

University of Louisville

## ThinkIR: The University of Louisville's Institutional Repository

---

Faculty Scholarship

---

4-2017

### Diverse stellar haloes in nearby Milky Way mass disc galaxies.

Benjamin Harmsen

*University of Michigan-Ann Arbor*

Antonela Monachesi

*Max Planck Institut fur Astrophysik*

Eric F. Bell

*University of Michigan-Ann Arbor*

Roelof S. de Jong

*Leibniz-Institut fur Astrophysik Potsdam*

Jeremy Bailin

*University of Alabama - Tuscaloosa*

*See next page for additional authors*

Follow this and additional works at: <https://ir.library.louisville.edu/faculty>



Part of the [Astrophysics and Astronomy Commons](#)

---

#### Original Publication Information

Harmsen, Benjamin, et al. "Diverse Stellar Haloes in Nearby Milky Way Mass Disc Galaxies." 2017. *Monthly Notices of the Royal Astronomical Society* 466(2): 1491-1512.

This Article is brought to you for free and open access by ThinkIR: The University of Louisville's Institutional Repository. It has been accepted for inclusion in Faculty Scholarship by an authorized administrator of ThinkIR: The University of Louisville's Institutional Repository. For more information, please contact [thinkir@louisville.edu](mailto:thinkir@louisville.edu).

---

**Authors**

Benjamin Harmsen, Antonela Monachesi, Eric F. Bell, Roelof S. de Jong, Jeremy Bailin, David J. Radburn-Smith, and Benne W. Holwerda

# Diverse stellar haloes in nearby Milky Way mass disc galaxies

Benjamin Harmsen,<sup>1★</sup> Antonela Monachesi,<sup>2★</sup> Eric F. Bell,<sup>1★</sup> Roelof S. de Jong,<sup>3</sup>  
Jeremy Bailin,<sup>4,5</sup> David J. Radburn-Smith<sup>6</sup> and Benne W. Holwerda<sup>7</sup>

<sup>1</sup>Department of Astronomy, University of Michigan, 311 West Hall, 1085 South University Ave., Ann Arbor, MI 48109-1107, USA

<sup>2</sup>Max Planck Institut für Astrophysik, Karl-Schwarzschild-Str 1, Postfach 1317, D-85741 Garching, Germany

<sup>3</sup>Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, D-14482 Potsdam, Germany

<sup>4</sup>Department of Physics and Astronomy, University of Alabama, Box 870324, Tuscaloosa, AL 35487-0324, USA

<sup>5</sup>National Radio Astronomy Observatory, PO Box 2, Green Bank, WV 24944, USA

<sup>6</sup>Department of Astronomy, University of Washington, 3910 15th Ave NE, Seattle, WA 98195, USA

<sup>7</sup>Leiden Observatory, Niels Bohrweg 2, NL-2300 CA Leiden, the Netherlands

Accepted 2016 November 15. Received 2016 November 15; in original form 2016 June 21

## ABSTRACT

We have examined the resolved stellar populations at large galactocentric distances along the minor axis (from 10 kpc up to between 40 and 75 kpc), with limited major axis coverage, of six nearby highly inclined Milky Way (MW) mass disc galaxies using *Hubble Space Telescope* data from the Galaxy haloes, Outer discs, Substructure, Thick discs, and Star clusters (GHOSTS) survey. We select red giant branch stars to derive stellar halo density profiles. The projected minor axis density profiles can be approximated by power laws with projected slopes of  $-2$  to  $-3.7$  and a diversity of stellar halo masses of  $1-6 \times 10^9 M_{\odot}$ , or 2–14 per cent of the total galaxy stellar masses. The typical intrinsic scatter around a smooth power-law fit is 0.05–0.1 dex owing to substructure. By comparing the minor and major axis profiles, we infer projected axis ratios  $c/a$  at  $\sim 25$  kpc between 0.4 and 0.75. The GHOSTS stellar haloes are diverse, lying between the extremes charted out by the (rather atypical) haloes of the MW and M31. We find a strong correlation between the stellar halo metallicities and the stellar halo masses. We compare our results with cosmological models, finding good agreement between our observations and accretion-only models where the stellar haloes are formed by the disruption of dwarf satellites. In particular, the strong observed correlation between stellar halo metallicity and mass is naturally reproduced. Low-resolution hydrodynamical models have unrealistically high stellar halo masses. Current high-resolution hydrodynamical models appear to predict stellar halo masses somewhat higher than observed but with reasonable metallicities, metallicity gradients, and density profiles.

**Key words:** galaxies: evolution – galaxies: general – galaxies: haloes – galaxies: individual: NGC 253, NGC 891, NGC 3031, NGC 4565, NGC 4945, NGC 7814 – galaxies: stellar content.

## 1 INTRODUCTION

The current favoured cosmological model  $\Lambda$ -cold dark matter ( $\Lambda$ CDM) is hierarchical, predicting that dark matter haloes are assembled over time through the collisionless accretion and mergers of smaller haloes. Stars form in the centres of the larger dark matter haloes (White & Rees 1978; Moore et al. 1999), where the number of stars that form in low mass haloes is dramatically suppressed compared to larger haloes (likely owing to feedback from super-

novae; e.g. Dekel & Silk 1986; Cole 1991; Wheeler et al. 2014). As satellite haloes merge with the main halo, their (typically meager) stellar components tidally disrupt and are spread into a diffuse and structured stellar halo (Bullock, Kravtsov & Weinberg 2001; Bullock & Johnston 2005; Cooper et al. 2010). The resulting stellar haloes are expected to exhibit steep density profiles, have abundant substructure, and show considerable halo-to-halo variation in their properties, all of which are expected to be closely tied with their merger histories. The goal of this work is to carefully characterize the density profiles, projected axis ratios, stellar masses and substructure of the stellar haloes of six nearby roughly Milky Way (MW) mass disc galaxies with resolved stellar population measurements from the *Hubble Space Telescope* (HST; Radburn-Smith et al. 2011; Monachesi et al. 2016a).

\* E-mail: [benharms@umich.edu](mailto:benharms@umich.edu) (BH); [antonela@mpa-garching.mpg.de](mailto:antonela@mpa-garching.mpg.de) (AM); [ericbell@umich.edu](mailto:ericbell@umich.edu) (EFB)

While models in which stellar haloes are composed of the tidal debris from dwarf galaxy disruption alone share a number of broad qualitative predictions – diffuse, centrally concentrated, highly structured stellar haloes whose metallicities reflect the metallicities of the disrupted dwarf galaxies – some of the quantitative predictions for stellar halo properties vary considerably from model to model. Stellar halo masses and metallicities of MW mass disc galaxies vary considerably more from halo to halo in the Cooper et al. (2010) and Gómez et al. (2012) models than in Bullock & Johnston (2005), likely from a wider range of satellite accretion histories (the importance of input satellite metallicity distributions is emphasized by Tumlinson 2010). Stellar haloes in the  $N$ -body only models of Cooper et al. (2010) are triaxial, whereas haloes in Bullock & Johnston (2005) are oblate. Bailin et al. (2014) show that oblate haloes are the natural result of the growth of a stellar halo in a potential where the baryons are allowed to grow into a galaxy with a prominent disc, suggesting that the presence of a disc potential (incorporated in Bullock & Johnston 2005 but absent in Cooper et al. 2010) is a key driver of stellar halo oblateness. Furthermore, while the Bullock & Johnston (2005) models all lack a significant metallicity gradient, the metallicity gradients of the Cooper et al. (2010) models vary considerably, from no gradient to relatively rapid changes in metallicity with radius. This diversity in model predictions signals the strength of observations to test and guide the models.

In addition to the stars accreted from disrupted satellites, the inner parts of stellar haloes are predicted to have a considerable population of *in situ* stars formed in the main galaxy potential (Zolotov et al. 2009; Font et al. 2011; Tissera et al. 2014; Pillepich, Madau & Mayer 2015). The physical ingredients of such models have significant uncertainties, e.g. the modelling of feedback processes, stellar winds, star formation recipes, and gas dynamics; all of which have significant impact on how stars populate dwarf satellites, the shape of the potential of the main galaxy, the fraction of *in situ* stars, and even whether the *in situ* stars are a common feature of all haloes (Bailin et al. 2014). The mass and extent of expected *in situ* haloes are predicted to vary by large factors, ranging from being dominant at radii of even 30 kpc (e.g. Font et al. 2011) to being dominant only at <5 kpc (e.g. Pillepich et al. 2015). The prominence of *in situ* stars is expected to be a function of position with respect to the disc (more prominent along the major axis, and less detectable along the minor axis; Monachesi et al. 2016b), galaxy mass, and merger history (Zolotov 2011).

All of these considerations motivate the careful characterization of a sizeable sample of stellar haloes. Yet, owing to the observational challenge of detecting low surface brightness and diffuse features in more distant galaxies, the stellar populations and shapes of the main body of the haloes of only the MW and Andromeda have been studied in depth to date (e.g. Ivezić et al. 2000; Newberg et al. 2007; Bell et al. 2008; Gilbert et al. 2012; Deason et al. 2013; Gilbert et al. 2014; Ibata et al. 2014). While both haloes are richly substructured and qualitatively agree with the  $\Lambda$ CDM paradigm of galaxy formation, they display significant differences. The MW halo has a weak to no metallicity gradient (Sesar, Jurić & Ivezić 2011; Xue et al. 2015) and its stellar density distribution can be described by a broken power law – within 25–30 kpc, it follows an oblate,  $\rho \propto r^{-\gamma}$  power-law distribution with index  $\gamma \sim 2.5$ –3 (Yanny et al. 2000; Bell et al. 2008; Jurić et al. 2008; van Vledder et al. 2016) whereas a more rapidly declining stellar density is detected beyond  $\sim 30$  kpc, with  $\gamma \sim 3.5$  (Deason, Belokurov & Evans 2011; Sesar et al. 2011; Deason et al. 2014; Cohen, Sesar & Banholzer 2015; Slater et al. 2016). M31, on the other hand, has a clear metallicity gradient with a 1 dex

variation in [Fe/H] from 10 to  $\sim 100$  kpc (Gilbert et al. 2014; Ibata et al. 2014) and its stellar density distribution can be described by a single power law with  $\gamma \sim 3.3$  (Guhathakurta et al. 2005; Gilbert et al. 2012). In order to test model predictions and quantify the halo-to-halo variations such as differences in metallicity profiles, fraction of stellar halo created *in situ* and accreted, stellar halo morphology, etc., we need to observe the stellar halo properties of more similar mass galaxies. In particular, the stellar halo density profiles and shapes can provide important constraints on the merging and accretion history of a galaxy (Johnston et al. 2008; Deason et al. 2013; Amorisco 2017).

Over the last decades, a number of efforts have sought to characterize the diffuse stellar envelopes around galaxies. Integrated light studies of nearby galaxies show that stellar streams (thought to be from the disruption of dwarf galaxies) are reasonably common (Malin & Hadley 1997; Shang et al. 1998; Mihos et al. 2005; Tal et al. 2009; Martínez-Delgado et al. 2010a; Paudel et al. 2013; Watkins, Mihos & Harding 2015; Merritt et al. 2016). Such studies are challenging, requiring excellent control of scattered light (Slater, Harding & Mihos 2009; Abraham & van Dokkum 2014). Although it is often possible to control scattered light well enough to uncover tidal streams as local enhancements in surface brightness, it is extremely challenging to control scattered light well enough to convincingly and correctly recover the brightness profile of the larger scale ‘aggregate’ (sometimes, somewhat misleadingly termed ‘smooth’) stellar halo (de Jong 2008, although see D’Souza et al. 2014; Trujillo & Fliri 2016 and Merritt et al. 2016 for encouraging progress). Such broad scale and diffuse structures are possible (with substantial observational cost) to detect and characterize by resolving individual (typically red giant) stars in nearby galaxy stellar haloes. Using such methods, diffuse stellar haloes have been detected and characterized around a number of nearby galaxies (Mouhcine et al. 2005; Monachesi et al. 2016a; M81: Barker et al. 2009, Monachesi et al. 2013; NGC 253: Bailin et al. 2011, Greggio et al. 2014; NGC 891: Mouhcine, Ibata & Rejkuba 2010; Cen A: Harris & Harris 2002, Rejkuba et al. 2014, Crnojević et al. 2016; NGC 3115: Peacock et al. 2015; NGC 3379: Harris et al. 2007b; NGC 3377: Harris et al. 2007a). Yet, quantitative analysis of the density profiles has proven challenging. For the disc-dominated galaxies that we focus on in this paper, such analyses were carried out only for NGC 253 (Bailin et al. 2011; Greggio et al. 2014) and M81 (Barker et al. 2009) where flattened ( $0.4 < c/a < 0.6$ ), steeply declining power-law density profiles were determined. In addition, in common with the integrated light studies, substantial substructure in stellar haloes (streams or shells) has been uncovered in many cases (e.g. Bailin et al. 2011; Greggio et al. 2014; Crnojević et al. 2016). These efforts illuminate the path towards quantifying halo properties with resolved stellar populations, but have not yet yielded a sizeable sample of galaxies with quantified stellar halo properties.

In this paper, we present the projected red giant branch (RGB) star density profiles out to projected radii  $\sim 40$ –75 kpc along the disc’s minor and major axis profiles out to smaller radii of the stellar haloes of six massive nearby disc galaxies using *HST* observations from the Galaxy haloes, Outer discs, Substructure, Thick discs, and Star clusters (GHOSTS) survey (Radburn-Smith et al. 2011; Monachesi et al. 2016a). We focus on the massive galaxy subset of the GHOSTS survey both because their stellar haloes are prominent and straightforward to characterize in our data set and because many models of stellar halo properties focus on galaxies in this stellar mass range ( $4 \times 10^{10} M_{\odot} < M_{*} < 8 \times 10^{10} M_{\odot}$ ). With these data, we estimate the stellar halo density profiles, degree of substructure,

projected axis ratio, and halo stellar masses of this sample, and can combine these measurements with stellar halo population measurements from Monachesi et al. (2013) and Monachesi et al. (2016a) to explore correlations between stellar halo structures and stellar populations. Section 2 provides an overview of the GHOSTS survey. The data reduction and photometry are summarized in Section 3. We present our results for the RGB star density profiles, surface brightnesses, estimated shapes, degree of substructure, and stellar halo masses in Section 4. We build intuition about how results from our sparsely sampled data might compare with global halo properties using models in Section 5. With this intuition in hand, we discuss the results for each galaxy in more depth and compare with previous results in Section 6. In Section 7, we explore the correlations between stellar halo and galaxy properties and compare with theoretical predictions; readers interested primarily in the big picture results are invited to skip directly to Section 7. We summarize and conclude in Section 8.

## 2 OBSERVATIONS: THE GHOSTS SURVEY

The Galaxy haloes, Outer discs, Substructure, Thick discs, and Star clusters (GHOSTS) survey, (Radburn-Smith et al. 2011) is a large *HST* programme designed to resolve the stars in the outskirts of 18 local volume disc galaxies of different masses, luminosities, and inclinations – the largest such study to date. Fields along the principal axes of each galaxy were observed, reaching projected galactocentric distances as large as  $\sim 75$  kpc. GHOSTS observations provide star counts and colour–magnitude diagrams (CMDs) reaching typically  $\sim 2$ – $3$  magnitudes below the tip of the red giant branch (TRGB). Using the RGB stars as tracers of the stellar halo population, we are able to study the size and shape of each stellar halo as well as the properties of their stellar populations such as age and metallicity. A more detailed description of the survey can be found in Radburn-Smith et al. (2011) and Monachesi et al. (2016a).

We selected six galaxies for study in this paper. These galaxies were chosen because they are the most massive among those in GHOSTS, comparable in stellar mass and rotation velocity to the MW. The galaxies are also highly inclined or edge-on, which ensures minimal to no disc contamination when observing along the minor axis beyond 10 kpc.

Fields were chosen to lie on the galactic discs (to study disc structure, dust, and flaring; e.g. de Jong et al. 2007; Radburn-Smith et al. 2012, 2014; Streich et al. 2016) or much further out along the major and minor axes of the main body of the galaxy (to study outer discs and stellar haloes; Monachesi et al. 2013, 2016a), with a few pointings exploring intermediate position angles. In practice, the outermost major axis fields have stellar populations and structures indicative of being dominated by stellar halo, permitting the projected shape of the stellar halo to be estimated if one assumes alignment between the principal axes of the disc and halo (Bullock & Johnston 2005; Bailin et al. 2014; Pillepich et al. 2015, we examine this assumption later in Section 5). The locations of the fields extend out to distances of  $\sim 40$ – $75$  kpc from the centre of the galaxy along the minor axis, depending on the galaxy, as shown in Fig. 1.

## 3 DATA REDUCTION AND PHOTOMETRY

We summarize in this section the main data reduction steps and stellar photometry performed for each exposure using the GHOSTS pipeline. We refer the interested reader to Radburn-Smith et al. (2011) and Monachesi et al. (2016a) where the pipeline for the

data is described for *HST*/Advanced Camera for Surveys (ACS) and *HST*/Wide Field Camera 3 (WFC3) respectively.

We downloaded the ACS  $*\_flc$  FITS images from the Hubble Data Archive the Mikulski Archive for Space Telescopes (MAST),<sup>1</sup> which have been bias-subtracted, cosmic ray flagged and removed, flat fielded, and corrected for charge transfer efficiency (CTE; Anderson & Bedin 2010). For the WFC3 images, we have generated the  $*\_flc$  FITS images locally from the  $*\_raw$  FITS images downloaded from MAST, using a code provided by Space Telescope Science Institute (STScI), since the pixel-based CTE correction is not yet a part of the WFC3/UVIS pipeline. For each field, we combine the individual FLC images per filter using AstroDrizzle (Gonzaga 2012). The resulting image per field and filter is a drizzled DRC FITS image, which has been corrected for geometric distortion.

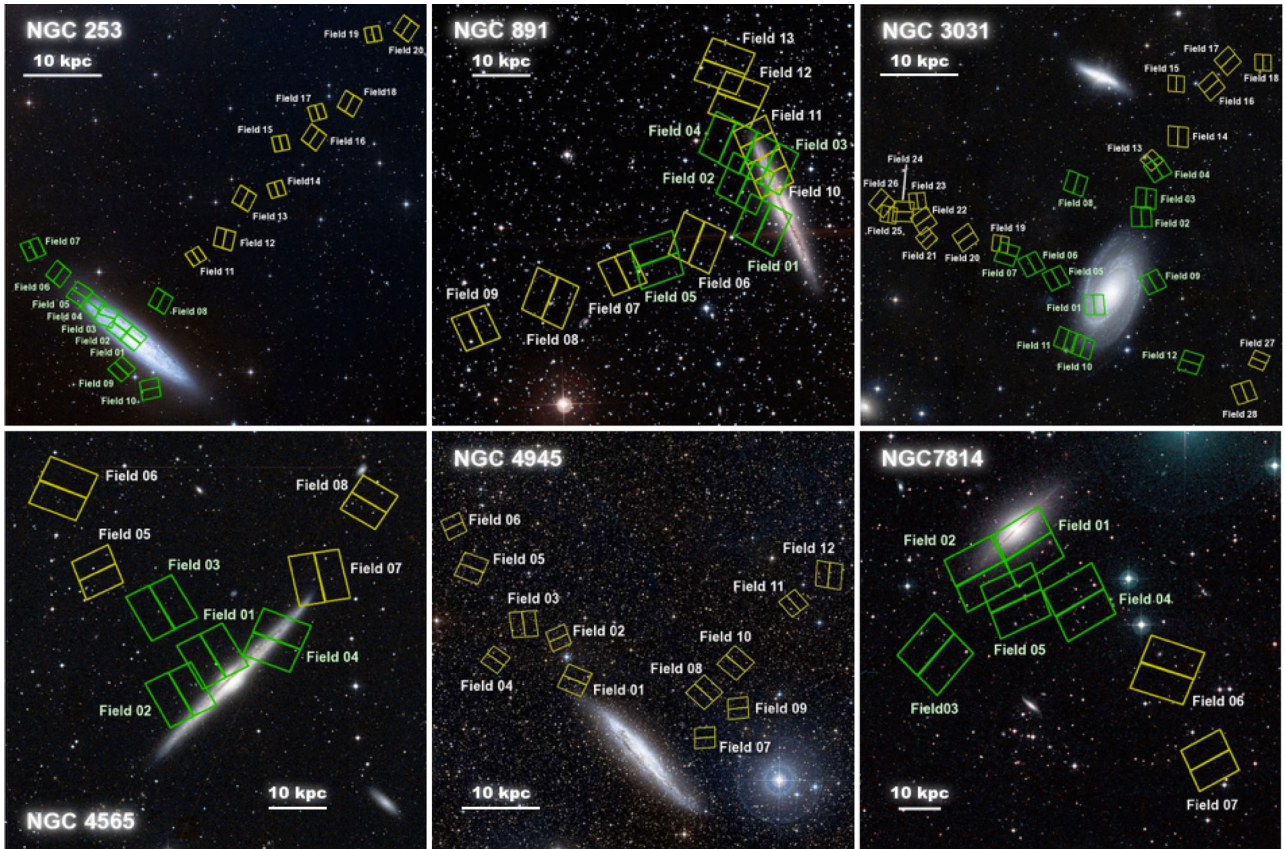
We used DOLPHOT, an updated version of HSTPHOT (Dolphin 2000) for ACS and WFC3 images, to perform simultaneous point-spread function (PSF) fitting photometry on all the individual FLC exposures per field. The DOLPHOT parameters used for the GHOSTS fields are given in table A2 of Monachesi et al. (2016a). DOLPHOT provides the position of each star relative to the *F814W* drizzled image, together with the instrumental *HST* magnitudes in the VEGAMag system already corrected for CTE loss and with aperture corrections calculated using isolated stars. The DOLPHOT output includes various diagnostic parameters that are used to discriminate between PSF-like stars and non-PSF-shaped detections such as cosmic rays and background galaxies.

When attempting to measure the number of faint stars in sparsely populated (with tens to hundreds of stars) *HST* fields, compact background galaxies are the most important source of contamination. We impose several selection criteria to the ACS and WFC3 catalogues, termed ‘culls’ by Radburn-Smith et al. (2011) and Monachesi et al. (2016a), using diagnostic parameters such as sharpness and crowding to distinguish between PSF-shaped sources and sources more or less extended than the PSF. These culls were applied to the photometry output, which removed  $\sim 95$  per cent of the contaminants. The different culls and details on how they were obtained for the ACS and WFC3 data can be found in Radburn-Smith et al. (2011) and Monachesi et al. (2016a) respectively. In addition, we used SExtractor (Bertin & Arnouts 1996) to construct a mask for all extended sources for each field, which include both background galaxies as well as bright foreground MW stars. Detected sources that fall in the masks were removed from the photometry output file. The shorter observations of the WFC3 fields of our closest galaxies (all WFC3 fields in NGC 3031 and NGC 4945 as well as Field 14 in NGC 253) have only one exposure in the *F606W*-band image. Because our pipeline was unable to remove the cosmic rays in these single exposure *F606W* images, many cosmic rays, which are as compact or more compact than real stars, remain in the *F606W* images. Following Monachesi et al. (2016a), we performed an iterative analysis where objects were detected in the *F606W* and *F814W* images; those objects which are much too bright in *F606W* to be real stars were masked out and the photometry recomputed. These masked cosmic rays were added to the SExtractor mask, and were used to reject spurious sources.

We note that contamination from MW foreground stars was not significant within the colour and magnitude range of interest for four of our six sample galaxies at high galactic latitude. Foreground contamination was more severe for NGC 4945 and NGC 0891 fields

<sup>1</sup> <http://archive.stsci.edu>





**Figure 1.** Location of the GHOSTS HST ACS/WFC and WFC3/UVIS fields overlaid on DSS coloured images of each galaxy. Green fields were introduced in Radburn-Smith et al. (2011) whereas yellow fields were presented by Monachesi et al. (2016a). North is up and east is to the left. Fields were placed mostly along the principal axes with some at intermediate position angles. This strategy allows us to both probe their haloes out to projected distances of  $R \sim 40\text{--}75$  kpc along the minor axis from the galactic centre as well as to measure the halo structure and stellar population differences where different regions are observed. For our purposes, not all GHOSTS fields were used as some lie along intermediate axes or are too close to the disc; the list of fields that we have analysed here is given in table 1 of Monachesi et al. (2016a).

since these galaxies are at a low galactic latitude. Based on the CMDs and colour distributions of fields simulated by TRILEGAL<sup>2</sup> (Girardi et al. 2005) and Besançon<sup>3</sup> (Robin et al. 2003) models, we adopted a colour cut for NGC 4945 CMDs to remove most MW contaminants.

In order to assess the completeness of our data and to quantify the photometric uncertainties, we have performed extensive artificial star tests (ASTs) on each exposure, as described by Radburn-Smith et al. (2011). Approximately 2000 000 fake stars were injected in each exposure with realistic colours and magnitudes and distributed such that they follow the observed stellar density gradient. We run DOLPHOT on each fake star at a time and we applied the same culls as in the real output photometry catalogues. Artificial stars that did not pass the culls were considered to be lost. The completeness level was calculated as the ratio of recovered-to-injected number of artificial stars at a given colour and magnitude bin.

Examples of a resulting CMD per galaxy, after the masks and culls were applied, are shown in Fig. 2. The mean distance of each field from the galaxy centre is indicated in each panel. These CMDs are largely free from background and foreground sources. The 50 per cent completeness level of each field is indicated with

a dashed red line. Several CMDs for each galaxy are shown in Radburn-Smith et al. (2011) and Monachesi et al. (2016a). All the CMDs for the entire survey can be found in the GHOSTS website at <http://vo.aip.de/ghosts/>.

## 4 RESULTS

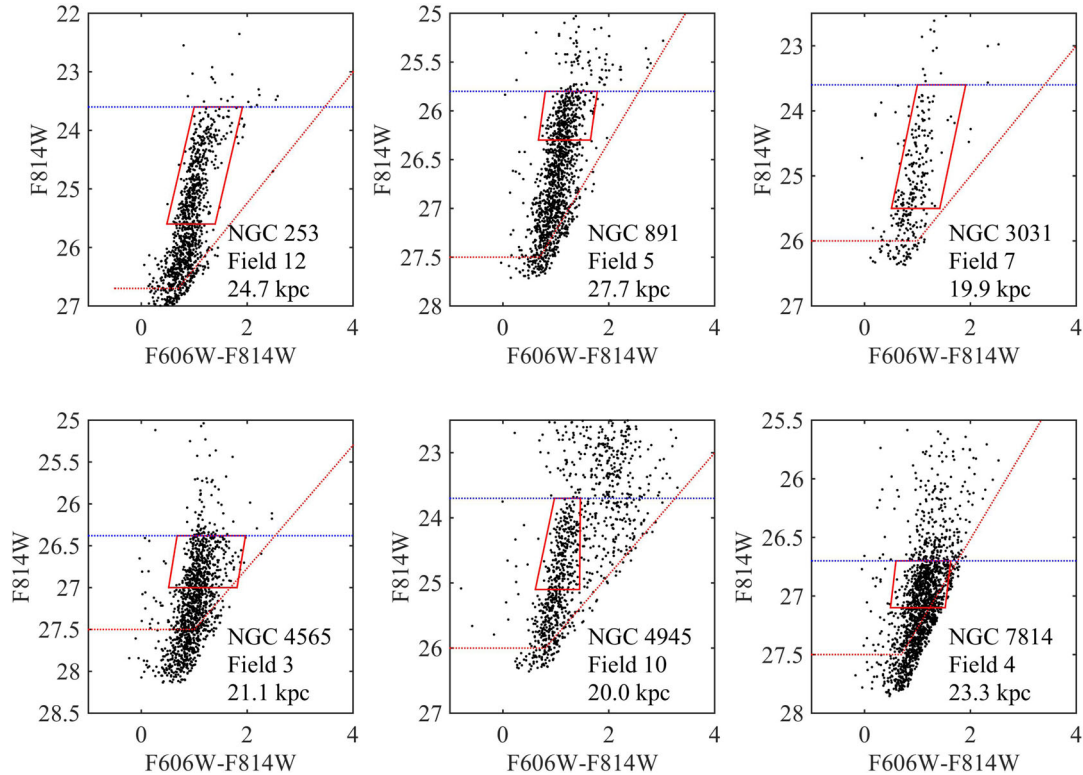
In this section, we present the methods used to calculate and fit the RGB density profiles of the six galaxies in our sample, and the results of those fits. More detailed discussion of the fits on a galaxy-by-galaxy basis, along with tests of our methods and comparison with models is presented later in Sections 6 and 7.

### 4.1 Stellar density profiles

In order to characterize the stellar haloes of our six sample galaxies, we choose to select and analyse stars with the colours and magnitudes of relatively metal-poor RGB stars at the distances of each of the target galaxies. As tracers of the stellar halo, RGB stars offer a number of advantages. The RGB is a prominent feature of the CMD of essentially all intermediate-age and old stellar populations. RGB stars are relatively numerous, offering a large sample of stars to characterize and study. They also have a well-defined maximum luminosity (e.g. Bellazzini, Ferraro & Pancino 2001), allowing measurement of the stellar halo distances and an important

<sup>2</sup> <http://stev.oapd.inaf.it/cgi-bin/trilegal>

<sup>3</sup> <http://model.obs-besancon.fr/>



**Figure 2.** Representative  $F606W-F814W$  versus  $F814W$  GHOSTS CMDs for each of our six target galaxies. The red boxes represent the selection cuts (see Section 4). Only detections selected to be stars following the Radburn-Smith et al. (2011) and Monachesi et al. (2016a) photometric cuts are shown. The TRGB is indicated by a dotted blue line; the 50 per cent completeness limits, as determined from ASTs, are represented by a dotted red line. The field number as well as their projected radial distance from the galaxy centre is indicated in each panel.

check that the stars under consideration indeed belong to the target galaxy. RGB stars of the metallicities and ages thought to dominate the bulk of stellar halo populations, with ages in excess of a few billion years and metallicities below  $1/3$  solar, have a moderately well-defined range of colours, making their identification and characterization relatively straightforward. Finally, RGB star colours do vary somewhat as a function of population parameters (primarily metallicity, see e.g. Hoyle & Schwarzschild 1955; Sandage & Smith 1966), offering insight into the stellar populations of the target stellar haloes.

Candidate RGB stars were selected by making cuts in colour-magnitude space. We select candidate RGB stars to have magnitudes between the TRGB, as presented in Monachesi et al. (2016a, see their table C1), and a limit chosen to lie above the 50–70 per cent completeness limits as determined by the results of the ASTs; this limit is between 0.5 and 1.5 magnitudes fainter than the TRGB. In practice, this limit depends primarily on the distance to the galaxy (which set our depth compared to the TRGB), where more distant galaxies tend to have shallower CMDs and therefore smaller magnitude ranges for RGB star selection. The colour limits at the blue end are designed to prevent contamination from main sequence and helium-burning stars,<sup>4</sup> while the colour limits at the red end are designed to prevent contamination from very metal-rich disc or MW foreground stars as well as to ensure the 50–70 per cent completeness level of the selected stars. In all cases, the colour

selection encompasses the vast majority of halo stars at all relevant radii (minor axis radii  $>5$  kpc and major axis radii  $>20$  kpc). A representative CMD for each galaxy is presented in Fig. 2 along with the adopted selection cuts. For a given galaxy, the same RGB selection cuts were used for all fields. Only stars inside the selection cuts were used to compute the stellar density profiles.

Candidate RGB stars in each field were divided into bins based on their radial distance from the centre of the galaxy. The bins were chosen for each field to balance counting statistics on one hand with a fine enough radial sampling to allow detection of density gradients within a field and substructure, if it exists. Due to the sparse nature of the outer fields, fewer bins were typically used at greater radial distances. In order to minimize contamination by disc stars, we use only stars with radial distances greater than 5 kpc for the minor axis fields and 20 kpc for the major axis fields; the full list of fields that we analyse is given in table 1 of Monachesi et al. (2016a). In each bin, the results from the ASTs were used to correct the star counts for photometric incompleteness. NGC 7814 presents a unique case of severe crowding in the innermost fields. This crowding results in significant undercounting of stars. Deep IRAC  $3.6 \mu\text{m}$  imaging from S4G (Muñoz-Mateos et al. 2015) detects extended light out to radii of  $\sim 9$  and  $\sim 23$  kpc along the minor and major axis, respectively, where our data are crowded; accordingly, we use those surface brightnesses as additional data points. Further description can be found in Section 6.6.

To estimate the area in which stars can be reliably detected in each bin, we need to account for the regions of the images that are discarded. The mask generated using SExtractor (see Section 3) was used to remove any detections near the locations of unresolved background galaxies, bright foreground stars, bad pixels, or globular

<sup>4</sup> In NGC 3031’s case, this meant a blue colour limit that is very close to the RGB (see Fig. 2, as there are substantial numbers of blue stars in M81’s outskirts; Okamoto et al. 2015).



clusters. Detections that fell within a certain distance (25 pixels for ACS and 5 pixels for WFC3) of any masked area were removed to ensure that our star catalogues are minimally contaminated by spurious detections in the vicinity of contaminating objects. The area of each bin was calculated by counting the pixels inside the previously determined radial bins and then subtracting out unused pixels from the mask.

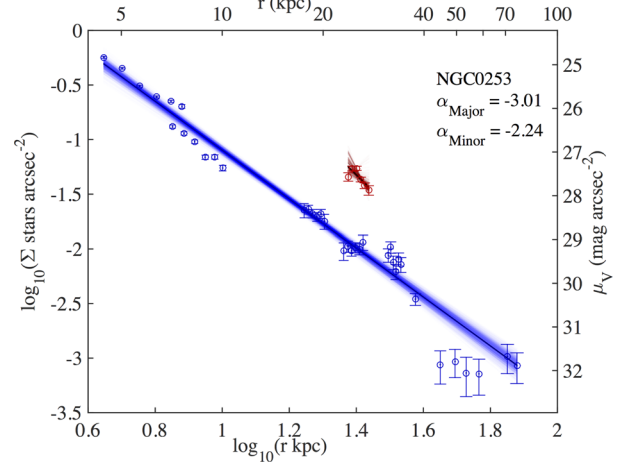
A major advantage of using resolved stellar populations to study low surface brightness stellar haloes is that the fore/background object counts correspond to a faint limiting surface brightness. Accordingly, the effects of fore/background subtraction are worth accounting for, but are only of modest importance. For most of our galaxies, our sparsest regions (typically in the outermost pointings) appear to have CMDs consistent with foreground MW stars plus the few unresolved background galaxies left by the culls (Radburn-Smith et al. 2011; Monachesi et al. 2016a). We choose to designate these areas as representing a fore/background density of RGB-coloured unresolved sources that should be subtracted from the area density for every pointing. For NGC 253, the outermost fields are well populated by RGB stars, and we estimate (using the high-latitude control fields of Radburn-Smith et al. 2011 and Monachesi et al. 2016a) that  $1/3 \pm 1/6$  of the RGB-coloured unresolved sources in the outermost fields (e.g. Field 20) are contaminants, and we adopt that density as an estimate of the fore/background density. For NGC 891, we adopt the number density of detections in the outermost Field 9 as the fore/background estimate. For NGC 3031 and NGC 4565, the lowest stellar density measurement was used as an estimate of the fore/background. For NGC 4945, the outermost fields appear to have a significant population of RGB stars, and we estimate that  $1/2 \pm 1/4$  of Field 12's detected density is fore/background. For NGC 7814, Field 6 has a CMD consistent with mostly MW foreground stars, and is adopted as an estimate of fore/background. The uncertainty in the background was determined to be the square root of the number of stars except in the cases of NGC 253 and 4945, where we adopted an uncertainty of 50 per cent of the adopted background. For every RGB density measurement, the uncertainty on the background value was added in quadrature to the uncertainty for each data point. These corrections produce very modest effects on our final inferences. We have tested this by carrying out a full analysis without fore/background subtraction; all final measurements change by less than their quoted random error bars (as most of the inferences are driven by the higher surface brightness inner parts of the haloes), except for the minor axis power-law slopes, which change by  $\Delta\alpha \sim 0.1\text{--}0.5$ , which is of the order of the systematic uncertainties in their power-law slope.

Figs 3–8 show the stellar density profiles of each galaxy along their major (red symbols) and minor (blue symbols) axes. Given that many of the profiles appear to behave approximately as a power law with substantial scatter around that profile, we maximum-likelihood fit a three parameter power-law model to each of the major and minor axis data sets for each galaxy. The fit is weighted based on the uncertainties in radial bin densities. We assume that the area density of RGB stars at a given projected radius  $r$  can be drawn from the following distribution:

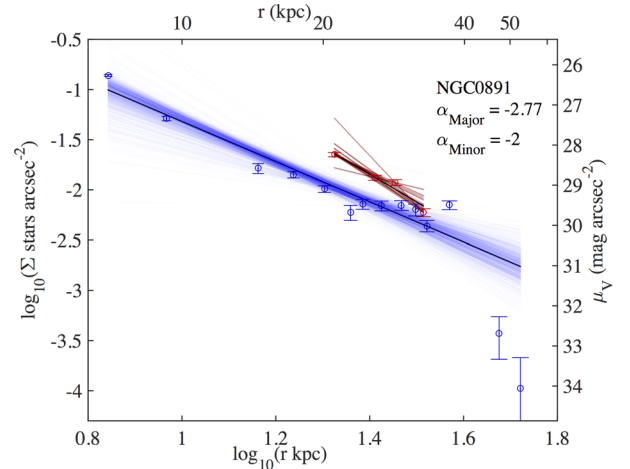
$$P(\log_{10} \Sigma(r)) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\log_{10} \Sigma(r) - \log_{10} \Sigma'(r))^2}{2\sigma^2}}, \quad (1)$$

where the expectation for the RGB star area density at that radius  $\Sigma'(r)$  is given by:

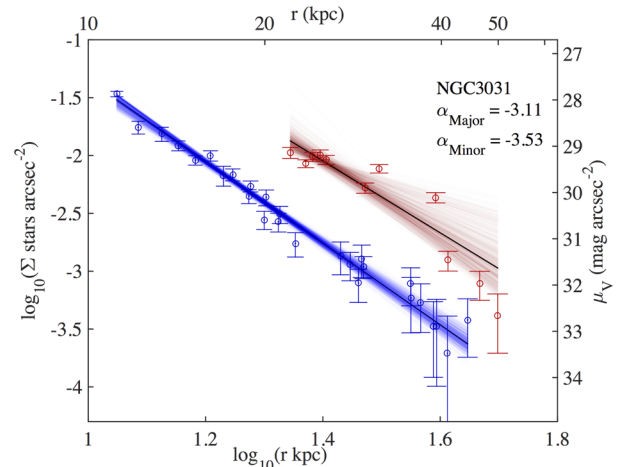
$$\log_{10} \Sigma'(r) = \log_{10} \Sigma_0(r_0) - \alpha \times \log_{10}(r/r_0), \quad (2)$$



**Figure 3.** Stellar density profile for NGC 253's halo along the minor (blue) and major (red) axes. The line resulting from a maximum likelihood fit is displayed with its corresponding slope representing a best-fitting power law for the halo. The translucent lines are the fits resulting from bootstrapping the data.

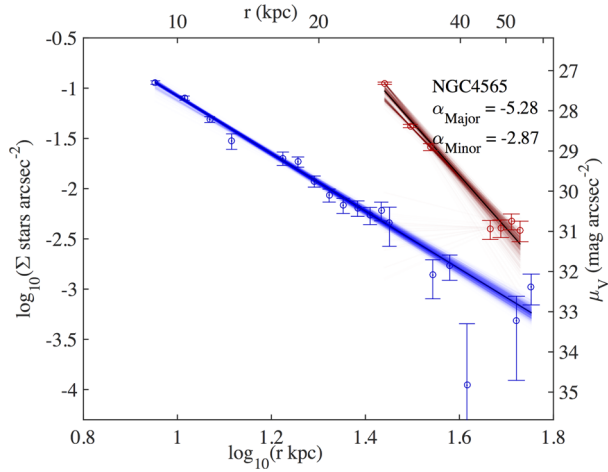


**Figure 4.** Stellar density profile for NGC 891's halo, for a general description, see Fig. 3.

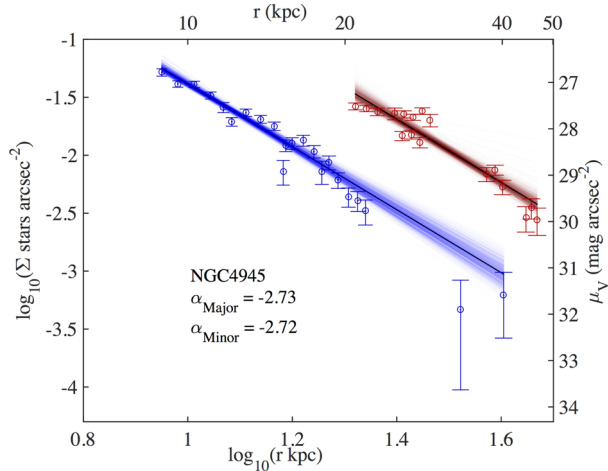


**Figure 5.** Stellar density profile for NGC 3031's halo, for a general description, see Fig. 3. Part of the stellar halo of M82 can be seen in the major axis profile between  $\sim 25$  and 40 kpc as a significant overdensity.

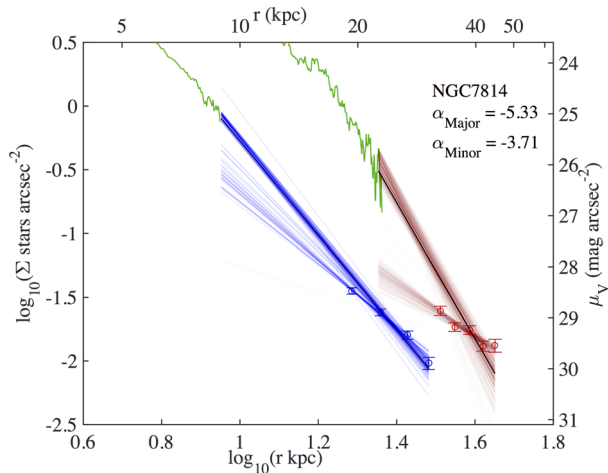




**Figure 6.** Stellar density profile for NGC 4565’s halo, for a general description, see Fig. 3.



**Figure 7.** Stellar density profile for NGC 4945’s halo, for a general description, see Fig. 3.



**Figure 8.** Stellar density profile for NGC 7814’s halo, for a general description, see Fig. 3. The green lines show the S4G integrated surface brightness profiles along the major and minor axes, converted into the equivalent star counts using isochrones.

where  $\Sigma_0$  is the density at the characteristic radius  $\log r_0 = (\log r)$ ,  $\alpha$  is the power-law slope, and  $\sigma$  is the RMS (Table 1) of the data points around the expectation. The best fits are shown as the solid lines in Figs 3–8. Uncertainties were calculated by bootstrapping individual stellar density measurements, and the resulting bootstrapped fits are shown in Figs 3–8 using translucent red or blue lines. The parameters for these power-law fits are given in Table 1, and the star count values and isochrone-derived factors that we use to turn star counts into equivalent  $V$ -band surface brightness are given in Tables 2 and 3, respectively.

We find that the stellar density profiles of our sample of six roughly MW mass galaxies decline steeply, broadly characterized by a range of power-law functions  $\Sigma \propto r^{-\alpha}$ , where  $1.7 < \alpha < 5.3$ . For most galaxies, there is substantial scatter around a single power law that is not well described by measurement uncertainties alone, which is parametrized in this very simple model of a Gaussian scatter around the power-law fit of up to  $\sim 0.15$  dex. This scatter appears to be systematic in nature, with coherent bumps and wiggles in the profiles, indicative of stellar halo substructure in the target galaxies. There is diversity in the recovered power-law slopes in excess of the measurement uncertainties (the dispersion in slopes is substantially larger than the combined error for the slopes), indicative of real diversity in the stellar halo properties of these six roughly MW sized galaxies. It is clear that the choice of a power-law profile is an important oversimplification: a multipart profile would be a substantially better fit for at least the minor axis profiles of NGC 253 and NGC 891 and possibly NGC 4945, where the profiles appear to change slope at radii around  $\sim 30$  kpc.

The prominence of coherent brightness profile fluctuations (associated with recognized large-scale substructure in many galaxies; e.g. NGC 253 and NGC 891; see Section 6) acts to emphasize the importance of substructure in the study of stellar haloes. All of our observed and inferred characteristics – the surface brightness profiles, mass, power-law slope, intrinsic scatter, estimated axis ratio, and stellar populations – are influenced by these substructures. The properties of the stellar halo are best thought of as measurements of the ‘aggregate’ stellar halo. The issue is whether one’s survey has ‘fairly’ sampled the different lines of sight to converge towards a ‘representative’ measurement of stellar halo properties. Cognizant that this issue cannot be quantitatively settled without deep panoramic measurements for a large sample of galaxies (e.g. a survey like GHOSTS but with tens to hundreds of times more survey area), we provisionally estimate the magnitude of such effects using simulations in Section 5.

## 4.2 Stellar halo axis ratios

Given the sparse sampling of GHOSTS along two principal axes, we have a relatively limited ability to estimate projected axis ratio. Given that the major axis profiles typically sample substantially less dynamic range in radius (from  $\sim 20$  kpc to roughly  $\sim 40$  kpc) than the minor axis profiles, we estimate axis ratio by comparing the minor and major axis density profiles at a characteristic radius of 25 kpc.

This ‘indicative’ projected axis ratio  $c/a_{25 \text{ kpc}}$  is determined using the power-law fits obtained above, as described in Fig. 9. Given the interpolated (slightly extrapolated in the case of NGC 4565’s major axis) major and minor axis densities at 25 kpc, the mean of the values of  $\log_{10} \Sigma(25 \text{ kpc})$  on the major and minor axis is calculated,  $\log_{10} \Sigma_{\text{intermed}}(25 \text{ kpc})$ . For each axis, the radius at which the interpolated (extrapolated only for NGC 253’s major axis) densities reach  $\log_{10} \Sigma_{\text{intermed}}(25 \text{ kpc})$  is recorded ( $r_{\text{minor}}$  and  $r_{\text{major}}$ , as

**Table 1.** Table of values for GHOSTS galaxies.

	NGC 253	NGC 891	NGC 3031	NGC 4565	NGC 4945	NGC 7814
Minor axis						
Assumed distance <sup>b</sup> (Mpc)	3.5	9.2	3.6	11.4	4.0	14.4
$r_0$ (kpc)	19.2	23.3	23.1	22.9	16.0	20.0
$\log_{10} \Sigma_0$ ( $\log_{10} N \cdot \text{arcsec}^{-2}$ )	$-1.73^{+0.02}_{-0.02}$	$-2.06^{+0.04}_{-0.04}$	$-2.63^{+0.02}_{-0.02}$	$-2.11^{+0.02}_{-0.02}$	$-1.94^{+0.03}_{-0.03}$	$-1.39^{+0.04}_{-0.07}$
Power-law slope $\alpha$ ( $\pm 0.2$ )	$-2.24^{+0.07}_{-0.06}$	$-2.00^{+0.33}_{-0.33}$	$-3.53^{+0.18}_{-0.18}$	$-2.87^{+0.08}_{-0.07}$	$-2.72^{+0.16}_{-0.16}$	$-3.71^{+0.99}_{-0.09}$
Intrinsic scatter $\sigma$ ( $\pm 0.03$ )	$0.10^{+0.01}_{-0.01}$	$0.13^{+0.05}_{-0.05}$	$0.03^{+0.02}_{-0.02}$	$< 0.11$ (95 per cent)	$0.05^{+0.01}_{-0.02}$	$< 0.03^c$
Major axis <sup>a</sup>						
$r_0$ (kpc)	25.5	26.7	31.2	40.7	30.4	35.0
$\log_{10} \Sigma_0$ ( $\log_{10} N \cdot \text{arcsec}^{-2}$ )	$-1.34^{+0.02}_{-0.02}$	$-1.91^{+0.03}_{-0.03}$	$-2.34^{+0.08}_{-0.06}$	$-1.92^{+0.04}_{-0.05}$	$-1.94^{+0.03}_{-0.02}$	$-1.53^{+0.10}_{-0.13}$
Power-law slope $\alpha$ ( $\pm 0.2$ )	$-3.01^{+1.09}_{-1.55}$	$-2.77^{+0.73}_{-0.44}$	$-3.11^{+0.88}_{-0.48}$	$-5.28^{+0.47}_{-0.45}$	$-2.73^{+0.23}_{-0.23}$	$-5.33^{+3.34}_{-0.57}$
Intrinsic scatter $\sigma$ ( $\pm 0.03$ )	$< 0.05$ (95 per cent)	$< 0.03^c$	$0.14^{+0.04}_{-0.06}$	$< 0.03^c$	$0.09^{+0.01}_{-0.02}$	$< 0.03^c$
Projected $c/a_{25} \text{ kpc}$ axis ratio ( $\pm 0.1$ )	$0.55^{+0.04}_{-0.05}$	$0.74^{+0.04}_{-0.05}$	$0.61^{+0.03}_{-0.05}$	$0.42^{+0.02}_{-0.01}$	$0.52^{+0.02}_{-0.02}$	$0.59^{+0.14}_{-0.05}$
Stellar halo mass ( $M_{10-40}$ ) ( $\pm 30$ per cent)	$1.45^{+0.17}_{-0.10} \times 10^9$	$8.58^{+0.72}_{-0.50} \times 10^8$	$3.66^{+0.35}_{-0.22} \times 10^8$	$7.16^{+0.33}_{-0.31} \times 10^8$	$1.11^{+0.07}_{-0.06} \times 10^9$	$2.05^{+0.43}_{-0.26} \times 10^9$
Total stellar halo mass ( $M_{\text{halo}}$ ) ( $\pm 43$ per cent)	$4.53^{+0.53}_{-0.31} \times 10^9$	$2.69^{+0.23}_{-0.16} \times 10^9$	$1.14^{+0.10}_{-0.07} \times 10^9$	$2.24^{+0.10}_{-0.10} \times 10^9$	$3.47^{+0.19}_{-0.22} \times 10^9$	$6.41^{+1.34}_{-0.81} \times 10^9$
Total galaxy stellar mass ( $M_{\text{galaxy}}$ )	$5.5 \pm 1.4 \times 10^{10}$	$5.3 \pm 1.3 \times 10^{10}$	$5.6 \pm 1.4 \times 10^{10}$	$8.0 \pm 2.0 \times 10^{10}$	$3.8 \pm 0.95 \times 10^{10}$	$4.5 \pm 1.1 \times 10^{10}$
$V_{\text{rot}}$ ( $\text{km s}^{-1}$ )	194	212	224	245	167	231
Colour gradient <sup>b</sup> ( $\times 10^{-4} \text{ mag kpc}^{-1}$ )	$-7.0 \pm 6.1$	$-29 \pm 18$	$0.2 \pm 15$	$-39 \pm 12$	$-9.0 \pm 23$	$-38 \pm 23$

Notes. <sup>a</sup>Not used explicitly to estimate stellar halo mass.

<sup>b</sup>From Monachesi et al. (2016a).

<sup>c</sup>Three  $\sigma$  upper limit based on the formal fit uncertainties.

Systematic uncertainties are given in parentheses in the second column for the power-law slope  $\alpha$ , the intrinsic scatter estimate  $\sigma$ , the axis ratio, and estimates of stellar halo mass. The axis ratio for NGC 3031 has an additional asymmetric error as a result of the galaxy not being completely edge-on.

**Table 2.** Star count Data for the Six GHOSTS Galaxies Examined in this Paper. The whole Table set is available in the online version of this paper as Supporting Information.

Galaxy	Axis	Field	r (kpc)	N·Arcsec <sup>-2</sup>	-1σ	+1σ
NGC 253	Major	7	23.77	0.0456	0.0415	0.0496
			24.45	0.0542	0.0515	0.0570
			25.17	0.0546	0.0522	0.0569
			25.92	0.0431	0.0409	0.0452
			26.63	0.0378	0.0356	0.0401
	Minor	8	27.37	0.0343	0.0309	0.0377
			7.11	0.131	0.127	0.136
			7.69	0.114	0.110	0.118
			8.28	0.0955	0.0919	0.0992
			8.89	0.0689	0.0657	0.0720

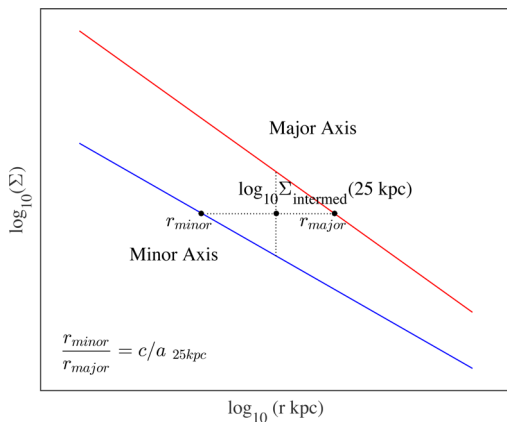
*Notes.* \*Data point derived from the 3.6μm S4G profile for 7814’s axes fits. Points marked with a (W) are derived from WFC3 data; all other fields are from ACS.

r (kpc) column shows radial distance from the galactic centre.

**Table 3.** Flux/star ratios for the galaxies examined in this paper.

Galaxy	Flux ratio (absolute V band mag · N <sup>-1</sup> )
NGC 253	$2.06 \times 10^{-10}$
NGC 891	$2.20 \times 10^{-10}$
NGC 3031	$2.09 \times 10^{-10}$
NGC 4565	$1.03 \times 10^{-10}$
NGC 4945	$3.66 \times 10^{-10}$
NGC 7814	$1.21 \times 10^{-10}$

*Note.* Conversion from star density to apparent magnitude:  $\mu_V = -2.5 \cdot \log_{10}(N \cdot \text{arcsec}^{-2} \cdot \text{fluxratio})$



**Figure 9.** An illustration of the procedure for estimating the projected axis ratio  $c/a$  at  $\sim 25$  kpc,  $c/a_{25\text{kpc}}$ . The average  $\log_{10}\Sigma$  value at 25 kpc is determined, and we estimate  $c/a_{25\text{kpc}}$  to be  $r_{\text{minor}}/r_{\text{major}}$ .

shown in Fig. 9), and we adopt  $c/a_{25\text{kpc}} = r_{\text{minor}}/r_{\text{major}}$  as our best estimate of  $c/a$ . Formal uncertainties in  $c/a_{25\text{kpc}}$  are calculated in concert with the power-law fits to each axis, and are typically small ( $<0.05$  in axis ratio), even when small extrapolations were necessary to estimate the value. In practice, there is considerable uncertainty in translating  $c/a_{25\text{kpc}}$  into  $c/a$ , particularly in cases where the power-law profiles of the major and minor axes differ considerably, indicating a radially varying  $c/a$ . We explore sources of systematic uncertainty in  $c/a$  values using stellar halo models (Section 5), finding a typical systematic uncertainty of  $\Delta c/a \sim 0.1$ , except in one case (out of 11) where there is a large misalignment

between the model galaxy’s principal axes and the stellar halo’s principal axes. We adopt a systematic uncertainty of  $\Delta c/a \sim 0.1$  in what follows. The stellar halo axis ratio estimates  $c/a_{25\text{kpc}}$  for the GHOSTS MW-mass galaxies range from  $c/a \sim 0.4$  to  $\sim 0.75$  (Table 1).

### 4.3 Stellar halo masses and surface brightnesses

We determine the stellar halo mass  $M_{10-40}$  between minor axis equivalent radii of 10–40 kpc, corresponding to  $(10-40)[c/a_{25\text{kpc}}]^{-1}$  kpc along the major axis, using numerical integration. When determining the mass estimates, the choice of lower bound is particularly significant considering the divergent nature of a power-law fit. We chose 10 kpc as the inner bound since this is the closest galactocentric distance along the minor axis for which there is minimal to no disc contamination for the less highly inclined galaxies, such as NGC 3031. The choice of outer bound has a relatively small effect; little mass lies outside 40 kpc for the halo profiles characteristic of GHOSTS galaxies. We first integrate the minor axis power-law profile over the area of the halo within 10–40 kpc, using elliptical annuli with a constant axis ratio of  $c/a_{25\text{kpc}}$  to obtain the number of RGB stars within that area  $N_{\text{RGB},10-40}$ . We use stellar halo models in Section 5 to calibrate this measurement (which can be carried out equally well on our data and with models) and estimate how  $N_{\text{RGB},10-40}$  and  $M_{10-40}$  may be expected to compare to total stellar halo mass.

We then use stellar evolution models to estimate the amount of mass and light represented by each detected RGB star. Our halo CMDs appear broadly consistent with old metal-poor populations; accordingly, we choose to adopt a fiducial 10 Gyr old Padova isochrone (Bressan et al. 2012; Chen et al. 2014; Tang et al. 2014) with a metallicity  $Z = 0.0016$  ( $[\text{Fe}/\text{H}] = -1.2$  dex) – similar to the average metallicity for our data set – to represent the bulk of the halo population. We adopted a Chabrier (2003) stellar initial mass function (IMF). A well-populated model CMD was constructed, and the number of RGB stars in the selection region of the CMD (see red box in Fig. 2) per unit initial stellar mass and V-band luminosity is calculated for each galaxy. The right-hand axis in Figs 3–8 shows the  $\mu_V$  profile in units of  $V\text{-mag arcsec}^{-2}$ .

Scaling of star counts to total surface brightness using stellar population models is a common technique (e.g. Ibata et al. 2014). Nonetheless, it is useful to cross-validate our inferred surface brightness profiles with previously published values. Such cross-validation is challenging owing to the difficulty in finding systems with low enough surface brightness for the resolved stellar populations to remain uncrowded while remaining well measured in integrated light (V-band surface brightnesses of  $\sim 27$  mag arcsec<sup>-2</sup>). In addition, we wish to target metal-poor regions, as our star counts focus on metal-poor stars.

We can compare our measurements of isochrone-scaled star counts with integrated surface brightness estimates for three systems in the GHOSTS sample: NGC 253, NGC 891, and NGC 4565. We compare our inferred V-band major axis surface brightness profile for NGC 253 with the J-band surface brightness profile of Greggio et al. (2014, from star counts scaled to J-band brightness where their profiles overlapped), assuming  $V-J \sim 1.7$  for a  $[\text{Fe}/\text{H}] \sim -1$ , 10 Gyr old stellar population following Bruzual & Charlot (2003), finding agreement within  $\Delta\mu \sim 0.1$  mag arcsec<sup>-2</sup>. We compare our inferred V-band minor axis surface brightness profile for NGC 891 at 6–9 kpc with the R-band brightness profile of Miller (1996) converted to V-band assuming a  $[\text{Fe}/\text{H}] \sim -1$ , 10 Gyr old stellar population with  $V-R \sim 0.52$  following Bruzual & Charlot (2003),

finding agreement within  $\Delta\mu \sim 0.3$  mag arcsec $^{-2}$ . Turning to NGC 4565, a  $<10$  kpc extrapolation of our inferred minor axis  $V$ -band brightness agrees within  $\Delta\mu \sim 0.2$  mag arcsec $^{-2}$  with the  $V$ -band surface brightness of 27 mag arcsec $^{-2}$  at 8 kpc minor axis distance from Naeslund & Joersaeter (1997). We conclude that our surface brightness measurements appear to be accurate, with no sign of a systematic offset at the 0.3 mag arcsec $^{-2}$  level.

The isochrones give estimates of initially formed stellar mass, which must be corrected to present-day mass by accounting for stellar mass loss by multiplying the initially formed mass by 0.56 (following Bruzual & Charlot 2003). The present-day stellar halo mass  $M_{10-40}$  is then calculated by dividing the total number of detected RGB stars between minor axis equivalent radii of 10–40 kpc  $N_{\text{RGB}, 10-40}$  by the number of RGB stars per unit present day stellar mass. Our resulting stellar halo mass estimates  $M_{10-40}$  are presented in Table 1. We note that the random uncertainties (determined from bootstrapping) presented in Table 1 do not include a contribution from systematic uncertainty about the halo stellar populations or isochrone uncertainties; we varied ages and metallicities by  $\pm 30$  per cent in age and a factor of three in metallicity, and this changes the final masses by  $\pm 30$  per cent or less. These are included in the systematic error budget in Table 1.

We also indicate in Table 1 the total stellar mass of each galaxy, estimated using  $K$ -band luminosities in concert with a  $K$ -band mass to light ratio of  $M/L = 0.6$ , typical of massive spiral galaxies, following Bell & de Jong (2001) using a universally applicable Chabrier (2003) stellar IMF. Luminosities were calculated using  $K$ -band total magnitudes from Jarrett et al. (2003), in conjunction with the distances presented in Table 1. Such masses carry at least 30 per cent uncertainties, and potentially suffer from larger systematic error if assumptions underlying their calculation are incorrect, e.g. if the stellar IMF varies from galaxy to galaxy. Despite these uncertainties, these masses are useful in order to build intuition about how these galaxies compare to larger samples of galaxies, e.g. from the Sloan Digital Sky Survey (SDSS) (e.g. Kauffmann et al. 2003) that have stellar mass estimates but lack accurate measures of rotation velocity.

## 5 HOW GENERALIZABLE ARE OUR INFERENCES FROM THE DATA? GENERATING INTUITION THROUGH ANALYSIS OF STELLAR HALO MODELS

Before examining the results for individual galaxies, intercomparing them, and comparing our observations with theoretical models, it is important to generate intuition about how our results might generalize to the bulk properties of a realistically structured stellar halo. As articulated earlier, the key concern is the degree of systematic error caused by sparse sampling stellar halo structure in a highly structured aggregate halo; a secondary concern is the influence of stellar population variations in the stellar halo on our inferences.

In the absence of panoramic imaging as deep or deeper than our data [e.g. future wide-area surveys with *Wide-Field Infrared Survey Telescope* and the Large Synoptic Survey Telescope (LSST)], it is necessary to use simulations to explore this issue. While any simulation could be used in principle, we choose to analyse the 11 halo realizations from the Bullock & Johnston (2005) simulations.<sup>5</sup> These stellar halo models are built through the disruption and accretion of satellite galaxies in a cosmological context. Star particles

in subhaloes were generated using high-resolution  $N$ -body simulations and painted on to dark matter particles such that their luminosity function follows a King profile. A cosmologically motivated semi-analytic galaxy formation model was used to assign stellar properties to the painted particles (see also Robertson et al. 2005; Font et al. 2006). We converted the star particles into RGB stars and generated projected RGB maps of stars as explained in Monachesi et al. (2013). For these haloes, we emulated ACS observations by choosing square sections of 202 arcsec on a side along the major and minor axes. The different galaxy distances and colour–magnitude cuts that correspond to each of the six massive GHOSTS galaxies were used to examine the models. This allows us to determine how representative our data are for each galaxy.

We choose to analyse 10 ACS-like fields per galaxy, 5 on the minor axis and 5 on the major axis. While clearly the number of pointings per galaxy varies from case to case (see Fig. 1), this is close to the average number of independent pointings per galaxy. The simulated ACS-like fields were treated identically to the real ACS observations. A best-fitting power law was calculated and integrated over an ellipse between 10 and 40 kpc using an axis ratio derived at an ‘indicative’ radius of 25 kpc, and the stellar mass of the models was found using the same process that was applied to the data. We compared the results from these simulated ACS observations to the true values for each model for the power-law slopes, axis ratio, and stellar halo mass as described below.

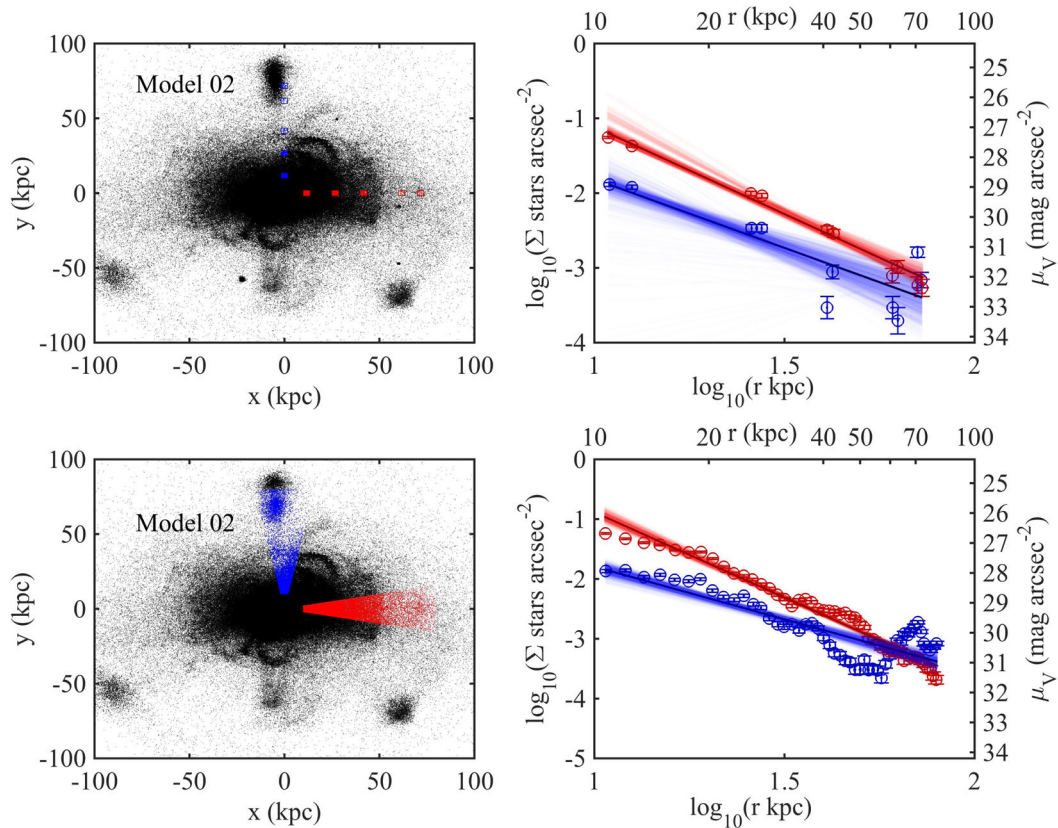
In order to find the true power-law slope for the stellar density profile of the model, we selected stars within wedges of 1/8 radian half-width around the major and minor axes, between 10 and 80 kpc from the centre, as illustrated in the bottom-left panel of Fig. 10. Each of these regions was divided into 50 radial sections and we constructed projected stellar density profiles of the modelled RGB stars on the minor and major axes. An example of the wedge density profiles for Bullock & Johnston Halo 02 can be seen in the bottom-right panel of Fig. 10. The resulting power-law slopes that best fit the profiles were taken to be the ‘true’ values in order to measure the accuracy of the simulated ACS observations. The top panel of Fig. 10 shows the ACS-like fields corresponding to the same Bullock & Johnston Halo 02 model as well as the resulting density profiles. Comparing the results obtained using these two methods, we find that our sparse sampling method produces power-law slope estimates accurate to about  $\pm 0.2$ .

To find the ‘true’ axis ratio of the models at 25 kpc, we fit the RGB stars distribution of the models using an iterative method. We select RGB stars within an elliptical annulus with a geometric mean distance of 25 kpc using an initial guess for axis ratio and assuming alignment between the major axis and the long axis of the initial ellipse. The second moment tensor of the distribution was calculated, giving improved estimates of axis ratios and position angle. This process was repeated until it converged to within 0.001 in axis ratio. We find the axis ratio we calculate based on sparse-sampled *HST* fields is accurate to within  $\Delta c/a \sim 0.1$ , except in one case where there is a large amount of substructure (1 case out of 11) where our method recovered  $c/a \sim 1$  for a halo with actual  $c/a_{\text{intrinsic}} \sim 0.5$  owing to a misalignment between the actual position angle of the halo and the major axis of the galaxy. Given the sparse survey strategy that we adopted (constrained by the amount of available telescope time), it is difficult to guard against position angle differences between the halo and principal axes of the main body of the galaxy; given that this happens at the 1/11 level in simulations, we expect the bulk of our axis ratios to be accurate to  $\Delta c/a \sim 0.1$ .

These models also offer an important end-to-end test of our survey strategy’s ability to infer reliable stellar halo masses. For a

<sup>5</sup>The stellar halo models are available at <http://user.astro.columbia.edu/kvj/halos/>.





**Figure 10.** Top: results from Bullock & Johnston Halo 02 inferred using sparse sampling mimicking that of the data. ACS-like fields were placed along the minor and major axes as illustrated in the left-hand panel, giving the stellar density profiles shown in the right-hand panel. The cuts in space and colour–magnitude were the same as those used in NGC 253. Bottom: Stellar density profile (right-hand panel) and 2D map of RGB stars (left-hand panel) for Bullock & Johnston Halo 02. The resulting density profiles along wedges on the major and minor axes are obtained from colour–magnitude cuts as those used in NGC 253. The slopes for the major and minor axes as well as the masses calculated using the two different methods were in agreement within 10 per cent.

range of distances corresponding to our sample galaxies, we choose colour–magnitude selections appropriate to each galaxy and calculate the mass between minor axis equivalent radii of 10–40 kpc using the method described above (using the minor axis profile and the indicative axis ratio at  $\sim 25$  kpc). In concert, we calculate the true mass between 10 and 40 kpc in an elliptical annulus with the correct position angle and ellipticity (the ‘true’ 10–40 kpc mass) and the total stellar halo mass. Our observational and analysis techniques give estimates of  $M_{10-40}$  which are  $97 \pm 22$  per cent of the ‘true’ 10–40 kpc mass; our estimates of  $M_{10-40}$  correspond to  $32 \pm 10$  per cent of the total RGB stars for model stellar haloes from Bullock & Johnston (2005).

M81 presents a unique case as it has an inclination of  $60^\circ$ . We rotate the models to simulate its orientation and find that the power-law slopes vary by typically less than 0.2 in power-law slope, the masses by 10 per cent or less, and the axis ratios increase typically by 0.2 compared to a perfectly edge-on model. Accordingly, we include an extra systematic uncertainty of  $^{+0.0}_{-0.2}$  in  $c/a$  for M81 in Table 1.

We incorporate estimates of these systematic uncertainties in Table 1.

## 6 NOTES ON INDIVIDUAL GALAXIES

Table 1 presents our estimates of the stellar halo properties – power-law slope, normalization, intrinsic scatter around a power-law profile, indicative axis ratio and mass between minor axis equivalent

radii of 10 and 40 kpc, for each of the galaxies studied. In this section, we discuss our results for individual galaxies and compare our estimates of halo properties, determined using our strategy which obtains deep high-quality detections on relatively few sparse pointings, with other work typically derived from wide-field ground-based studies. In what follows, we will often quote random and systematic uncertainties separately.

### 6.1 NGC 253

The minor axis density profile for NGC 253 is well measured out to more than 75 kpc, following a power law with slope  $-2.24^{+0.07}_{-0.06} \pm 0.2$  (random and systematic errors, respectively) reasonably well out to  $\sim 50$  kpc as can be seen in Fig. 3. We note that this detection of stellar halo stars at  $>75$  kpc is somewhat remarkable – only three galaxies, the MW, M31, and Centaurus A (Rejkuba et al. 2014; Crnojević et al. 2016) have halo stars detected to such radii.

There is significant scatter around the fitted power-law profile, with a best-fitting intrinsic RMS of  $0.10 \pm 0.01 \pm 0.03$  dex (random and systematic errors, respectively). These deviations are systematic, with coherent overdensities compared to the power-law fit at  $\sim 30$  kpc, and coherent underdensities at  $\sim 10$  kpc and most notably outside  $\sim 40$  kpc, where the profile is significantly depressed compared to smaller radii and appears to become flat. Comparison with the single halo-dominated major axis field in the GHOSTS survey yields a rough estimate of  $cl/a \sim 0.55^{+0.04}_{-0.05} \pm 0.1$  (random and systematic errors, respectively) for the projected axis ratio, though the

uncertainties may be larger owing to our use of only a single major axis field.

There are two existing estimates of the power-law slope and axis ratio of the stellar halo of NGC 253 from panoramic ground-based imaging: Greggio et al. (2014) used Visible and Infrared Survey Telescope for Astronomy (VISTA) wide-area near-infrared imaging to determine a slope of  $\sim -1.6$  and an axis ratio  $b/a \sim 0.4$ , and Bailin et al. (2011) measured a power-law slope of  $-2.8 \pm 0.6$  and  $b/a \sim 0.35$  using IMACS data for the southwest quadrant of NGC 253's stellar halo. Our power-law slopes are intermediate to these estimates, and our axis ratio is rather larger than both of these estimates. These works, along with our own and that of Davidge (2010), all show clear evidence of substantial substructure in NGC 253's stellar halo. Two significant overdensities have been reported: a prominent 'shelf' in the southwestern quadrant of the inner part of NGC 253's halo (Beck, Hutschenreiter & Wielebinski 1982; Davidge 2010; Bailin et al. 2011; Greggio et al. 2014), and an overdensity along the northern minor axis at  $\sim 30$  kpc best visualized in figs 16 and 21 of Greggio et al. (2014). Our minor axis profile intersects the northern minor axis overdensity, and it is clearly visible in Fig. 3 as an overdensity at 30 kpc, beyond which the star count profile drops precipitously. We interpret the significant differences in stellar halo parameters reported by our work, Greggio et al. (2014) and Bailin et al. (2011), to stem in large part from the prominent substructure in NGC 253's stellar halo (this was also emphasized by Bailin et al. 2011 and Greggio et al. 2014 as their main source of systematic uncertainty); such differences may indicate the level of variation expected from study to study owing to substructure in stellar haloes.

Only one estimate of stellar halo mass has been published to date:  $2.5 \pm 1.5 \times 10^9 M_{\odot}$  outside of minor axis radius of 5 kpc (4.5 per cent of the galaxy stellar mass) from Bailin et al. (2011) using wide-area coverage of the southwestern quadrant of the inner parts of NGC 253's halo. Our halo mass estimate is  $1.45^{+0.17}_{-0.10} \pm 0.5 \times 10^9 M_{\odot}$  (random and systematic errors, respectively) between minor axis equivalent radii of 10–40 kpc; recall in Section 5 we use models to suggest that this likely implies a three times larger total stellar halo mass, implying a total stellar halo mass of roughly  $4.5 \pm 1.9 \times 10^9 M_{\odot}$  ( $8 \pm 3$  per cent of the galaxy stellar mass). These estimates agree to within their uncertainties.

## 6.2 NGC 891

As far as we are aware, our measurement is the first quantitative measurement of the stellar halo density profile, axis ratio, and mass for NGC 891. In particular, the mass of the stellar halo between 10 and 40 projected minor axis equivalent kpc is  $8.6^{+0.7}_{-0.5} \pm 2.6 \times 10^8 M_{\odot}$  (random and systematic errors, respectively). This corresponds to an estimated total stellar halo mass of  $2.7 \pm 1.2 \times 10^9 M_{\odot}$ , corresponding to  $5 \pm 2$  per cent of NGC 891's total stellar mass. NGC 891 has been imaged using Subaru's Suprime-Cam (Mouhcine et al. 2010), leading to the discovery of extensive stellar streams and a relatively dense 'cocoon' of stars in the inner parts of NGC 891's stellar halo (their fig. 1). We clearly detect the stream and cocoon (towards Mouhcine et al. 2010's positive Z direction) on the minor axis fields between 25 and 40 kpc, where the density profile is close to flat. This overdensity, and relatively dramatic drop in density outside 40 kpc, drive both a relatively uncertain minor axis power-law slope ( $-2.00^{+0.33}_{-0.23} \pm 0.2$ ) and one of our largest values of intrinsic scatter ( $0.13 \pm 0.05 \pm 0.03$  dex). One could arbitrarily fit the density profile with a double power law broken at  $\sim 40$  kpc, in which case the best-fitting slopes are  $\sim -2$  inside 40 kpc and  $\sim -7$

(but with huge uncertainty) outside 40 kpc. We do not adopt the parameters of this fit in this work, nor do we show it in Fig. 4; such a fit would be too specific to the particular density profile seen in Fig. 4 and would hinder fair comparison with other galaxies or with simulations (most of which use single power-law fits to broadly characterize the density distribution).

## 6.3 NGC 3031/M81

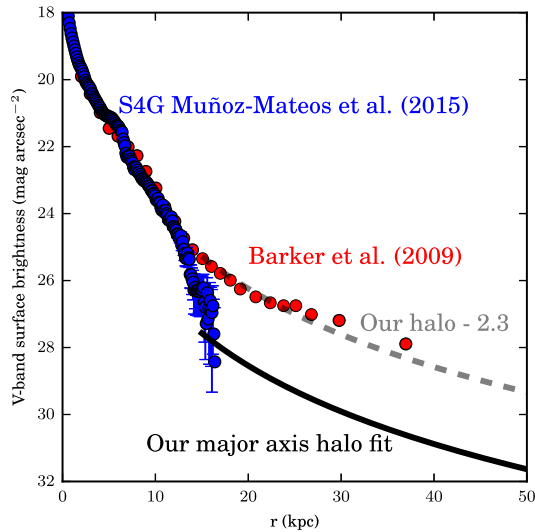
Power-law fits over a dynamic range of a factor of 4 in radius for NGC 3031/M81 along the minor axis and a factor of nearly 2.5 in the major axis show that the metal-poor RGB stars show a steeply declining roughly power-law profile with slopes  $-3.53^{+0.18}_{-0.15} \pm 0.2$  and  $-3.11^{+0.88}_{-0.48} \pm 0.2$  respectively. The scatter is  $0.03 \pm 0.02 \pm 0.03$  and  $\sim 0.14^{+0.04}_{-0.06} \pm 0.03$  dex along the minor and major axes, respectively, and axis ratio is  $0.61^{+0.03}_{-0.05} \pm 0.1^{+0.0}_{-0.2}$ , where the last error term accounts for the possible increase of projected axis ratio compared to the intrinsic axis ratio owing to M81's intermediate inclination. This yields a stellar halo mass between minor axis equivalent radii of 10 and 40 kpc of  $3.7^{+0.4}_{-0.2} \pm 1.1 \times 10^8 M_{\odot}$ . This corresponds to an estimated total stellar halo mass of  $1.1 \pm 0.5 \times 10^9 M_{\odot}$ , corresponding to  $2 \pm 0.9$  per cent of M81's total stellar mass.

Many of our values appear to be in significant conflict with the only other estimates of the properties of M81's stellar halo from Barker et al. (2009) using ground-based Suprime-Cam observations. While our axis ratio estimate of  $\sim 0.6$  is in agreement with the axis ratio of  $\sim 0.5$  assumed by Barker et al. (2009) when analysing the inner part of M81's stellar halo, our other measurements disagree with those of Barker et al. (2009). Our power-law slopes are  $\sim -3.5$ , whereas those of Barker et al. (2009) are  $\sim -2$ , and most prominently, our estimated total stellar halo mass of  $1.1 \pm 0.5 \times 10^9 M_{\odot}$  (corresponding to  $\sim 2$  per cent of M81's total stellar mass) appears to differ by almost an order of magnitude with their claim that M81's halo contains 10–15 per cent of the luminosity of M81.

We explore this discrepancy in depth in Fig. 11, which shows the major axis V-band surface brightness profile M81/NGC 3031 from fig. 17 of Barker et al. (2009) in red, and our major axis halo fit in black. These are clearly discrepant at the radii at which they overlap, but are not grossly different in shape, as evidenced by the dashed grey line, which shows our major axis profile fit offset by 2.3 magnitudes to approximately overlap with Barker et al. (2009).<sup>6</sup> This brightness offset (coupled with minor differences in extrapolations to total stellar halo mass and luminosity) accounts for the difference between our and their halo luminosity estimates.

How is such a large difference in calibration possible? We attempt to shed light on this issue by comparing these brightness profiles with the 3.6  $\mu\text{m}$  surface brightness profile from S4G (Muñoz-Mateos et al. 2015) scaled to V-band by matching the inner parts of Barker et al.'s surface brightness profile. S4G (Sheth et al. 2010) is sensitive to relatively faint levels, and is much more immune to low surface brightness Galactic cirrus emission than optical light (clearly visible in fig. 2 of Barker et al. 2009). The S4G brightness profile – well measured out to  $\sim 17$  kpc – clearly continues to decline with an exponential profile well outside of  $\sim 12$ – $14$  kpc where Barker et al. (2009) claim a transition in the integrated brightness profile to a shallower power law. As discussed in their section 6

<sup>6</sup> The difference in power-law slope is visible by a 'drift' in the best offset between the two data sets of about 0.5 mag between 20 and 40 kpc, in the sense that the brightness profile of Barker et al. (2009) is flatter than ours.



**Figure 11.** Major axis V-band surface brightness profile M81/NGC 3031. We show the  $3.6\ \mu\text{m}$  surface brightness profile from S4G (Muñoz-Mateos et al. 2015) scaled to V band by matching the inner parts of Barker et al.’s surface brightness profile, the final combined integrated light+star counts profile from Barker et al. (2009) in red, and our major axis halo fit in black, along with our fit rescaled 2.3 mag arcsec $^{-2}$  brighter to approximately match the Barker et al. (2009) profile in their outer parts in dashed-grey line. It is clear that the Barker et al. (2009) star counts were erroneously calibrated to a surface brightness around  $10\times$  higher than the isochrone-derived calibrations we use, and predict far more surface brightness than either we or Muñoz-Mateos et al. (2015) observe at deprojected major axis radii in excess of 14 kpc.

and shown in their fig. 17, Barker et al. (2009) use their 14–17 kpc integrated light profile (which we believe to have been erroneously bright owing to Galactic foreground cirrus, and as evidenced by the S4G data) to calibrate their star counts. By enforcing that they overlapped, Barker et al. (2009) calibrated their star counts to have a surface brightness nearly a factor of 10 brighter than those that they would have derived by doing ASTs plus calibration with isochrones.

It is interesting to ask what drives the difference in power-law slope reported by Barker et al. (2009) ( $\Sigma \sim r^{-2}$ ) and our derived slopes along minor and major axes (both  $\Sigma \sim r^{-3}$ ). Part of this difference could well be crowding in the ground-based data. Bailin et al. (2011) show for NGC 253 (with a very similar distance to M81) that within radial distances of  $\sim 20$  kpc the ground-based data were very crowded, resulting in the detection of only a tiny fraction of the real RGB stars, particularly at brighter surface brightnesses, artificially flattening their star counts.<sup>7</sup> Furthermore, real substructure may be responsible for some of the difference in power-law slope: inspection of the major axis profile in Fig. 5 between major axis radii of 20 and 40 kpc (the range covered by Barker et al. 2009) shows a considerably flatter profile ( $\Sigma \propto r^{-2}$  or somewhat shallower) than the profile between 20 and 50 kpc ( $\Sigma \propto r^{-3}$ ). Recently, Okamoto et al. (2015) showed that there is an extensive fan of debris between M82 and M81 (we expect composed largely of material tidally liberated from M82). This is clearly detected along our major axis fields, and is prominent in much of the area probed by Barker

<sup>7</sup> Incidentally, we also attribute the lack of a strong change in slope in the Barker et al. star count profile within  $\sim 17$  kpc, where S4G predicts a transition from an exponential to power-law profile, to the effects of crowding.

et al. (2009). We propose that this drives the density profile derived by Barker et al. (2009) towards  $\Sigma \propto r^{-2}$ . We claim that our measurement of a projected density profile  $\Sigma \propto r^{-3}$  is somewhat more representative of M81’s stellar halo not only because it is derived from the minor axis where there appears to be no material stripped off of M82 or NGC 3077 (Okamoto et al. 2015), but also because it draws from a larger range of major axis radii, showing a return to a  $\Sigma \propto r^{-3}$  profile outside of projected major axis radii of 40 kpc.

We conclude that the shape of the brightness profiles from our work and Barker et al. (2009) are largely consistent, given the importance of both crowding and substructure on the Barker et al. result. Our more reliable isochrone-based luminosity and stellar mass calibration differs strongly from Barker et al.’s, and with the benefit of deeper uncrowded HST data and S4G’s deep integrated light profile we conclude that Barker et al.’s brightness calibration and luminosity estimate appears to be in error, owing to an unfortunate limitation in how their star counts were converted into an estimate of the V-band surface brightness.

#### 6.4 NGC 4565

As far as we are aware, this work, Monachesi et al. (2016a), and preliminary results from de Jong, Radburn-Smith & Sick (2009) are the first reported detection and characterization of NGC 4565’s resolved stellar halo. NGC 4565’s stellar halo is detected out to 60 kpc along the minor axis, and more than 50 kpc along the major axis. NGC 4565’s minor axis density profile has a power-law slope of  $-2.87^{+0.08}_{-0.07} \pm 0.2$  (random and systematic uncertainties, respectively), and while the fit prefers  $\sim 0.11 \pm 0.03$  dex of intrinsic scatter, owing to the magnitude of the error bars in the outer parts of NGC 4565 it is also consistent with having no intrinsic scatter around its minor axis profile. NGC 4565 has a substantially steeper major axis profile, with a power-law slope of  $-5.28^{+0.47}_{-0.45} \pm 0.2$ . If interpreted in terms of a halo with changing projected axis ratio, the axis ratio would vary from  $c/a_{25\text{kpc}} \sim 0.44$  to  $c/a_{40\text{kpc}} \sim 0.56$  (determined from the outermost points of the major axis profile, comparing them to the points of equal density along the minor axis at radii  $\sim 30$  kpc). The calculated stellar mass between 10 and 40 minor axis equivalent kpc is  $7.2 \pm 0.3 \pm 2.2 \times 10^8 M_{\odot}$ , corresponding to an estimated total stellar halo mass of  $2.2 \pm 0.9 \times 10^9 M_{\odot}$  or  $2.8 \pm 1.2$  per cent of the total stellar mass of NGC 4565.

We note that Field 06 (the outermost minor axis field in NGC 4565) has a significant overdensity (the high data point at minor axis radius  $\sim 57$  kpc), which we interpret to be a relatively thin stellar stream (also discussed in Monachesi et al. 2016a). The width of the overdensity is  $\sim 2\text{--}3$  kpc.

#### 6.5 NGC 4945

As far as we are aware, this work (together with Monachesi et al. 2016a) is the first reported detection and characterization of NGC 4945’s stellar halo. Owing to the substantial foreground contamination, we detect the minor axis to distances of only  $\sim 40$  kpc and the major axis to  $\sim 45$  kpc. The power-law slopes for the density profiles along the minor and major axes are consistent with each other at  $-2.72 \pm 0.17 \pm 0.2$  and  $-2.73 \pm 0.23 \pm 0.2$ , respectively. The scatter around the minor axis profile is  $0.05^{+0.01}_{-0.02} \pm 0.03$  dex, and around the major axis is  $0.09^{+0.01}_{-0.02} \pm 0.03$  dex. Owing to the similarity of the density profiles along the major and minor axes, the axis ratio appears to be  $c/a \sim 0.5 \pm 0.1$  with little radial dependence (although it is measured only out to minor axis equivalent radii of  $\sim 22$  kpc). The resulting halo stellar mass between 10 and



40 kpc is  $1.11^{+0.07}_{-0.06} \pm 0.33 \times 10^9 M_{\odot}$ , corresponding to a rough estimate of total stellar halo mass of  $3.5 \pm 1.5 \times 10^9 M_{\odot}$  which is roughly  $9 \pm 4$  per cent of the total stellar mass of NGC 4945.

## 6.6 NGC 7814

NGC 7814 is the most distant galaxy in our sample; as can be seen in Fig. 2, the CMDs are relatively shallow and are limited in their colour coverage towards the red fainter than  $F814W \sim 27.2$  by a relatively shallow  $F606W$  limit. Accordingly, we caution that our stellar counts may be somewhat less reliable than for our other galaxies, and may represent a lower limit to the true value. Furthermore, the innermost fields in GHOSTS suffer from significant crowding. Reliable GHOSTS star counts exist only between  $\sim 19$  and  $\sim 35$  kpc along the minor axis and beyond  $\sim 32$  kpc along the major axis. We supplemented the minor (major) axis profile with an equivalent star count value at  $\sim 9$  ( $\sim 23$ ) kpc derived from  $3.6 \mu\text{m}$  imaging data from S4G, with a normalization derived using the same isochrone used to convert the RGB star counts into  $V$ -band surface brightness and stellar mass. NGC 7814 has power slopes of  $-3.71^{+0.99}_{-0.09} \pm 0.2$  along its minor axis, and  $-5.33^{+3.34}_{-0.57} \pm 0.2$  along its major axis. The implied axis ratio is  $0.59^{+0.14}_{-0.05} \pm 0.1$ . The implied stellar mass between 10 and 40 kpc is  $2.05^{+0.43}_{-0.26} \pm 0.6 \times 10^9 M_{\odot}$ , or a estimated total stellar halo mass of  $6.41^{+1.34}_{-0.81} \times 10^9 M_{\odot}$ , corresponding to  $\sim 14 \pm 6$  per cent of its total stellar mass. This is the largest stellar halo in our sample. NGC 7814 has a wide-format imaging from observations with small robotic telescopes (Martínez-Delgado et al. 2010b) and does not show any signs of tidal streams (Javanmardi et al. 2016).

## 7 DISCUSSION AND MODEL COMPARISONS

### 7.1 Comparison of stellar halo properties between galaxies

With the inferred stellar halo masses, axis ratios, and power-law slopes in hand, along with an idea of the likely sources of systematic uncertainty, we turn to exploring correlations among the GHOSTS galaxies between these halo properties and their halo stellar populations (Monachesi et al. 2016a) as well as compare with the bulk properties of the MW's and M31's stellar haloes.

We restrict our comparisons to quantities which are well constrained by the data in hand. Accordingly, we choose to characterize the stellar halo mass using the mass between 10 and 40 minor axis equivalent kpc,  $M_{10-40}$ ; minor axis power-law slopes are measured over a similar range. We characterize the stellar halo metallicity by quoting a derived  $[\text{Fe}/\text{H}]$  value at 30 kpc along the minor axis following the observational calibration of  $[\text{Fe}/\text{H}]$  as a function of RGB colours for globular clusters (Streich et al. 2014) assuming  $[\alpha/\text{Fe}] = 0.3$ . Instead of presenting minor axis metallicity gradients, we choose to present minor axis RGB colour gradients per kpc; RGB colour is related in a highly non-linear way to metallicity, making RGB colour gradient more robust to possible future changes in RGB colour calibration than an inferred metallicity gradient.

The properties of M31 and the MW are compiled from a variety of sources. Stellar masses for the MW and M31 are assumed to be  $6.1 \pm 1.1 \times 10^{10}$  and  $10.3 \pm 2.3 \times 10^{10} M_{\odot}$ , taken from Licquia & Newman (2015) and Sick et al. (2015), respectively. Rotation velocities for the MW and M31 are adopted from Bovy et al. (2012);  $V_c = 218 \pm 6 \text{ km s}^{-1}$  and HyperLEDA<sup>8</sup>  $V_{c,M31} = 257 \pm 6 \text{ km s}^{-1}$ ,

respectively. The stellar halo mass for M31 outside 27 kpc is estimated to be  $\sim 1.1 \times 10^{10} M_{\odot}$  and has a 3D (2D) density slope of roughly  $\sim -3.7$  ( $-2.7$ ) (Ibata et al. 2014). Extrapolation of the profile inside 27 kpc is obviously uncertain; we assume a total mass  $M_{\text{halo}, M31} = 1.5 \pm 0.5 \times 10^{10} M_{\odot}$  in what follows, with a corresponding estimate of  $M_{10-40, M31}$  three times smaller at  $M_{10-40, M31} \sim 5 \pm 2 \times 10^9 M_{\odot}$ . The total stellar halo mass for the MW has been estimated to be  $M_{\text{halo}, MW} = 4-7 \times 10^8 M_{\odot}$  (Bland-Hawthorn & Gerhard 2016, following Bell et al. 2008); we assume a  $M_{10-40, MW} = 1.75 \pm 0.5 \times 10^8 M_{\odot}$  three times smaller than the total stellar halo mass. There is evidence that the 3D halo density profile changes slope from  $-2.5$  to  $-3.5$  at halo radii of around 25 kpc (Bell et al. 2008; Sesar et al. 2011; Xue et al. 2015), accordingly we assume a 3D power-law slope of  $-3 \pm 0.5$ . Following Monachesi et al. (2016a), the metallicity at 30 kpc of M31's halo is from Gilbert et al. (2014), correcting their values to an assumed alpha enhancement  $[\alpha/\text{Fe}] = 0.3$ . The metallicity at 30 kpc of the MW's halo is the mean metallicity between the values reported in Sesar et al. (2011) and Xue et al. (2015), i.e.  $[\text{Fe}/\text{H}] = -1.7$ . Colour gradients for the MW and M31 are estimated from the metallicities at 10 and 40 kpc from Xue et al. (2015) and Gilbert et al. (2014), respectively, using the Streich et al. (2014) relationship between metallicity and RGB colour, assuming  $[\alpha/\text{Fe}] = 0.3$ .

We note that resolved stellar populations data in the halo of the large S0 galaxy NGC 3115 (Peacock et al. 2015) exist along its minor axis, and at very large radii in the elliptical galaxy Centaurus A (Rejkuba et al. 2014). We choose not to include these galaxies in our comparisons at this time owing to important differences between the halo profile and mass estimation techniques. In the case of NGC 3115, Peacock et al. (2015) lack a measurement along NGC 3115's major axis, making it impossible to estimate the stellar halo axis ratio. Furthermore, Peacock et al. identify a low-metallicity tail with  $[\text{Fe}/\text{H}] < -1$  as NGC 3115's stellar halo, whereas we identify all material at large minor axis radius as stellar halo. Our approach makes sense for galaxies dominated by a geometrically thin stellar disc, but may be less appropriate for a bulge-dominated galaxy such as NGC 3115. In the case of Centaurus A, Rejkuba et al. (2014) derive no estimate of total stellar halo mass, and again it is unclear how to proceed with quantitative halo analysis in an elliptical galaxy, where one expects a large number of stars from the central parts of the galaxy to have been scattered to large radii during the violent relaxation that shapes the main body of the galaxy. We note that both galaxies have metallicities  $[\text{Fe}/\text{H}]_{30 \text{ kpc}} > -0.6$  and appear to have fairly substantial masses (in excess of  $10^{10} M_{\odot}$ ) between 10 and 40 kpc, so appear to be qualitatively consistent with the trends discussed here using the GHOSTS sample augmented with the MW and M31.

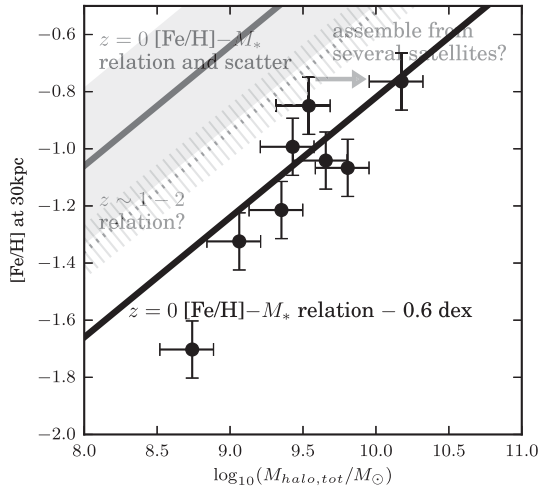
In panel a of Fig. 12, we show the ratio of the stellar halo mass to the total galaxy stellar mass, as a function of total galaxy stellar mass. This sample of MW mass disc-dominated galaxies with total stellar masses between  $4 \times 10^{10}$  and  $10^{11} M_{\odot}$  shows a remarkably large range of stellar halo mass fractions, varying by a factor of  $\sim 7$  in GHOSTS galaxies alone, with variations of a factor of  $\sim 15$  in stellar halo mass fractions with the addition of the MW and M31 to the sample.

In panel d of Fig. 12, we show the inferred stellar halo minor axis 3D density power-law slope as a function of inferred stellar halo mass between 10 and 40 kpc. The stellar haloes in nearby MW mass disc galaxies have masses between 10 and 40 kpc that range between  $\sim 10^{8.2}$  and  $10^{9.7} M_{\odot}$ , a factor of 30 range in mass. These galaxies show a range in halo power-law slopes between 10 and 40 kpc, with 3D equivalent minor axis power-law slopes

<sup>8</sup> <http://leda.univ-lyon1.fr/>; see also Makarov et al. (2014).







**Figure 14.** The  $[\text{Fe}/\text{H}] - M_{\text{halo, tot}}$  relation for the GHOSTS galaxies, the MW and M31. Total halo masses are estimated by multiplying the masses between 10 and 40 kpc by a factor of 3. The  $[\text{Fe}/\text{H}]$  is characterized by the  $[\text{Fe}/\text{H}]$  at 30 kpc along the minor axis. Overplotted is the  $z = 0$  galaxy stellar metallicity–stellar mass relation in grey (with  $\sim 0.2$  dex scatter), and a relation offset by  $-0.6$  dex in black to give a rough estimate in the offset between the galaxy and stellar halo metallicity–mass relations. The dashed area is the  $z \sim 2$  gas metallicity–mass relation, shifted by  $\sim -0.3$  dex from the present day relation. One can see that the relationship between stellar halo mass and metallicity has a shape that is broadly consistent with the galaxy metallicity–mass relation offset to slightly lower metallicities, as would be broadly expected if haloes are assembled from the debris of the disruption of multiple dwarf galaxies at intermediate redshifts ( $z = 1-2$ ).

Panels c and f of Fig. 12 show the minor axis stellar halo RGB colour gradient (a proxy for metallicity gradient) as a function of stellar halo mass between 10 and 40 kpc (panel c) and rotation velocity (panel f). As highlighted in Monachesi et al. (2016a), MW mass disc galaxies appear to have a range of behaviours in terms of their metallicity gradients, where  $\sim 1/2$  of the sample have no or a weak metallicity gradient, and the other half have strong metallicity gradients. The gradient does not appear to vary systematically with stellar halo mass. In our data set, halo metallicity gradient appears tentatively to favour steeper negative gradients for galaxies with higher  $v_{\text{max}}$ , but with little statistical significance (Pearson correlation coefficient of  $-0.62$ , corresponding to a 10 per cent chance of being drawn from an uncorrelated data set).

## 7.2 A correlation between stellar halo metallicity and stellar halo mass

The most significant correlation between stellar halo properties in our data set is a correlation between the mass of stellar haloes and their metallicity at 30 kpc along the minor axis (panel b of Fig. 12). In Fig. 14, we scale Fig. 12 to ‘total’ stellar halo masses, estimated by multiplying the  $M_{10-40}$  values by a factor of 3, following Section 5. A very rough fit to the data (using orthogonal distance regression) is:  $[\text{Fe}/\text{H}]_{30\text{kpc}} \sim (-1.45 \pm 0.1) + (0.7 \pm 0.15)(\log_{10} M_{\text{halo, tot}} - 9)$ ; the slope is changed significantly by the MW’s stellar halo metallicity estimate, and is accordingly rather uncertain. In order to build intuition about how to interpret this relation, we show in Fig. 14 the local galaxy stellar metallicity–stellar mass relation (obtained by tying together the stellar metallicity–stellar mass relations of Kirby et al. 2013 and Gallazzi et al. 2005) in grey along with an assumed scatter of  $\sim 0.2$  dex which is comparable to the  $< 0.15$  dex scatter of

Kirby et al. (2013) and the scatter at high stellar masses of Gallazzi et al. (2005). The black solid line is the local galaxy metallicity–mass relation offset by  $-0.6$  dex, which is scaled to go through the majority of the data points. The galaxy metallicity–mass relation is consistent with but slightly ( $< 2\sigma$ ) shallower than the stellar halo metallicity–mass relation.

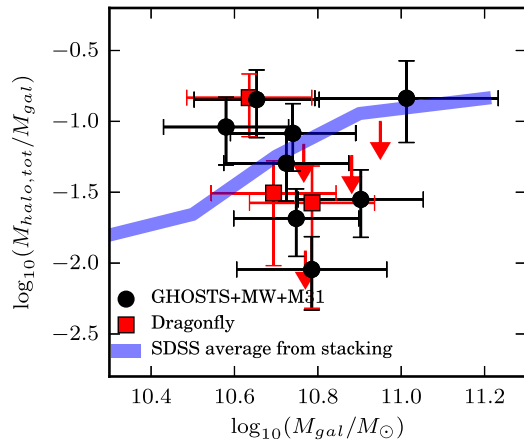
We argue the broad similarity of the slope of the stellar halo metallicity–halo mass relation with the galaxy metallicity–mass relation is no coincidence, and that its offset is broadly as expected in a cosmological context. Accretion-only models of halo formation in a cosmological context predict that most of the mass in stellar haloes should come from the few most massive progenitors (e.g. Deason, Mao & Wechsler 2016), and that the main epoch of halo building is at  $z \sim 1-2$  (Bullock et al. 2001; Bullock & Johnston 2005; Cooper et al. 2010). As a (very rough) guide to what kind of offset one might expect at earlier times from the metallicity–mass relation, Erb et al. (2006) find that the  $z \sim 2$  gas metallicity–mass relation is offset by  $\sim -0.3$  dex from the present day relation (shown as a dashed area in Fig. 14). Noting that most of the mass in stellar haloes should be formed from the disruption of the largest progenitors, the remaining offset should be interpreted as a largely horizontal offset (the arrow in Fig. 14), with lower mass galaxies being incorporated into a larger halo, where the relations are broadly consistent with the idea that the few largest progenitors provide most of the mass of a stellar halo at the present day, at a given metallicity characteristic of their metallicity at the time of accretion. While this argument is both necessarily rough and approximate, it is clear that there is an intuitive and quantitatively reasonable accretion-only framework in which to interpret the relation between the present-day stellar halo metallicity along the minor axis and the stellar halo mass (see Deason et al. 2016 for an in-depth discussion of this issue). This accretion-only framework appears also consistent with recent results from hydrodynamical simulations, where the *in situ* halo component is expected to be negligible at  $R \gtrsim 15$  kpc along the minor axis (Pillepich et al. 2015; Monachesi et al. 2016b). Thus, characterizing the halo metallicity with its metallicity along the minor axis yields a robust measurement of the properties of the dwarf satellites that were tidally disrupted by the central galaxy.

## 7.3 Comparison of stellar halo mass fractions with observations

A number of estimates or upper limits for stellar halo masses have recently become available from integrated light studies (D’Souza et al. 2014; Merritt et al. 2016). While there are no galaxies in common between the data sets, one can fruitfully compare the fraction of mass in stellar haloes for MW mass galaxies between these studies.

In Fig. 15, we show the overall good level of agreement between estimated fractions of mass in stellar haloes from GHOSTS, MW, and M31 (this paper; black) and two other studies based on integrated light.

Red symbols denote detections or upper limits of stellar haloes of galaxies with stellar masses above  $4 \times 10^{10} M_{\odot}$  from deep imaging with the Dragonfly telescope array (Merritt et al. 2016). We interpret galaxies with halo fraction error bars that intersect zero to be non-detections, and take the upper error bar from their table 1 to be the upper limit to the halo mass fraction. An additional complication is that their estimated stellar halo mass fraction refers to light measured outside 5 half-mass radius  $R_h$  only. We correct their estimates to an estimate of total stellar halo mass following Section 5. We convert major axis  $R_h$  into minor axis equivalent



**Figure 15.** A comparison of estimates of the fraction of stars in MW mass galaxies that are in its stellar halo. Black points denote the measurements from GHOSTS from this paper, red data points show three estimates and four upper limits from Merritt et al. (2016), and blue shows an estimate of average accreted fraction for low concentration galaxies using stacking of SDSS images (D’Souza et al. 2014, private communication).

**Table 4.** Estimated ‘total’ stellar halo mass fraction from Dragonfly imaging from the values presented in Merritt et al. (2016).

Galaxy	Stellar halo fraction $f_{\text{halo,tot}}$
NGC 1084	$0.15 \pm 0.07$
NGC 2903	$0.031 \pm 0.021$
NGC 3351	$<0.068$
NGC 3368	$<0.099$
NGC 4220	$0.027 \pm 0.022$
NGC 4258	$<0.057$
M101	$<0.012$

values using axis ratio estimates from the 2MASS Large Galaxy Atlas Jarrett et al. (2003) when possible or from the NASA/IPAC Extragalactic Database. For most of their sample,  $5R_h$  corresponds to a minor axis radius of  $\sim 10$  kpc, and we adopt the factor of 3 correction determined in Section 5 to scale their measurements to ‘total’ mass. NGC 4220 has a smaller minor axis equivalent radius of  $\sim 4$  kpc, corresponding to a rough factor of two extrapolation to total mass. Merritt et al.’s imaging for M101, owing to its low inclination and large disc extent, is sensitive to a stellar halo only outside of  $\sim 30$  kpc, and following Section 5 we estimate a factor of ten extrapolation to total stellar mass. The adopted stellar halo mass fractions or limits are tabulated in Table 4. We find that the treatment of Dragonfly non-detections as upper limits, and a model-motivated extrapolation to total stellar halo fraction brings the measurements of Merritt et al. (2016) into line with ours. These studies paint a consistent picture of there being around a factor of  $\sim 15$  full range in stellar halo mass fractions for roughly MW mass galaxies, with the Dragonfly non-detections being clustered towards the lower side of our observed range of stellar halo fractions.

Estimates of the average accreted stellar mass fraction for low-concentration disc-dominated galaxies from stacking of the SDSS (D’Souza et al. 2014) are given in blue. D’Souza et al. (2014) give an estimate of the outer light fraction; we adjust it downwards by 0.1 dex to account for stellar M/L differences between the bluer outer envelopes of galaxies and their redder, higher stellar M/L cores, using a typical colour difference of  $\Delta(g-r) \sim 0.1$  from

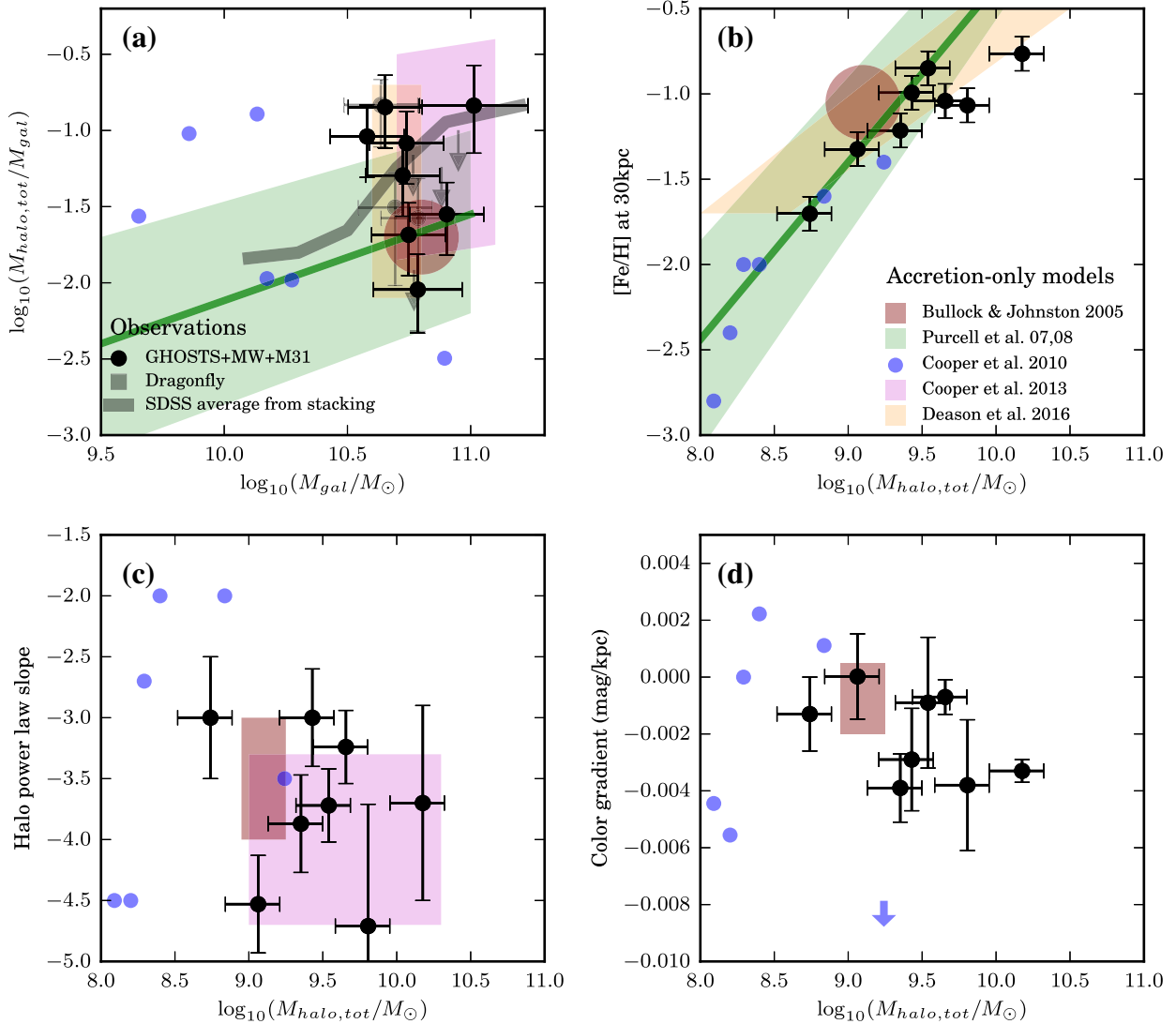
D’Souza et al. (2014) and estimates of stellar M/L from Bell et al. (2003). Furthermore, D’Souza (private communication) has found using stacking of mock images from the Illustris hydrodynamical simulations (Vogelsberger et al. 2014) that the outer light fraction is  $\sim 0.1$  dex systematically higher than the fraction of accreted stars in those simulations. Accordingly, following D’Souza’s recommendation, we adjust their accreted light fraction down by another 0.1 dex. This line, denoted in blue, represents the best available estimate from SDSS stacking of the accreted fraction of stars in low-concentration galaxies. This average is in good accord with the typical stellar halo fractions measured by GHOSTS or Dragonfly. By focusing on individual haloes, GHOSTS and Dragonfly enrich the results of D’Souza et al. (2014) by showing that galaxies have considerable scatter in stellar halo mass around that relation, with a full order of magnitude or more range in halo mass fraction at stellar masses comparable to the MW.

## 7.4 Comparison between stellar halo observations and models of stellar halo formation

Given the diversity of halo properties seen in Figs 12 and 13, and the strength of the correlation between halo metallicity along the minor axis and stellar halo mass seen in Figs 12 and 14, we turn now to a preliminary and relatively rudimentary comparison between these observational results with expectations from models of galaxy formation in a cosmological context. We make this comparison as fair as possible, noting that the majority of the predicted quantities, such as metallicity and metallicity gradients, are typically presented as spherically averaged properties in models whereas we measure them as projected along the disc minor axis, thus some discrepancies may arise due to the different methodology used. We note that models of stellar halo formation in a cosmological context is a rapidly developing field; in this spirit, we chart out some broad patterns and ideas, but leave detailed comparisons to future works and works by those modelling stellar halo formation in a cosmological context. In what follows, we focus on comparing expectations and predictions for the stellar haloes of disc galaxies with roughly the MW’s total (dark plus baryonic) mass of  $0.5\text{--}2 \times 10^{12} M_\odot$  with our observations, cognizant of the considerable difficulty in measuring the dark halo mass of our and nearby galaxies.

### 7.4.1 Accretion-only models

We first compare with models which account for the build-up of stellar haloes from the tidal disruption of dwarf satellites only (accretion-only models) in Fig. 16. Light green shaded areas denote the region occupied by roughly 95 per cent of the accretion-only haloes from the halo occupation models of Purcell, Bullock & Zentner (2007, 2008). The Purcell et al. (2008) metallicity–stellar halo mass relation is approximated using their Fig. 7, assuming that their galaxies in  $\sim 10^{12} M_\odot$  haloes have stellar masses  $\sim 10^{10.75} M_\odot$ . In order to convert metallicity into  $[\text{Fe}/\text{H}]$ , we assume  $[\text{Fe}/\text{H}]_{\text{Purcell}} \sim \log_{10} Z/Z_\odot - 0.2$ , assuming  $[\alpha/\text{Fe}] \sim 0.3$ . Brick red denotes the results from the Bullock & Johnston (2005) models (see also Font et al. 2006); these haloes have a narrow  $\sim \times 2$  range in stellar mass. Blue symbols denote individual haloes from high resolution  $N$ -body simulations by Cooper et al. (2010), where we again estimate iron abundance by subtracting 0.2 dex from the predicted metallicity. Magenta regions enclose the 68 per cent range of the structural properties of haloes expected to host galaxies with discs ( $B/T < 0.9$ ) modelled using the Millennium-II  $N$ -body simulations



**Figure 16.** Comparison to accretion-only models – Panel a: ratio of ‘total’ stellar halo mass and total stellar mass, as a function of total stellar mass. Panel b: stellar halo metallicity at 30 kpc as a function of ‘total’ stellar halo mass. Panel c: inferred 3D stellar halo density power-law slope as a function of ‘total’ stellar halo mass. Panel d: stellar halo colour gradient (a proxy for metallicity gradient) as a function of ‘total’ stellar halo mass. The observational data are shown in black and grey. Models: brick red area: Bullock & Johnston (2005), light green+line: Purcell et al. (2007, 2008), blue: Cooper et al. (2010), magenta: Cooper et al. (2013), orange: Deason et al. (2016).

by Cooper et al. (2013). Orange regions enclose the range of predictions from Deason et al. (2016).

Accretion-only models, for the most part, predict a reasonably realistic range of stellar halo masses (panel a of Fig. 16). Similarly, the broad range of power-law slopes of the accretion-only models seen in panel c of Fig. 16 appears to be in reasonable accord with the data, with the possible exception of the Cooper et al. (2010) models, which may have too wide a diversity in halo density profiles. Panel d of Fig. 16 suggests that the weakness and uniformity of metallicity gradients of the Bullock & Johnston (2005) models is in conflict with the observations – the variability of stellar halo metallicity gradients from galaxy to galaxy in the Cooper et al. (2010) models appears to match more accurately the observational data.

Most notably, and as foreshadowed in Section 7.2, the accretion-only models predict a strong stellar halo mass–metallicity relation,

in striking accord with the data. While the form of the relation was argued to be inevitable, given the stellar mass–stellar metallicity relation of the progenitor satellites (which is typically baked into the accretion-only models as a constraint) and given the almost negligible contribution from *in situ* halo stars at the radius along the minor axis where the [Fe/H] was derived (see Pillepich et al. 2015; Monachesi et al. 2016b), the normalization of the relation compared to the galaxy metallicity–mass relation is less trivial and appears to have been predicted accurately by the models. This match in normalization suggests that the accretion-only models have assembled their stellar haloes from appropriate progenitors with reasonable metallicities at the time of satellite in fall. Note that a broad relation between halo metallicity and total galaxy luminosity was suggested by Mouhcine et al. (2005); the observations (and models) predict that a broad relation of this type is expected, but it is driven by a much more fundamental stellar halo mass–metallicity



relation – bigger stellar haloes are assembled from bigger, more metal-rich pieces (Deason et al. 2016, see also Amorisco 2016).

This overall impressive level of agreement is encouraging. On one hand, the disagreements are now subtle enough that differences between observational and model metrics will matter, motivating an effort by modellers and observers to compare more consistently with each other in an effort to refine the models and better interpret the observations. On the other hand, this agreement motivates the use of stellar haloes to quantitatively explore galactic merger histories: Deason et al. (2016) show that there is a relationship between the stellar halo mass, metallicity, and the properties of the most massive progenitor, suggesting that we can gain precious insight into the most massive mergers that affected many of our nearby galactic neighbours.

#### 7.4.2 Hydrodynamical models

We now compare with hydrodynamical models, which account for both the accreted plus *in situ* build-up of stellar haloes, in Fig. 17. Hydrodynamical models attempt to more completely capture the gas, stellar, and feedback physics of galaxy formation; they are both complex and computationally expensive. Importantly, hydrodynamical simulations also predict that an *in situ* stellar halo should exist, with contributions from both early chaotically distributed star formation during the assembly of the galaxy and stars kicked up from deeper in the potential well (see Cooper et al. 2015 for a thoughtful discussion of the origin of the *in situ* stellar halo). These *in situ* haloes vary from model to model in prominence (Cooper et al. 2015 emphasize the importance of different choices for sub-grid physics on the properties of the *in situ* halo), but typically carry  $>5 \times 10^9 M_{\odot}$  (Zolotov et al. 2009; Cooper et al. 2015; Pillepich et al. 2015). The *in situ* halo appears to show a metallicity difference from the accreted stars in most models, although both components have broad distributions and the distribution of *in situ* metallicities is likely to be sensitive to input physics (see e.g. fig. 6 in Cooper et al. 2015 and the illuminating and forthright discussion in Zolotov et al. 2010). Pillepich et al. (2015) and Monachesi et al. (2016b) emphasize that the *in situ* halo may be highly flattened, and their metallicity signatures may be difficult to discern along the minor axis.

The broad properties of hydrodynamical models vary considerably from model to model (Fig. 17), but the models appear for the most part to overpredict stellar halo masses (panel a; e.g. light cyan for Font et al. 2011, green for Tissera, White & Scannapieco 2012; Tissera et al. 2014, red for Cooper et al. 2015, and yellow for Pillepich et al. 2015). This overprediction of stellar halo mass seems particularly acute – nearly an order of magnitude – for lower resolution hydrodynamical simulations. While we show examples from Font et al. (2011) and Tissera et al. (2012, 2014), similar behaviour is seen in the simulated haloes of e.g. Zolotov et al. (2009), Bailin et al. (2014), and Pillepich et al. (2014). Such a large discrepancy implies that both the accreted and *in situ* parts of the halo are overproduced in such lower resolution models, as both components separately violate observational constraints. More current simulations with substantially higher resolution (Cooper et al. 2015; Pillepich et al. 2015) have stellar haloes that are in rather better accord with the observations, and appear to overpredict halo mass by a factor of less than 3, where systematic differences in how stellar haloes are defined between simulations and the observations may play an important role (see e.g. the discussion in Pillepich et al. 2015).

Hydrodynamical model stellar halo metallicities at 30 kpc (panel b of Fig. 17) are similar to the observational estimates at  $[Fe/H] \sim -1.2$ , but would be rather low for their halo stellar mass. Lower resolution hydrodynamical models (cyan and green) do not reproduce the observed strong correlation between stellar halo mass and stellar halo metallicity. More current, higher resolution simulations (red and yellow) have similar metallicities, but owing to their lower and somewhat more realistic stellar masses lie closer to the observed galaxies in the stellar halo mass–stellar halo metallicity plane. With only four simulations to explore, it is too early to say if high-resolution hydrodynamical simulations naturally reproduce the observed stellar halo metallicities–stellar halo mass correlation. Like Pillepich et al. (2015) and Monachesi et al. (2016b), we caution that with the moderate level of discrepancy now seen between high-resolution hydrodynamical models and the observations the choice of observational metric matters; accordingly, it will be important to estimate halo masses and metallicities of the simulated haloes in ways that connect well with these observations.

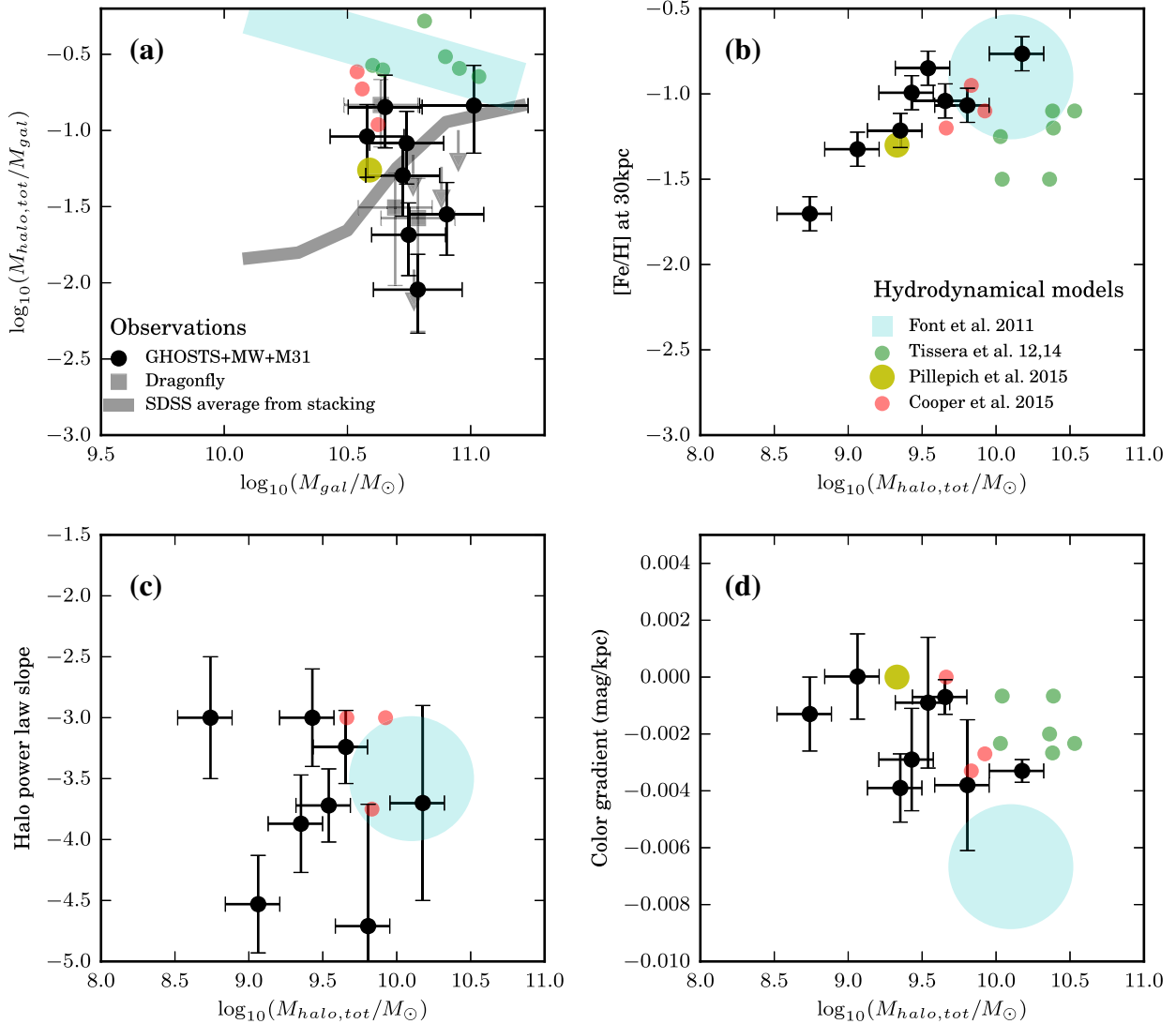
Only two simulations predict halo power-law slopes, and these simulations broadly match the observational constraints (panel c). In addition, and not shown, their axis ratios tend to be  $c/a \sim 0.6$  but with considerable range in axis ratio (e.g. McCarthy et al. 2012 for the Font et al. 2011 models), in reasonable agreement with the observations.

Finally, a diversity of stellar halo metallicity gradients was predicted by current high-resolution hydrodynamical simulations – Cooper et al. (2015) and Pillepich et al. (2015) – and appears to be in accord with the data (panel d of Fig. 17). The simulations of Tissera et al. (2012, 2014) also reproduce the observed diversity of metallicity gradients, remembering that their overall stellar halo masses are generally dramatically overpredicted. Observations appear to rule out ubiquitous and large metallicity gradients of the kind predicted by Font et al. (2011). We caution that our minor-axis metallicity gradients may be relatively insensitive to exploring the fraction of stars from *in situ* formation of halo stars; instead, major axis metallicity gradients may prove a more decisive probe of the importance of *in situ* stars in stellar haloes, owing to their predicted high degree flattening in current high-resolution simulations (Pillepich et al. 2015; Monachesi et al. 2016b).

#### 7.4.3 Degree of substructure

While metrics of the number and prominence of individual streams or shells have been proposed and discussed (e.g. Johnston et al. 2008; Atkinson, Abraham & Ferguson 2013; Amorisco 2015; Hendel & Johnston 2015), fewer works have quantified the degree of substructure of the ‘aggregate’ stellar halo (e.g. Bell et al. 2008; Amorisco 2017). Our prime observational measure of the degree of substructure is the intrinsic scatter that must be included to our power-law model in order to make it an acceptable fit to the data, similar in spirit to the RMS/total estimates of Bell et al. (2008) and the tidal parameter estimates of elliptical galaxies by Tal et al. (2009). For our sample, the typical intrinsic scatter around the fit is 0.05–0.1 dex, ranging from undetectable to 0.14 dex for the highly structured major axis profile of NGC 3031.

This metric for the degree of substructure of the ‘aggregate’ halo has been very rarely calculated by modellers. Our own fits to the Bullock & Johnston (2005) simulations (hybrid disc+bulge+dark halo potential plus *N*-body satellites) yield values of  $\sim 0.1$  dex, in broad accord with the observations. We do note that it is likely that some models would fail to match the observational constraints:



**Figure 17.** Comparison to hydrodynamical models – Panel a: ratio of ‘total’ stellar halo mass and total stellar mass, as a function of total stellar mass. Panel b: stellar halo metallicity at 30 kpc as a function of ‘total’ stellar halo mass. Panel c: inferred 3D stellar halo density power-law slope as a function of ‘total’ stellar halo mass. Panel d: stellar halo colour gradient (a proxy for metallicity gradient) as a function of ‘total’ stellar halo mass. The observational data are shown in black and grey. Hydrodynamical models: light cyan: Font et al. (2011), green: Tissera et al. (2012, 2014), yellow: Pillepich et al. (2015), red: Cooper et al. (2015).

the models of Cooper et al. (2010) have strong substructure, far in excess of that seen in the MW or in the Bullock & Johnston (2005) models (compare the RMS/total measures of Helmi et al. 2011 with Bell et al. 2008). Bailin et al. (2014) attribute the high degree of substructure in pure  $N$ -body only models (which display triaxial, very structured stellar haloes) to the lack of a potential from the main body of the galaxy; the potential from the main body of the galaxy appears to lead to precession which erases substructure and produces a more oblate halo, in better accord with the data. We strongly encourage simulators to produce quantitative estimates of ‘aggregate’ stellar halo substructure, as a crucial test of their input model physics.

## 8 SUMMARY AND CONCLUSIONS

We have examined the halo stellar populations of six galaxies from the GHOSTS survey (Radburn-Smith et al. 2011). HST/ACS and

HST/WFC3 data were used from fields observed along the major and minor axes of each halo. We construct CMDs from these observations and select RGB stars above the 50 per cent completeness limit to trace the stellar halo populations in these galaxies. We use the selected RGB stars to derive a stellar density profile for each halo. From the density profiles we estimate a best-fitting power-law slope and intrinsic scatter around a smooth power law, axis ratio, and stellar mass of each halo.

We find a diversity of stellar halo masses between minor axis equivalent 10–40 kpc of  $3\text{--}21 \times 10^8 M_{\odot}$  and projected power-law slopes of between  $-2$  and  $-3.7$  along the minor axes. Owing to substructure in stellar haloes (particularly prominent for example along NGC 891’s minor axis), we measure a typical intrinsic scatter around a smooth power-law fit of 0.05–0.1 dex. By comparing the densities of the minor axis and major axis profiles for each galaxy at distances around  $\sim 25$  kpc, we infer axis ratios at  $\sim 25$  kpc ranging from 0.4 to 0.75.

Using the 11 halo realizations from the Bullock & Johnston (2005) models as a guide for interpreting a richly structured halo using sparse pointings, we estimate systematic uncertainties for the inferred stellar halo masses, power-law slopes, and projected axis ratios for the GHOSTS galaxies. In particular, using the stellar halo mass measurements within 10–40 kpc for the Bullock & Johnston (2005) models and comparing them to the total stellar halo mass, we expect that the above 10–40 kpc halo masses should be around 30–40 per cent of the total stellar halo mass for an accretion-dominated halo. Consequently, we expect that the GHOSTS galaxies have ‘total’ stellar masses of around  $1\text{--}6 \times 10^9 M_{\odot}$ .

In conjunction with measurements for the stellar halo properties of the MW and M31, we find that the GHOSTS stellar haloes lay in between the extremes charted out by the (rather atypical) haloes of the MW and M31. Galaxies with stellar masses similar to the MW have an order of magnitude range in stellar halo mass, factors of several differences in characteristic minor axis halo metallicities, power-law profiles with best-fitting slopes varying between  $-2$  and  $-3.7$ , and a variety of metallicity gradients, where  $\sim 1/2$  of the sample have little to no measurable metallicity gradient. The sample shows a strong correlation between stellar halo metallicity and stellar halo mass.

We compare our observational results with the results of models of stellar halo formation in a cosmological context. We find good agreement between accretion-only models, where the stellar haloes are formed by the disruption of dwarf satellites, and the observations. In particular, the strong observed correlation between stellar halo metallicity and stellar halo mass is naturally reproduced by the models as the result of a strong metallicity–mass relation of the satellite progenitors, plus the tendency for more massive stellar haloes to have been formed by the disruption of larger progenitors. Low-resolution hydrodynamical models have unrealistically high stellar halo masses. Current high-resolution hydrodynamical models predict stellar halo masses somewhat higher than observed but in better accord with the data, with reasonable metallicities, metallicity gradients, and density profiles. The level of the differences between predictions and observations may be small enough that differences in definition between our observational and model metrics may be important.

## ACKNOWLEDGEMENTS

We thank the referee for their helpful comments and suggestions. We appreciate helpful conversations and feedback and insights from Sarah Loebman, Monica Valluri, Andrei Kravtsov, Nicolas Martin, and Oleg Gnedin. This work was supported by NSF grant AST 1008342 and HST grants GO-11613 and GO-12213 provided by NASA through a grant from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. Additionally, some of the data presented in this paper were obtained from the Mikulski Archive for Space Telescopes (MAST). STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. Support for MAST for non-HST data is provided by the NASA Office of Space Science via grant NNX09AF08G and by other grants and contracts. We acknowledge the usage of the HyperLeda data base (<http://leda.univ-lyon1.fr>). This research has made use of the NASA/IPAC Extragalactic Database (NED) which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. This work has made use of the IAC-STAR Synthetic CMD

computation code. IAC-STAR is supported and maintained by the IACs IT Division. This work used the astronomy & astrophysics package for MATLAB (Ofek 2014). This research has made use of NASA’s Astrophysics Data System Bibliographic Services. This research made use of Astropy, a community-developed core Python package for Astronomy (Astropy Collaboration et al. 2013).

## REFERENCES

- Abraham R. G., van Dokkum P. G., 2014, *PASP*, 126, 55  
 Amorisco N. C., 2015, *MNRAS*, 450, 575  
 Amorisco N. C., 2016, in Gil de Paz A., Lee J. C., Knapen J. H., eds, *Proc. IAU Symp. 321, Formation and Evolution of Galaxy Outskirts*. Cambridge Univ. Press, Cambridge  
 Amorisco N. C., 2017, *MNRAS*, 464, 2882  
 Anderson J., Bedin L. R., 2010, *PASP*, 122, 1035  
 Astropy Collaboration et al., 2013, *A&A*, 558, A33  
 Atkinson A. M., Abraham R. G., Ferguson A. M. N., 2013, *ApJ*, 765, 28  
 Bailin J., Bell E. F., Chappell S. N., Radburn-Smith D. J., de Jong R. S., 2011, *ApJ*, 736, 24  
 Bailin J., Bell E. F., Valluri M., Stinson G. S., Debattista V. P., Couchman H. M. P., Wadsley J., 2014, *ApJ*, 783, 95  
 Barker M. K., Ferguson A. M. N., Irwin M., Arimoto N., Jablonka P., 2009, *AJ*, 138, 1469 (B09)  
 Beck R., Hutschenreiter G., Wielebinski R., 1982, *A&A*, 106, 112  
 Bell E. F., de Jong R. S., 2001, *ApJ*, 550, 212  
 Bell E. F., McIntosh D. H., Katz N., Weinberg M. D., 2003, *ApJS*, 149, 289  
 Bell E. F. et al., 2008, *ApJ*, 680, 295  
 Bellazzini M., Ferraro F. R., Pancino E., 2001, *ApJ*, 556, 635  
 Bertin E., Arnouts S., 1996, *A&AS*, 117, 393  
 Bland-Hawthorn J., Gerhard O., 2016, *ARA&A*, 54, 529  
 Bovy J. et al., 2012, *ApJ*, 759, 131  
 Bressan A., Marigo P., Girardi L., Salasnich B., Dal Cero C., Rubele S., Nanni A., 2012, *MNRAS*, 427, 127  
 Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000  
 Bullock J. S., Johnston K. V., 2005, *ApJ*, 635, 931 (BJ05)  
 Bullock J. S., Kravtsov A. V., Weinberg D. H., 2001, *ApJ*, 548, 33  
 Chabrier G., 2003, *PASP*, 115, 763  
 Chen Y., Girardi L., Bressan A., Marigo P., Barbieri M., Kong X., 2014, *MNRAS*, 444, 2525  
 Cohen J., Sesar B., Banholzer S., 2015, *IAU General Assembly*, 22, 2255152  
 Cole S., 1991, *ApJ*, 367, 45  
 Cooper A. P. et al., 2010, *MNRAS*, 406, 744  
 Cooper A. P., D’Souza R., Kauffmann G., Wang J., Boylan-Kolchin M., Guo Q., Frenk C. S., White S. D. M., 2013, *MNRAS*, 434, 3348  
 Cooper A. P., Parry O. H., Lowing B., Cole S., Frenk C., 2015, *MNRAS*, 454, 3185  
 Crojević D. et al., 2016, *ApJ*, 823, 19  
 de Jong R. S., 2008, *MNRAS*, 388, 1521  
 de Jong R. S. et al., 2007, *ApJ*, 667, L49  
 de Jong R. S., Radburn-Smith D. J., Sick J. N., 2009, in Andersen J., Nordström B., Bland-Hawthorn J., eds, *Proc. IAU Symp. 254, The Galaxy Disk in Cosmological Context*. Cambridge Univ. Press, Cambridge, p. 19  
 D’Souza R., Kauffman G., Wang J., Vegetti S., 2014, *MNRAS*, 443, 1433  
 Davidge T. J., 2010, *ApJ*, 725, 1342  
 Deason A. J., Belokurov V., Evans N. W., 2011, *MNRAS*, 416, 2903  
 Deason A. J., Belokurov V., Evans N. W., Johnston K. V., 2013, *ApJ*, 763, 113  
 Deason A. J., Belokurov V., Koposov S. E., Rockosi C. M., 2014, *ApJ*, 787, 30  
 Deason A. J., Mao Y.-Y., Wechsler R. H., 2016, *ApJ*, 821, 5  
 Dekel A., Silk J., 1986, *ApJ*, 303, 39  
 Dolphin A. E., 2000, *PASP*, 112, 1383  
 Erb D. K., Shapley A. E., Pettini M., Steidel C. C., Reddy N. A., Adelberger K. L., 2006, *ApJ*, 644, 813

- Font A. S., Johnston K. V., Bullock J. S., Robertson B. E., 2006, *ApJ*, 638, 585
- Font A. S., McCarthy I. G., Crain R. A., Theuns T., Schaye J., Wiersma R. P. C., Dalla Vecchia C., 2011, *MNRAS*, 416, 2802
- Gallazzi A., Charlot S., Brinchmann J., White S. D. M., Tremonti C. A., 2005, *MNRAS*, 362, 41
- Gilbert K. M. et al., 2012, *ApJ*, 760, 76
- Gilbert K. M. et al., 2014, *ApJ*, 796, 76
- Girardi L., Groenewegen M. A. T., Hatziminaoglou E., da Costa L., 2005, *A&A*, 436, 895
- Gómez F. A., Coleman-Smith C. E., O’Shea B. W., Tumlinson J., Wolpert R. L., 2012, *ApJ*, 760, 112
- Gonzaga S., 2012, *The DrizzlePac Handbook*. Space Telescope Science Institute (STScI), Baltimore, Maryland
- Greggio L., Rejkuba M., Gonzalez O. A., Arnaboldi M., Iodice E., Irwin M., Neeser M. J., Emerson J., 2014, *A&A*, 562, A73
- Guhathakurta P. et al., 2005, *BAAS*, 207, 1386
- Harris W. E., Harris G. L. H., 2002, *AJ*, 123, 3108
- Harris W. E., Harris G. L. H., Layden A. C., Stetson P. B., 2007a, *AJ*, 134, 43
- Harris W. E., Harris G. L. H., Layden A. C., Wehner E. M. H., 2007b, *ApJ*, 666, 903
- Helmi A., Cooper A. P., White S. D. M., Cole S., Frenk C. S., Navarro J. F., 2011, *ApJ*, 733, L7
- Hendel D., Johnston K. V., 2015, *MNRAS*, 454, 2472
- Hoyle F., Schwarzschild M., 1955, *ApJS*, 2, 1
- Ibata R. A. et al., 2014, *ApJ*, 780, 128
- Ivezić Ž. et al., 2000, *AJ*, 120, 963
- Jarrett T. H., Chester T., Cutri R., Schneider S. E., Huchra J. P., 2003, *AJ*, 125, 525
- Javanmardi B. et al., 2016, *A&A*, 588, A89
- Johnston K. V., Bullock J. S., Sharma S., Font A., Robertson B. E., Leitner S. N., 2008, *ApJ*, 689, 936
- Jurić M. et al., 2008, *ApJ*, 673, 864
- Kauffmann G. et al., 2003, *MNRAS*, 341, 33
- Kirby E. N., Cohen J. G., Guhathakurta P., Cheng L., Bullock J. S., Gallazzi A., 2013, *ApJ*, 779, 102
- Licquia T. C., Newman J. A., 2015, *ApJ*, 806, 96
- Makarov D., Prugniel P., Terekhova N., Courtois H., Vauglin I., 2014, *A&A*, 570, A13
- Malin D., Hadley B., 1997, *PASA*, 14, 52
- Martínez-Delgado D. et al., 2010a, *AJ*, 140, 962
- Martínez-Delgado D. et al., 2010b, *AJ*, 140, 962
- McCarthy I. G., Schaye J., Font A. S., Theuns T., Frenk C. S., Crain R. A., Dalla Vecchia C., 2012, *MNRAS*, 427, 379
- Merritt A., van Dokkum P., Abraham R., Zhang J., 2016, *ApJ*, 830, 62
- Mihos J. C., Harding P., Feldmeier J., Morrison H., 2005, *ApJ*, 631, L41
- Miller E., 1996, *Electronic Thesis or Dissertation*
- Monachesi A. et al., 2013, *ApJ*, 766, 106
- Monachesi A., Bell E. F., Radburn-Smith D. J., Bailin J., de Jong R. S., Holwerda B., Streich D., Silverstein G., 2016a, *MNRAS*, 457, 1419
- Monachesi A., Gómez F. A., Grand R. J. J., Kauffmann G., Marinacci F., Pakmor R., Springel V., Frenk C. S., 2016b, *MNRAS*, 459, L46
- Moore B., Ghigna S., Governato F., Lake G., Quinn T., Stadel J., Tozzi P., 1999, *ApJ*, 524, L19
- Mouhcine M., Ferguson H. C., Rich R. M., Brown T. M., Smith T. E., 2005, *ApJ*, 633, 821
- Mouhcine M., Ibata R., Rejkuba M., 2010, *ApJ*, 714, L12
- Muñoz-Mateos J. C. et al., 2015, *ApJS*, 219, 3
- Naeslund M., Joersaeter S., 1997, *A&A*, 325, 915
- Newberg H. J., Yanny B., Cole N., Beers T. C., Re Fiorentin P., Schneider D. P., Wilhelm R., 2007, *ApJ*, 668, 221
- Ofek E. O., 2014, *MATLAB package for astronomy and astrophysics*, Astrophysics Source Code Library, record ascl:1407.005
- Okamoto S., Arimoto N., Ferguson A. M. N., Bernard E. J., Irwin M. J., Yamada Y., Utsumi Y., 2015, *ApJ*, 809, L1
- Paudel S. et al., 2013, *ApJ*, 767, 133
- Peacock M. B., Strader J., Romanowsky A. J., Brodie J. P., 2015, *ApJ*, 800, 13
- Pillepich A. et al., 2014, *MNRAS*, 444, 237
- Pillepich A., Madau P., Mayer L., 2015, *ApJ*, 799, 184
- Purcell C. W., Bullock J. S., Zentner A. R., 2007, *ApJ*, 666, 20
- Purcell C. W., Bullock J. S., Zentner A. R., 2008, *MNRAS*, 391, 550
- Radburn-Smith D. J. et al., 2011, *ApJS*, 195, 18 (RS11)
- Radburn-Smith D. J. et al., 2012, *ApJ*, 753, 138
- Radburn-Smith D. J. et al., 2014, *ApJ*, 780, 105
- Rejkuba M., Harris W. E., Greggio L., Harris G. L. H., Jerjen H., Gonzalez O. A., 2014, *ApJ*, 791, L2
- Robertson B., Bullock J. S., Font A. S., Johnston K. V., Hernquist L., 2005, *ApJ*, 632, 872
- Robin A. C., Reylé C., Derrière S., Picaud S., 2003, *A&A*, 409, 523
- Sandage A., Smith L. L., 1966, *ApJ*, 144, 886
- Sesar B., Jurić M., Ivezić Ž., 2011, *ApJ*, 731, 4
- Shang Z. et al., 1998, *ApJ*, 504, L23
- Sheth K. et al., 2010, *PASP*, 122, 1397
- Sick J., Courteau S., Cuillandre J.-C., Dalcanton J., de Jong R., McDonald M., Simard D., Tully R. B., 2015, in *Capellari M., Courteau S., eds, Proc. IAU Symp. 311, Galaxy Masses as Constraints of Formation Models*. Cambridge Univ. Press, Cambridge, p. 82
- Slater C. T., Harding P., Mihos J. C., 2009, *PASP*, 121, 1267
- Slater C. T., Nidever D. L., Munn J. A., Bell E. F., Majewski S. R., 2016, *ApJ*, preprint ([arXiv:1601.02034](https://arxiv.org/abs/1601.02034))
- Streich D., de Jong R. S., Bailin J., Goudfrooij P., Radburn-Smith D., Vlajic M., 2014, *A&A*, 563, A5
- Streich D., de Jong R. S., Bailin J., Bell E. F., Holwerda B. W., Minchev I., Monachesi A., Radburn-Smith D. J., 2016, *A&A*, 585, A97
- Tal T., van Dokkum P. G., Nelan J., Bezanson R., 2009, *AJ*, 138, 1417
- Tang J., Bressan A., Rosenfield P., Slemer A., Marigo P., Girardi L., Bianchi L., 2014, *MNRAS*, 445, 4287
- Tissera P. B., White S. D. M., Scannapieco C., 2012, *MNRAS*, 420, 255
- Tissera P. B., Beers T. C., Carollo D., Scannapieco C., 2014, *MNRAS*, 439, 3128
- Trujillo I., Fliri J., 2016, *ApJ*, 823, 123
- Tumlinson J., 2010, *ApJ*, 708, 1398
- van Vledder I., van der Vlugt D., Holwerda B. W., Kenworthy M. A., Bouwens R. J., Trenti M., 2016, *MNRAS*, 458, 425
- Vogelsberger M. et al., 2014, *MNRAS*, 444, 1518
- Watkins A. E., Mihos J. C., Harding P., 2015, *ApJ*, 800, L3
- Wheeler C., Phillips J. I., Cooper M. C., Boylan-Kolchin M., Bullock J. S., 2014, *MNRAS*, 442, 1396
- White S. D. M., Rees M. J., 1978, *MNRAS*, 183, 341
- Xue X.-X., Rix H.-W., Ma Z., Morrison H., Bovy J., Sesar B., Janesh W., 2015, *ApJ*, 809, 144
- Yanny B. et al., 2000, *ApJ*, 540, 825
- Zolotov A., 2011, *PhD thesis*, New York University
- Zolotov A., Willman B., Brooks A. M., Governato F., Brook C. B., Hogg D. W., Quinn T., Stinson G., 2009, *ApJ*, 702, 1058
- Zolotov A., Willman B., Brooks A. M., Governato F., Hogg D. W., Shen S., Wadsley J., 2010, *ApJ*, 721, 738

## SUPPORTING INFORMATION

Supplementary data are available at [MNRAS](https://www.mnras.org) online.

**Table 2.** Star count data for the six GHOSTS galaxies examined in this article.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.