

Sub-millimetre galaxies with Webb

Near-infrared counterparts and multi-wavelength morphology

Steven Gillman^{1,2}, Bitten Gullberg^{1,2}, Gabe Brammer^{1,3}, Aswin P. Vijayan^{1,2}, Minju Lee^{1,2}, David Blánquez^{1,2}, Malte Brinch^{1,2}, Thomas R. Greve^{1,2,4}, Iris Jermann^{1,2}, Shuowen Jin^{1,2}, Vasily Kokorev⁵, Lijie Liu^{1,2}, Georgios Magdis^{1,2,3}, Francesca Rizzo^{1,3}, and Francesco Valentino^{1,3,6}

¹ Cosmic Dawn Center (DAWN), Denmark

² DTU-Space, Elektrovej, Building 328, 2800, Kgs., Lyngby, Denmark

e-mail: srigi@space.dtu.dk

³ Niels Bohr Institute, University of Copenhagen, Jagtvej 128, 2200 Copenhagen N, Denmark

⁴ Dept. of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK

⁵ Kapteyn Astronomical Institute, University of Groningen, PO Box 800 9700 AV Groningen, The Netherlands

⁶ European Southern Observatory, Karl-Schwarzschild-Str. 2, 85748 Garching bei München, Germany

Received 28 March 2023 / Accepted 6 June 2023

ABSTRACT

We utilised the unprecedented depth and resolution of recent early-release science (ERS) JWST observations to define the near-infrared counterparts of sub-millimetre galaxies (SMGs). We identified 45 SCUBA-2 SMG positions within the Cosmic Evolution Early Release Science (CEERS) survey JWST/NIRCam fields. Through an analysis of multi-wavelength p -values, NIRCam colours and predicted SCUBA-2 fluxes, we define 43 JWST/NIRCam counterparts to the SCUBA-2 SMGs, finding a 63 per cent agreement with those identified in prior HST studies. Using EaZy-py, we fitted the available HST and JWST observations to quantify the photometric redshifts of the NIRCam-SMGs, establishing a broad range of redshift from $z \approx 0.2$ –5.4 with a median of $z \approx 2.29$, in agreement with other studies of SMGs. We identified significant variations in the morphology of the NIRCam-SMGs from isolated discs and spheroidal galaxies to irregular interacting systems. We analysed their rest-frame optical and near-infrared morphological properties (e.g. effective radius (R_e), Sérsic index (n), concentration (C), asymmetry (A), clumpiness (S), as well as the Gini and M_{20} parameters), finding, on average, late-type disc-like morphologies with large scatter into the intermediate and merger regions of the non-parametric parameter space. For the non-merging galaxies, we find a median rest-frame optical size and Sérsic index (and 1σ scatter) of $R_e = 3.10 \pm 1.67$ kpc and $n = 0.96 \pm 0.66$. Whilst in the rest-frame near-infrared, we establish more compact, higher Sérsic index morphologies ($R_e = 1.64 \pm 0.97$, $n = 1.85 \pm 0.63$). We further establish that both the rest-frame optical and near-infrared effective radii correlate negatively (at a 2σ level) with redshift, whilst the Sérsic index remains constant with cosmic time. Our results are consistent with the picture of inside-out galaxy evolution, with more centrally concentrated older stellar populations, and more extended, younger star-forming regions whose stellar emission is heavily attenuated in the central regions.

Key words. galaxies: high-redshift – galaxies: structure – galaxies: evolution – submillimeter: galaxies

1. Introduction

Sub-millimetre galaxies (SMGs), defined by their excess ($\gtrsim 1$ –2 mJy) sub-millimetre (450, 850 μ m) flux, were first identified in ground-based wide-field single-dish observations from telescopes such as the Sub-millimetre Common-User Bolometer Array (SCUBA; e.g. Smail et al. 1997; Eales et al. 1999)¹. In the last few decades, continued analysis has shown these sub-millimetre bright systems constitute a unique, massive ($\sim 10^{11} M_\odot$; Wardlow et al. 2011; Simpson et al. 2014) galaxy population that is heavily dust-obscured and far-infrared (FIR) luminous, commonly exhibiting intense star formation rates ($\sim 10^2$ – $10^3 M_\odot$ yr $^{-1}$; Blain et al. 2002; Chapman et al. 2005; Swinbank et al. 2014).

Reproducing these extremely dusty, massive systems in simulations has been a long-standing problem (e.g. Baugh et al. 2005; Davé et al. 2010; McAlpine et al. 2019). Oftentimes sim-

ulations have to be tuned to reproduce the sub-millimetre number counts and redshift distribution, invoking novel modelling assumptions (e.g. top-heavy initial mass function; Cowley et al. 2019). More recent simulations with complex dust models or sub-millimetre flux scaling relations (e.g. Lovell et al. 2021; Hayward et al. 2021) have had more success and established that the sub-millimetre bright phases are a natural consequence of massive galaxy evolution in a Λ CDM universe (e.g. Long et al. 2022; Lower et al. 2023; Cochrane et al. 2023). Despite these recent successes, constraining the multi-wavelength properties of sub-millimetre bright galaxies across cosmic time still poses many observational challenges due to their inherent faintness at optical to infrared wavelengths (e.g. Dey et al. 1999; Weiß et al. 2009; Hodge et al. 2012; Smail et al. 2021).

To ascertain the cosmic evolution of SMGs, their full multi-wavelength (UV to far-infrared) properties need to be quantified. In the far-infrared, wide-field single-dish observations, although often incredibly deep (1σ depth of 2 mJy beam $^{-1}$; Smail et al. 1997) which is crucial for identifying SMGs, suffer from large beam sizes and poor resolution. For instance, the Sub-millimetre

¹ We note SMGs are often also classified as IR-selected dusty star-forming galaxies (DSFG; see Casey et al. 2014 for review), but here we focus on the original single-dish selected definition of SMGs.

Common-User Bolometer Array 2 (SCUBA-2; Dempsey et al. 2013) at 850 μm has a 14.5'' full width half maximum (FWHM; Holland et al. 2013). Thus identifying the individual multi-wavelength counterparts to the sub-millimetre selected galaxies is very challenging (e.g. Ivison et al. 2007; Biggs et al. 2011; Chen et al. 2016; An et al. 2018). This is compounded at optical–near-infrared wavelengths where SMGs are inherently faint due to their extremely dusty nature, and telescopes such as the *Hubble* Space Telescope (HST) provide sub-arcsecond resolution revealing a multitude of potential counterparts (e.g. Zavala et al. 2018; Ling & Yan 2022; Shim et al. 2022).

In recent years, high-resolution continuum observations from the Atacama Large Millimetre/sub-millimetre Array (ALMA) have followed up on single-dish surveys, identifying individual galaxy counterparts at sub-millimetre wavelengths (see Hodge & da Cunha 2020 for a review of ALMA SMG studies). A large fraction of these observations have identified that SMGs commonly reside in over-dense regions (e.g. Simpson et al. 2014; Hodge et al. 2016; Miller et al. 2018; Oteo et al. 2018), with 10–80% of bright single-dish sub-millimetre sources being found to be compromised of two or more individual SMGs when observed at sub-mJy rms depths (e.g. Karim et al. 2013; Simpson et al. 2014; Stach et al. 2018; An et al. 2019). A single bright sub-millimetre source, however, can be comprised of multiple SMGs over a broad range of redshifts. Both the negative K -correction, which effectively fixes the luminosity as a function of redshift (e.g. Hill et al. 2018), or alternatively strong gravitational lensing (e.g. Weiß et al. 2013; Díaz-Sánchez et al. 2017; Harrington et al. 2021) can contribute to this ‘line-of-sight’ blending.

Studies such as Hodge et al. 2016; Gullberg et al. 2019; Cochrane et al. 2021, which utilised sub-arcsecond resolution ($\leq 0''.2$) observations with ALMA, have revealed compact (\sim few kpc) dusty galaxies, often with observed-frame far-infrared disc-like morphologies (i.e. $n \approx 1$) and kinematics (e.g. Lelli et al. 2021; Rizzo et al. 2021). In contrast, rest-frame ultra-violet HST studies of $z \approx 1$ –3 SMGs identified extended ($R_{\text{dust}}/R_{\text{stellar}} \approx 0.6$) irregular morphologies with multiple components (e.g. Swinbank et al. 2010; Chen et al. 2015; Gómez-Guijarro et al. 2018; Zavala et al. 2018; Lang et al. 2019; Ling & Yan 2022).

Detecting and resolving SMGs’ rest-frame near-infrared emission, which reflects the bulk of the stellar population, has, however, previously been incredibly challenging. Observations with the Infrared Array Camera (IRAC; Fazio et al. 2004) on the *Spitzer* space telescope, whilst providing the required near-infrared wavelength coverage from 3.6–8 μm , lacks the spatial (sub-arcsecond) resolution required to constrain the underlying morphologies of the stellar emission in SMGs (e.g. Krick et al. 2021). With the advent of recent high-resolution observations from JWST, the near-infrared counterparts of sub-millimetre bright galaxies can be identified and their subsequent properties (e.g. stellar morphology) quantified. With JWST observations, we can robustly constrain the total mass budget of this unique galaxy population and constrain their evolution across cosmic time.

To this end, in this paper, we present an analysis of the infrared counterparts to SCUBA-2 450 and 850 μm selected SMGs detected in public JWST observations from the early-release science programme; The Cosmic Evolution Early Release Science (CEERS; programme ID:1345, PI: Finkelstein; Bagley et al. 2023)² survey. In Sect. 2, we define the sample of SCUBA-2 SMGs that have been observed with JWST and com-

pile the public multi-wavelength data. Through an analysis of the near-infrared NIRCam colours, multi-wavelength (850 μm , 24 μm , 4.44 μm) p -values, and predicted spectral energy distribution (SED) fluxes, we first identified the NIRCam counterparts to the SMGs in Sect. 3 before analysing their rest-frame multi-wavelength morphology in Sect. 4 and summarising our analysis in Sect. 5. Throughout the paper, we assume a Λ CDM cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, and $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$. All quoted magnitudes are on the AB system and stellar masses were calculated assuming a Chabrier initial mass function (IMF; Chabrier 2003).

2. Sample selection

To identify the near-infrared counterparts to the SMGs observed with JWST, we first compile a list of previously characterised SMGs in the literature, including samples from Zavala et al. 2017; Geach et al. 2017. We then cross-match this sample with the JWST/NIRCam observations from the Cycle 1 ERS programme CEERS, resulting in 45 SCUBA-2 pointings falling within the field of view of the NIRCam observations. This sample of 45 is comprised of both 850 μm and 450 μm SCUBA-2 pointings from the Zavala et al. (2017) sample³, as we describe in the next section and present in Table 1.

2.1. SCUBA-2 observations

We utilise the sample of SMGs presented in Zavala et al. (2017) which were observed with SCUBA-2 camera on the *James Clark Maxwell* Telescope (JCMT) between 2012 and 2015 as part of the SCUBA-2 Cosmic Legacy Survey (S2CLS; Geach et al. 2017) in the extragalactic deep field Extended Growth Strip (EGS; Davis et al. 2007). The SMG sample defined in Zavala et al. (2017) derive from the ‘deep tier’ observations of the survey, whilst the ‘wider tier’ observations are presented in Geach et al. (2017). For details of the observations, data reduction and SMG identification we refer the reader to Zavala et al. (2017) and references therein.

In short, observations were carried out at 450 and 850 μm in the EGS field with 8 and 14.5'' FWHM beams respectively. The 450 (850) positions, signal-to-noise (S/N) ratio and fluxes of the SMGs are reported in Appendix A of Zavala et al. (2017) and summarised in Table 1. We use these target positions and 450 (850) beam sizes as positions and search radii, respectively, within JWST/NIRCam observations to identify near-infrared counterparts to the SCUBA-2 SMGs. If the SCUBA-2 SMG is detected in both 450 and 850 μm observations, we use the 450 position and beam due to its smaller FWHM. The 850 μm selected SMGs have a deboosted median (and 16th–84th quartile) flux of $S_{850} = 1.62^{+0.81}_{-0.79} \text{ mJy beam}^{-1}$ whilst the 450 μm selected SMGs have a median (and 16th–84th quartile) deboosted flux of $S_{450} = 5.2^{+3.15}_{-1.50} \text{ mJy beam}^{-1}$.

2.2. HST and JWST/NIRCam observations

To identify the near-infrared counterparts of the 45 SMGs outlined above we analyse the CEERS Epoch 1 JWST/NIRCam imaging. The NIRCam imaging covers 34.7 sq. arcmin of the EGS field with observations taken in both short wavelength

³ We note six of the galaxies in our sample are also studied in Chen et al. (2022) with JWST/CEERS observations, where their selection is based on combined 450 and 850 μm detections and the Zavala et al. (2018) HST counterpart identification.

² <https://ceers.github.io>

Table 1. Summary of the SCUBA-2 SMGs identified in CEERS Epoch 1 NIRCam Imaging.

850 ID	450 ID	RA	Dec.	Deboosted S_{850} (mJy beam $^{-1}$)	S/N_{850}	Deboosted S_{450} (mJy beam $^{-1}$)	S/N_{450}	HST 1.6 μm	Radio 1.4 GHz
		SCUBA-2	SCUBA-2						
850.003	450.05	214.917	52.891	5.24 ± 0.31	23.2	13.3 ± 2.0	9.9	✓	✓
850.004	450.25	214.947	52.910	3.87 ± 0.31	17.0	6.0 ± 2.2	4.7	✓	
850.017	450.11	214.900	52.852	2.70 ± 0.37	10.1	9.7 ± 2.4	6.5	✓	
850.019	450.41	214.972	52.958	3.43 ± 0.51	9.6	7.3 ± 3.5	3.9	✓	
850.025	450.29	214.852	52.866	2.68 ± 0.42	9.0	6.6 ± 2.9	4.5	✓	
850.026	450.31	214.890	52.894	2.40 ± 0.38	8.8	5.9 ± 2.6	4.3		
850.028	450.69	214.878	52.853	2.45 ± 0.42	8.2	3.9 ± 2.5	3.2	✓	
850.030	450.10	214.879	52.877	2.14 ± 0.38	8.0	11.1 ± 2.4	7.2	✓	✓
850.031	450.49	214.823	52.873	2.87 ± 0.36	8.0	6.0 ± 3.3	3.8		
850.034	450.66	214.916	52.952	2.14 ± 0.40	7.5	4.3 ± 2.8	3.3		
850.038	450.27	214.865	52.899	2.08 ± 0.41	7.2	7.1 ± 2.9	4.6	✓	
850.043	450.73	214.950	52.937	1.68 ± 0.35	6.9	3.6 ± 2.4	3.1	✓	
850.059	450.24	214.857	52.849	1.62 ± 0.51	4.9	8.3 ± 3.2	4.8	✓	
850.065	450.18	214.877	52.867	1.35 ± 0.38	5.1	8.4 ± 2.5	5.5	✓	✓
850.070	450.34	214.970	52.930	1.15 ± 0.34	4.9	—	—	✓	
850.027	—	214.910	52.937	2.25 ± 0.36	8.7	—	—		
850.035	—	214.901	52.943	2.07 ± 0.39	7.5	—	—		
850.036	—	214.905	52.922	1.89 ± 0.36	7.3	—	—		
850.047	—	214.859	52.861	1.91 ± 0.42	6.5	—	—	✓	
850.054	—	214.874	52.843	1.95 ± 0.51	5.7	—	—	✓	
850.055	—	214.880	52.912	1.56 ± 0.39	5.6	—	—		
850.058	—	214.848	52.852	1.89 ± 0.51	5.6	—	—		
850.063	—	214.867	52.883	1.40 ± 0.39	5.2	—	—		
850.064	—	214.947	52.924	1.19 ± 0.33	5.2	—	—		
850.067	—	214.831	52.893	1.55 ± 0.46	5.1	—	—	✓	
850.068	—	214.893	52.933	1.30 ± 0.38	5.0	—	—		
850.075	—	214.835	52.869	1.26 ± 0.46	4.3	—	—	✓	
850.076	—	214.917	52.956	0.83 ± 0.50	3.3	—	—		
850.080	—	214.925	52.934	0.91 ± 0.36	4.0	—	—	✓	
850.086	—	214.945	52.929	0.60 ± 0.36	3.1	—	—		
850.091	—	214.900	52.838	0.94 ± 0.71	3.6	—	—		
850.093	—	214.930	52.920	0.72 ± 0.38	3.5	—	—		
850.098	—	214.923	52.928	0.69 ± 0.39	3.4	—	—		
850.099	—	214.944	52.946	0.73 ± 0.41	3.3	—	—		✓
850.101	—	214.838	52.859	0.83 ± 0.72	3.3	—	—		
850.108	—	214.866	52.871	0.68 ± 0.43	3.1	—	—		
—	450.35	214.904	52.863	—	—	5.2 ± 2.4	4.2		
—	450.48	214.910	52.927	—	—	4.9 ± 2.8	3.8	✓	
—	450.68	214.957	52.947	—	—	4.0 ± 2.6	3.2		
—	450.70	214.906	52.933	—	—	3.9 ± 2.5	3.2		
—	450.42	214.905	52.851	—	—	4.6 ± 2.5	3.8	✓	✓
—	450.77	214.844	52.894	—	—	3.8 ± 2.5	3.1		
—	450.81	214.934	52.906	—	—	3.1 ± 1.9	3.1		
—	450.88	214.875	52.888	—	—	3.4 ± 2.3	3.0		
—	450.89	214.916	52.852	—	—	3.1 ± 1.9	3.0		

Notes. The 850, 450 IDs, SCUBA-2 RA and Dec and 850, 450 S/N and deboosted fluxes (Cols. 1–8) are from Zavala et al. (2017). Column 9 indicates if an optical (HST/F160W) counterpart is identified in Zavala et al. (2018) and Col. 10 indicates if a 1.4 GHz radio counterpart is reported in Ivison et al. (2007).

($F115W$, $F200W$) and long wavelength ($F277W$, $F356W$, $F444W$) filters. The level-2 data products were downloaded from STScI portal⁴ and processed using Grizli pipeline (Brammer & Matharu 2021; Brammer et al. 2022) with routines to derive accurate NIRCam photometric zeropoints and correct for cosmic rays and stray light (e.g. Boyer et al. 2022; Nardiello et al. 2022; Bradley et al. 2022a). Further incorporat-

ing available optical and near-infrared HST observations from the Complete *Hubble* Archive for Galaxy Evolution (CHARGE; Kokorev et al. 2022), the final images were aligned to *Gaia* DR3 (Gaia Collaboration 2021) and drizzled (Fruchter & Hook 2002) to a resolution of 0.04'' per pixel.

Sources were then identified using SEP⁵ (Barbary et al. 2016), a python-wrapped Source Extractor (Bertin & Arnouts 1996)

⁴ <https://mast.stsci.edu>

⁵ <https://github.com/kbarbary/sep>

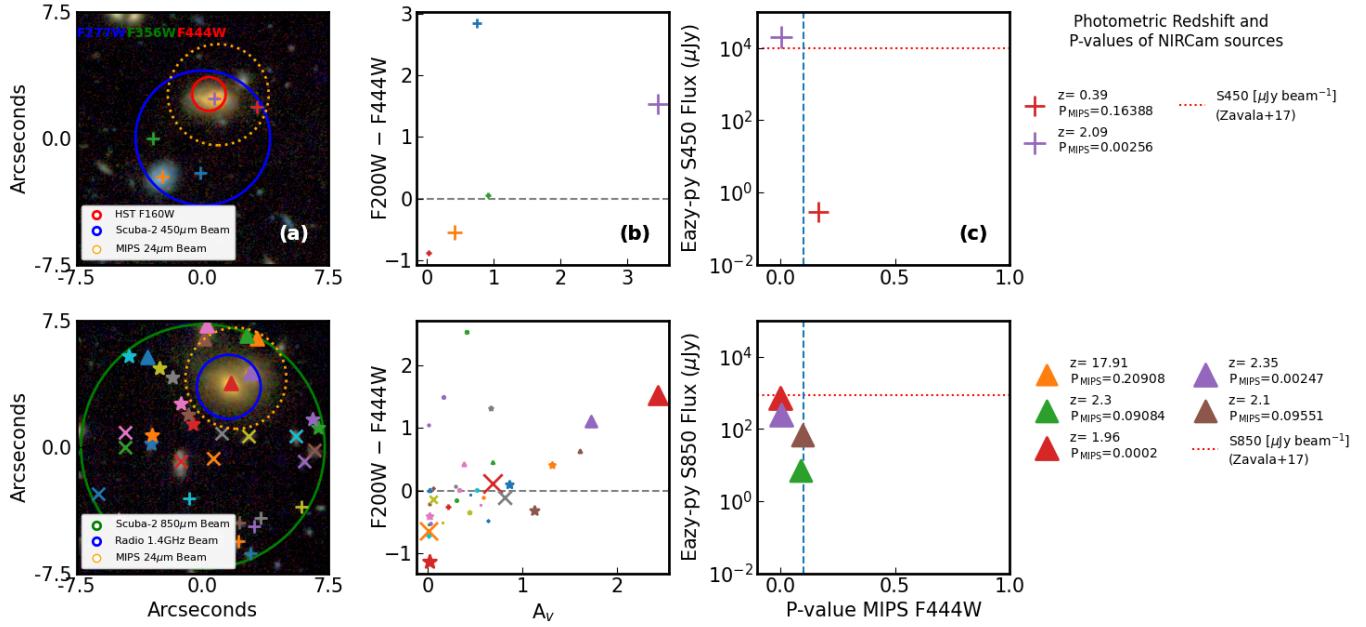


Fig. 1. Two examples of the multi-wavelength data used to identify the NIRCam-SMGs for 450 μm (upper) and 850 μm (lower) detected SCUBA-2 SMGs. We show a $F277W$, $F356W$, $F444W$ three colour NIRCam image centred on the SCUBA-2 position (a). In panel b we correlate the NIRCam $F200W - F444W$ colour with the EaZy-py A_V values for the NIRCam sources, where the marker size reflects the MIPS to $F444W$ p -value of the NIRCam source (lower p -value, larger marker) and the grey-dashed line indicates a colour of zero. Finally in panel c the EaZy-py predicted fluxes as a function of MIPS to $F444W$ p -values are shown. We also indicate the classical p -value < 0.1 threshold (vertical blue-dashed line) and observed 450 μm (850 μm) deboosted SCUBA-2 fluxes (red-dashed horizontal lines). NIRCam SMGs are identified, following the decision tree (Appendix A), by having high A_V , red colour, low p -values and EaZy-py predicted fluxes comparable to the SCUBA-2 fluxes in Zavala et al. (2017).

run on a noise-weighted combined long-wavelength channel NIRCam image ($F277W + F356W + F444W$). Aperture photometry was performed on each source using 0''.5 circular apertures corrected to ‘total’ values within an elliptical Kron aperture as defined in (Kron 1980)⁶. The aperture correction is derived on the long-wavelength detection image and applied to all bands. For the HST wavelength bands we apply an additional correction of order 10 per cent to correct for flux outside the Kron aperture, derived from curve of growth analysis of the PSFs⁷.

For each source identified in the CEERS Epoch 1 observations, we run the EaZy-py⁸ (Brammer et al. 2008) SED fitting code on the available HST and JWST photometry to derive photometric redshifts. We use thirteen templates from the Flexible Stellar Populations Synthesis code (FSPS; Conroy & Gunn 2010) described in Kokorev et al. (2022) linearly combined to allow for maximum flexibility. We note however given the uncertainties associated with SED fitting (e.g. Pacifici et al. 2023) we only focus on the photometric redshift, dust attenuation and 450 and 850 μm predicted fluxes for our analysis. Further defining the full SED of the NIRCam-SMGs will be presented in future papers. Examples of the EaZy-py SED fitting are given in Appendix B. We now use the positions, fluxes, and photometric redshift information for all the sources in CEERS NIRCam observations, in addition to the ancillary multi-wavelength data described in the next section, to identify the SMG counterparts.

2.3. Ancillary data

In addition to the JWST/NIRCam and SCUBA-2 observations detailed in prior sections, we take advantage of other publicly available multi-wavelength observations of the EGS field to aid the identification of the near-infrared counterparts of the SMGs. For each of the SCUBA-2 positions outlined in Table 1, we compile the available Radio 1.4 GHz data products from the All Wavelength Extended Groth Strip International Survey (AEGIS20; Ivison et al. 2007) and MIPS 24 μm data products (images and source catalogue) from the Far-Infrared Deep Extragalactic Legacy (FIDEL; PI: Dickinson & FIDEL Team 2007) survey. We identify five sources with 1.4 GHz counterparts to the sub-millimetre detection, within the SCUBA-2 (450/850) beam as listed in Table 1. All the SCUBA-2 SMG pointings are identified to have MIPS 24 μm imaging with the majority (40/45) having 24 μm sources identified in the SCUBA-2 beam, as derived by the FIDEL consortium⁹. Using this multi-wavelength dataset for each of the SCUBA-2 positions we proceed to first identify the near-infrared counterpart of the SMGs and then measure their multi-wavelength morphological properties.

3. Near-Infrared counterparts

Defining the near-infrared counterparts of the SMGs requires us to analyse all the available multi-wavelength data for each SCUBA-2 SMG. Firstly, for the SCUBA-2 SMG positions detailed in Table 1, we make $15'' \times 15''$ NIRCam $F444W$ cutouts centred on the 850 μm or 450 μm SCUBA-2 source position. As shown in Fig. 1, we then overlay all the potential NIRCam

⁶ The photometry catalogue and mosaics are available online [here](#).

⁷ See <https://www.stsci.edu/hst/instrumentation/wfc3/data-analysis/psf> for details.

⁸ <https://github.com/gbrammer/eazy-py>

⁹ <http://benjaminmagnelli.weebly.com/downloads.html>

sources which fall within the beam. Using a combination of p -values, near-infrared colours, EaZy-py predicted SCUBA-2 fluxes, and dust attenuation A_V , as described in the following sections, we define a decision tree to identify the near-infrared counterparts of the SMGs (see Appendix A).

3.1. NIRCam colours and dust attenuation

It is well known that SMGs will exhibit very red optical-to-infrared colours due to their extremely dusty nature (e.g. Dey et al. 1999; Weiß et al. 2009; Walter et al. 2012). Thus, to aid our identification of the SMGs in the JWST/NIRCam observations, we calculate the $F200W$ – $F444W$ colour for each potential counterpart within the SCUBA-2 beam. We correlate this NIRCam colour with the dust attenuation (A_V) from EaZy-py SED fitting as detailed in Sect. 2.2. Given the degeneracy between infrared colour and dust attenuation, we expect the SMGs’ NIRCam counterparts to occupy the top-right of this plane with ‘red’ NIRCam colour and high A_V value (e.g. Martis et al. 2022). We show an example of the correlation between $F200W$ – $F444W$ colour and A_V in Fig. 1.

3.2. Multi-wavelength p -values and predicted fluxes

Classically the SMG-counterpart selection is achieved through p -value calculations. This method is typically used to identify radio or far-infrared counterparts. In addition to the well-known FIR-radio correlation (e.g. Barger et al. 2014), the FIR source density is often low and the FWHM of the observations is \sim few arcsec, thus reducing the probability of source mis-identification. First introduced in Downes et al. (1986), the p -value corresponds to the probability the counterpart identified is a ‘by-chance’ association. The p -value is a function of the magnitude of the potential counterpart and its proximity to the beam centre, whereby brighter and closer counterparts will have lower p -values. Typically counterparts with $p(m, r) < 0.1$ are used as the selection criteria (e.g. Ivison et al. 2007; Biggs et al. 2011; Zavala et al. 2018).

For our analysis, we calculate the p -value to go from the SCUBA-2 450 μm (850 μm) beam to MIPS 24 μm imaging and then MIPS 24 μm to $F444W$ NIRCam imaging. For the five sources without MIPS coverage we calculate the p -value to go from SCUBA-2 450 μm (850 μm) to $F444W$ NIRCam directly. In addition to the p -values of the NIRCam sources, we can also use their predicted 450 μm (850 μm) SCUBA-2 fluxes from the EaZy-py code (Sect. 2.2), as further information in the counterpart identification¹⁰. For each SMG position, we plot the EaZy-py predicted SCUBA-2 450 μm (850 μm) fluxes as a function of the MIPS 24 μm to $F444W$ p -values, an example of which is shown in Fig. 1. We also overlay the observed SCUBA-2 450 μm (850 μm) fluxes from Zavala et al. (2017). We expect the NIRCam counterpart to the SCUBA-2 SMG to exhibit low p -values with SCUBA-2 fluxes comparable to those reported in Zavala et al. (2017). However, we do not use p -values as the only criteria for the NIRCam-SMG selection given the sub-arcsecond resolution of the JWST observations.

3.3. Near-infrared counterparts

To identify the near-infrared counterpart(s) in the NIRCam $F444W$ imaging, we construct a decision tree as shown in

Appendix A. The decision tree is designed to consider all of the available information for the SCUBA-2 positions, including the p -values, NIRCam colours, EaZy-py outputs (fluxes, A_V), and MIPS coverage to define the most likely SMG counterpart. To make this process more robust, our team have made independent selections of the 45 SMG counterpart(s) using the information shown in Fig. 1 and following the decision tree. We remove eight SCUBA-2 SMGs from the sample, due to poor multi-wavelength data resulting in inconclusive counterpart selection. From the remaining 37 SCUBA-2 SMGs, we identify 43 NIRCam counterparts with 1, 4, and 32 SCUBA-2 SMGs having 3, 2, and 1 NIRCam counterparts respectively. Of the 43 NIRCam counterparts, 23 have 100 per cent agreement in all independent selections. We label these quality 1 (Q1). Counterparts with $\geq 80\%$ agreement (12/43) are labelled quality 2 (Q2) and $< 80\%$ as quality 3 (Q3) (8/43). In the following analysis, we focus on Q1 and Q2 counterparts, but it is important to note that the inclusion of Q3 counterparts does not alter our conclusions. In Fig. 2 we show $6'' \times 6''$ NIRCam colour images at the positions of the SCUBA-2 SMGs with the NIRCam-SMGs labelled by source ID, which in a number of cases contain multiple NIRCam sources.

For just under half (19/43) of the SCUBA-2 SMGs in our sample, the optical counterparts have previously been identified in HST $F160W$ observations using similar analysis techniques (see Zavala et al. 2018). To validate our counterpart selection we compare to those identified in Zavala et al. (2018), finding agreement with 63% of previously identified sources. We note that given the increased depth and sensitivity of the JWST/NIRCam observations compared to HST/WFC3 used in the Zavala et al. (2018) selection, as well as the difference in observed wavelength (4.44 μm compared to 1.6 μm), our selection is more sensitive to optically faint, redder objects that are not identified in Zavala et al. (2018). Furthermore, studies indicate that a non-negligible fraction sub-mm detected sources are expected to have no counterpart below 3.6 μm (e.g. Franco et al. 2018; Manning et al. 2022) and thus be HST ($F160W$) dark.

In Fig. 3 we show the photometric redshift distribution of SCUBA-2 SMG NIRCam counterparts (hereafter NIRCam-SMGs) identifying a median redshift (and 16th–84th quartile) of $\bar{z} = 2.29^{+1.29}_{-0.94}$ which is comparable to that identified in Zavala et al. (2018) for 850 μm selected SMGs. The redshift distribution is also comparable to other studies of single-dish detected SMGs (e.g. Stach et al. 2019; Dudzevičiūtė et al. 2020) indicating the NIRCam-SMGs do not represent a new population of higher redshift SMGs, rather a population undetected by HST. For 15 NIRCam-SMGs we also have a spectroscopic redshift from archival observations. In the inset panel of Fig. 3 we show the correlation between spectroscopic and photometric redshift identifying a median ratio (and 16th–84th quartile) of $z_{\text{phot}}/z_{\text{spec}} = 1.0^{+0.10}_{-0.06}$. If a spectroscopic redshift is available for the NIRCam-SMG we adopt this redshift, as opposed to the EaZy-py photometric redshift, in all following analysis. In future papers, we will present both the complete SED properties of the NIRCam-SMGs (Lee & Vijayan et al. in prep.) and their close environment (Gullberg et al., in prep.). In the following section, we focus on quantifying the multi-wavelength morphology of the NIRCam-SMGs.

4. Multi-wavelength morphology

As shown in Fig. 2, the NIRCam-SMGs exhibit a diverse range of stellar morphology, from compact spheroidal galaxies to extended spiral structures and even potentially merging systems.

¹⁰ We note however, these can be uncertain when using EaZy-py for the FIR. Discussion of this will be presented in future papers with more sophisticated SED fitting codes.

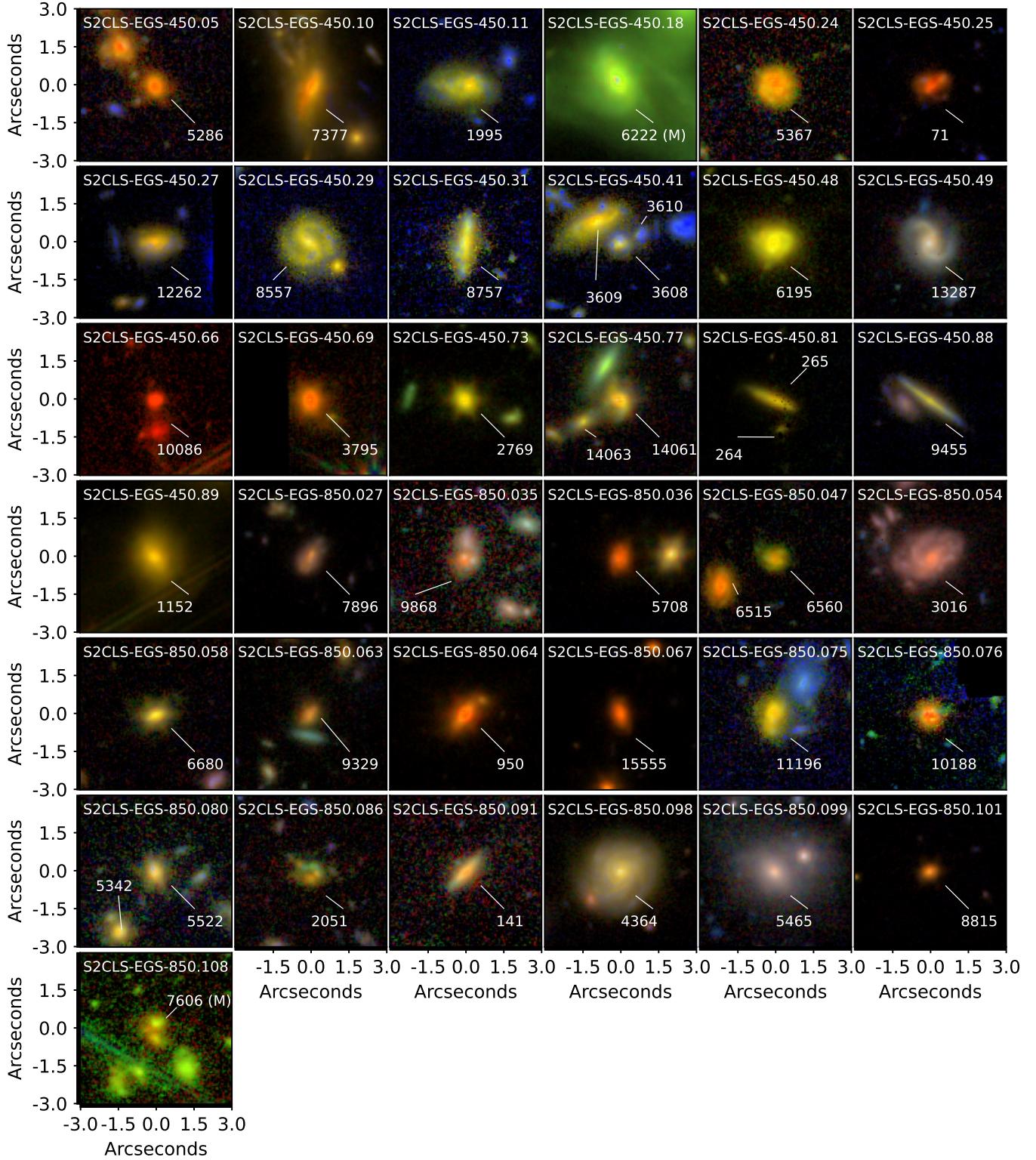


Fig. 2. $6'' \times 6''$ three-band ($F115W/F200W, F277W/F356W, F444W$) colour images for each SCUBA-2 SMGs with the NIRCam-SMGs labelled. In each image we indicate the ID of the NIRCam counterpart(s) with 1, 4, and 32 SCUBA-2 SMGs having 3, 2 and 1 NIRCam counterparts respectively. It is clear that the sample of NIRCam-SMGs contains a broad range of morphologies from isolated disc and spheroidal galaxies to disturbed interacting systems. We indicate the 11 galaxies classified non-parametrically (and visually) as Mergers from Fig. 5 with an (M).

As a first attempt to quantify the morphology of the galaxies, we employ a non-parametric approach, in which no prior assumptions about the galaxies' structure are made.

Throughout this section, the analysis is carried out on $0''.04$ per pixel 10 arcsec cutouts in each filter centred on the NIRCam-

SMG identified in Sect. 3¹¹. We use the segmentation map

¹¹ This cutout size is chosen to optimise the background estimation and modelling of the sources whilst minimising contamination from neighbours and computation time.

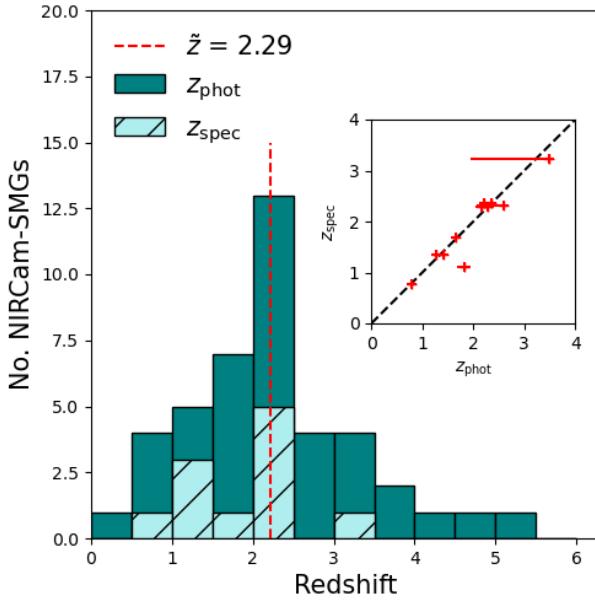


Fig. 3. Photometric redshift distribution from EaZy-py for the NIRCam counterparts, with a median redshift of $\bar{z} = 2.29$ and range from $z \approx 0.21\text{--}5.4$. The inset shows the spectroscopic redshift, where available, compared to the photometric redshift (and 16th–84th percentiles from the $p(z)$). We derive a median ratio (and 16th–84th percentiles) of $z_{\text{phot}}/z_{\text{spec}} = 1.0^{+0.10}_{-0.06}$.

generated from SEP (Sect. 2.2) to mask other sources in the cutout, which we dilate using the binary dilation routine in `photutils` (Bradley et al. 2022b) to ensure full masking of contaminants. Unmasked contaminants would alter the non-parametric measures of morphology, thus we visually inspect the galaxies growth curves to ensure full masking. We further use `photutils` to model (and remove) the background level in each cutout as well as to quantify the root mean-square (rms) noise. In the following sections ‘cutout’ refers to this 10-arcsec, background subtracted masked cutout and is used in the morphological analysis that follows.

Prior to measuring the morphology of the galaxies we derive the point spread function (PSF) in each NIRCam and HST band. For NIRCam we first use `WebbPSF` (Perrin et al. 2014) to generate PSF models across the NIRCam detectors for both short- and long-wavelength channels which are calibrated with wavefront models at the epoch of the CEERS observations. For HST we use well-defined PSF models for WFC3 incorporated into `grizli`¹². The PSFs for each filter are then inserted into individual exposures of the final mosaic and drizzled to the final world coordinate system (wcs) solution. For each NIRCam-SMG we evaluate the PSF model at the position of the galaxy, resulting in a position-dependent PSF that accurately resembles the PSF of the final mosaics used in our analysis and accounts for the detector level variations.

4.1. Non-parametric morphology

To constrain the multi-wavelength morphology of the NIRCam-SMGs, we first derive the galaxies’ non-parametric morphology via two independent methods. We first perform a curve of growth analysis in each of the NIRCam and HST WFC3 wavelength bands above 1 μm for the 43 NIRCam-SMGs. This is achieved

by fitting a Gaussian profile to the cutout of each galaxy, allowing the centroid (x, y), axis ratio (b/a) and position angle (PA) to vary. We note the original centroid of the cutout is derived from the SEP source detection on the stacked long-wavelength NIRCam bands (Sect. 2.2), and thus may not be the true centre of the galaxy at shorter (bluer) wavelengths. A curve of growth is then derived in each band using ellipses which align to the galaxies axis ratio (b/a) and position angle (PA).

From the curve of growth, examples of which are shown in Fig. 4, we measure the convolved 20, 50, and 80 per cent radii of each galaxy. The intrinsic radii of the galaxies’ are derived by de-convolving the sizes with the PSF size in each band, measured through a similar curve of growth analysis. To quantify the uncertainty on the intrinsic radius of a galaxy, we perform bootstrapping over the rms noise of the cutout with 1000 iterations. The final intrinsic 20, 50, and 80% radii (and uncertainty) of each galaxy are defined as the median (and standard deviation) of these 1000 iterations.

We derive a median (and 16th–84th quartile) axis ratio of $b/a = 0.67^{+0.21}_{-0.14}$ in the NIRCam $F444W$ band. This is comparable to the axis ratio of $b/a \sim 0.6$ expected for a population of triaxial ellipsoids and higher than that expected for randomly orientated discs with exponential light profiles (e.g. Law et al. 2012; van der Wel et al. 2014a; Zhang et al. 2019; Robertson et al. 2023; Kartaltepe et al. 2023). The dust (850 μm) morphology of SMGs have also been shown to exhibit similar axis ratios, with Gullberg et al. (2019) reporting a median axis ratio of $b/a = 0.62$. The median half-light radius in $F444W$ band is $R_h = 3.26^{+1.51}_{-1.19}$ kpc, which is comparable to that derived from similar studies of SMGs (e.g. Zavala et al. 2018; Ling & Yan 2022; Cheng et al. 2023), although we note the rest-frame wavelength probed by these studies various.

We further measure the concentration index (C) of the galaxies stellar light profile in each wavelength band using Eq. (4) from Conselice (2014)¹³ defined as,

$$C = 5 \times \log_{10} \left(\frac{r_{\text{outer}}}{r_{\text{inner}}} \right), \quad (1)$$

where C_{28} uses $r_{\text{inner}} = 20\%$ and $r_{\text{outer}} = 80\%$. A higher concentration indicates a larger fraction of the galaxies light is contained within the central regions. For the NIRCam-SMGs, we derive a median (and 16th–84th quartile) concentration index of $C_{28} = 3.77^{+1.12}_{-0.66}$ in the NIRCam $F444W$ band. The 16th–84th quartile is comparable to the range derived for local late-type discs and ellipticals from optical R -band imaging (0.65 μm ; Conselice et al. 2003). However, at the median redshift of our sample the $F444W$ band probes the rest-frame $\sim 1 \mu\text{m}$ emission, so this is not a direct comparison, as we analyse in Sect. 4.1.

To provide an alternative measure of half-light radius and concentration index, as well as more quantitative morphological indicators, we use the `statmorph`¹⁴ code (Rodriguez-Gomez et al. 2019). We run `statmorph` on the NIRCam and HST WFC3 wavelength bands above 1 μm for each of the 43 NIRCam-SMGs, using the same segmentation and PSFs as for the growth curve analysis described above. The `statmorph` code derives the concentration, asymmetry and clumpiness (CAS; Abraham et al. 2003; Lotz et al. 2008; Conselice 2014) parameters which quantify how concentrated, asymmetrical and clumpy the galaxies’ surface brightness profiles are, with higher values indicating more concentrated, asymmetric, clumpier light profiles.

¹³ See also Abraham et al. (2003).

¹⁴ <https://statmorph.readthedocs.io/en/latest/>

¹² <https://github.com/gbrammer/grizli-psf-library>

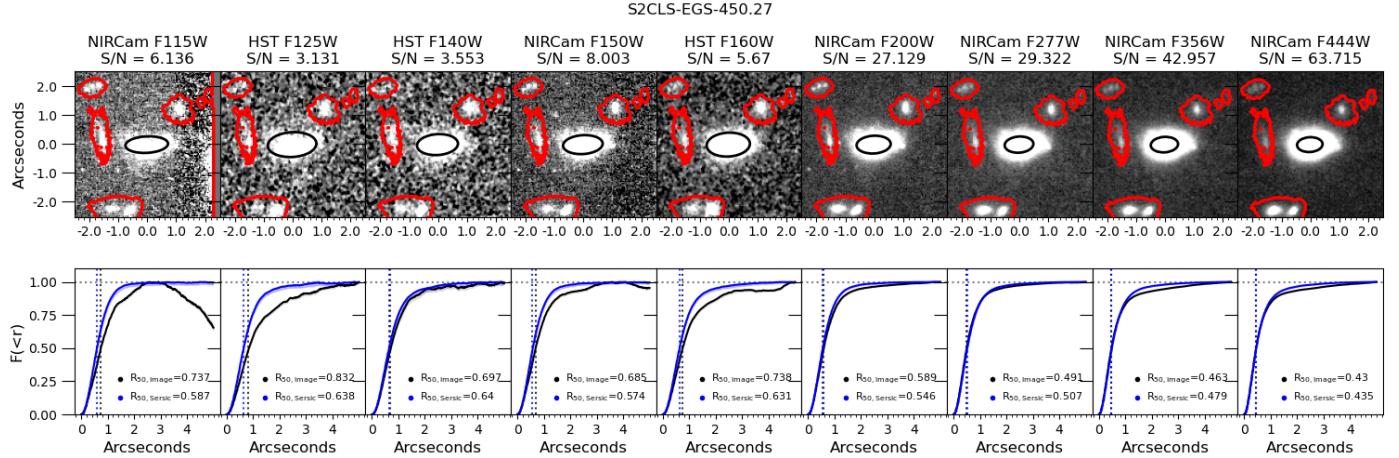


Fig. 4. Non-parametric growth-curve analysis for one of the NIRCam-SMGs. For each wavelength band, we show the 5 arcsec cutout, masked regions (red contours) and derived half-light radius (black ellipse) in the upper panel. In the lower panel we show the growth curve from the image (black curve) and growth from the parametric model (blue curve) derived in Sect. 4.2.

In addition, the Gini and M_{20} parameters are also derived (for full definitions see Lotz et al. 2004; Snyder et al. 2015). In short, the Gini parameter defines the pixel distribution of the galaxy’s light, where $G = 1$ means all of the light is concentrated in one pixel whilst $G = 0$ indicates each pixel contributes equally. The M_{20} parameter measures the moment of the galaxy’s brightest regions containing 20% of the total flux. This is then normalised by the total light moment for all pixels. Highly negative values indicates a high concentration of light, not necessarily at the centre of the galaxy.

To quantify the uncertainty on the non-parametric measures of morphology, we run `statmorph` 1000 times, each time adding random noise to the cutout of magnitude equal to the original rms noise measured in the cutout. The final properties (and uncertainty) for each galaxy are defined as the median (and standard deviation) of these 1000 iterations. To validate the robustness of the `statmorph` measurements we compare the half-light radius and concentration index to that derived from our growth curve analysis. We derive a median (and 16th–84th quartile) growth curve to `statmorph` concentration index ratio of $C_{28,\text{GC}} / C_{28,\text{statmorph}} = 1.08^{+0.34}_{-0.76}$ and median half-light radius ratio of $R_{\text{h,GC}} / R_{\text{h,statmorph}} = 1.0^{+0.18}_{-0.10}$, indicating good agreement between the independent morphological measurements¹⁵.

Having quantified the non-parametric morphology of the NIRCam-SMGs we now have multi-wavelength morphological measurements for the 43 NIRCam-SMGs from 1.15 to 4.44 μm . This sample however encapsulates a broad range of rest-frame wavelengths given the redshift range of the NIRCam-SMGs (see Fig. 3) and therefore can not be directly compared. To allow a more accurate comparison, for each NIRCam-SMG, we determine the rest-frame wavelength probed by the multi-wavelength imaging using the galaxy’s redshift, separating the wavelengths into rest-frame optical (0.25 to 0.75 μm) and near-infrared (1.0 to 1.5 μm) samples. In addition, `statmorph` returns the S/N per pixel for each cutout as well as a quality flag of the non-parametric measurements, with flag=0 indicating robust measurements were obtained.

If a galaxy has multiple optical or near-infrared measurements we select the measurement with the highest S/N per pixel

Table 2. Median (and $1 - \sigma$) non-parametric morphology of the NIRCam-SMGs at rest-frame optical and near-infrared wavelengths.

Property	Optical ($0.25 < \lambda < 0.75 \mu\text{m}$)	Near-IR ($1.0 < \lambda < 1.5 \mu\text{m}$)
C_{28}	2.90 ± 0.45	2.91 ± 0.39
A	0.19 ± 0.10	0.17 ± 0.13
S	0.03 ± 0.07	0.03 ± 0.06
Gini	0.51 ± 0.05	0.52 ± 0.05
M_{20}	-1.50 ± 0.33	-1.77 ± 0.23

and further remove any measurements with `statmorph` flag ≥ 1 and S/N per pixel < 2.5 to ensure the most robust morphological parameters are analysed for each galaxy. In total, this results in 32 and 26 robust non-parametric morphological measurements in the optical and near-infrared respectively for the NIRCam-SMGs for which we summarise the median properties in Table 2.

The concentration and asymmetry of the NIRCam-SMGs at both rest-frame optical and near-infrared wavelengths (as quantified in Table 2) are comparable to that identified in local irregular galaxies from R -band ($0.65 \mu\text{m}$) imaging of $C_{28} = 2.7 \pm 0.2$ and $A = 0.17 \pm 0.11$ (Conselice 2014). However, the local irregular galaxies studied in Conselice (2014) have much higher clumpiness ($S = 0.40 \pm 0.20$) than the NIRCam-SMGs. We suspect this is driven by resolution effects, smoothing out the morphology of the high-redshift galaxies.

In Fig. 5 we correlate the different non-parametric morphology measures, showing the Gini– M_{20} (left), clumpiness (S) – concentration (middle) and asymmetry – concentration (right) correlations for the optical and near-infrared samples. In each panel, we overlay the different morphological classifications which distinguish between mergers, disc-like galaxies and ellipticals in the local Universe (e.g. Lotz et al. 2008; Conselice et al. 2003; Bershady et al. 2000). On average both optical and near-infrared measurements lie in the disc-like region of the different planes, with some scatter towards intermediate, elliptical-like systems, reflecting the range of morphology seen in Fig. 2. This is typical of the rest-frame optical properties of galaxies identified in the CEERS survey from $z = 3$ to $z = 9$ (Kartaltepe et al. 2023) as well as other studies of SMGs non-parametric morphology (e.g. Ling & Yan 2022). The average optical

¹⁵ We note however we expect some variation between the two methods due to the definitions of centroid and total fluxes used in `statmorph` (see Lotz et al. 2004).

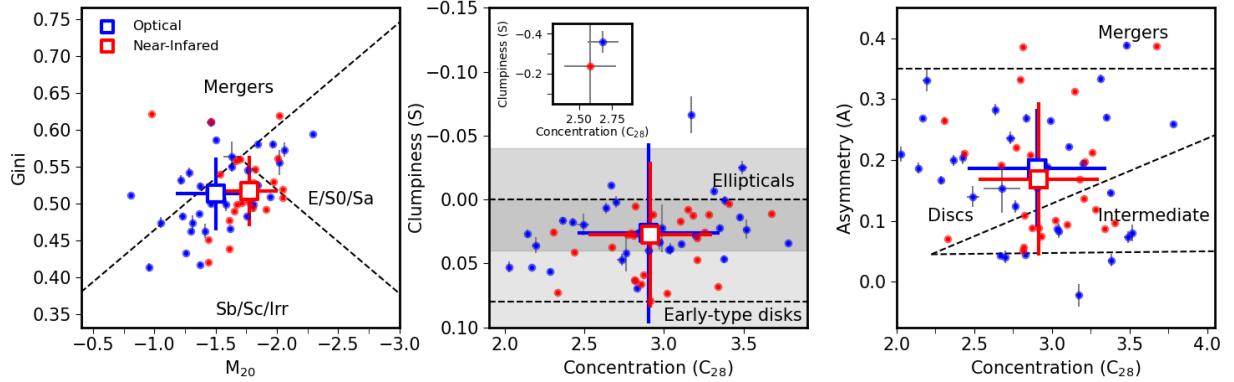


Fig. 5. Non-parametric rest-frame optical and near-infrared morphology of the NIRCam-SMGs. We show the Gini and M_{20} plane (left). The dashed black lines indicate the division between Mergers, Elliptical and Disc galaxies as defined by Lotz et al. (2008). In the middle panel we show the correlation between clumpiness (S) and concentration (C_{28}), with the parameter space of local galaxies from Conselice et al. (2003) overlaid. Finally, in the right panel we show the asymmetry (A) concentration (C_{28}) plane with dashed lines showing the distinction between Mergers, Disc, Intermediate and Ellipticals from Bershady et al. (2000) and Conselice et al. (2003). In all three panels, the optical and near-infrared samples occupy the region with disc-like morphology.

measurement has a larger M_{20} value at fixed Gini parameter, pushing it closer to the merger region of the Gini– M_{20} diagram. We infer this could be driven by the optical measurements probing more recent events of star formation (e.g. spiral arms, star-forming clumps) as opposed to the near-infrared measurements that probe the bulk of the older stellar population. We, therefore, expect the optical morphology to be more clumpy and asymmetrical compared to the near-infrared measurements.

To investigate this further, for galaxies which have both an optical and near-infrared measurement in Fig. 5 (26 galaxies), we quantify the ratio between the two measurements for each non-parametric parameter. On average we identify no significant offset between the optical and near-infrared concentration or Gini parameters with median ratios of $C_{28,\text{NIR}}/C_{28,\text{OPT}} = 1.02$ and $\text{Gini}_{\text{NIR}}/\text{Gini}_{\text{OPT}} = 1.00$ respectively but with an offset in the M_{20} parameter of $M_{20,\text{NIR}}/M_{20,\text{OPT}} = 1.1$. Whilst the optical measurements, for individual galaxies, have on average larger asymmetry and clumpiness with $A_{\text{NIR}}/A_{\text{OPT}} = 0.95$ and $S_{\text{NIR}}/S_{\text{OPT}} = 0.73$ respectively, reflecting the fact that the optical morphologies trace recent star-formation.

There are a number of NIRCam-SMGs (9) that fall into the non-parametric merger classification in Fig. 5. These objects are indicated by the ‘M’ in Fig. 2. Visual inspection of their morphology further indicates the presence of potential galaxy interactions and we also flag two other NIRCam-SMGs as mergers, S2CLS-EGS-450.18 and S2CLS-EGS-850.108, from visual inspection. Previous studies of SMGs have suggested they commonly reside in over-dense environments and are prone to galaxy interactions (e.g. Conselice et al. 2003; Oteo et al. 2018; An et al. 2019; Cardona-Torres et al. 2023; Calvi et al. 2023). However, this is in contrast to more recent JWST studies of ALMA-detected SMGs that on average identify isolated systems with stable disc-like morphologies (e.g. Chen et al. 2022, 2023). To investigate the NIRCam-SMGs morphology further, we undertake a parametric analysis of the non-merging systems, as discussed in the next section, using the non-parametric measures (e.g. x,y , PA, and b/a) as priors.

4.2. Parametric morphology

The parametric modelling of the SMG counterparts is carried out using GalfitM¹⁶ (Häußler et al. 2013). GalfitM is a multi-

wavelength wrapper for Galfit (Peng et al. 2010), in which a range of parametric models can be fitted to a galaxy’s morphology as a function of wavelength. The model parameters can be constrained to be independent at each wavelength provided or vary following some functional form (Häußler et al. 2022).

For our analysis, we fitted¹⁷ all wavelength bands above $1\mu\text{m}$ (with a S/N per pixel > 2.5) from HST WFC3 and JWST NIRCam with a single Sérsic model, convolved with the PSF, using GalfitM. We fix the centroid (x, y), axis ratio (b/a), and position angle (PA) to be constant across the wavelength bands, allowing the effective radius (R_e) and Sérsic index (n) to vary linearly whilst the magnitude is modelled quadratically as a function of wavelength (Häußler et al. 2022). We also provide GalfitM with a sigma image for each source generated from the ERR¹⁸ extension of the final mosaics.

We identify a median (and 16th–84th quartile) ratio of GalfitM to statmorph Sérsic index and half-light radius of $n_{\text{GalfitM}}/n_{\text{statmorph}} = 0.98^{+0.14}_{-0.22}$ and $R_{\text{h,GalfitM}}/R_{\text{h,statmorph}} = 0.97^{+0.08}_{-0.16}$ respectively, indicating strong agreement between the two independent measures of parametric morphology. Despite the agreement between the two parametric codes, it does not guarantee that the galaxies’ underlying morphology is well-modelled by a single Sérsic model. Given the complex morphologies seen in Fig. 2 and the broad scatter in non-parametric morphology in Fig. 5, as well as the diverse morphology identified in previous SMG studies (e.g. Conselice et al. 2003; Chang et al. 2018; Calvi et al. 2023), the NIRCam-SMGs morphology may in fact be more accurately described by a two-component model.

To investigate this further, we perform two-component multi-wavelength Sérsic fits using GalfitM. We constrain the first component to be an exponential disc ($n = 1$) and allow the second component’s Sérsic index to vary. This is similar to the approach taken in studies of intermediate redshift galaxies with HST, whose resolved morphologies often show complex multi-component features (e.g. dos Reis et al. 2020). We further constrain the centroid of the two components to be co-located within ± 3 pixels ($0''.12$), whilst the axis ratio, position angle and magnitude of the models are constrained in the same way as the single Sérsic fits described above.

¹⁷ We do not perform simultaneous modelling of galaxies since we have masked out all other objects (see Sect. 4).

¹⁸ The ERR extension is derived from the Grizli pipeline and contains the read and Poisson noise of the observations.

¹⁶ <https://www.nottingham.ac.uk/astronomy/megamorph/>

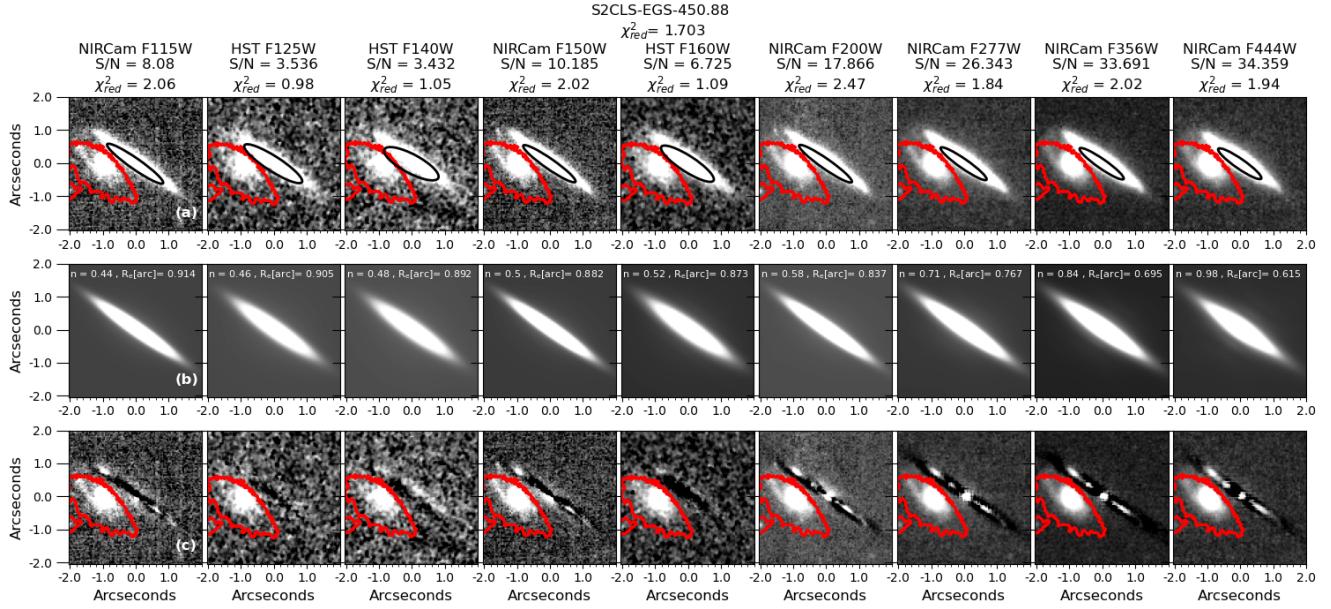


Fig. 6. Parametric morphological analysis for a NIRCam-SMG (S2CLS-EGS-450.88) for which we have coverage in every wavelength band. For each filter, we show a 5'' cutout (a) with masked regions (red contours) and a black ellipse which indicates the growth curve derived half-light radius (Sect. 4.1). The multi-wavelength `galfitM` model and residual are shown in panels (b) and (c) respectively with masked regions indicated as before. For each filter, we indicate the `statmorph` S/N per pixel as well as the reduced chi-squared per band, and overall, from `GalfitM`.

To assess whether the two-component fitting is more appropriate, we quantify the quality of `GalfitM` modelling using both the multi-wavelength reduced chi-squared ($\chi^2_{\text{red.}}$) and the Bayesian Information Criteria (BIC; Liddle 2007) which are defined as,

$$\chi^2 = \sum_i^{N_d} r_i^2 \quad (2)$$

$$\chi^2_{\text{red.}} = \chi^2 / (N_d - N_{\text{varys}}) \quad (3)$$

$$\text{BIC} = \chi^2 + N_{\text{varys}} \times \ln(N_d) \quad (4)$$

where $\sum_i^{N_d} r_i^2$ is the sum of the residual image, N_d is the number of data points, and N_{varys} is the number of variable parameters. Both of these statistical parameters reflect the goodness of fit whilst the BIC parameter penalises models with a large number of parameters and is commonly used to determine between two parametric models (e.g. Head et al. 2014; Lange et al. 2016). The model with the lower BIC values and $\chi^2_{\text{red.}}$ close to unity is preferred.

For the single Sérsic fitting we identify a median $\chi^2_{\text{red.}} = 1.76$ with a 16th–84th quartile range of $\chi^2_{\text{red.}} = 1.44$ –4.41. Whilst for two components we establish a median $\chi^2_{\text{red.}} = 1.69$ with a 16th–84th quartile range of $\chi^2_{\text{red.}} = 1.38$ –3.60. However, on average the two-component models have a higher BIC value than the single Sérsic models, with $\langle \Delta \text{BIC} \rangle = \langle \text{BIC}_{\text{single}} - \text{BIC}_{\text{two}} \rangle = -172$. This indicates that although the two-component model produces smaller residuals, it is over-fitting the morphology of the galaxies and not providing new physical information. We note also two-component models with two free Sérsic indexes produce similar results. We, therefore, proceed with the single Sérsic models to represent the NIRCam-SMGs parametric morphology, an example of which is shown in Fig. 6.

We again split the sample into rest-frame optical and near-infrared measurements as before, removing galaxies with a

Table 3. Median (and $1 - \sigma$) parametric morphology of the NIRCam-SMGs at rest-frame optical and near-infrared wavelengths.

Property	Optical ($0.25 < \lambda < 0.75 \mu\text{m}$)	Near-IR ($1.0 < \lambda < 1.5 \mu\text{m}$)
$R_e [\text{kpc}]$	3.10 ± 1.67	1.64 ± 0.97
n	0.96 ± 0.66	1.85 ± 0.63
b/a	0.64 ± 0.20	0.73 ± 0.18
$\chi^2_{\text{red.}}$	1.77 ± 0.54	1.72 ± 0.50

single-band reduced chi-squared of $\chi^2_{\text{red.}} > 3.0$ and those objects that occupy the non-parametric merger regions in Fig. 5. In total, this results in 25 and 14 robust parametric morphological measurements in the optical and near-infrared of the NIRCam-SMGs respectively. The image, model and residuals for each of the galaxies for this final sample are shown in Appendix C, for which we summarise the median parametric morphology properties in Table 3.

On average we identify almost twice as large optical effective radius (~ 3 kpc) than in the near-infrared (~ 1.64 kpc). Whilst for the Sérsic index, the optical measurements more resemble exponential discs ($n = 1$) with the rest-frame near-infrared observations having higher Sérsic index of order $n \sim 2$. This is similar to the trends seen in the non-parametric morphology and further reflects the near-infrared measurements probing older, less dust-attenuated, stellar populations in the process of bulge-formation (i.e. more compact), whilst rest-frame optical observations are more sensitive to recent star-formation, and consequently dust attenuation, thus seen in the extended regions of galaxies.

To analyse the cosmic evolution of the rest-frame optical and near-IR morphology of the NIRCam-SMGs, in Fig. 7 we correlate the effective radius of the `GALFITM` model and the Sérsic index with the galaxies' redshift. We also show a running median (of 5 points) and standard deviation in each panel (squares) along

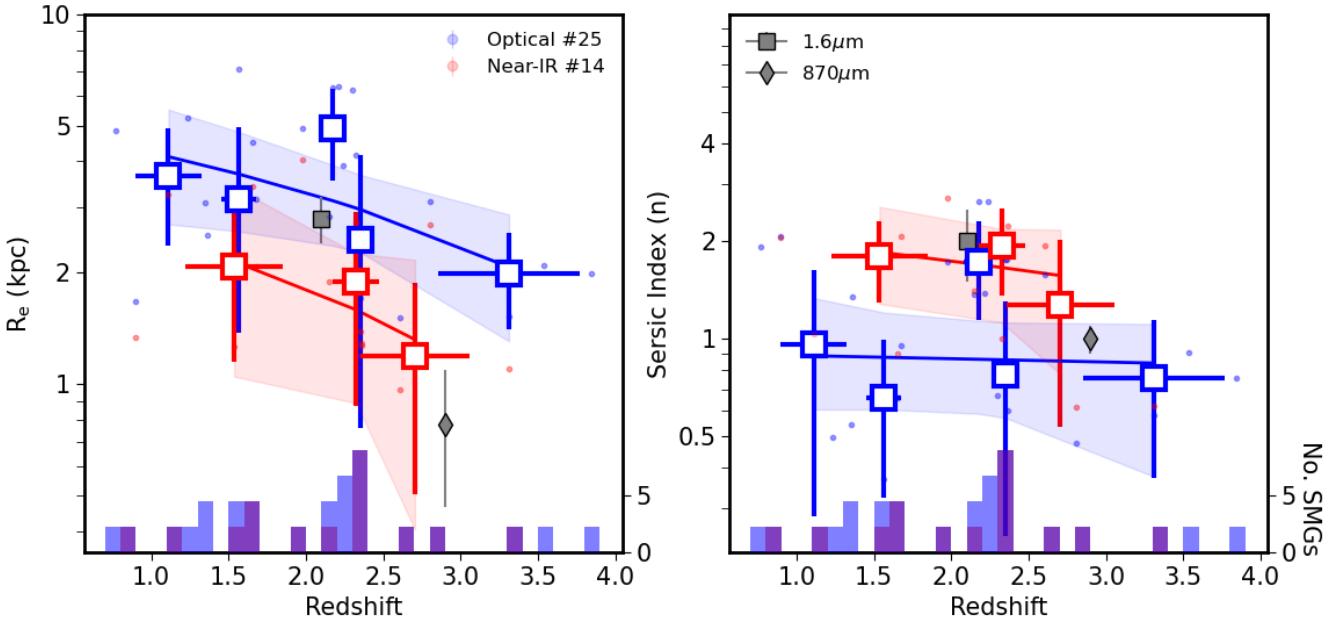


Fig. 7. Median rest-frame effective radius (R_e ; left) and Sérsic index (n ; right) for optical and near-infrared parametric sub-samples as a function of redshift. The squares represent the median (and $1 - \sigma$ scatter) for redshift and morphological property (R_e, n) in a running median. On the x -axis we show the distribution of redshift for each optical and near-infrared sub-samples. For each sub-sample, in each panel, we show a linear parametric fit and $1 - \sigma$ scatter. We also indicate the median size and Sérsic index from literature studies of SMGs (Swinbank et al. 2014; Gullberg et al. 2019). We identify larger (lower) effective radius (Sérsic index) at optical wavelengths compared to near-infrared measurements, with general a negative ($\sim 2\sigma$) correlation between the SMGs size and redshift whilst no strong correlation is identified with Sérsic index.

with a linear parametric fit and $1 - \sigma$ uncertainty (solid line and shaded region). On each axis, we indicate the redshift distribution of the optical and near-infrared samples.

For comparison we show the median $870\text{ }\mu\text{m}$ effective radius and Sérsic index ($R_e \sim 0.8\text{ kpc}$, $n \sim 1$) from Gullberg et al. (2019) for 153 bright SMGs with a median redshift of $z \approx 2.9$. We also indicate the median H -band ($1.6\text{ }\mu\text{m}$) effective radius and Sérsic index ($R_e \sim 2.8\text{ kpc}$ and $n \sim 2$) from Swinbank et al. (2014) for 25 SMGs with a median redshift of $z \approx 2.1$.

At both optical and near-infrared wavelengths, we identify a negative correlation of order $\sim 2\sigma$ between galaxy effective radius and redshift. In the optical we identify a linear slope of $\alpha = -0.92 \pm 0.45$ whilst for the near-infrared the linear slope is $\alpha = -0.74 \pm 0.34$. This evolution can be attributed to the expected growth of galaxies with cosmic time (e.g. van der Wel et al. 2014b), with the larger optical size reflecting the extended star-forming regions probed in the rest-frame optical.

Our optical and near-infrared continuum sizes are larger than the far-infrared (dust) sizes derived in Gullberg et al. (2019), as expected from the compact nature of the far-IR dust component of these galaxies (e.g. Gullberg et al. 2018; Hodge et al. 2019; Lang et al. 2019). The rest-frame optical sizes are consistent with those reported by Zavala et al. (2018) with $R_h = 4.8 \pm 0.4\text{ kpc}$ as well as other studies of the SMG morphology with HST (Swinbank et al. 2014; Chen et al. 2015; Ling & Yan 2022). These studies, on average, are probing the rest-frame UV, and thus should only be compared to our optical sample. A similar offset between optical and far-IR, is found in resolved $\text{H}\alpha$ studies of $z \sim 2$ SMGs (e.g. Chen et al. 2020) as well as in hydrodynamical simulations of typical high-redshift ($1 < z < 5$) main-sequence galaxies (e.g. Popping et al. 2022) which both identify significantly larger optical sizes compared to the galaxies' far-infrared dust sizes.

Using Illustris The Next Generation 50 (TNG50; Nelson et al. 2019) simulation, Popping et al. (2022) identify for high-redshift ($2 < z < 3$) massive ($\log_{10}(M_*) > 10$) galaxies, a FIR ($850\text{ }\mu\text{m}$) to $1.6\text{ }\mu\text{m}$ size ratio of ≤ 0.5 . At $z = 2.9$, our linear parametric fit predicts a rest-frame optical size of 2.45 kpc . Taking the $850\text{ }\mu\text{m}$ size as that derived in Gullberg et al. (2019) (0.8 kpc), we identify an $850\text{ }\mu\text{m}$ to optical size ratio of 0.33 . This is comparable to that predicted by Popping et al. (2022), although we note this is not a direct comparison due to the different selection functions across the samples.

In the right-hand panel of Fig. 7 we show the Sérsic index correlation with redshift, identifying a linear slope of $\alpha = -0.05 \pm 0.23$ in the rest-frame optical whilst for the near-infrared the linear slope is $\alpha = -0.25 \pm 0.44$. This indicates there is no correlation with redshift for Sérsic index both at optical and near-infrared wavelengths for the non-merging SMGs in our sample. The elevated near-infrared Sérsic index, in comparison to the optical Sérsic index, might be driven by the compact dusty nature of the SMGs. The dust will heavily attenuate the short-wavelength (optical) emission in the central regions of the galaxies, resulting in centrally less-peaked surface brightness profiles. As a consequence the observed size and Sérsic index would be higher (lower) in the rest-frame optical, than for the intrinsic (unattenuated) light profiles.

A similar trend has been identified in recent cosmological simulations which identify that massive, intrinsically compact galaxies appear significantly more extended when the effects of dust are taken into account (e.g. Roper et al. 2022, 2023). Our sample is intrinsically selected to have high levels of dust attenuation (Sect. 3), and therefore dust is potentially driving the trends seen in Fig. 7. Recent studies of high-redshift ($2 < z < 6$) optically faint galaxies have identified high-dust content and compact ($R_e \sim 1\text{--}2\text{ kpc}$) near-infrared morphologies (e.g. Nelson et al. 2023). Whilst the dominance of the disc-like morphologies in

the rest-frame optical has been shown to continue out to $z=8$ (e.g. Ferreira et al. 2022). These studies are in agreement with our morphological analysis of non-merging NIRCam-SMGs, with compact (extended) near-infrared (optical) sizes and disc-dominated optical morphologies out to $z=5.4$.

5. Conclusions

In this paper we have presented an analysis of the near-infrared JWST/NIRCam counterparts to 45 previously observed SCUBA-2 sub-mm selected galaxies (Table 1). Through a combination of the multi-wavelength p -values, NIRCam colours and Eazy-py predicted fluxes we construct a decision tree to identify the NIRCam counterpart to each sub-mm selected SCUBA-2 source (Fig. 1). We establish 43 NIRCam counterparts, from 37 SCUBA-2 sources, which cover a broad range of redshift from $z \approx 0.21$ – 5.4 with a median of $\tilde{z} = 2.29$.

We then proceed to quantify the rest-frame optical and near-infrared morphology of these 43 NIRCam-SMGs, utilising a variety of parametric and non-parametric morphological measures and taking advantage of the unprecedented depth and resolution of the CEERS JWST NIRCam observations (Fig. 2). We establish, on average, the non-parametric measures of the NIRCam-SMGs morphology indicate late-type discs but with a broad scatter with some galaxies exhibiting galaxy interactions and mergers, whilst others resemble elliptical spheroidal systems (Fig. 5). For individual NIRCam-SMGs, we establish the rest-frame optical morphology is on average more asymmetrical and clumpy in comparison to near-infrared morphology, with the largest offset towards higher M_{20} parameter (Table 2). We infer this is likely driven by the rest-frame optical morphology reflecting recent and on-going star-formation whilst the near-infrared probes the older stellar populations.

The parametric morphology of the galaxies is quantified using GALFITM and is well-modelled by a single component Sérsic model, as we identify two-component models result in over fitting of the data. On average the optical stellar continuum half-light radii exhibit larger sizes (Table 3) than the near-infrared measurements, likely driven by dust attenuation at UV and optical wavelengths. We establish a negative correlation with effective radius and redshift (i.e. larger sizes at later cosmic times; Fig. 7) reflecting the general growth of galaxies and build-up of stellar mass with cosmic time. The median Sérsic index at optical wavelengths is identified to be a factor of 2 smaller than at near-infrared wavelengths, with both exhibiting minimal evolution with cosmic time (Fig. 7). The implications of dust on the observed properties of the NIRCam-SMGs are likely driving the offsets between near-infrared and optical morphologies in addition to the age of the stellar populations probed. Our results are consistent with the picture of inside-out galaxy evolution, with more centrally concentrated older stellar populations, and more extended, younger star-forming regions whose stellar emission is heavily attenuated in the central regions.

Overall we have shown the morphology of single-dished detected SMGs at optical and near-infrared wavelengths is diverse, ranging from compact spheroidal galaxies, grand design spirals and interacting merging systems. We have identified an evolution in morphology from small, disc-dominated structures to more extended galaxies with disc-like optical morphologies and more centrally concentrated (late-type disc) near-infrared morphologies. Future quantification of the morphology and underlying stellar population of SMGs with JWST/MIRI imaging and spectroscopy will further build upon this picture and the wider cosmic evolution of SMGs.

Acknowledgements. We thank the referee for a constructive review and also Ian Smail, Stephen Wilkins, Julie Wardlow, and Oliver Newton for helpful discussions on this project. We would also like to thank the CEERS team for designing and executing the Early Release Science observations upon which this work is based. The observations analysed in this work are made with the NASA/ESA/CSA *James Webb* Space Telescope (DOI: 10.17909/z7p0-8481). This work made use of the following open-source software; Astropy (Astropy Collaboration 2013, 2018), Photutils (Bradley et al. 2022b), Source Extractor (Bertin & Arnouts 1996) SEP (Barbary et al. 2016) Eazy-py (Brammer & Matharu 2021) GriZli (Brammer et al. 2022) GalFitM (Häubler et al. 2013) The Cosmic Dawn Center (DAWN) is funded by the Danish National Research Foundation (DNRF) under grant No. 140. B.G. acknowledges support from the Carlsberg Foundation Research Grant CF20-0644 ‘Physical pRoperties of the InterStellar Medium in Luminous Infrared Galaxies at High redshiftT: PRISM- LIGHT’. A.P.V., I.J. and T.R.G. are grateful for support from the Carlsberg Foundation via grant No. CF20-0534. G.M. acknowledges the Villum Fonden research grants 13160 and 37440. S.J. is supported by the European Union’s Horizon Europe research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 101060888. V.K. acknowledges funding from the Dutch Research Council (NWO) through the award of the Vici Grant VI.C.212.036. Cloud-based data processing and file storage for this work is provided by the AWS Cloud Credits for Research programme. This work was partially supported by DeiC National HPC (g.a. DeiC-DTU-L-20210103) and by the dtu_00026 project.

References

- Abraham, R. G., van den Bergh, S., & Nair, P. 2003, *ApJ*, 588, 218
- An, F. X., Stach, S. M., Smail, I., et al. 2018, *ApJ*, 862, 101
- An, F. X., Simpson, J. M., Smail, I., et al. 2019, *ApJ*, 886, 48
- Astropy Collaboration (Robitaille, T. P., et al.) 2013, *A&A*, 558, A33
- Astropy Collaboration (Price-Whelan, A. M., et al.) 2018, *AJ*, 156, 123
- Bagley, M. B., Finkelstein, S. L., Koekemoer, A. M., et al. 2023, *ApJ*, 946, L12
- Barbary, K., Boone, K., McCully, C., et al. 2016, <https://doi.org/10.5281/zenodo.159035>
- Barger, A. J., Cowie, L. L., Chen, C. C., et al. 2014, *ApJ*, 784, 9
- Baugh, C. M., Lacey, C. G., Frenk, C. S., et al. 2005, *MNRAS*, 356, 1191
- Bershady, M. A., Jangren, A., & Conselice, C. J. 2000, *AJ*, 119, 2645
- Bertin, E., & Arnouts, S. 1996, *A&AS*, 117, 393
- Biggs, A. D., Ivison, R. J., Ibar, E., et al. 2011, *MNRAS*, 413, 2314
- Blain, A. W., Smail, I., Ivison, R. J., Kneib, J. P., & Frayer, D. T. 2002, *Phys. Rep.*, 369, 111
- Boyer, M. L., Anderson, J., Gennaro, M., et al. 2022, *Res. Notes Am. Astron. Soc.*, 6, 191
- Bradley, L. D., Coe, D., Brammer, G., et al. 2022a, *ApJ*, submitted [arXiv:2210.01777]
- Bradley, L., Sipőcz, B., Robitaille, T., et al. 2022b, <https://doi.org/10.5281/zenodo.6825092>
- Brammer, G., & Matharu, J. 2021, <https://doi.org/10.5281/zenodo.5012699>
- Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, *ApJ*, 686, 1503
- Brammer, G., Strait, V., Matharu, J., & Momcheva, I. 2022, <https://doi.org/10.5281/zenodo.6672538>
- Calvi, R., Castignani, G., & Dannerbauer, H. 2023, *A&A*, accepted [arXiv:2302.10323]
- Cardona-Torres, L., Artxaga, I., Montaña, A., Zavala, J. A., & Faber, S. M. 2023, *MNRAS*, 520, 5446
- Casey, C. M., Narayanan, D., & Cooray, A. 2014, *Phys. Rep.*, 541, 45
- Chabrier, G. 2003, *PASP*, 115, 763
- Chang, Y.-Y., Ferraro, N., Wang, W.-H., et al. 2018, *ApJ*, 865, 103
- Chapman, S. C., Blain, A. W., Smail, I., & Ivison, R. J. 2005, *ApJ*, 622, 772
- Chen, C.-C., Smail, I., Swinbank, A. M., et al. 2015, *ApJ*, 799, 194
- Chen, C.-C., Smail, I., Swinbank, A. M., et al. 2016, *ApJ*, 831, 91
- Chen, C.-C., Harrison, C. M., Smail, I., et al. 2020, *A&A*, 635, A119
- Chen, C.-C., Gao, Z.-K., Hsu, Q.-N., et al. 2022, *ApJ*, 939, L7
- Cheng, C., Huang, J.-S., Smail, I., et al. 2023, *ApJ*, 942, L19
- Cochrane, R. K., Best, P. N., Smail, I., et al. 2021, *MNRAS*, 503, 2622
- Cochrane, R. K., Hayward, C. C., Anglés-Alcázar, D., & Somerville, R. S. 2023, *MNRAS*, 518, 5522
- Conroy, C., & Gunn, J. E. 2010, *ApJ*, 712, 833
- Conselice, C. J. 2014, *ARA&A*, 52, 291
- Conselice, C. J., Chapman, S. C., & Windhorst, R. A. 2003, *ApJ*, 596, L5
- Cowley, W. I., Lacey, C. G., Baugh, C. M., et al. 2019, *MNRAS*, 487, 3082
- Davé, R., Finlator, K., Oppenheimer, B. D., et al. 2010, *MNRAS*, 404, 1355
- Davis, M., Guhathakurta, P., Konidaris, N. P., et al. 2007, *ApJ*, 660, L1
- Dempsey, J. T., Friberg, P., Jenness, T., et al. 2013, *MNRAS*, 430, 2534

- Dey, A., Graham, J. R., Ivison, R. J., et al. 1999, [ApJ](#), **519**, 610
 Díaz-Sánchez, A., Iglesias-Groth, S., Rebolo, R., & Dannerbauer, H. 2017, [ApJ](#), **843**, L22
 Dickinson, M., & FIDEL Team 2007, [Am. Astron. Soc. Meet. Abstr.](#), **211**, 52.16
 dos Reis, S. N., Buitrago, F., Papaderos, P., et al. 2020, [A&A](#), **634**, A11
 Downes, A. J. B., Peacock, J. A., Savage, A., & Carrie, D. R. 1986, [MNRAS](#), **218**, 31
 Dudzevičiūtė, U., Smail, I., Swinbank, A. M., et al. 2020, [MNRAS](#), **494**, 3828
 Eales, S., Lilly, S., Gear, W., et al. 1999, [ApJ](#), **515**, 518
 Fazio, G. G., Hora, J. L., Allen, L. E., et al. 2004, [ApJS](#), **154**, 10
 Ferreira, L., Conselice, C. J., Sazonova, E., et al. 2022, [ApJ](#), submitted [arXiv:[2210.01110](#)]
 Franco, M., Elbaz, D., Béthermin, M., et al. 2018, [A&A](#), **620**, A152
 Fruchter, A. S., & Hook, R. N. 2002, [PASP](#), **114**, 144
 Gaia Collaboration (Brown, A. G. A., et al.) 2021, [A&A](#), **649**, A1
 Geach, J. E., Dunlop, J. S., Halpern, M., et al. 2017, [MNRAS](#), **465**, 1789
 Gómez-Guijarro, C., Toft, S., Karim, A., et al. 2018, [ApJ](#), **856**, 121
 Gullberg, B., Swinbank, A. M., Smail, I., et al. 2018, [ApJ](#), **859**, 12
 Gullberg, B., Smail, I., Swinbank, A. M., et al. 2019, [MNRAS](#), **490**, 4956
 Harrington, K. C., Weiss, A., Yun, M. S., et al. 2021, [ApJ](#), **908**, 95
 Häufler, B., Bamford, S. P., Vika, M., et al. 2013, [MNRAS](#), **430**, 330
 Häufler, B., Vika, M., Bamford, S. P., et al. 2022, [A&A](#), **664**, A92
 Hayward, C. C., Sparre, M., Chapman, S. C., et al. 2021, [MNRAS](#), **502**, 2922
 Head, J. T. C. G., Lucey, J. R., Hudson, M. J., & Smith, R. J. 2014, [MNRAS](#), **440**, 1690
 Hill, R., Chapman, S. C., Scott, D., et al. 2018, [MNRAS](#), **477**, 2042
 Hodge, J. A., & da Cunha, E. 2020, [R. Soc. Open Sci.](#), **7**
 Hodge, J. A., Carilli, C. L., Walter, F., et al. 2012, [ApJ](#), **760**, 11
 Hodge, J. A., Swinbank, A. M., Simpson, J. M., et al. 2016, [ApJ](#), **833**, 103
 Hodge, J. A., Smail, I., Walter, F., et al. 2019, [ApJ](#), **876**, 130
 Holland, W. S., Bintley, D., Chapin, E. L., et al. 2013, [MNRAS](#), **430**, 2513
 Ivison, R. J., Greve, T. R., Dunlop, J. S., et al. 2007, [MNRAS](#), **380**, 199
 Karim, A., Swinbank, A. M., Hodge, J. A., et al. 2013, [MNRAS](#), **432**, 2
 Kartaltepe, J. S., Rose, C., Vanderhoof, B. N., et al. 2023, [ApJ](#), **946**, L15
 Kokorev, V., Brammer, G., Fujimoto, S., et al. 2022, [ApJS](#), **263**, 38
 Krick, J. E., Lowrance, P., Carey, S., et al. 2021, [AJ](#), **161**, 177
 Kron, R. G. 1980, [ApJS](#), **43**, 305
 Lang, P., Schinnerer, E., Smail, I., et al. 2019, [ApJ](#), **879**, 54
 Lange, R., Moffett, A. J., Driver, S. P., et al. 2016, [MNRAS](#), **462**, 1470
 Law, D. R., Steidel, C. C., Shapley, A. E., et al. 2012, [ApJ](#), **745**, 85
 Lelli, F., Di Teodoro, E. M., Fraternali, F., et al. 2021, [Science](#), **371**, 713
 Liddle, A. R. 2007, [MNRAS](#), **377**, L74
 Ling, C., & Yan, H. 2022, [ApJ](#), **929**, 40
 Long, A. S., Casey, C. M., Lagos, C. d. P., et al. 2022, [ApJ](#), submitted [arXiv:[2211.02072](#)]
 Lotz, J. M., Primack, J., & Madau, P. 2004, [AJ](#), **128**, 163
 Lotz, J. M., Davis, M., Faber, S. M., et al. 2008, [ApJ](#), **672**, 177
 Lovell, C. C., Geach, J. E., Davé, R., Narayanan, D., & Li, Q. 2021, [MNRAS](#), **502**, 772
 Lower, S., Narayanan, D., Li, Q., & Davé, R. 2023, [ApJ](#), **950**, 94
 Manning, S. M., Casey, C. M., Zavala, J. A., et al. 2022, [ApJ](#), **925**, 23
 Martis, N. S., Marchesini, D. M., Muzzin, A., Willott, C. J., & Sawicki, M. 2022, [MNRAS](#), **518**, 4961
 McAlpine, S., Smail, I., Bower, R. G., et al. 2019, [MNRAS](#), **488**, 2440
 Miller, T. B., Chapman, S. C., Aravena, M., et al. 2018, [Nature](#), **556**, 469
 Nardiello, D., Bedin, L. R., Burgasser, A., et al. 2022, [MNRAS](#), **517**, 484
 Nelson, D., Pilipich, A., Springel, V., et al. 2019, [MNRAS](#), **490**, 3234
 Nelson, E. J., Suess, K. A., Bezanson, R., et al. 2023, [ApJ](#), **948**, L18
 Oteo, I., Ivison, R. J., Dunne, L., et al. 2018, [ApJ](#), **856**, 72
 Pacifici, C., Iyer, K. G., Mobasher, B., et al. 2023, [ApJ](#), **944**, 141
 Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2010, [AJ](#), **139**, 2097
 Perrin, M. D., Sivaramakrishnan, A., Lajoie, C. P., et al. 2014, [SPIE Conf. Ser.](#), **9143**, 91433X
 Popping, G., Pilipich, A., Calistro Rivera, G., et al. 2022, [MNRAS](#), **510**, 3321
 Rizzo, F., Vegetti, S., Fraternali, F., Stacey, H. R., & Powell, D. 2021, [MNRAS](#), **507**, 3952
 Robertson, B. E., Tacchella, S., Johnson, B. D., et al. 2023, [ApJ](#), **942**, L42
 Rodriguez-Gomez, V., Snyder, G. F., Lotz, J. M., et al. 2019, [MNRAS](#), **483**, 4140
 Roper, W. J., Lovell, C. C., Vijayan, A. P., et al. 2022, [MNRAS](#), **514**, 1921
 Roper, W. J., Lovell, C. C., Vijayan, A. P., et al. 2023, ArXiv e-prints [arXiv:[2301.05228](#)]
 Shim, H., Lee, D., Kim, Y., et al. 2022, [MNRAS](#), **514**, 2915
 Simpson, J. M., Swinbank, A. M., Smail, I., et al. 2014, [ApJ](#), **788**, 125
 Smail, I., Ivison, R. J., & Blain, A. W. 1997, [ApJ](#), **490**, L5
 Smail, I., Dudzevičiūtė, U., Stach, S. M., et al. 2021, [MNRAS](#), **502**, 3426
 Snyder, G. F., Torrey, P., Lotz, J. M., et al. 2015, [MNRAS](#), **454**, 1886
 Stach, S. M., Smail, I., Swinbank, A. M., et al. 2018, [ApJ](#), **860**, 161
 Stach, S. M., Dudzevičiūtė, U., Smail, I., et al. 2019, [MNRAS](#), **487**, 4648
 Swinbank, A. M., Smail, I., Chapman, S. C., et al. 2010, [MNRAS](#), **405**, 234
 Swinbank, A. M., Simpson, J. M., Smail, I., et al. 2014, [MNRAS](#), **438**, 1267
 van der Wel, A., Chang, Y.-Y., Bell, E. F., et al. 2014a, [ApJ](#), **792**, L6
 van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2014b, [ApJ](#), **788**, 28
 Walter, F., Decarli, R., Carilli, C., et al. 2012, [Nature](#), **486**, 233
 Wardlow, J. L., Smail, I., Coppin, K. E. K., et al. 2011, [MNRAS](#), **415**, 1479
 Weiβ, A., Ivison, R. J., Downes, D., et al. 2009, [ApJ](#), **705**, L45
 Weiβ, A., De Breuck, C., Marrone, D. P., et al. 2013, [ApJ](#), **767**, 88
 Zavala, J. A., Artxaga, I., Geach, J. E., et al. 2017, [MNRAS](#), **464**, 3369
 Zavala, J. A., Artxaga, I., Dunlop, J. S., et al. 2018, [MNRAS](#), **475**, 5585
 Zhang, H., Primack, J. R., Faber, S. M., et al. 2019, [MNRAS](#), **484**, 5170

Appendix A: Decision tree

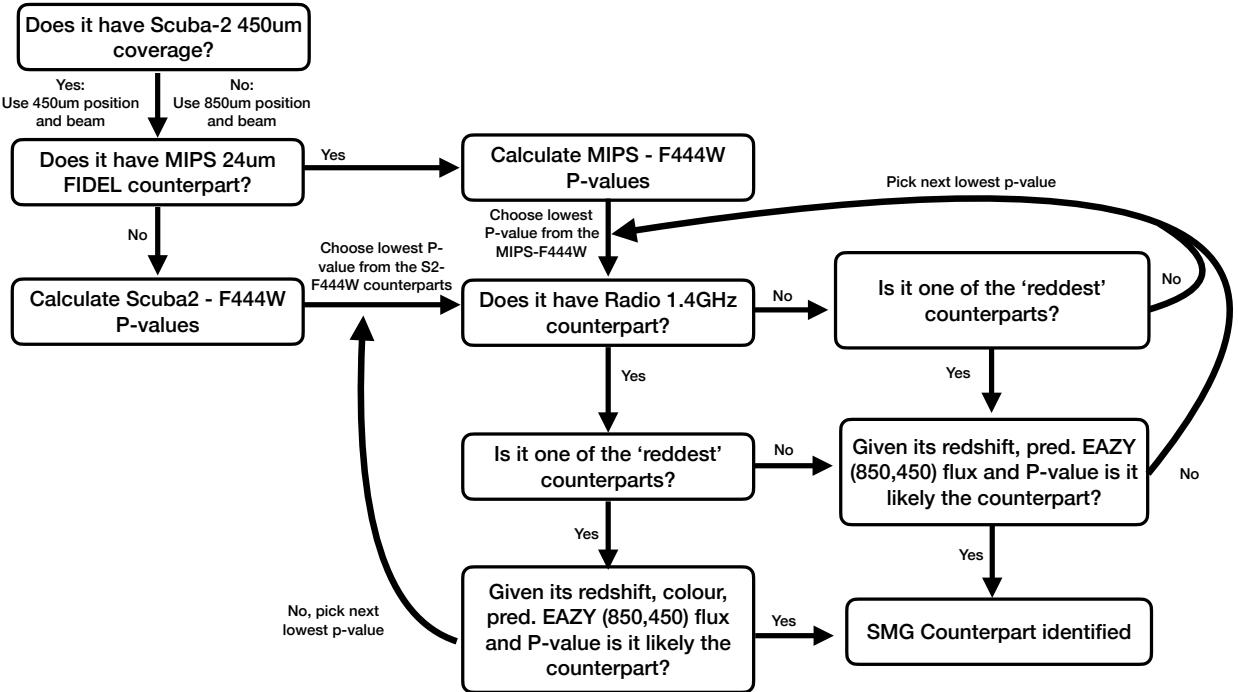


Fig. A.1. Schematic of the decision tree used to determine the NIRCam counterpart(s) to the SCUBA-2 selected SMGs. If a MIPS counterpart has been identified in the SCUBA-2 beam then p -values from the MIPS 24 μ m image to NIRCam F444W are calculated, else directly from SCUBA-2 to NIRCam F444W. Then considering the NIRCam colours and EaZy-py outputs as shown in Figure 1, a decision is made.

Appendix B: SED fitting

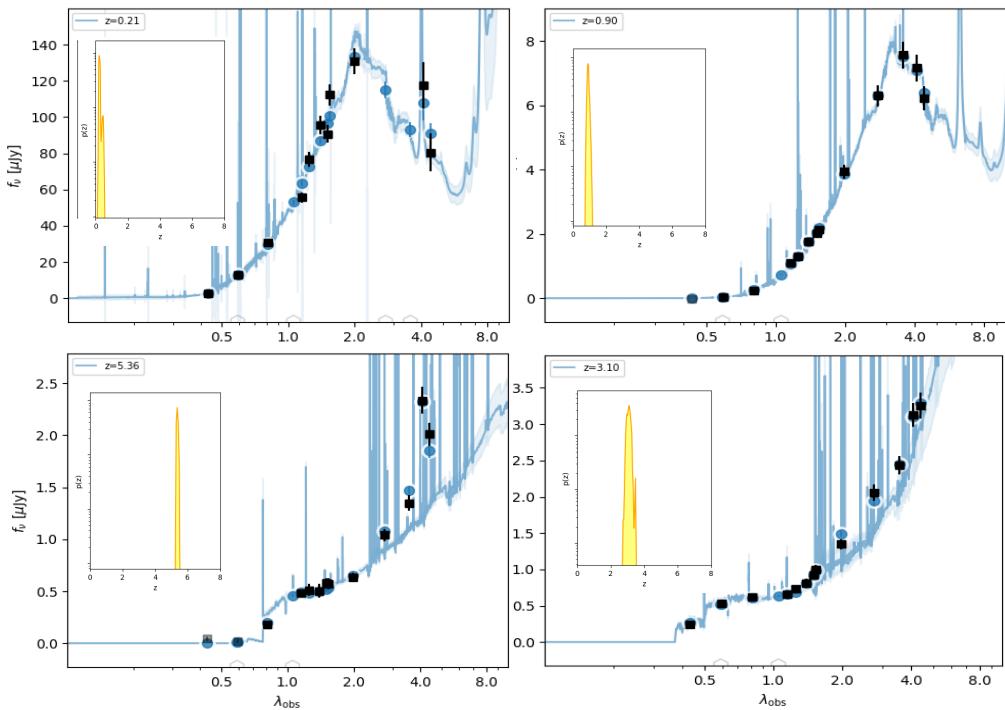


Fig. B.1. Eazy-py fitting results for four of the NIRCam counterparts across a broad range of redshift. The inset panel indicates the $p(z)$ distribution.

Appendix C: GALFITM modelling

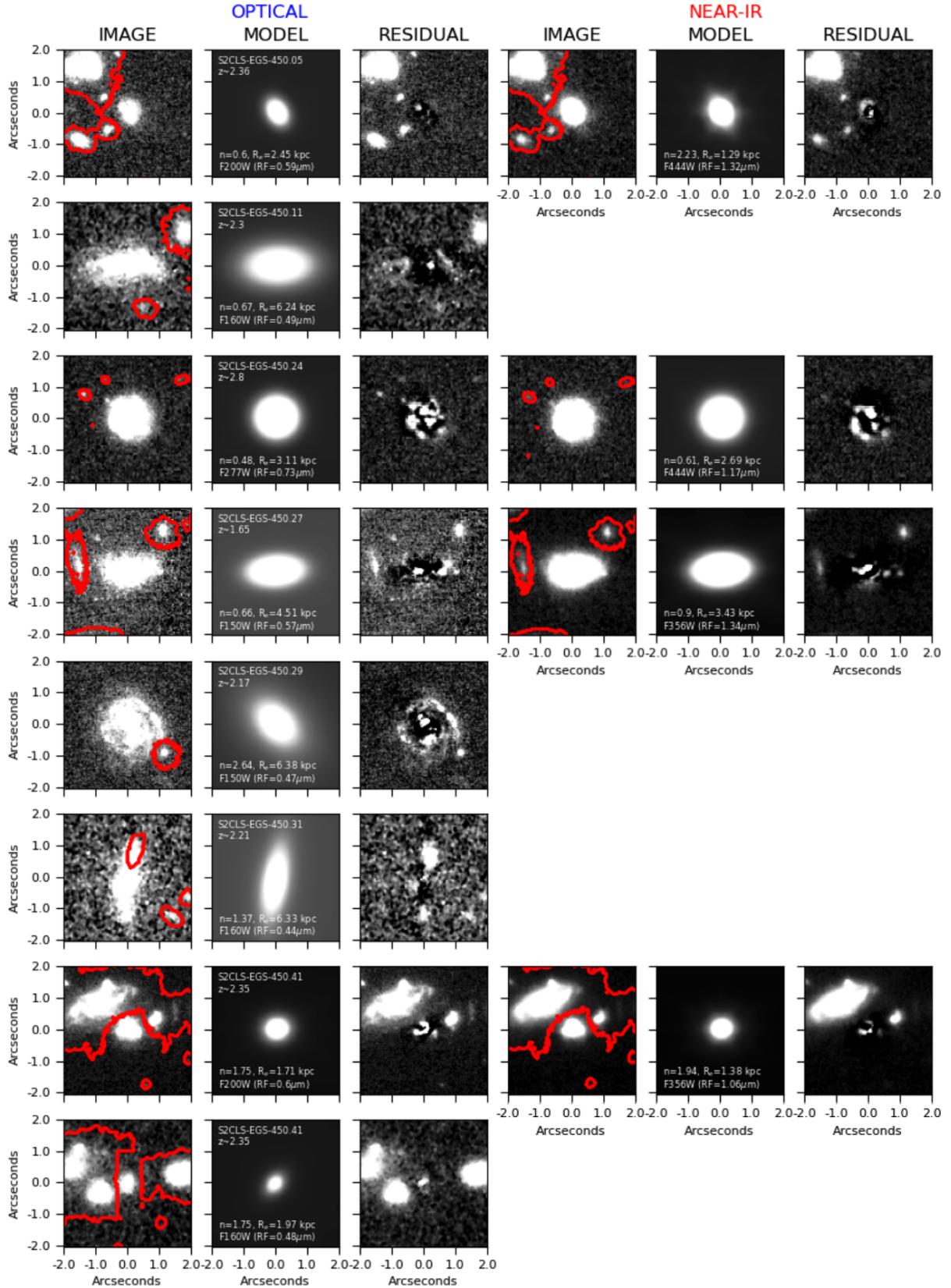
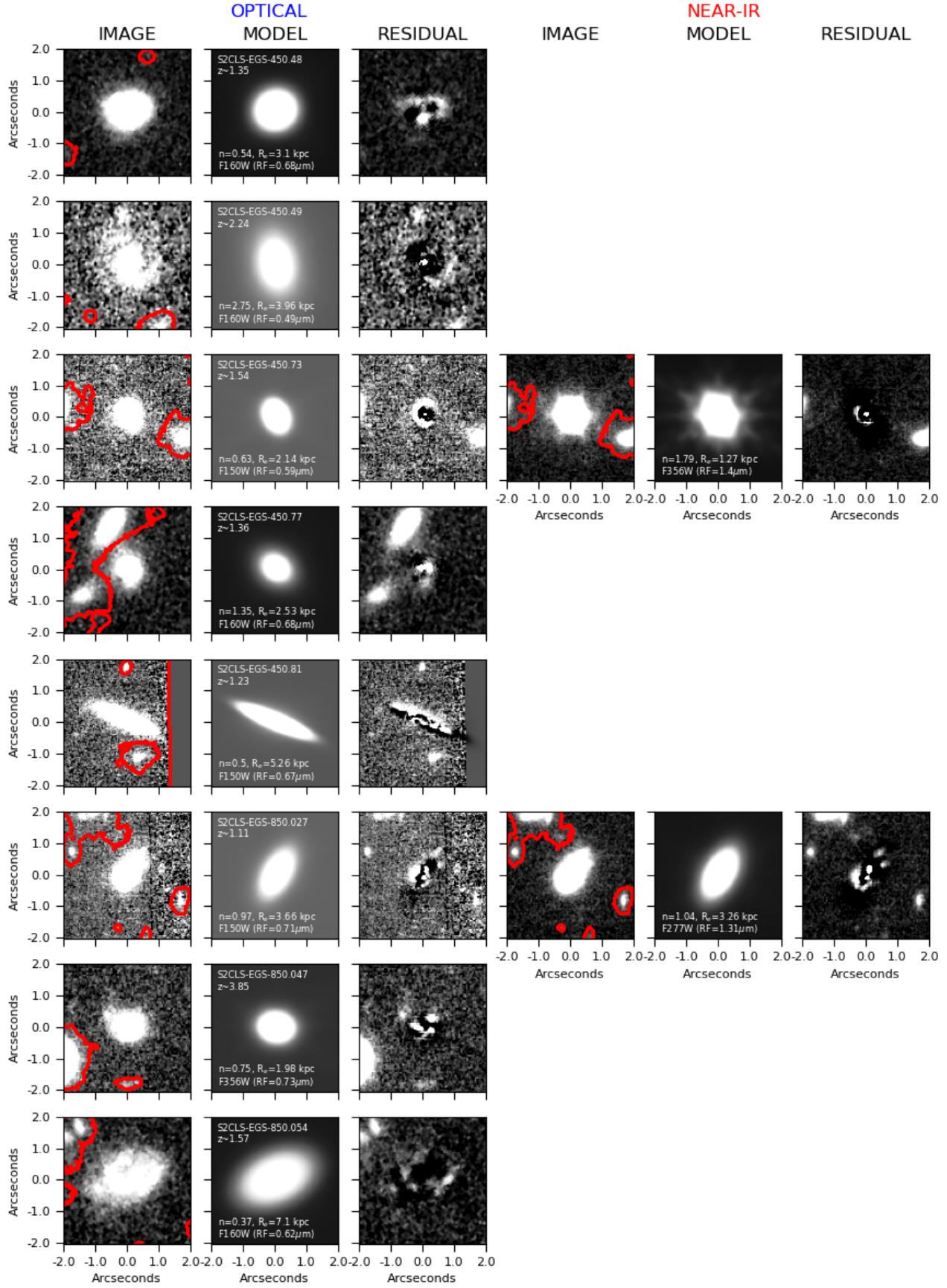
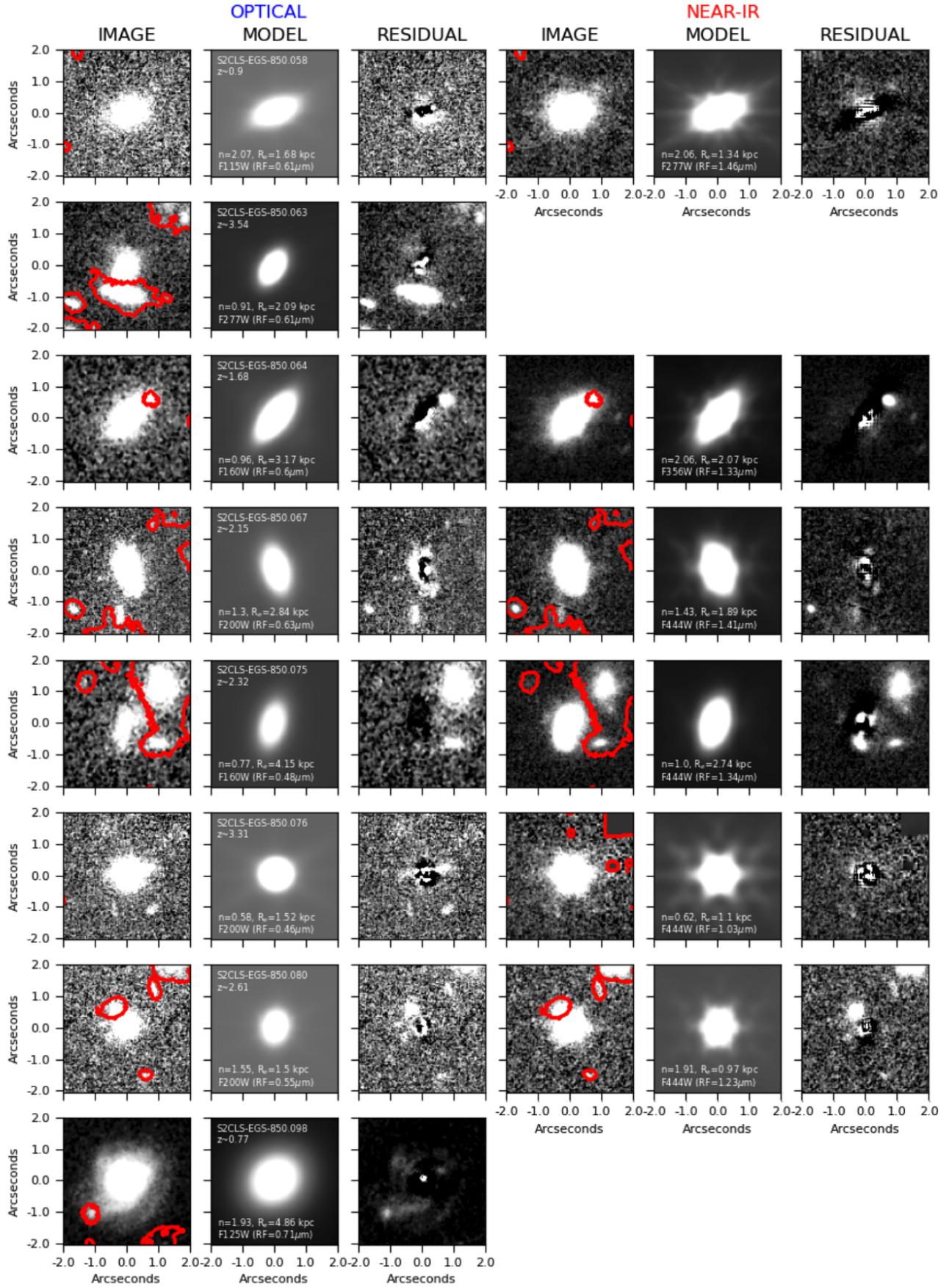


Fig. C.1. Rest-frame optical and near-infrared GALFITM modelling of the final rest-frame sample presented in Figure 7. For each NIRCam-SMG we present the 4-arcsecond optical image, GALFITM model and residual. We further indicate the mask regions of nearby objects (red contours) and indicate for each model the redshift, derived Sérsic index, effective radius, instrument band and rest-frame wavelength probed. Where available (14/25) we display the near-infrared image, model and residual.

**Fig. C.1.** continued.

**Fig. C.1.** continued.

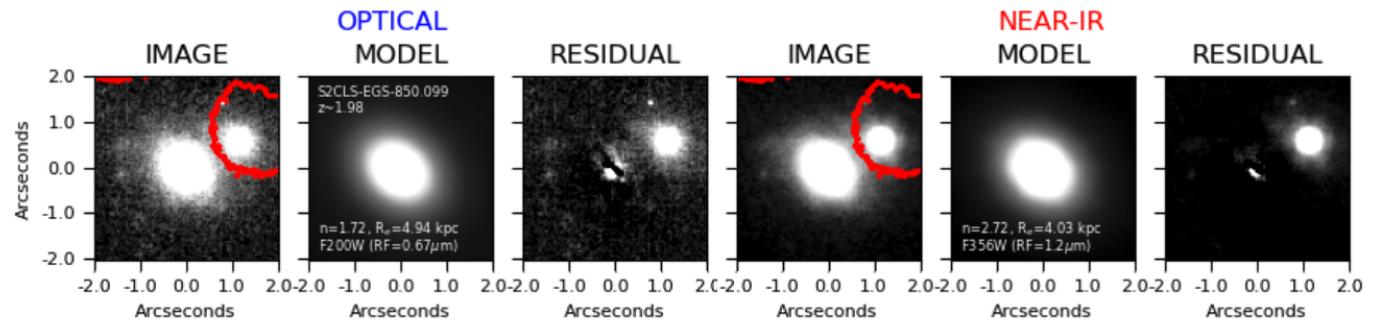


Fig. C.1. continued.