

Timing the earliest quenching events with a robust sample of massive quiescent galaxies at $2 < z < 5$

A. C. Carnall,^{1*} S. Walker¹, R. J. McLure,¹ J. S. Dunlop,¹ D. J. McLeod,¹ F. Cullen¹, V. Wild², R. Amorin^{3,4}, M. Bolzonella,⁵ M. Castellano,⁶ A. Cimatti,^{7,8} O. Cucciati⁵, A. Fontana,⁶ A. Gargiulo,⁹ B. Garilli,⁹ M. J. Jarvis^{10,11}, L. Pentericci,⁶ L. Pozzetti,⁵ G. Zamorani,⁵ A. Calabro,⁶ N. P. Hathi¹² and A. M. Koekemoer¹²

¹Scottish Universities Physics Alliance, Institute for Astronomy, University of Edinburgh, Royal Observatory, Edinburgh EH9 3HJ, UK

²Scottish Universities Physics Alliance, School of Physics and Astronomy, University of St Andrews, North Haugh, St Andrews KY16 9SS, UK

³Instituto de Investigación Multidisciplinar en Ciencia y Tecnología, Universidad de La Serena, Raúl Bitrán 1305, La Serena, Chile

⁴Departamento de Física y Astronomía, Universidad de La Serena, Av. Juan Cisternas 1200 Norte, La Serena, Chile

⁵INAF – OAS Bologna, Via P. Gobetti 93/3, I-40129 Bologna, Italy

⁶INAF – Osservatorio Astronomico di Roma, Via Frascati 33, I-00078 Monteporzio Catone, Italy

⁷University of Bologna, Department of Physics and Astronomy (DIFA), Via Gobetti 93/2, I-40129 Bologna, Italy

⁸INAF – Osservatorio Astrofisico di Arcetri, Largo E. Fermi 5, I-50125 Firenze, Italy

⁹INAF – IASF Milano, Via A. Corti 12, I-20133 Milano, Italy

¹⁰Department of Physics, University of Oxford, Keble Road, Oxford OX1 3RH, UK

¹¹Department of Physics and Astronomy, University of the Western Cape, Private Bag X17, Bellville, Cape Town 7535, South Africa

¹²Space Telescope Science Institute, 3700 San Martin Dr, Baltimore, MD 21218, USA

Accepted 2020 May 26. Received 2020 May 18; in original form 2020 February 3

ABSTRACT

We present a sample of 151 massive ($M_* > 10^{10} M_\odot$) quiescent galaxies at $2 < z < 5$, based on a sophisticated Bayesian spectral energy distribution fitting analysis of the CANDELS UDS and GOODS-South fields. Our sample includes a robust sub-sample of 61 objects for which we confidently exclude low-redshift and star-forming solutions. We identify 10 robust objects at $z > 3$, of which 2 are at $z > 4$. We report formation redshifts, demonstrating that the oldest objects formed at $z > 6$; however, individual ages from our photometric data have significant uncertainties, typically ~ 0.5 Gyr. We demonstrate that the UVJ colours of the quiescent population evolve with redshift at $z > 3$, becoming bluer and more similar to post-starburst galaxies at lower redshift. Based upon this, we construct a model for the time evolution of quiescent galaxy UVJ colours, concluding that the oldest objects are consistent with forming the bulk of their stellar mass at $z \sim 6-7$ and quenching at $z \sim 5$. We report spectroscopic redshifts for two of our objects at $z = 3.440$ and 3.396 , which exhibit extremely weak Ly α emission in ultra-deep VANDELS spectra. We calculate star formation rates based on these line fluxes, finding that these galaxies are consistent with our quiescent selection criteria, provided their Ly α escape fractions are > 3 and > 10 per cent, respectively. We finally report that our highest redshift robust object exhibits a continuum break at $\lambda \sim 7000 \text{ \AA}$ in a spectrum from VUDS, consistent with our photometric redshift of $z_{\text{phot}} = 4.72^{+0.06}_{-0.04}$. If confirmed as quiescent, this object would be the highest redshift known quiescent galaxy. To obtain stronger constraints on the times of the earliest quenching events, high-SNR spectroscopy must be extended to $z \gtrsim 3$ quiescent objects.

Key words: methods: statistical – galaxies: evolution – galaxies: star formation.

1 INTRODUCTION

Understanding the processes that shaped the most massive galaxies in the local Universe is one of the major challenges in modern astrophysics. Historically, the age–metallicity–dust degeneracy that plagues photometric studies (e.g. Conroy 2013) has limited

* E-mail: adamc@roe.ac.uk

our ability to extract detailed physical information from galaxy spectral energy distributions (SEDs). However, since the advent of large spectroscopic surveys it has been firmly established that local massive galaxies exhibit little ongoing star-formation activity, possessing old, passively evolving stellar populations (e.g. Kauffmann et al. 2003; Brinchmann et al. 2004; Gallazzi et al. 2005).

Even with spectroscopic data it is still challenging to determine the redshift at which the bulk of the stellar mass in local quiescent galaxies formed, owing to the extremely slow time evolution of old stellar populations. Instead, as instrumental capabilities have expanded, efforts have been made to learn about the formation histories of these objects by studying their assumed progenitors, in the form of high-redshift massive galaxies. These are now routinely studied in detail during their star-forming phases, both in the rest-frame UV optical (e.g. Kriek et al. 2015; McLure et al. 2018a; Cullen et al. 2019) and far infrared (e.g. Dunlop et al. 2017; Stach et al. 2019; Williams et al. 2019). However, once star formation ceases galaxies rapidly become fainter and redder, making the extension of spectroscopic studies to quiescent objects at the highest redshifts challenging.

Spectroscopic studies at $z \sim 1\text{--}2$ have long demonstrated that old, passively evolving galaxies already exist at this epoch (e.g. Dunlop et al. 1996; Cimatti et al. 2004, 2008; Daddi et al. 2005; Whitaker et al. 2013) and recent work has extended these studies to robust statistical samples (e.g. Belli, Newman & Ellis 2019; Carnall et al. 2019a). This finding suggests a gap in our knowledge between the most massive star-forming galaxies at $z \gtrsim 3$ and already-old quiescent galaxies at $z \lesssim 2$. In order to understand this transition in detail, it is imperative that large samples of spectroscopic data are collected for objects at $z > 2$ that show signs of being the first galaxies in the Universe to quench their star formation activity. Such studies would allow the precise characterization of the onset times and time-scales of quenching for old massive quiescent galaxies at $z \lesssim 2$, and also provide constraints on the growth of such systems through mergers post-quenching.

Recently, preliminary spectroscopic studies at $z > 3$ have begun to confirm the existence of quiescent objects at these redshifts (e.g. Glazebrook et al. 2017; Schreiber et al. 2018; Forrest et al. 2019; Valentino et al. 2019), for which strong photometric evidence has been available for several years (e.g. Nayyeri et al. 2014; Straatman et al. 2014, 2016). These studies involve attempts to observe Balmer absorption features in the near-infrared H and K bands; however, the limitations of existing instrumentation mean that such studies must be confined to the most extreme objects, and cannot reach the depths required to place strong constraints on physical properties such as stellar ages and metallicities.

As we move into the 2020s new instruments, such as NIRSpec on-board the *James Webb Space Telescope* and HARMONI at the Extremely Large Telescope, will provide the ability to observe the highest redshift quiescent galaxies with sufficient signal-to-noise ratio (SNR) to extract subtle physical properties. This will afford us a new view of the ultimate origins of local massive galaxies, as well as a deeper understanding of the quiescent galaxy population at high redshift, which is currently poorly reproduced by cosmological simulations (e.g. Davé, Rafieferantsoa & Thompson 2017; Cecchi et al. 2019).

In this work, we seek to define a robust and representative sample of massive quiescent galaxies at $z > 2$ from the extremely high-quality photometric data available in CANDELS UDS and GOODS-South (see Section 2.1). This analysis will allow us to set preliminary constraints on the number densities and physical properties of these

objects, and serve as a basis for large, detailed spectroscopic follow-up programmes with current and future instruments.

Whilst similar attempts have been made for many years (e.g. Chen & Marzke 2004; Fontana et al. 2009; Mancini et al. 2009), several significant challenges remain. First, the age–metallicity–dust degeneracy leads to significant uncertainties in the samples selected, with contemporary studies still experiencing contamination at levels of $\gtrsim 20$ per cent (Schreiber et al. 2018). A specific issue is the presence of strong emission lines in high-redshift dusty star-forming galaxies, in particular [O III] at 5007 Å, which can mimic the near-infrared colours associated with the strong Balmer breaks of recently quenched galaxies (e.g. Merlin et al. 2018).

In addition, it is also known that the specifics of the modelling assumptions made when performing SED fitting significantly impact the physical parameter values obtained (e.g. Carnall et al. 2019b; Leja et al. 2019a), with many analyses basing their conclusions on codes optimized for photometric redshift determination rather than detailed physical parameter recovery. Finally, there is no consensus on the optimal method for separating star-forming and quiescent galaxies, particularly at high redshift, with a variety of colour and specific star formation rate (sSFR) criteria in use (e.g. Fontana et al. 2009; Williams et al. 2009; Pacifici et al. 2016; Merlin et al. 2018; Girelli, Bolzonella & Cimatti 2019; Leja, Tacchella & Conroy 2019b).

In this work, we use Bayesian Analysis of Galaxies for Physical Inference and Parameter ESTimation (BAGPIPES; Carnall et al. 2018) to address these issues. BAGPIPES is one of a new generation of spectral modelling and fitting tools, which combines highly sophisticated physical models for galaxy spectra with a fully Bayesian-fitting approach. One of the major advantages of BAGPIPES is that the observed redshift can be fitted in parallel with sophisticated physical models. This, combined with the implementation of the MULTINEST nested sampling algorithm (Feroz & Hobson 2008; Feroz, Hobson & Bridges 2009; Feroz et al. 2013), allows detailed recovery of full posterior distributions for physical parameters, including clear identification of secondary low-redshift or star-forming solutions for candidate high-redshift quiescent galaxies.

We begin in Section 2 by introducing the data, fitting methodology and selection criteria we employ to obtain our sample. In Section 3, we report inferred UVJ colours, formation redshifts, and number densities. In Section 4, we present an analysis designed to understand when the oldest galaxies in our sample quenched their star formation. In Section 5, we focus on our individual robustly identified objects at $z > 3$, reporting probable spectroscopic redshifts for three objects. We present our conclusions in Section 6. All magnitudes are quoted in the AB system. For cosmological calculations, we adopt $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$, and $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$. All times, t , are measured forwards from the beginning of the Universe. We assume a Kroupa (2001) initial mass function.

2 METHOD

2.1 Photometric catalogues

In this work, we analyse data from CANDELS UDS and GOODS-South, two of the best-studied extragalactic legacy fields. Both have deep near-ultraviolet to near-infrared *Hubble Space Telescope* (*HST*) data from CANDELS (Grogin et al. 2011; Koekemoer et al. 2011) and other programmes, as well as a wide range of ground-based and *Spitzer*-IRAC multiwavelength ancillary data. Full details

of the catalogues used can be found in Galametz et al. (2013) and Guo et al. (2013) for UDS and GOODS-S, respectively.

These two fields in particular were chosen for this work due to the availability of ultra-deep K -band imaging from the Hawk-I UDS and GOODS Survey (HUGS; Fontana et al. 2014). At $z \gtrsim 3$, the Balmer/4000 Å break falls between the H and K bands, meaning the extremely deep imaging in these bands provided by CANDELS and HUGS, respectively, is critical for the robust identification of older stellar populations at these redshifts. For this reason, we supplement the Guo et al. (2013) GOODS-S catalogue with additional, deeper K -band data from the final HUGS data release.

Before fitting we select only objects with $H_{160} < 27$, leaving 61 837 objects. We additionally remove all objects with missing observations in more than two bands (~ 3 per cent in UDS; ~ 5 per cent in GOODS-S). When fitting we impose a maximum SNR (or error floor; e.g. Muzzin et al. 2013) to account for the error budget becoming dominated by systematic rather than random errors at high SNRs. We use a maximum SNR of 20 for all photometric bands except IRAC channels 1 and 2 for which we use a maximum SNR of 10 and IRAC channels 3 and 4 for which we use a maximum SNR of 5. The median SNR of the catalogue is 12.0 in the H band, 5.0 in IRAC channel 1, and 1.2 in IRAC channel 3.

2.2 Fitting with BAGPIPES

Our main objective when designing the model to be fitted was to allow sufficient flexibility to permit all potential solutions over a wide range of observed redshifts and sSFRs. As such, we use a nine-parameter model including flexible prescriptions for galaxy star formation histories (SFHs) and dust attenuation curve shapes. The parameters of our model are detailed below and summarized in Table 1.

We use the default BAGPIPES stellar population models, which are the 2016 updated version of the Bruzual & Charlot (2003) models¹ using the MILES stellar spectral library (Falcón-Barroso et al. 2011) and updated stellar evolutionary tracks of Bressan et al. (2012) and Marigo et al. (2013).

We parametrize the SFHs of our galaxies using the double power-law form of Carnall et al. (2018), which has been shown to reproduce well the SFHs of massive quiescent galaxies from the MUFASA simulation (Davé, Thompson & Hopkins 2016). We assume a single metallicity for all stars in our galaxies with scaled Solar abundances. This metallicity is allowed to vary with a uniform prior in logarithmic space over the range from $-1 < \log_{10}(Z_*/Z_{\odot}) < 0.4$, where Z_{\odot} is defined as 0.02. The total stellar mass formed by each galaxy up to the observed redshift is fitted with a uniform prior in logarithmic space from $0 < \log_{10}(M_*/M_{\odot}) < 13$.

We model dust attenuation in our galaxies using the flexible model of Salim et al. (2018), which parametrizes the dust curve shape in terms of a power-law deviation, δ , from the Calzetti et al. (2000) law (see also Noll et al. 2009). As we are primarily interested in massive galaxies our prior is centred on the Calzetti et al. (2000) curve ($\delta = 0$), which has been found to be a robust average attenuation curve shape for massive galaxies (e.g. Cullen et al. 2017, 2018; McLure et al. 2018a). We set a Gaussian prior on δ with a standard deviation, $\sigma = 0.1$, and set a uniform prior on the 2175 Å bump strength, B , from 0 to 5, where the Milky Way law (e.g. Cardelli, Clayton & Mathis 1989) has $B = 3$. We assume a uniform prior on A_V from 0 to 8 mag. Whilst quiescent galaxies are

the focus of our analysis, this wide range of A_V values is critical for correct identification of intermediate redshift ($z \sim 1-3$) dusty star-forming galaxies, which, if misidentified, are known to contaminate high-redshift quiescent galaxy samples (e.g. Schreiber et al. 2018). We further assume light from stars younger than 10 Myr and resulting nebular line and continuum emission experiences double the attenuation experienced by the rest of the stellar population.

Our model includes nebular line and continuum emission by post-processing of the stellar models using the CLOUDY photoionization code (Ferland et al. 2017), following a method based on that of Byler et al. (2017), as described in section 3.1.3 of Carnall et al. (2018). The nebular metallicity is assumed to follow the stellar metallicity, and a fixed ionization parameter of $\log_{10}(U) = -3$ is also assumed.

We additionally include dust emission assuming energy balance between light attenuated and light re-radiated and using the dust emission models of Draine & Li (2007). We assume values of 2 for Q_{pah} , the percentage of dust mass in polycyclic aromatic hydrocarbons, 1 for U_{min} , the minimum starlight intensity to which the dust is exposed, and 0.01 for γ_e , the fraction of the incident starlight at U_{min} . BAGPIPES also includes attenuation due to the intergalactic medium using the model of Inoue et al. (2014). We fit the observed redshift of each galaxy assuming a uniform prior.

2.3 Selection of high-redshift quiescent galaxies

Having fitted the catalogues of Section 2.1 using the model of Section 2.2, we calculate posterior percentiles for the physical parameters of each object, as well as minimum reduced chi-squared values, χ_v^2 . We then exclude objects with $\chi_v^2 > 5.3$ in UDS and $\chi_v^2 > 6.8$ in GOODS-S. These cuts correspond to 5σ upper limits on the expected χ_v^2 values under the assumption that the fitted model correctly describes the data (the UDS catalogue has 19 bands, whereas the GOODS-S catalogue has 16, giving 9 and 6 degrees of freedom when fitting with our 9 parameter model). These cuts remove the worst-fitted ~ 13 per cent of objects. The majority of these are contaminated due to close proximity with bright sources, whilst a minority have SEDs consistent with stars. We did not apply a point-source selection due to concerns about wrongly excluding high-redshift quiescent galaxies, which are known to be extremely compact (e.g. McLure et al. 2013). We then select objects with posterior median redshifts, $z > 2$, and stellar masses, $M_* > 10^{10} M_{\odot}$.

The best method of separating star forming and quiescent galaxies at any given epoch is still a subject of much debate. The two most common approaches are selection by rest-frame UVJ colours (e.g. Williams et al. 2009) and selection by sSFR (e.g. Fontana et al. 2009; Gallazzi et al. 2014; Pacifici et al. 2016; Merlin et al. 2018), with the selection criteria usually a function of observed redshift in both cases. In this work, we use the most commonly applied sSFR selection criterion, requiring that galaxies have posterior median sSFR $< 0.2/t_H$, where t_H is the age of the Universe at the posterior median observed redshift. In Carnall et al. (2018, 2019a), we demonstrate good agreement between samples selected by this criterion and the Williams et al. (2009) $z = 0$ UVJ criterion of $U - V > 0.88(V - J) + 0.69$. This will be further explored in Section 3.1. To summarize, we require

- (i) $\chi_v^2 < 6.8$ for GOODS-S or $\chi_v^2 < 5.3$ for UDS
- (ii) $z_{50} > 2$
- (iii) $M_{*50} > 10^{10} M_{\odot}$
- (iv) $\text{sSFR}_{50} < 0.2/t_H$.

At this point, we visually checked the imaging data and posterior distributions for all remaining objects. We identified 21 objects as

¹https://www.bruzual.org/~gbruzual/bc03/Updated_version_2016

Table 1. Parameters and priors for the model we fit to our data (see Section 2.2). For Gaussian priors, μ is the mean and σ the standard deviation of the prior distribution. The upper limit on the τ parameter, t_{obs} , is the age of the Universe at redshift z . Logarithmic priors are uniform in log base ten of the parameter. We use a double power-law SFH model and the dust attenuation law of Salim, Boquien & Lee (2018).

| Parameter | Symbol/unit | Range | Prior | Hyperparameters |
|---|-------------------|--------------------------|-------------|--------------------------|
| Redshift | z | (0, 10) | Uniform | |
| Stellar mass formed | M_*/M_\odot | (1, 10^{13}) | Logarithmic | |
| Stellar metallicity | Z/Z_\odot | (0.1, 2.5) | Logarithmic | |
| Double power-law falling slope | α | (0.01, 1000) | Logarithmic | |
| Double power-law rising slope | β | (0.01, 1000) | Logarithmic | |
| Double power-law turnover time | τ/Gyr | (0.1, t_{obs}) | Uniform | |
| V-band attenuation | A_V/mag | (0, 8) | Uniform | |
| Deviation from Calzetti et al. (2000) slope | δ | (−0.3, 0.3) | Gaussian | $\mu = 0$ $\sigma = 0.1$ |
| Strength of 2175 Å bump | B | (0, 5) | Uniform | |

having very high posterior median redshifts ($5.4 < z < 8.3$). All of these objects have secondary lower redshift dusty star-forming solutions, and several are required to be extremely massive at their posterior median redshift ($M_* \gtrsim 10^{11.5} M_\odot$). We therefore remove these 21 objects as probable low-redshift dusty interlopers. We also identify 10 poorly fitted objects that are not removed by our χ^2_v criterion (e.g. due to individual strongly discrepant photometry points), which are also rejected. The resulting parent sample of 151 objects is provided in Table 2.

2.4 Selection of a robust sub-sample

It should be stressed that inclusion in our parent sample, shown in Table 2, does not constitute strong evidence that these objects are $z > 2$ quiescent galaxies. Whilst the criteria laid out in Section 2.3 require that >50 per cent of the posterior is consistent with the object being a massive quiescent galaxy at $z > 2$, it is still possible that objects have secondary low-redshift or star-forming solutions. We therefore define a further ‘robust’ sub-sample, for which we can confidently exclude such secondary solutions.

Objects in our robust sub-sample are required to have >97.5 per cent of their sSFR posterior below $\text{sSFR} = 0.2/t_H$ and <2.5 per cent of their redshift posterior below $z = 2$. These criteria correspond to 2σ upper and lower limits, respectively, for a Gaussian posterior distribution. To summarize, in addition to the criteria laid down in Section 2.3, objects in our robust sub-sample must have

- (i) $z_{2.5} > 2$
- (ii) $\text{sSFR}_{97.5} < 0.2/t_H$.

The 61 resulting objects are flagged as ‘robust’ in Table 2.

Fig. 1 shows two example objects from our parent sample that do (top) and do not (bottom) meet our robust selection criteria. Both are at $z \sim 2.3$ – 2.4 ; however, the sSFR posterior in the case of the upper object can be seen to be confined more strictly to values below our chosen threshold, which is marked with a dashed black vertical line. The sSFR posterior is better constrained in the case of the upper object first because it is brighter and hence detected at higher SNR, and secondly because the lower object is dustier, permitting greater uncertainty in the level of ongoing star formation. In addition, the deeper H , K , and IRAC data in GOODS-S allows us to place stronger constraints on physical parameters, weighting our robust sub-sample slightly in favour of objects in the GOODS-S field.

3 RESULTS

3.1 Quiescent selection at $z > 2$: UVJ versus sSFR

Our parent sample is shown on the UVJ diagram in Fig. 2, split into three redshift bins. Star-forming galaxies that obey the criteria of Section 2.3 excepting the sSFR threshold are also shown. The standard $z = 0$ UVJ selection box of Williams et al. (2009) is shown with black solid lines. The dashed black lines carry on the diagonal beyond the commonly applied vertical and horizontal colour cuts. Following Williams et al. (2009), we have adopted the U - and V -band filter curves of Bessell (1990) and the J -band filter curve of Tokunaga, Simons & Vacca (2002). The median uncertainty on the derived rest-frame colours shown in Fig. 2 is 0.05 mag in $U - V$ and 0.09 mag in $V - J$.

As in Carnall et al. (2018, 2019a), good agreement can be seen between the Williams et al. (2009) $z = 0$ UVJ selection box and the evolving sSFR criterion we have applied. 121 of the 151 galaxies in our (sSFR-selected) parent sample fall within the solid boxes in Fig. 2, along with 1 additional object we identify as star forming. This suggests the solid box identifies an 80 per cent complete sample of quiescent galaxies at $2 < z < 5$, with a very low contamination rate.

By contrast, ~ 90 per cent agreement was found between these selection criteria at $0.25 < z < 2$ by Carnall et al. (2018). This difference is due to an increased fraction of objects at $z > 2$ above the diagonal but below the horizontal cut ($U - V > 1.3$) when compared to lower redshift samples (10 of 151 objects; 7 per cent), in agreement with the results of Merlin et al. (2018). This region is strongly associated with post-starburst galaxies at lower redshifts (e.g. Wild et al. 2014). When the horizontal colour cut is removed we recover a similar level of agreement to that observed at $z < 2$ (131 of 151 objects; 87 per cent), whilst admitting three additional objects we identify as star-forming contaminants.

Extending the solid UVJ selection box in Fig. 2 by relaxing the diagonal colour cut to the $z > 1$ criterion proposed by Williams et al. (2009) of $U - V > 0.88(V - J) + 0.49$ admits an additional 15 objects in our parent sample (10 per cent), at the cost of including an additional 35 galaxies that do not meet our sSFR criterion. This implied contamination rate of 20 per cent is in agreement with the results of Schreiber et al. (2018) using the same selection criteria.

In the left-hand panel of Fig. 3, we show only our robust sub-sample on the UVJ diagram. It can be seen that 60 of the 61 objects fall above the solid/dashed diagonal selection criterion. All but two of these fall within the solid box, with the remaining objects below the horizontal selection criterion.

Table 2. Coordinates, magnitudes, and BAGPIPES physical parameters (16–84th percentiles) for our parent $z > 2$ massive quiescent galaxy sample, selected in Section 2.3. The full table, including additional columns, is available as supplementary online material.

| ID | RA | Dec. | $F160W$ | K_s | Redshift | $t_{\text{form}}/\text{Gyr}$ | $\log_{10}(M_*/M_{\odot})$ |
|--------------|-----------|------------|---------|-------|------------------------|------------------------------|----------------------------|
| GOODSS-00402 | 53.125301 | -27.934856 | 23.4 | 23.2 | $2.41^{+0.06}_{-0.05}$ | $2.3^{+0.1}_{-0.1}$ | $10.38^{+0.05}_{-0.04}$ |
| GOODSS-00954 | 53.161680 | -27.918738 | 23.3 | 22.9 | $2.54^{+0.05}_{-0.04}$ | $2.0^{+0.1}_{-0.4}$ | $10.42^{+0.05}_{-0.03}$ |
| GOODSS-01000 | 53.136832 | -27.917348 | 24.5 | 23.6 | $2.30^{+0.15}_{-0.15}$ | $0.7^{+0.5}_{-0.4}$ | $10.61^{+0.07}_{-0.06}$ |
| GOODSS-01086 | 53.137180 | -27.915837 | 23.7 | 23.3 | $2.70^{+0.05}_{-0.05}$ | $2.0^{+0.1}_{-0.1}$ | $10.26^{+0.05}_{-0.05}$ |
| GOODSS-01119 | 53.169052 | -27.915846 | 22.0 | 21.6 | $2.12^{+0.05}_{-0.05}$ | $2.4^{+0.2}_{-0.7}$ | $11.00^{+0.06}_{-0.04}$ |

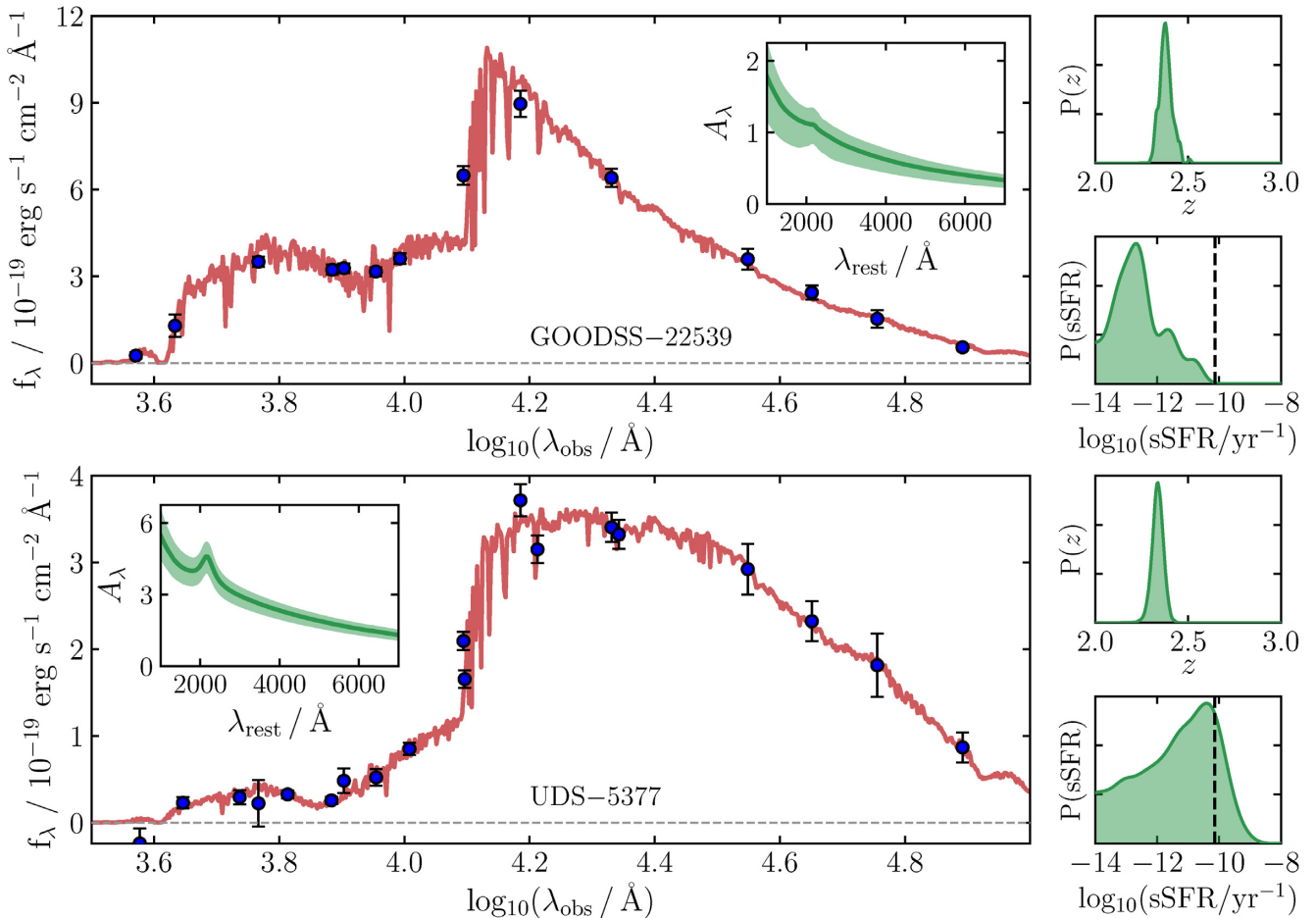


Figure 1. Example objects from our parent sample that do (top) and do not (bottom) meet our robust criteria (see Section 2.4). Data are shown in blue, with the posterior median model shown in red. Posterior distributions for redshift, sSFR, and dust curve shape are shown in green in the side and inset panels. The dashed vertical lines show our quiescent selection criterion. For the top galaxy, >97.5 per cent of the sSFR posterior falls below the limit, meaning this object is part of our robust sub-sample. For the bottom object >50 per cent of the sSFR posterior falls below the limit; however, >2.5 per cent falls above. The sSFR of the bottom object is less well constrained as it is fainter and dustier. The sSFR posteriors have tails out to lower values driven by our SFH priors, which are not shown.

3.2 Redshift evolution of quiescent galaxy colours

As well as playing a role in the separation of the star-forming and quiescent populations, rest-frame UVJ colours are increasingly becoming used as a general tool for understanding galaxy evolution and quenching (e.g. Fang et al. 2018; Morishita et al. 2019; Wu et al. 2019). In particular, Belli et al. (2019, hereafter B19) find evidence for a strong age trend within the quiescent population across the UVJ selection box at $1.5 < z < 2.5$, with younger galaxies found towards the bottom-left of the box, and older galaxies found in the

upper-right region (see also Whitaker et al. 2013). This implies that (assuming no significant redshift evolution in dust attenuation) the upper right region of the UVJ box should become depopulated at higher redshifts, as not enough time has elapsed since the big bang for galaxies to grow old enough to have these colours.

The left-hand panel of Fig. 2 demonstrates (in agreement with e.g. Straatman et al. 2016) that galaxies fully populate the diagonal edge of the UVJ selection box by $2 < z < 3$. However, in the central-hand and right-hand panels, galaxies appear to be confined

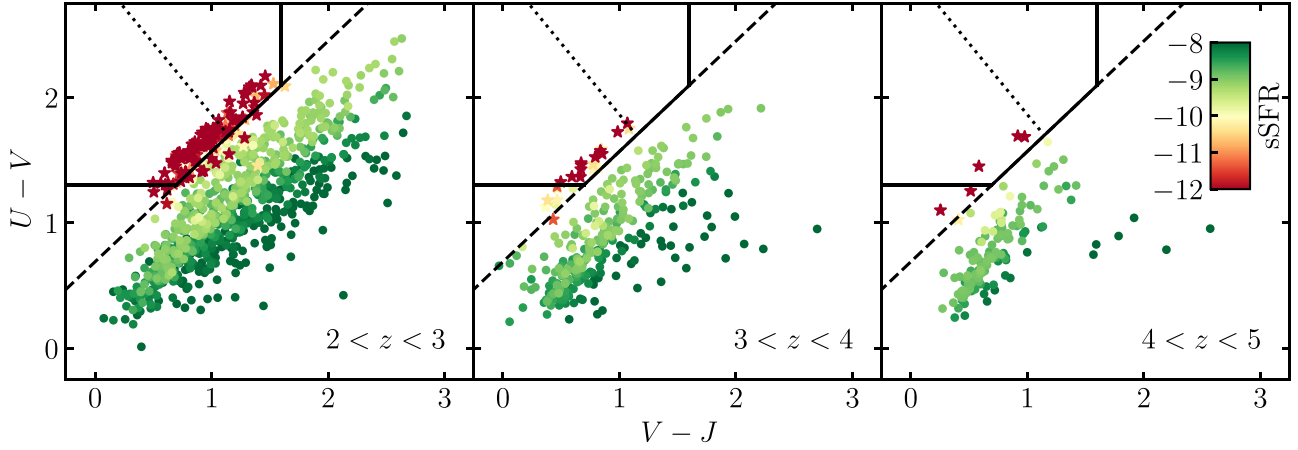


Figure 2. *UVJ* diagrams showing the distribution of massive galaxies ($M_* > 10^{10} M_\odot$) at $2 < z < 5$, coloured by sSFR. Our parent quiescent sample is shown with stars, whilst star-forming galaxies that meet the criteria of Section 2.3 excepting the sSFR cut are shown with circles. The standard $z = 0$ *UVJ* selection criteria of Williams et al. (2009) are shown with solid black lines, the dashed black line shows the continuation of the diagonal beyond the vertical and horizontal cuts. The dotted line (equation 1) bisects the solid diagonal line. At $2 < z < 3$, 25 per cent of our parent sample falls above the bisector, whereas at $z > 3$ only 1 of 41 objects falls above this line.

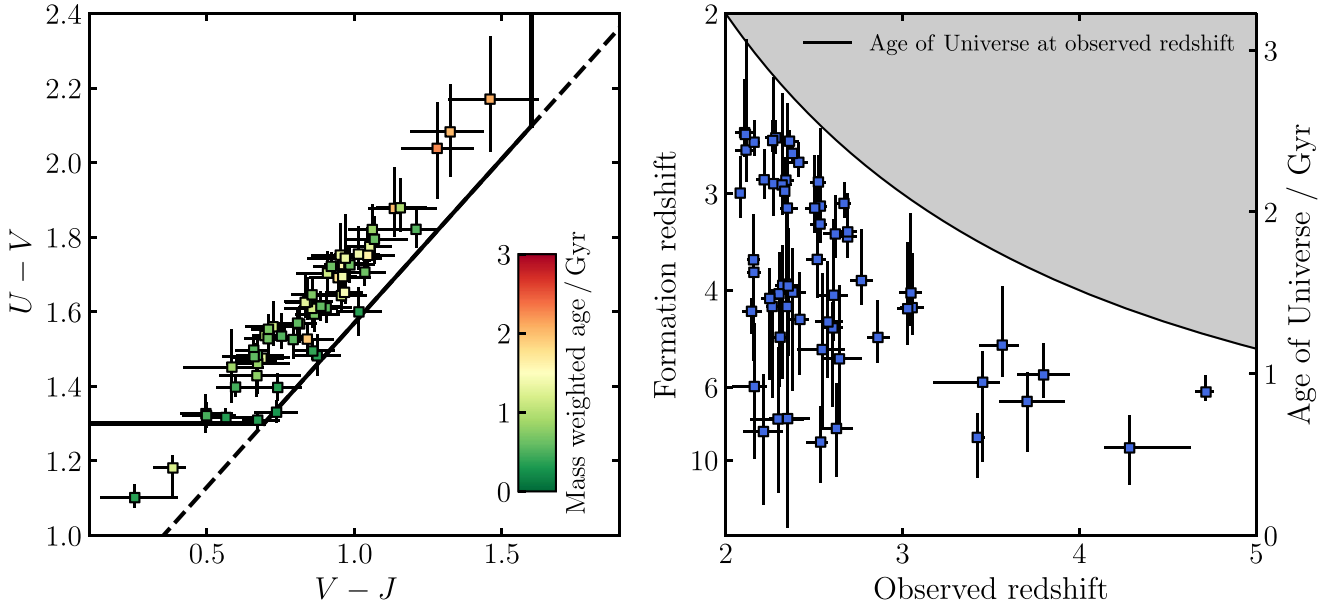


Figure 3. Ages and formation redshifts for galaxies in our robust sub-sample. In the left-hand panel, galaxies are shown on the *UVJ* diagram coloured by mass-weighted age. The *UVJ* selection box is as shown in Fig. 2. In the right-hand panel, formation redshifts are shown as a function of observed redshift. A range of formation redshifts are observed, from $z \gtrsim 6$ to immediately preceding the observed redshift.

to the lower half of the diagonal edge. In order to test whether this difference is significant, we construct the perpendicular bisector for the diagonal edge of the solid *UVJ* selection box shown in Fig. 2, given by

$$U - V = -\frac{1}{0.88}(V - J) + 3, \quad (1)$$

which is shown with a black dotted line in Fig. 2.

In the left-hand panel of Fig. 2, at $2 < z < 3$, a total of 27 objects out of 110 in our parent quiescent sample (25 per cent) fall above this bisecting line. If the distribution of colours were the same above $z = 3$, we would expect to observe 10 ± 3 objects above this line out of a total of 41. However, as can be seen from Fig. 2, only 1

out of 41 objects in our parent sample at $z > 3$ falls above this bisecting line. We therefore find strong evidence that the quiescent population shifts towards bluer colours at $z > 3$, consistent with the scheme put forward by B19.

3.3 Stellar ages and formation redshifts

In this section, we report the ages we infer based on our photometric analysis. It should be noted that, as discussed in Section 1, age-dating stellar populations based on photometric data is challenging, owing to the age–metallicity–dust degeneracy. The flexible model described in Section 2.2, combined with careful exploration of parameter space using MULTINEST, allows us to fully characterize the

Table 3. Number counts and number densities for objects in our sample ($M_* > 10^{10} M_\odot$) from $2 < z < 4$. The r subscripts denote the robust sub-sample selected in Section 2.4, as opposed to the parent sample defined in Section 2.3.

| Redshift | N | N_r | n/Mpc^{-3} | n_r/Mpc^{-3} |
|-----------------|-----|-------|------------------------------|------------------------------|
| $2.0 < z < 2.5$ | 85 | 33 | $1.4 \pm 0.2 \times 10^{-4}$ | $5.3 \pm 0.9 \times 10^{-5}$ |
| $2.5 < z < 3.0$ | 38 | 18 | $6.2 \pm 1.0 \times 10^{-5}$ | $2.9 \pm 0.7 \times 10^{-5}$ |
| $3.0 < z < 4.0$ | 22 | 8 | $1.9 \pm 0.4 \times 10^{-5}$ | $6.8 \pm 2.4 \times 10^{-6}$ |

uncertainties on our inferred ages by marginalizing over solutions across a wide range of metallicities and dust contents (neither of these properties are strongly constrained by our photometric data). The average uncertainty is ~ 0.5 Gyr, with the largest uncertainties for the oldest objects.

Age determinations for quiescent galaxies using our double power-law SFH model were extensively validated in Carnall et al. (2018); however, no such test has been performed for star-forming galaxies. We therefore limit the discussion in this section to our robust sub-sample. We calculate formation times, t_{form} , as the average time at which the stars in each galaxy formed,

$$t_{\text{form}} = \frac{\int_0^{t_{\text{obs}}} t \text{SFR}(t) dt}{\int_0^{t_{\text{obs}}} \text{SFR}(t) dt}. \quad (2)$$

This corresponds to the mass-weighted age of the galaxy. We then calculate the redshift corresponding to t_{form} , which we call the formation redshift. In the left-hand panel of Fig. 3, we show the mass-weighted ages of galaxies in our robust sub-sample as a function of position on the UVJ diagram. In accordance with the results of Section 3.1, it can be seen that we recover a broadly similar age trend across the UVJ box to B19, albeit with large scatter due to the significant uncertainties on our photometric age determinations.

The right-hand panel of Fig. 3 shows the formation redshifts we infer as a function of observed redshift. It can be seen that our galaxies exhibit a diverse range of SFHs, with formation redshifts ranging from $z \sim 6$ –10 to immediately preceding the observed redshift. At $z < 3$, the oldest objects are already $\gtrsim 2$ Gyr old, making it challenging to set upper limits on their formation redshifts. At $z > 3$, objects are younger and their ages are hence better constrained, with posterior median formation redshifts ranging from $5 < z < 9$. We will further address the question of when the oldest galaxies quenched their star formation in Section 4.

3.4 Number density

The number density of quiescent galaxies at high redshift provides an important constraint on quenching models in simulations of galaxy formation (e.g. Davé et al. 2017; Cecchi et al. 2019; Merlin et al. 2019). In this section, we report number densities for our parent sample, as well as our robust sub-sample, which serve as reliable lower limits. We only consider objects at $z < 4$, for which our stellar mass limit of $10^{10} M_\odot$ constitutes a conservative mass-completeness limit in both fields (e.g. Mortlock et al. 2015; McLeod et al. in preparation). The number densities we calculate are given in Table 3, along with uncertainties estimated as the Poisson uncertainty on the number of objects.

Schreiber et al. (2018) report a number density of $2.0 \pm 0.3 \times 10^{-5} \text{Mpc}^{-3}$ for a spectroscopic sample composed of quiescent galaxies with $K_s < 24.5$ at $3 < z < 4$. In our parent sample, 20 galaxies meet these criteria, and we hence derive a number density of $1.7 \pm 0.4 \times 10^{-5} \text{Mpc}^{-3}$, in good agreement

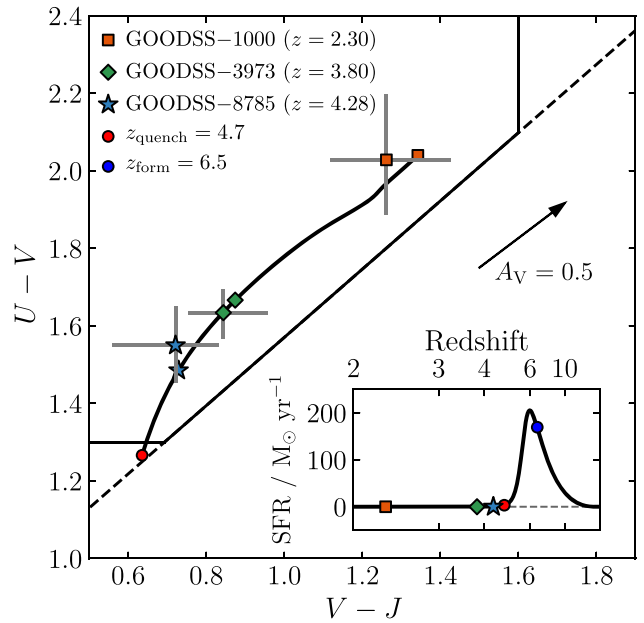


Figure 4. A model for the time-evolution of the UVJ colours of the reddest galaxies in our robust sub-sample in integer redshift bins. These objects, corrected to a common $A_V = 0.58$, are shown as symbols with grey error bars. The black curved line shows a model for the time-evolution of the UVJ colours of these objects. The corresponding SFH is shown in the inset panel. Symbols superimposed on the line show predicted UVJ colours at the appropriate redshifts. The UVJ selection box is as shown in Fig. 2. The Calzetti et al. (2000) dust reddening vector is also shown.

with their result. This suggests that our photometrically selected sample is not strongly contaminated.

4 WHEN DID THE FIRST GALAXIES QUENCH?

As discussed in Section 3.3, the individual ages we derive from our photometric data are highly uncertain. In this section, we therefore attempt to use a subset of the most extreme objects in our robust sub-sample to jointly constrain the formation and quenching times of the first galaxies in the Universe to cease star formation.

As demonstrated in Section 3.2, the UVJ colours of the reddest quiescent galaxies evolve with redshift across the range from $2 < z < 5$, becoming bluer at higher redshift as the cosmic time available for them to passively age reduces. We hence select the reddest objects in our robust sub-sample (by $U - V$ colour) in each of the redshift bins shown in Fig. 2 and attempt to construct a model SFH that traces these objects back to a common redshift of formation.

The three objects selected are shown on the UVJ diagram in Fig. 4. As these objects all have different levels of dust attenuation, their positions have been corrected to a common A_V of 0.58, which is the average value for our robust sub-sample, using the Calzetti et al. (2000) law. This average attenuation is higher than is typically found at low redshift, in agreement with other recent analyses (e.g. Gobat et al. 2018). The dust attenuation curve slope has little effect on the magnitude of the dust vector shown in Fig. 4, instead resulting in a rotation, which moves objects perpendicular to the diagonal edge of the UVJ box. As we are mainly interested in the evolution of galaxy colours along the age sequence parallel to the diagonal edge of the UVJ box, we do not consider variations in dust attenuation curve shape in this section.

For our model, we assume the same constant A_V of 0.58 and further assume fixed Solar metallicity. The model forms a total stellar mass of $10^{11} M_\odot$, although this has no effect on the resulting *UVJ* colours. We then vary the SFH shape to attempt to match the *UVJ* colours of the objects shown in Fig. 4 at the appropriate redshifts.

Our best model is also shown in Fig. 4. The SFH is shown in the inset panel, whilst the black curved line shows the time-evolution of the associated *UVJ* colours. The symbols plotted along this line correspond to predictions for the *UVJ* colours of the three objects shown at their observed redshifts. The model SFH has a formation redshift (see Section 3.3) of $z = 6.5$. The model meets our quiescent sSFR selection criterion (see Section 2.3) at $z = 4.7$. This point is marked with a red circle in Fig. 4. The formation redshifts found through our SED-fitting analysis for the objects shown are also consistent with this model.

From this analysis, we conclude that the oldest objects in our parent sample are consistent with having formed the bulk of their stellar mass at $z \sim 6-7$ and quenched at $z \sim 5$. Earlier formation and quenching is not required to explain our data; however neither is this possibility ruled out. As a final comment, it should be noted that the CANDELS data covers relatively small areas, and it is a clear possibility that more extreme objects may be found in larger area surveys.

5 MASSIVE QUIESCENT GALAXIES AT $Z > 3$

Our robust sub-sample contains 10 objects at $z > 3$, including 2 at $z > 4$, both of which are at higher redshift than the most distant spectroscopically confirmed quiescent galaxies. As demonstrated in Section 3.3, the younger ages of quiescent galaxies at $z > 3$ make them ideal for placing strong constraints on quenching physics at the highest redshifts. Given this key role in our understanding of galaxy formation, spectroscopic follow-up of such objects is clearly of significant interest.

In this section, we focus on our robust sub-sample at $z > 3$. We first compare our results with samples from the literature and then discuss currently available spectroscopic data. The faintness of these objects means that these spectroscopic data are somewhat sparse and of low SNR. However with the next generation of telescopes, we can aspire to place strong constraints on the physical properties of such galaxies through high-SNR spectroscopic observations.

Fitted SEDs for our 10 robust $z > 3$ objects are shown in Fig. 5. *HST*, *K_s* band, and IRAC imaging data for each object are shown in Fig. 6. These objects can be seen to be generally compact and isolated in the *HST* imaging.

5.1 Comparisons with recent studies

Comparing with other recent work, the $z > 3$ robust sub-sample shown in Fig. 5 contains 6 of the 10 $z > 3$ quiescent galaxy candidates reported by Merlin et al. (2018) in GOODS-S, with a further 3 of their objects included in our parent sample (which contains 17 objects at $z > 3$ in GOODS-S). The $z > 3$ robust object we identify in UDS is not part of the sample reported by Merlin et al. (2019); however, 7 of their 16 UDS candidates feature in our parent sample (which contains 11 objects at $z > 3$ in UDS).

Santini et al. (2019) report an analysis of archival ALMA data for the sample of Merlin et al. (2018), aiming to confirm the passive nature of these objects by placing upper limits on their SFRs. They report non-detections for 3 of our robust $z > 3$ objects (GOODSS-2782, 8785, and 9209) and $<3\sigma$ detections for a further

two (GOODSS-3973 and 18180). The implied sSFRs for the two ALMA-detected objects are above our selection threshold; however, they are consistent to within 2σ and 1σ , respectively.

The single UDS $z > 3$ object in our robust sub-sample was not observed by Schreiber et al. (2018), who studied $3 < z < 4$ massive quiescent galaxies in the CANDELS UDS, COSMOS, and EGS fields with Keck-MOSFIRE. However, three $z > 3$ objects from our parent sample were observed. The authors report a spectroscopic redshift of $z = 3.543$ for UDS-8682, as well as reporting sSFRs consistent with our quiescent selection criterion for all three objects.

5.2 Spectroscopic observations

We finally search the many available spectroscopic surveys within UDS and GOODS-S for observations of our robust $z > 3$ objects. We find two spectra: one of which is from VANDELS (McLure et al. 2018b; Pentericci et al. 2018) and one of which is from VUDS (Le Fèvre et al. 2015; Tasca et al. 2017). These objects will be discussed in the following sections. We will also discuss a further $z > 3$ object from our parent sample that was observed as part of VANDELS.

5.2.1 GOODSS-2782

VANDELS is a uniquely deep ESO Public spectroscopic survey targeting 2106 high-redshift galaxies with 20, 40, and 80 h integrations using the VIMOS instrument, providing wavelength coverage from ~ 5000 to 10000 \AA at spectral resolution, $R \sim 600$. Whilst VANDELS nominally targets star-forming galaxies at $z > 2.4$, the chosen criterion of $\text{sSFR} > 0.1 \text{ Gyr}^{-1}$ permits some overlap with our sample at $z > 3$.

GOODSS-2782 is a member of our robust sub-sample, and was targeted by VANDELS for 80 h of integration. The 1D and 2D spectra for this object are shown in the top panel of Fig. 7. A marginal detection of an emission line can be seen at 5400 \AA . For this object we recover a photometric redshift of $z = 3.45^{+0.28}_{-0.10}$. This, combined with a CANDELS photometric redshift of $z_{\text{CANDELS}} = 3.47$, strongly implies this detection is weak Ly α emission at $z = 3.440$.

Under this assumption we measure the Ly α flux by first summing the pixels in the 1D spectrum within $\pm 30 \text{ \AA}$ of 5400 \AA . We then correct this for continuum emission by subtracting the summed flux of pixels in the 60 \AA region directly redwards ($5430-5490 \text{ \AA}$). The raw flux within the Ly α region is $2.9 \pm 0.6 \times 10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2}$, whereas the corrected line flux is $2.1 \pm 0.9 \times 10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2}$.

It is possible to turn this line flux into an approximation of the SFR for this galaxy, assuming an intrinsic Ly α /H α ratio of 8.7 (from case B recombination theory; see Osterbrock 1989) and the Kennicutt & Evans (2012) relationship between H α luminosity and SFR. This calculation produces an SFR of $0.14 \pm 0.07/f_{\text{esc, Ly } \alpha} M_\odot \text{ yr}^{-1}$, where $f_{\text{esc, Ly } \alpha}$ is the escape fraction for Ly α . Dividing by the stellar mass posterior yields $\log_{10}(\text{sSFR yr}^{-1}) = -11.5^{+0.3}_{-0.2} - \log_{10}(f_{\text{esc, Ly } \alpha})$. The sSFR threshold for inclusion in our quiescent sample at $z = 3.44$ is $\log_{10}(\text{sSFR yr}^{-1}) < -9.95$, meaning this object meets our quiescent selection criterion provided that $f_{\text{esc, Ly } \alpha} > 0.03$. Contemporary studies find average Ly α escape fractions of $\sim 5-10$ per cent in this redshift range (e.g. Sobral et al. 2017, 2018), implying that the strength of Ly α in this spectrum is consistent with this galaxy meeting our sSFR selection criterion.

Santini et al. (2019) report that this galaxy is undetected in ALMA band 7 observations, with an observed flux of $0.04 \pm 0.11 \text{ mJy beam}^{-1}$, implying a 1σ upper limiting SFR of $22.5 M_\odot \text{ yr}^{-1}$. To

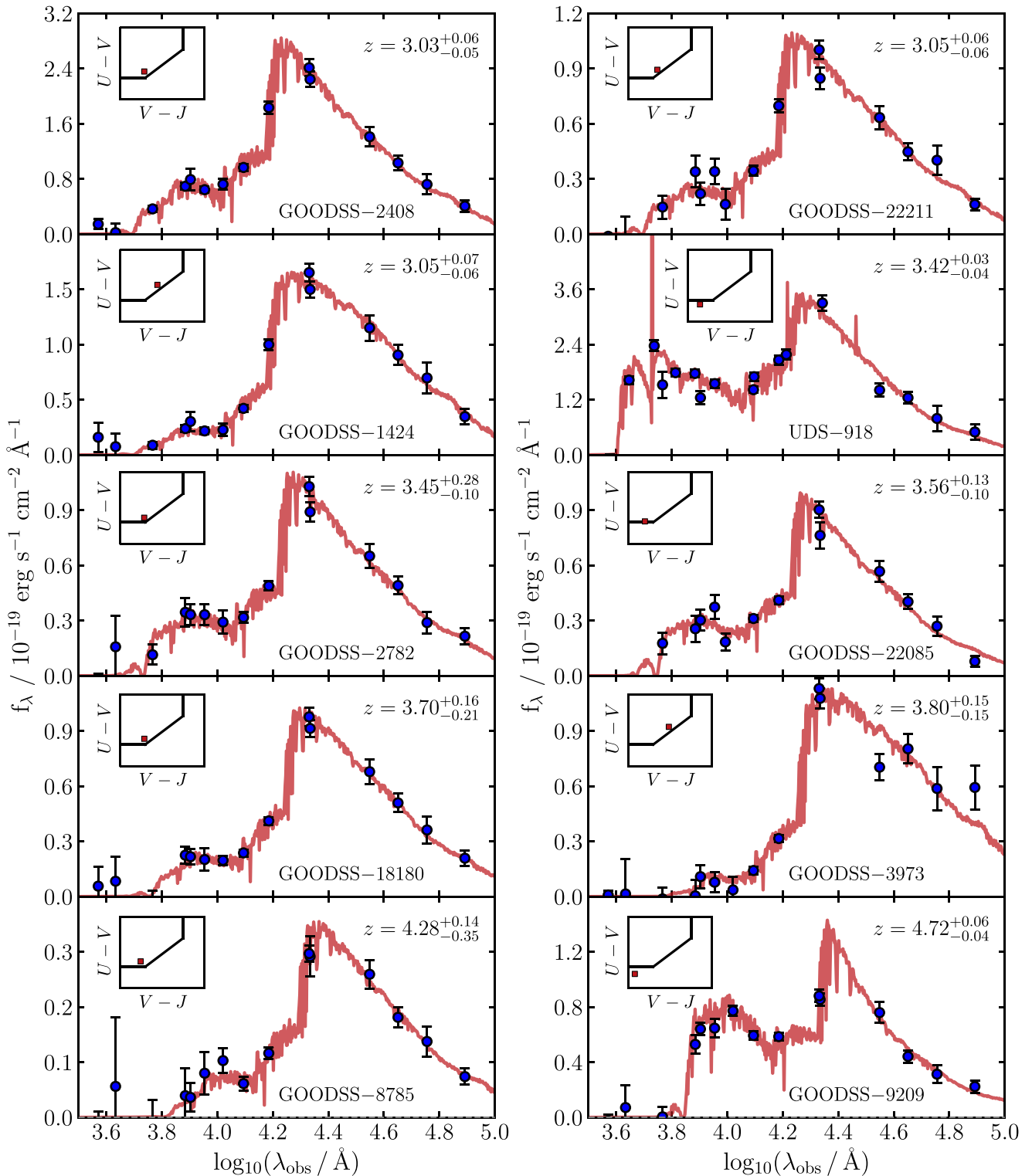


Figure 5. Fitted SEDs for the 10 objects in our robust sub-sample at $z > 3$ in ascending order of redshift. Photometric data are shown in blue, whereas our posterior median fitted model is shown in red at our posterior median fitted redshift. The inset panels show the positions of each object on the UVJ diagram, with the solid black lines the same as those shown in Fig. 2. The object GOODSS-2782 was observed for 80 h as part of VANDELS. The resulting spectrum is shown in the top panel of Fig. 7 and is discussed in Section 5.2.1. The highest redshift object, GOODSS-9209, was observed for 14 h as part of VUDS. The resulting spectrum is shown in Fig. 8 and is discussed in Section 5.2.3.

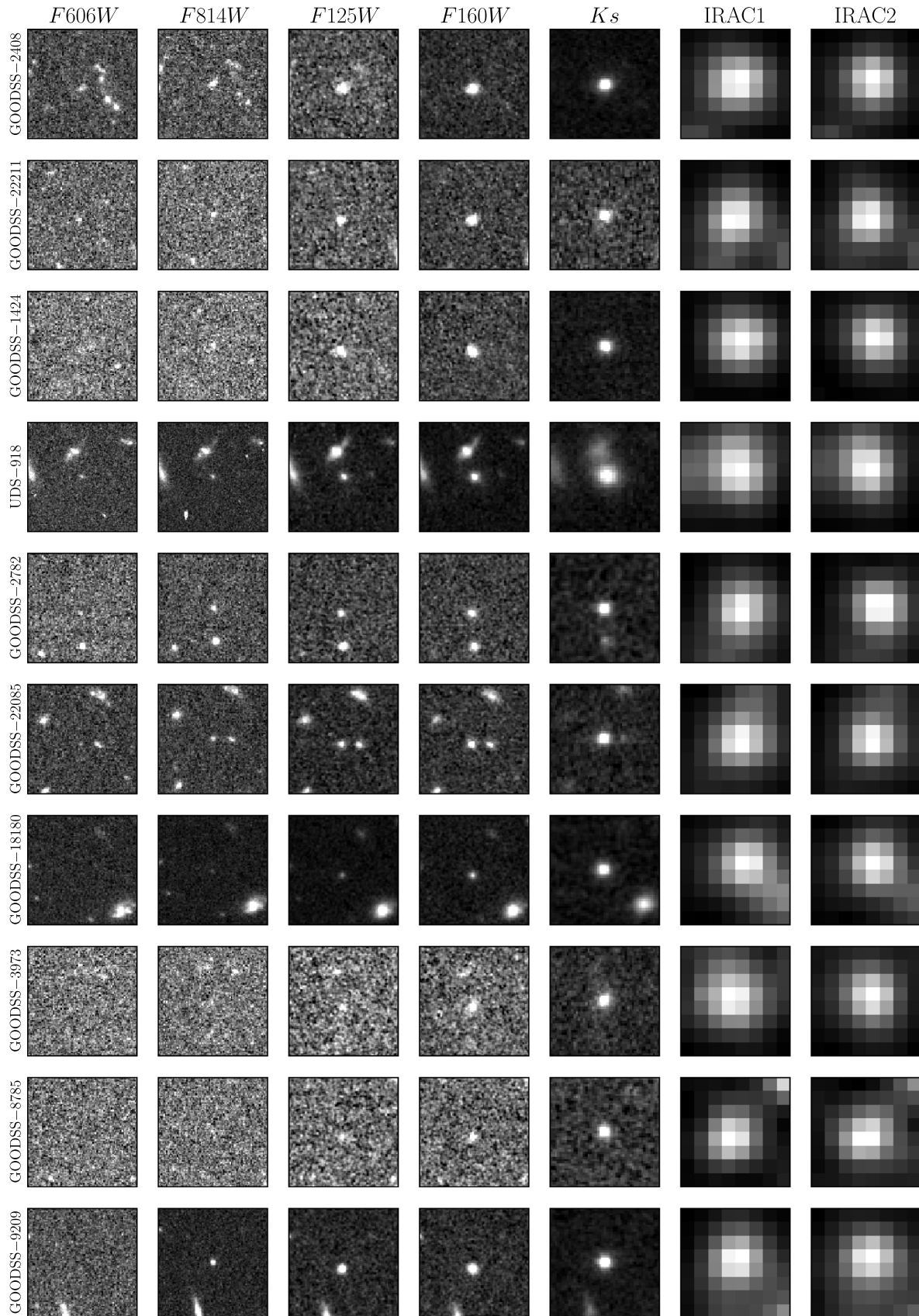


Figure 6. 5 arcsec \times 5 arcsec cut-outs for each of our 10 robust objects at $z > 3$. *HST* data were taken from CANDELS and Hubble Legacy Field mosaics (Teplitz et al. 2013; Rafelski et al. 2015; Whitaker et al. 2019). IRAC data is from SEDS and S-CANDELS (Ashby et al. 2015).

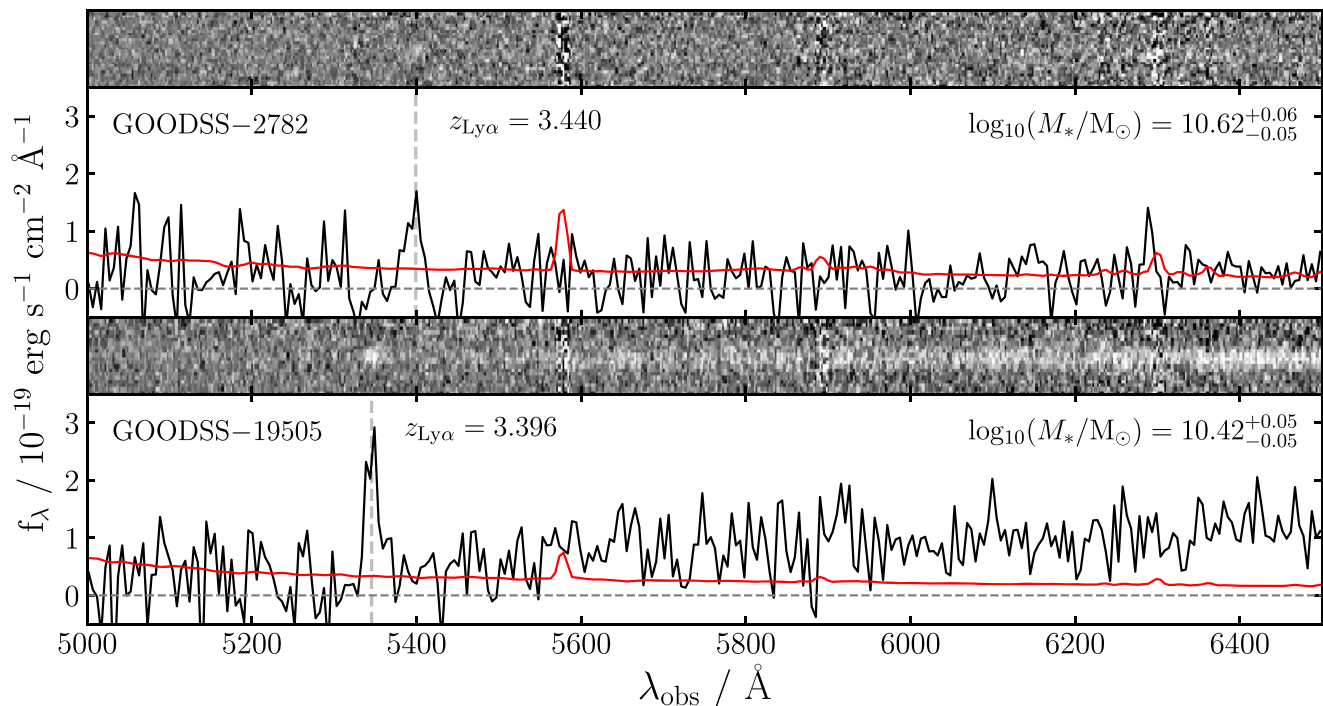


Figure 7. VANDELS spectra for two $z \sim 3.4$ galaxies from our sample. The 1D spectra are shown in black with error spectra shown in red. The 2D spectra are shown at the top of each panel. The upper and lower objects received 80 and 40 h of integration, respectively. The upper object is a member of our robust sub-sample, the lower is not. It can be seen that the lower object exhibits stronger Ly α despite its lower mass, implying that our robust selection criteria have been successful in identifying the objects with the lowest sSFRs.

put this into context, a galaxy of this redshift and stellar mass on the star-forming main sequence (SFMS) would have $\text{SFR} \sim 250 M_{\odot} \text{ yr}^{-1}$ (Speagle et al. 2014), suggesting this galaxy has an SFR suppressed by a factor of $\gtrsim 11$ relative to the SFMS.

5.2.2 GOODSS-19505

GOODSS-19505 is a member of our parent sample but not a member of our robust sub-sample, with 28 per cent of the BAGPIPES sSFR posterior above the $0.2/t_{\text{H}}$ threshold. This object was targeted by VANDELS for 40 h of integration. The resulting spectrum is shown in the bottom panel of Fig. 7. This spectrum shows a much clearer emission line, this time at 5350 \AA , as well as continuum emission at longer wavelengths. The photometric redshift we recover of $z = 3.49^{+0.09}_{-0.27}$ and CANDELS photometric redshift of $z_{\text{CANDELS}} = 3.47$ are again consistent with the identification of this line as Ly α , this time at $z = 3.396$.

Following the same procedure as Section 5.2.1 we find a raw line flux of $6.2 \pm 0.6 \times 10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2}$ and a corrected line flux of $4.4 \pm 0.8 \times 10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2}$. This implies a SFR of $0.28 \pm 0.05/f_{\text{esc, Ly}\alpha} M_{\odot} \text{ yr}^{-1}$ and therefore $\log_{10}(\text{sSFR}/\text{yr}^{-1}) = -11.0^{+0.2}_{-0.3} - \log_{10}(f_{\text{esc, Ly}\alpha})$. An escape fraction, $f_{\text{esc, Ly}\alpha} > 0.1$ is therefore required for this object to be consistent with our quiescent selection criterion.

Our finding of stronger Ly α and an implied higher sSFR for this object compared to GOODSS-2782 is encouraging, suggesting that our robust selection criteria are correctly identifying objects with the most suppressed sSFRs.

Santini et al. (2019) also report that this galaxy is undetected in ALMA band 6 observations, with an observed flux of $0.02 \pm 0.4 \text{ mJy beam}^{-1}$, implying a 1σ upper limiting SFR of $20.6 M_{\odot} \text{ yr}^{-1}$. A galaxy of this redshift and stellar mass on the SFMS would have $\text{SFR} \sim 180 M_{\odot} \text{ yr}^{-1}$, meaning that the SFR of this galaxy is suppressed by a factor of $\gtrsim 9$ relative to the SFMS.

5.2.3 GOODSS-9209

GOODSS-9209 is the highest redshift object in our robust sub-sample, shown in the bottom right-hand panel of Fig. 5 and the bottom row of Fig. 6. This object was observed for 14 h as part of VUDS, the resulting 2D and 1D spectra are shown in Fig. 8. The spectrum exhibits a continuum break around $\lambda \sim 7000 \text{ \AA}$, consistent with a redshift of $z \sim 4.75$. The VUDS team measure a redshift of $z = 4.657$, with a quality flag of 1, indicating 50 per cent probability of correct identification. The spectrum exhibits no obvious Ly α emission; however, the redshift implied by the continuum break is consistent with the photometric redshift we derive with BAGPIPES of $z = 4.72^{+0.06}_{-0.04}$. If confirmed, this would be the highest redshift known quiescent galaxy.

Santini et al. (2019) report that this galaxy is undetected in ALMA band 6 observations, with an observed flux of $-0.03 \pm 0.16 \text{ mJy beam}^{-1}$, implying a 1σ upper limiting SFR of $41.6 M_{\odot} \text{ yr}^{-1}$. A galaxy of the same stellar mass on the SFMS at this redshift would have $\text{SFR} \sim 300 M_{\odot} \text{ yr}^{-1}$. This implies the SFR of this galaxy is suppressed by a factor of $\gtrsim 7.5$ relative to the main sequence.

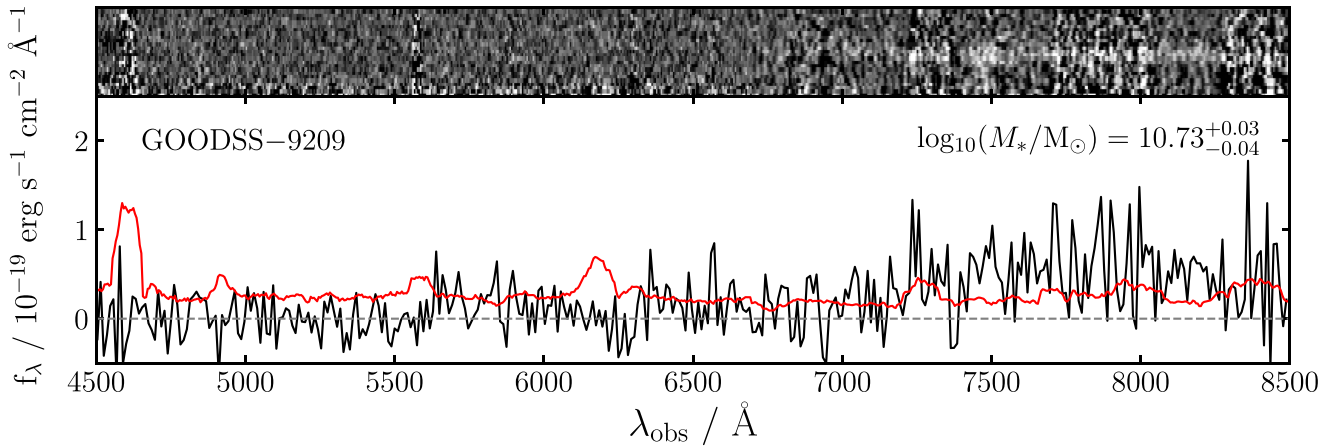


Figure 8. VUDS spectrum for GOODSS-9209, the highest redshift object in our robust sub-sample. The 1D spectrum is shown in black with the error spectrum shown in red. The 2D spectrum is shown at the top. This object received 14 h of integration. The continuum break in the spectrum at $\lambda \sim 7000 \text{ \AA}$ implies a redshift of $z \sim 4.75$, in good agreement with our photometric redshift of $z = 4.72^{+0.06}_{-0.04}$.

6 CONCLUSION

In this work, we have selected a sample of 151 massive ($M_* > 10^{10} M_\odot$) quiescent galaxies at $2 < z < 5$ from high-quality photometric data in the CANDELS UDS and GOODS-South fields. We apply a sophisticated Bayesian fitting approach, using the BAGPIPES code (Carnall et al. 2018) to fit a nine-parameter model to our data, including freedom in the observed redshift, a flexible dust curve shape and double power-law SFH model. This approach is designed to identify all potential solutions across a wide range of sSFRs and observed redshifts. Our sample includes a robust sub-sample of 61 objects for which we can exclude star-forming and low-redshift solutions with high confidence. Our robust sub-sample includes 10 objects at $z > 3$, of which 2 are at $z > 4$. A comparison of number densities between our sample and the spectroscopic sample of Schreiber et al. (2018) at $3 < z < 4$ yields good agreement, implying that our BAGPIPES fitting approach has been successful and confirming that our sample is not strongly contaminated.

We report rest-frame *UVJ* colours for our sample, demonstrating that the distribution of *UVJ* colours for the quiescent population becomes bluer with increasing redshift at $z \gtrsim 3$. This is predicted by the model of Belli et al. (2019), in which galaxies quench close to the lower left edge of the box, then passively age upwards and to the right. In addition, we identify a substantial population of quiescent objects below the standard $U - V = 1.3$ colour cut, a region that is sparsely populated at low redshift, and is strongly associated with post-starburst galaxies (e.g. Wild et al. 2014).

We report individual formation redshifts for the objects in our robust sub-sample, demonstrating that the $z > 2$ quiescent population is diverse, with formation redshifts ranging from $z \sim 6$ –10 to immediately preceding the redshift of observation. However, the individual ages and SFHs we obtain from our photometric data are significantly uncertain, particularly for older objects at $z < 3$, with an average uncertainty of ~ 0.5 Gyr on our inferred formation times.

These findings demonstrate the need for high-SNR continuum spectroscopic data to constrain the formation epochs and maximum historical SFRs of these galaxies (e.g. Carnall et al. 2019a). Such analyses would allow us to better constrain high-redshift quenching physics and to quantify the widely discussed connection between high-redshift massive quiescent galaxies and rapidly

star-forming submillimetre galaxies (e.g. Stach et al. 2019; Wild et al. 2020).

In an attempt to more strongly constrain the formation and quenching times of the oldest galaxies from photometry, we select the reddest and hence oldest, objects in our robust sub-sample in integer redshift bins and construct a model for the time-evolution of their *UVJ* colours. We find that the *UVJ* colours of these oldest objects can be explained by a model with a formation redshift of $z = 6.5$ and quenching redshift of $z = 4.7$. We find no evidence that quenching before $z = 5$ is required to explain the stellar populations we observe, though earlier formation is not ruled out by these data.

We report spectroscopic redshifts for two galaxies in our sample at $z \sim 3.4$, one of which is a member of our robust sub-sample. These galaxies exhibit weak Ly α emission in ultra-deep, 40 and 80 h optical spectra from VANDELS (McLure et al. 2018b; Pentericci et al. 2018). We estimate the SFRs implied by their Ly α fluxes, demonstrating that they are consistent with our quiescent galaxy selection criterion provided that these objects have Ly α escape fractions, $f_{\text{esc, Ly}\alpha} > 0.03$ and $f_{\text{esc, Ly}\alpha} > 0.1$ for the robust and non-robust objects, respectively. This suggests our robust selection criteria have been successful in identifying the objects with the lowest sSFRs. These objects are reported as being undetected in ALMA observations by Santini et al. (2019), implying SFRs suppressed by a factor of $\gtrsim 10$ compared to the star-forming main sequence at this redshift.

We finally present a spectrum from VUDS for the highest redshift object in our robust sub-sample, for which we find a photometric redshift of $z = 4.72^{+0.06}_{-0.04}$. This object exhibits a continuum break at $\lambda \sim 7000 \text{ \AA}$, implying a redshift of $z \sim 4.75$, consistent with our photometric redshift. This object is also undetected in ALMA observations, suggesting an SFR suppressed by a factor of $\gtrsim 7.5$ compared to the SFMS at our photometric redshift. If spectroscopically confirmed this would be the highest redshift known quiescent galaxy.

ACKNOWLEDGEMENTS

ACC and FC acknowledge the support of the UK Science and Technology Facilities Council. AC acknowledges the grant nos ASI n.2018-23-HH.0, PRIN MIUR 2015, and PRIN MIUR 2017

– 20173ML3WW_001. This work is based on data products from observations made with ESO Telescopes at La Silla Paranal Observatory under ESO programme ID 194.A-2003(E-Q). This work is based in part on observations made with the *Spitzer Space Telescope*, which is operated by the Jet Propulsion Laboratory, California Institute of Technology under a NASA contract. This research made use of ASTROPY, a community-developed core PYTHON package for Astronomy (Astropy Collaboration et al. 2013).

REFERENCES

Ashby M. L. N. et al., 2015, *ApJS*, 218, 33
 Astropy Collaboration et al., 2013, *A&A*, 558, A33
 Belli S., Newman A. B., Ellis R. S., 2019, *ApJ*, 874, 17
 Bessell M. S., 1990, *PASP*, 102, 1181
 Bressan A., Marigo P., Girardi L., Salasnich B., Dal Cero C., Rubele S., Nanni A., 2012, *MNRAS*, 427, 127
 Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, *MNRAS*, 351, 1151
 Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000
 Byler N., Dalcanton J. J., Conroy C., Johnson B. D., 2017, *ApJ*, 840, 44
 Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, *ApJ*, 533, 682
 Cardelli J. A., Clayton G. C., Mathis J. S., 1989, *ApJ*, 345, 245
 Carnall A. C. et al., 2019a, *MNRAS*, 490, 417
 Carnall A. C., McLure R. J., Dunlop J. S., Davé R., 2018, *MNRAS*, 480, 4379
 Carnall A. C., Leja J., Johnson B. D., McLure R. J., Dunlop J. S., Conroy C., 2019b, *ApJ*, 873, 44
 Cecchi R., Bolzonella M., Cimatti A., Girelli G., 2019, *ApJ*, 880, L14
 Chen H.-W., Marzke R. O., 2004, *ApJ*, 615, 603
 Cimatti A. et al., 2004, *Nature*, 430, 184
 Cimatti A. et al., 2008, *A&A*, 482, 21
 Conroy C., 2013, *ARA&A*, 51, 393
 Cullen F. et al., 2018, *MNRAS*, 476, 3218
 Cullen F. et al., 2019, *MNRAS*, 487, 2038
 Cullen F., McLure R. J., Khochfar S., Dunlop J. S., Dalla Vecchia C., 2017, *MNRAS*, 470, 3006
 Daddi E. et al., 2005, *ApJ*, 626, 680
 Davé R., Thompson R., Hopkins P. F., 2016, *MNRAS*, 462, 3265
 Davé R., Rafieferantsoa M. H., Thompson R. J., 2017, *MNRAS*, 471, 1671
 Draine B. T., Li A., 2007, *ApJ*, 657, 810
 Dunlop J. S. et al., 2017, *MNRAS*, 466, 861
 Dunlop J., Peacock J., Spinrad H., Dey A., Jimenez R., Stern D., Windhorst R., 1996, *Nature*, 381, 581
 Falcón-Barroso J., Sánchez-Blázquez P., Vazdekis A., Ricciardelli E., Cardiel N., Cenarro A. J., Gorgas J., Peletier R. F., 2011, *A&A*, 532, A95
 Fang J. J. et al., 2018, *ApJ*, 858, 100
 Ferland G. J. et al., 2017, *RMxAA*, 53, 385
 Feroz F., Hobson M. P., 2008, *MNRAS*, 384, 449
 Feroz F., Hobson M. P., Bridges M., 2009, *MNRAS*, 398, 1601
 Feroz F., Hobson M. P., Cameron E., Pettitt A. N., 2013, preprint ([arXiv:1306.2144](https://arxiv.org/abs/1306.2144))
 Fontana A. et al., 2009, *A&A*, 501, 15
 Fontana A. et al., 2014, *A&A*, 570, A11
 Forrest B. et al., 2020, *ApJ*, 890, L1
 Galametz A. et al., 2013, *ApJS*, 206, 10
 Gallazzi A., Charlot S., Brinchmann J., White S. D. M., Tremonti C. A., 2005, *MNRAS*, 362, 41
 Gallazzi A., Bell E. F., Zibetti S., Brinchmann J., Kelson D. D., 2014, *ApJ*, 788, 72
 Girelli G., Bolzonella M., Cimatti A., 2019, *A&A*, 632, A80
 Glazebrook K. et al., 2017, *Nature*, 544, 71
 Gobat R. et al., 2018, *Nat. Astron.*, 2, 239
 Grogin N. A. et al., 2011, *ApJS*, 197, 35
 Guo Y. et al., 2013, *ApJS*, 207, 24

Inoue A. K., Shimizu I., Iwata I., Tanaka M., 2014, *MNRAS*, 442, 1805
 Kauffmann G. et al., 2003, *MNRAS*, 341, 33
 Kennicutt R. C., Evans N. J., 2012, *ARA&A*, 50, 531
 Koekemoer A. M. et al., 2011, *ApJS*, 197, 36
 Kriek M. et al., 2015, *ApJS*, 218, 15
 Kroupa P., 2001, *MNRAS*, 322, 231
 Le Fèvre O. et al., 2015, *A&A*, 576, A79
 Leja J., Carnall A. C., Johnson B. D., Conroy C., Speagle J. S., 2019a, *ApJ*, 876, 3
 Leja J., Tacchella S., Conroy C., 2019b, *ApJ*, 880, L9
 Mancini C., Matute I., Cimatti A., Daddi E., Dickinson M., Rodighiero G., Bolzonella M., Pozzetti L., 2009, *A&A*, 500, 705
 Marigo P., Bressan A., Nanni A., Girardi L., Pumo M. L., 2013, *MNRAS*, 434, 488
 McLure R. J. et al., 2013, *MNRAS*, 428, 1088
 McLure R. J. et al., 2018a, *MNRAS*, 476, 3991
 McLure R. J. et al., 2018b, *MNRAS*, 479, 25
 Merlin E. et al., 2018, *MNRAS*, 473, 2098
 Merlin E. et al., 2019, *MNRAS*, 490, 3309
 Morishita T. et al., 2019, *ApJ*, 877, 141
 Mortlock A. et al., 2015, *MNRAS*, 447, 2
 Muzzin A. et al., 2013, *ApJS*, 206, 8
 Nayyeri H. et al., 2014, *ApJ*, 794, 68
 Noll S. et al., 2009, *A&A*, 499, 69
 Osterbrock D. E., 1989, *Astrophysics of Gaseous Nebulae and Active Galactic Nuclei*. University Science Books, California, USA
 Pacifici C. et al., 2016, *ApJ*, 832, 79
 Pentericci L. et al., 2018, *A&A*, 616, A174
 Rafelski M. et al., 2015, *AJ*, 150, 31
 Salim S., Boquien M., Lee J. C., 2018, *ApJ*, 859, 11
 Santini P. et al., 2019, *MNRAS*, 486, 560
 Schreiber C. et al., 2018, *A&A*, 618, A85
 Sobral D. et al., 2017, *MNRAS*, 466, 1242
 Sobral D., Santos S., Matthee J., Paulino-Afonso A., Ribeiro B., Calhau J., Khostovan A. A., 2018, *MNRAS*, 476, 4725
 Speagle J. S., Steinhardt C. L., Capak P. L., Silverman J. D., 2014, *ApJS*, 214, 15
 Stach S. M. et al., 2019, *MNRAS*, 487, 4648
 Straatman C. M. S. et al., 2014, *ApJ*, 783, L14
 Straatman C. M. S. et al., 2016, *ApJ*, 830, 51
 Tasca L. A. M. et al., 2017, *A&A*, 600, A110
 Teplitz H. I. et al., 2013, *AJ*, 146, 159
 Tokunaga A. T., Simons D. A., Vacca W. D., 2002, *PASP*, 114, 180
 Valentino F. et al., 2020, *ApJ*, 889, 92
 Whitaker K. E. et al., 2013, *ApJ*, 770, L39
 Whitaker K. E. et al., 2019, *ApJS*, 244, 16
 Wild V. et al., 2014, *MNRAS*, 440, 1880
 Wild V. et al., 2020, *MNRAS*, 494, 529
 Williams C. C. et al., 2019, *ApJ*, 884, 154
 Williams R. J., Quadri R. F., Franx M., van Dokkum P., Labbé I., 2009, *ApJ*, 691, 1879
 Wu P.-F. et al., 2020, *ApJ*, 888, 2

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

table2.ascii.txt

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

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