

# Reaching for the Edge I: Probing the Outskirts of Massive Galaxies with HSC, DECaLS, SDSS and Dragonfly

Jiaxuan Li<sup>1,2\*</sup>, Song Huang<sup>1,7</sup>, Alexie Leauthaud<sup>1</sup>, John Moustakas<sup>3</sup>, Shany Danieli<sup>4,5,6</sup>, Pieter van Dokkum<sup>4</sup>, Jenny Greene<sup>7</sup>, Felipe Ardila<sup>1</sup>, Roberto Abraham<sup>8</sup>, Erin Kado-Fong<sup>7</sup>, Deborah Lokhorst<sup>8</sup>, Robert Lupton<sup>7</sup>, Paul Price<sup>7</sup>, Lamiya Mowla<sup>4</sup>

<sup>1</sup> Department of Astronomy and Astrophysics, University of California Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA

<sup>2</sup> Department of Astronomy, Peking University, 5 Yiheyuan Road, Haidian District, Beijing 100871, China

<sup>3</sup> Department of Physics and Astronomy, Siena College, 515 Loudon Road, Loudonville, NY 12211, USA

<sup>4</sup> Department of Astronomy, Yale University, New Haven, CT 06511, USA

<sup>5</sup> Department of Physics, Yale University, New Haven, CT 06520, USA

<sup>6</sup> Yale Center for Astronomy and Astrophysics, Yale University, New Haven, CT 06511, USA

<sup>7</sup> Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA

<sup>8</sup> Department of Astronomy and Astrophysics, University of Toronto, Toronto ON, M5S 3H4, Canada

Accepted XXX. Received YYY; in original form ZZZ

## ABSTRACT

In this paper, we compare measurements of the outer light profiles (stellar haloes) of massive galaxies using four independent data sets: the Hyper Suprime-Cam survey, the Dark Energy Camera Legacy Survey (DECaLS), the Sloan Digital Sky Survey (SDSS), and the Dragonfly Telephoto Array survey (Dragonfly). We discuss the impact of methodology, sky subtraction, and masking of satellite galaxies in the recovery of light profiles at  $R = 100$  kpc and beyond. At  $z < 0.1$ , Dragonfly has the best control of systematics reaching surface brightness levels of  $\mu_r \sim 30$  mag/arcsec<sup>2</sup>. At  $z \sim 0.35$ , HSC can reliably recover light profiles to  $\mu_r \sim 28.5$  mag/arcsec<sup>2</sup> reaching  $R = 100\text{--}150$  kpc. The median profile of galaxy ensembles in HSC reaches  $R > 200$  kpc without bias. A customized pipeline is required to measure stellar haloes with DECaLS. DECaLS light profiles show good agreement with HSC but are noisier at large radii. At  $z \sim 0.35$ , DECaLS and HSC measurements of the stellar mass contained within 100 kpc agree within 0.05 dex. The tests and results from the paper represent an important step forward for accurate measurements of the outer light profiles of massive galaxies.

**Key words:** galaxies: photometry – galaxies: elliptical and lenticular, cD – galaxies: structure – galaxies: formation

## 1 INTRODUCTION

There is significant interest in understanding exactly how massive galaxies grow with time. The inner regions of massive galaxies have been well studied at low redshifts (Kormendy et al. 2009; Lauer et al. 2014; Greene et al. 2015, 2019; Ene et al. 2019). However, accurate measurements of the outer extended envelopes of massive galaxies are not as straightforward. This is because the diffuse stellar haloes of massive galaxies typically have low surface brightness (LSB) and are difficult to measure accurately.

The low surface brightness nature of the extended outskirts of massive galaxies makes them extremely sensitive to background subtraction effects. For example, even a small over-subtraction of the background can lead to a sudden truncation of the outer surface brightness profiles and an underestimation of the total luminosity (e.g. Blanton et al. 2011; Fischer et al. 2017; Bernardi et al. 2017). Moreover, the outer-wing of the point spread function (PSF, e.g. Tal & van Dokkum 2011; D’Souza et al. 2014) and the diffuse emission of the dust in the Milky Way (known as the Galactic cirrus; e.g. Duc et al. 2015; Román et al. 2019) adds complexity to the analysis of galaxy outskirts.

\* Email: jiaxuan\_li@pku.edu.cn

Furthermore, choices regarding photometric methods

can also greatly influence stellar halo studies. When images are not deep enough, stacking techniques are commonly employed (e.g. Zibetti et al. 2005; Tal & van Dokkum 2011; D’Souza et al. 2014). Stacking enables statistically significant detection of light in the extreme outskirts of massive galaxies ensembles. However, the results are not free of systematics. When images are deep enough, stellar haloes can be studied on an individual galaxy basis (e.g., Huang et al. 2018c). Here choices involve either using 2-D model fitting or 1-D surface brightness profile analysis. In principle, the 2-D fitting approach has several key advantages (e.g. Huang et al. 2013). Model fitting takes the PSF convolution into account and it is straightforward to integrate the results to get total luminosity. The 2-D approach can also account for isophotal twists. Yet, in reality, it also faces many challenges. For example, simple 2-D models like the de Vaucouleurs or single-Sérsic model cannot reliably describe the light distribution in massive galaxies. 2-D model fitting can therefore be subject to model fitting bias. For these reasons, 1-D surface brightness profiles based on isophotes remains a popular and robust method for exploring the properties of stellar halos (Huang et al. 2018c; van Dokkum et al. 2014; Gilhuly et al. 2019; Iodice et al. 2017, 2019).

Improved measurements of the outer regions of massive galaxies are important for many reasons. First, they are needed for correct estimations of total stellar mass (e.g. Bernardi et al. 2013). Second, at large radii, the long dynamical time means that fossil evidence of the assembly history may be preserved (e.g. tidal features from recent mergers, see Tal et al. 2009; Kado-Fong et al. 2018; Smercina et al. 2019). Third, stellar halos show clear dependence on both stellar mass and halo mass and can therefore provide insight about the galaxy-halo connection (e.g. Krick & Bernstein 2007; Huang et al. 2018a; Montes & Trujillo 2019). Finally, stellar halos are expected to show strong redshift evolution which can be used as a strong constraint for galaxy evolution model (e.g. van Dokkum & Conroy 2012; Buitrago et al. 2017).

The goal of this research is to determine how far in radius current surveys and techniques can reliably measure the outer regions of massive galaxies. We focus on four independent data sets, each of which has been used in various analysis of massive galaxies and to study LSB features. The four data sets are: the Hyper Suprime-Cam Subaru Strategic Program Survey (HSC, Aihara et al. 2018), which offers deep images with good seeing but limited sky coverage; the Dark Energy Camera Legacy Survey (DECaLS, Dey et al. 2019), with images that are very wide, and deeper than SDSS but shallower than HSC; the Dragonfly Wide Field Survey (Dragonfly, Abraham & van Dokkum 2014; Danieli et al. 2019), which is optimised for LSB detection but has poor resolution and small sky coverage currently; and the Sloan Digital Sky Survey (SDSS) (Abazajian et al. 2009), which has very large sky coverage and has been used in many different studies. We compare the light profiles extracted from these four surveys as a means of empirically determining the robustness of outer light profile measurements. We study a low redshift ( $z < 0.05$ ) sample using SDSS, Dragonfly, HSC, and DECaLS, and an intermediate redshift sample at  $0.19 < z < 0.50$  using HSC and DECaLS.

The layout of this paper is as follows. The data and sample selection are described in Section 2. In Sections 3, 4

and 5, we describe our image reduction and surface brightness profile measurement methods. The main results of this work are presented in Section 6. Implications and future directions are discussed in Section 7. Section 8 presents our summary and conclusions. We adopt a  $\Lambda$ CDM cosmology with  $\Omega_m = 0.3$ ,  $\Omega_\Lambda = 0.7$  and  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ . We use the AB system (Oke & Gunn 1983) for magnitudes.

## 2 DATA AND SAMPLE SELECTION

We consider two samples at different redshifts, namely, a low redshift galaxy sample and an intermediate redshift galaxy sample. Dragonfly and SDSS can only study stellar halos at low redshifts. For this reason, we define a low redshift sample that will be used to compare the surface brightness profiles between the four different data sets under consideration. However, given the current area of the Dragonfly survey, the number of massive galaxies is small (5 galaxies). We also study a higher redshift sample that contains a much larger number of galaxies (2171 galaxies). This intermediate redshift sample is used to study the profiles of galaxy ensembles. In this section, we describe our data and sample selection.

### 2.1 Imaging Data

The Hyper Suprime-Cam Subaru Strategic Program Survey (Aihara et al. 2018; hereafter ‘HSC survey’) is based on the Hyper Suprime-Cam (Miyazaki et al. 2012, 2018), which is a prime-focus camera mounted on the 8.2-m Subaru telescope. The WIDE layer of the HSC survey aims to provide deep images in five broad-bands (*grizy*) over  $1000 \text{ deg}^2$ . The depth of the WIDE layer is  $g = 26.6 \text{ mag}$ ,  $r = 26.2 \text{ mag}$  and  $i = 26.2 \text{ mag}$  ( $5\sigma$  point source). The responses of the HSC *gri*-band filters are almost the same as those of SDSS (Kawanomoto et al. 2018). HSC has a pixel scale of  $0.168''/\text{pixel}$  and a typical seeing of  $0.8''$  in the *r*-band, giving it the best resolution among large sky surveys. The HSC footprint overlaps with many spectroscopic surveys, such as SDSS Baryon Oscillation Spectroscopic Survey (BOSS, Dawson et al. 2013), Galaxy and Mass Assembly (GAMA, Driver et al. 2011), and the Visible Multi-Object Spectrograph (VIMOS) Very Large Telescope (VLT) Deep Survey (VVDS, Le Fèvre et al. 2005).

HSC data are processed using `hscPipe` (Bosch et al. 2018), which is a modified version of the Large Synoptic Survey Telescope (LSST) pipeline (Jurić et al. 2017). `hscPipe` performs forced multiband photometry of objects using `cModel`, which fits each object with a combination of de Vaucouleurs and exponential components considering PSF effects. At fainter magnitudes, `cModel` photometry is accurate down to  $i \sim 25.0 \text{ mag}$  (Huang et al. 2018b) but it often underestimates fluxes in the outskirts of extended massive early-type galaxies (ETGs) (Huang et al. 2018a). In previous data releases, sky background over-subtraction severely affected the analysis of galaxy outskirts. Data release S18A, also known as HSC Public Data Release 2 (Aihara et al. 2019), fixed this issue with `hscPipe` 6.7 and provided sky objects for calibrating the sky background (see Section 3.2 and Appendix A for details). In this paper, we use the WIDE layer data of the HSC survey internal data release S18A.

The Dark Energy Camera Legacy Survey is a public survey based on the data from Dark Energy Camera (DECam, DePoy et al. 2008), which is mounted on the Blanco 4-m telescope. It covers a  $9000 \text{ deg}^2$  region within  $-20^\circ \leq \text{DEC} \leq 32^\circ$  in  $grz$  bands with a  $5\sigma$  point source detection limit of  $g = 24.0$ ,  $r = 23.4$  and  $z = 22.5$ . DECaLS is also an important part of Legacy Surveys (Dey et al. 2019). Since the current data release DR8<sup>1</sup> still suffers from sky over-subtraction, we use a custom background subtraction at the CCD level and generate co-added images (coadds) manually. We describe this procedure in Section 4.

The Dragonfly Telephoto Array ('Dragonfly' for short, Abraham & van Dokkum 2014; Danieli et al. 2019) is an imaging system consisting of 48 Canon 400 mm f/2.8 IS II telephoto lenses, equivalent to a 1 m aperture,  $f/0.4$  refractor. The system is designed and optimized for detecting LSB features. Each lens is mounted to a CCD camera, providing a  $2.6^\circ \times 1.9^\circ$  instantaneous field of view with  $2.8''$  native pixel scale, and a full-width-half-maximum (FWHM) seeing  $\approx 5''$ . Half of the lenses are equipped with SDSS- $g$  filters and the other half with SDSS- $r$  filters. The all-refractive design and 'sub-wavelength' anti-reflection coatings of Canon telephoto lenses ensure that Dragonfly's PSF is very well-controlled with low power distributed in far wings. Dragonfly is capable of detecting surface brightness down to  $\mu = 32 \text{ mag/arcsec}^2$  in 1-D surface brightness profiles (van Dokkum et al. 2014). In this paper we use data from the Dragonfly Wide Field Survey (DWFS; Danieli et al. 2019). The footprint of DWFS covers  $330 \text{ deg}^2$  and overlaps with Stripe 82 and the GAMA fields to maximize its scientific potential with other multi-wavelength data sets. The methods used for Dragonfly data reduction are further described in Section 5.

We also use images from the Sloan Digital Sky Survey (Gunn et al. 2006). SDSS provides  $ugriz$  images for  $35,000 \text{ deg}^2$  with a  $5\sigma$  point source detection limit of  $g = 23.13$ ,  $r = 22.70$ ,  $i = 22.20$ , and  $z = 20.71$ . We generate SDSS coadds using `SWarp` (Bertin et al. 2002) based on SDSS DR7 images (Abazajian et al. 2009) and using the SDSS DR12 image mosaic tool<sup>2</sup>.

Response curves for the same bands can differ among surveys, not only because of filter differences but also due to the quantum efficiency (QE) curves of detectors, which vary a lot from camera to camera. The difference in filters can incur small shifts in surface brightness profiles and deviations in the mass measurement of galaxies. We describe our filter correction method in Section 3.4.

## 2.2 Low Redshift Galaxy Sample

The DWFS footprint covers Stripe82 and some of the GAMA fields. The HSC S18A WIDE layer footprint overlaps with Dragonfly in the G09, G12, and G15 GAMA fields. We cross-match the GAMA catalogue (Taylor et al. 2011) with both HSC and Dragonfly data, then select ETGs with  $z \leq 0.04$  and  $\log(M_*^{\text{GAMA}}/M_\odot) > 10.5$  as an initial sample. For HSC, a major concern is that very nearby massive galaxies are typically saturated in  $r$ - or  $i$ -band. Since ETGs

Name	R.A.	Dec.	$z$	$\log M_*/M_\odot$
low- $z$ -A	175.02673	-0.83771	0.020	10.53
low- $z$ -B	177.40144	-1.45549	0.019	10.64
low- $z$ -C	177.37548	-1.08661	0.019	10.93
low- $z$ -D	140.00904	1.03827	0.017	11.04
low- $z$ -E	222.48908	0.55829	0.040	11.05

**Table 1.** Name, position, redshift, and stellar mass of galaxies in the low- $z$  sample. The stellar mass values are retrieved from the GAMA catalogue (Taylor et al. 2011).

are typically brighter in  $i$ -band and thus are prone to be saturated, we focus only on  $r$ -band data in this paper (in addition, DECaLS does not have  $i$ -band). We apply the bright star mask and also exclude  $r$ -band saturated galaxies in the initial sample. We select the top five brightest galaxies as our low redshift galaxy sample (the 'low- $z$  sample'). Their mass is in the range of  $10^{10.5} M_\odot$  to  $10^{11.1} M_\odot$ ; hence they are less massive compared with the intermediate redshift sample (see Section 2.3). Basic information about the sample is listed in Table 1. Figure 1 shows tri-colour images of our low- $z$  sample. The tri-colour images for HSC, DECaLS, and SDSS are generated using three bands ( $gri$  for HSC and SDSS,  $grz$  for DECaLS) according to the colour scheme described in Lupton et al. (2004). Since Dragonfly has only  $g$ - and  $r$ -band data, we make synthetic colour images using  $g$ -band intensity  $I(g_{\text{DF}})$  as the blue channel,  $r$ -band intensity  $I(r_{\text{DF}})$  as the red channel, and  $0.5 I(g_{\text{DF}}) + 0.5 I(r_{\text{DF}})$  as the green channel.

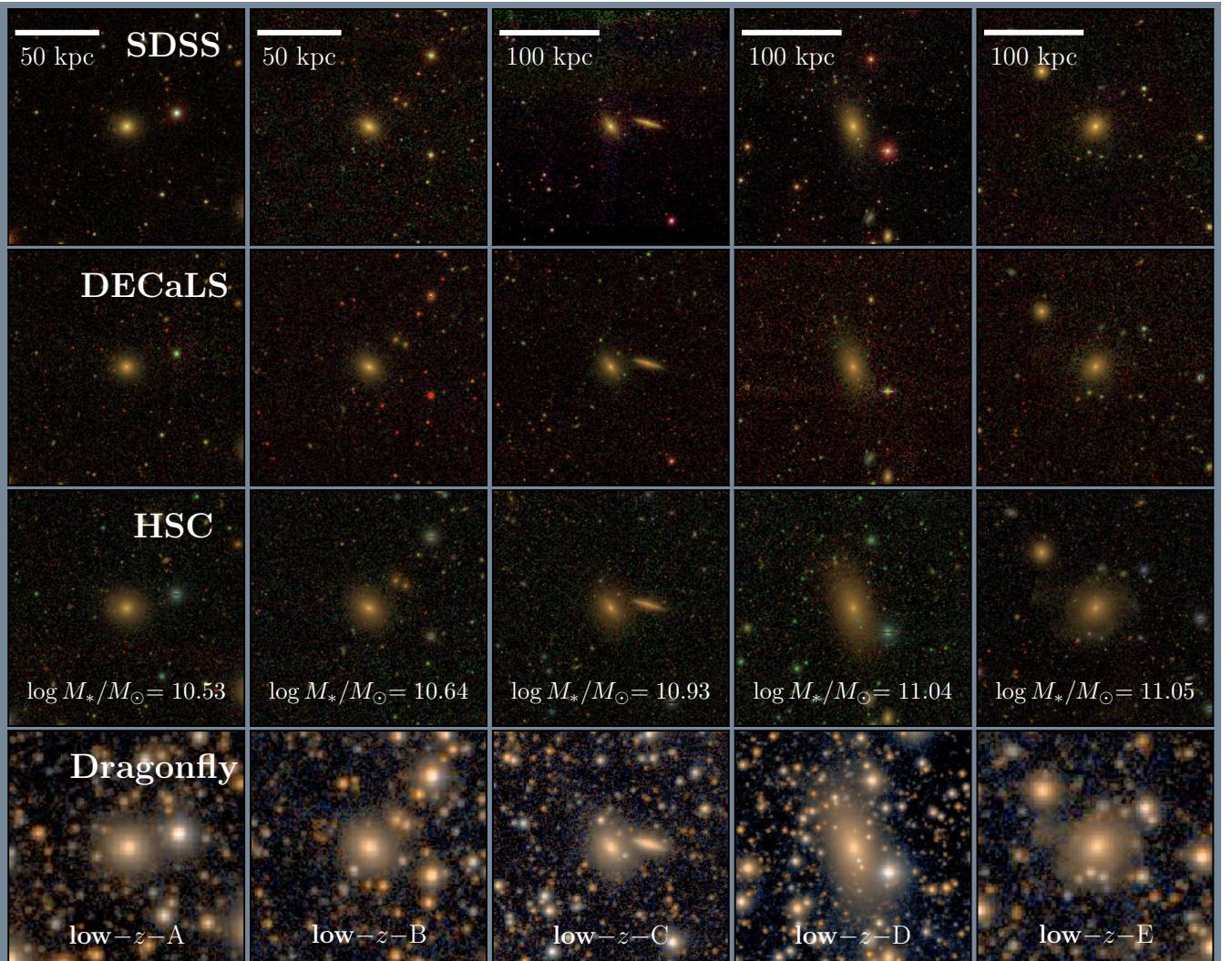
## 2.3 Intermediate Redshift Galaxy Sample

We construct an intermediate redshift sample that will be used to compare DECaLS and HSC. This sample is very similar to the sample used in Huang et al. (2018a). We select galaxies in the HSC footprint with spectroscopic redshifts  $0.19 < z < 0.50$  and stellar mass  $\log M_*^{\text{max}} > 11.6$ , where  $M_*^{\text{max}}$  is the maximum 1-D stellar mass derived from the surface brightness profile using the HSC data release S15B (see Huang et al. 2018c for details). For these ETGs, the maximum 1-D stellar mass is typically larger than the cModel mass by  $0.2 - 0.3$  dex. This redshift range ensures that sky background subtraction and cosmological dimming are not major issues for studying galaxy outskirts and that the redshift evolution of galaxies across the bin can be neglected. For each galaxy, we detect objects around it using `Source Extractor` (Bertin & Arnouts 1996), and we remove galaxies with very close satellites or strong contamination by bright stars and artefacts. After filtering galaxies with the S18A bright star mask and full-colour full-depth mask (Ai-hara et al. 2019), we obtain 2171 galaxies in the intermediate redshift sample ('intermediate- $z$  sample'). This sample has reliable spectroscopic redshift measurements (from BOSS, SDSS, or GAMA). The distributions of redshift and stellar mass of the intermediate- $z$  sample are shown in Figure 2. Five randomly selected galaxies from the intermediate- $z$  sample are shown in Figure 3.

To study redshift-dependent effects, we also divide the intermediate- $z$  sample into two redshift bins:  $0.19 < z < 0.35$  and  $0.35 < z < 0.50$ . The cut at  $z = 0.35$  is arbitrary. The first bin contains 870 galaxies, the second bin contains 1301 galaxies. We neglect redshift-dependent effects within

<sup>1</sup> <http://legacysurvey.org/dr8/description/>

<sup>2</sup> <https://dr12.sdss.org/mosaics>



**Figure 1.** Tri-colour images of nearby ETGs in the low- $z$  sample. Images from top to bottom correspond to SDSS, DECaLS, HSC, and Dragonfly. The two galaxies on the left have an image size of 150 kpc  $\times$  150 kpc. The three images on the right have an image size of 250 kpc  $\times$  250 kpc. Further information on these galaxies can be found in Table 1. The HSC point source detection is  $\sim 2.6$  mag deeper than DECaLS and  $\sim 3.5$  mag deeper than SDSS. Of the four surveys, Dragonfly has the worst resolution and seeing but an excellent ability to map the LSB features of low- $z$  galaxies.

each redshift bin. We compare the two redshift bins as a test for redshift dependant effects. Figure 2 shows that the two redshift bins have similar stellar mass distributions. Unless noted otherwise, the stellar mass of an intermediate- $z$  galaxy corresponds to the maximum 1-D stellar mass for the intermediate- $z$  sample hereinafter.

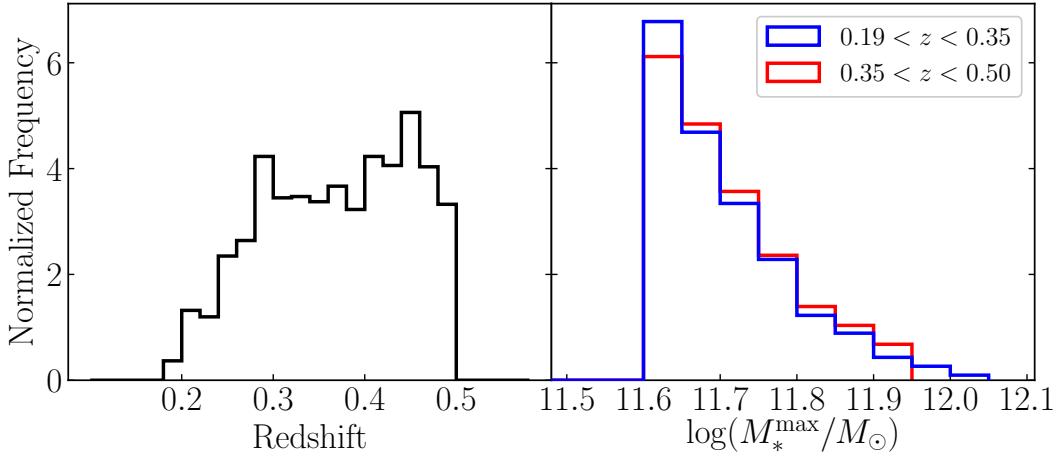
### 3 HSC METHODOLOGY

In this section, we describe our methodology for extracting 1-D surface brightness profiles using HSC data. We also describe 2-D modelling and sky background correction. Our methods are validated using a variety of tests with mock galaxies in Section 3.2.

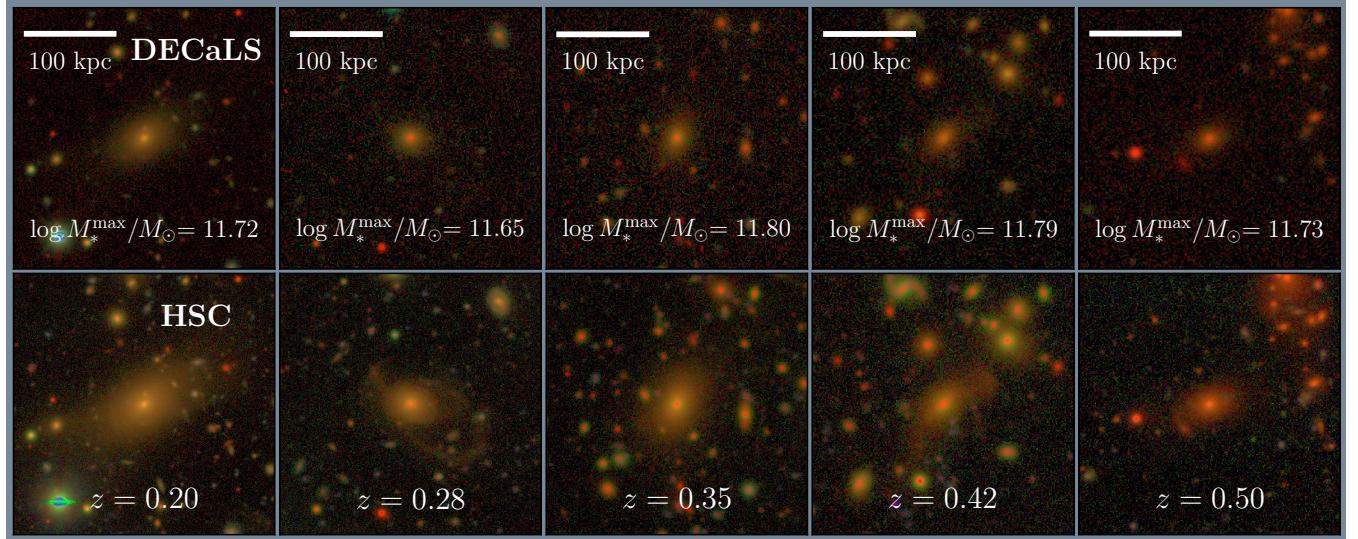
#### 3.1 Measurements of 1-D surface brightness profiles

We use the elliptical isophote fitting program `Ellipse` (Jedrzejewski 1987) in the Image Reduction and Analysis Facility (IRAF) to measure surface brightness profiles. We follow the basic method used in Huang et al. (2018c) with some modifications. We review some of the main steps here and refer the reader to Huang et al. (2018c) for more details.

We download  $r$ -band coadd cutouts with 1000 pixel  $\times$  1000 pixel sizes for intermediate- $z$  galaxies. This corresponds to 320 kpc at  $z = 0.19$  and 770 kpc at  $z = 0.50$  in radius, which is sufficient for the purpose of this research. We do not choose a larger size in consideration of the computing time consumption. The largest cutout image generated for galaxies in the low- $z$  sample has a size of 5000  $\times$  5000 pixels, corresponding to 320 kpc  $\times$  320 kpc for the nearest galaxy in the low- $z$  sample. The cutout images are large enough for us to measure the surface brightness profile out to a very far



**Figure 2.** Redshift (left) and stellar mass distributions (right) of the intermediate- $z$  sample. The blue and red lines in the right panel correspond to the two redshift bins. The stellar mass distributions of them are similar.



**Figure 3.** Tri-colour images for five randomly selected galaxies in the intermediate- $z$  sample from DECaLS and HSC. Images are generated following the same format as in Figure 1. Images are 300 kpc  $\times$  300 kpc, with maximum 1-D stellar mass labelled. Deep images reveal extended outskirts and evidence of recent mergers.

radius and estimate the local sky background if needed. HSC S18A also provides bitmasks indicating bad pixels, cosmic rays, edges of CCDs and pixels with source detection (Aihara et al. 2019).

We detect bright objects and mask them out using `sep` (Barbary 2016), a Python version of `Source Extractor` (Bertin & Arnouts 1996). Specifically, we extract objects above a  $4.5\sigma$  threshold and convolve the segmentation map with a 15-pixel radius Gaussian kernel to increase the mask size. We aggressively mask out very bright objects ( $r_{\text{mag}} < 18.0$ ) by making an elliptical mask with the shape and size from `SEExtractor` but enlarged by a factor of 10. We combine this mask with the HSC bright star bitmask<sup>3</sup> to shield the light from bright stars. Considering that `SEExtractor`

`tractor` cannot recognize irregular extended objects (such as tidal tails) in the image, we also incorporate part of the detection mask of the HSC image with a `SEExtractor` mask. In this way, almost all bright objects, bleeding tails, and tidal disruption debris are masked out except our target object. To attest the robustness of our masking method, we adjusted several parameters (such as the detection threshold and the Gaussian kernel radius) several times, with negligible changes in profiles. Next, we run `Ellipse` on the cutout image with a mask generated as above. We first fit isophotes with a fixed center but allow shapes (i.e. ellipticity and position angle) to vary freely. We estimate the mean ellipticity and position angle within  $20 \text{ kpc} \leq R \leq 50 \text{ kpc}$  for the intermediate- $z$  sample, where  $R$  is the radial distance along the semi-major axis of elliptical isophotes. For galaxies in low- $z$  sample (stellar mass below  $10^{11.10} M_\odot$ ), we set this interval to  $20 \text{ kpc} \leq R \leq 30 \text{ kpc}$ . Then we measure the sur-

<sup>3</sup> <https://hsc-release.mtk.nao.ac.jp/doc/index.php/bright-star-masks-2/>

face brightness profile again along the semi-major axis with a fixed center and shape. The mean ellipticity and position angle from the first run are taken as the fixed isophote shape in the second run.

We use 0.1 dex in logarithm as the intensity step along the semi-major axis, and use the median pixel value over the elliptical annulus after rejecting outlying pixels by applying sigma-clipping 3 times, with lower clipping at  $3\sigma$  and upper clipping at  $2.5\sigma$ . These choices are made to ensure the final surface brightness profile is less affected by any nearby compact objects which are not identified and masked out. We present all the surface brightness profiles with  $R^{1/4}$  scale to emphasize the outskirts.

For this paper, we manually apply some corrections to the sky background subtraction. Background corrections are applied directly onto the profile, instead of at the image level. We do not find any notable difference between these two approaches. `Ellipse` also gives the measurement error of intensity along the semi-major axis, but the errors estimated by `Ellipse` are usually underestimated. In this work, the median profile is calculated at the intensity level first, then converted to surface brightness in the unit of mag/arcsec<sup>2</sup>. We do not exclude any negative intensity values in any single profile before calculating the median profile. We discuss the caveat of using magnitude as the unit of surface brightness in Section 6.3.

### 3.2 Tests on the Robustness of HSC Profiles to Background Subtraction

In measuring the surface brightness profile of massive galaxies, sky subtraction plays a critical role (e.g. Huang et al. 2013, 2018c). In previous HSC data releases, `hscPipe` measured the sky by fitting a map of ‘super-pixels’ (256 pixels  $\times$  256 pixels) with a sixth-order 2-D Chebyshev polynomial in each exposure. In this way, the sky model is discontinuous at the boundaries of CCDs and could easily be biased by nearby bright objects. Therefore, the old `hscPipe` tends to over-subtract the sky background around bright objects, leaving ‘dark rings’ around them. This issue impedes our ability to reach deeper in the LSB regime; hence Huang et al. (2018c) only consider the surface brightness profile within 100 kpc as reliable. In data release S18A, `hscPipe` implements a new method that measures the sky background globally on the entire focal plane and correct response difference of the instruments (Aihara et al. 2019). S18A data have much better sky subtraction and preserve LSB features very well (see Figure 5 in Aihara et al. 2019). In S18A, `hscPipe` also detects void regions in which no object is detected and no surrounding objects are detected in any of the five bands. Then the fluxes within these voids are measured on the coadd level with different aperture sizes. These void regions are called ‘Sky Objects’ (`SkyObjs`). They are good proxies for the local sky background, but they can be biased low compared to the regular sky background.

We conduct a series of mock tests to check the quality of sky subtraction in HSC S18A and verify the rationality of our methods on extracting surface brightness profiles and correcting sky background. Since there are typically no objects within `SkyObjs`, it is convenient to put mock galaxies in the positions of `SkyObjs` to mimic real observation re-

Name	$z$	$r_{\text{mag}}$	$n$	$R_e$ (arcsec)	$e$	P.A.
Model-0	0.240	16.81	4.514	5.86	0.166	-57.17
Model-1	0.309	17.42	4.613	3.95	0.178	-16.18
Model-2	0.383	18.05	5.211	2.67	0.226	40.90

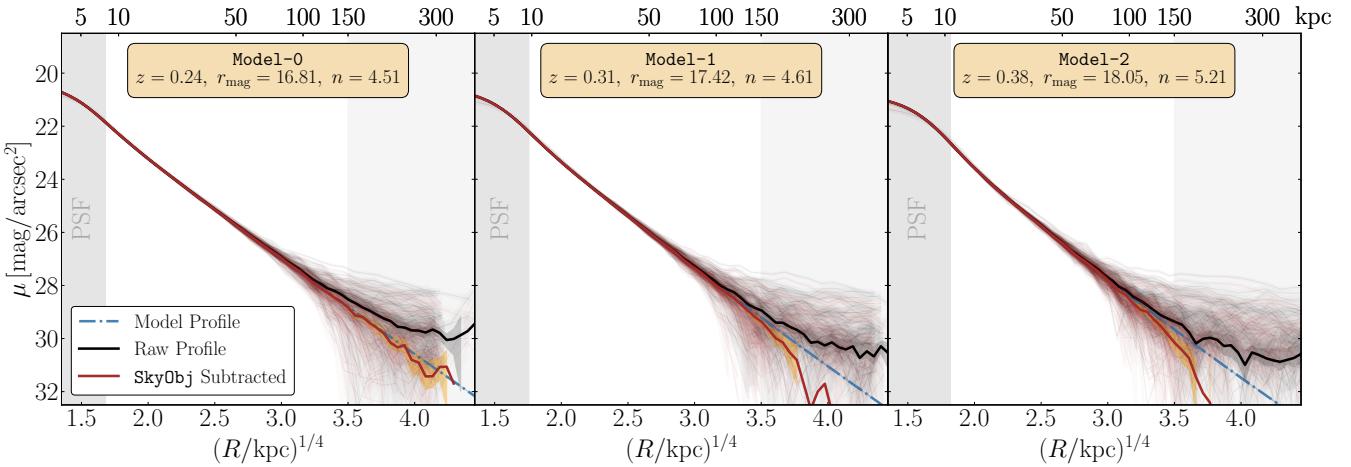
**Table 2.** Three different galaxy models used in mock tests, which are from single-Sérsic modelling in GAMA survey (Kelvin et al. 2012). They span a large range of redshift,  $r$ -band magnitude, Sérsic index and shape.

sults. We generate mock images and test sky subtraction as follows.

First, we make  $r$ -band cutouts with centers located in the positions of randomly selected `SkyObjs`. Then we generate fake galaxies with a given model using `GalSim` (Rowe et al. 2015) and add the fake galaxies onto the cutouts. The PSF model of a given position is retrieved from `hscPipe` as an input of `GalSim`<sup>4</sup>. Studies have shown that ETGs can be described by single-Sérsic profiles (e.g. Kormendy et al. 2009); but recent literature has found significant deviations from single-Sérsic model exist for some massive ETGs, especially in galaxy outskirts, and has suggested multiple-component models (e.g. Huang et al. 2013; Oh et al. 2017). However, to test the sky subtraction by generating fake galaxies, we simply make use of single-Sérsic fitting results of the GAMA survey (Kelvin et al. 2012) as a reference. We choose three different models with relatively large Sérsic indexes ( $n > 4$ ) to mimic the extended outskirts of real massive ETGs. Their information is listed in Table 2. In total, we generate mock images for 60 randomly selected positions with three different models located at different redshifts.

Then we measure the 1-D surface brightness profiles for the mock galaxy in the mock image using the same method described in Section 3.1. Their profiles are shown in Figure 4 as dim black lines. The thick black line is the median profile of all mock galaxies, and the blue dash-dotted lines are fiducial model profiles. The flattened outfit of the median profile indicates that the sky of HSC S18A is very likely under-subtracted, instead of over-subtracted as seen in previous data releases (Huang et al. 2018c). The under-subtraction might be caused by the fact that `hscPipe` subtracts the sky background on a relatively large scale in S18A (Aihara et al. 2019). This kind of sky residual is also filter-dependent. Confronted with the under-subtracted sky, we subtract the image by the mean value of nearby `SkyObjs` and eliminate the under-subtraction issue to a large extent. The mean correction for the sky background is typically  $0.002 - 0.008 \mu\text{Jy}/\text{arcsec}^2$  ( $0.001 - 0.004 \text{ counts/pixel}$ ). Based on our tests, the median profile recovers the model best when subtracting the mean value of `SkyObjs` matched within  $[1', 4']$ . The thick red line in Figure 4 is the median profile after `SkyObjs` correction using  $8.4''$  aperture `SkyObjs`, which recovers the model profile to a very large extent ( $> 150$  kpc). `SkyObjs` with  $8.4''$  aperture size works best for bright low- $z$  galaxies, whereas  $5.7''$  aperture size is better for faint high- $z$  galaxies. However, Section 6.3 and Appendix A support the validity of choosing `SkyObjs` with  $8.4''$  aperture using the real data of the intermediate- $z$  sample, ruling out  $5.7''$  aperture `Sky-`

<sup>4</sup> <https://github.com/GalSim-developers/GalSim>



**Figure 4.** Tests on the accuracy of HSC sky estimation using `SkyObjs` with synthetic galaxies. We test three different galaxy models (see Table 2). The blue dash-dotted line indicates the fiducial model profile. Thin grey lines are individual realizations of synthetic galaxies. The thick line is the median profile of 60 mock galaxies. The value of the sky is estimated using the mean value of `SkyObjs` with  $8.4''$  aperture which are selected within an annulus between  $1'$  and  $4'$ . The sky subtracted profile is shown by the red line. These tests demonstrate two things. First, small sky residual exists in HSC S18A images. Second, the use of `SkyObjs` allows us to reliably estimate the sky background and to extract profiles to  $\sim 150$  kpc and  $\mu = 29.0 - 29.5 \text{ mag/arcsec}^2$ .

`Objs`. We continue to use `SkyObjs` with  $8.4''$  throughout this paper. Appendix A includes more details on mock tests and `SkyObjs`.

We also find that the low- $z$  bright galaxies still suffer from over-subtraction in S18A. The raw surface brightness profiles all truncate dramatically at the outskirts of nearby galaxies. Furthermore, sky objects around these galaxies are rare compared with intermediate- $z$  galaxies and are typically far away from the over-subtracted ring. `SkyObjs` are no longer good proxies of the over-subtracted sky under this circumstance. Thus we measure the sky background manually and correct over-subtraction at 1-D profile level using the following method. We first aggressively mask out every object in the image by combining high-threshold ( $\sim 5\sigma$ ) and low-threshold ( $\sim 2\sigma$ ) source extractions. Then we bin the masked image with a  $6'' \times 6''$  box (which is larger than the local fluctuation of the image), and we find that the intensity distribution of these ‘super-pixels’ can be well-described by a Gaussian profile. Thus, we adopt the median of the distribution as the local sky background value. We also estimate the standard error of the median (referred to as the *uncertainty of background*) to be  $\sigma/(f\sqrt{N})$ , where  $\sigma$  is the standard deviation of super-pixel values,  $N$  is the number of super-pixels and  $f$  is the binning factor. One can also calculate this *uncertainty of background* by bootstrap resampling, which gives almost the same results as  $\sigma/(f\sqrt{N})$  in our cases. The median and its standard error help us construct the upper-limit and lower-limit profiles in Section 6.1.

### 3.3 Modelling Neighbouring Stars and Galaxies

Satellite galaxies, stars and foreground/background galaxies contaminate the surface brightness profile of target galaxies. In Section 3.1, we describe how we use a binary mask to remove these contaminants. We have performed tests using synthetic galaxies to show that we can recover profiles accurately using this binary mask. This method recovers surface

brightness profiles well in mock tests with the help of `SkyObjs`. Another method is to model surrounding objects using 2-D models and to subtract these models from the image, leaving the target galaxy in the center. Here we test this method on mock galaxies.

The `Tractor`<sup>5</sup> is a probability-based photometry and model fitting tool (Lang et al. 2016). We use the `Tractor` to build 2-D models of neighbouring objects and subtract them from the image. We first use an HSC bright star mask to mask out bright or saturated stars that cannot be modeled well by simple models. Then all the objects in the image are extracted by `sep`, along with their shapes and fluxes. We need to assign prior model types to objects to run the `Tractor`; available models contain point source, de Vaucouleurs profile, exponential profile and composite profile (de Vaucouleurs + exponential). DECaLS provides the `Tractor` fitting results (e.g. model type, shape, flux) for each object within its footprint. We match objects in the image with DECaLS catalogues and use only the model types from DECaLS. The shapes and fluxes of models are still based on our measurements using `sep`. For objects that are not matched in DECaLS catalogues, we assign their type to be ‘de Vaucouleurs’. The central target object is assigned to be ‘composite’. Fitting hundreds of objects together at the same time is very time consuming; for this reason, we first fit the 40 brightest objects together, then add more objects to the optimized model and fit again. This process is repeated until the model of every object is added and optimized. Figure 5 demonstrates the result of the `Tractor`. The satellites, stars and other galaxies around the target galaxy have been modelled and removed as shown in the right panel. Nevertheless, many galaxies cannot be simply modelled by ‘de Vaucouleurs’ or ‘exponential’, as residuals show up as bright dots and negative rings. We mask out the remaining residu-

<sup>5</sup> <https://github.com/dstndstn/tractor>

als before extracting surface brightness profiles of the central target.

Although **the Tractor** behaves well for individual galaxies, it is not clear that it will improve the measurement of surface brightness profiles. Mock tests are needed to verify this effect. Following the same procedure as in Section 3.2, we generate 70 mock images with the first model ( $z = 0.24$ ,  $r_{\text{mag}} = 16.81$ ) in  $r$ -band. Then we subtract neighbours using **the Tractor** and extract surface brightness profiles. Figure 6 shows that **the Tractor** performs as well as a binary mask with **SkyObj** correction for intermediate- $z$  galaxies. The 2-D modelling method also subtracts the residual sky background from the image. Given that the binary mask method could accurately recover the profile and considering the heavy time consumption of running **the Tractor**, we extract surface brightness profiles using only binary masks in the following results.

In a companion paper (F. Ardila et al., in preparation, Paper II), we have also tested satellite removals using mock images from Illustris (Vogelsberger et al. 2014a,b) and IllustrisTNG (Springel et al. 2018; Pillepich et al. 2018) simulations. The results of those tests show that residual light from satellites has a negligible impact on 1-D profiles.

### 3.4 Filter Corrections

To compare the photometry of HSC, DECaLS, SDSS and Dragonfly, we need to account for filter differences among them. Here we calculate the differences using synthetic stellar photometry. Given the response curves of each band and a spectrum, we calculate the AB magnitude in each band of this star. We use spectra from the Gunn-Stryker stellar spectra atlas (Gunn & Stryker 1983) and calculate AB magnitudes using **sedpy**. By assuming that galaxies and stars with the same colour have *nearly* the same spectral energy distribution (SED), we obtain the filter correction for each galaxy.

For the intermediate- $z$  sample, we apply corrections to DECaLS magnitudes to match HSC values. We fit a third-order polynomial between the source colour  $g_{\text{HSC}} - i_{\text{HSC}}$  and magnitude difference  $\Delta_r = r_{\text{HSC}} - r_{\text{DECaLS}}$ . HSC magnitudes are given by **cModel**. Then we use the HSC colour ( $g_{\text{HSC}} - i_{\text{HSC}}$ ) to calculate the magnitude difference  $\Delta_r$  according to the fitted polynomial and add it to the DECaLS profile. For low- $z$  sample, we calibrate other surveys to SDSS in the same way, but using  $g_{\text{SDSS}} - r_{\text{SDSS}}$  colour instead. The filter correction term is typically around 0.15 mag for the intermediate- $z$  sample and is less than 0.10 mag for the low- $z$  sample. We have applied filter corrections to all the profiles in this paper.

However, this method has limitations. It assumes that the SED of a galaxy is the same as the SED of a star with the same colour, which is not necessarily true. A better analysis using SED fitting is needed to derive the filter correction term for each galaxy. Furthermore, we ignore the colour gradient in a galaxy when applying the filter correction term to the surface brightness profile. These corrections are, however, beyond the scope of this paper.

## 4 DECaLS METHODOLOGY

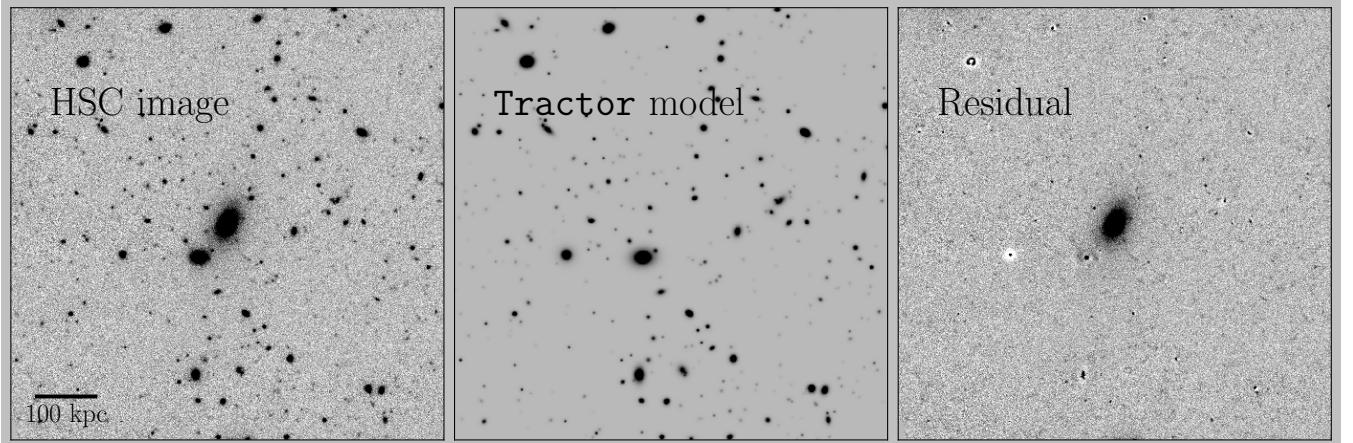
The DECaLS public data release (here, DR8) includes tools to generate coadds from all the CCD imaging available at a given position in the survey footprint (separately, of course, for each  $grz$  bandpass). The algorithm used to model and subtract the sky-background in DR8, however, systematically over-subtracts the signal from the low surface brightness, outer envelopes of the galaxies in our sample (see Dey et al. 2019 and J. Moustakas et al., in preparation). To mitigate this issue, we use a custom algorithm to subtract the small-scale background<sup>6</sup> from each individual CCD before building the DECaLS mosaics and performing our photometric measurements.

First, we determine which CCDs overlap an area of the sky in fixed physical (comoving) coordinates based on the central coordinates and redshift of each galaxy. For the intermediate- $z$  sample we use a  $500 \text{ kpc} \times 500 \text{ kpc}$  area, while for the low- $z$  sample we use  $300 \text{ kpc} \times 300 \text{ kpc}$ . For reference, the typical angular diameter of these cutouts is  $\approx 1.5 \text{ arcmin}$  ( $\approx 370$  pixels) for the intermediate- $z$  sample and  $\approx 20 \text{ arcmin}$  ( $\approx 5000$  pixels) for the low- $z$  sample, while the area of a single DECam CCD is roughly  $9 \times 18 \text{ arcmin}^2$ . Next, for each central galaxy and CCD, we define a circular annulus whose inner and outer radii are a multiple of the cluster radius (i.e., 250 kpc and 150 kpc for the intermediate- $z$  and low- $z$  samples, respectively), and measure the median sky background in that annulus after aggressively and iteratively rejecting pixels contaminated by astrophysical sources. Finally, we subtract that pedestal sky value from the whole CCD.

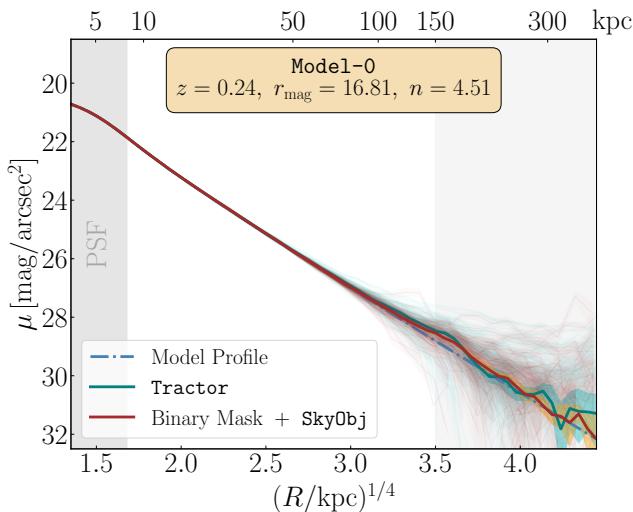
To determine the optimal choice of sky annulus, we select a sample of 100 intermediate- $z$  galaxies with well-determined HSC  $r$ -band surface brightness profiles. We investigate a range of annuli, 0.5-1, 0.5-2, 1-2, 1-3, 1-4, 2-3, 2-4, 2-5, and 3-5 times the cluster radius, and compare the resulting surface brightness profiles to one another and to the fiducial HSC surface brightness profile. Based on these tests, we find that a sky annulus which is 2-4 times the cluster radius (i.e., 500 – 1000 kpc and 300 – 600 kpc from the central galaxy for the intermediate- $z$  and low- $z$  samples, respectively) yields the most robust outer surface brightness profiles and galaxy curves of growth.

We take the following step to measure the surface brightness profile of the central galaxy. First, we project the background-subtracted CCD images onto a common tangent plane and pixel scale ( $0.^{\prime\prime}262$  per pixel) using a Lanczos-3 (sinc) interpolation kernel, and build the image stack (in each bandpass) using inverse variance weighting. Next, we use **the Tractor** to model and subtract from the mosaic all the sources in the field *except* the central galaxy of interest (see Section 3.3, Dey et al. 2019, and J. Moustakas et al., in preparation). In general, the fitting of the surrounding galaxies (including the cluster satellites) and foreground stars in the DECaLS imaging is reliable; nevertheless, to account for model mismatch and other systematic residu-

<sup>6</sup> The large-scale, inter-CCD sky background, including the pupil ghost and other large-scale additive background terms, have already been subtracted from the data at this point in the pipeline, leaving only the small-scale, intra-CCD background (see Dey et al. 2019 for details).



**Figure 5.** Modelling and subtracting neighbouring stars and galaxies with the **Tractor**. The central galaxy is drawn from the intermediate- $z$  sample and has  $z = 0.294$ ,  $r_{\text{cModel}} = 18.17$ ,  $\log M_*^{\max} = 11.67$ . The middle panel shows the **Tractor** models for surrounding galaxies (excluding the central galaxy). The right panel shows the image after subtracting neighbours. Small residuals are masked out before extracting 1-D surface brightness profiles.



**Figure 6.** Median profile of mock galaxies after subtracting neighbouring stars and galaxies using the **Tractor** is shown in green. The red line shows the results of the binary mask with the **SkyObj** correction. The two methods agree with each other very well, but the **Tractor** takes much more time to run.

als, we mask outlier pixels identified in the residual images (excluding the central galaxy). Finally, we fix the geometry (position angle and ellipticity) of each elliptical isophote to the values derived from the HSC imaging (see Section 3.1) and measure the median surface brightness (iteratively clipping  $3\sigma$  outliers) along the semi-major axis using the ellipse-fitting tools in **photutils**.<sup>7</sup>

## 5 DRAGONFLY METHODOLOGY

In this section we briefly describe the sky background modelling and subtraction of the Dragonfly Wide Field Survey (DWFS; Danieli et al. 2019), which is the Dragonfly data set we use in this paper. We refer the interested reader to Danieli et al. 2019 and Jielai Zhang’s PhD thesis<sup>8</sup> for more details.

The sky modelling and subtraction is an important step of Dragonfly data reduction pipeline and is designed to conserve as much low surface brightness emission as possible. A two-stage sky subtraction is applied to each frame before the final stacking. In the first stage, SExtractor is used to create a background map of each science exposure with a background mesh size of  $128 \times 128$  (BACK\_SIZE= 128). Then the background map is fitted by a third-order polynomial and subtracted from each individual exposure. Next, the images are registered to a common grid with pixel scale 2.5''/pixel, using SCAMP (Bertin 2006) and SWarp (Bertin et al. 2002) and scaled to a common flux level. A median coadd image is generated by taking the median of all sky-subtracted frames. The  $g+r$  median coadd thus represents the full detection power of the data set. In the second stage, the sky background of a single frame is modelled in the same way but has an input weight map when running SExtractor. The weight map is made from the corresponding area in the  $g+r$  median coadd image, in which sky pixels have the value of 1 and source pixels (to be masked) are 0. Masking individual science frames according to the deep coadd ensures that all sources are masked down to their very low surface brightness edges. Then a sky model is fitted and subtracted from each frame. Cosmic rays and satellite trails are also removed from each frame. At last, single frames are co-added together in a weighted-average way to optimize the signal-to-noise ratio of the final image, where the weight of each frame is inversely proportional to both the sky background brightness and the scaling of zeropoint. The

<sup>7</sup> Extensive tests have shown that the surface brightness profile calculated by **photutils** agrees with IRAF **Ellipse** result to a very high precision (deviation less than 0.1%).

<sup>8</sup> [https://jielai.zhang.github.io/files/Zhang\\_Jielai\\_201811\\_PhD\\_Thesis\\_excludech4.pdf](https://jielai.zhang.github.io/files/Zhang_Jielai_201811_PhD_Thesis_excludech4.pdf)

co-added image is also projected with pixel scale  $2.5''/\text{pixel}$ . Under this meticulous sky subtraction, DWFS achieves a  $1\sigma$  surface brightness detection limit of  $\mu_r = 29.6 \text{ mag}/\text{arcsec}^2$ ,  $\mu_g = 29.2 \text{ mag}/\text{arcsec}^2$  on scales of  $10''$ .

The Dragonfly Telephoto Array has already shown the tremendous ability to detect low surface brightness phenomena (van Dokkum et al. 2014; Merritt et al. 2016a,b; Zhang et al. 2018; Cohen et al. 2018). However, Dragonfly has the worst seeing ( $\text{FWHM} \sim 5 \text{ arcsec}$ ) and lowest resolution ( $2.5''/\text{pixel}$ ) among the four surveys, which makes blending a major issue in the Dragonfly data. For the study of low surface brightness outskirts, the scattered lights from nearby stars and galaxies often masquerade as low surface brightness features. Hence, it is crucial to accurately subtract compact objects and their associated scattered lights from the Dragonfly images. We mitigate these issues using multi-resolution filtering<sup>9</sup> (MRF, van Dokkum et al. 2019a). The algorithm models and subtracts compact objects based on high-resolution images [e.g. Canada France Hawaii Telescope (CFHT), DECaLS, SDSS], conserving target galaxies and low surface brightness features below a certain magnitude level. The algorithm also models star haloes by stacking bright stars. MRF has already shown its potential in Dragonfly data (van Dokkum et al. 2019b; Gilhuly et al. 2019; Danieli et al. 2019). We apply MRF to the DWFS images, mask out residual pixels, and extract 1-D surface brightness profiles using the same methodology as in Section 3.1. This method successfully subtracted scattered lights from neighbouring stars and galaxies, improving the 1-D profile of a galaxy's outskirts.

## 6 RESULTS

We compare the light profiles and stellar mass of massive galaxies measured from HSC, DECaLS, SDSS and Dragonfly. We then highlight issues that can arise when studying 1-D surface brightness profiles in magnitude units and advocate using fluxes instead.

### 6.1 Profiles of the Low Redshift Sample

We compare SDSS, DECaLS, HSC and Dragonfly on five low- $z$  galaxies as described in Section 2.2. The 1-D surface brightness profiles for low- $z$  sample is extracted and corrected for filter differences based on the methodology described in Section 3, 4 and 5. We use only the binary mask method instead of 2-D modelling (see Section 3.3) to remove satellite contamination. We remind our readers that we locally measure the sky background (see Section 3.2) and correct it at the profile level here, instead of using `SkyObjs`.

We noticed that even a slight over-subtraction of sky background (often the case for low- $z$  bright galaxies) can lead to negative surface intensity in the outskirts and sudden change of the profile expressed in  $\text{mag}/\text{arcsec}^2$ . To make the profile more reasonable and convey more information, we define an upper-limit profile by adding the  $1\sigma$  standard error of median sky value (i.e. the uncertainty of background,

as mentioned in Section 3.2) to the surface brightness profile. A lower-limit profile can also be defined in the same way, but it will be ill-defined when the true flux is smaller than the uncertainty of background, in which case we are concerned only about the upper-limit profile. A measured surface brightness profile can be in any place between the upper-limit and lower-limit profiles.

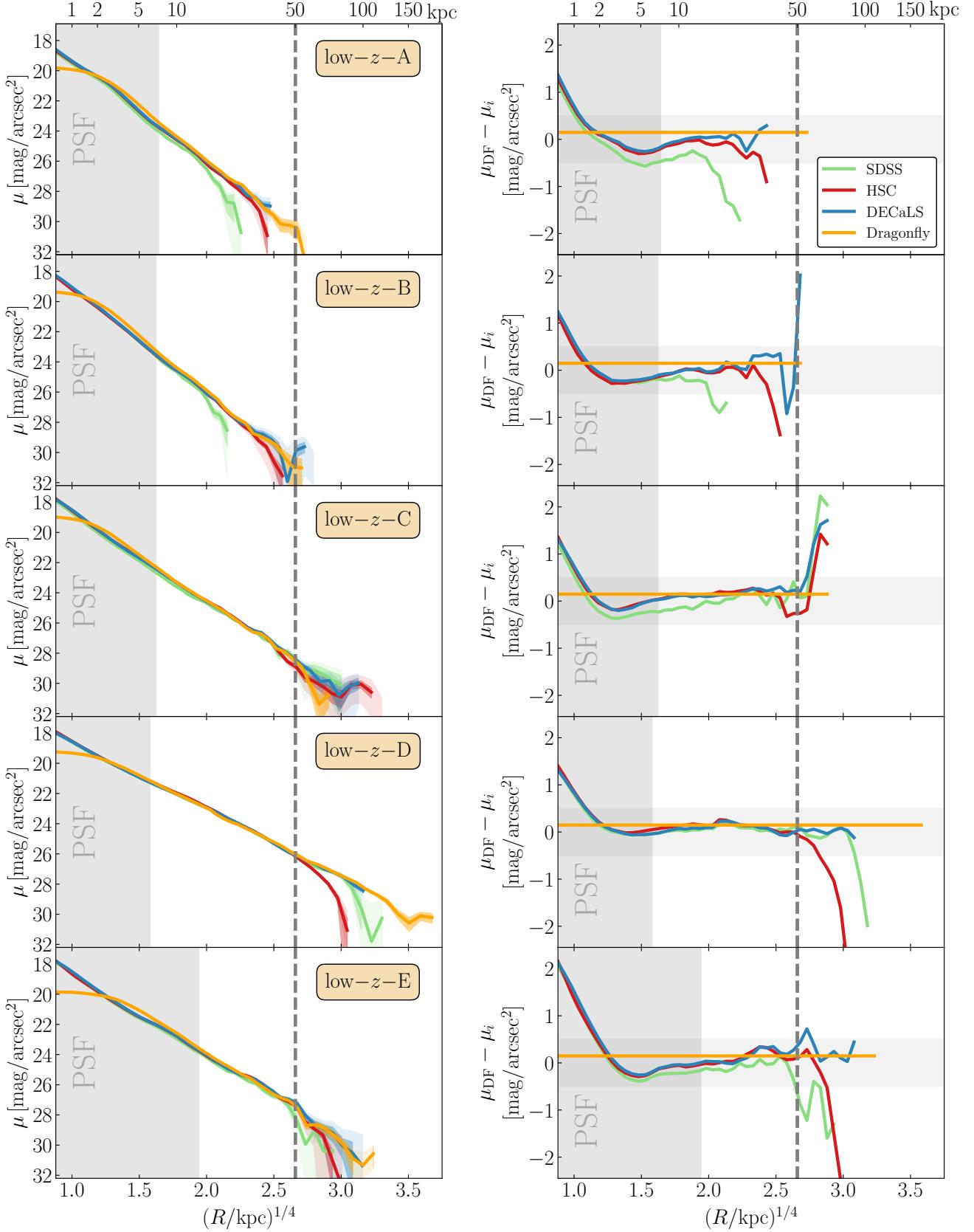
The 1-D surface brightness profile of each galaxy is shown in the left panels of Figure 7, colour-coded by surveys. The dark-shaded area around each profile indicates the uncertainties of surface brightness given by `Ellipse`. The area between the upper-limit and lower-limit profile is shown with a lighter shade. The right panels show the surface brightness differences between surveys with respect to the Dragonfly profiles. Since PSF smears the intrinsic inner profile, we highlight a region that corresponds to three times the typical FWHM of the Dragonfly PSF in radius ( $r = 3 \times \text{FWHM}$ ) using a grey shaded region. Among the four surveys, Dragonfly has the worst seeing condition and hence shows a significantly flattened inner profile compared to the others. The difference within the PSF-dominated region should be interpreted with caution. Although the inner profiles using SDSS, HSC and DECaLS show excellent agreement, we would like to draw the reader's attention to the outskirt in this work.

The key challenge for large sky surveys on imaging low- $z$  bright galaxies is the over-subtraction of the sky background. As shown in Figure 7, Dragonfly images are the most robust ones against the over-subtraction of sky background and often provide the most extended surface brightness profiles without a sign of sudden nonphysical truncation. Dragonfly's profiles agree with DECaLS and HSC profiles very well down to  $29 - 30 \text{ mag}/\text{arcsec}^2$ . Meanwhile, SDSS profiles have the most serious over-subtraction among these surveys. The sudden truncation of HSC profiles indicates that over-subtraction still exists for nearby bright galaxies in HSC S18A and is also very hard to compensate by empirical sky correction only at the coadd level. We attempted to correct the background level using `SkyObjs` around each galaxy but did not find noticeable improvement.

DECaLS is much shallower in terms of imaging depth measured by point source detection limit (see Section 2.1), but DECaLS displays outer profiles that are as extended as the HSC profiles and we find that DECaLS is not affected by the over-subtraction issue too much. It is probable that manually generated coadds with sky background well-considered at CCD level (see Section 4) enhanced the detection ability of DECaLS. Despite the poor seeing condition and worse spatial resolution, Dragonfly has an outstanding ability to detect low surface brightness features to a level deeper than  $30 \text{ mag}/\text{arcsec}^2$  owing to the well-controlled PSF, globally modelled sky and dedicated reduction pipeline including the MRF technique.

Based on the analysis above, DECaLS with a customized pipeline is an appropriate choice to reach down to  $\sim 29 \text{ mag}/\text{arcsec}^2$  in  $r$ -band for the purpose of studying outer profiles of  $z < 0.1$  bright galaxies. Dragonfly can further reach down to  $\sim 30 \text{ mag}/\text{arcsec}^2$  in  $r$ -band. However, for HSC, the over-subtraction around nearby galaxies impedes its ability. Both DECaLS and Dragonfly are reliable references for studies of low redshift massive galaxies. With the completion of DWFS, more comparisons between Drag-

<sup>9</sup> <https://github.com/AstroJacobLi/mrf>



**Figure 7.** Surface brightness profiles of five galaxies in low- $z$  sample for four different surveys (left panels) and the surface brightness differences with respect to Dragonfly results (right panels). The area between upper-limit and lower-limit profiles is shown with light shade. HSC, DECaLS and Dragonfly show excellent consistency on measuring 1-D surface brightness profiles for nearby massive galaxies down to  $29 - 30$  mag/arcsec<sup>2</sup>, whereas SDSS has serious over-subtraction for low- $z$  galaxies. The outermost parts of the profiles are noisy but consistent with each other considering the sky uncertainty and filter differences.

only and other higher resolution surveys can be done on various topics.

## 6.2 Profiles of the Intermediate Redshift Sample

Taking five nearby massive galaxies as examples, we have found good and robust agreements among different surveys on measuring 1-D profile. Now we enlarge our scope and study the profiles of intermediate- $z$  sample using HSC and DECaLS data. We divide the intermediate- $z$  sample into two redshift bins as  $0.19 < z < 0.35$  and  $0.35 < z < 0.50$ , as well as two mass bins  $11.6 < \log(M_*/M_\odot) < 11.8$  and  $11.8 < \log(M_*/M_\odot) < 12.1$ . Here we use the maximum 1-D stellar mass as  $M_*$ . HSC and DECaLS surface brightness profiles are extracted according to the methodology described in Section 3 and Section 4. For each galaxy, the DECaLS profile shares the same fixed shape (ellipticity and position angle, which are determined from the HSC profile) with the HSC profile. As described in Section 3.2, Section 6.3 and Appendix A, we correct all HSC profiles afterwards with  $8.4''$  aperture `SkyObjs`. The filter difference is also corrected for every DECaLS profile.

The left panels of Figure 8 show the results for different redshift bins and mass bins. Five randomly selected profiles in each redshift bin and mass bin are shown in coloured lines, along with the median profiles in each bin shown in thick black lines. HSC profiles are shown in solid lines, whereas DECaLS profiles are in dashed lines. The shaded region around a single profile is uncertainty given by `Ellipse`. We estimate the uncertainty of the median profile by bootstrap resampling. The right panels of Figure 8 show the surface brightness differences  $\mu_{\text{HSC}} - \mu_{\text{DECaLS}}$  of the two median profiles on the left. When displaying the median profile, we discard those data points where more than 30% of the profiles are undefined (i.e. `NaN`).

The typical seeing of DECaLS  $r$ -band is  $1.2''$  corresponding to  $4.0$  kpc at  $z = 0.20$  and  $7.3$  kpc at  $z = 0.50$ . To avoid the smearing effect of seeing, three times FWHM of DECaLS seeing (in radius) is shown as a shaded region in Figure 8. DECaLS has larger seeing (FWHM  $\sim 1.2''$  in  $r$ -band) than HSC (FWHM  $\sim 0.7''$  in  $r$ -band), thus the large PSF distributes lights to the outer region and makes the inner profile flatter than the HSC inner profile. PSF far-wing structures could also affect outskirts profiles (e.g. [de Jong 2008](#); [Tal & van Dokkum 2011](#)). We discuss the impact of PSF wings on the study of low surface brightness features in Section 7.

Individual HSC profiles can be studied well to  $\sim 150$  kpc, and  $100$  kpc is a conservative choice. A single DECaLS profile can reach to  $\sim 100$  kpc but also fluctuates more in the outskirts. We are confident that HSC median profiles reach farther than  $200$  kpc without sign of over- or under-subtraction. HSC and DECaLS median profiles agree with each other well to  $\sim 100$  kpc ( $\sim 28$  mag/arcsec $^2$ ). However, the surface brightness measured by HSC tends to be brighter in a statistical sense than DECaLS surface brightness in the outskirts ( $R > 100$  kpc). The deviation has almost the same trend and magnitude in different stellar mass and redshift bins. From the right panels, we do not see strong redshift or stellar mass dependence of the median  $\mu_{\text{HSC}} - \mu_{\text{DECaLS}}$ . This suggests that slight over-subtraction occurs in the DECaLS data, or slight under-subtraction still exists in the HSC data.

Nonetheless, we exclude the latter possibility in Section 6.3 by demonstrating the median HSC intensity profile.

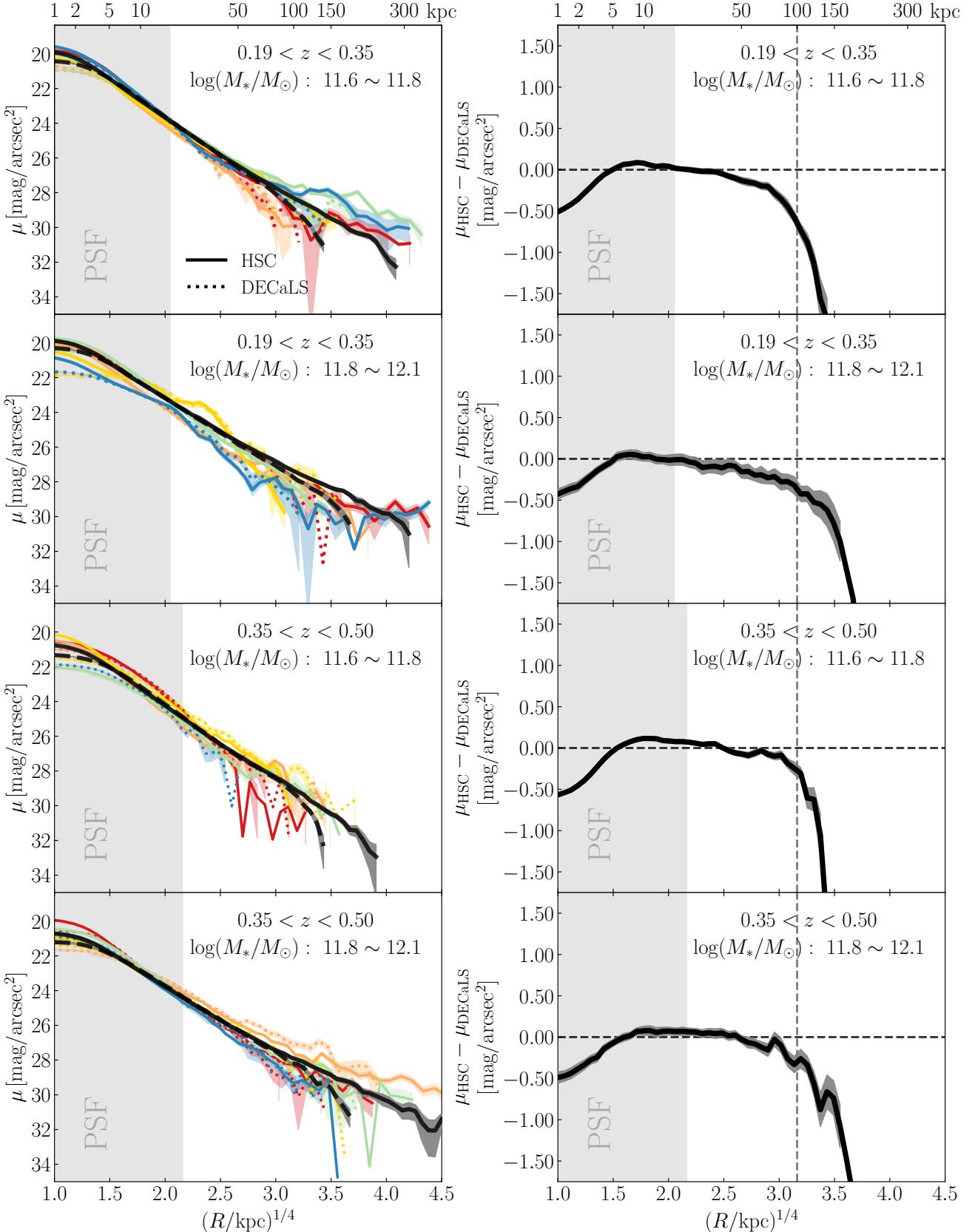
Although HSC and DECaLS median profiles are consistent with each other within  $\sim 100$  kpc, individual profiles do not necessarily agree well with each other. Instead, the surface brightness difference profile ( $\mu_{\text{HSC}} - \mu_{\text{DECaLS}}$  as a function of  $R$ ) of a single galaxy is very noisy in the outskirts. Figure 9 demonstrates the median (thick brown lines) and the scatter of surface brightness difference profiles for all galaxies in two redshift bins. Three different shades indicate  $0.5\sigma$ ,  $1\sigma$  and  $2\sigma$  deviations of the difference profile distributions (not the deviation of median profile). The deviations of DECaLS median profiles from HSC median profiles are small (less than  $0.25$  mag) above  $\mu_r \sim 28$  mag/arcsec $^2$ . But the scatters of  $\mu_{\text{HSC}} - \mu_{\text{DECaLS}}$  are quite large below this surface brightness limit. Other than intrinsic scatters and other observational systematics, the involvement of `SkyObjs` might introduce extra systematics to HSC profiles. Although the median profiles robustly extend to more than  $100$  kpc, it is still very hard to measure an individual 1-D profile reliably to very far ( $R > 100$  kpc), whether it is an HSC profile or a DECaLS profile.

## 6.3 Median profiles in linear scale

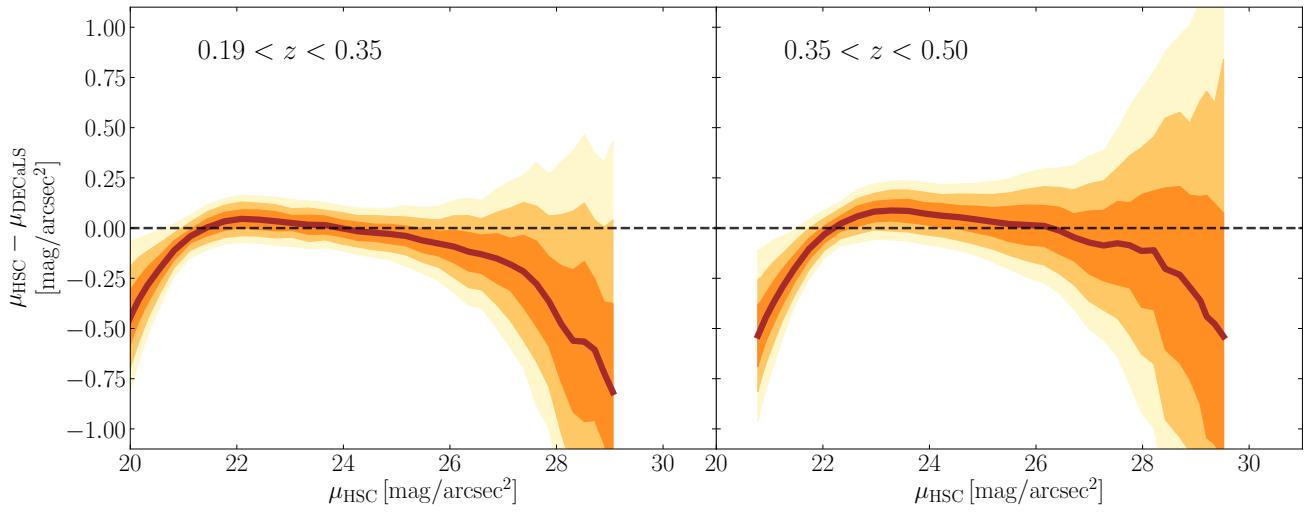
In this subsection we argue that showing the surface brightness profile in magnitude could be misleading in low surface brightness studies and bias the understanding of observation results.

In the reduction of deep images, over-subtraction and large sky background uncertainty can result in negative pixels, but they are not non-physical and do provide information on both galaxy outskirts and the sky background. The biggest problem of logarithmic magnitude is that it discards negative pixels because of the undefined behavior of `log` function at negative values. If magnitude is used, all useful information contained by negative pixels gets lost both at the single profile level and at the statistical level. If the median profile of a galaxy ensemble is calculated at the magnitude level, it would be biased to be brighter in the outskirts, since lots of negative values are discarded by logarithm. Furthermore, the sensitivity (i.e. derivative) of the logarithmic magnitude is different at different surface brightness. Slight changes of pixel counts in the bright regions will not alter the magnitude too much, but it could cause notable changes of magnitude in the low surface brightness region. Hence, a large scatter of magnitude (mag/arcsec $^2$ ) in the outskirts does not necessarily imply a large scatter of intensity (such as  $\mu\text{Jy}/\text{arcsec}^2$ ).

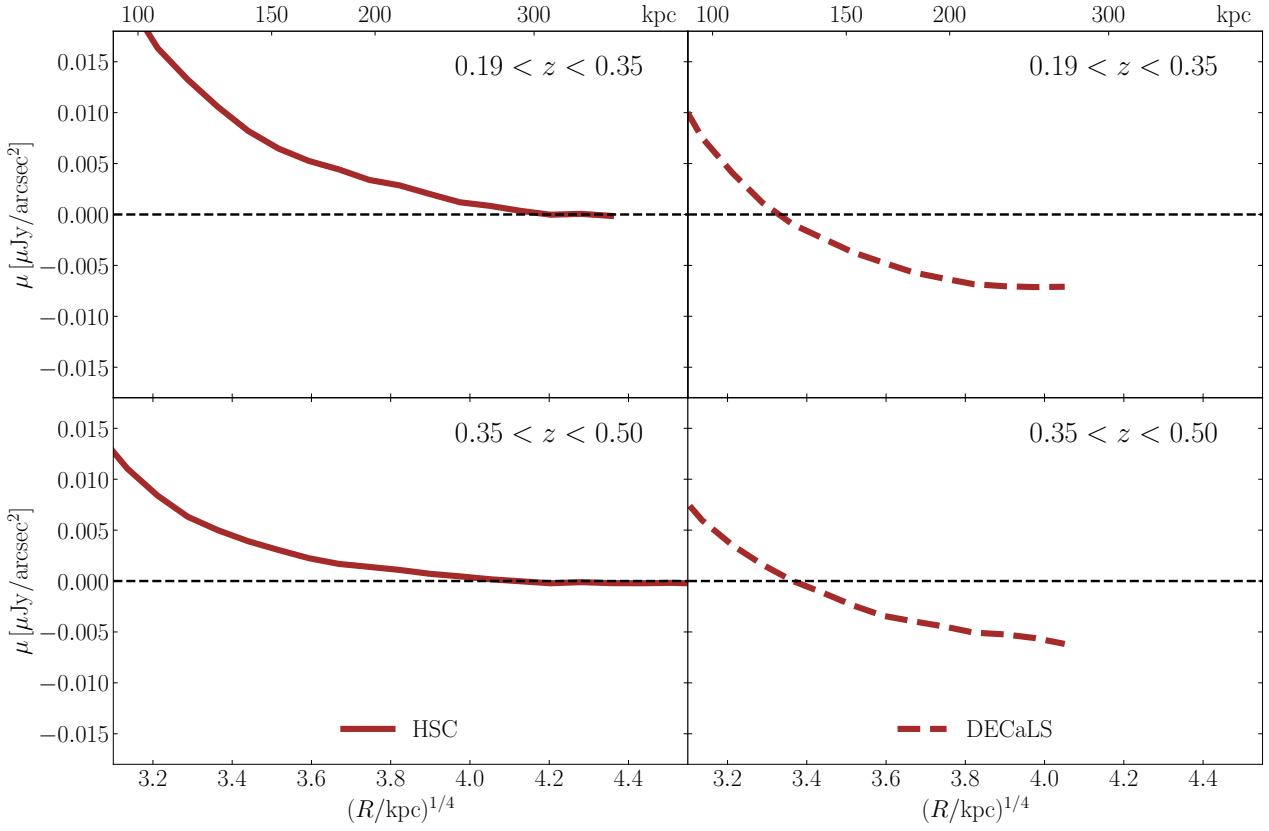
Therefore, we suggest expressing 1-D surface brightness profiles in the unit of intensity, especially for the study of low surface brightness outskirts. We show an example in Figure 10, which uses  $\mu\text{Jy}/\text{arcsec}^2$  as the unit of surface brightness. Since the median profiles of HSC and DECaLS coincide well before  $100$  kpc (see Figure 8 and Figure 9), here we show only the region beyond  $100$  kpc. The median profile is shown as the thick brown line, and three transparencies indicate  $0.5\sigma$ ,  $1\sigma$  and  $2\sigma$  standard deviation of the profile distribution. We still use  $8.4''$  `SkyObjs` for HSC sky residual correction here. By plotting intensity profiles, we do not discard any negative intensities in any single profile before calculating the median, thereby maintaining complete statistical information.



**Figure 8.** Surface brightness profiles of intermediate- $z$  massive galaxies for HSC and DECaLS (left panels) and the difference of two median profiles (right panels). Five randomly selected profiles are shown in coloured lines, together with the median profiles shown in thick black lines. The median HSC profiles smoothly reach farther than 100 kpc. Good agreements (in a statistical sense) between HSC and DECaLS profiles are achieved before  $\sim 100$  kpc, but the median HSC profiles are higher than DECaLS median profiles in the outskirts, suggesting little over-subtraction in DECaLS data even after subtracting sky carefully at CCD level.



**Figure 9.** Median of surface brightness difference profile (shown in thick brown lines) and the distributions of individual surface brightness profile. Three different shades indicate  $0.5\sigma$ ,  $1\sigma$  and  $2\sigma$  deviations of individual  $\mu_{\text{HSC}} - \mu_{\text{DECaLS}}$  profile distributions. Although HSC and DECaLS median profiles coincide well to  $\sim 100$  kpc, the scatter beyond  $\sim 28$  mag/arcsec $^2$  is quite large. One should be very careful when interpreting the outer parts of individual surface brightness profile of intermediate- $z$  massive galaxies.



**Figure 10.** Surface brightness profiles of the intermediate- $z$  sample shown in intensity unit  $\mu\text{Jy}/\text{arcsec}^2$ . The median profile is shown in thick brown line, the three transparencies indicate  $0.5\sigma$ ,  $1\sigma$  and  $2\sigma$  standard deviation of the profile distribution. HSC sky residual can be well eliminated by using an 8.4-arcsec aperture `Sky0bj`s, as the median profiles converge to zero. However, over-subtraction in DECaLS data can be clearly seen, whereas it is hard to diagnose when viewing profiles in logarithmic magnitude.

Assuming perfect sky background subtraction, the median profile should gradually converge to zero intensity, and the distribution of profiles should be symmetrically centered at zero. In Figure 10, the HSC median profile converges to zero very well, whereas DECaLS median profile shows strong evidence of over-subtraction by  $\sim 0.075 \mu\text{Jy}/\text{arcsec}^2$ . Hence, even using customised pipeline for co-addition, it is still challenging to avoid over-subtraction in DECaLS images. We also tested `SkyObj`s with different aperture sizes, and it turns out correction with  $8.4''$  aperture `SkyObj`s gives the best convergence to zero. The good performance of HSC median profiles in linear scale affirms the plausibility and robustness of `SkyObj`s corrections again.

#### 6.4 Robustness of Mass Estimates

One of the main reasons to extract a surface brightness profile is to derive the luminosity and stellar mass of a galaxy. Here we examine the consistency of stellar mass measurements between HSC and DECaLS for the intermediate- $z$  sample. First, we measure the apparent magnitude of a galaxy within physical radius  $R$  by integrating the surface brightness profile within  $R$  for HSC and DECaLS; then we apply filter difference corrections (see Section 4) and calculate the galaxy's absolute magnitudes based on the distance given by redshift. We also calculate the absolute magnitudes of the Sun under HSC and DECaLS filters according to Willmer (2018). Since we only care about the stellar mass difference, the K-corrections for one certain galaxy cancels out directly if we assume that a little filter difference does not induce large deviation of K-correction. Therefore, we obtain the luminosity difference between HSC and DECaLS by converting the absolute magnitude difference to luminosity ratio. Because we are considering the luminosity and stellar mass difference of a certain galaxy, the mass-to-light ratio term also cancels out after logarithm. In this way, we obtain the stellar mass difference between HSC and DECaLS.

Stellar mass within 10 kpc and 100 kpc (e.g.  $M_{10 \text{ kpc}}^*$  and  $M_{100 \text{ kpc}}^*$ ) could reveal the past growth of ETGs (e.g. Huang et al. 2018a). Therefore, we compare the HSC and DECaLS values for  $M_{10 \text{ kpc}}^*$  and  $M_{100 \text{ kpc}}^*$  to check the robustness of mass measurements, as shown in Figure 11. On average, the stellar mass estimated by the two surveys is almost the same for individual galaxies, with median deviations less than 0.1 dex. Figure 11 also indicates that the stellar mass measurements based on HSC and DECaLS is quite robust: The scatter (standard deviation) of stellar mass difference is small (less than 0.05 dex). Stellar mass measured from HSC data is larger than that from DECaLS data for both  $M_{10 \text{ kpc}}^*$  and  $M_{100 \text{ kpc}}^*$ . The larger DECaLS PSF scatters inner lights into outer region, making  $M_{10 \text{ kpc}}^*$  measured by DECaLS slightly lower. For  $M_{100 \text{ kpc}}^*$ , over-subtraction of DECaLS data significantly reduces the measured stellar mass in the outskirts. We also tested the measurements of stellar mass within 50 kpc to 100 kpc as an indicator of outskirts mass. Typically, HSC measures 0.07 dex more stellar mass within 50 kpc – 100 kpc than does DECaLS. We also note that wrong filter correction causes notable shifts of the median mass difference  $\Delta M$ . Detailed filter correction based on SED-fitting of each galaxy is needed to make a better comparison of stellar mass measurement, but this is beyond the scope of this paper.

## 7 DISCUSSION

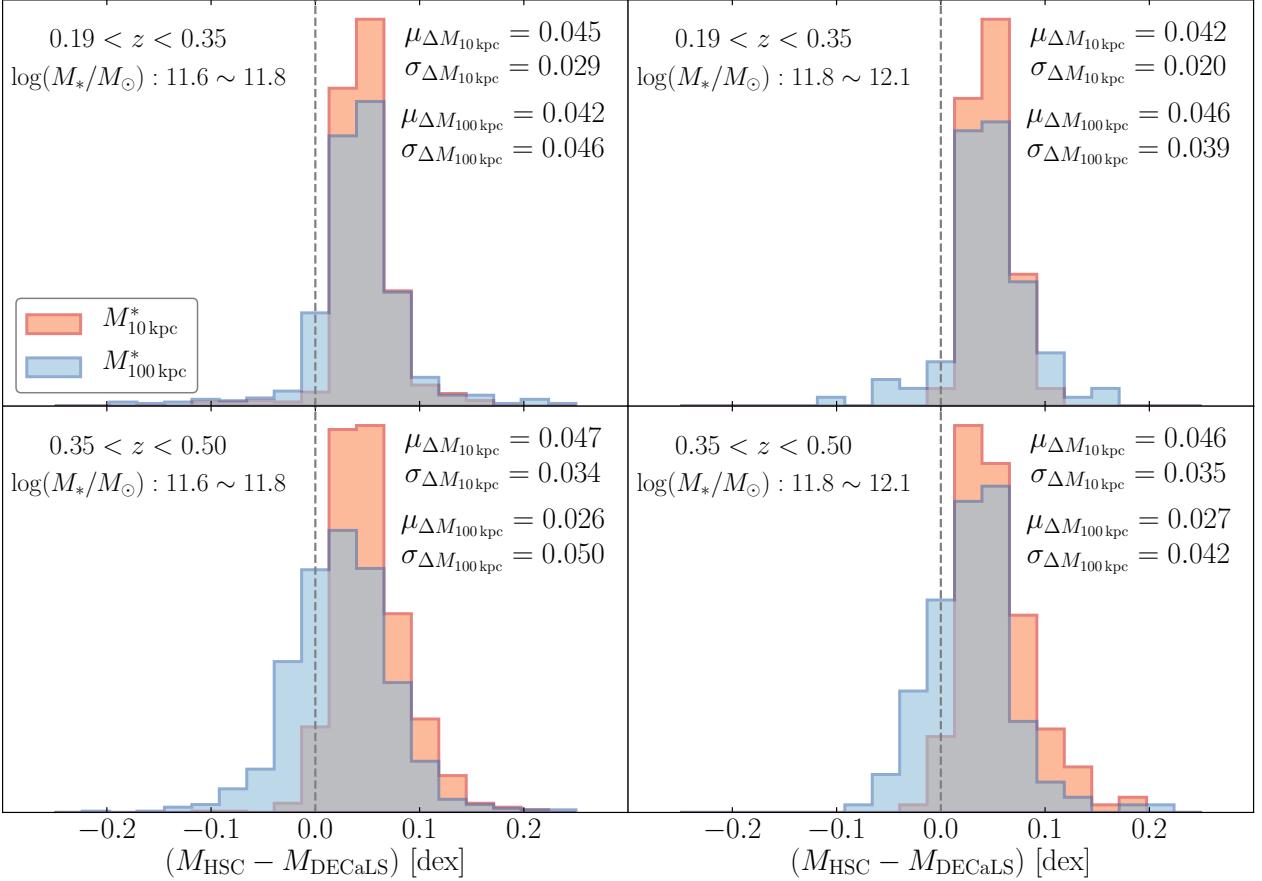
Here we briefly discuss several systematic issues concerning the use of deep imaging surveys to study the low surface brightness outskirts of galaxies along with a few possible future research directions.

### 7.1 Further Systematic Concerns

- **Sky background uncertainty:** background estimation remains the most important systematic when mapping the outskirts of galaxies. We show that the HSC S18A data can recover the *average* profiles of massive galaxies at  $0.2 < z < 0.5$  out to  $> 150$  kpc thanks to the global sky correction adopted by the HSC pipeline (Aihara et al. 2019). However, the profile of an individual galaxy in HSC can still be affected at  $r > 100$  by background over- or under-subtraction. The data reduction pipelines of modern imaging surveys are usually not designed to preserve the extended structure around a bright object as doing so increases the difficulty for an image deblender and decreases the useful area for the main scientific goal (e.g. cosmology). A carefully designed survey strategy and reduction pipeline (e.g. the Dragonfly survey) or specially tailored pipeline to redo the background subtraction (e.g. the DECaLS survey) is often required to mitigate background-related systematics. As shown in this work, we also highly recommend mock tests and cross-survey comparison before reaching conclusions regarding galaxy outskirts.

- **Low surface brightness structure of the PSF:** PSFs from modern imaging surveys often display a complex outer structure due to issues related to the detector (e.g., Michard 2002; Wu et al. 2005; Sirianni et al. 1998), the structure of the telescope or the optical system (e.g. Racine 1996; Bernstein 2007; Slater et al. 2009), and even the atmosphere (e.g. DeVore et al. 2013). An ill-characterized PSF ‘wing’ can lead to a wrong or biased conclusion in the LSB regime (Tal & van Dokkum 2011; Sandin 2014). Because we use a red-sensitive thick-CCD, we often assume the PSF from HSC, DECaLS and LSST should show a much less prominent outer structure than the outer structures shown by a fully depleted thin-CCD (e.g. SDSS) in the same filter. For HSC images, PSF wings are not a concern when studying the profile of low- $z$  massive galaxies at  $r < 29 \text{ mag}/\text{arcsec}^2$  (e.g., Wang et al. 2019). However, the outer PSF structures in HSC and DECaLS are still not carefully characterized and could hinder us from taking further advantage of increasingly deep imaging surveys. We will investigate this in future work by stacking non-saturated stars and via comparison to images with a much better controlled outer PSF wing (e.g. Dragonfly, Abraham & van Dokkum 2014, Q. Liu et al., in preparation).

- **Galactic cirrus:** It has recently become clear that the diffuse Milky Way dust emission known as the Galactic cirrus is a major challenge when exploring the low surface brightness Universe (Duc et al. 2015; Román et al. 2019). At  $\mu_V > 26 \sim 27 \text{ mag}/\text{arcsec}^2$ , Galactic cirrus start to contaminate extragalactic observation even at high Galactic latitude  $b > 75^\circ$  (Mihos et al. 2017), and it could cover up to 80% of the whole sky down to  $\mu_V = 30 \text{ mag}/\text{arcsec}^2$ . In light of the next-generation deep imaging surveys (e.g. LSST) and projects designed for LSB studies (e.g., Dragon-



**Figure 11.** Histograms of stellar mass differences  $\Delta M = M_{\text{HSC}} - M_{\text{DECaLS}}$  measured by HSC and DECaLS. On average, the stellar mass estimated by the two surveys is almost the same (median deviation less than 0.05 dex), and the scatter of stellar mass difference is also small (less than 0.05 dex). This ensures the credibility of stellar mass measured by DECaLS even though it has over-subtraction.

fly, MESSIER surveyor (Valls-Gabaud & MESSIER Collaboration 2017), Australian Space Eye (Horton et al. 2016), and Huntsman Telescope (Longbottom 2019)), it is a pressing task to develop the capability of disentangling Galactic cirrus from low surface brightness extragalactic structures.

## 7.2 Future Directions

- **Exploring the intra-cluster light:** In this work, we describe the outer profiles of massive early-type galaxies. Meanwhile, massive central galaxies of galaxy clusters often host an extremely extended LSB envelope known as the intra-cluster light (ICL, e.g., Carlberg et al. 1997; Lin & Mohr 2004; Gonzalez et al. 2005; Mihos et al. 2005; Mihos 2016). An ICL consists of stars that follow the gravitational potential of the dark matter halo of the host cluster and could represent a crucial piece of assembly history of the massive halo (e.g., Rudick et al. 2006, 2011; Cui et al. 2014; Cooper et al. 2015). Using stacked 1-D profiles from the Dark Energy Survey (DES), Zhang et al. (2019) recently revealed a large amount of ICL in the very outskirts of low- $z$  brightness cluster galaxies (BCGs). If true, this result means the stellar mass around these BCGs can double itself outside the radii that can be probed by HSC ( $> 150$  kpc) and could have profound implications for the stellar-halo connection of

massive galaxies. We will examine this result by comparing the 1-D profiles of common BCGs in HSC and DES, and by performing careful image stacking analysis using HSC BCG sample. DECaLS shares the same camera with DES and can also provide an independent check using a large sample of nearby clusters (J. Moustakas et al., in preparation). In a companion paper (F. Ardila et al., in preparation), we also derive a scheme for correcting an individual stellar mass measurement based on hydrodynamical simulations.

- **Extending to multi-filters:** In this work, we use only single-band images to explore the outskirts of massive galaxies. Meanwhile, rest-frame optical colour of the stellar halo contains important information about the underlying stellar population (Carollo et al. 1993; La Barbera et al. 2012; D’Souza et al. 2014) and helps us convert observed the surface brightness profile into stellar mass density. Current stellar mass estimates of massive galaxies often use an average colour and assume constant  $M/L$  while knowing it is not true. Including a radial  $M/L$  profile could further help us improve the stellar mass measurements. However, due to the difference of background subtraction and seeing between two bands, accurate recovery of colour in the LSB outskirts of massive galaxies is extremely difficult and requires the consideration of most of the systematics mentioned in the previous section (e.g., Wang et al. 2019). We will first eval-

ate the reliability of the *average* colour profile recovered by HSC using mock galaxies.

- **The quest of better ‘total’ stellar mass:** Accurate and precise stellar mass of massive galaxies are essential for building galaxy-halo connection models and using them as cosmological probes in the future (e.g. LSST, DESI). Here we show that the stellar mass of massive galaxies measured within 100 kpc using DECaLS images is consistent with the stellar mass measurements using deeper HSC images to  $\sim 0.05$  dex level using a customized reduction pipeline. This careful treatment will also benefit the DESI survey, which will include many massive galaxies in the Luminous Red Galaxy (LRB) and low- $z$  Bright Galaxy Survey (BGS) samples. At the same time, we do know that the stellar mass distribution of massive galaxies extends beyond 100 or 150 kpc and that the 1-D profile approach has its limitation. In the future quest for an even more accurate ‘total’ stellar mass for these massive galaxies, we will rely on stacked 1-D profiles using HSC and hydro-simulations to explore a reliable way to extrapolate beyond 150 kpc (J. Li et al., in preparation). We will also test the potential of 2-D forward modeling to overcome some of the difficulties faced by 1-D approach (e.g. seeing effect, cannot deal with mergers) and include many systematics issues in the model. We hope that a combination of these methods will help us reach to the edge of massive galaxies and paint a complete picture of their stellar mass distributions.

## 8 SUMMARY AND CONCLUSIONS

In this paper, we have explored the ability of different datasets to extract the outer light profiles (also known as the stellar haloes) of massive galaxies. We have performed a variety of tests on the robustness of such measurements with HSC. We then present comparisons between the light profiles extracted with HSC, DECaLS, SDSS and Dragonfly. Our main results are summarized below.

- Several iterations on masking and sky subtraction were necessary to get results to match-up between different datasets. This cross-survey comparison was a useful exercise and helped to identify a number of fixable issues along the way.

- The largest systematic impacting outer surface brightness profiles is sky subtraction. The HSC S18A release (or ‘PDR2’) uses a global background calibration approach. This approach is superior to the method that was employed in previous data releases. However, in S18A, `hscPipe` still over-subtracts the background around bright low- $z$  galaxies. On average, there is also a small residual background in the  $r$ -band. Researchers should be aware of changes in the sky subtraction schemes that occur for different HSC data releases.

- We used HSC, DECaLS, SDSS and Dragonfly to compare profiles for nearby low- $z$  massive galaxies. SDSS suffers from sky subtraction for low- $z$  galaxies and has lower signal to noise in the outskirts.

- Thanks to its excellent control of scattered light, a carefully designed strategy for observations and a dedicated data reduction pipeline, Dragonfly is better than HSC and DECaLS at detecting low surface brightness features for nearby massive galaxies (reaching  $\mu_r \sim 30$  mag/arcsec $^2$ ).

The special DECaLS pipeline can also recover the surface brightness profiles of low redshift galaxies when sky subtraction is performed at the individual CCD level.

- However, Dragonfly’s poor spatial sampling (2.5''/pixel) and large seeing disc mean that Dragonfly cannot accurately probe galaxies at higher redshifts or the inner regions of galaxies at low redshifts. The current DWFS (330 deg $^2$ ) does not probe enough volume to sample many massive galaxies (limited by survey volume). DECaLS covers a larger volume and hence will provide larger samples of low redshift massive galaxies.

- We used HSC and DECaLS to compare intermediate- $z$  massive galaxy profiles. We have shown that the light profiles of massive galaxies in HSC can be reliably recovered to  $\mu_r \sim 28.5$  mag/arcsec $^2$  using a series of tests with mock galaxies. The light profiles of individual galaxies can be studied to  $\sim 150$  kpc and 100 kpc is a conservative choice. HSC profiles on individual galaxies are noisy at large radii but unbiased.

- For intermediate- $z$  massive galaxies, we can also recover an unbiased average (median) profile out to  $> 200$  kpc for galaxies with similar outer profiles (i.e. similar stellar mass and redshift). This average profile in HSC is unbiased on these radial scales (Figure 10). In HSC, the average light profiles in two stellar mass bins extend smoothly to  $\sim 300$  kpc without any sign of artificial truncation or flattening due to residual background issues (Figure 8).

- A single DECaLS profile can go to 100 kpc for massive galaxies at  $0.19 < z < 0.50$ , but DECaLS profiles are noisier than HSC profiles in the outskirts. The average DECaLS profile also reaches to  $\sim 100$  kpc.

- Cross-checks with DECaLS data were also useful. HSC and DECaLS average profiles agree with each other to  $\sim 100$  kpc ( $\mu_r \sim 28$  mag/arcsec $^2$ ). But DECaLS profiles are often truncated at  $r \sim 100$  kpc. The sky background of DECaLS is probably over-subtracted (Figure 10) and thus biases the profiles. However, if the customised pipeline by J. Moustakas is applied, DECaLS can recover the stellar mass within 100 kpc from HSC to within 0.05 dex.

Looking forward, Paper II (F. Ardila et al., in preparation) in this series will explore techniques for robustly measuring the light profiles of massive galaxies above 100 kpc. Paper III (J. Li et al., in preparation) will explore image stacking techniques to perform measurements out to larger radii and will delve more deeply into the effect of PSF wings.

## ACKNOWLEDGEMENTS

The research reported in this paper is supported by the scholarship from China Scholarship Council (CSC) under Grant CSC # 201806010530 during JL’s stay at University of California, Santa Cruz. This material is based on work supported by the U.D Department of Energy, Office of Science, Office of High Energy Physics under Award Number DE-SC0019301. AL acknowledges support from the David and Lucille Packard foundation, and from the Alfred P. Sloan foundation. JM gratefully acknowledges funding support from the National Science Foundation grant number AST-1616414.

The Hyper Suprime-Cam (HSC) collaboration includes the astronomical communities of Japan and Taiwan, and

Princeton University. The HSC instrumentation and software were developed by National Astronomical Observatory of Japan (NAOJ), Kavli Institute for the Physics and Mathematics of the Universe (Kavli IPMU), University of Tokyo, High Energy Accelerator Research Organization (KEK), Academia Sinica Institute for Astronomy and Astrophysics in Taiwan (ASIAA) and Princeton University. Funding was contributed by the FIRST program from Japanese Cabinet Office, Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan Society for the Promotion of Science (JSPS), Japan Science and Technology Agency (JST), Toray Science Foundation, NAOJ, Kavli IPMU, KEK, ASIAA and Princeton University.

This project used data obtained with the Dark Energy Camera (DECam), which was constructed by the Dark Energy Survey (DES) collaboration. Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundacao Carlos Chagas Filho de Amparo, Financiadora de Estudos e Projetos, Fundacao Carlos Chagas Filho de Amparo a Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Cientifico e Tecnologico and the Ministerio da Ciencia, Tecnologia e Inovacao, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey. The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energeticas, Medioambientales y Tecnologicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenossische Technische Hochschule (ETH) Zurich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciencias de l'Espai (IEEC/CSIC), the Institut de Fisica d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universitat Munchen and the associated Excellence Cluster Universe, the University of Michigan, the National Optical Astronomy Observatory, the University of Nottingham, the Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, and Texas A&M University.

We gratefully acknowledge NSF grant AST-1613582, which supports the Dragonfly Wide Field Survey.

Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, National Science Foundation, the U.S. Department of Energy, National Aeronautics and Space Administration, the Japanese Monbukagakusho Scholarship, the Max Planck Society, and the Higher Education Funding Council for England. The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The Partic-

ipating Institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, University of Cambridge, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle Astrophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, the Max Planck Institute for Astronomy (MPIA), the Max Planck Institute for Astrophysics (MPA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the U.S. Naval Observatory, and University of Washington.

This research made use of: `NumPy`, a fundamental package for scientific computing with Python (van der Walt et al. 2011); `SciPy`, an open source scientific tool for Python (Jones et al. 01 ); `Matplotlib`, a 2-D plotting library for Python (Hunter 2007); `Astropy`, a community-developed core Python package for Astronomy (Astropy Collaboration et al. 2013); `sep`, a Python library for Source Extraction and Photometry (Bertin & Arnouts 1996; Barbary 2016); `SWarp`, a program that resamples and co-adds FITS images using any arbitrary astrometric projection defined in the WCS standard (Bertin et al. 2002); `SCAMP`, a program for automatic astrometric and photometric calibration (Bertin 2006); `sedpy`, a Python package for astronomical SEDs; `The Tractor`, a tool for optimizing or sampling from models of astronomical objects (Lang et al. 2016); `GalSim`, the modular galaxy image simulation toolkit (Rowe et al. 2015); `Photutils`, an Astropy package for detection and photometry of astronomical sources (Bradley et al. 2019); `mrf`, a method for isolating faint, extended emission in Dragonfly data and other low-resolution images (van Dokkum et al. 2019a).

## APPENDIX A: MOCK TESTS ON HSC SKY BACKGROUND

The mock test results shown in Section 3.2 verify the rationality of our methodology of extracting surface brightness profiles from HSC images. In this appendix, we briefly summarize the `SkyObj` detection method, ways to use `SkyObj` to correct sky background residual in HSC data release S18A and caveats.

`hscPipe` (version 6.7) detects void areas in the single CCD images and uses different aperture sizes to evaluate the flux in these areas. `hscPipe` first masks out pixels with detection, then randomly selects void regions in the remaining sky; the regions are called `SkyObjs`. The pipeline picks 100 random `SkyObjs` in one patch (about  $4000 \times 4000$  pixels, i.e.  $12' \times 12'$ ). There are six aperture diameters:  $2.0'', 3.0'', 4.0'', 5.7'', 8.4''$  and  $11.8''$ . The flux inside these sky objects represents the local sky background value; hence they can help us characterize the sky better. In using `SkyObj` around a certain galaxy, we have two parameters: aperture size and matching range. In the following tests we show that matching `SkyObjs` between  $1'$  and  $4'$ , and using  $8.4''$  aperture behaves best on HSC S18A data.

We randomly select 300 `SkyObjs` as the position to put fake galaxies. To evaluate the local sky value around these

fake galaxies, we match all surrounding `SkyObjs` and calculate the mean values. First, fixing the matching range to be between  $1'$  and  $4'$ , we match all `SkyObjs` around one fake galaxy with a given aperture size, and then calculate their mean value. To exclude outlying `SkyObjs`, we execute a five-time sigma-clipping with a  $3\sigma$  threshold. Then we change the aperture size and go through all the fake galaxies. The red violin plots in Figure A1 show the distributions of mean `SkyObjs` value per pixel and their relative deviations with respect to the mean `SkyObjs` value of  $8.4''$  aperture size. The black dot shows the median of the violin plot and the error bar represents  $1\sigma$  deviation. The trends in the two left panels are the same: the mean `SkyObjs` value with  $8.4''$  aperture size is the highest, while `SkyObjs` with  $3.0''$  aperture size has the lowest mean value. Considering that the FWHM of  $r$ -band seeing disc is roughly  $1''$ , the aperture size used to estimate sky value should not be smaller than  $3''$ . The relative deviations of mean `SkyObjs` values with different aperture sizes are quite obvious, as the lower left panel shows.

With aperture size fixed to  $8.4''$ , we study how the matching range affects the mean value of `SkyObjs` using the same method as above. Also considering the possibility that sky objects too close to the galaxy might be contaminated by the light of the galaxy itself, we match only `SkyObjs` outside a circle with  $1'$  radius. This excludes `SkyObjs` within 190 kpc for  $z = 0.19$  and 370 kpc for  $z = 0.50$ , where they are far enough from the central galaxy to be less likely to be contaminated. We vary the upper limit of matching from  $2'$  to  $6'$ . The blue violin plots in Figure A1 show the distributions of mean `SkyObjs` values within different matching ranges. The medians stay nearly constant with the change of matching range. The violin plot with the upper limit of matching range equal to  $2'$  has very large scatter, which might be caused by the small number of `SkyObjs` within this range. With the angular size of a single CCD of HSC about  $6' \times 12'$ , the larger matching range does not make sense since the variation of sky background is typically CCD-level.

Hence, we fix the matching range to be  $1' < D < 4'$  and analyze the mock test results from Section 3.2 as follows. We subtract the profiles for 60 mock images with the mean values of `SkyObj` with different aperture sizes, then plot the median profiles in Figure 4. The mean value of `SkyObjs` increases as the aperture size increases, since larger `SkyObjs` include faint sources and lights from nearby bright objects. This trend is consistent with the violin plot. By comparing the median profiles with the model profiles, we find that sky correction using  $8.4''$  aperture could recover the model best for relatively low- $z$  (brighter) galaxies, whereas  $5.7''$  aperture is a better choice for sky correction around galaxies at higher redshift (fainter).

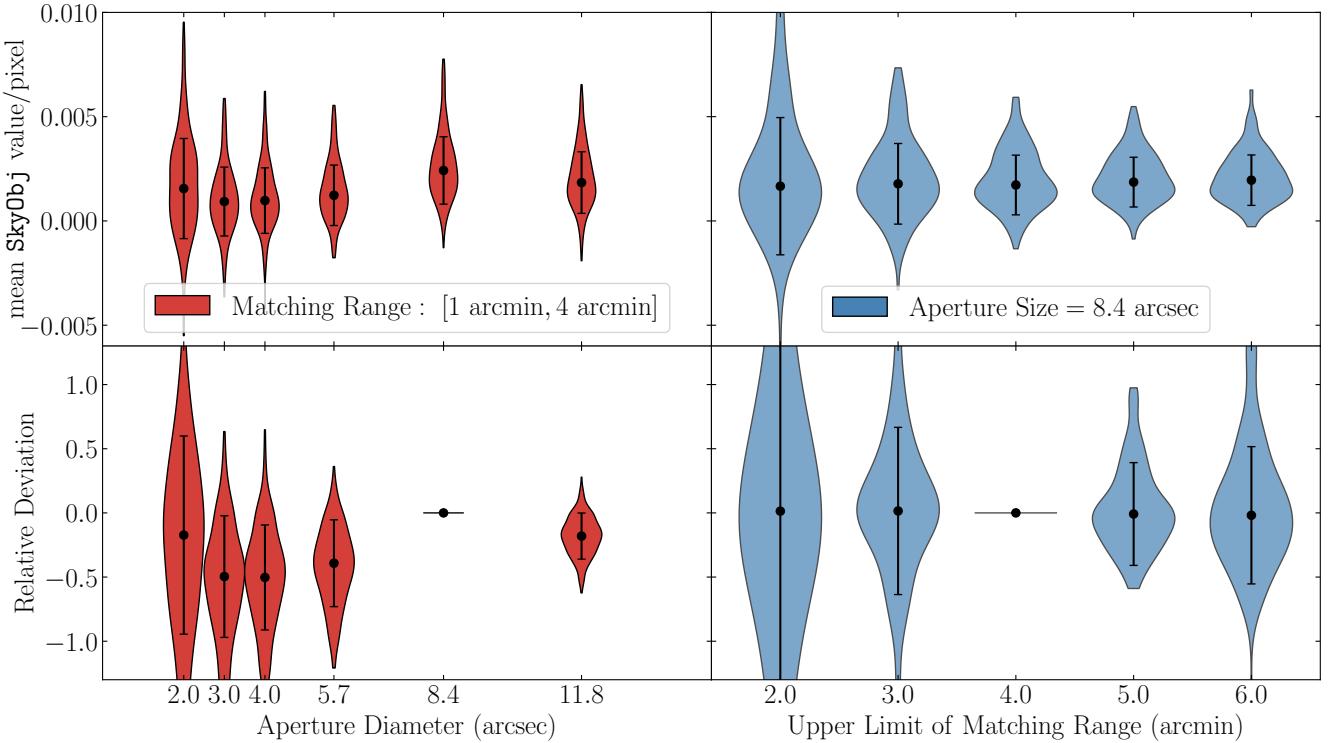
However, we use a simplistic model of massive galaxies in the mock tests that may not represent the real massive galaxies perfectly. In the future iterations, we will develop more realistic galaxy models to improve the robustness of the mock test. On the other hand, in real data, the sky subtraction will be affected by the injected mock galaxy. Therefore, a mock test shows only one aspect of sky subtraction, and tests on real data are needed to affirm the soundness of using `SkyObjs` to correct the sky residual. Figure 10 in Section 6.3 shows converged median profiles and symmetric distribution around zero intensity when the sky is corrected using an  $8.4''$  aperture `SkyObjs`. We also tested correcting the sky resid-

ual by using a  $5.7''$  aperture `SkyObjs` on the intermediate- $z$  sample, but we found a slight under-subtraction. This attests that  $8.4''$  aperture size is the most reasonable choice for intermediate- $z$  galaxies regardless of stellar mass, since the median profiles converge to zero smoothly without any sign of over- or under-subtraction.

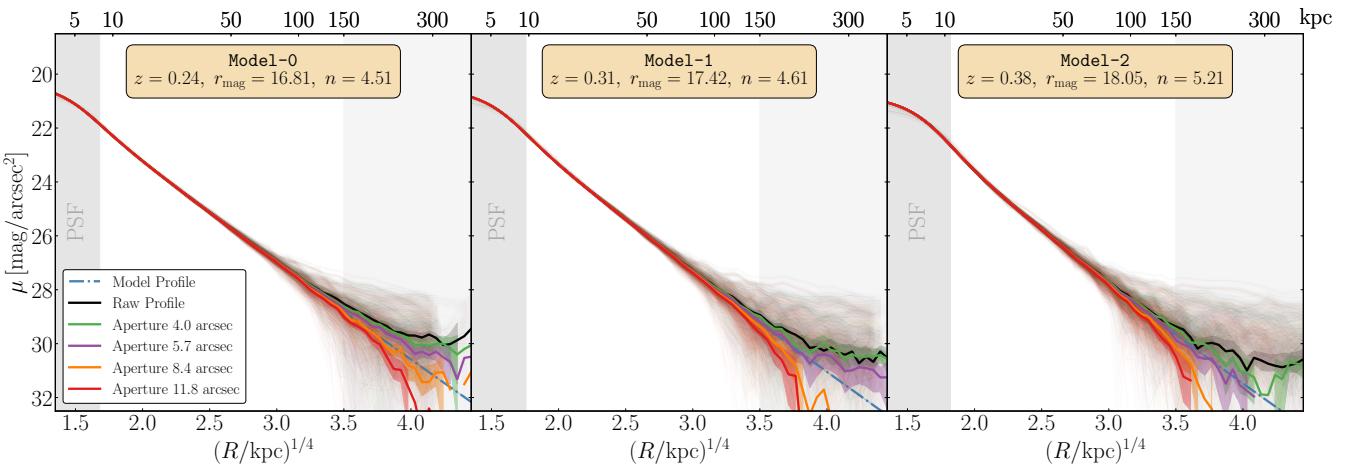
Although S18A avoids over-subtraction around intermediate- $z$  galaxies, the sky residual makes the footprints of bright objects too large, which slows down the deblender dramatically, causes over-deblending issues around them and sometimes crashes the pipeline. If we use a large bright object mask that is too large, we see a drastic decrease of galaxy sample size, damaging the study of cosmology. In the future data releases, `hscPipe` will detect objects on locally subtracted images and make catalogues, but co-added globally subtracted images will also be retained for future use.

## REFERENCES

- Abazajian K. N., et al., 2009, *ApJS*, **182**, 543  
 Abraham R. G., van Dokkum P. G., 2014, *PASP*, **126**, 55  
 Aihara H., et al., 2018, *Publications of the Astronomical Society of Japan*, **70**, S4  
 Aihara H., et al., 2019, arXiv e-prints, p. arXiv:1905.12221  
 Astropy Collaboration et al., 2013, *A&A*, **558**, A33  
 Barbary K., 2016, *Journal of Open Source Software*, **1**(6), 58  
 Bernardi M., Meert A., Sheth R. K., Vikram V., Huertas-Company M., Mei S., Shankar F., 2013, *MNRAS*, **436**, 697  
 Bernardi M., Fischer J. L., Sheth R. K., Meert A., Huertas-Company M., Shankar F., Vikram V., 2017, *MNRAS*, **468**, 2569  
 Bernstein R. A., 2007, *ApJ*, **666**, 663  
 Bertin E., 2006, in Gabriel C., Arviset C., Ponz D., Enrique S., eds, *Astronomical Society of the Pacific Conference Series Vol. 351, Astronomical Data Analysis Software and Systems XV*. p. 112  
 Bertin E., Arnouts S., 1996, *A&AS*, **117**, 393  
 Bertin E., Mellier Y., Radovich M., Missonnier G., Didelon P., Morin B., 2002, in Bohlender D. A., Durand D., Handley T. H., eds, *Astronomical Society of the Pacific Conference Series Vol. 281, Astronomical Data Analysis Software and Systems XI*. p. 228  
 Blanton M. R., Kazin E., Muna D., Weaver B. A., Price-Whelan A., 2011, *AJ*, **142**, 31  
 Bosch J., et al., 2018, *PASJ*, **70**, S5  
 Bradley L., et al., 2019, *astropy/photutils: v0.6*, doi:10.5281/zenodo.2533376, <https://doi.org/10.5281/zenodo.2533376>  
 Buitrago F., Trujillo I., Curtis-Lake E., Montes M., Cooper A. P., Bruce V. A., Pérez-González P. G., Cirasuolo M., 2017, *MNRAS*, **466**, 4888  
 Carlberg R. G., Yee H. K. C., Ellingson E., 1997, *ApJ*, **478**, 462  
 Carollo C. M., Danziger I. J., Buson L., 1993, *MNRAS*, **265**, 553  
 Cohen Y., et al., 2018, *ApJ*, **868**, 96  
 Cooper A. P., Gao L., Guo Q., Frenk C. S., Jenkins A., Springel V., White S. D. M., 2015, *MNRAS*, **451**, 2703  
 Cui W., et al., 2014, *MNRAS*, **437**, 816  
 D'Souza R., Kauffmann G., Wang J., Vegetti S., 2014, *MNRAS*, **443**, 1433  
 Danieli S., et al., 2019, arXiv e-prints, p. arXiv:1910.14045  
 Dawson K. S., et al., 2013, *AJ*, **145**, 10  
 DePoy D. L., et al., 2008, in *Ground-based and Air-borne Instrumentation for Astronomy II*. p. 70140E, doi:10.1117/12.789466



**Figure A1.** SkyObj analysis results. We test the statistical differences of choosing aperture size and matching range. This figure shows that the mean value of SkyObj around an object is less variable with the change of matching range, but changes significantly with aperture size. Hence, we use the mock test results to test which aperture size is more reasonable, as shown in Figure A2.



**Figure A2.** The median profiles of mock test profiles after we subtract the mean value of SkyObj with different aperture size. The matching range is fixed to  $1' < D < 4'$ . This figure clearly shows the behaviors and differences between methods and suggests that subtracting the mean value of SkyObj with  $8.4''$  aperture (for relatively low redshift) and  $5.7''$  aperture (for relatively high redshift) give us the best results.

- DeVore J. G., Kristl J. A., Rappaport S. A., 2013, *Journal of Geophysical Research (Atmospheres)*, **118**, 5679  
 Dey A., et al., 2019, *AJ*, **157**, 168  
 Driver S. P., et al., 2011, *MNRAS*, **413**, 971  
 Duc P.-A., et al., 2015, *MNRAS*, **446**, 120  
 Ene I., Ma C.-P., McConnell N. J., Walsh J. L., Kempki P., Greene J. E., Thomas J., Blakeslee J. P., 2019, *ApJ*, **878**, 57  
 Fischer J. L., Bernardi M., Meert A., 2017, *MNRAS*, **467**, 490  
 Gilhuly C., et al., 2019, arXiv e-prints, p. arXiv:1910.05358

- Gonzalez A. H., Zabludoff A. I., Zaritsky D., 2005, *ApJ*, **618**, 195  
 Greene J. E., Janish R., Ma C.-P., McConnell N. J., Blakeslee J. P., Thomas J., Murphy J. D., 2015, *ApJ*, **807**, 11  
 Greene J. E., et al., 2019, *ApJ*, **874**, 66  
 Gunn J. E., Stryker L. L., 1983, *ApJS*, **52**, 121  
 Gunn J. E., et al., 2006, *AJ*, **131**, 2332  
 Horton A., et al., 2016, in Proc. SPIE. p. 99041Q (arXiv:1606.06960), doi:10.1111/12.2232467  
 Huang S., Ho L. C., Peng C. Y., Li Z.-Y., Barth A. J., 2013, *ApJ*,

- 766, 47
- Huang S., et al., 2018a, arXiv e-prints, p. arXiv:1811.01139
- Huang S., et al., 2018b, *PASJ*, **70**, S6
- Huang S., Leauthaud A., Greene J. E., Bundy K., Lin Y.-T., Tanaka M., Miyazaki S., Komiyama Y., 2018c, *MNRAS*, **475**, 3348
- Hunter J. D., 2007, *Computing in Science Engineering*, **9**, 90
- Iodice E., et al., 2017, *ApJ*, **839**, 21
- Iodice E., et al., 2019, *A&A*, **623**, A1
- Jedrzejewski R. I., 1987, *Monthly Notices of the Royal Astronomical Society*, **226**, 747
- Jones E., Oliphant T., Peterson P., et al., 2001–, SciPy: Open source scientific tools for Python, <http://www.scipy.org/>
- Jurić M., et al., 2017, in Lorente N. P. F., Shortridge K., Wayth R., eds, *Astronomical Society of the Pacific Conference Series Vol. 512, Astronomical Data Analysis Software and Systems XXV*. p. 279 (arXiv:1512.07914)
- Kado-Fong E., et al., 2018, *ApJ*, **866**, 103
- Kawanomoto S., et al., 2018, *Publications of the Astronomical Society of Japan*, **70**
- Kelvin L. S., et al., 2012, *MNRAS*, **421**, 1007
- Kormendy J., Fisher D. B., Cornell M. E., Bender R., 2009, *ApJS*, **182**, 216
- Krick J. E., Bernstein R. A., 2007, *AJ*, **134**, 466
- La Barbera F., Ferreras I., de Carvalho R. R., Bruzual G., Charlot S., Pasquali A., Merlin E., 2012, *MNRAS*, **426**, 2300
- Lang D., Hogg D. W., Mykytyn D., 2016, The Tractor: Probabilistic astronomical source detection and measurement, *Astrostatistics Source Code Library* (ascl:1604.008)
- Lauer T. R., Postman M., Strauss M. A., Graves G. J., Chisari N. E., 2014, *ApJ*, **797**, 82
- Le Fèvre O., et al., 2005, *A&A*, **439**, 845
- Lin Y.-T., Mohr J. J., 2004, *ApJ*, **617**, 879
- Longbottom F., 2019, in *Linking Galaxies from the Epoch of Initial Star Formation to Today*. p. 82, doi:10.5281/zenodo.2635363
- Lupton R., Blanton M. R., Fekete G., Hogg D. W., O'Mullane W., Szalay A., Wherry N., 2004, *PASP*, **116**, 133
- Merritt A., van Dokkum P., Abraham R., Zhang J., 2016a, *ApJ*, **830**, 62
- Merritt A., van Dokkum P., Danieli S., Abraham R., Zhang J., Karachentsev I. D., Makarova L. N., 2016b, *ApJ*, **833**, 168
- Michael R., 2002, *A&A*, **384**, 763
- Mihos J. C., 2016, in Bragaglia A., Arnaboldi M., Rejkuba M., Romano D., eds, *IAU Symposium Vol. 317, The General Assembly of Galaxy Halos: Structure, Origin and Evolution*. pp. 27–34 (arXiv:1510.01929), doi:10.1017/S1743921315006857
- Mihos J. C., Harding P., Feldmeier J., Morrison H., 2005, *ApJ*, **631**, L41
- Mihos J. C., Harding P., Feldmeier J. J., Rudick C., Janowiecki S., Morrison H., Slater C., Watkins A., 2017, *ApJ*, **834**, 16
- Miyazaki S., et al., 2012, *Proc. SPIE*, **8446**, 8446
- Miyazaki S., et al., 2018, *PASJ*, **70**, S1
- Montes M., Trujillo I., 2019, *MNRAS*, **482**, 2838
- Oh S., Greene J. E., Lackner C. N., 2017, *ApJ*, **836**, 115
- Oke J. B., Gunn J. E., 1983, *ApJ*, **266**, 713
- Pillepich A., et al., 2018, *MNRAS*, **475**, 648
- Racine R., 1996, *PASP*, **108**, 699
- Román J., Trujillo I., Montes M., 2019, arXiv e-prints, p. arXiv:1907.00978
- Rowe B. T. P., et al., 2015, *Astronomy and Computing*, **10**, 121
- Rudick C. S., Mihos J. C., McBride C., 2006, *ApJ*, **648**, 936
- Rudick C. S., Mihos J. C., McBride C. K., 2011, *ApJ*, **732**, 48
- Sandin C., 2014, *A&A*, **567**, A97
- Sirianni M., et al., 1998, in D'Odorico S., ed., *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 3355, Proc. SPIE*. pp 608–612, doi:10.1117/12.316832
- Slater C. T., Harding P., Mihos J. C., 2009, *PASP*, **121**, 1267
- Smercina A., et al., 2019, arXiv e-prints, p. arXiv:1910.14672
- Springel V., et al., 2018, *MNRAS*, **475**, 676
- Tal T., van Dokkum P. G., 2011, *ApJ*, **731**, 89
- Tal T., van Dokkum P. G., Nelan J., Bezanson R., 2009, *AJ*, **138**, 1417
- Taylor E. N., et al., 2011, *MNRAS*, **418**, 1587
- Valls-Gabaud D., MESSIER Collaboration 2017, in Gil de Paz A., Knapen J. H., Lee J. C., eds, *IAU Symposium Vol. 321, Formation and Evolution of Galaxy Outskirts*. pp 199–201, doi:10.1017/S1743921316011388
- Vogelsberger M., et al., 2014a, *MNRAS*, **444**, 1518
- Vogelsberger M., et al., 2014b, *Nature*, **509**, 177
- Wang W., et al., 2019, *MNRAS*, **487**, 1580
- Willmer C. N. A., 2018, *ApJS*, **236**, 47
- Wu H., Shao Z., Mo H. J., Xia X., Deng Z., 2005, *ApJ*, **622**, 244
- Zhang J., Abraham R., van Dokkum P., Merritt A., Janssens S., 2018, *ApJ*, **855**, 78
- Zhang Y., et al., 2019, *ApJ*, **874**, 165
- Zibetti S., White S. D. M., Schneider D. P., Brinkmann J., 2005, *MNRAS*, **358**, 949
- de Jong R. S., 2008, *MNRAS*, **388**, 1521
- van Dokkum P. G., Conroy C., 2012, *ApJ*, **760**, 70
- van Dokkum P. G., Abraham R., Merritt A., 2014, *ApJ*, **782**, L24
- van Dokkum P., Lokhorst D., Danieli S., Li J., Merritt A., Abraham R., Gilhuly C., Greco J. P., 2019a, arXiv e-prints, p. arXiv:1910.12867
- van Dokkum P., et al., 2019b, *ApJ*, **883**, L32
- van der Walt S., Colbert S. C., Varoquaux G., 2011, *Computing in Science Engineering*, **13**, 22