Towards studying hierarchical assembly in real time: a Milky Way progenitor galaxy at $z = 2.36$ under the microscope

Tom O. Zick,[1]★ Daniel R. Weisz [1] Bruno Ribeiro,[2] Mariska T. Kriek,[1] Benjamin D. Johnson,[3] Xiangcheng Ma [1] and Rychard Bouwens[2]

[1]*Astronomy Department, University of California, Berkeley, CA 94720, USA*
[2]*Leiden Observatory, Leiden University, NL-2300 RA Leiden, the Netherlands*
[3]*Center for Astrophysics, Harvard & Smithsonian, Cambridge, MA 02138, USA*

ABSTRACT

We use *Hubble Space Telescope* (*HST*) imaging and near-infrared spectroscopy from Keck/Multi-Object Spectrometer For Infra-Red Exploration (MOSFIRE) to study the substructure around the progenitor of a Milky Way-mass galaxy in the *Hubble Frontier Fields* (HFF). Specifically, we study an $r_e = 40^{+70}_{-30}$ pc, $M_\star \sim 10^{8.2}$ M$_\odot$ rest-frame ultraviolet luminous 'clump' at a projected distance of $\sim$100 pc from a $M_\star \sim 10^{9.8}$ M$_\odot$ galaxy at $z = 2.36$ with a magnification $\mu = 5.21$. We measure the star formation history of the clump and galaxy by jointly modelling the broad-band spectral energy distribution from *HST* photometry and Hα from MOSFIRE spectroscopy. Given our inferred properties (e.g. mass, metallicity, dust) of the clump and galaxy, we explore scenarios in which the clump formed *in situ* (e.g. a star-forming complex) or *ex situ* (e.g. a dwarf galaxy being accreted). If it formed *in situ*, we conclude that the clump is likely a single entity as opposed to a aggregation of smaller star clusters, making it one of the most dense star clusters catalogued. If it formed *ex situ*, then we are witnessing an accretion event with a 1:40 stellar mass ratio. However, our data alone are not informative enough to distinguish between *in situ* and *ex situ* scenarios to a high level of significance. We posit that the addition of high-fidelity metallicity information, such as [O III] 4363 Å, which can be detected at modest signal-to-noise ratio with only a few hours of *James Webb Space Telescope* (*JWST*)/Near-Infrared Spectrograph (NIRSpec) time, may be a powerful discriminant. We suggest that studying larger samples of moderately lensed substructures across cosmic time can provide unique insight into the hierarchical formation of galaxies like the Milky Way.

Key words: globular clusters: general – galaxies: high-redshift – Local Group.

1 INTRODUCTION

Galactic archaeology of the Milky Way (MW) is among the primary test beds for hierarchical galaxy assembly (e.g. Freeman & Bland-Hawthorn 2002; Madau et al. 2008; Bovy et al. 2012; Brown et al. 2012). The ages, abundances, and kinematics of individual stars and star clusters (globular and otherwise) encode the entire formation history of the MW back to the dawn of star formation. Increasingly detailed studies suggest that our MW had an active accretion and star formation history (SFH) in the early Universe, but has mainly undergone secular evolution in the last several Gyr (e.g. Belokurov et al. 2018; Haywood et al. 2018; Helmi et al. 2018).

A number of studies have extended such archaeological techniques beyond the MW. For example, results from the Pan-Andromeda Archaeological Survey (PAndAS) suggest that M31 has had a more active recent accretion history than the MW (McConnachie et al. 2009). Similarly, other ambitious efforts with observations of resolved stars and/or integrated light are in the process of revealing the formation histories of MW analogues throughout the Local Volume (e.g. Mouhcine et al. 2005; Merritt et al. 2016; Monachesi et al. 2016).

Ideally, it should be possible to connect this galactic archaeology approach directly to observations of accretion and star formation of MW-like ancestors at higher redshifts. Though indirect studies relying on galaxy number densities have shed light on hierarchical formation (Leja, van Dokkum & Franx 2013; van Dokkum et al. 2013), establishing direct links has proven challenging. This is primarily because even fairly substantial accretion events in the

★ E-mail: tzick@berkeley.edu

history of the MW (e.g. Gaia-Enceledus; Helmi et al. 2018) would amount to a 1:10 minor merger, which lies at the mass ratio limit of current high-redshift merger classification schemes (Lotz et al. 2008, 2011; Ribeiro et al. 2017).

At the same time, observations of the *Hubble Frontier Fields* (HFF; Lotz et al. 2017) are poised to transform our capacity to study substructure (e.g. minor accretion events, star cluster formation) in and around the progenitors of MW-like galaxies at high redshift. Compared to blank fields, the magnification power of gravitational lensing in the HFF provides both the spatial resolution and sensitivity to identify and characterize luminous features with sizes and luminosities comparable to star-forming regions and bright satellite galaxies in the local Universe over most of cosmic time (e.g. Kawamata et al. 2015, 2018; Swinbank et al. 2015; Bouwens et al. 2017a,b).

Intermediate-redshift galaxies ($1 < z < 3$) have long been known to have higher degrees of substructure (usually referred to as 'clumps') than their lower redshift counterparts in rest-frame optical wavelengths (e.g. Förster Schreiber et al. 2011; Genzel et al. 2011; Elmegreen, Malhotra & Rhoads 2012; Wuyts et al. 2012; Shibuya et al. 2016). Pioneering studies in the rest-frame optical have even characterized the properties of these clumps as a function of spatial distribution within their host galaxies (e.g. Guo et al. 2018; Zanella et al. 2019). However, the higher spatial resolution afforded by lensing has shown ~1 kpc clumps to either be multiple more compact clumps, or simply smaller clumps with stellar masses overestimated by up to an order of magnitude (e.g. Dessauges-Zavadsky et al. 2017; Cava et al. 2018). In principle, such observations can begin to reveal the degree of substructure around high-redshift galaxies, allowing us to directly explore hierarchical galaxy formation scenarios in the early Universe.

However, even with improvements in spatial resolution due to lensing, it remains challenging to discern the true nature of these clumps (i.e. *in situ* star formation versus accretion). For example, Zanella et al. (2015) identified an off-centre clump in [O III] Integral Field Unit spectroscopy data that was not detected in broad-band filter observations of a $z = 1.9$ galaxy. Though the [O III] implied it was a young cluster, the spatial resolution of the line emission identified clump was ~500 pc, which is too coarse for even a star-forming complex (see Krumholz, McKee & Bland-Hawthorn 2019, and references therein). Additionally, Rujopakarn et al. (2019) used Atacama Large Millimeter/submillimeter Array (ALMA)'s 30 mas capabilities to study dust clumping on 200 pc scales in $L_\star$ galaxies identified in the *Hubble Ultra Deep Field* (HUDF). However, the resolution of the HUDF imaging is coarser than that of ALMA, meaning that ~100 pc counterparts in the rest-frame optical would likely be smoothed out to much larger effective sizes (Tamburello et al. 2017; Gullberg et al. 2018), creating a mismatch between the dust and galaxy scales. Finally, at $z \sim 6$, Vanzella et al. (2019) combined HFF photometry with the Multi Unit Spectroscopic Explorer (MUSE) spectroscopy to study Lyman α (Lyα) emission from a clump consistent with $r_e < 15$ pc within a 100 pc dwarf galaxy. However, due to intergalactic medium (IGM) absorption, Lyα is of limited use as a star formation rate (SFR) or kinematics indicator. Spectroscopic SFRs (e.g. from Hα) are necessary to probe star formation within tens of Myr (i.e. the formation time-scales of star clusters), but such observations are largely lacking.

In this work, we undertake a joint HFF and Keck spectroscopic study of substructure around the progenitor of a MW mass galaxy. Specifically, we combine *K*-band Keck 1/Multi-Object Spectrometer For Infra-Red Exploration (MOSFIRE) spectroscopy with deep HFF photometry to study an off-centre (~100 pc from the galaxy centre) $r_e = 40^{+70}_{-30}$ pc, $M_\star \sim 10^{8.2}$ M$_\odot$ clump within a $r_e = 1.1^{+0.1}_{-0.3}$ kpc, MW progenitor mass ($M_\star \sim 10^{9.8}$ M$_\odot$) galaxy at $z = 2.36$, with a magnification of $\mu = 5.21$. We jointly model the photometric spectral energy distribution (SED) and Hα to determine the SFH of the host galaxy and clump, assuming scenarios in which the 1:40 mass ratio clump is forming *in situ* and in which it is being accreted. Our joint analysis of broad-band photometry and near-infrared spectroscopy illustrates the scientific promise of the data that *Hubble* and the *James Webb Space Telescope* (*JWST*) will provide at most redshifts for studying substructure, while also highlighting some of the outstanding challenges that must be addressed to increase our understanding of hierarchical galaxy formation at high redshifts.

2 METHODOLOGY

We selected targets for spectroscopic follow-up from the HFF MACS J1149, MACS J0717, and Abell 370 clusters, using the foreground subtracted Shipley et al. (2018) catalogues with coverage in F275W, F336W, F435W, F606W, F814W, F105W, F125W, F140W, F160W, and Infrared Array Camera (IRAC)/Multiband Imaging Photometer (MIPS) from *Spitzer*.

Initial sizes for our catalogue were derived with an updated version of methodology described in Bouwens et al. (2017a). In short, we mask nearby sources and apply a lensed Sérsic fit for the intrinsic brightness of the source and its size simultaneously using a Markov chain Monte Carlo (MCMC) methodology. As delensing an image introduces error and bias depending on the magnification map used, we instead employ a forward modelling approach. Namely, we account for magnification in our models by distorting each pixel by the requisite shear and amplification factors necessary to mimic the effects of lensing. For this to be computationally tractable, we fix the total magnification for our models to the median value from the CATS (Jullo & Kneib 2009; Richard et al. 2014; Jauzac et al. 2015a,b), Bradač et al. (2009), GLAFIC (Oguri 2010; Ishigaki et al. 2015; Kawamata et al. 2016), and Johnson et al. (2014) lensing maps. We discuss the updated size measurement for this source in Section 3.

When possible, we further selected for sources with photometrically inferred high equivalent width in [O III] or Hα, depending on redshift. This was done by running the SED fitting code, FAST (Kriek et al. 2009, 2018), on our sources and computing the line contamination to the continuum fit from [O III] or Hα. Specifically, we fit the photometry excluding the band containing the relevant line to measure the stellar continuum. We then convolve the appropriate filter with the best-fitting SED and used the difference between this continuum value and the measured photometry to infer line contamination in the band. For sources outside the proper redshift range, we instead selected for high SFRs from the SED fit. This SED fit corresponds to the photometry of the full galaxy, and was only used for selecting sources. We discuss the improved SED fit for this source as well as the SED fits for the clump at the end of this section and in Section 4.

The galaxy we highlight in this work is faint by traditional spectroscopic targeting standards (F160W = 25, not correcting for magnification). However, it met our selection criteria for a potentially observable emission line. An early variation of our size measurement routine was biased towards detecting clumps rather than extended sources, which serendipitously resulted in this interesting source falling into our sample. We discuss the full spectroscopic sample in a follow-up paper.
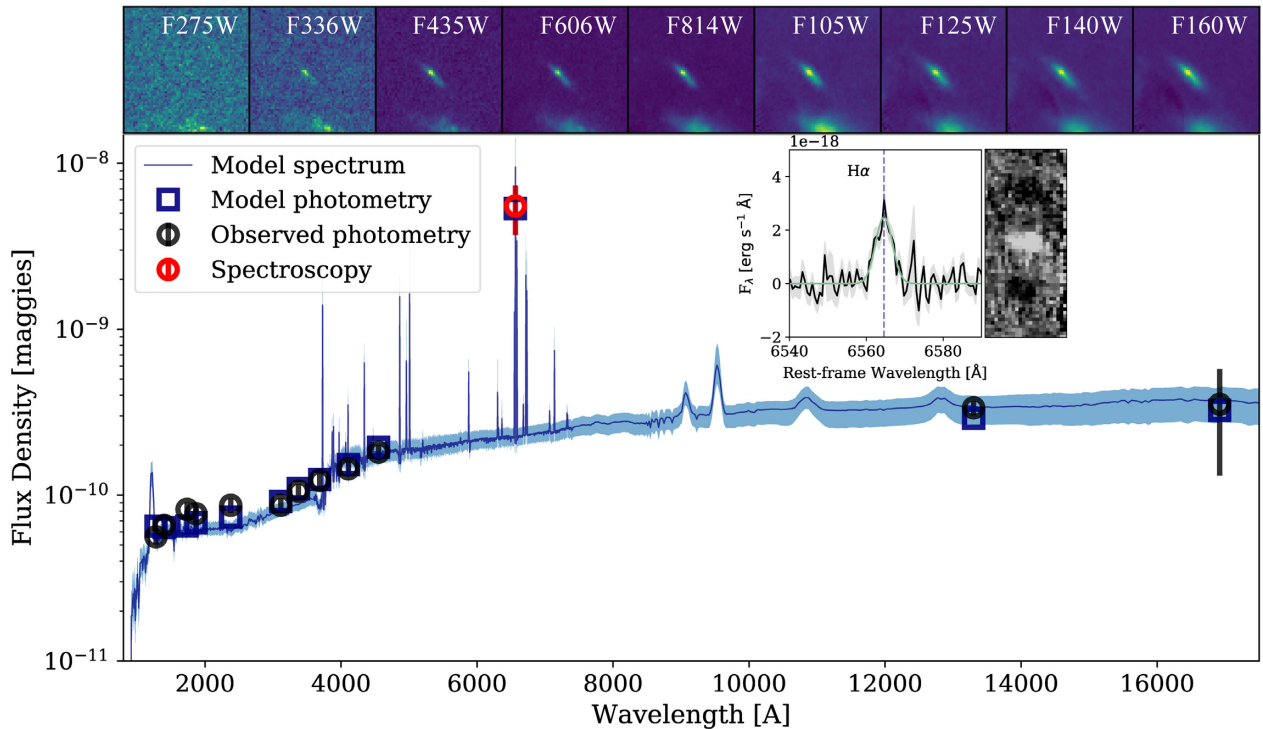
Figure 1. The *HST* images, SED, and MOSFIRE spectrum of the galaxy plus clump. Main panel: PROSPECTOR non-parametric SED fit including constraints from our MOSFIRE Hα measurement (shown in orange). The 16th and 84th percentile fits to the SED are shown in light blue. Inset panel: MOSFIRE Hα measurement, where the black line is our raw data, the grey is our error, and the teal is our fit to the line. Top panels: non-point spread function (PSF) matched photometry in each *HST* filter considered.

The spectroscopic Hα detection of this object shown in Fig. 1 is the result of 6 h of integration split over two half-nights. The reduction was conducted with the MOSFIRED pipeline, an updated version of the reduction procedure described in Kriek et al. (2015). As the source is entirely unresolved, we use an aperture correction for a point source, using a star observed in one of our MOSFIRE slits. No differential aperture correction was required for flux calibration purposes. We then measure a spectroscopic redshift by fitting a Gaussian to the Hα line and taking the mean and standard deviation of 10 000 realizations perturbed within the noise as our measurement and error, respectively.

We measure the velocity dispersion of this source by fitting a Gaussian to the Hα emission line and correcting for the instrument resolution using skylines in the 200 Å redwards and bluewards of Hα. We again derive errors through a Monte Carlo method.

To measure a non-parametric SFH for our source, we simultaneously fit our MOSFIRE spectrum, and 12 bands of photometry using the Bayesian spectral fitting code PROSPECTOR (Leja et al. 2017; Johnson et al. 2019). To minimize the noise introduced by the lack of continuum detection in our spectrum, we incorporate it into the fit by creating a synthetic narrow-band filter over solely Hα and including it as an additional band in our photometric fitting, correcting for magnification effects. For the non-parametric fit, we use 13 age bins that are uniformly logarithmically sampled at early times, with finer resolution in the most recent 1 Gyr. To avoid biasing our SFH with our choice of age bins, we include more bins than our projected resolution on SFH, and apply a continuity prior. Following Leja et al. (2019), we apply a stellar-mass–stellar-metallicity prior using a version of the Gallazzi et al. (2005) relation where we have doubled the width of the confidence interval to accommodate any redshift evolution.

Table 1. Summary of priors used to fit the SED for the whole galaxy. We fit mass and metallicity using a joint prior, based on the Gallazzi et al. (2005) mass–metallicity relation with artificially inflated scatter to account for redshift effects. We use the Kriek & Conroy (2013) dust parametrization and therefore fit for dust index and A_v. We also include gas parameters in our model. Though we do not expect our observations to constrain these, including them in our model allows us to treat them as nuisance parameters rather than fixing them to set values.

| Parameter	Prior type	Prior range
Mass/metallicity	Gaussian	$7 < M < 12.5\,\mathrm{M_\odot}/-1.99 < \log(Z) < 0.0$
Attenuation (A_v)	Top-hat	$0 < A_v < 1.5$
Dust index (δ)	Top-hat	$-2.0 < \delta < 0.5$
Gas log(Z)	Top-hat	$-1.99 < \mathrm{gas}\,\log(Z) < 0$
Ionization	Top-hat	$-4 < \log(U) < -1$

We then use the Bayesian nested sampling code DYNESTY to sample the posterior (Speagle 2020), with priors reported in Table 1. Our most probable parameters are reported in Table 2, where errors reflect the 16th and 84th percentiles of the posterior for each parameter, with the uncertainty due to magnification added in quadrature. The magnification uncertainties are modest owing to the small magnification of this source and the geometric constraints afforded by the spectroscopy.

We show the SFH from our PROSPECTOR fit in Fig. 2. We find the star formation is consistent with a rising τ over a period of a Gyr, with a potential underlying older stellar population at earlier times. However, the amplitude of the uncertainties precludes a more detailed interpretation of the older SFH.

We quantify the information added by the Hα measurement by modelling the SED without it. As expected, the net effect is that

Table 2. Summary of source and clump properties form our PROSPECTOR fits to the photometry from the full galaxy and the clump with *ex situ* and *in situ* priors.

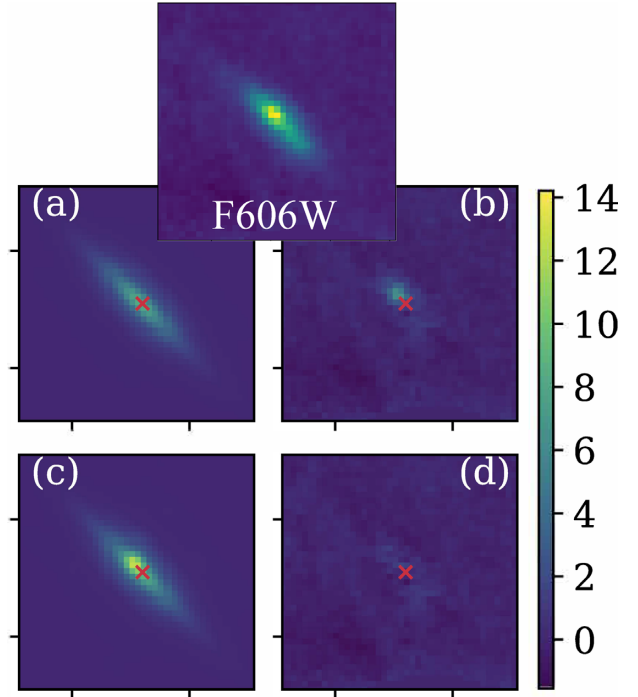| $\mu = 5.14$	Full galaxy	Clump only (*ex situ*)	Clump only (*in situ*)
Redshift	2.3695 ± 0.0005	–	–
$\log(M)$ ($M_\odot$)	$9.8^{+0.3}_{-0.2}$	$8.2^{+0.3}_{-0.2}$	$8.1^{+0.2}_{-0.1}$
A_v	$0.44^{+0.13}_{-0.12}$	$0.15^{+0.05}_{-0.08}$	$0.37^{+0.08}_{-0.02}$
$\log(Z)$	$-0.22^{+0.30}_{-0.35}$	$-1.35^{+0.75}_{-0.43}$	$-0.18^{+0.16}_{-0.23}$
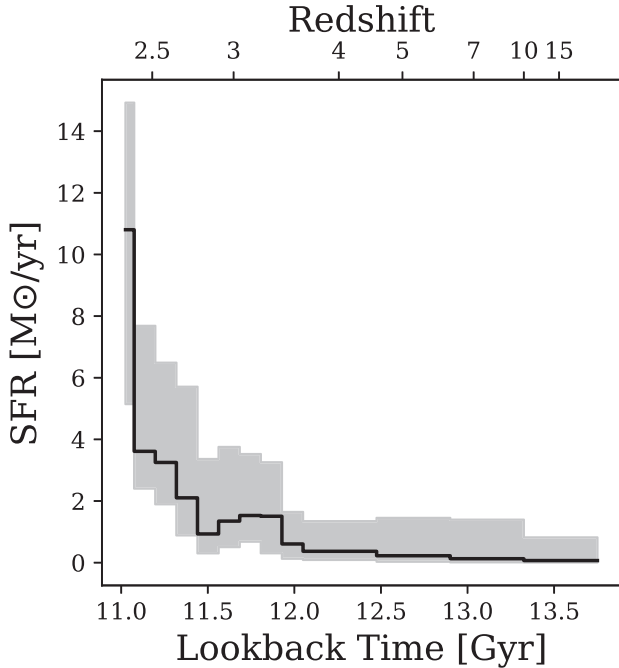
Figure 2. SFH from a non-parametric PROSPECTOR fit. We use 13 age bins in our fit with higher time resolution in the most recent 1 Gyr bins. To avoid biasing our SFH with our choice of age bins, we include more bins than our projected resolution on SFH and apply a continuity prior.

Figure 3. Example of size measurement procedure for F606W image. (a) Best-fitting Sérsic $n = 1$ model. (b) The residual for the $n = 1$ sérsic model with an 'x' demarcating the centre of the galaxy accounting for shear. The projected distance from the clump to the centre of the source is ∼100 pc. (c) Best-fitting Gaussian + Sérsic $n = 1$ model. (d) Residuals for the Gaussian + Sérsic $n = 1$ model. It is apparent from the last panel that there is an off-centre clump in the galaxy.

the inclusion of Hα provides a tighter constraint on star formation within the most recent 50 Myr, compared to an SED-only fit. Star formation at early times is also modestly more constrained (i.e. the probability distribution function corresponding to bins beyond 1.5 Gyr is narrower) when our Hα measurement is included.

3 SIZE MEASUREMENTS

With a redshift in hand, we are able to improve our estimate of the clump's size. To do this, we fix our models to the spectroscopic redshift of the source. This constraint allows us to measure a more precise magnification, decreasing the original size measured for this object by 5 per cent over the photometric redshift-derived size.

To measure the size of the source we began with an unconstrained Sérsic profile and found that we could not simultaneously fit both the diffuse light component and the clump for any Sérsic index from 1 to 4. In an attempt to constrain both components simultaneously, we tested an exponential disc + Sérsic model, an Sérsic model + Gaussian model, and an exponential disc with a flux-scaled PSF. We find that the latter results in the best residuals by far for all HFF bands, indicating that the clump is essentially unresolved. A

conservative upper limit on the size of the clump is therefore the size of the PSF itself, which is 108 pc in F606W.

For a more detailed estimate of the clump size, we fix the Sérsic profile of the galaxy and simultaneously compute the residuals for a grid of model parameters for a Gaussian clump (size and brightness). We considered values of $r \approx 0.02$ pixels to $r = 0.5$ pixels for the radius of the Gaussian component and show the resulting model derived magnitudes in Table B1. The resulting probability distribution allowed us to constrain the maximum likelihood size of the clump to $r_e = 40^{+70}_{-30}$ pc in F606W.

Using these model-derived parameters, we compare to the background subtracted image within a 1 arcsec aperture, and find the 16σ detection shown in panel (b) of Fig. 3. We measure the projected distance of the clump from the centre of the Sérsic fit to the diffuse component to be 100^{+70}_{-30} pc. As the centre of the Sérsic fit to the diffuse component is well constrained, magnification uncertainties dominate our reported uncertainties. This galaxy appears edge-on due to lensing shear, we therefore assume a typical inclination of

$45°$ to estimate the true distance (d) of the clump to be 140 pc. Comparing this to the effective radius (R_e) of the galaxy results in a $d/R_e \sim 0.1$.

A caveat to measuring size in bluer bands is that they are biased towards more recent star formation (e.g. Stark 2016). Though we do not expect substantial colour gradients in a galaxy of this mass, we assess the robustness of our measured clump by comparing the maximum a posteriori value of the MCMC fit for the clump as measured from the F160W to F606W bands. For this clump, the size changes by the resolution difference between F606W and F160W. Thus, this source appears more compact in bluer bands primarily due to its point source-like nature, rather than bias towards younger populations. As higher redshift studies are limited to probing similarly blue rest-frame wavelengths in determining sizes, our measured clump size can be directly compared.

4 MODELLING THE SED FOR *IN SITU* VERSUS ACCRETION SCENARIOS

There are two primary mechanisms that can explain the presence of a 1:40 mass ratio for a clump in a proto-MW galaxy: (i) *in situ* formation due to disc turbulence and gravitational collapse; (ii) star formation triggered by an accretion event with a neighbouring satellite, which may be analogous to the accretion of an Small Magellanic Cloud progenitor and a MW progenitor at $z \sim 2$ (Weisz et al. 2013). While the spatial resolution necessary to extract the dynamical information to distinguish these two scenarios will likely remain out of reach until the era of 30-m telescopes, we can gain some insight as to the likelihood of each scenario by combining spectral and photometric fitting with informed priors.

In the case of *in situ* disc fragmentation, we expect the metallicity parameters, stellar and nebular, to be similar to the host galaxy parameters. Though the SFH of the clump may differ (possibly dramatically) from the host galaxy, the fact that it is forming out of the galaxy's cold gas supply implies the metallicity of the clump should be similar to that of its host. To model this scenario, we take the metallicity prior on the clump to be within the 16–84th percentiles of the posterior distribution to the metallicity of the galaxy itself. We leave all other parameters as shown in Table 1. We assume that though the metallicity of the clump and host galaxy may vary, the metallicity of an *in situ* clump should not be outside the host galaxy's metallicity range. We then compare this to a 'free' fit to the clump, where the priors are described in Table 1, with the exception of the mass–metallicity prior, which does not apply at such low masses. We therefore instead use separate flat mass and metallicity priors but leave their ranges as shown in Table 1. The resulting SED fits for the two sets of priors are shown in panels (a) and (b) of Fig. 4, where the former corresponds to an accretion event and the latter to *in situ* formation.

We find that the model with *ex situ* priors yields smaller residuals than the *in situ* priors. However, the Bayesian evidence criterion is of order unity that does not signify a robust difference between the two sets of priors. We further show the corresponding SFH for each set of priors in panels (c) and (d). Though these appear qualitatively different, they both essentially correspond to an extremely young stellar population. We discuss future prospects for this archaeological approach to characterize substructure in Section 5.

5 DISCUSSION

In this section, we explore several plausible clump formation scenarios given the analysis presented in the previous sections. We attempt to identify features that may allow us to discriminate between accretion and *in situ* formation mechanisms and discuss how future observations and targeted theoretical studies will improve our ability to cleanly associate observed substructure with a formation scenario.

First, we consider the case that this clump formed *in situ* via turbulence-induced disc fragmentation (e.g. Dekel et al. 2009, 2013). Cosmological zoom-in simulations have characterized such clumps to have baryonic masses up to 10^9 M$_\odot$ and radii up to 1 kpc. Typical clumps with a baryonic mass corresponding to our measured stellar mass have characteristic radii of $\sim$250 pc (Mandelker et al. 2017; Oklopčić et al. 2017). However, as we are limited to measuring stellar mass (we have no gas measurement), our observed clump likely corresponds to a larger baryonic mass in simulations and therefore a larger R_e ($\sim$500 pc). The mass surface density we measure for the clump is above 10^3 M$_\odot$ pc^{-2}, which is denser than simulated clumps. Non-cosmological, hydrodynamic simulations described in Tamburello et al. (2015, 2017) do examine the stellar mass and radii of clumps directly as a function of resolution. We find good agreement with our clump mass and radii when comparing to their high-resolution (100 pc) mock observations.

We can also compare the mass fraction of the clump within the galaxy to simulations. Mandelker et al. (2014) find that *ex situ* clumps have a characteristically higher mass fraction than *in situ* ones. Comparing to their updated cosmological zoom-in simulations including feedback, we find that our 1:40 mass ratio is on the very massive end of their clump to host mass ratio distribution (Mandelker et al. 2017).

The location of our clump, so near the centre of its host, remains atypical for clumps formed *in situ*. Both theoretical (Mandelker et al. 2017; Oklopčić et al. 2017) and observational (Guo et al. 2012; Shibuya et al. 2016) studies of clump formation find clumps as blue and young as the one we observe further out than $d/R_e \sim$ 0.1. The d/R_e value we infer is smaller than the bulk of simulated clumps, but approaches the distance ratio seen in simulations for clumps older than our measured age (e.g. 1 Gyr in Mandelker et al. 2017) Finally, disc fragmentation in simulations ubiquitously leads to the formation of multiple clumps (Mandelker et al. 2014, 2017) and though we have sub-kpc resolution for the entire galaxy, we only identify one. Although it is possible that there are other clumps that remain obscured due to projection effects, it would require an unlikely geometry to hide all of them from discovery in our data. Given that this clump is detected at 16σ, it is unlikely that we are missing similarly massive clumps due to surface brightness limits.

A second *in situ* formation scenario is that this clump is not a single star-forming entity, but instead an unresolved blend of multiple smaller star clusters. We gain some insight into this possibility by comparing our clump to resolved extreme star-forming regions in the local Universe. For example, the central star-forming cluster of 30 Doradus, NGC 2070, is commonly referred to as an analogue for high-redshift star cluster formation (Ochsendorf et al. 2017; Leroy et al. 2018). Integrating the mass of all clusters in a 200-pc region centred on NGC 2070 yields a total stellar mass of $M_\star \sim 10^5$ M$_\odot$ (Cignoni et al. 2015), which is 1000 times less massive than our clump. This is perhaps not surprising given the differences in the host galaxy properties and relative star-forming conditions in the low- and high-redshift Universe. Nevertheless, it does illustrate just how massive a star-forming region our clump may be.

Perhaps a more apt comparison is to the merging Antennae system. The Antennae yields gas pressures analogous to those expected in high-redshift star formation (e.g. Cabrera-Ziri et al. 2016).
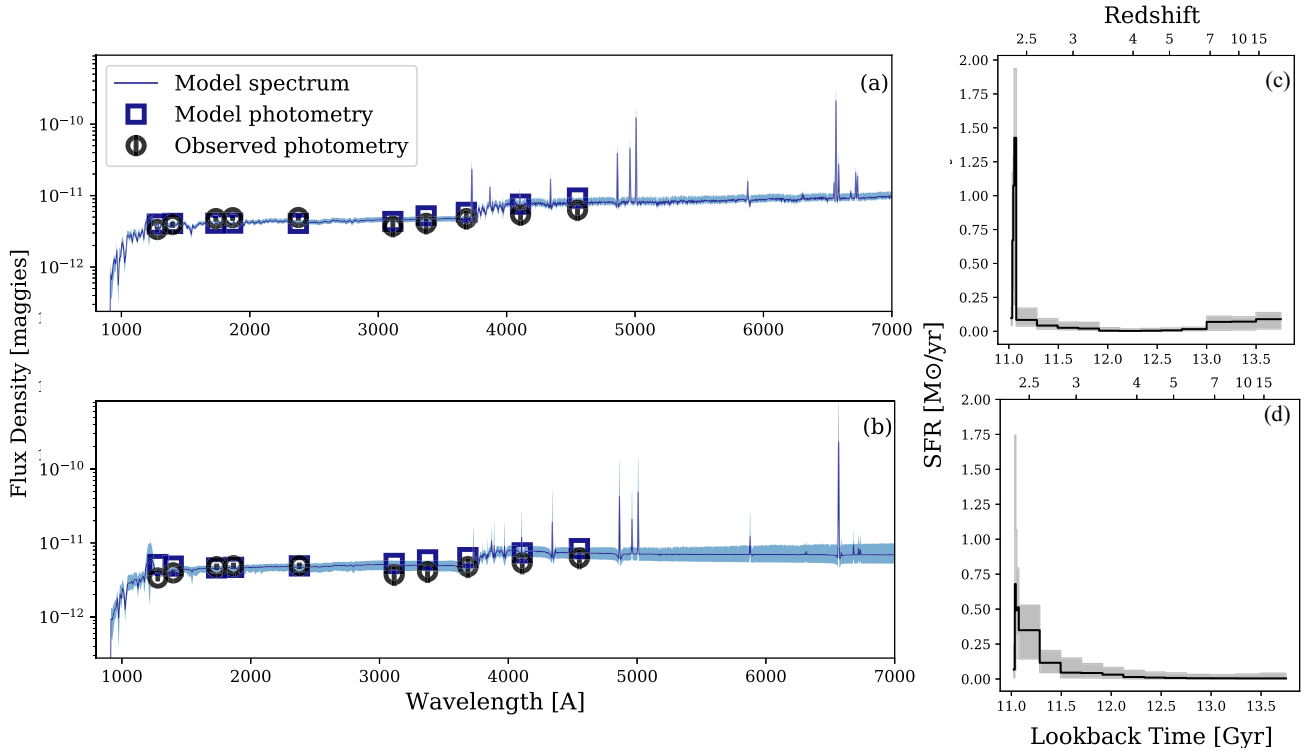
Figure 4. The SED fits and SFH for the model extracted clump photometry. (a) The SED fit corresponding to the *in situ* event priors, with the 16th and 84th percentile models shown in light blue. (b) The SED fit to the accretion formation scenario priors, plotted with the same conventions as panel (a). (c) The SFH corresponding to the *in situ* scenario. The grey show the 16th and 84th percentiles and the black line shows the median model. (d) The SFH corresponding to the accretion scenario plotted with the same conventions as panel (c). The accretion scenario is a slightly better fit to the data according to the smaller residuals and higher Bayesian evidence criteria. The SFHs resulting from both sets of priors indicate an extremely young stellar population. The burst of the accretion scenario seems to line up with the highest SFRs in the host galaxy, however, as our first SFH bin corresponds to 50 Myr (in concordance with the fit to the host galaxy), the difference in this small time bin is likely not well constrained by either model.

Integrating the stellar mass for the three most massive clusters in the Antennae yields a mass of $M_\star \sim 10^6$ M$_\odot$ (Johnson et al. 2015). Even assuming the 84th percentile radius for our clump, we would require a mass contamination from the host to be factor of 10 larger than inferred from high-resolution simulations (Tamburello et al. 2017) or nearby young molecular clouds (Hollyhead et al. 2015). Given our high spatial resolution and model-extracted photometry, we do not expect anywhere near as severe of a mass contamination. It is therefore unlikely that our clump is a set of very tightly packed massive star clusters.

A third possibility is that an off-centre star-forming clump may be an indication of an ongoing merger, or pristine-gas accretion event (e.g. Hopkins et al. 2012; D'Souza & Bell 2018). A comparison of the SFH we infer for the observed clump in Fig. 3 to the SFH of the main galaxy in Fig. 2 shows an extremely young stellar population whose formation corresponds to a peak in star formation in the main galaxy. This concurrence can be interpreted as consistent with an *in situ* formation scenario or a gas accretion scenario over an accretion event.

However, though both sets of model priors we used for the clump result in a fit consistent with a single burst of star formation, our *ex situ* priors yield a fit that allows for a more extended (~1 Gyr) clump SFH. Furthermore, it is difficult to compare our results to the literature on mergers as the redshift and mass range of our source has not been explored at similar resolutions. Morphological criteria such as *Gini*–M_{20} (Lotz, Primack & Madau 2004) or concentration and asymmetry (Conselice et al. 2014) have primarily

been used at lower redshifts (Lotz et al. 2011) or higher stellar masses (Man, Zirm & Toft 2016). Recent work by Nevin et al. (2019) uses linear decomposition of non-parametric morphological criteria in simulations to find an optimal merger identifying criteria. Although work in this vein is promising, it is unclear how well these criteria hold up at higher redshift (Thompson et al. 2015). Current criteria for mergers are categorically less sensitive to minor mergers (1:4–1:10 stellar mass ratios) and are insensitive to mergers with larger mass disparity. This means that a MW–Large Magellanic Cloud merger, for example, would likely not be identified with current criteria and a MW–Sagittarius merger would be completely undetectable. With more work in lensed fields, where substructure resolution to sub ~100 pc is possible, studying such minor mergers should be within reach.

Though we cannot conclusively determine the nature of this off-centre star-forming clump from our current data set, there are promising paths forward. For example, a larger sample of galaxies and wider spectral coverage with the *JWST* could enable the modelling technique presented in this work. A statistical approach targeting similarly low-magnification sources ($\mu < 10$) would allow us to disentangle line-of-site effects that limit our comparisons to simulations (i.e. number of clumps per source). Current explorations of clumps in the HFF generally require observations to approach surface brightness detection limits, with the sensitivity of Near-Infrared Camera (NIRCam), it should be possible to detect more clumps than observed in current HFF observations. Likewise, the spectral range, sensitivity, and angular resolution of Near-Infrared

Spectrograph (NIRSpec) will allow more efficient spectral analysis of clumps out to higher redshifts than available with current ground-based instruments.

Current estimates using the *JWST*/NIRSpec exposure time calculator show that Hα, Hβ, and [O III] can be detected with S/N > 5, in under 2 h for this galaxy, with faint lines like [O III] 4363 Å detected in 5 h. Such faint lines are easier to measure in more magnified and extreme star-forming sources, such as the $z = 2.69$, $\mu \sim 8$, dwarf galaxy identified in Gburek et al. (2019), for which *JWST*/NIRSpec would detect [O III] 4363 Å with S/N > 5 in under an hour of exposure time. These lines enable robust dust attenuation measurement and vastly improved constraints on metallicity. These can be used in conjunction with simulations to elucidate the origin of clumps. Furthermore, the multiplexing abilities of *JWST*/NIRSpec combined with the resolving power afforded by even modestly lensed fields enable a statistical approach to characterize clumps in lower mass and higher redshift systems than currently feasible.

6 CONCLUSIONS

We are entering an era of increased sensitivity that enables us to study substructure at $z > 2$ on ever smaller spatial and mass scales. Our current MOSFIRE study previews the potential of spectroscopy to connect substructure and SFH, as well as the challenges in interpreting even high spatial resolution photometry.

In this work, we present a combined spectral and photometric analysis of a lensed, $\mu = 5.14$, MW-mass progenitor galaxy at $z = 2.36$ with an off-centre star-forming clump. We have used precise redshifts to measure robust sizes for the clump and the main galaxy. Finally, we investigate the origin of the clump by leveraging our measurement of the main galaxy to inform priors on the fit to the clump. Our findings are as follows.

(i) A $\log(M_\star) = 8.2^{+3}_{-2}$ M$_\odot$ clump with $r_e = 40^{+70}_{-30}$ pc located $\sim$100 pc away from the galaxy centre. We find that the host galaxy has a mass of $\log(M_\star) = 9.8^{+0.3}_{0.2}$ M$_\odot$.

(ii) Comparing our inferred clump/galaxy masses, SFRs, and sizes to low-redshift analogues and simulations, we find that this clump is unlikely to be an aggregate of multiple less massive clumps. Given our measured stellar mass surface density of order 10^3 M$_\odot$ pc^{-2}, we conclude that if this clump formed *in situ*, it is one of the densest star-forming regions confirmed to date. If the clump formed *ex situ*, this system is undergoing a 1:40 mass ratio merger.

(iii) With the advent of *JWST* NIRCam and NIRSpec, we find that this analysis can be reasonably undertaken statistically. The sensitivity and broad-band capabilities of NIRSpec will allow measurement of important dust and metallicity indicators. While its angular resolution will allow a more robust comparison between clump and galaxy properties that can be used to infer clump origins.

ACKNOWLEDGEMENTS

REFERENCES

Belokurov V., Erkal D., Evans N. W., Koposov S. E., Deason A. J., 2018, MNRAS, 478, 611

Bouwens R. J., Illingworth G. D., Oesch P. A., Atek H., Lam D., Stefanon M., 2017a, ApJ, 843, 41

Bouwens R. J., Illingworth G. D., Oesch P. A., Maseda M., Ribeiro B., Stefanon M., Lam D., 2017b, preprint (arXiv:1711.02090)

Bovy J., Rix H.-W., Liu C., Hogg D. W., Beers T. C., Lee Y. S., 2012, ApJ, 753, 148

Bradač M. et al., 2009, ApJ, 706, 1201

Brown T. M. et al., 2012, ApJ, 753, L21

Cabrera-Ziri I., Lardo C., Davies B., Bastian N., Beccari G., Larsen S. S., Hernandez S., 2016, MNRAS, 460, 1869

Cava A., Schaerer D., Richard J., Pérez-González P. G., Dessauges-Zavadsky M., Mayer L., Tamburello V., 2018, Nat. Astron., 2, 76

Cignoni M. et al., 2015, ApJ, 811, 76

Conselice C. J., Bluck A. F. L., Mortlock A., Palamara D., Benson A. J., 2014, MNRAS, 444, 1125

Dekel A. et al., 2009, Nature, 457, 451

Dekel A., Zolotov A., Tweed D., Cacciato M., Ceverino D., Primack J. R., 2013, MNRAS, 435, 999

Dessauges-Zavadsky M., Schaerer D., Cava A., Mayer L., Tamburello V., 2017, ApJ, 836, L22

D'Souza R., Bell E. F., 2018, Nat. Astron., 2, 737

Elmegreen B. G., Malhotra S., Rhoads J., 2012, ApJ, 757, 9

Förster Schreiber N. M. et al., 2011, ApJ, 739, 45

Freeman K., Bland-Hawthorn J., 2002, ARA&A, 40, 487

Gallazzi A., Charlot S., Brinchmann J., White S. D. M., Tremonti C. A., 2005, MNRAS, 362, 41

Gburek T. et al., 2019, ApJ, 887, 168

Genzel R. et al., 2011, ApJ, 733, 101

Gullberg B. et al., 2018, ApJ, 859, 12

Guo Y., Giavalisco M., Ferguson H. C., Cassata P., Koekemoer A. M., 2012, ApJ, 757, 120

Guo Y. et al., 2018, ApJ, 853, 108

Haywood M., Di Matteo P., Lehnert M., Snaith O., Fragkoudi F., Khoperskov S., 2018, A&A, 618, A78

Helmi A., Babusiaux C., Koppelman H. H., Massari D., Veljanoski J., Brown A. G. A., 2018, Nature, 563, 85

Hollyhead K., Bastian N., Adamo A., Silva-Villa E., Dale J., Ryon J. E., Gazak Z., 2015, MNRAS, 449, 1106

Hopkins P. F., Kereš D., Murray N., Quataert E., Hernquist L., 2012, MNRAS, 427, 968

Ishigaki M., Kawamata R., Ouchi M., Oguri M., Shimasaku K., Ono Y., 2015, ApJ, 799, 12

Jauzac M. et al., 2015a, MNRAS, 446, 4132

Jauzac M. et al., 2015b, MNRAS, 452, 1437

Johnson T. L., Sharon K., Bayliss M. B., Gladders M. D., Coe D., Ebeling H., 2014, ApJ, 797, 48

Johnson K. E., Leroy A. K., Indebetouw R., Brogan C. L., Whitmore B. C., Hibbard J., Sheth K., Evans A. S., 2015, ApJ, 806, 35

Johnson B. D., Leja J. L., Conroy C., Speagle J. S., 2019, Astrophysics Source Code Library, record ascl:1905.025

Jullo E., Kneib J.-P., 2009, MNRAS, 395, 1319

Kawamata R., Ishigaki M., Shimasaku K., Oguri M., Ouchi M., 2015, ApJ, 804, 103

Kawamata R., Oguri M., Ishigaki M., Shimasaku K., Ouchi M., 2016, ApJ, 819, 114

Kawamata R., Ishigaki M., Shimasaku K., Oguri M., Ouchi M., Tanigawa S., 2018, ApJ, 855, 4

Kriek M., Conroy C., 2013, ApJ, 775, L16

Kriek M., van Dokkum P. G., Labbé I., Franx M., Illingworth G. D., Marchesini D., Quadri R. F., 2009, ApJ, 700, 221
Kriek M. et al., 2015, ApJS, 218, 15
Kriek M. et al., 2018, Astrophysics Source Code Library, record ascl:1803.008
Krumholz M. R., McKee C. F., Bland-Hawthorn J., 2019, ARA&A, 57, 227
Leja J., van Dokkum P., Franx M., 2013, ApJ, 766, 33
Leja J., Johnson B. D., Conroy C., van Dokkum P. G., Byler N., 2017, ApJ, 837, 170
Leja J., Carnall A. C., Johnson B. D., Conroy C., Speagle J. S., 2019, ApJ, 876, 3
Leroy A. K. et al., 2018, ApJ, 869, 126
Lotz J. M., Primack J., Madau P., 2004, AJ, 128, 163
Lotz J. M. et al., 2008, ApJ, 672, 177
Lotz J. M., Jonsson P., Cox T. J., Croton D., Primack J. R., Somerville R. S., Stewart K., 2011, ApJ, 742, 103
Lotz J. M. et al., 2017, ApJ, 837, 97
McConnachie A. W. et al., 2009, Nature, 461, 66
Madau P., Kuhlen M., Diemand J., Moore B., Zemp M., Potter D., Stadel J., 2008, ApJ, 689, L41
Man A. W. S., Zirm A. W., Toft S., 2016, ApJ, 830, 89
Mandelker N., Dekel A., Ceverino D., Tweed D., Moody C. E., Primack J., 2014, MNRAS, 443, 3675
Mandelker N., Dekel A., Ceverino D., DeGraf C., Guo Y., Primack J., 2017, MNRAS, 464, 635
Merritt A., van Dokkum P., Abraham R., Zhang J., 2016, ApJ, 830, 62
Monachesi A., Bell E. F., Radburn-Smith D. J., Bailin J., de Jong R. S., Holwerda B., Streich D., Silverstein G., 2016, MNRAS, 457, 1419
Mouhcine M., Ferguson H. C., Rich R. M., Brown T. M., Smith T. E., 2005, ApJ, 633, 810
Nevin R., Blecha L., Comerford J., Greene J., 2019, ApJ, 872, 76
Ochsendorf B. B., Zinnecker H., Nayak O., Bally J., Meixner M., Jones O. C., Indebetouw R., Rahman M., 2017, Nat. Astron., 1, 784
Oguri M., 2010, PASJ, 62, 1017
Oklopčić A., Hopkins P. F., Feldmann R., Kereš D., Faucher-Giguère C.-A., Murray N., 2017, MNRAS, 465, 952
Ribeiro B. et al., 2017, A&A, 608, A16
Richard J. et al., 2014, MNRAS, 444, 268
Rujopakarn W. et al., 2019, ApJ, 882, 107
Shibuya T., Ouchi M., Kubo M., Harikane Y., 2016, ApJ, 821, 72
Shipley H. V. et al., 2018, ApJS, 235, 14
Speagle J. S., 2020, MNRAS, 493, 3132
Stark D. P., 2016, ARA&A, 54, 761
Swinbank A. M. et al., 2015, ApJ, 806, L17
Tamburello V., Mayer L., Shen S., Wadsley J., 2015, MNRAS, 453, 2490
Tamburello V., Rahmati A., Mayer L., Cava A., Dessauges-Zavadsky M., Schaerer D., 2017, MNRAS, 468, 4792
Thompson R., Davé R., Huang S., Katz N., 2015, preprint (arXiv:1508.01851)
van Dokkum P. G. et al., 2013, ApJ, 771, L35
Vanzella E. et al., 2019, MNRAS, 483, 3618
Weisz D. R., Dolphin A. E., Skillman E. D., Holtzman J., Dalcanton J. J., Cole A. A., Neary K., 2013, MNRAS, 431, 364
Wuyts S. et al., 2012, ApJ, 753, 114
Zanella A. et al., 2015, Nature, 521, 54
Zanella A. et al., 2019, MNRAS, 489, 2792

APPENDIX A: SED FITS AND INFERRED SFH WITH VARYING STELLAR MODELS

There is currently no standard universally excepted stellar modelling framework in the galaxies literature, potentially leading to systematic offsets in inferred properties across different studies. We investigate alternate isochrone and spectral libraries and their impact on our analysis. We examine both the BPASS models with binaries, as well as the popular BC03 that we replicate by combining the BASEL spectral libraries with the GENEVA isochrones. We show the results of our PROSPECTOR fits in Fig. A1. We find that both BPASS and BC03 produce a more peaked median SFH than the MIST models (see Fig. 4). However an extremely young burst of star formation is favoured by all three models.
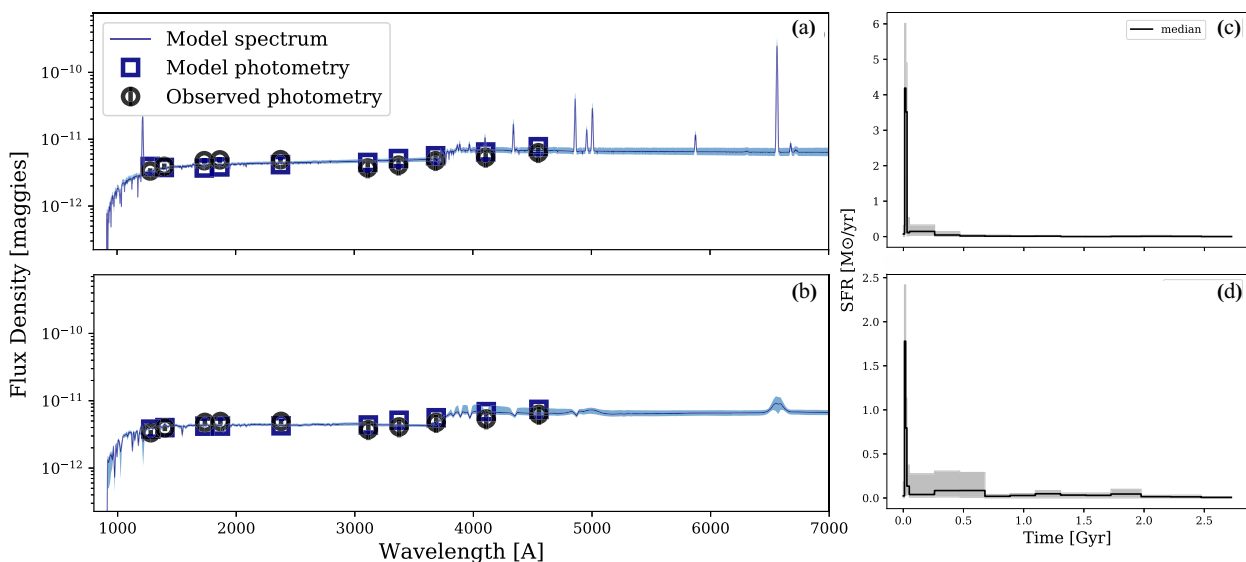

Figure A1. The SED fits and SFH for the model extracted clump photometry with alternate stellar libraries. In all plots the shaded regions correspond to the 16th and 84th percentiles. (a) SED fit using the BPASS models with binaries. (b) SED fit using BC03 models. (c) SFH corresponding to BPASS model fit. (d) SFH corresponding to the BC03 models. Both models show a more peaked median clump SFH, however all models are consistent with a young star-forming burst.

APPENDIX B: INTRINSIC MAGNITUDES MEASURED FOR THE GALAXY AND CLUMP

Table B1. Measured intrinsic magnitude for the full galaxy and the clump. The uncertainties on the clump flux are from the model fit and represent the 16th and 84th percentiles of the posterior with the uncertainty in magnification added in quadrature. The uneven uncertainties are due to magnification uncertainties.

Filter	Full galaxy	Clump
F336W	$26.67^{+0.35}_{-0.14}$	$28.99^{+1.51}_{-2.50}$
F435W	$25.63^{+0.33}_{-0.05}$	$28.81^{+0.67}_{-1.65}$
F475W	$25.46^{+0.33}_{-0.05}$	$28.63^{+0.52}_{-1.50}$
F606W	$25.21^{+0.33}_{-0.04}$	$28.33^{+0.41}_{-1.39}$
F625W	$25.28^{+0.33}_{-0.08}$	$28.34^{+0.49}_{-2.78}$
F814W	$25.15^{+0.33}_{-0.04}$	$28.35^{+0.49}_{-1.45}$
F105W	$25.152^{+0.33}_{-0.04}$	$28.41^{+0.45}_{-1.48}$
F110W	$24.94^{+0.33}_{-0.05}$	$28.56^{+0.40}_{-4.38}$
F125W	$24.78^{+0.33}_{-0.04}$	$28.05^{+0.41}_{-1.39}$
F140W	$24.60^{+0.33}_{-0.04}$	$27.98^{+0.38}_{1.36}$
F160W	$24.35^{+0.33}_{-0.04}$	$27.95^{+0.43}_{-1.41}$
IRAC 2	$23.68^{+0.33}_{-0.06}$	–
IRAC 3	$23.63^{+0.76}_{-0.68}$	–

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