subsection, we derive the radial distributions of the mass–size outliers, compare them with literature results, and discuss possible implications.

We calculate the radial distribution as follows. For each host, we count the number of UDGs (UPGs) in each radial bin scaled by $R_{\rm vir}$ and calculate the number density of UDGs (UPGs) within each radial annulus. Since foreground and background contaminants are distributed quite homogeneously in the sky, their contribution to the observed radial distribution can be subtracted based on the average contaminant density and the angular area occupied by the host. Then we average over all hosts and correct for completeness. The final radial distributions are shown in Figure 2. Red squares are the number of UDGs (UPGs) per radial bin, and the error includes the Poisson, contamination subtraction, and completeness errors. The radial distributions of both UDGs and UPGs turn over at $R < 0.2\,R_{\rm vir}$ (shown in lighter colors). Completeness may well drop with proximity to the host due to the blending between the UDG and the host galaxy or due to sky subtraction issues. Although we have corrected each object for completeness based on size and surface brightness, it is still possible that we have not fully corrected for radial incompleteness within a halo. As we mentioned above, it is also possible that UDGs and UPGs are depleted because of tidal disruption.
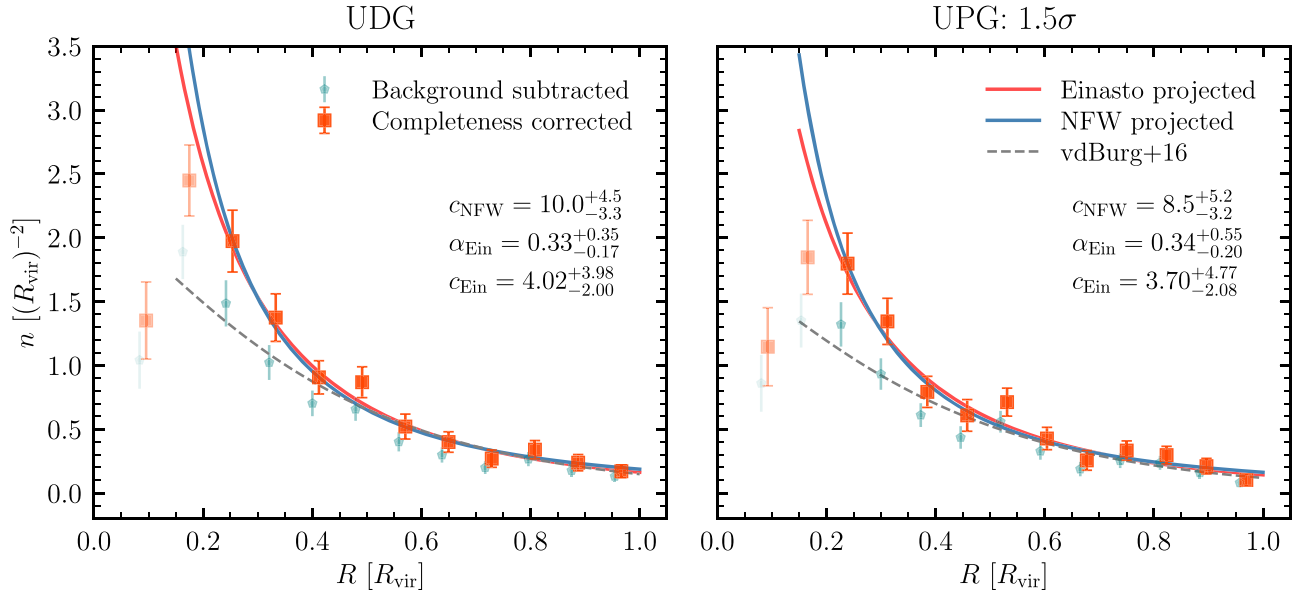
Figure 2. Radial distributions of UDGs (left) and UPGs (right). We show the radial distributions after subtracting the foreground and background contamination (shown in teal) and correcting for completeness in each radial distance bin (shown in red). After excluding $R < 0.2R_{vir}$ (shown in lighter colors), we fit projected NFW and Einasto profiles to the radial distributions and find good fits for both profiles. The radial distributions of UDGs and UPGs are similar in shape. The best-fit NFW concentration is consistent with an MW-mass halo, whereas the best-fit Einasto profiles have a smaller concentration than MW-mass halos.

We fit the radial distributions of UDGs (UPGs) with the projected NFW and Einasto (Einasto 1965) profiles for $R > 0.2R_{vir}$ using colossus[8] (Diemer 2018). The Einasto profile introduces an extra "shape" parameter, α_{Ein}, and has been argued to be a better description of dark matter halo profiles in simulations (e.g., Navarro et al. 2004, 2010; Gao et al. 2008; Dutton & Macciò 2014). The Einasto profile also makes concentration estimates less sensitive to the radial range fitted. The best-fit parameters are obtained using the least-squares method, and the parameter uncertainty is from the estimated covariance matrix of the parameters. The best-fit NFW (blue) and Einasto (red) profiles are shown in Figure 2 as solid lines. The best-fit Einasto profile from van der Burg et al. (2016) is shown as the dashed gray line.

Compared with van der Burg et al. (2016), there are more UDGs at smaller radial distances in our sample; thus, the UDG radial distribution is more concentrated, even though we exclude $R < 0.2R_{vir}$. Van der Burg et al. (2016) found that the radial distribution of UDGs cannot be well described by an NFW profile, but an Einasto profile does provide a good fit to the data. Unlike van der Burg et al. (2016), we find that both the NFW and Einasto profiles describe the radial distribution of UDGs (UPGs) quite well. The concentration of the best-fit NFW profile is $c_{NFW,UDG} = 10.0^{+4.5}_{-3.3}$ for UDGs and $c_{NFW,UPG} = 8.5^{+5.2}_{-3.2}$ for UPGs. The best-fit Einasto profile has $\alpha_{Ein,UDG} = 0.33^{+0.35}_{-0.17}$, $c_{Ein,UDG} = 4.02^{+3.98}_{-2.00}$ for UDGs and $\alpha_{Ein,UPG} = 0.34^{+0.55}_{-0.20}$, $c_{Ein,UPG} = 3.70^{+4.77}_{-2.08}$ for UPGs. The radial distribution profiles of UDGs and UPGs are found to be very similar in shape. Our best-fit Einasto profiles have lower α and higher concentration compared with van der Burg et al. (2016), where they have $\alpha_{Ein,UDG} = 0.92^{+0.08}_{-0.18}$, $c_{Ein,UDG} = 1.83^{+0.13}_{-0.12}$.

According to the concentration–mass relation of dark matter halos, the concentration of an MW-like halo is $c_{NFW} \sim 10$ (e.g., Bullock et al. 2001; Duffy et al. 2008; Dutton & Macciò 2014; Diemer & Joyce 2019). The Einasto shape parameter for an

MW-like halo is $\alpha_{Ein} \sim 0.15$ at $z = 0$, and its concentration is $c_{Ein} \sim 8$–10 (Gao et al. 2008; Dutton & Macciò 2014). From Figure 2, the best-fit NFW profiles for the radial distributions agree with the MW-like halo in terms of concentration, but the best-fit Einasto profiles have a lower concentration compared with MW-like halos. This agrees with the fact that the concentration of the satellite radial distribution is lower than that of the halo (e.g., McDonough & Brainerd 2022). We notice that the errors of the best-fit parameters are large due to the sample size and also the fact that concentration is most sensitive to data at smaller radial distances, which are excluded from the analysis. A larger, cleaner, and more complete sample is needed to better constrain the radial distributions.

4. Quenching

As argued in Section 1, the star-forming properties of mass–size outliers are an interesting subject given their diffuse stellar components and unknown formation history. In this section, we first discuss how to define "quenching" for our sample and present the distribution of mass–size outliers on the mass–size plane (Section 4.1). Then we study the quiescent fractions for UDGs and UPGs as a function of stellar mass and host properties and compare them with other observations and simulations (Section 4.2). We discuss how these results would imply the formation and quenching of mass–size outliers in Section 5.

4.1. Quenching Criteria

In principle, direct estimates of the current star formation rate (SFR) from UV or Hα or estimates for the H I gas are needed to define whether a satellite is quenched. For example, the SAGA survey (Geha et al. 2017; Mao et al. 2021) measures the Hα equivalent width from spectra to define quenched satellites. There are also many studies trying to probe H I for star-forming satellites in the Local Group (e.g., Grcevich & Putman 2009; Spekkens et al. 2014; Putman et al. 2021;
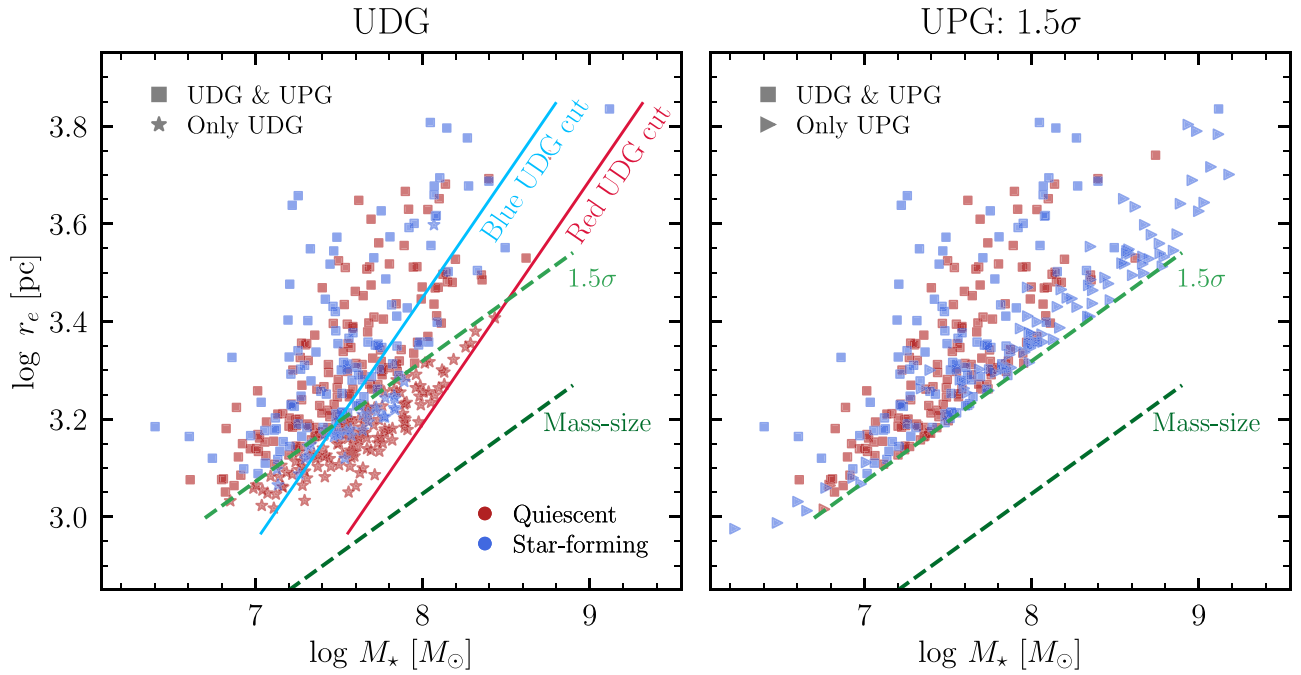
Figure 3. Distribution of UDGs (left) and UPGs (right) on the mass–size plane. We classify galaxies into quiescent and star-forming based on a mass-dependent color cut. The average mass–size relation and 1.5σ above it are shown in green. The solid lines are constant surface brightness cuts at $\overline{\mu}_{\rm eff}(g) = 25$ mag arcsec^{-2} for two different colors, $g - i = 0.4$ (blue) and 0.8 (red). The UDG sample includes a significant number of galaxies that are not mass–size outliers (falling below the 1.5σ line) that are red in color because the surface brightness cuts are different on the mass–size plane for blue and red galaxies.

Karunakaran et al. 2022; Zhu & Putman 2023) and derive the quiescent fraction. However, due to the low surface brightness and low-luminosity nature of our mass–size outliers, obtaining such measurements is extremely expensive over the full footprint of our sample. Therefore, we seek other indicators for SFR, including the broadband colors and morphology.

The ELVES survey (Carlsten et al. 2022a) searches for satellite galaxies of 30 MW-like hosts in the Local Volume ($D < 12$ Mpc). Carlsten et al. (2022a) visually inspected all satellites in the ELVES survey and classified them into early-type (red and smooth) and late-type (blue, asymmetric, and clumpy). Using this sample, they found that a mass-dependent color cut $(g - i)_Q = -0.067 \cdot M_V - 0.23$ could divide the sample into two subsets that are nearly identical to the early and late types based on morphology. Carlsten et al. (2022a) derived the quenched fraction of the ELVES satellites using such a color cut and showed that the resulting quenched fractions are very similar to the morphology-based quenched fractions. Furthermore, Font et al. (2022) showed that this color cut effectively separates star-forming galaxies from quiescent ones in simulations that are not fine-tuned to match the observed properties of ELVES galaxies. Therefore, we use this color cut to define quenched galaxies among our mass–size outliers; galaxies that are redder than $(g - i)_Q$ are defined as quenched, whereas bluer galaxies are defined as star-forming. To apply the color cut, we derive the V-band apparent magnitude following $V = g - 0.5784 \cdot (g - r) - 0.0038$[9] and convert it to absolute magnitude M_V. For comparison with ELVES and SAGA (Figures 4–6), we use the same color cut to maintain consistency.

Before calculating the quenched fraction, we first look at the distributions of mass–size outliers on the mass–size plane, as shown in Figure 3. This figure is similar to Figure 5 in L22, but here we highlight the quiescent galaxies in red and star-forming galaxies in blue to emphasize how the star formation properties differ between the two samples. The average mass–size relation from Carlsten et al. (2021) and the 1.5σ line above it are shown in green. The blue and red solid lines show the constant surface brightness line ($\overline{\mu}_{\rm eff}(g) = 25$ mag arcsec^{-2}) for galaxies with $g - i = 0.4$ and 0.8, respectively. The two colors are chosen as examples to highlight the dependence of surface brightness cut on galaxy color; for a given stellar mass, a bluer galaxy must have a larger size to be considered a UDG. Compared with the UPG sample, we find that the UDG sample includes many galaxies that are below the 1.5σ line and quiescent. Red galaxies dominate the region below the 1.5σ line because blue galaxies do not satisfy the surface brightness cut in the UDG definition. In contrast, the UPG sample is comprised of many star-forming galaxies at the high-mass end.

4.2. Quenched Fractions

In this section, we study the quenched fraction as a function of satellite stellar mass. We divide the stellar mass range $6.4 < \log M_\star / M_\odot < 9.0$ into eight bins and calculate the fraction of quiescent galaxies for each bin. Both the denominator and numerator are weighted by the importance weight w_k, which represents the likelihood of not being a contaminant (see Section 2.2). We also apply completeness corrections to the denominator and numerator. Not applying the importance weights will make the quenched fraction lower, since the background contaminants are bluer, on average, than UDGs and UPGs.

We plot the quenched fractions f_q of the full UDG (gray) and UPG (green) samples in Figure 4. The error bars correspond to

---

[9] This relation was derived for SDSS filters. We neglect the small difference between the SDSS and HSC filter systems: http://classic.sdss.org/dr4/algorithms/sdssUBVRITransform.html#Lupton2005.
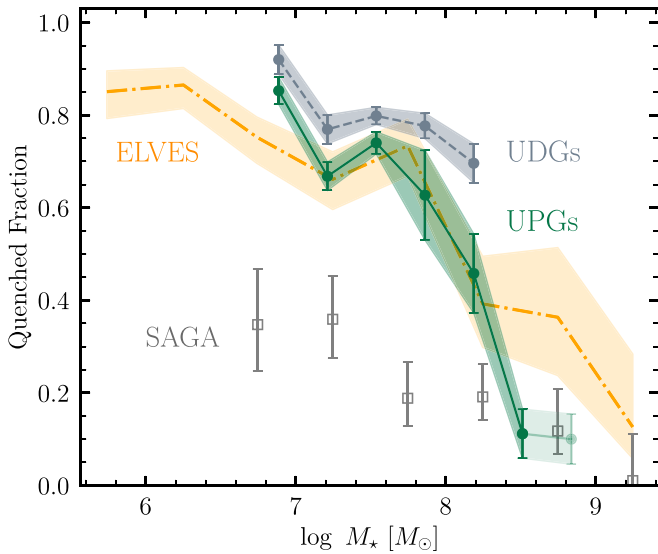
Figure 4. Quenched fractions of UDGs (gray) and UPGs (green) as a function of their stellar masses. The results from ELVES (Carlsten et al. 2022a) and SAGA (Mao et al. 2021) are also shown for comparison, where the quenched fractions are all defined based on the mass-dependent color cut and corrected for incompleteness following Mao et al. (2021). We find that UDGs have a high quenched fraction that changes little over a wide range of stellar mass. On the contrary, the quenched fraction of UPGs is lower and decreases with increasing satellite stellar mass. Given their large sizes, UPGs have a similar quenched fraction to the normal satellites of MW analogs in ELVES.

the 1σ Bernoulli standard error. We plot the UPG quenched fraction above $M_\star > 10^{8.5} M_\odot$ with lighter colors to emphasize where we extrapolate the mass–size relation in Carlsten et al. (2021) to define UPGs. The quenched fraction of UDGs is high ($f_q \sim 0.8$) and weakly depends on the stellar mass. Goto et al. (2023) also found a high quenched fraction for UDGs among the satellites of MW analogs in the SMUDGes survey. On the contrary, for UPGs, the quenched fraction is lower than that of UDGs and decreases more rapidly with increasing stellar mass. Lower-mass UPGs ($M_\star \sim 10^7 M_\odot$) are mostly quiescent, roughly half of the intermediate-mass UPGs ($M_\star \approx 10^8 M_\odot$) are quenched, and higher-mass UPGs ($M_\star > 10^{8.5} M_\odot$) are mostly star-forming. Such distinctions in quenched fractions highlight that UDGs and UPGs select different subsets of the satellite population. In Section 5.1, we argue that the high quenched fraction of UDGs is merely an artifact of the UDG definition.

We compare the quenched fractions of mass–size outliers (UDGs and UPGs) with those of the satellites of MW analogs in the Local Volume (ELVES; orange line) and the nearby Universe (SAGA; gray squares). Following Carlsten et al. (2022a), the quenched fractions of ELVES and SAGA are calculated based on the same color cut as we use for mass–size outliers.[10] Thus, the quenched fraction of SAGA in Figure 4 is slightly lower than the values in Figure 11 of Mao et al. (2021), where their quenched fraction is defined based on the detection of Hα. The quenched fractions of SAGA and ELVES are all corrected for incompleteness. We note that the error bars of the SAGA quenched fraction in Figure 4 are smaller than the ones in Mao et al. (2021). The tip of the light green error bar in

Figure 11 of Mao et al. (2021) is derived assuming that all potential satellites without a redshift detection are real quenched satellites. In our case, we adopt the satellite probability for each redshift failure from Mao et al. (2021) and then apply a color cut to both spectroscopically confirmed and nonconfirmed dwarfs, resulting in a smaller upper bound of the quenched fraction. Overall, the SAGA quenched fraction is lower than that of ELVES and the MW. As suggested by Carlsten et al. (2022a), the difference in both luminosity function and quenching could be naturally explained if the SAGA sample is preferentially missing ~1–1.5 satellites per host, and quiescent galaxies are fainter than star-forming ones at a fixed stellar mass (see also Font et al. 2022; Greene et al. 2023).

From Figure 4, the quenched fraction of UDGs is higher than that of normal satellites in ELVES and SAGA, especially at the high-mass end ($M_\star \sim 10^8 M_\odot$). However, the quenched fraction of UPGs agrees with ELVES reasonably well at $7 < \log M_\star/M_\odot < 8.2$. At the higher-mass end ($\log M_\star/M_\odot > 8.4$), the UPG quenched fraction agrees with SAGA but is lower than ELVES, with a caveat that the mass–size relation we use might not be valid in this mass range. Overall, it is surprising that UPGs, being mass–size outliers, have a similar quenched fraction as the normal-sized satellites of MW analogs.

We further divide the sample based on the projected radial distance to the host $R/R_{\rm vir}$ and the host $g - i$ color. As shown in the top panels of Figure 5, we find that UDGs (UPGs) that are closer to their hosts have higher quenched fractions. The average difference in quenched fractions between the two radial bins is ~0.1 for UDGs and ~0.2 for UPGs. This trend agrees with the findings in Greene et al. (2023) and Karunakaran & Zaritsky (2023) for normal satellites in the ELVES survey. From the bottom panels of Figure 5, we find that UDGs (UPGs) hosted by redder hosts also have a higher quenched fraction. The difference in quenched fractions between the two host color bins is ~0.2 for UDGs. But the host color has a smaller impact than the distance to the host on the quenched fraction of UPGs. Additionally, we divide the samples based on the host stellar mass and redshift but do not find a significant difference in quenched fractions.

As a sanity check, we select UPGs in the ELVES sample and calculate the quenched fraction, shown as the brown line in the left panel of Figure 6. Out of the 313 ELVES satellites that have secure distances around hosts that are mapped out to $R > 200$ kpc, there are 27 UPGs (see Greene et al. 2023). Despite the number of UPGs in ELVES being small, the quenched fraction of ELVES UPGs agrees with that of our UPG sample (green line) quite well. We also do not find a significant difference between the quenched fraction of the whole ELVES sample from that of ELVES UPGs.

We compare the quenched fraction of UPGs with the simulation results from ARTEMIS (purple lines in Figure 6; Font et al. 2022) and FIRE-2 (pink lines in Figure 6; Samuel et al. 2022), which focus on studying normal-sized satellites of MW analogs. The quenched galaxies in these simulations are defined as those with no instantaneous (for ARTEMIS) or recent (<200 Myr for FIRE-2) star formation.[11] We select ARTEMIS and FIRE-2 among other simulations because they roughly cover the highest and lowest quenched fraction for

---

[10] We note that some galaxies in the ELVES and SAGA data only have $g - r$ color because certain surveys might not have i-band data. Carlsten et al. (2022a) converted $g - r$ color to $g - i$ following their Equation (1) in Carlsten et al. (2021) that is derived from simple stellar population models.

[11] Font et al. (2022) found a similar quenched fraction for ARTEMIS when using the same color cut as we do.
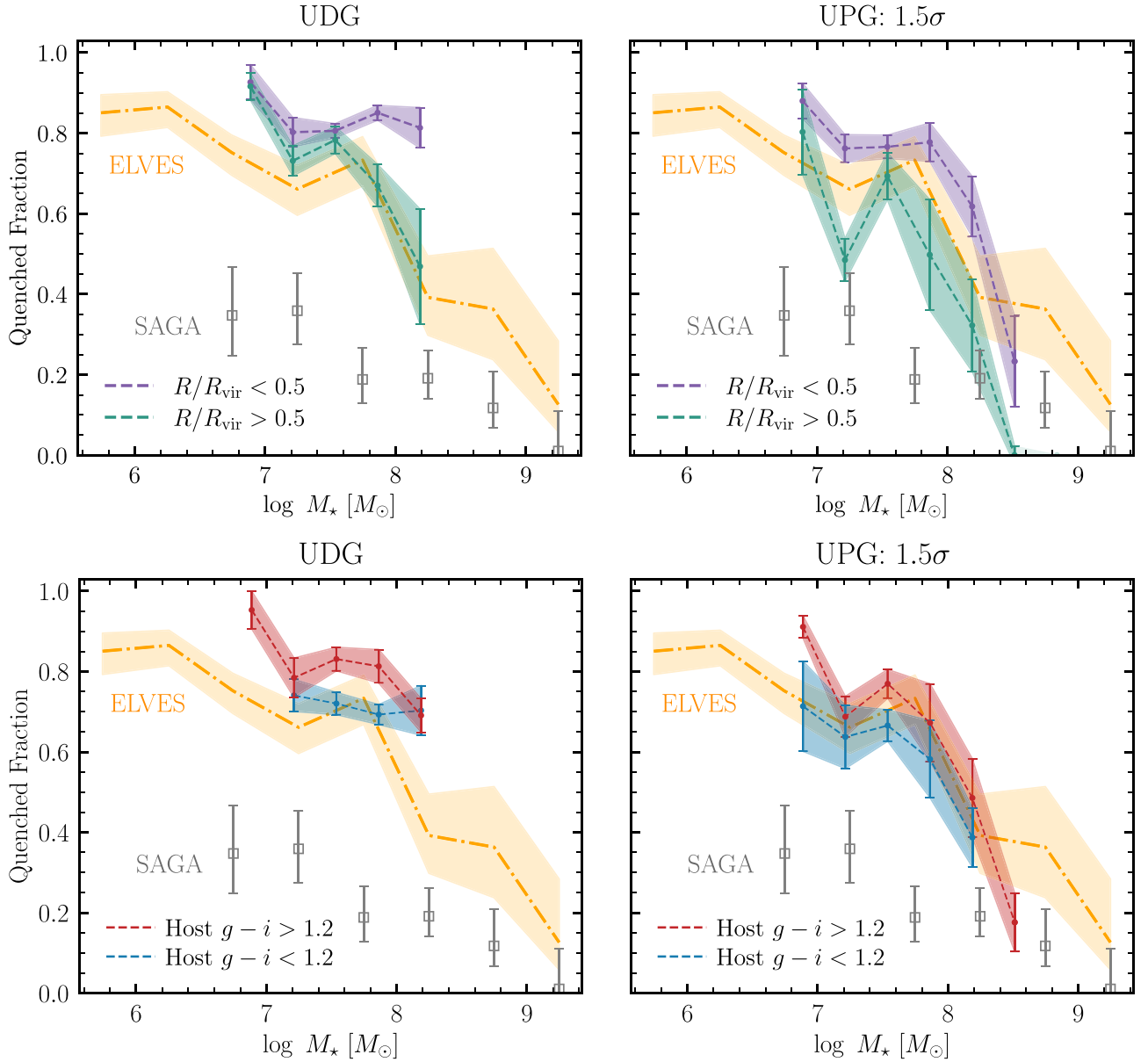
Figure 5. Same as Figure 4, but the samples are divided into subsets based on the projected distance to the host R/R_{vir} (top panels) and host $g - i$ color (bottom panels), shown as dashed red and blue lines. The quenched fraction is larger for those associated with redder hosts and those that are closer to the host.

MW satellites produced in simulations (see Figure 13 in Samuel et al. 2022 for a compilation). Since none of these data sets are designed to study mass–size outliers, their quenched fractions are presumably dominated by normal-sized satellites. As shown in the right panel of Figure 6, our UPG quenched fraction agrees best with ARTEMIS but is also consistent with the FIRE-2 result. We discuss the implications of these results in Section 5.2.

5. Discussion

5.1. The High Quenched Fraction of UDGs Is an Artifact

In L22, we proposed the concept of ultra-puffy galaxies (UPGs) as a means to robustly study mass–size outliers. Compared with the UDG selection, UPGs are not selected via a constant cut in half-light radius. The UPG selection considers the fact that the average sizes of satellite galaxies increase with stellar mass following a mass–size relation. There are many

UDGs that are not outliers with respect to the mass–size relation. Thus, the UPGs better represent the tail of the satellite size distribution. Moreover, for a given stellar mass and size, blue galaxies have brighter surface brightness than red galaxies because red galaxies have a higher mass-to-light ratio. As a consequence, blue galaxies will be preferentially excluded if one applies a hard cut on surface brightness, as in the definition of UDG. As seen in Figure 3, the UDG sample includes a number of galaxies below the 1.5σ line, and these galaxies are mostly red and quenched, therefore giving rise to a high quenched fraction in Figure 4. This high quenched fraction of UDGs compared to normal-sized dwarfs is just an artifact of the UDG definition. The hard size cut and surface brightness cut make it difficult to directly compare the red and blue galaxy populations and hamper the study of diffuse galaxies. The cut at 1.5σ above the average mass–size relation does not discriminate blue or red galaxies but just selects size outliers at a given stellar mass. The UPGs are physically motivated and
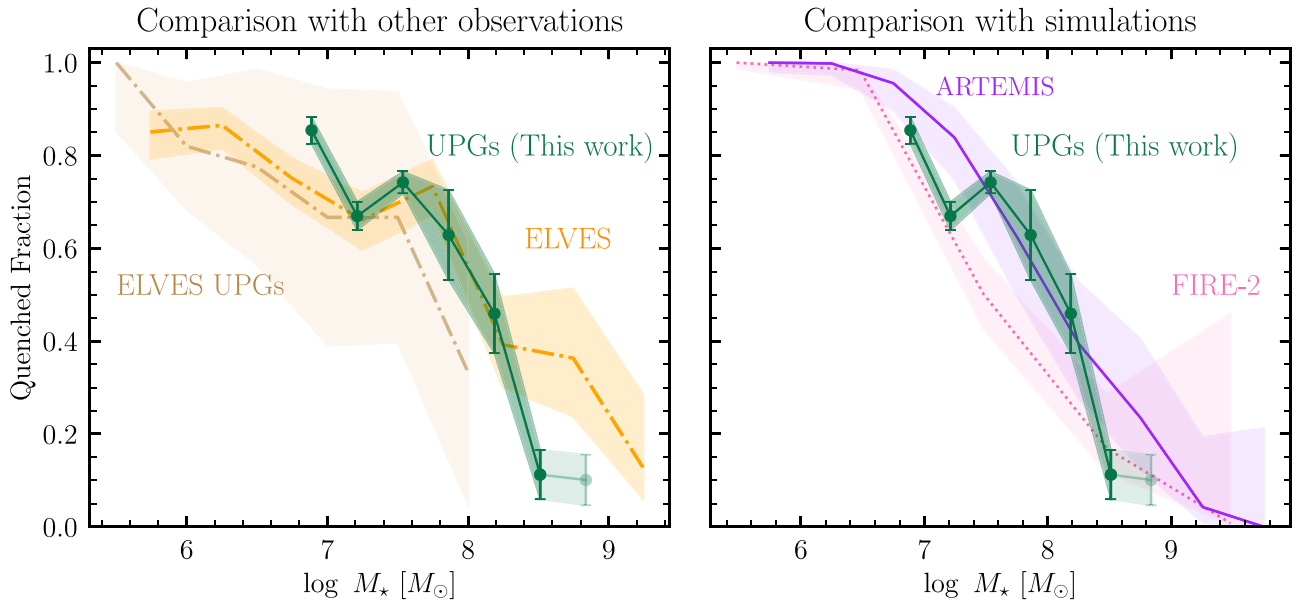
Figure 6. Left: quenched fractions of UPGs in this work and the ELVES survey. We select UPGs from the ELVES sample and calculate the quenched fraction, shown as the brown line. The quenched fraction of UPGs in this work is highlighted in green, while the quenched fraction of the whole ELVES sample is shown in orange. We do not find a significant dependence of the quenched fraction on the size of the satellites by comparing UPGs with the bulk of the satellites. Right: comparisons with simulations of satellites in MW analogs. The purple and pink lines show the quenched fractions produced in the ARTEMIS (Font et al. 2022) and FIRE-2 (Samuel et al. 2022) simulations, respectively. The quenched fraction of UPGs agrees quite well with that of normal-sized satellites in simulations.

naturally represent the large-size tail of satellite galaxies in MW analogs.

We note that Lim et al. (2020) defined UDGs in the Virgo cluster as the outliers of the scaling relations of Virgo satellites. However, in clusters, the difference between such a definition and the fiducial UDG definition in van Dokkum et al. (2015) becomes minor because nearly all UDGs are red and quenched. Nevertheless, we advocate studying outliers of scaling relations, including UPGs, which provide an unbiased perspective on the formation and evolution of diffuse dwarf galaxies.

5.2. Formation and Quenching of Mass–Size Outliers

In this paper, we study the mass–size outliers associated with MW analogs in the nearby Universe. We calculate the quenched fraction of UPGs in L22 and compare it with the ELVES survey, the SAGA survey, and numerical simulations. One key finding is that UPGs have a similar quenched fraction as those normal-sized satellites of MW analogs in the Local Volume (ELVES; Figure 4) and numerical simulations (ARTEMIS and FIRE-2; Figure 6). We also check the quenched fraction of UPGs among ELVES satellites and find similar results (Figure 6). In addition, the radial distribution of mass–size outliers can be well described by an NFW (or Einasto) profile at $R > 0.2R_{\rm vir}$. Thus, it is intriguing that despite being selected as outliers in size, UPGs are acting like "normal" satellites in terms of quenched fraction and radial distribution. Furthermore, Carlsten et al. (2021) found that the mass–size relation of satellites in MW analogs does not depend on the morphology or color of the satellites. All of these pieces of evidence suggest that, for satellites in MW analogs, quenching is not tied to being a mass–size outlier. Quenching and morphological transformation might involve very mild size evolution.

We are left with the question of what physical processes are responsible for puffing up satellites and quenching them such

that quenching and size growth happen separately. If the mass–size outliers were normal-sized before falling into the group, then tidal interactions are believed to puff them up. Jiang et al. (2019) found that in simulations, the UDGs in groups are accreted as either UDGs from the field or normal dwarfs but tidally heated near the orbital pericenter. Tremmel et al. (2020) studied simulated UDGs in the cluster environment and found that UDGs with higher mass ($M_\star > 10^8\, M_\odot$) are puffed up suddenly due to tidal heating at the pericenter, but UDGs with lower mass ($M_\star \sim 10^{7.5}\, M_\odot$) are gradually puffed up due to adiabatic expansion as a response to the mass loss from tidal and ram pressure stripping. In this scenario, normal-sized star-forming dwarf galaxies are accreted from the field, then consequently puffed up by strong tidal interactions. If so, we would expect a higher quenched fraction for mass–size outliers, since they must have undergone violent ram pressure stripping near the pericenter or lost gas due to tidal stripping. This contradicts our results in Figure 4.

On the other hand, the mass–size outliers can be puffed up when they are still in the field. For example, bursty stellar feedback can cause the stellar component of the dwarf galaxy to expand (e.g., Di Cintio et al. 2017; Chan et al. 2018; Carleton et al. 2019; Jiang et al. 2019; Martin et al. 2019). Analytical models also support a scenario where UDGs originate from a population of dwarf galaxies residing in halos with higher spin (Dalcanton et al. 1997; Amorisco & Loeb 2016; Rong et al. 2017; Liao et al. 2019). Wright et al. (2021) showed that field UDGs can also be formed from major mergers that cause star formation to migrate outward. If mass–size outliers are already puffed up prior to infall, in order to achieve a similar quenched fraction as normal satellites, they must remain star-forming after being puffed up such that they can be quenched together with normal satellites. Wright et al. (2021) showed that an early merger does not significantly change the total SFR. Samuel et al. (2022) showed that preprocessing can boost the quenched fraction by 20% at

$M_\star \approx 10^8 M_\odot$ and merely change the quenched fraction for lower-mass satellites. Although it is not clear whether the high halo spin is the cause or the consequence of UPG formation, it is probable that a higher halo spin does not quench a UPG. If any of these mechanisms can only increase the size of normal dwarfs to make UPGs but not quench them when they are still in the field, it provides a plausible way to produce the observed trend in Figures 4 and 6.

Mass–size outliers and normal satellites might be quenched together after falling into the group. It is not clear whether they are quenched mainly by tidal or ram pressure stripping. In the tidal stripping scenario, the dark matter needs to be removed before a significant fraction of stars and/or gas can be stripped by tidal forces, indicating a longer quenching timescale. Jiang et al. (2019) argued that if tidal stripping dominates the quenching of UDGs, the satellites closer to the host will have a smaller stellar mass and effective radius, since stars are also stripped. They found no such trend in simulations of group environments. In Section 3.1, we split the UDG/UPG samples into two radial distance bins but do not find a significant change in size or stellar mass distribution as the mass–size outliers get closer to the host. This suggests that tidal stripping does not dominate the quenching of UDGs and UPGs in MW analogs.

Many studies support the idea that ram pressure stripping is the dominant quenching mechanism for normal-sized satellites with $M_\star < 10^{8.0} M_\odot$ (e.g., Tonnesen & Bryan 2009; Jaffé et al. 2015; Simpson et al. 2018; Akins et al. 2021; Wang et al. 2021; Samuel et al. 2022). In the Auriga simulations, for instance, Simpson et al. (2018) showed that the quenched fraction of satellites $M_\star < 10^8 M_\odot$ have a strong dependence on the distance to the hosts, and they considered ram pressure stripping as the main quenching mechanism. Using the FIRE-2 simulation, Samuel et al. (2022) also showed that the quenched fraction increases as the distance to the host decreases, and hosts with higher CGM mass have more quenched satellites. Such trends are also found in observations (e.g., Greene et al. 2023; Karunakaran et al. 2023). In Figure 4, by comparing UDGs and UPGs at two radial bins, we find a very similar trend that satellites that are closer to the host are more quiescent. Although we are not able to derive a quenching timescale to better compare with simulations, we consider our results to be consistent with the ram pressure stripping scenario.

In Figure 5, we also find that UDGs and UPGs associated with redder hosts are more quiescent than those in bluer hosts, although the trend is less significant for UPGs. This trend agrees with "galaxy conformity," which refers to the excess of early-type satellites in the vicinity of early-type central galaxies compared to satellites around late-type centrals (Weinmann et al. 2006). Wang & White (2012) argued that galaxy conformity arises because quiescent centrals occupy more massive halos than star-forming centrals. More massive halos have more hot gas and stronger tidal fields, leading to more efficient quenching. If quiescent centrals are in older halos, their satellites tend to be accreted earlier and thus have more time to be quenched. Based on such models, the galaxy conformity signals from UDGs and UPGs make sense if they are puffed up before the infall and quenched together with normal satellites due to ram pressure stripping.

The efficiency of ram pressure stripping also depends on the stellar mass density, which provides the restoring force and the gas mass density. The diffuse stellar component might make mass–size outliers more vulnerable to ram pressure and tidal stripping. However, Kado-Fong et al. (2022) found in observations that although UDGs have a more diffuse stellar component, they harbor the same amount of H I gas as normal dwarfs with similar stellar sizes. Equivalently, for a given stellar mass, UDGs (UPGs) have higher H I mass than normal dwarfs. Thus, even if UDGs and UPGs are more vulnerable to ram pressure stripping, they have more gas to be stripped. It is probable that the net effect of low stellar mass density and high H I mass is that UPGs have a similar quenched fraction as normal satellites. Of course, the origin and evolution of the mass–size outliers in MW analogs could also be a mixture of many scenarios. Using our observations, it is hard to sort out the detailed mechanisms for quenching and size growth. More detailed observations and simulations are needed to solve this intriguing puzzle.

6. Summary

In this work, we present a statistical analysis of mass–size outliers among the satellite galaxies of MW analogs at $0.01 < z < 0.04$ based on the sample from L22. Besides the ultra-diffuse galaxies (UDGs), we study "ultra-puffy galaxies" (UPGs), which are defined to lie 1.5σ above the average mass–size relation and better represent the large-size tail of the dwarf galaxy population. We derive the size distribution (Section 3.1), radial distribution (Section 3.2), and quenched fractions (Section 4) of mass–size outliers (UDGs and UPGs) and discuss the implications on their formation and evolution (Section 5). We summarize our main findings and prospects as follows.

1. The size distribution of UDGs follows a power law $dn/d \log r_e \propto r_e^{-1.45\pm0.20}$, which is shallower than the UDG size distribution in larger groups and clusters (e.g., van der Burg et al. 2016, 2017). Since the UPG sample includes fewer small satellites, its size distribution is suppressed at the smaller-size end (Figure 1). We do not find a significant dependence of size distribution on the host stellar mass or the distance to the host.

2. The radial distributions of UDGs and UPGs can be described by projected NFW and Einasto profiles (Figure 2). The UDGs and UPGs follow similar radial distributions. We find that the best-fit NFW profiles have concentrations similar to that of an MW-mass halo. The best-fit Einasto profiles are less concentrated than MW-mass halos.

3. The UDGs have a very high quenched fraction ($\sim$70%) that remains roughly constant over 1 dex in satellite stellar mass (Figure 4). However, for UPGs, the quenched fraction shows a strong dependence on satellite stellar mass; 80% of UPGs are quiescent at the lower-mass end, $M_\star = 10^{6.8} M_\odot$, but the fraction drops to 20% at $M_\star = 10^{8.5} M_\odot$. Almost half of the UPGs are quenched at $M_\star \approx 10^8 M_\odot$. The quenched fractions of both UDGs and UPGs increase as satellites get closer to the hosts, agreeing with the trends of normal-sized satellites (Greene et al. 2023; Karunakaran et al. 2023). Redder hosts also harbor more quenched mass–size outliers, showing evidence for "galaxy conformity."

4. Surprisingly, although UPGs have a much larger size than "normal" satellites of similar stellar mass, they show a

similar quenched fraction as the normal-sized satellites in MW analogs (Figure 6). The UPG quenched fraction is consistent with the results from both observations (ELVES) and simulations (ARTEMIS, FIRE-2) for normal-sized satellites. Such a null trend of the quenched fraction on the size of satellites indicates that the quenching and size evolution of mass–size outliers happen separately in MW-like environments. It is plausible that these mass–size outliers are puffed up in the field before accretion and quenched together with normal satellites (Section 5.2).

5. We argue that the high and constant quenched fraction for UDGs is merely an artifact of the UDG definition, where the constant surface brightness cut preferentially selects more red galaxies (Figure 3). We demonstrate that the more physically motivated UPG selection does not introduce artifacts in the mass–size distribution and better represents the large-size tail of the dwarf galaxy population.

Colossus (see footnote 8; Diemer 2018), scarlet[18] (Melchior et al. 2018), unagi.[19]

ORCID iDs

Jiaxuan Li (李嘉轩) ⓘ https://orcid.org/0000-0001-9592-4190
Jenny E. Greene ⓘ https://orcid.org/0000-0002-5612-3427
Johnny Greco ⓘ https://orcid.org/0000-0003-4970-2874
Rachael Beaton ⓘ https://orcid.org/0000-0002-1691-8217
Shany Danieli ⓘ https://orcid.org/0000-0002-1841-2252
Andy Goulding ⓘ https://orcid.org/0000-0003-4700-663X
Song Huang (黄崧) ⓘ https://orcid.org/0000-0003-1385-7591
Erin Kado-Fong ⓘ https://orcid.org/0000-0002-0332-177X

References

Aihara, H., Arimoto, N., Armstrong, R., et al. 2018, PASJ, 70, S4
Akins, H. B., Christensen, C. R., Brooks, A. M., et al. 2021, ApJ, 909, 139
Amorisco, N. C., & Loeb, A. 2016, MNRAS, 459, L51
Applebaum, E., Brooks, A. M., Christensen, C. R., et al. 2021, ApJ, 906, 96
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Baxter, D. C., Cooper, M. C., & Fillingham, S. P. 2021, MNRAS, 503, 1636
Behroozi, P. S., Conroy, C., & Wechsler, R. H. 2010, ApJ, 717, 379
Benavides, J. A., Sales, L. V., Abadi, M. G., et al. 2021, NatAs, 5, 1255
Benson, A. J., Lacey, C. G., Baugh, C. M., Cole, S., & Frenk, C. S. 2002, MNRAS, 333, 156
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Blanton, M. R., Kazin, E., Muna, D., Weaver, B. A., & Price-Whelan, A. 2011, AJ, 142, 31
Blanton, M. R., Lupton, R. H., Schlegel, D. J., et al. 2005, ApJ, 631, 208
Bryan, G. L., & Norman, M. L. 1998, ApJ, 495, 80
Buck, T., Macciò, A. V., Dutton, A. A., Obreja, A., & Frings, J. 2019, MNRAS, 483, 1314
Budzynski, J. M., Koposov, S. E., McCarthy, I. G., McGee, S. L., & Belokurov, V. 2012, MNRAS, 423, 104
Bullock, J. S., & Boylan-Kolchin, M. 2017, ARA&A, 55, 343
Bullock, J. S., Kolatt, T. S., Sigad, Y., et al. 2001, MNRAS, 321, 559
Bullock, J. S., Kravtsov, A. V., & Weinberg, D. H. 2000, ApJ, 539, 517
Carleton, T., Errani, R., Cooper, M., et al. 2019, MNRAS, 485, 382
Carlsten, S. G., Greene, J. E., Beaton, R. L., Danieli, S., & Greco, J. P. 2022a, ApJ, 933, 47
Carlsten, S. G., Greene, J. E., Beaton, R. L., & Greco, J. P. 2022b, ApJ, 927, 44
Carlsten, S. G., Greene, J. E., Greco, J. P., Beaton, R. L., & Kado-Fong, E. 2021, ApJ, 922, 267
Carlsten, S. G., Greene, J. E., Peter, A. H. G., Greco, J. P., & Beaton, R. L. 2020, ApJ, 902, 124
Chabrier, G. 2003, PASP, 115, 763
Chan, T. K., Kereš, D., Wetzel, A., et al. 2018, MNRAS, 478, 906
Dalcanton, J. J., Spergel, D. N., Gunn, J. E., Schmidt, M., & Schneider, D. P. 1997, AJ, 114, 635
Di Cintio, A., Brook, C. B., Dutton, A. A., et al. 2017, MNRAS, 466, L1
Diemer, B. 2018, ApJS, 239, 35
Diemer, B., & Joyce, M. 2019, ApJ, 871, 168
Duffy, A. R., Schaye, J., Kay, S. T., & Dalla Vecchia, C. 2008, MNRAS, 390, L64
Dutton, A. A., & Macciò, A. V. 2014, MNRAS, 441, 3359
Einasto, J. 1965, TrAlm, 5, 87
El-Badry, K., Quataert, E., Wetzel, A., et al. 2018, MNRAS, 473, 1930
Emerick, A., Mac Low, M.-M., Grcevich, J., & Gatto, A. 2016, ApJ, 826, 148
Fillingham, S. P., Cooper, M. C., Pace, A. B., et al. 2016, MNRAS, 463, 1916
Font, A. S., McCarthy, I. G., Belokurov, V., Brown, S. T., & Stafford, S. G. 2022, MNRAS, 511, 1544
Gao, L., Navarro, J. F., Cole, S., et al. 2008, MNRAS, 387, 536
Garrison-Kimmel, S., Wetzel, A., Hopkins, P. F., et al. 2019, MNRAS, 489, 4574
Geha, M., Blanton, M. R., Yan, R., & Tinker, J. L. 2012, ApJ, 757, 85
Geha, M., Wechsler, R. H., Mao, Y.-Y., et al. 2017, ApJ, 847, 4

Goto, H., Zaritsky, D., Karunakaran, A., Donnerstein, R., & Sand, D. J. 2023, arXiv:2303.00774
Grcevich, J., & Putman, M. E. 2009, ApJ, 696, 385
Greco, J. P., Greene, J. E., Strauss, M. A., et al. 2018, ApJ, 857, 104
Greene, J. E., Danieli, S., Carlsten, S., et al. 2023, ApJ, 949, 94
Gunn, J. E., & Gott, J. R. I. 1972, ApJ, 176, 1
Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Natur, 585, 357
Hearin, A. P., Campbell, D., Tollerud, E., et al. 2017, AJ, 154, 190
Hunter, J. D. 2007, CSE, 9, 90
Jaffé, Y. L., Smith, R., Candlish, G. N., et al. 2015, MNRAS, 448, 1715
Jahn, E. D., Sales, L. V., Wetzel, A., et al. 2022, MNRAS, 513, 2673
Jiang, F., Dekel, A., Freundlich, J., et al. 2019, MNRAS, 487, 5272
Jones, E., Oliphant, T., Peterson, P., et al. 2001, SciPy: Open Source Scientific Tools for Python, http://www.scipy.org/
Joshi, G. D., Pillepich, A., Nelson, D., et al. 2021, MNRAS, 508, 1652
Kado-Fong, E., Greene, J. E., Huang, S., & Goulding, A. 2022, ApJ, 941, 11
Karunakaran, A., Sand, D. J., Jones, M. G., et al. 2023, MNRAS, 524, 5314
Karunakaran, A., Spekkens, K., Carroll, R., et al. 2022, MNRAS, 516, 1741
Karunakaran, A., Spekkens, K., Oman, K. A., et al. 2021, ApJL, 916, L19
Karunakaran, A., & Zaritsky, D. 2023, MNRAS, 519, 884
Kravtsov, A. V., Gnedin, O. Y., & Klypin, A. A. 2004, ApJ, 609, 482
Li, J., Greene, J. E., Greco, J. P., et al. 2022a, arXiv:2210.14994
Li, J., Huang, S., Leauthaud, A., et al. 2022b, MNRAS, 515, 5335
Liao, S., Gao, L., Frenk, C. S., et al. 2019, MNRAS, 490, 5182
Lim, S., Côté, P., Peng, E. W., et al. 2020, ApJ, 899, 69
Mancera Piña, P. E., Peletier, R. F., Aguerri, J. A. L., et al. 2018, MNRAS, 481, 4381
Mao, Y.-Y., Geha, M., Wechsler, R. H., et al. 2021, ApJ, 907, 85
Martin, G., Kaviraj, S., Laigle, C., et al. 2019, MNRAS, 485, 796
McCarthy, I. G., Frenk, C. S., Font, A. S., et al. 2008, MNRAS, 383, 593
McDonough, B., & Brainerd, T. G. 2022, ApJ, 933, 161
Melchior, P., Moolekamp, F., Jerdee, M., et al. 2018, A&C, 24, 129
Miyazaki, S., Komiyama, Y., Kawanomoto, S., et al. 2018, PASJ, 70, S1
Miyazaki, S., Komiyama, Y., Nakaya, H., et al. 2012, Proc. SPIE, 8446, 8446
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493
Navarro, J. F., Hayashi, E., Power, C., et al. 2004, MNRAS, 349, 1039
Navarro, J. F., Ludlow, A., Springel, V., et al. 2010, MNRAS, 402, 21
Oke, J. B., & Gunn, J. E. 1983, ApJ, 266, 713
Pan, Y., Simpson, C. M., Kravtsov, A., et al. 2023, MNRAS, 519, 4499

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A13
Putman, M. E., Zheng, Y., Price-Whelan, A. M., et al. 2021, ApJ, 913, 53
Rodriguez-Gomez, V., Snyder, G. F., Lotz, J. M., et al. 2019, MNRAS, 483, 4140
Rong, Y., Guo, Q., Gao, L., et al. 2017, MNRAS, 470, 4231
Sales, L. V., Navarro, J. F., Lambas, D. G., White, S. D. M., & Croton, D. J. 2007, MNRAS, 382, 1901
Sales, L. V., Wetzel, A., & Fattahi, A. 2022, NatAs, 6, 897
Samuel, J., Wetzel, A., Santistevan, I., et al. 2022, MNRAS, 514, 5276
Samuel, J., Wetzel, A., Tollerud, E., et al. 2020, MNRAS, 491, 1471
Simons, R. C., Peeples, M. S., Tumlinson, J., et al. 2020, ApJ, 905, 167
Simpson, C. M., Grand, R. J. J., Gómez, F. A., et al. 2018, MNRAS, 478, 548
Somerville, R. S. 2002, ApJL, 572, L23
Spekkens, K., Urbancic, N., Mason, B. S., Willman, B., & Aguirre, J. E. 2014, ApJL, 795, L5
Tal, T., Wake, D. A., & van Dokkum, P. G. 2012, ApJL, 751, L5
Tanoglidis, D., Drlica-Wagner, A., Wei, K., et al. 2021, ApJS, 252, 18
Tollerud, E. J., & Peek, J. E. G. 2018, ApJ, 857, 45
Tonnesen, S., & Bryan, G. L. 2009, ApJ, 694, 789
Tremmel, M., Wright, A. C., Brooks, A. M., et al. 2020, MNRAS, 497, 2786
van den Bosch, F. C., Ogiya, G., Hahn, O., & Burkert, A. 2018, MNRAS, 474, 3043
van den Bosch, F. C., Yang, X., Mo, H. J., & Norberg, P. 2005, MNRAS, 356, 1233
van der Burg, R. F. J., Hoekstra, H., Muzzin, A., et al. 2017, A&A, 607, A79
van der Burg, R. F. J., Muzzin, A., & Hoekstra, H. 2016, A&A, 590, A20
van Dokkum, P. G., Abraham, R., Merritt, A., et al. 2015, ApJL, 798, L45
Wang, J., Staveley-Smith, L., Westmeier, T., et al. 2021, ApJ, 915, 70
Wang, W., Sales, L. V., Henriques, B. M. B., & White, S. D. M. 2014, MNRAS, 442, 1363
Wang, W., & White, S. D. M. 2012, MNRAS, 424, 2574
Weinmann, S. M., van den Bosch, F. C., Yang, X., & Mo, H. J. 2006, MNRAS, 366, 2
Wetzel, A. R., Deason, A. J., & Garrison-Kimmel, S. 2015a, ApJ, 807, 49
Wetzel, A. R., Tollerud, E. J., & Weisz, D. R. 2015b, ApJL, 808, L27
Wright, A. C., Tremmel, M., Brooks, A. M., et al. 2021, MNRAS, 502, 5370
Zaritsky, D., Donnerstein, R., Karunakaran, A., et al. 2021, ApJS, 257, 60
Zhu, J., & Putman, M. E. 2023, MNRAS, 521, 3765