



# Radii of young star clusters in nearby galaxies

Gillen Brown and Oleg Y. Gnedin

Department of Astronomy, University of Michigan, Ann Arbor, MI 48109, USA

Accepted 2021 October 1. Received 2021 September 16; in original form 2021 June 23

## ABSTRACT

We measure the projected half-light radii of young star clusters in 31 galaxies from the Legacy Extragalactic UV Survey (LEGUS). We implement a custom pipeline specifically designed to be robust against contamination, which allows us to measure radii for 6097 clusters. This is the largest sample of young star cluster radii currently available. We find that most (but not all) galaxies share a common cluster radius distribution, with the peak at around 3 pc. We find a clear mass–radius relation of the form  $R_{\text{eff}} \propto M^{0.24}$ . This relation is present at all cluster ages younger than 1 Gyr, but with a shallower slope for clusters younger than 10 Myr. We present simple toy models to interpret these age trends, finding that high-mass clusters are more likely to be not tidally limited and expand. We also find that most clusters in LEGUS are gravitationally bound, especially at older ages or higher masses. Lastly, we present the cluster density and surface density distributions, finding a large scatter that appears to decrease with cluster age. The youngest clusters have a typical surface density of  $100 \text{ M}_\odot \text{ pc}^{-2}$ .

**Key words:** galaxies: star clusters: general – galaxies: star formation.

## 1 INTRODUCTION

Young star clusters bridge the small scales of star formation and the large scales of galaxy formation. They are easily detected in nearby star-forming galaxies and contain the majority of massive stars (Krumholz, McKee & Bland-Hawthorn 2019). Their almost universal luminosity function, and corresponding mass function (Adamo et al. 2020), is used both as a test of molecular cloud collapse models (Ballesteros-Paredes et al. 2020) and as a constraint on the star formation modelling in cosmological simulations (Li, Gnedin & Gnedin 2018). However, our understanding of the origin and long-term evolution of star clusters is still hindered by lack of reliable measurements of their density distribution.

Depending on the initial density, young massive clusters may evolve into globular clusters or may dissolve into the smooth stellar field (Portegies Zwart, McMillan & Gieles 2010). As clusters form in giant molecular clouds (GMCs), feedback from massive stars ejects the residual gas, making the cluster expand to re-establish virial equilibrium (Goodwin & Bastian 2006). Additionally, tidal shocks from encounters with other GMCs or spiral arms kinematically heat the cluster and lead to its disruption (Spitzer 1958). The density of young clusters determines whether they can survive this intense phase as bound stellar systems.

To calculate the density of young star clusters, one needs to measure their radii. Currently published measurements are limited to only a few galaxy samples. Kharchenko et al. (2013) measured the radii for  $\sim 1100$  open clusters (OCs) in the Milky Way, while samples from external galaxies include Bastian et al. (2012) and Ryon et al. (2015), who measured radii for several hundred clusters in M83, about 3800 clusters in M51 measured by Chandar et al. (2016), and 514 clusters in M31 measured by the PHAT survey (Johnson et al.

2012; Fouesneau et al. 2014). Of particular note for this paper, Ryon et al. (2017) (hereafter R17) measured the radii of several hundred clusters spread between two galaxies in the Legacy Extragalactic UV Survey (LEGUS).

In this paper, we measure the projected half-light radius (effective radius) of clusters in the 31 galaxies with publicly available cluster catalogues from LEGUS. Our method for fitting the radii is described in Section 2. In Section 3, we describe our findings of a cluster radius distribution common to most galaxies and a clear cluster mass–radius relation. In Section 4, we discuss how selection effects affect our results, calculate cluster densities, and present a toy model of cluster evolution. We close with a summary in Section 5.

## 2 METHODS

### 2.1 The LEGUS sample

We use the publicly available LEGUS data set<sup>1</sup> to extend the sample of clusters with uniformly measured radii and densities to the 31 galaxies with currently available cluster catalogues. We summarize some of the key details of LEGUS in this section (see Calzetti et al. 2015, for more on the LEGUS survey description, Adamo et al. 2017 and Cook et al. 2019 for more on the cluster catalogues).

LEGUS is a Cycle 21 Treasury programme on *HST* that collected imaging with WFC3/UVIS to supplement archival ACS/WFC imaging, producing five-band coverage from the near-UV to the *I* band for 50 galaxies. Within each field, a uniform process was used to identify cluster candidates.

First, SourceExtractor (Bertin & Arnouts 1996) is used to find sources in the white-light images (a combination of the images in all five filters, weighted by S/N). Next, a user selects a training set of

\* E-mail: [gillenb@umich.edu](mailto:gillenb@umich.edu)

<sup>1</sup><https://archive.stsci.edu/prepds/legus/dataproducts-public.html>

objects that are clearly clusters or stars. The pipeline calculates the concentration index (CI), which is the magnitude difference between apertures of radius 1 pixel and 3 pixels (Holtzman et al. 1992; Whitmore et al. 2010). The user then selects a value of the CI that separates stars from clusters. The aperture for science photometry is chosen as the integer number of pixels containing at least 50 per cent of the cluster flux, based on the curve of growth of the clusters. This science photometry is done with the same aperture in all filters, using a sky annulus located at 7 pixels with a width of 1 pixel.

Next, aperture corrections for each filter are determined using an averaged method and a CI-based method. In this study, we use catalogues using the averaged aperture correction method, so we describe that here. The correction is the difference between the magnitude of the source obtained using an aperture of 20 pixels and the magnitude obtained from the science aperture. The average aperture correction for a user-defined set of well-behaved clusters is used for all clusters. This photometry is corrected for galactic foreground extinction (Schlafly & Finkbeiner 2011). Sources that are detected in at least four filters with a photometric error of less than 0.3 mag, have an absolute  $V$ -band magnitude brighter than  $-6$ , and have a larger CI than the limit determined earlier, are then visually inspected by the LEGUS team. Sources that do not pass some of these cuts can be manually added by LEGUS team members, but the number is small.

Three or more team members visually inspect each cluster candidate, classifying it into one of the following four classes. Class 1 objects are compact and centrally concentrated with a homogeneous colour. Class 2 clusters have slightly elongated density profiles and a less symmetric light distribution. Class 3 clusters are likely compact associations, having asymmetric profiles or multiple peaks on top of diffuse underlying wings. Class 4 objects are stars or artefacts.

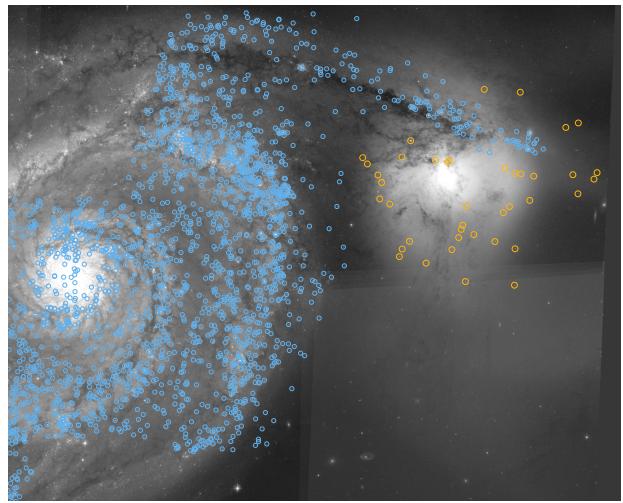
The age, mass, and extinction of each cluster are determined by the spectral energy distribution (SED) fitting code YGGDRASIL (Zackrisson et al. 2011). Versions of the catalogues are created with different stellar tracks and extinction laws, but in this paper we select the catalogues that use the MW extinction law (Cardelli, Clayton & Mathis 1989) and Padova-AGB tracks available in STARBURST99 (Leitherer et al. 1999; Vázquez & Leitherer 2005).

The survey targeted 50 galaxies. Currently, public cluster catalogues are available for 31 of these galaxies. Some galaxies are observed with multiple fields, resulting in 34 total fields with cluster catalogues. Table 1 lists these fields along with some key properties of the galaxies.

## 2.2 Outline of the measurement procedure

To fit the cluster radii, we implement a custom pipeline written in PYTHON. We choose to implement our own method to have full control over the fitting process and better quantify the distribution of errors of the fit parameters. It is analogous to that in the popular package GALFIT (Peng et al. 2002, 2010), but with automated masking and several other features described in Section 2.4 that make it more robust against contamination from other nearby sources and ensure a good estimate of the local background. We assume an E $\mathrm{FF}$  density profile for young clusters (Elson, Fall & Freeman 1987; Larsen 1999; McLaughlin & van der Marel 2005), and then convolve it with the empirically measured point spread function (PSF) before comparing to the data.

We use the F555W images in 25 of the 34 fields, but for the other 9 fields that were not observed in F555W, we instead use F606W. All images have a pixel scale of  $39.62 \text{ mas pixel}^{-1}$ . The LEGUS mosaics have units of  $\text{e}^{-\text{s}^{-1}}$ , and we multiply by the exposure time



**Figure 1.** A portion of the F555W image of the LEGUS field with NGC 5194 and NGC 5195, illustrating our selection of cluster belonging to NGC 5194 (blue circles) and NGC 5195 (orange circles).

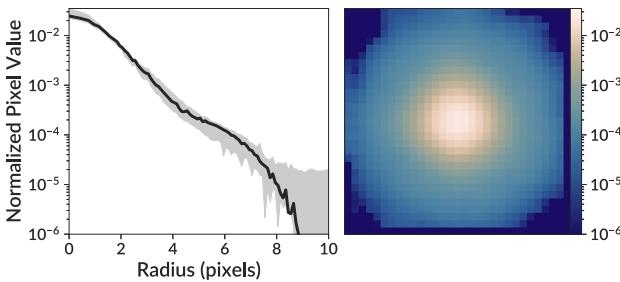
to convert to the electron count. We use the recommended LEGUS cluster catalogues that adopt the MW extinction, Padova stellar evolutionary tracks, and the averaged aperture correction method (Adamo et al. 2017). The LEGUS team created visual classification tags by visually inspecting each cluster with multiple team members. We select clusters that were identified as being concentrated and either symmetric (class 1) or with some degree of asymmetry (class 2). We use the mode of the classifications from multiple team members. Additionally, machine learning classifications are available for several galaxies (Grasha 2018; Grasha et al. 2019). For NGC 5194 and NGC 5195, we use the human classifications for clusters where those are available, and supplement with machine learning classifications for clusters not inspected by humans. In NGC 1566, we use the hybrid classification system created by the LEGUS team, where some clusters are inspected by humans only, some by machine learning only, and some with a machine learning classification verified by humans. We do not use the machine learning classifications for NGC 4449, as we find that the machine-classified clusters have significantly different CI and radius distributions from the rest of the sample. This selection produces a final sample of 7242 clusters. As NGC 5194 and NGC 5195 are in the same LEGUS field, we manually identify their host galaxy. Fig. 1 shows this selection.

## 2.3 PSF creation

Like GALFIT, our method convolves the PSF with the model image, and compares the result with the observed data. To produce this PSF, we use Photutils (Bradley et al. 2019), an Astropy package (Astropy Collaboration 2013, 2018). We manually select bright isolated stars in each field, and then use the EPSFBuilder class of Photutils to create a separate PSF for each field. EPSFBuilder follows the prescription of Anderson & King (2000). The final PSF images are 15 pixel (0.59 arcsec) wide, and we spatially subsample the PSF by a factor of 2, producing a PSF with twice the spatial resolution of the input image. We do not choose higher values, as this significantly increases the computational cost of the fitting procedure (particularly the convolution). As shown in Section 3.1, our results are consistent with those of R17 for NGC 1313 and NGC 628, even for the smallest clusters, indicating that this oversampling factor

**Table 1.** List of the LEGUS fields included in the sample, with some key properties of their cluster populations. We use the TRGB-based distances from Sabbi et al. (2018) except for NGC 1566 – see the discussion at the end of Section 2.4. Stellar masses and specific star formation rates are from Calzetti et al. (2015), who obtained SFR from *GALEX* far-UV corrected for dust attenuation, as described in Lee et al. (2009), and stellar mass from extinction-corrected *B*-band luminosity and colour information, as described in Bothwell, Kennicutt & Lee (2009) and using the mass-to-light ratio models of Bell & de Jong (2001). Note that NGC 5194 and NGC 5195 are in the same field, so they share a distance and PSF, but we split them for clarity.

| LEGUS field | <i>N</i> | $\log M_*$ ( $M_\odot$ ) | $\log \text{sSFR}$ ( $\text{yr}^{-1}$ ) | Distance (Mpc)  | PSF size (pc) | Cluster $R_{\text{eff}}$ : 25–50–75th percentiles |
|-------------|----------|--------------------------|---|-----------------|---------------|---|
| IC 4247     | 5        | 8.08                     | -10.18                                  | $5.11 \pm 0.4$  | 1.48          | 2.30 – 3.52 – 4.36                                |
| IC 559      | 21       | 8.15                     | -10.45                                  | $10.0 \pm 0.9$  | 2.74          | 4.05 – 4.85 – 6.34                                |
| NGC 1313-e  | 137      | 9.41                     | -9.35                                   | $4.30 \pm 0.24$ | 1.26          | 1.41 – 2.35 – 3.00                                |
| NGC 1313-w  | 276      | 9.41                     | -9.35                                   | $4.30 \pm 0.24$ | 1.26          | 1.13 – 2.54 – 3.58                                |
| NGC 1433    | 112      | 10.23                    | -10.80                                  | $9.1 \pm 1.0$   | 2.53          | 1.11 – 1.79 – 3.20                                |
| NGC 1566    | 881      | 10.43                    | -9.68                                   | $15.6 \pm 0.6$  | 4.33          | 3.00 – 4.30 – 6.28                                |
| NGC 1705    | 29       | 8.11                     | -9.07                                   | $5.22 \pm 0.38$ | 1.35          | 2.73 – 3.26 – 4.20                                |
| NGC 3344    | 237      | 9.70                     | -9.76                                   | $8.3 \pm 0.7$   | 2.37          | 1.72 – 2.40 – 3.53                                |
| NGC 3351    | 19       | 10.32                    | -10.13                                  | $9.3 \pm 0.9$   | 2.75          | 0.88 – 2.55 – 5.17                                |
| NGC 3738    | 142      | 8.38                     | -9.54                                   | $5.09 \pm 0.40$ | 1.52          | 2.17 – 3.25 – 4.30                                |
| NGC 4242    | 12       | 9.04                     | -10.04                                  | $5.3 \pm 0.3$   | 1.54          | 1.28 – 2.47 – 3.11                                |
| NGC 4395-n  | 20       | 8.78                     | -9.25                                   | $4.54 \pm 0.18$ | 1.24          | 1.34 – 1.84 – 2.42                                |
| NGC 4395-s  | 95       | 8.78                     | -9.25                                   | $4.54 \pm 0.18$ | 1.33          | 0.49 – 0.82 – 1.76                                |
| NGC 4449    | 425      | 9.04                     | -9.07                                   | $4.01 \pm 0.30$ | 1.17          | 1.68 – 2.40 – 3.32                                |
| NGC 45      | 14       | 9.52                     | -9.97                                   | $6.8 \pm 0.5$   | 1.84          | 2.41 – 3.25 – 4.61                                |
| NGC 4656    | 184      | 8.60                     | -8.90                                   | $7.9 \pm 0.7$   | 2.11          | 2.35 – 3.32 – 4.15                                |
| NGC 5194    | 2921     | 10.38                    | -9.54                                   | $7.40 \pm 0.42$ | 2.16          | 1.29 – 2.17 – 3.29                                |
| NGC 5195    | 40       | 10.36                    | -10.82                                  | $7.40 \pm 0.42$ | 2.16          | 2.69 – 3.39 – 4.88                                |
| NGC 5238    | 8        | 8.15                     | -10.15                                  | $4.43 \pm 0.34$ | 1.29          | 2.29 – 2.63 – 2.94                                |
| NGC 5253    | 57       | 8.34                     | -9.34                                   | $3.32 \pm 0.25$ | 0.98          | 1.09 – 1.86 – 2.66                                |
| NGC 5474    | 143      | 8.91                     | -9.48                                   | $6.6 \pm 0.5$   | 1.95          | 2.28 – 3.25 – 4.13                                |
| NGC 5477    | 14       | 7.60                     | -9.12                                   | $6.7 \pm 0.5$   | 1.98          | 2.40 – 3.64 – 4.78                                |
| NGC 628-c   | 691      | 10.04                    | -9.48                                   | $8.8 \pm 0.7$   | 2.59          | 1.70 – 2.43 – 3.44                                |
| NGC 628-e   | 172      | 10.04                    | -9.48                                   | $8.8 \pm 0.7$   | 2.49          | 1.95 – 2.64 – 4.05                                |
| NGC 6503    | 167      | 9.28                     | -9.77                                   | $6.3 \pm 0.5$   | 1.69          | 1.31 – 2.12 – 3.34                                |
| NGC 7793-e  | 108      | 9.51                     | -9.79                                   | $3.79 \pm 0.20$ | 0.92          | 0.60 – 1.02 – 2.20                                |
| NGC 7793-w  | 135      | 9.51                     | -9.79                                   | $3.79 \pm 0.20$ | 1.11          | 0.66 – 1.26 – 2.20                                |
| UGC 1249    | 48       | 8.74                     | -9.56                                   | $6.4 \pm 0.5$   | 1.88          | 1.83 – 2.58 – 3.26                                |
| UGC 4305    | 45       | 8.36                     | -9.28                                   | $3.32 \pm 0.25$ | 0.97          | 0.69 – 1.15 – 1.91                                |
| UGC 4459    | 7        | 6.83                     | -8.99                                   | $3.96 \pm 0.30$ | 1.23          | 1.83 – 2.51 – 3.13                                |
| UGC 5139    | 9        | 7.40                     | -9.10                                   | $3.83 \pm 0.29$ | 1.12          | 1.07 – 1.54 – 4.00                                |
| UGC 685     | 11       | 7.98                     | -10.13                                  | $4.37 \pm 0.34$ | 1.31          | 1.77 – 2.23 – 2.96                                |
| UGC 695     | 11       | 8.26                     | -9.95                                   | $7.8 \pm 0.6$   | 2.08          | 1.39 – 6.21 – 7.49                                |
| UGC 7408    | 35       | 7.67                     | -9.67                                   | $7.0 \pm 0.5$   | 2.08          | 3.75 – 4.41 – 5.68                                |
| UGCA 281    | 11       | 7.28                     | -8.98                                   | $5.19 \pm 0.39$ | 1.51          | 3.12 – 3.50 – 4.25                                |
| Total       | 7242     | –                        | –                                       | –               | –             | 1.53 – 2.48 – 3.69                                |



**Figure 2.** Visualization of one of our PSFs, from the NGC 1313-e field. The solid line in the left-hand panel shows the azimuthally averaged radial profile of this PSF, while the shaded region shows the range of the PSF profiles from all other fields, normalized to integrate to unity. The right-hand panel shows an image of the NGC 1313-e PSF.

is adequate. We use *Photutils*’s ‘quadratic’ smoothing kernel, which is the polynomial fit with degree =2 to  $5 \times 5$  array of zeros

with 1 at the centre. We found that the other options gave unphysically non-smooth PSFs. Fig. 2 illustrates our created PSFs.

#### 2.4 Fitting cluster parameters

We fit the clusters with the EFF surface brightness profile, as it accurately describes the light profiles of young star clusters (Elson et al. 1987; Larsen 1999; McLaughlin & van der Marel 2005; Bastian et al. 2012; Ryon et al. 2015, 2017; Cuevas-Otahola et al. 2020). Assuming circular symmetry, it takes the form

$$\mu(r) = \mu_0 \left( 1 + \frac{r^2}{a^2} \right)^{-\eta}, \quad (1)$$

where  $\mu$  is the surface brightness,  $a$  is the scale radius, and  $\eta$  is the power-law slope. As real clusters are typically not circularly symmetric, we include ellipticity as follows:

$$\mu(x, y) = \mu_0 \left( 1 + \left[ \frac{x'(x, y)}{a} \right]^2 + \left[ \frac{y'(x, y)}{aq} \right]^2 \right)^{-\eta}, \quad (2)$$

where  $q$  is the ratio of the minor to major axes of the ellipse ( $0 < q \leq 1$ ). We have rotated the image coordinate system by angle  $\theta$  about the cluster centre  $(x_c, y_c)$  to new coordinates  $(x', y')$  as follows:

$$x'(x, y) = (x - x_c) \cos \theta + (y - y_c) \sin \theta \quad (3)$$

$$y'(x, y) = -(x - x_c) \sin \theta + (y - y_c) \cos \theta. \quad (4)$$

Here,  $x'$  is aligned with the cluster major axis, while  $y'$  is aligned with the minor axis. This gives seven cluster parameters:  $\mu_0, x_c, y_c, a, q, \theta$ , and  $\eta$ . We also leave the local background  $f_{\text{BG}}$  as a free parameter, giving eight total parameters to fit.

We perform this fit on a  $30 \times 30$  pixel snapshot centred on the cluster, following R17. We tested larger snapshot sizes (40 and 50 pixels) but found no significant differences in fitted cluster radii, even for the biggest clusters where a larger snapshot could potentially allow for a better determination of the local background. These larger snapshot sizes also included more contaminating sources, leading to more catastrophic fit failures. The 30 pixel snapshot minimizes these failures while still performing well on the largest clusters.

To account for contaminating sources inside this  $30 \times 30$  pixel snapshot, we mask star-like sources identified by the `IRAFStarFinder` class of `Photutils`. Any pixels within  $2 \times \text{FWHM}$  of the stars are masked. However, we discard any stars whose masked region would extend within 3 pixels of the cluster centre, as well as any sources with a peak pixel value less than 2 times the local sky background identified by `IRAFStarFinder`. This second criterion was added to stop the masking of substructure in the most extended clusters. We also mask any pixels that are within 6 pixels of another star cluster, in cases where two star clusters are close to each other.

However, an automated masking system cannot solve all issues with contamination. To make our fitting method robust to potential contamination, we have made substantial modifications compared to a GALFIT-like method.

Our best-fitting parameters maximize the posterior distribution, defined as

$$\log P_{\text{posterior}} = -\frac{1}{2} \sum_{x, y} w(x, y) \left| \frac{f_d(x, y) - f_m(x, y)}{\sigma(x, y)} \right| + \log P_{\text{prior}}, \quad (5)$$

where  $x$  and  $y$  are pixel coordinates,  $w$  are pixel weights,  $f_d$  is the data value at pixel  $(x, y)$ ,  $f_m$  is the model,  $\sigma(x, y)$  is the pixel uncertainty, and  $P_{\text{prior}}$  is the prior distribution. We will expand on each of these components in turn. Note that we use the absolute value of the differences between the model and data rather than the more typical square. As the square weights large differences more heavily, it has the effect of increasing the attention the fit pays to unmasked contamination, as these pixels have large deviations. Using the absolute value instead produces fits that are less affected by contamination.

In addition, pixel weights are used to reduce the effect of contamination, particularly at large distances from the cluster. We weight each pixel proportional to  $1/r$  so that each annulus has the same weight. To avoid giving dramatically more weight to the most central pixels, all pixels within 3 pixels from the centre receive the same weight. We use the distance from the centre of the cluster to determine the radius, giving  $r^2 = (x - x_c)^2 + (y - y_c)^2$  and

$$w(r) = \begin{cases} 1 & \text{if } r \leq 3 \\ 3/r & \text{if } r > 3. \end{cases} \quad (6)$$

Giving equal weight to each annulus stops the large number of pixel values at large radius from dominating the fit, effectively increasing the focus on the cluster at the centre.

We break the pixel uncertainty  $\sigma(x, y)$  into two components: image-wide sky noise plus Poisson noise from individual sources:

$$\sigma^2(x, y) = \sigma_{\text{sky}}^2 + f_d(x, y), \quad (7)$$

where  $f_d(x, y)$  is the pixel value in electrons, and equals the Poisson variance. To calculate the global sky noise, we use the  $3\sigma$  clipped standard deviation of the pixel values of the entire image.

The model component  $f_m(x, y)$  is the convolution of the underlying cluster model with the PSF plus the local background  $f_{\text{BG}}$ , which we assume to be constant over the fitting region. We subsample the pixels for both the model and the empirical PSF, and then rebin the resulting model image to the same scale as the data:

$$f_m(x, y) = \sum_{x_s \in x} \sum_{y_s \in y} (\text{PSF} * \mu)(x_s, y_s) + f_{\text{BG}}, \quad (8)$$

where  $x_s$  and  $y_s$  represent subpixel positions that are not integer pixel values like  $x$  and  $y$ ,  $*$  represents convolution, and  $\mu$  is the functional form of the fitted profile given by equation (2).

The last component of equation (5) is the prior distribution. We employ a prior on the local background. This is needed because at very low values of  $\eta$  (shallow power-law slopes), the background can be incorrectly fit by this cluster component rather than a truly flat background. This attributes light to the cluster that should be attributed to the background, incorrectly inflating the enclosed light and therefore  $R_{\text{eff}}$ . Additionally,  $\eta$  is strongly degenerate with  $a$ , so as to give the same value of  $\mu$  at some typical radius. Low values of  $\eta$  caused by incorrect background fits also lead to unphysically small values for  $a$ . We find that constraining the background addresses these issues. We first estimate the background and background uncertainty by using sigma clipping to calculate the mean ( $\mu_{\text{BG}}$ ) and standard deviation ( $\sigma_{\text{BG}}$ ) of all pixels farther than 6 pixels from the cluster centre. The mean is used as the mean value of a Gaussian prior on the background. As the background becomes an issue for low values of  $\eta$ , we condition the width of our prior on it. We use a logistic function that produces a tight prior  $\sigma_{\text{prior}} = 0.1\sigma_{\text{BG}}$  for low values of  $\eta$ , while giving a looser constraint  $\sigma_{\text{prior}} = \sigma_{\text{BG}}$  for higher values of  $\eta$ . This takes the form

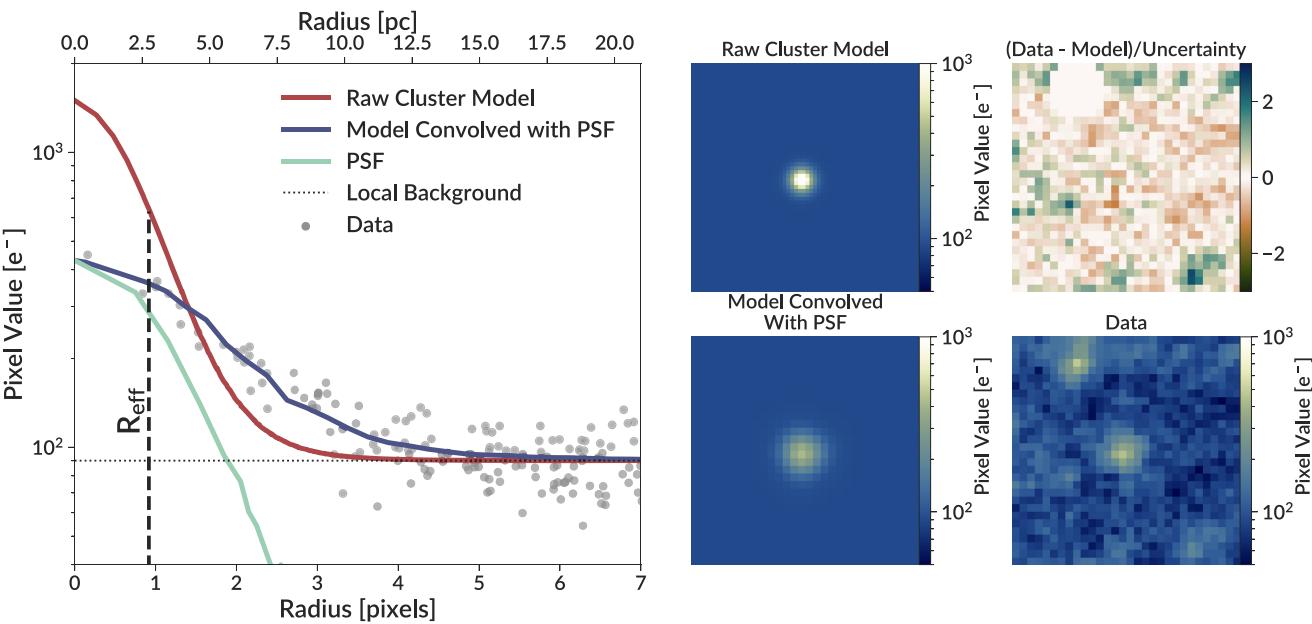
$$\log P_{\text{prior}} = -\frac{1}{2} \left( \frac{f_{\text{BG}} - \mu_{\text{BG}}}{\sigma_{\text{prior}}} \right)^2, \quad (9)$$

where

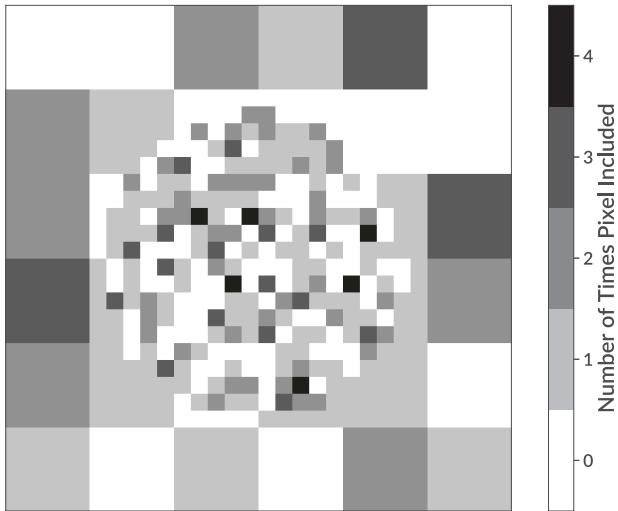
$$\sigma_{\text{prior}}(\eta) = \sigma_{\text{BG}} \left( 0.1 + \frac{0.9}{1 + e^{10(1-\eta)}} \right). \quad (10)$$

The combined effect of using the absolute value of differences, the pixel weights, and the prior on the background is to produce cluster fits that more closely match the cluster itself. They prevent the fit from being drawn towards any surrounding structures, while also allowing the local background to be fitted appropriately. This produces trustworthy cluster parameter values. Lastly, to ensure full numerical convergence of the fit, we use multiple sets of initial values for the fit parameters, selecting the final parameter set corresponding to the highest likelihood. An example of a cluster fitted with our method is shown in Fig. 3.

To determine the distribution of errors of the fit parameters, we perform bootstrapping on the pixels in the snapshot. We split the snapshot into two regions: pixels closer than 9 pixels from the cluster centre and pixels outside this region. For the inner region, we resample individual pixels with replacement. For the outer region, we group the pixels into  $5 \times 5$  blocks, and then resample those blocks with replacement. Using blocks in the outskirts does a better job accounting for any missed or faint contaminating sources. If we



**Figure 3.** An example of our fitting process. The right set of panels shows (in counterclockwise order, starting from the top left) the 2D pixel values for the raw cluster model, the cluster model after being convolved with the PSF, the data, and then the residuals after subtracting the PSF-convolved model from the data. Note the masking of the contaminating object in the top right panel. The left-hand panel shows the radial profiles for these components (with the PSF normalization adjusted accordingly). Note that the radial profile is included for illustrative purposes only; the fitting is done using the 2D images.



**Figure 4.** An example of the pixels included in one randomly selected bootstrap iteration. The central region is resampled on a pixel-by-pixel basis, while the outskirts are resampled on  $5 \times 5$  pixel blocks.

were to use individual pixels, at least some of the pixels from these sources would be included in a given resampling, while using blocks allows us to exclude these sources completely in certain iterations, giving a better estimate of how these sources affect the cluster fit. Using individual pixels in the centre is necessary as the cluster itself may be roughly the size of the  $5 \times 5$  chunk. Fig. 4 shows an example of the pixels included in one randomly selected bootstrap realization.

We run the bootstrap realizations until convergence of the fit parameters. Every 20 iterations, we calculate the standard deviation of the distributions of all eight fit parameters in the accumulated iterations, and then compare them to the standard deviations from

the last time it was calculated. We stop bootstrapping when the standard deviation of each parameter changes by less than 10 percent. Most clusters required 100–140 iterations to converge. Our reported uncertainties on  $R_{\text{eff}}$  are marginalized over all other parameters. We use  $R_{\text{eff}}$  calculated using the original snapshot as the best-fitting value, and then take the 16–84th percentile range of  $R_{\text{eff}}$  from the bootstrap iterations as the uncertainty.

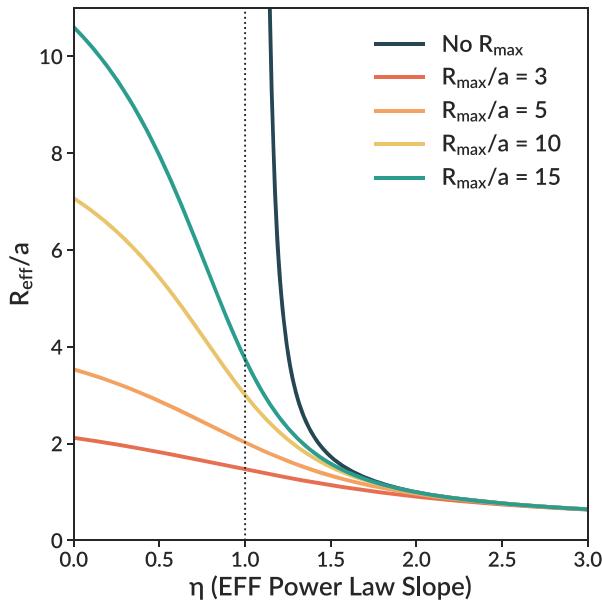
As we measure everything in pixel values, we need to convert to physical length units. To do this, we use the TRGB distances to all LEGUS galaxies provided by Sabbi et al. (2018). That work provides independent estimates of distance for each field. For galaxies split between two fields, we use the mean of the two distance estimates for both fields. Lastly, NGC 1566 has an unreliable distance estimate. The TRGB was too faint to be detected in Sabbi et al. (2018), and available values in the literature span a wide range (from 6 to 20 Mpc). However, NGC 1566 was identified as being part of the group centred on NGC 1553, which has a measured distance and group radius (Tully, Courtois & Sorce 2016; Kourkchi & Tully 2017). For NGC 1566, we adopt the distance to NGC 1553 with uncertainty of the group radius.

## 2.5 Converting to effective radius

The effective radius is defined to be the circular radius that contains half of the projected light of the cluster profile. For a circularly symmetric EFF profile, this is

$$R_{\text{eff}} = a \sqrt{2^{\frac{1}{\eta-1}} - 1}. \quad (11)$$

However, this equation asymptotically approaches infinity as  $\eta$  approaches unity, and the total light of the profile is infinite for  $\eta \leq 1$ . As some cluster fits prefer values of  $\eta$  near or below unity, we implement a maximum radius for the cluster profiles, removing this infinity and allowing the effective radius to be well defined for any value of  $\eta$ . We choose the size of our box (15 pixel radius) as  $R_{\text{max}}$ .



**Figure 5.** Comparison of the effective radius when calculated with (equation 12) or without (equation 11) a maximum radius. We show this for several representative values of  $R_{\max}/a$ , as clusters have values across this full range.

When using a maximum radius, the effective radius for the circularly symmetric EFF profile is

$$2 \left[ 1 + \left( \frac{R_{\text{eff}}}{a} \right)^2 \right]^{1-\eta} = 1 + \left[ 1 + \left( \frac{R_{\max}}{a} \right)^2 \right]^{1-\eta}. \quad (12)$$

For  $\eta \gtrsim 1.5$ , this agrees well with equation (11), as shown in Fig. 5.

A correction is required for an elliptical profile. We empirically determine this correction for the EFF profile as a function of  $\eta$  and  $q$ , by performing numerical integration of elliptical EFF profiles. We determine the circular aperture that contains half of the total light. We use a circular maximum radius  $R_{\max}/a = 10$  for this integration, although we find that the results do not depend strongly on the chosen  $R_{\max}$ . At a given  $\eta$ , we find that the relation between true effective radius and the effective radius calculated assuming a circularly symmetric profile (equation 12) is linear with  $q$ , so we parametrize the correction as

$$\frac{R_{\text{eff,true}}}{R_{\text{eff,circ}}} = 1 + m(q - 1), \quad (13)$$

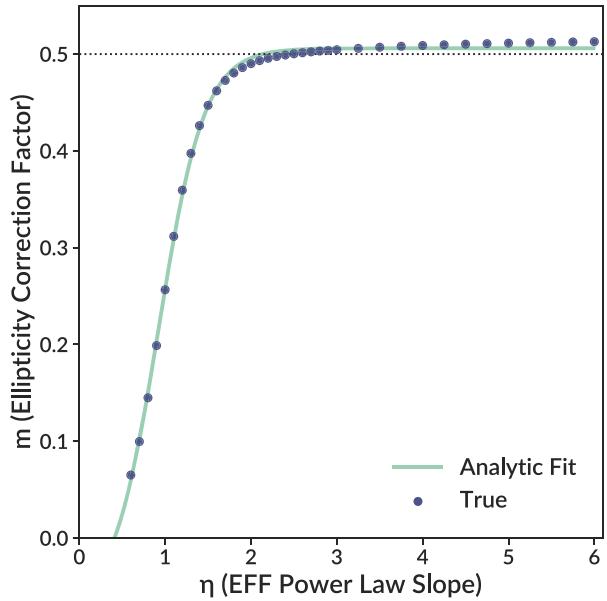
where  $m$  is a function of  $\eta$ . With  $m = 0.5$ , we obtain the commonly used correction

$$\frac{R_{\text{eff,true}}}{R_{\text{eff,circ}}} = 0.5(1 + q) \quad (14)$$

as found in the ISHAPE manual (Larsen 1999). We measure  $m$  as a function of  $\eta$ , and find it is well fitted by a logistic-type function, with an RMS deviation of only 0.0043. This gives a final correction of the form

$$\frac{R_{\text{eff,true}}}{R_{\text{eff,circ}}} = 1 + \left( \frac{0.579}{1 + \exp(\frac{0.924-\eta}{0.266})} - 0.073 \right) (q - 1). \quad (15)$$

The results of the numerical integration and the fit are shown in Fig. 6.



**Figure 6.** Numerical calculations of the ellipticity correction factor  $m$  (see equation 13), along with the analytical fit (equation 15). The horizontal dotted line shows  $m = 0.5$ , as used in equation (14).

## 2.6 Cluster fit quality

While our fitting procedure is designed to be robust, it does not perform perfectly on all clusters. We exclude clusters with unrealistic parameter values, which we define as a scale radius  $a < 0.1$  pixels,  $a > 15$  pixels, or an axial ratio  $q < 0.3$ . We also exclude any clusters where the fitted centre is more than 2 pixels away from the central pixel identified by LEGUS. This eliminates 6.7 per cent of the sample.

Additionally, to quantitatively evaluate which clusters have poor fits, we implement a quality metric based on a comparison of the cumulative light profiles of the cluster data and the model (after subtracting the best-fitting background from both the data and model). As our primary goal is to evaluate the reliability of  $R_{\text{eff}}$ , the cumulative profile is a strong indicator as it probes all light enclosed within a given radius.

Specifically, our metric uses the cumulative light profile to estimate the half-light radius of the cluster non-parametrically, and then compares the enclosed light of the model and data within this radius. The relative difference is

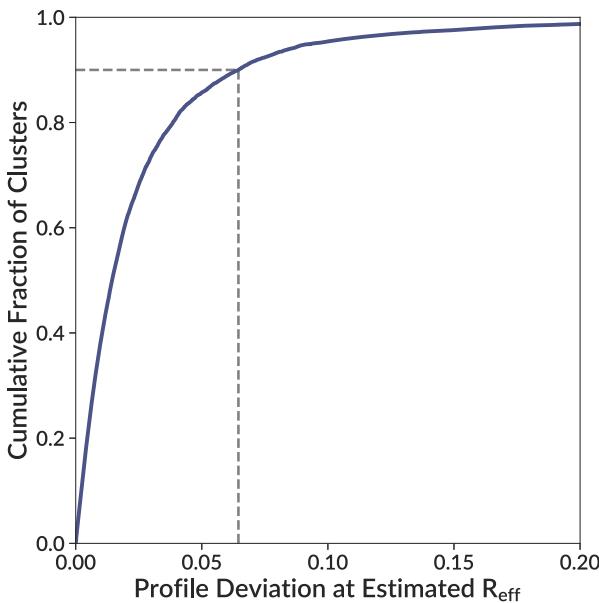
$$d = \left| \frac{F_{\text{model}}(< R_{1/2}) - F_{\text{data}}(< R_{1/2})}{F_{\text{data}}(< R_{1/2})} \right|, \quad (16)$$

where the non-parametric radius  $R_{1/2}$  is defined by

$$F_{\text{data}}(< R_{1/2}) = 0.5 F_{\text{data}}(< 15 \text{ pixels}). \quad (17)$$

Here,  $F(< R)$  is the cumulative flux enclosed within a circular radius  $R$ . We use 15 pixels as the maximum radius as it is the radius of the individual cluster snapshots.

We then calculate the distribution of this metric  $d$  for clusters that pass the cuts mentioned at the first paragraph of this section, shown in Fig. 7. The knee of the cumulative distribution is at approximately the 90th percentile, so we use that percentile as our cut. Any cluster above the 90th percentile will be excluded from the analysis in the rest of this paper. This results in a final sample of 6097 clusters with reliable radius measurements. This 90th percentile cut corresponds



**Figure 7.** Cumulative distribution for the deviation in the cumulative light profile as described in Section 2.6. Our analysis excludes clusters that fall above the 90th percentile of this metric, corresponding to approximately 6.5 per cent deviation of the integrated light within non-parametric  $R_{1/2}$ .

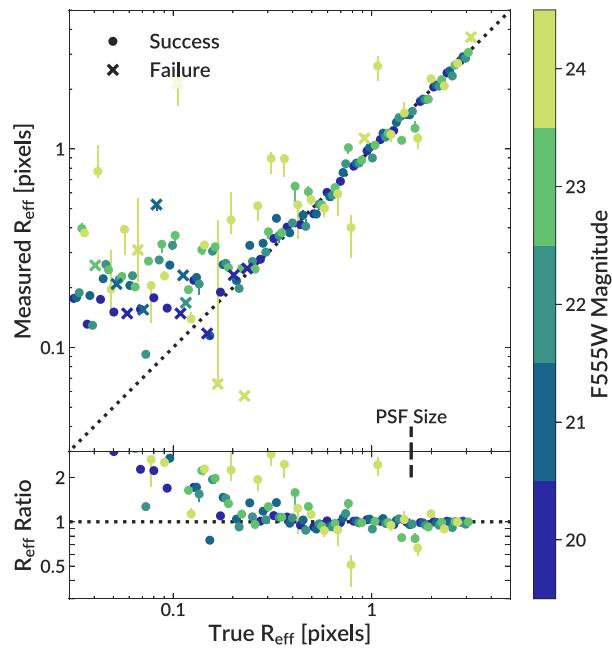
to about 6.5 per cent error on the light enclosed within the estimated effective radius, indicating the high quality of the fits we keep.

## 2.7 Artificial cluster tests

To test the ability of the pipeline to recover the effective radius of small clusters, we perform artificial cluster tests. We generate 150 synthetic clusters following the EFF profile. These clusters have magnitudes from 20 to 24,  $1.25 \leq \eta \leq 2.5$ , and  $0.03 \text{ pixels} \leq R_{\text{eff}} \leq 3 \text{ pixels}$ . All parameters are uniformly distributed within these ranges. To calculate the cluster magnitude, we follow the LEGUS pipeline as described in Adamo et al. (2017). We use a circular aperture with a radius of 4 pixels (chosen as the integer pixel value that contains 50 per cent of the flux of typical clusters) and a local sky annulus located at 7 pixels with a width of 1 pixel, and then apply the average aperture corrections for the NGC 628-c field. The artificial clusters span the range of the magnitudes of real clusters in this field.

We convolve these models with the PSF for the NGC 628-c field, add Poisson noise, and insert these artificial clusters into the NGC 628-c field. We then run our pipeline on this new image to measure their effective radii. The results of this test are shown in Fig. 8.

The pipeline is able to accurately measure cluster radii down to about 0.3 pixels. Below this point, the pipeline systematically overestimates the true radius. The performance of the pipeline does depend on magnitude, as faint clusters with magnitude 24 (the limit of the clusters in the LEGUS catalogue for NGC 628-c) have a much wider dispersion than brighter clusters, even at larger radii. A visual examination of these fits shows that contamination and noise are the primary causes of this dispersion. Faint sources rise above the background less than bright sources, so variations in the background can influence the fit more. Additionally, the Poisson pixel noise of the source itself can influence the fit, even for artificial clusters placed in a region where the background is smooth. Due to the nature of Poisson noise, this affects faint clusters the most.



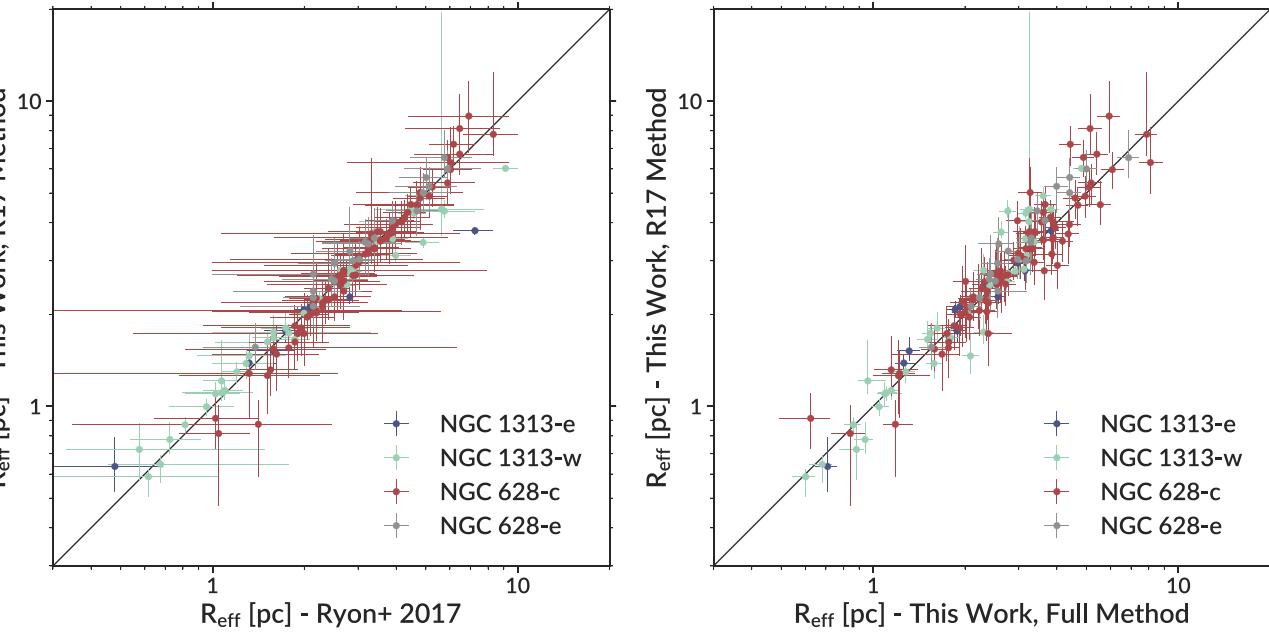
**Figure 8.** Results of artificial cluster tests. The top panel shows a comparison of the true effective radius to that measured by our pipeline. Solid circles show fits that the pipeline identified as successful, while the crosses show failures. The bottom panel shows the ratio of the measured effective radius to the true effective radius. Only successful fits are shown in this bottom panel. The dashed line spanning both panels indicates the PSF size in pixels for the NGC 628-c field, which is the field into which clusters were inserted.

Fig. 8 also shows the ability of the pipeline to detect when clusters have poor fits. Many of the catastrophic failures are correctly identified as failures. However, for the very compact clusters the pipeline identifies many fits that actually overestimate  $R_{\text{eff}}$  as reliable. This is likely because when  $R_{\text{eff}}$  is much smaller than the PSF, the observed cluster is not very different from the PSF. A slightly larger model is still very similar to the data, making the pipeline identify the fit as a success. This also means that generally the error bars may be underestimated for clusters with  $R_{\text{eff}} < 0.3$  pixels. Nevertheless, for most clusters our pipeline gives reliable measurements. Importantly, if  $R_{\text{eff}}$  is measured to be small, then it really is small. These trends are true for all values  $\eta \gtrsim 1.25$ . For values of  $\eta \sim 1$ , we found larger scatter.

## 3 RESULTS

### 3.1 Comparison to Ryon et al. (2017)

R17 used GALFIT to measure the effective radii of clusters in NGC 1313 and NGC 628, two of the galaxies in the LEGUS sample. To validate our method, we compare our measurements to those of R17. When making this comparison, we perform a separate round of fitting with several modifications to our method to match what was done in R17. We do not mask any contaminating sources, do not use radial weighting, and use the square of differences rather than the absolute value. When post-processing these results, we do not use a maximum radius to calculate the effective radii (instead using equation 11), we use the simple ellipticity correction (equation 14), and we use the same distances to NGC 1313 and NGC 628 as R17 did (Jacobs et al. 2009; Olivares E. et al. 2010). These changes ensure consistency with the R17 method.



**Figure 9.** Comparison of our cluster effective radii using different fitting methods for NGC 628 and NGC 1313. The left-hand panel shows a comparison to those of Ryon et al. (2017) using the same fitting method. The right-hand panel shows a comparison of this method to our full method used in the rest of the paper.

The left-hand panel of Fig. 9 shows the results of that comparison. Following R17, only clusters with  $\eta > 1.3$  are shown in this plot, as  $R_{\text{eff}}$  given by equation (11) is unreliable for lower values. To quantify the deviation, we use the RMS error, defined as

$$\text{RMS} = \sqrt{\frac{1}{N} \sum \frac{(R_{\text{eff},\text{R17}} - R_{\text{eff}})^2}{\sigma_{\text{R17}}^2 + \sigma^2}}, \quad (18)$$

where  $\sigma$  and  $\sigma_{\text{R17}}$  are the error on  $R_{\text{eff}}$  in this work and R17, respectively. This RMS deviation is 0.55, indicating excellent agreement.

In the right-hand panel of Fig. 9, we show a comparison of our full method to our R17-like method. These two methods show good agreement across the full radius range, with no significant deviations. The RMS deviation here is 1.62. This higher value is primarily driven by the error bars, which are smaller than those obtained in R17. In addition, a small number of clusters have significantly different radii between the two methods. The majority of these discrepant clusters have a small value for  $\eta$  (typically just above the cut-off of  $\eta > 1.3$  for inclusion in this plot), where the use of a maximum radius has the greatest effect on  $R_{\text{eff}}$  (see Fig. 5).

In the rest of this paper, we analyse the results obtained with our full fitting method.

### 3.2 Cluster radius distribution

In Fig. 10, we show the distribution of effective radii measured in the entire LEGUS sample with our full fitting method. For most galaxies, the distributions are remarkably similar. They have a peak at  $R_{\text{eff}} \approx 3$  pc, an extended tail to below 1 pc, and a sharper cut-off at large radii. This peak at  $R_{\text{eff}} \approx 3$  pc has been seen consistently in other studies of young clusters (e.g. Meurer et al. 1995; Larsen 1999; Scheepmaker et al. 2007; Bastian et al. 2012; Ryon et al. 2015, 2017; Cantat-Gaudin et al. 2018). Galaxies with similar distributions are shown in the left-hand panel of Fig. 10, while two galaxies NGC 1566 and NGC 7793 with discrepant shapes are shown in the right-hand panel and will be

discussed below. While we only show several galaxies individually in the left-hand panel of Fig. 10 for clarity, an examination of all galaxies shows that they have very similar distributions. In Table 1, we include the quartiles of the cluster radius distribution of each galaxy as another method of quantifying their distributions.

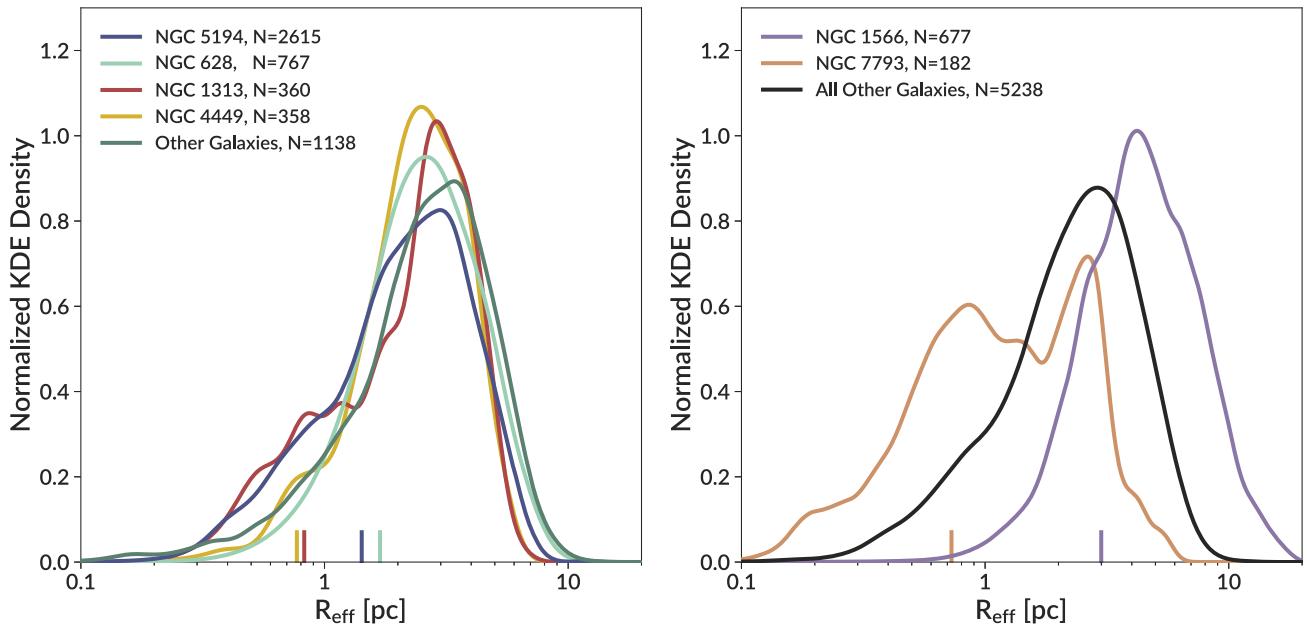
To characterize this common distribution, we create a stacked distribution of clusters from all galaxies other than the two discrepant galaxies. The distribution is shown as the comparison line in the right-hand panel of Fig. 10 and has a sharp peak at 2.9 pc. We compared it to several common analytical distributions and found that neither normal nor lognormal functions are good fits, due to the asymmetric shape of the observed distribution. Instead, we find that the Weibull distribution produces an excellent match:

$$\frac{dN}{dR_{\text{eff}}} = \frac{k}{\lambda} \left( \frac{R_{\text{eff}} - R_0}{\lambda} \right)^{k-1} \exp \left[ - \left( \frac{R_{\text{eff}} - R_0}{\lambda} \right)^k \right] \quad (19)$$

with  $k = 2.22$ ,  $\lambda = 3.64$  pc, and  $R_0 = 0.185$  pc.

To quantitatively test whether the individual galaxy samples are statistically consistent with being drawn from the same distribution, we employ the one-sided Kolmogorov–Smirnov test. We find that of the 31 galaxies, 16 have  $p$ -value  $> 0.01$  (13 with  $p$ -value  $> 0.05$ ), indicating that they are not inconsistent with the stacked distribution. However, of the 13 galaxies with more than 50 clusters, only 4 have  $p > 0.01$  (2 with  $p > 0.05$ ). The large number of clusters in these galaxies provides high statistical significance to formally distinguish the distributions. Still, the individual distributions exhibit strong visual similarity.

Two galaxies show cluster distributions significantly different from the rest. NGC 1566 appear shifted to larger radii than other galaxies. It has less low-radius clusters, a peak at larger radii (4.2 pc compared to 2.9 pc for the stacked distribution), and more high-radius clusters than any other galaxy. Selection effects may be partly responsible. At the adopted distance of NGC 1566 of 15.6 Mpc, 1 pixel covers 3 pc and our PSF model has an effective radius of 4.3 pc. Small



**Figure 10.** Kernel density estimation of the cluster radius distributions of the galaxies with the most clusters. The line for each galaxy shows the summed Gaussian kernels representing its clusters, where we use a width equal to twice each cluster’s radius error. Each curve is normalized to the same area for comparison purposes. The left-hand panel shows galaxies with similar cluster distributions, while the right-hand panel shows two galaxies with different distributions. Note that the ‘Other Galaxies’ in the left-hand panel do not include the two discrepant galaxies NGC 1566 and NGC 7793 shown in the right-hand panel, while the ‘All Other Galaxies’ in the right-hand panel include all galaxies shown in the left-hand panel. The tick marks at the bottom show the pixel size in parsecs for the images of each galaxy.

clusters may not be resolved and therefore not included in the LEGUS catalogue. While a full characterization of the LEGUS selection effects is beyond the scope of this paper, Adamo et al. (2017) examined the completeness as a function of cluster radius in NGC 628 at a distance of 8.8 Mpc. The CI cut excluded roughly 50 per cent of clusters with  $R_{\text{eff}} = 1$  pc. This selection effect is not significant for most galaxies, as the peak of the cluster radius distribution is at higher radii, and most galaxies are closer than NGC 628. However, since NGC 1566 is at approximately twice the distance of NGC 628, we can expect that its observations will be incomplete below 2 pc. This could explain the dearth of small clusters in NGC 1566, but would not explain the overabundance of large clusters. Our adopted distance could be responsible for this. As mentioned at the end of Section 2.4, the distance to NGC 1566 is uncertain, with distance estimates ranging from 6 to 20 Mpc. If our adopted distance of 15.6 Mpc is an overestimate, our cluster radius measurements will also be overestimated. Adopting a distance of 11 Mpc rather than 15.6 Mpc would bring it in line with the distributions of other galaxies, thus effectively treating this distribution like a standard ruler (Jordán et al. 2005). Future distance measurements may be able to resolve this and determine whether the cluster radius distribution in NGC 1566 is significantly different than that of other galaxies.

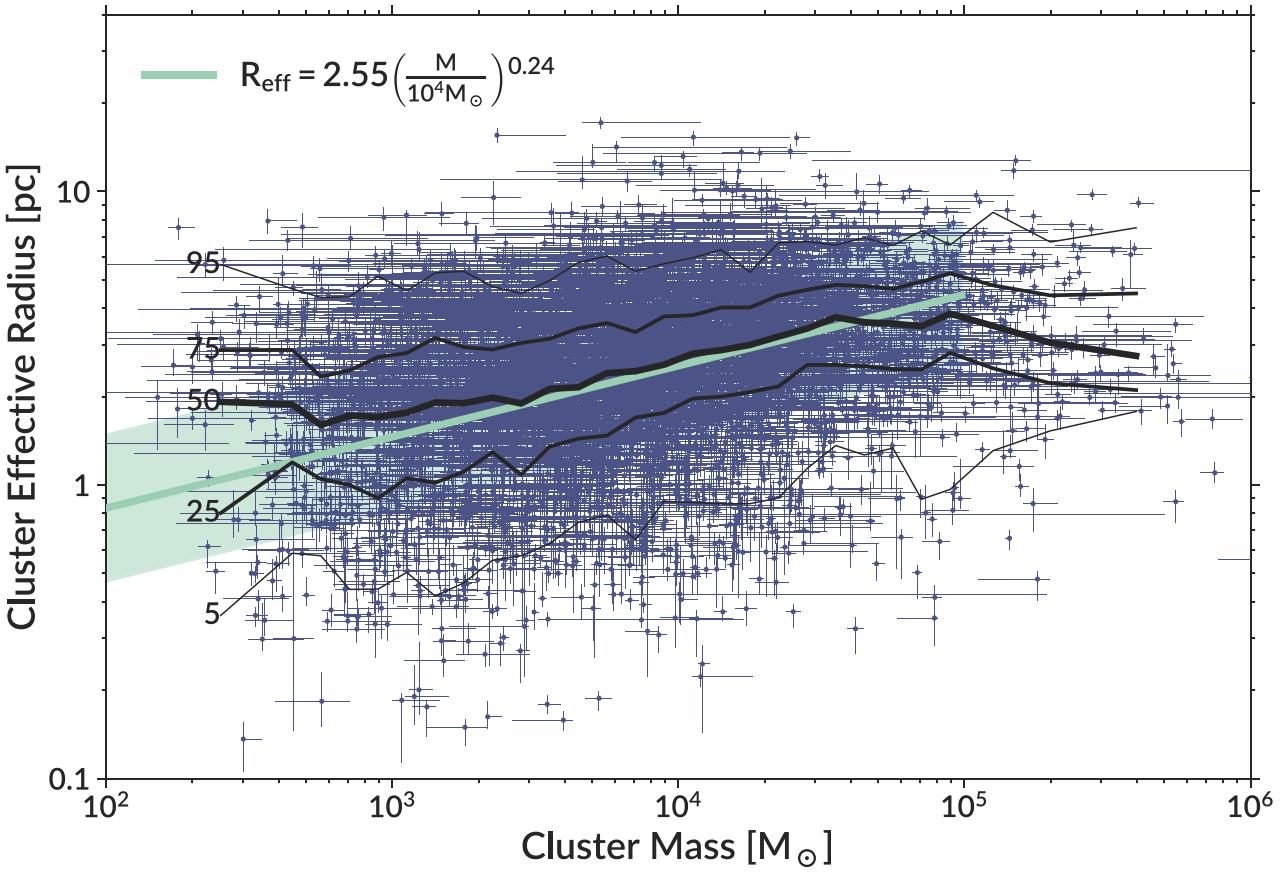
The other discrepant galaxy, NGC 7793, has a double-peaked distribution that is much broader than that in other galaxies. One peak is near the 3 pc peak seen in other galaxies, while another is at  $\sim 0.8$  pc. The reason for this is unclear. While NGC 7793 is split into two fields, both fields show the same double-peaked distribution. Its specific star formation rate is within the range of other galaxies. It is closer than most other galaxies in the sample, meaning the smallest clusters are more likely to be included, but other galaxies

with similar distances do not show this bimodal distribution. A visual examination of the spatial distribution within NGC 7793 of the clusters belonging to each peak does not show any striking trends. An examination of the age and mass of the clusters shows that, compared to other galaxies, NGC 7793 has more small young clusters and no large young clusters. Additionally, the age distribution is bimodal, with a deficit of intermediate-age clusters. Roughly speaking, this results in the low-radius peak being mostly young, low-mass clusters, while the high-radius peak is mostly old, high-mass clusters. Future detailed studies of NGC 7793 may be needed to understand its cluster population in more detail.

### 3.3 Cluster mass–radius relation for all LEGUS galaxies

Fig. 11 shows the mass–radius relation for the clusters in our sample. A clear mass–radius relation is visible, albeit with a shallow slope. To guide the eye, lines indicate the 5, 25, 50, 75, and 95th percentiles of the radius at a given mass.

This plot shows the full mass range of the LEGUS clusters. However, masses below  $5000 M_{\odot}$  measured by the deterministic method used in LEGUS may be unreliable, as the assumption of a fully sampled IMF is no longer valid (Maíz Apellániz 2009; Silva-Villa & Larsen 2011; Krumholz et al. 2015). To account for this, previous work with the LEGUS data, such as Adamo et al. (2017) and R17, restricted to clusters with masses above  $5000 M_{\odot}$ . As LEGUS is not complete for  $5000 M_{\odot}$  clusters at old ages, these papers selected clusters younger than 200 Myr. This produces a complete sample using clusters with the most reliable masses. We take a different approach in this paper, and use the full mass and age ranges of clusters with good SED fits, adjusting their errors to account for the systematic errors in the SED fitting. If we were to consider only



**Figure 11.** The mass–radius relation for the clusters in LEGUS. Clusters of all ages are included in this plot. The black lines show the 5, 25, 50, 75, and 95th percentiles of radii at a given mass. The solid line shows the best-fitting linear relation, with the shaded region showing the intrinsic scatter.

clusters more massive than  $5000 M_{\odot}$ , it would exclude about a third of the complete sample and greatly decrease the dynamic range of the mass–radius relation. Section 4.1 below discusses possible effects of incompleteness.

As discussed in Krumholz et al. (2015), clusters have poor fits – as quantified by the  $Q(\chi^2)$  statistic – when the assumption of a fully sampled IMF is violated. To avoid these unreliable low-mass clusters, we restrict ourselves to clusters with  $Q(\chi^2) > 10^{-3}$ . Of the 6097 clusters with successfully measured radii, 5105 pass this further cut. In addition to low-mass clusters, we also find qualitatively that this cut removes many clusters with high masses ( $> 10^5 M_{\odot}$ ) and very small radii ( $R_{\text{eff}} < 1 \text{ pc}$ ) that were outliers from the mass–radius relation due to their unreliable mass.

In addition, Krumholz et al. (2015) show that the deterministic LEGUS mass uncertainties are likely underestimated for low-mass clusters. To correct for this, we compare the mass uncertainty in Krumholz et al. (2015) to the uncertainty in the LEGUS catalogues. For clusters below  $5000 M_{\odot}$ , the median difference in uncertainty is 0.16 dex. We add this 0.16 dex correction to the mass uncertainty of all clusters below  $5000 M_{\odot}$  when performing our fits.

This produces a sample across the full mass range that includes only the most reliable low-mass clusters and adjust their errors to account for the systematic error in the deterministic LEGUS SED fitting. Our resulting sample may not be complete (as we will be missing old, low-mass clusters), but we will discuss these selection effects throughout the rest of the paper.

We fit this mass–radius distribution assuming a power-law relation:

$$\hat{R}_{\text{eff}}(M) = R_4 \left( \frac{M}{10^4 M_{\odot}} \right)^{\beta} \quad (20)$$

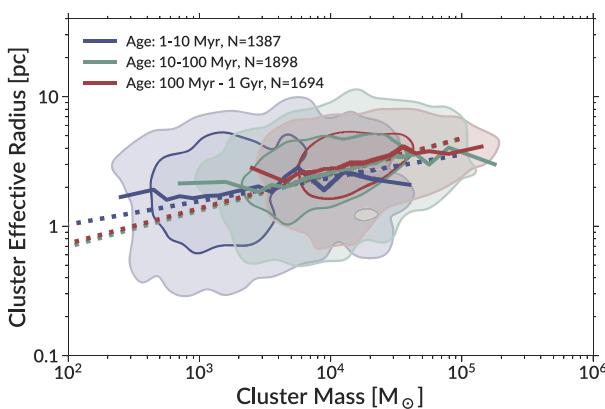
such that the normalizing factor  $R_4$  is the effective radius at  $10^4 M_{\odot}$ . We incorporate errors in both mass and radius and account for the intrinsic scatter by minimizing projected displacements of data points from the fit line as outlined in Hogg, Bovy & Lang (2010). In Appendix A, we describe this in more detail and compare it to a hierarchical Bayesian model that includes a treatment of selection effects. We found that neither the fitting method nor the inclusion of selection effects made a difference in the fit parameters, so we use this simpler method. We restrict the fitting to clusters below  $10^5 M_{\odot}$ , as the relation appears to flatten above that mass, possibly because of small-number statistics. We determine errors on the fit via bootstrapping. For this full sample, our best-fitting power-law slope is  $\beta = 0.24$ , with an intrinsic scatter of 0.25 dex. Restricting to clusters with ages younger than 1 Gyr produces the same result, primarily because most clusters in LEGUS are younger than 1 Gyr.

With this large sample, we can investigate the cluster mass–radius relation for various subsamples of the data. For all cases below, we restrict our sample to clusters younger than 1 Gyr and less massive than  $10^5 M_{\odot}$ . Table 2 shows the fitting parameters for all the subsets discussed below.

Fig. 12 shows the mass–radius relation split by cluster age. The mass range of the three bins is clearly different, due to LEGUS’s

**Table 2.** Fits to the cluster mass–radius relation for various subsets of the data. All fits were done using clusters with masses below  $10^5 M_\odot$ , while all fits other than the first row were done using clusters younger than 1 Gyr. Note that the fits including other data sets are sensitive to how these data sets are weighted, adding systematic uncertainties to the fit parameters.

| Selection                      | $N$  | $\beta$ : Slope   | $R_4: R_{\text{eff}}$ (pc) at $10^4 M_\odot$ | Intrinsic scatter | log $M$ percentiles: 1–99 |
|--------------------------------|------|-------------------|--|-------------------|---------------------------|
| Full LEGUS Sample              | 5105 | $0.242 \pm 0.010$ | $2.548 \pm 0.022$                            | $0.250 \pm 0.003$ | 2.57–5.40                 |
| 1 Myr–1 Gyr                    | 4979 | $0.246 \pm 0.010$ | $2.554 \pm 0.022$                            | $0.250 \pm 0.003$ | 2.57–5.26                 |
| Age: 1–10 Myr                  | 1387 | $0.180 \pm 0.028$ | $2.365 \pm 0.106$                            | $0.319 \pm 0.006$ | 2.40–4.89                 |
| Age: 10–100 Myr                | 1898 | $0.279 \pm 0.021$ | $2.506 \pm 0.035$                            | $0.238 \pm 0.005$ | 2.91–5.24                 |
| Age: 100 Myr–1 Gyr             | 1694 | $0.271 \pm 0.027$ | $2.558 \pm 0.048$                            | $0.198 \pm 0.005$ | 3.46–5.40                 |
| LEGUS + MW                     | 6158 | $0.296 \pm 0.002$ | $2.555 \pm 0.022$                            | $0.225 \pm 0.003$ | 0.93–5.22                 |
| LEGUS + External Galaxies      | 5874 | $0.229 \pm 0.008$ | $2.561 \pm 0.020$                            | $0.244 \pm 0.003$ | 2.36–5.31                 |
| LEGUS + MW + External Galaxies | 7053 | $0.292 \pm 0.002$ | $2.567 \pm 0.020$                            | $0.222 \pm 0.003$ | 0.97–5.26                 |



**Figure 12.** The mass–radius relation for the clusters in LEGUS, split by age. Contours enclose 50 and 90 per cent of the clusters in each bin, and are smoothed by a kernel of 0.08 dex. The dashed lines show the fits, while the solid lines show the running median in each age bin.

absolute  $V$ -band magnitude cut. Evolutionary fading results in only massive clusters being detected at older ages. To demonstrate this, we include the mass range spanned by the 1–99th percentile of each sample in Table 2. In all three bins, we detect a significantly non-zero slope of the mass–radius relation. The running medians of each panel are quite similar, especially for the two oldest age bins, which deviate from the 1–10 Myr bin at  $M > 10^4 M_\odot$ . This matches what we find in the formal fit, where the two older bins have a slope and normalization indistinguishable from each other, while the 1–10 Myr bin has a significantly shallower slope. In addition, the intrinsic scatter decreases with age.

Additionally, we supplement the LEGUS sample with several other large samples for young star clusters, mostly from the compilation of Krumholz et al. (2019). In all the samples below, we restrict to clusters with an age less than 1 Gyr and masses below  $10^5 M_\odot$ , as done for our main fits. This age cut means that we do not include any globular clusters. Additionally, some of the samples in Krumholz et al. (2019) are for galaxies already included in this paper (namely NGC 628, NGC 1313, and NGC 5194), so we do not include them again here.

We include Milky Way OCs within 2 kpc of the Sun from Kharchenko et al. (2013), who measured King parameters for these clusters. Following Krumholz et al. (2019), we calculate mass using equation (3) of Piskunov et al. (2007), Oort constants from Bovy (2017), and the distance from the Sun to the Galactic Centre from

Bland-Hawthorn & Gerhard (2016). We also include the sample of 12 Milky Way YMCs compiled in Krumholz et al. (2019).

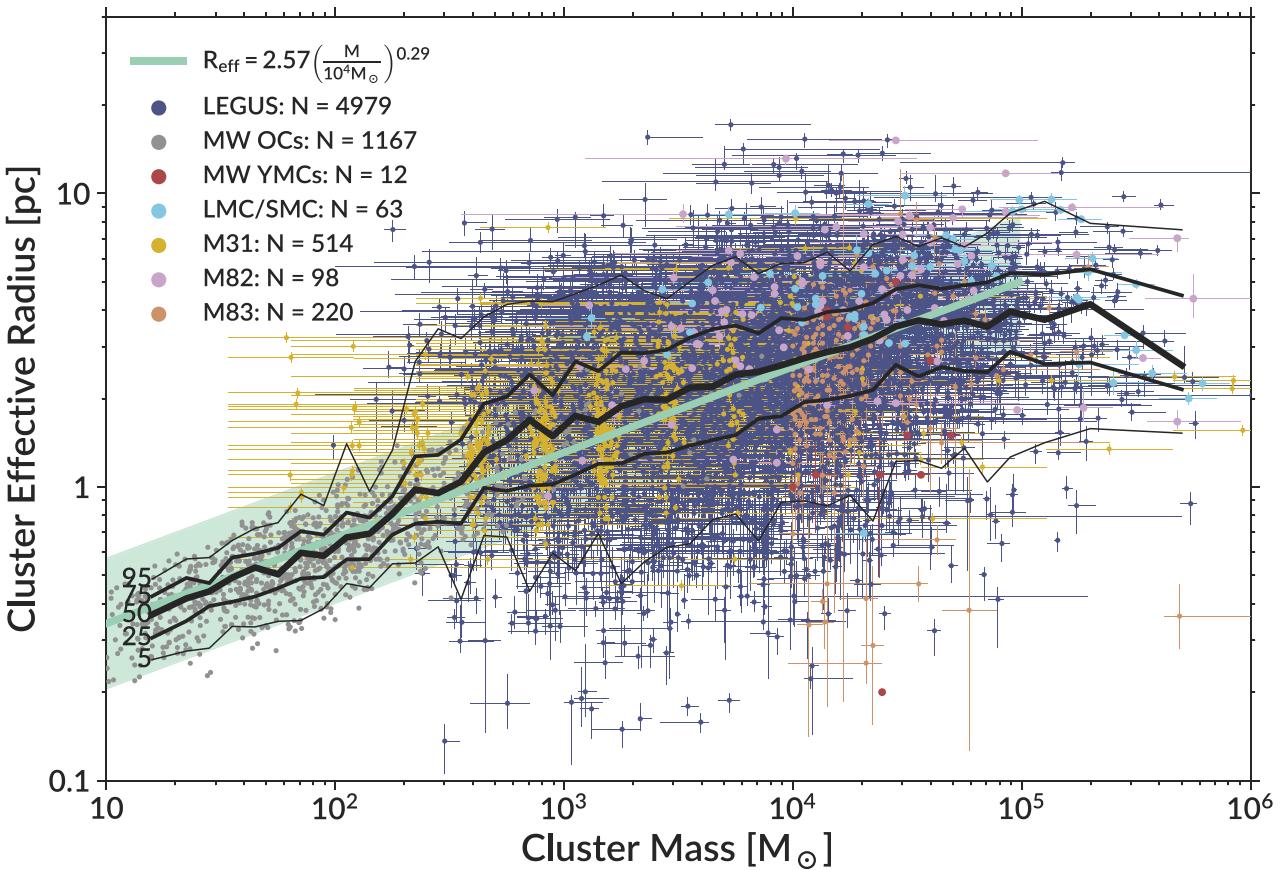
We additionally include samples from several external galaxies. Mackey & Gilmore (2003a, b) measured radii for 53 clusters in the LMC and 10 clusters in the SMC. EFF profiles were fitted to the surface brightness profiles of clusters. These surface brightness profiles were also used to obtain the total luminosity of each cluster, which was converted into the cluster mass by using mass-to-light ratios.

The Panchromatic Hubble Andromeda Treasury (PHAT) survey identified stellar clusters in M31 (Johnson et al. 2012; Fouesneau et al. 2014). Half-light radii were determined by interpolating the flux profile to find the radius in arcseconds that includes half of the light (Johnson et al. 2012). We then use a distance of 731 kpc to M31 (Wagner-Kaiser et al. 2015) to convert the radii to parsecs. Masses were determined using a Bayesian SED fitting method that explicitly accounts for the stochastic sampling of the IMF (Fouesneau et al. 2014).

In a series of papers, Cuevas-Otahola et al. (2020, 2021) calculated structural parameters for 99 star clusters in M82. They fit EFF, King, and Wilson profiles to the surface brightness profiles, finding that the EFF profile best represents the clusters in their sample. Similarly to Mackey & Gilmore (2003a, b), masses are determined by applying a mass-to-light ratio to fitted luminosities.

We also include the clusters in M83 measured by Ryon et al. (2015). Radii are measured by fitting an EFF profile to the 2D light profile using GALFIT, as in R17. Masses are derived in Silva-Villa et al. (2014) and are done using an SED-fitting method similar to that used in LEGUS (Adamo et al. 2010).

Fig. 13 shows the mass–radius relation including these data sources. In the fit shown in this figure, we give each cluster equal weight, no matter which data set it comes from. The addition of the MW OCs in particular extends the mass–radius relation down to very small masses, while the other samples overlap nicely with LEGUS. Including the MW clusters produces a steeper slope than the fit using only LEGUS clusters ( $\beta = 0.296$ ). In addition, the intrinsic scatter decreases, likely due to the smaller scatter in the MW OC data. Including the data from external galaxies produces a shallower slope ( $\beta = 0.229$ ), likely due to fewer low-mass clusters with small radii in the M31 PHAT sample. Including all data produces a fit similar to the fit for LEGUS + MW, likely due to the leverage and large numbers provided by the low-mass clusters in the MW sample. In all cases,  $R_4$  is consistent with that measured in the LEGUS sample alone. Note that due to the asymmetric shape of the cluster radius distributions,  $R_4$  may be different from the peak value of the distribution quoted in Section 3.2.



**Figure 13.** The mass–radius relation for the clusters in LEGUS as well as external data sets, as described in the text. Only clusters with ages less than 1 Gyr are shown here. The black lines show the 5, 25, 50, 75, and 95th percentiles of radii at a given mass. Note that a small random shift has been applied to the original discrete M31 masses for visual purposes.

### 3.4 Are clusters gravitationally bound?

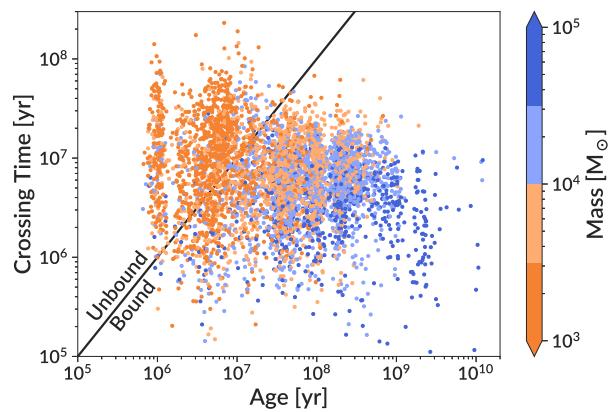
Gieles & Portegies Zwart (2011) suggest a distinction between bound clusters and unbound associations, where bound clusters are older than their instantaneous crossing time:

$$t_{\text{cross}} = 10 \left( \frac{R_{\text{eff}}^3}{GM} \right)^{1/2}. \quad (21)$$

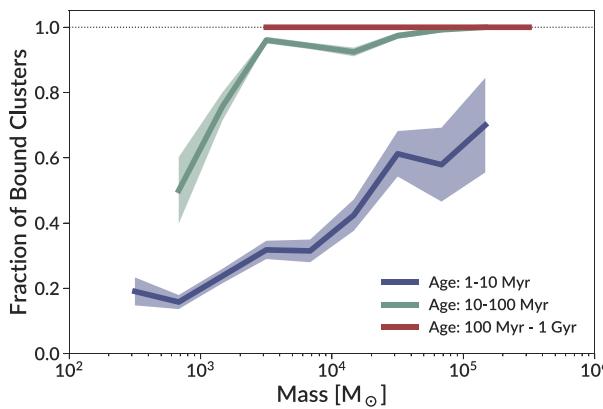
Clusters with  $t_{\text{age}} > t_{\text{cross}}$  have remained together for their lifetimes, indicating that they are gravitationally bound. Unbound objects expand with time, causing the crossing time to increase proportionally.

Fig. 14 shows the comparison of these time-scales for the clusters in the LEGUS sample. The majority of clusters are bound. The only unbound clusters are young (less than 10 Myr) and tend to be less massive. At a given age, the less massive clusters are more likely to be unbound. We find that 78 per cent of all clusters, 92 per cent of clusters with  $M > 5000 M_{\odot}$ , and 97 per cent of clusters older than 10 Myr are bound.

Fig. 15 shows how the fraction of clusters that are bound changes with age and mass. 100 per cent of objects older than 100 Myr are bound, while in the other age bins a clear trend with mass is seen. For the 10–100 Myr bin, nearly all clusters above  $3000 M_{\odot}$  are bound, while the youngest clusters show a steadily increasing fraction of bound clusters with mass. This confirms that the LEGUS pipeline selects gravitationally bound objects, especially for clusters with higher masses or older ages.



**Figure 14.** A comparison of the crossing time to the age of clusters in LEGUS. Clusters are colour coded according to their mass, and a small random offset was added to the discrete ages for visual purposes. The black line shows where these times are equal. Clusters where the age is longer than the crossing time are likely bound, while those where the age is smaller are likely unbound.



**Figure 15.** The fraction of clusters that are older than their crossing times (indicating that they are bound objects) as a function of mass. Shaded regions show the 68 per cent confidence region.

## 4 DISCUSSION

### 4.1 Selection effects

When selecting clusters, the LEGUS survey used an absolute magnitude cut, selecting clusters with a V-band magnitude brighter than  $-6$  mag. As massive stars in clusters die, the cluster fades. This means that older clusters must be more massive to be detected, producing a significant selection effect in the LEGUS sample.

This is particularly visible in Fig. 12. In the oldest age bin, there are nearly no clusters below a few  $10^3 M_\odot$ , where the bulk of the youngest clusters are. The mass ranges seen in Fig. 12 should not be simply interpreted as evidence for disruption of low-mass clusters, as old low-mass clusters would not be detected even if they existed. This also complicates an examination of cluster evolution, as without old low-mass clusters to compare, it is difficult to test predictions of low-mass cluster evolution.

In the above results and discussion that follows, we present results using all LEGUS clusters. We want to be clear that this is not a complete sample. Where relevant, we discuss how these selection effects may bias the results presented.

LEGUS also uses a cut in CI (with a value that varies for each galaxy). This cut may result in the removal of the smallest clusters. This will depend on galaxy distance, as the smallest clusters will be possible to resolve in nearby galaxies but not distant ones. As mentioned earlier in Section 3.2, this is not likely to affect many of our galaxies. Adamo et al. (2017) examined the completeness as a function of radius for NGC 628 at a distance of 8.8 Mpc, finding that LEGUS includes roughly 50 per cent of clusters with  $R_{\text{eff}} = 1$  pc. As most of our galaxies are closer than NGC 628, they will be less affected. The radius distribution shows a clear peak significantly above the radius where we may be incomplete (Fig. 10), showing that the potential removal of small clusters likely will not dramatically change our results.

We also note that because of the inability of our pipeline to pick up extremely small objects (smaller than 0.3 pixels; see Fig. 8), the smallest objects may be even more compact than reported.

The inclusion of small, low-mass clusters would have the effect of steepening the mass–radius relation. Interestingly, a comparison of the slopes in Figs 11 and 13 shows that a steeper slope better matches the MW OCs. If the true mass–radius relation is steeper

than we measure here, it may make our results more consistent with the measurements in the MW.

Lastly, we note that the radial coverage of each galaxy varies. About half of the LEGUS galaxies are compact enough that they can be covered with one *HST* WFC3/UVIS pointing, but larger galaxies are not completely covered. Some galaxies include only central regions (e.g. NGC 1566), while others include the central regions and part of the disc (e.g. NGC 628). See fig. 3 of Calzetti et al. (2015) for the full footprints for all LEGUS galaxies. This uneven coverage may bias our results somewhat if cluster populations vary throughout galaxies. This could happen if they are tidally bound, as the tidal radius would change with galactocentric radius. We defer a detailed examination of this for a future paper.

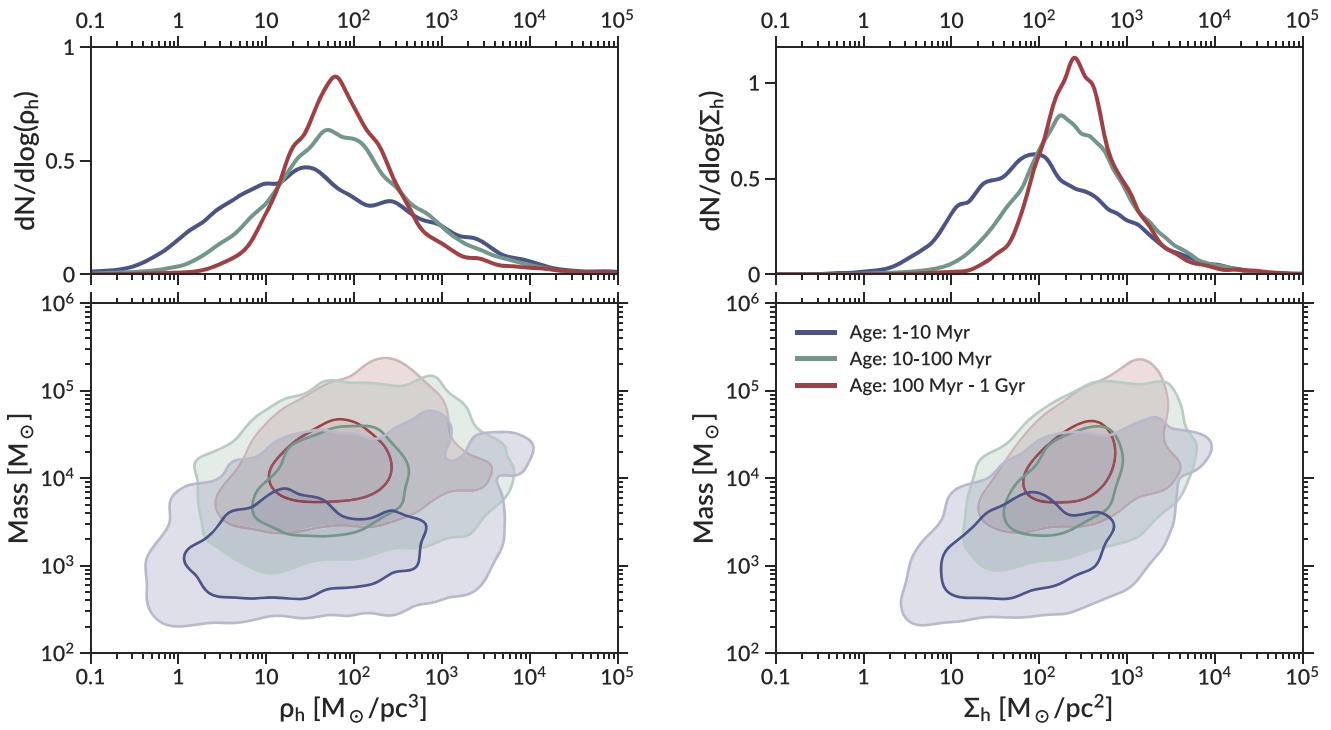
### 4.2 Mass–radius relation

The mass–radius relation shown in Fig. 11 has a relatively shallow slope and significant intrinsic scatter. Nevertheless, a relation is clearly present, even when splitting by age (Fig. 12).

In the Milky Way, observations of GMCs show roughly constant surface densities (Larson 1981). From this, we can expect a mass–radius relation  $R = (M/\pi\Sigma)^{1/2} \propto M^{1/2}$ . Measurements of clumps have found a range of slopes from 0.3 to 0.6 (Roman-Duval et al. 2010; Urquhart et al. 2018; Mok, Chandar & Fall 2021). These relations are steeper than the relation we measure for young clusters, which presumably form from such clouds. However, we note that the hierarchical structure of the ISM makes determining the size of a clump more challenging than measuring the radius of a star cluster, so these radii might not be directly comparable (Colombo et al. 2015). We examine cluster density further in Section 4.3. The analytical model of Choksi & Kruijssen (2019) also predicts a mass–radius relation of the form  $R_{\text{eff}} \propto (M/\Sigma_g)^{1/2}$ , where  $\Sigma_g$  is the gas surface density. After accounting for the fact that massive clusters can only form in high-density environments, they find a lower slope that is more consistent with this work.

Many previous studies have found inconclusive evidence of a correlation between the mass and radius of young clusters. Zepf et al. (1999) found a shallow mass–luminosity relation, with later studies finding little evidence of a strong mass–radius relation (Bastian et al. 2005, 2012; Scheepmaker et al. 2007; Ryon et al. 2015, 2017). Some studies have found a mass–radius relation for the most massive clusters above  $10^6 M_\odot$  (Kissler-Patig, Jordán & Bastian 2006; Bastian et al. 2013), but that mass range is not sampled in our results. The large sample size, uniform LEGUS selection, and uniform fitting procedure presented here allow us to clearly detect this relation. Our result is similar to that of Cuevas-Otahola et al. (2021), who find a power-law slope of 0.29. However, the normalization of their relation is higher than ours (see Fig. 13).

Interestingly, we also find less evolution in the cluster radius with age than seen in other studies. Bastian et al. (2012) fit the cluster radius distribution as a bivariate function of age and mass, finding that age is the stronger driver of cluster radius than mass. Ryon et al. (2015) found a significant age–radius relation. They also found a steepening of the mass–radius relation with time, although their age bins are different from the bins used in this paper. Chandar et al. (2016) found that clusters with age of 100–400 Myr have radii 4 times larger than clusters of similar mass with ages younger than 10 Myr. Our results stand in contrast, as we find no significant evolution after 10 Myr in the mass ranges probed.



**Figure 16.** Density  $\rho_h$  (left-hand panels) and surface density  $\Sigma_h$  (right-hand panels) of clusters within the half-light radius. The top panels show kernel density estimation of the density distributions, where clusters are smoothed by a Gaussian kernel with a width equal to their measurement error. Each curve is normalized to integrate to the same area. The bottom panels show the dependence of densities on mass. The contours enclose 50 and 90 per cent of the data, and are smoothed by 0.15 dex. In all panels, we split the sample by age.

#### 4.3 Density distribution

Using our measured radii, we calculate the average density and surface density of the LEGUS clusters:

$$\rho_h = \frac{3M}{8\pi R_{\text{eff}}^3}, \quad \Sigma_h = \frac{M}{2\pi R_{\text{eff}}^2}. \quad (22)$$

In this section, we will use  $\rho_h$  and  $\Sigma_h$  when referring to those quantities, respectively, and use the generic term ‘density’ when referring to both of them. In Fig. 16, the top panels show the distributions of these densities split by cluster age. Younger clusters have wider ranges and extend to lower densities than old clusters. The bottom panels show the distribution of densities as a function of mass.

We find that the distributions shown in the top panels of Fig. 16 are well described by lognormal functions:

$$\frac{dN}{d \log \rho_h} = \frac{1}{\sqrt{2\pi\sigma_\rho^2}} \exp \left[ -\frac{(\log \rho_h - \log \mu_\rho)^2}{2\sigma_\rho^2} \right] \quad (23)$$

and the equivalent for  $\Sigma_h$ , where  $\mu_\rho$  and  $\sigma_\rho^2$  are the mean and variance, respectively. We fit these distributions, and show the parameters in Table 3. We also include a fit to the entire distribution without splitting by age.

The decrease in the number of low-density clusters with age is likely to be a combination of selection effects and disruption of low-mass clusters. There appears to be a weak mass– $\rho_h$  relation, and a much stronger relation between mass and  $\Sigma_h$ . As old clusters are massive, they are more likely to have high  $\Sigma_h$ . However, in the narrow mass range where the age distributions overlap (around  $5 \times 10^4 M_\odot$ ), the youngest age bin extends to lower density than the older age bins. We also examined the density distributions for clusters in the mass

**Table 3.** Lognormal fits to the density distributions shown in Fig. 16. The log mean  $\mu$  and standard deviation  $\sigma$  are given for density  $\rho_h$  and surface density  $\Sigma_h$ . Note that the cluster mass ranges given in Table 2 apply to these fits too.

| Age           | $\log \mu_\rho$<br>( $M_\odot \text{pc}^{-3}$ ) | $\sigma_\rho$<br>(dex) | $\log \mu_\Sigma$<br>( $M_\odot \text{pc}^{-2}$ ) | $\sigma_\Sigma$<br>(dex) |
|---------------|---|------------------------|---|--------------------------|
| All           | 1.80  | 0.78                   | 2.36  | 0.61                     |
| 1–10 Myr      | 1.56  | 1.13                   | 1.98  | 0.82                     |
| 10–100 Myr    | 1.83  | 0.78                   | 2.37  | 0.60                     |
| 100 Myr–1 Gyr | 1.82  | 0.57                   | 2.44  | 0.43                     |

range shared by all age bins, again finding a larger spread at younger ages. This may indicate that disruption is responsible for removing these low-density clusters. At the same time, we cannot rule out that higher mass clusters simply form at higher density.

Observations of GMCs in nearby quiescent galaxies are consistent with a roughly constant surface density  $\Sigma_{\text{GMC}} \sim 100 M_\odot \text{pc}^{-2}$ , while in starbursting and high-redshift galaxies the normalization is higher  $\Sigma_{\text{GMC}} \sim 2000 M_\odot \text{pc}^{-2}$  (e.g. Dessauges-Zavadsky et al. 2019). Clumps within resolved Galactic clouds are also consistent with a nearly constant surface density  $\Sigma_{\text{GMC}} \gtrsim 1000 M_\odot \text{pc}^{-2}$  (Urquhart et al. 2018), although fixed column density may be a selection effect.

The density of the LEGUS clusters is similar to that of GMCs in nearby quiescent galaxies. The peak of the  $\Sigma_h$  distribution for young clusters is  $\approx 100 M_\odot \text{pc}^{-2}$ . However, there is a wide range in cluster densities, in contrast to the narrower range of  $\Sigma_{\text{GMC}}$ . This may be due to the fact that there is not a direct connection between the density of GMCs and clusters. Clusters form out of the densest clumps within GMCs, which in the Milky Way typically have  $\Sigma_{\text{GMC}}$  between 100

and a few  $10^4 M_\odot \text{ pc}^{-2}$  (Urquhart et al. 2018). After stars form out of these clumps, stellar feedback disperses the gas. This causes the cluster to increase in size and decrease in density. We note that we are measuring the radii of clusters at this phase of their evolution, after gas expulsion.

We also note that the LEGUS sample is from many galaxies with a range of star formation rates. This may produce a range of GMC properties that are partially responsible for explaining the scatter in cluster density (Sun et al. 2018). In future work, we will examine the dependence of cluster properties on their environment.

Taking full density distributions from the bottom panels of Fig. 16, we fit power-law relations and obtain  $\rho_h \propto M^{0.52 \pm 0.02}$  and  $\Sigma_h \propto M^{0.67 \pm 0.012}$ . The fitted intrinsic scatter in  $\rho_h$  is  $1.12 \pm 0.014$  dex, while for  $\Sigma_h$  it is  $0.74 \pm 0.009$  dex. Errors are determined by bootstrapping. As a consistency check, we can compare these fit slopes with those expected based on mass–radius relation fit. For the full LEGUS sample,  $R_{\text{eff}} \propto M^{0.242}$ . By rewriting that relation in terms of densities, the expected relations are  $\rho_h \propto M^{0.274}$  and  $\Sigma_h \propto M^{0.516}$ . These are less steep than the direct fits, especially for  $\rho_h$ . Any discrepancy in the slopes may be due to the very large intrinsic scatter in densities. Their dynamic range is larger than the dynamic range in mass. This scatter, along with the lack of a clear relation (particularly for the mass– $\rho_h$  relation), makes it difficult to fit a reliable slope.

#### 4.4 Cluster evolution

Previous literature indicated that clusters may expand with time (e.g. Bastian et al. 2012; Ryon et al. 2017). As stars within clusters lose mass, isolated clusters will slowly expand to maintain virial equilibrium, while at later times two-body relaxation can also increase the cluster radius (Gieles et al. 2010).

In Fig. 12, we see a statistically significant evolution in radius with age from the 1–10 Myr bin to the 10–100 Myr bin, with high-mass clusters slightly expanding. However, the magnitude of this increase is quite small. The fit parameters indicate typical clusters at  $10^4 M_\odot$  expand from 2.36 to 2.51 pc, while clusters at  $10^5 M_\odot$  expand from 3.58 to 4.76 pc. Notably, we only see this evolution between our first two age bins. We see no significant differences between the 10–100 Myr and 100 Myr–1 Gyr bins.

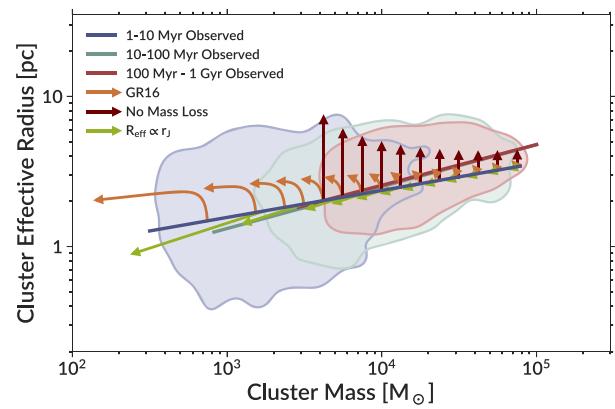
To start interpreting these results, we turn to an examination of the Jacobi radius, which sets the radius at which stars belong to a cluster when it is in a tidal field. Clusters that fill a larger fraction of the Jacobi radius are more vulnerable to mass-loss. For clusters with mass  $M$  in circular orbits with angular frequency  $\omega$  in a galaxy with a flat rotation curve, the Jacobi radius is defined as

$$r_J = \left( \frac{GM}{2\omega^2} \right)^{1/3}. \quad (24)$$

We do not directly calculate  $r_J$ , as obtaining  $\omega$  for each cluster is beyond the scope of this paper, and the assumption of a flat rotation curve may not be true for every galaxy. However, we can qualitatively examine how the ratio of the effective radius to the Jacobi radius scales with cluster mass:

$$\frac{R_{\text{eff}}}{r_J} \propto \frac{M^\beta}{M^{1/3}}, \quad (25)$$

where  $\beta$  is the slope of the mass–radius relation (equation 20 and Table 2). Note that this assumes no relation between  $M$  and  $\omega$ . For the full sample,  $\beta = 0.24$ , giving  $R_{\text{eff}}/r_J \propto M^{-0.09}$ . In the youngest age bin (1–10 Myr),  $\beta = 0.18$ , and  $R_{\text{eff}}/r_J \propto M^{-0.15}$ . In either case, high-mass clusters fill less of their Jacobi radii than low-mass clusters.



**Figure 17.** A comparison of three models of cluster evolution. Contours enclose 75 per cent of the clusters in each age bin, and are smoothed by 0.08 dex. The corresponding solid lines show the fit for these age bins from Fig. 12, but are restricted to the mass range spanned by the 1–99th percentile of masses in this age bin from Table 2. The sets of arrows show how three models for cluster evolution as described in Section 4.4 would change clusters lying on the 1–10 Myr relation after 300 Myr of evolution.

To examine the evolution of clusters with time, we also use the evolution model from Gieles & Renaud (2016), known hereafter as **GR16**. This model includes two processes that influence cluster evolution: tidal shocks and two-body relaxation. Tidal shocks increase the energy of the cluster and result in mass-loss, while two-body relaxation is assumed to only increase the energy of the cluster without causing mass-loss. To show how this model would change the clusters in the sample, we take clusters along the 1–10 Myr best-fitting relation and evolve them through this model for 300 Myr. Fig. 17 shows contours of the observed distribution of clusters in the three age bins, their best-fitting relations, and arrows illustrating how the **GR16** model affects clusters in this 300 Myr of evolution.

Fig. 17 also includes two other toy models of cluster evolution, designed to represent the tidally limited and not tidally limited extremes. When clusters are very small compared to Jacobi radii, any energy injection will increase the radius without causing substantial mass-loss. On the other hand, any energy from either two-body relaxation or shocks will cause mass-loss in tidally limited clusters, and the cluster’s effective radius will decrease along with the tidal radius.

To model these two cases, we modify the mass-loss prescriptions of **GR16**. They introduce a parameter  $f$  that relates mass-loss to energy injection (their equation 2):

$$\frac{dM}{M} = f \frac{dE}{E}. \quad (26)$$

For shocks, they set  $f_{\text{sh}} = 3$ , while setting  $f_{\text{rlx}} = 0$  turns off mass-loss from two-body relaxation.

We change these  $f$  values to produce our two cases. In the case where clusters are not tidally limited and mass-loss does not happen, we set both  $f_{\text{sh}} = f_{\text{rlx}} = 0$ . We then rederive the model, and its results are shown by the ‘No Mass-Loss’ arrows in Fig. 17. In the tidally limited case, we keep the original  $f_{\text{sh}} = 3$  but also allow for mass-loss from two-body relaxation with  $f_{\text{rlx}} = 0.2$  (Gieles et al. 2006). As the cluster loses mass, we require the radius to be proportional to the tidal radius:

$$\frac{R_{\text{eff}}}{R_{\text{eff},0}} = \frac{r_J}{r_{J,0}} = \left( \frac{M}{M_0} \right)^{1/3}. \quad (27)$$

The scaling relation presented above in equation (25) indicates that low-mass clusters are more likely to be tidally limited than high-mass clusters, and Fig. 17 qualitatively supports this conclusion. If high-mass clusters are not tidally limited, they will lose little mass and expand, matching the observations. While the mass range of old clusters prohibits a detailed examination of low-mass clusters, the mass–radius relation would steepen if the effective radius of low-mass clusters evolves proportional to the tidal radius.

However, it is clear that none of these models do a good job of quantitatively matching the full evolution. The GR16 model pushes clusters towards a mean relation of  $R_{\text{eff}} \propto M^{1/9}$ , making the relation shallower rather than steeper as required by the observations. The toy model with no mass-loss can increase the radius of massive clusters, but its effects are weakest for the highest mass clusters where the observations show the largest difference with age. The toy model that assumes the radius changes with the tidal radius may work for low-mass clusters, but for high-mass clusters it has nearly no effect.

Importantly, the time dependence of these models is in strong conflict with the observations. We see a change from the 1–10 Myr age bin to the 10–100 Myr bin, with no significant change afterwards. However, the models change clusters steadily with time, leading to little change in the first  $\sim 30$  Myr but large changes after 300 Myr.

In addition, models would need to match the change in scatter with time. As clusters evolve towards the mean relation of the GR16 model, the scatter decreases dramatically. We tested this by taking the full 1–10 Myr sample and putting each cluster through the GR16 model for 300 Myr. At this late time, the distribution of cluster radii has a much smaller scatter than seen in the observations of clusters at late times. While the observed scatter does decrease with time, it decreases less than this model predicts.

We note that stellar mass-loss is not included in the GR16 model and the modified versions presented here. In 1 Gyr, clusters can lose roughly 30 per cent of their mass through stellar evolution alone, and this can cause them to expand (Gieles et al. 2010). In addition, one should be careful comparing the cluster mass from the models to the observed mass. The models treat mass-loss from stars leaving the cluster as instantaneous, while in reality stars can remain in clusters for a long time before escaping through the Lagrangian points (Fukushige & Heggie 2000; Baumgardt 2001; Claydon et al. 2019). Once stars escape, their low velocity dispersion means that unbound stars can remain near the cluster (Küpper, MacLeod & Heggie 2008; Küpper, Lane & Heggie 2012; Webb et al. 2013). These unbound but nearby stars may still be included in the SED fit and radius fits. The observed mass of the cluster is therefore not necessarily the bound mass of the cluster, which is what the models present.

From all of this, it is clear that more work needs to be done to understand how cluster evolution models can be used to interpret this data set. The scaling relation presented earlier in equation (25) indicates that high-mass clusters may be less likely to be tidally bound, but beyond that we refrain from drawing definitive conclusions about cluster evolution.

## 5 CONCLUSIONS

We implemented a custom pipeline to measure the projected half-light radius  $R_{\text{eff}}$  of star clusters. This pipeline has several features designed to make it robust to contamination and accurately estimate the local background, producing reliable values of  $R_{\text{eff}}$  (Section 2). We applied this pipeline to the star clusters in 31 LEGUS galaxies, producing a uniformly measured sample of 7242 star cluster radii.

Of these, we identify 6097 as having reliable radii. This is currently the largest such catalogue of star cluster radii available.

We summarize the key results below:

(i) Most, but not all, galaxies share a common cluster radius distribution, with a peak at around 3 pc (Fig. 10). The shape of this distribution is asymmetric, with a tail to small radii, and is well described by a Weibull distribution (equation 19).

(ii) We find a clear but shallow mass–radius relation (Figs 11 and 13). This relation takes the form  $R_{\text{eff}} \propto M^{0.24}$ , with an intrinsic scatter of 0.25 dex (Table 2).

(iii) This mass–radius relation is present in clusters of all ages probed by LEGUS sample (Fig. 12). The slope of this relation is shallower at 1–10 Myr than at later times, but the slope does not evolve between the 10–100 Myr and 100 Myr–1 Gyr bins. The intrinsic scatter decreases with time. Selection effects cause the subsamples of different age to span different mass ranges, complicating interpretation (Section 4.1).

(iv) The majority of clusters identified in LEGUS are gravitationally bound (Fig. 14). The majority of unbound clusters are younger than 10 Myr and tend to be less massive (Fig. 15).

(v) The distributions of both average density and surface density of LEGUS clusters are well fitted by lognormal distributions (Fig. 16). The width of these distributions is large but decreases with cluster age. The peaks of the distributions for the youngest clusters are  $\rho_h \approx 30 M_\odot \text{ pc}^{-3}$  and  $\Sigma_h \approx 100 M_\odot \text{ pc}^{-2}$  (Table 3).

(vi) While we do not directly calculate the Jacobi radii for the LEGUS clusters, the shallow mass–radius relation implies that high-mass clusters fill less of their Jacobi radius than low-mass clusters (equation 25).

(vii) We create simple toy models of cluster evolution based on the model in Gieles & Renaud (2016) to interpret the trends we see with cluster age (Fig. 17). None of the models can successfully reproduce all aspects of the observed distributions.

## ACKNOWLEDGEMENTS

This research was supported in part by the US National Science Foundation through grants 1909063 and PHY-1748958. We are grateful to the Kavli Institute for Theoretical Physics in Santa Barbara for sponsoring an online programme *Globular Clusters at the Nexus of Star and Galaxy Formation* in 2020 May–June, which inspired this research.

We thank Arya Farahi for fruitful discussions that shaped our approach to incorporating selection effects in the hierarchical Bayesian model as described in Appendix A. We also thank the referee for suggestions that improved the paper.

## DATA AVAILABILITY

Cluster catalogues are available on Github at <https://gillenbrown.github.io/LEGUS-sizes/>. These catalogues include all quantities needed to reproduce the plots in this paper, including the fit parameters for each cluster. This repository also holds all the code used to fit the radii and generate the figures in this paper. The catalogues include properties such as mass and age that are originally from the LEGUS catalogues, which are found at <https://archive.stsci.edu/prepds/legus/dataproducts-public.html>.

## REFERENCES

Adamo A., Östlin G., Zackrisson E., Hayes M., Cumming R. J., Micheva G., 2010, *MNRAS*, 407, 870

Adamo A. et al., 2017, *ApJ*, 841, 131

Adamo A. et al., 2020, *Space Sci. Rev.*, 216, 69

Anderson J., King I. R., 2000, *PASP*, 112, 1360

Astropy Collaboration, 2013, *A&A*, 558, A33

Astropy Collaboration, 2018, *AJ*, 156, 123

Ballesteros-Paredes J. et al., 2020, *Space Sci. Rev.*, 216, 76

Bastian N., Gieles M., Lamers H. J. G. L. M., Scheepmaker R. A., de Grijs R., 2005, *A&A*, 431, 905

Bastian N. et al., 2012, *MNRAS*, 419, 2606

Bastian N., Cabrera-Ziri I., Davies B., Larsen S. S., 2013, *MNRAS*, 436, 2852

Baumgardt H., 2001, *MNRAS*, 325, 1323

Bell E. F., de Jong R. S., 2001, *ApJ*, 550, 212

Bertin E., Arnouts S., 1996, *A&AS*, 117, 393

Bland-Hawthorn J., Gerhard O., 2016, *ARA&A*, 54, 529

Bothwell M. S., Kennicutt R. C., Lee J. C., 2009, *MNRAS*, 400, 154

Bovy J., 2017, *MNRAS*, 468, L63

Bradley L. et al., 2019, Astropy/Photutils: v0.7.2

Calzetti D. et al., 2015, *AJ*, 149, 51

Cantat-Gaudin T. et al., 2018, *A&A*, 618, A93

Cardelli J. A., Clayton G. C., Mathis J. S., 1989, *ApJ*, 345, 245

Chandar R., Whitmore B. C., Dinino D., Kennicutt R. C., Chien L. H., Schinnerer E., Meidt S., 2016, *ApJ*, 824, 71

Choksi N., Kruijssen J. M. D., 2019, *MNRAS*, 507, 5492

Claydon I., Gieles M., Varri A. L., Heggie D. C., Zocchi A., 2019, *MNRAS*, 487, 147

Colombo D., Rosolowsky E., Ginsburg A., Duarte-Cabral A., Hughes A., 2015, *MNRAS*, 454, 2067

Cook D. O. et al., 2019, *MNRAS*, 484, 4897

Cuevas-Otahola B., Mayya Y. D., Puerari I., Rosa-González D., 2020, *MNRAS*, 492, 993

Cuevas-Otahola B., Mayya Y. D., Puerari I., Rosa-González D., 2021, *MNRAS*, 500, 4422

Dessauges-Zavadsky M. et al., 2019, *Nat. Astron.*, 3, 1115

Elson R. A. W., Fall S. M., Freeman K. C., 1987, *ApJ*, 323, 54

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, 125, 306

Fouesneau M. et al., 2014, *ApJ*, 786, 117

Fukushige T., Heggie D. C., 2000, *MNRAS*, 318, 753

Gieles M., Portegies Zwart S. F., 2011, *MNRAS*, 410, L6

Gieles M., Renaud F., 2016, *MNRAS*, 463, L103

Gieles M., Portegies Zwart S. F., Baumgardt H., Athanassoula E., Lamers H. J. G. L. M., Sipior M., Leenaarts J., 2006, *MNRAS*, 371, 793

Gieles M., Baumgardt H., Heggie D. C., Lamers H. J. G. L. M., 2010, *MNRAS*, 408, L16

Goodwin S. P., Bastian N., 2006, *MNRAS*, 373, 752

Grasha K., 2018, PhD thesis, Univ. Massachusetts

Grasha K. et al., 2019, *MNRAS*, 483, 4707

Hogg D. W., Bovy J., Lang D., 2010, preprint ([arXiv:1008.4686](https://arxiv.org/abs/1008.4686))

Holtzman J. A. et al., 1992, *AJ*, 103, 691

Jacobs B. A., Rizzi L., Tully R. B., Shaya E. J., Makarov D. I., Makarova L., 2009, *AJ*, 138, 332

Johnson L. C. et al., 2012, *ApJ*, 752, 95

Jordán A. et al., 2005, *ApJ*, 634, 1002

Kelly B. C., 2007, *ApJ*, 665, 1489

Kharchenko N. V., Piskunov A. E., Schilbach E., Röser S., Scholz R.-D., 2013, *A&A*, 558, A53

Kissler-Patig M., Jordán A., Bastian N., 2006, *A&A*, 448, 1031

Kourkchi E., Tully R. B., 2017, *ApJ*, 843, 16

Krumholz M. R. et al., 2015, *ApJ*, 812, 147

Krumholz M. R., McKee C. F., Bland-Hawthorn J., 2019, *ARA&A*, 57, 227

Küpper A. H. W., MacLeod A., Heggie D. C., 2008, *MNRAS*, 387, 1248

Küpper A. H. W., Lane R. R., Heggie D. C., 2012, *MNRAS*, 420, 2700

Larsen S. S., 1999, *A&AS*, 139, 393

Larson R. B., 1981, *MNRAS*, 194, 809

Lee J. C. et al., 2009, *ApJ*, 706, 599

Leitherer C. et al., 1999, *ApJS*, 123, 3

Li H., Gnedin O. Y., Gnedin N. Y., 2018, *ApJ*, 861, 107

Mackey A. D., Gilmore G. F., 2003a, *MNRAS*, 338, 85

Mackey A. D., Gilmore G. F., 2003b, *MNRAS*, 338, 120

McLaughlin D. E., van der Marel R. P., 2005, *ApJS*, 161, 304

Maíz Apellániz J., 2009, *ApJ*, 699, 1938

Meurer G. R., Heckman T. M., Leitherer C., Kinney A., Robert C., Garnett D. R., 1995, *AJ*, 110, 2665

Mok A., Chandar R., Fall S. M., 2021, *ApJ*, 911, 8

Olivares E. F. et al., 2010, *ApJ*, 715, 833

Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2002, *AJ*, 124, 266

Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2010, *AJ*, 139, 2097

Piskunov A. E., Schilbach E., Kharchenko N. V., Röser S., Scholz R. D., 2007, *A&A*, 468, 151

Portegies Zwart S. F., McMillan S. L. W., Gieles M., 2010, *ARA&A*, 48, 431

Roman-Duval J., Jackson J. M., Heyer M., Rathborne J., Simon R., 2010, *ApJ*, 723, 492

Ryon J. E. et al., 2015, *MNRAS*, 452, 525

Ryon J. E. et al., 2017, *ApJ*, 841, 92

Sabbi E. et al., 2018, *ApJS*, 235, 23

Scheepmaker R. A., Haas M. R., Gieles M., Bastian N., Larsen S. S., Lamers H. J. G. L. M., 2007, *A&A*, 469, 925

Schlafly E. F., Finkbeiner D. P., 2011, *ApJ*, 737, 103

Silva-Villa E., Larsen S. S., 2011, *A&A*, 529, A25

Silva-Villa E., Adamo A., Bastian N., Fouesneau M., Zackrisson E., 2014, *MNRAS*, 440, L116

Spitzer Lyman J., 1958, *ApJ*, 127, 17

Sun J. et al., 2018, *ApJ*, 860, 172

Tully R. B., Courtois H. M., Sorce J. G., 2016, *AJ*, 152, 50

Urquhart J. S. et al., 2018, *MNRAS*, 473, 1059

Vázquez G. A., Leitherer C., 2005, *ApJ*, 621, 695

Wagner-Kaiser R., Sarajedini A., Dalcanton J. J., Williams B. F., Dolphin A., 2015, *MNRAS*, 451, 724

Webb J. J., Harris W. E., Sills A., Hurley J. R., 2013, *ApJ*, 764, 124

Whitmore B. C. et al., 2010, *AJ*, 140, 75

Zackrisson E., Rydberg C.-E., Schaefer D., Östlin G., Tuli M., 2011, *ApJ*, 740, 13

Zepf S. E., Ashman K. M., English J., Freeman K. C., Sharples R. M., 1999, *AJ*, 118, 752

## APPENDIX A: METHODS FOR FITTING MASS–RADIUS RELATION

In the main text, we use the orthogonal fitting method described in sections 7 and 8 of Hogg et al. (2010) (hereafter H10), which we summarize here. In evaluating the Gaussian likelihood of each data point given a linear relation (with parameters of slope  $m$  and intercept  $b$ ), we use the displacements and variances projected perpendicular to the line being evaluated. The projected displacement is given by equation (30) of H10, and can also be written as

$$\Delta_i = \frac{y_i - (mx_i + b)}{\sqrt{1 + m^2}}, \quad (A1)$$

where in our case  $x$  is the log of the mass,  $y$  is the log of the effective radius,  $\beta = m$ , and we can calculate  $b$  given  $\beta$  and our pivot point  $R_4$ .

The projected variance is given by equation (31) of H10. In our case, the mass and radius errors are independent, so the off-diagonal terms of the covariance matrix are zero, allowing us to simplify that expression:

$$\sigma_i^2 = \frac{m^2 \sigma_{x,i}^2 + \sigma_{y,i}^2}{1 + m^2}. \quad (A2)$$

Then, we add an intrinsic scatter  $\sigma_{\text{int}}$  orthogonal to the line to the data variance, giving the likelihood for a single data point of

$$\mathcal{L}_i = \frac{1}{\sqrt{2\pi(\sigma_i^2 + \sigma_{\text{int}}^2)}} \exp\left(-\frac{\Delta_i^2}{2(\sigma_i^2 + \sigma_{\text{int}}^2)}\right). \quad (\text{A3})$$

The total data likelihood is the product of this over all data points. In the main text, we maximize this likelihood to produce final parameter values.

This method does not incorporate any selection effects. As these do exist in the LEGUS sample, we implemented an additional method to attempt to incorporate those selection effects. We use a hierarchical Bayesian model, following Kelly (2007). Each cluster has its observed mass and radius ( $M_{\text{obs},i}$ ,  $R_{\text{eff},i}$ ) with corresponding unobserved true quantities ( $m_i$ ,  $r_i$ ). We use the same relation as the main text (equation 20), but define it using the unobserved true quantities rather than the observed values:

$$\hat{r}_i(m_i) = r_4 \left( \frac{m_i}{10^4 M_{\odot}} \right)^{\beta} \quad (\text{A4})$$

such that the normalizing factor  $r_4$  is the underlying effective radius at  $10^4 M_{\odot}$ . In addition, we include an intrinsic lognormal scatter  $\sigma_{\text{int}}$ .

For a given cluster, our data likelihood takes the general form:

$$\begin{aligned} P(R_{\text{eff},i}, M_{\text{obs},i}, r_i, m_i | r_4, \beta, \sigma_{\text{int}}) &= P(R_{\text{eff},i} | r_i) \\ &\times P(M_{\text{obs},i} | m_i) \\ &\times P(r_i | m_i, r_4, \beta, \sigma_{\text{int}}) \\ &\times P(m_i), \end{aligned} \quad (\text{A5})$$

where the first two terms are the likelihoods of the observed values given the unobserved true values (which we treat as independent), the third term is the mass–radius relation, and the final term is the prior on the true mass. We model the radius distribution at a given mass as a lognormal distribution:

$$P(r_i | m_i, r_4, \beta, \sigma_{\text{int}}) = \frac{1}{\sigma_{\text{int}} \sqrt{2\pi}} \exp\left[-\frac{1}{2} \left( \frac{\log r_i - \log \hat{r}_i(m_i)}{\sigma_{\text{int}}} \right)^2\right], \quad (\text{A6})$$

where  $\log \hat{r}_i(m_i)$  is from equation (20). The normalizing factor here is important, as it includes a variable of interest  $\sigma_{\text{int}}$ .

We treat the mass and radius measurement errors as independent lognormal variables, with a width equal to the symmetrized observational uncertainties:

$$P(R_{\text{eff},i} | r_i) \propto \exp\left[-\frac{1}{2} \left( \frac{\log r_i - \log R_{\text{eff},i}}{\sigma_{R_{\text{eff},i}, \text{err}}} \right)^2\right] \quad (\text{A7})$$

$$P(M_{\text{obs},i} | m_i) \propto \exp\left[-\frac{1}{2} \left( \frac{\log m_i - \log M_{\text{obs},i}}{\sigma_{M_{\text{obs},i}, \text{err}}} \right)^2\right]. \quad (\text{A8})$$

This allows us to analytically marginalize over the unobserved radius  $r_i$ .

$$\begin{aligned} P(R_{\text{eff},i} | m_i, r_4, \beta, \sigma_{\text{int}}) &= \int P(R_{\text{eff},i} | r_i) P(r_i | m_i, r_4, \beta, \sigma_{\text{int}}) dr_i \\ &= \frac{1}{\sqrt{2\pi(\sigma_{\text{int}}^2 + \sigma_{R_{\text{eff},i}, \text{err}}^2)}} \\ &\times \exp\left[-\frac{1}{2} \frac{(\log R_{\text{eff},i} - \log \hat{r}_i(m_i))^2}{\sigma_{\text{int}}^2 + \sigma_{R_{\text{eff},i}, \text{err}}^2}\right] \end{aligned} \quad (\text{A9})$$

We also need to include the selection effects. There are two key selection variables: radius and  $V$ -band absolute magnitude.

To prevent contamination from unresolved sources, LEGUS selects clusters by examining the CI, the magnitude difference between

3 pixel and 1 pixel apertures. A hard boundary is drawn: anything above this is a cluster and anything below this is a star (Adamo et al. 2017). The  $V$ -band absolute magnitude cut is simple: selected clusters have absolute  $V$ -band magnitude brighter than  $-6$ . Unfortunately, this causes selection effects as a function of both mass and age, as dying massive stars go away, making clusters fade with age. To properly account for this, we need to include both age and  $V$ -band magnitude into our analysis. In what follows, we will use  $T$ ,  $\tau$  for observed and true ages, respectively, and  $V$ ,  $v$  for observed and true  $V$ -band magnitudes.

To renormalize our likelihood function, we need to determine the likelihood of a given cluster of a given true mass and age being selected:

$$\Phi_i(m_i, \tau_i, r_4, \beta, \sigma_{\text{int}}) = \int f(R_{\text{eff}}) f(V) P(R_{\text{eff}} | m_i, r_4, \beta, \sigma_{\text{int}}) P(V | m_i, \tau_i) dR_{\text{eff}} dV, \quad (\text{A10})$$

where  $f(R_{\text{eff}})$  and  $f(V)$  are the selection functions. For  $V$ , this is a simple step function:

$$f(V) = \begin{cases} 1 & \text{for } V < -6 \\ 0 & \text{for } V \geq -6. \end{cases} \quad (\text{A11})$$

For radius, this selection function is more complicated. Adamo et al. (2017) made a first attempt at this in the NGC 628c field, finding that clusters with effective radii of 2 pc were entirely recovered, while those with 1 pc were recovered roughly 50 per cent of the time. For this test, we simply represent the selection probability as

$$f(R_{\text{eff}}) = \max\left(\frac{R_{\text{eff}}}{0.05 \text{ arcsec}}, 1\right), \quad (\text{A12})$$

where 0.05 arcsec is roughly 2 pc at the distance of NGC 628. This gives a 100 per cent selection at 2 pc and 50 per cent selection at 1 pc. While this functional form is likely inaccurate, quantifying the selection effects in more detail is beyond the scope of this paper.

The final two terms in equation (A10) are the likelihood of a given radius being selected at that mass (given by equation A9) and the likelihood of observing a given  $V$ -band magnitude given a cluster's mass and age. We obtain these using the Yggdrasil models matching those used in LEGUS (Zackrisson et al. 2011). We can then represent the  $V$  band term as a normal distribution:

$$P(V_i | m_i, \tau_i) \propto \exp\left[-\frac{1}{2} \left( \frac{V_i - v(m_i, \tau_i)}{\sigma_{V, \text{err}}} \right)^2\right], \quad (\text{A13})$$

where  $v(m_i, \tau_i)$  is the expected  $V$ -band magnitude obtained from Yggdrasil and  $\sigma_{V, \text{err}}$  is the observed uncertainty in  $V$ -band magnitude. Note that since this term is for an arbitrary cluster with true mass  $m$  and true age  $\tau$ , we use typical errors of 0.04 mag.

We use this selection probability to renormalize the likelihood.  $\Phi_i$  is calculated separately for each cluster and used to renormalize that cluster's likelihood. In addition, we need to incorporate  $\tau$  into the likelihood as it enters the selection effects:

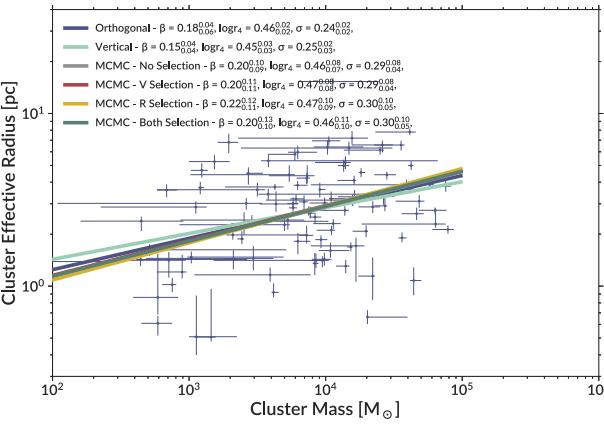
$$P(R_{\text{eff},i}, M_{\text{obs},i}, T_i, m_i, \tau_i | r_4, \beta, \sigma_{\text{int}}) = \quad (\text{A14})$$

$$\Phi_i^{-1}(m_i, \tau_i, r_4, \beta, \sigma_{\text{int}}) \quad (\text{A15})$$

$$\times P(R_{\text{eff},i} | m_i, r_4, \beta, \sigma_{\text{int}}) \quad (\text{A16})$$

$$\times P(M_{\text{obs},i} | m_i) P(T_i | \tau_i) P(m_i) P(\tau_i). \quad (\text{A17})$$

We treat the age in the same way we treat the mass and radius errors, as a lognormal distribution with a width equal to the symmetrized



**Figure A1.** The mass–radius relation fitted by several methods for a random sample of 100 clusters.

observational error:

$$P(T_i|\tau_i) \propto \exp \left[ -\frac{1}{2} \left( \frac{\log \tau_i - \log T_i}{\sigma_{T,\text{err}}} \right)^2 \right]. \quad (\text{A18})$$

As this likelihood is for one cluster, the total likelihood is the product of the likelihood for all clusters. We also use Bayes' Theorem to turn this into posterior likelihoods on our parameters:

$$P(r_4, \beta, \sigma_{\text{int}} | R_{\text{eff},i}, M_{\text{obs},i}, m_i) \propto P(r_4)P(\beta)P(\sigma_{\text{int}}) \times \prod_i P(R_{\text{eff},i}, M_{\text{obs},i}, m_i | r_4, \beta, \sigma_{\text{int}}). \quad (\text{A19})$$

We use flat priors on all parameters. The slope  $\beta$  is uniform between  $-1$  and  $1$ , the normalization  $r_4$  is uniform between  $0.01$  and  $100$  pc, and the intrinsic scatter  $\sigma$  is uniform between  $0$  and  $1$  dex.

We sample this posterior distribution using the EMCEE implementation of MCMC (Foreman-Mackey et al. 2013). This allows us to easily marginalize over  $m_i$  and  $\tau_i$  in post-processing of the MCMC chain.

As this implementation requires evaluating the integral in equation (A10) for each cluster at each step of the MCMC chain, it is computationally expensive, and scales with the number of clusters. We therefore tested this method using a random sample of only 100 clusters. Fig. A1 shows the fit parameters for regular least squares, our fiducial orthogonal least-squares method, and several variations of this method with different selection functions enabled.

This hierarchical Bayesian MCMC model produced larger error bars, but otherwise the results are consistent with those of the orthogonal fit. Additionally, removing the selection function terms from the fit does not change the result. Because of this, we decided to not use the hierarchical Bayesian method and instead use the simpler orthogonal fit described at the beginning of this appendix. Another reason to use the orthogonal fit is that we do not know the true selection effects in LEGUS, which would be needed to do a proper analysis of their impact. A full accounting of these is beyond the scope of this paper, making the functional form we assumed for the selection effects overly simplistic. We therefore choose to use the orthogonal fit as our method of choice in this paper.

This paper has been typeset from a TeX/LaTeX file prepared by the author.