

Research Paper

Massive high-redshift quiescent galaxies with JWST

Themiya Nanayakkara¹✉, James Esdaile, Karl Glazebrook, Juan M. Espejo Salcedo, Mark Durre and Colin Jacobs

Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Hawthorn, VIC 3122, Australia

Abstract

Recent ground-based deep observations of the Universe have discovered large populations of massive quiescent galaxies at $z \sim 3-5$. With the launch of the *James Webb Space Telescope (JWST)*, the on-board Near-Infrared Spectrograph (NIRSpec) instrument will provide continuous $0.6-5.3 \mu\text{m}$ spectroscopic coverage of these galaxies. Here we show that NIRSpec/CLEAR spectroscopy is ideal to probe the completeness of photometrically selected massive quiescent galaxies such as the ones presented by Schreiber et al. (2018b, A&A, 618, A85). Using a subset of the Schreiber et al. (2018b, A&A, 618, A85) sample with deep Keck/MOSFIRE spectroscopy presented by Esdaile J., et al. (2021b, ApJ, 908, L35), we perform a suite of mock JWST/NIRSpec observations to determine optimal observing strategies to efficiently recover the star formation histories (SFHs), element abundances, and kinematics of these massive quiescent galaxies. We find that at $z \sim 3$, medium resolution G235M/FL170LP NIRSpec observations could recover element abundances at an accuracy of $\sim 15\%$, which is comparable to local globular clusters. Mimicking ZFOURGE COSMOS photometry, we perform mock spectrophotometric fitting with Prospector to show that the overall shape of the SFHs of our mock galaxies can be recovered well, albeit with a dependency on the number of non-parametric SFH bins. We show that deep high-resolution G235H/FL170LP integral field spectroscopy with a $S/N \sim 7$ per spaxel is required to constrain the rotational properties of our sample at $>2\sigma$ confidence. Thus, through optimal grism/filter choices, JWST/NIRSpec slit and integral field spectroscopy observations would provide tight constraints to galaxy evolution in the early Universe.

Keywords: infrared observatories – high-redshift galaxies – quenched galaxies – galaxy evolution – chemical abundances

(Received 2 March 2021; revised 23 November 2021; accepted 16 December 2021)

1. Introduction

The development of sensitive near-infrared (NIR) imaging instruments such as Magellan/FourStar (Persson et al. 2013), VLT/HAWK-I (Kissler-Patig et al. 2008), ESO/VISTA (Sutherland et al. 2015), and UKIRT/WFCAM (Casali et al. 2007) opened a new window into the early Universe. Observations from these instruments detected a high abundance of red galaxies at $z \sim 3-5$ (e.g. Marchesini et al. 2010; Spitler et al. 2014; Straatman et al. 2014; Patel et al. 2017). This high abundance of massive quiescent galaxies in the early Universe posed a significant challenge to current cosmological simulations (Sparre et al. 2015; Wellons et al. 2015; Davé, Thompson, & Hopkins 2016; Merlin et al. 2019). The tensions between observed number densities and cosmological simulations arise due to the short evolutionary time between the Big Bang and $z \sim 3-5$. For example, a massive quiescent galaxy with a stellar mass of $\sim 2 \times 10^{11} M_{\odot}$ at $z \sim 3.7$ (e.g. Glazebrook et al. 2017) needs to have formed all its stellar mass and undergone subsequent cessation of star formation within the first ~ 1.5 billion years of the Universe. Galaxy evolution and mass build-up within such a short time frame have strong implications for cosmological and chemical evolutionary models of the Universe. These massive $z \sim 3-5$ quiescent galaxies are ideal laboratories to determine how galaxies grew and what mechanisms shut down star formation in the early Universe.

Spectroscopy of $z \sim 3-5$ quiescent galaxies can be used to address the challenges faced by current cosmological models of the Universe. Spectroscopic confirmations of photometrically selected quiescent candidates are vital to provide tight constraints to their abundance in the $z \sim 3-5$ Universe. In addition, a detailed understanding of the stellar population properties of the $z \sim 3-5$ massive quiescent galaxies and their formation mechanisms can only be obtained through the spectroscopic analysis of different elements. Glazebrook et al. (2017) used deep Keck/MOSFIRE observations to spectroscopically confirm the very first massive quiescent galaxy in the $z > 3$ Universe. Their results showed that this galaxy was likely formed in a major star formation event at $z > 5$, with a star formation rate (SFR) exceeding $1000 M_{\odot} \text{yr}^{-1}$, posing a significant challenge to models. Subsequent studies have now started to build up samples of spectroscopically confirmed massive quiescent galaxies at $z \sim 3-4$ (e.g. Marsan et al. 2017; Schreiber et al. 2018b; Tanaka et al. 2019; Forrest et al. 2020a, 2020b; Valentino et al. 2020).

Spectroscopic confirmations of several massive $z \sim 3-5$ quiescent galaxies have further strengthened the need for galaxy formation models to provide efficient mass build-up and subsequent quick quenching mechanisms (e.g. Davé et al. 2016; Merlin et al. 2019). Formation timescales of these galaxies are an important quantity to be constrained as extended formation allows the galaxy to be assembled gradually in a hierarchy of mergers relieving the tension with current models. Star formation history (SFH) reconstruction (from spectral energy distribution (SED) modelling) of three $\sim 10^{11} M_{\odot}$ $z \sim 4$ quiescent galaxies by Valentino et al. (2020) showed that the majority of the stellar mass was formed within a ~ 50 Myr window with intense star formation

Author for correspondence: Themiya Nanayakkara, e-mail: wnanayakkara@swin.edu.au or themiyananayakkara@gmail.com

Cite this article: Nanayakkara T, Esdaile J, Glazebrook K, Espejo Salcedo JM, Durre M and Jacobs C. (2022) Massive high-redshift quiescent galaxies with JWST. *Publications of the Astronomical Society of Australia* 39, e002, 1–16. <https://doi.org/10.1017/pasa.2021.61>

followed by an abrupt quenching event. SFH analysis from limited spectral features suggests average formation timescales of $\lesssim 200$ Myr, which poses a challenge to early galaxy formation models (Glazebrook *et al.* 2017), and hierarchical models cannot produce these massive galaxies in a single rapid event.

The formation mechanisms of these massive $z \sim 3-5$ quiescent galaxies may have significant implications to the mass build-up of the early Universe. For example, if these galaxies were built up soon after the Big Bang in a short sharp star formation episode followed by an abrupt quenching, they would be a fossil record of the first generation of stellar populations in the Universe. Studying their stellar populations would open up a unique observational window to probe the star formation processes in the $z > 6$ Universe. Analysing the chemical signatures from the stars can provide vital clues to the early star formation processes.

If these galaxies are α -enhanced (e.g. Kriek *et al.* 2016), this would suggest that the interstellar medium (ISM) was preferentially enriched by core-collapse supernovae (Nomoto *et al.* 2006). Core-collapse supernovae are end of the life products of short-lived ($\sim 10-50$ Myr) massive stars and produce α -elements through supernovae nucleosynthesis. Longer lived (> 0.1 Gyr) lower mass stars result in Type Ia supernova and produce heavier elements such as Fe through supernova nucleosynthesis (Kobayashi & Nomoto 2009). Therefore, when the star formation episodes are shorter than ~ 1 Gyr, the new stars formed from the gas in the enriched ISM would be α -enhanced and would lack Fe. However, if the initial mass functions (IMFs) of these galaxies vary from local values, the interpretation of the star formation timescales becomes complicated (e.g. Martín-Navarro 2016).

Furthermore, these galaxies could have been built within a very short timescale from low metallicity gas with high specific SFRs. Under these conditions, some simulations predict a preferential formation of more massive stars (e.g. Narayanan & Davé 2012), resulting in a top-heavy IMF (e.g. Chon, Omukai, & Schneider 2021) and a higher characteristic stellar mass (Sharda & Krumholz 2022). However, currently there are no observational constraints of the stellar metallicities, α -enhancements, or the IMF in the $z > 6$ Universe. For example, spectral fitting to current ground-based spectroscopy of $z \sim 4$ massive quiescent galaxies shows degeneracies between high and low stellar metallicity solutions (Saracco *et al.* 2020). The S/N and spectral coverage of current data at $z \sim 3-5$ are insufficient to develop linkages to $z \sim 2$ observations of α -enhancements (e.g. Kriek *et al.* 2016) and simultaneous α and Fe enhancements (Jafariyazani *et al.* 2020). Deep rest-UV continuum observations are required to constrain the stellar population properties and ISM conditions of star-forming galaxies at $z > 6$, which are likely progenitors of massive quiescent galaxies at $z \sim 3-5$.

The abundance, formation mechanisms, and the stellar population properties of $z \sim 3-5$ massive quiescent galaxies also have strong implications for the reionisation and chemical evolutionary history of the Universe. If the stellar populations of these galaxies were α -enhanced and Fe-deficient, the massive stars would have had less Fe blanketing and less stellar winds in the atmospheres (Steidel *et al.* 2016). Therefore, these stars would have produced higher amounts of ionising photons compared to solar α -abundance stars (Pauldrach, Hoffmann, & Lennon 2001). A top-heavy IMF in these galaxies could have resulted in higher rates of core-collapse supernovae. Core-collapse supernovae have higher IMF averaged ejecta mass compared to Type Ia supernovae, which leads to stronger feedback mechanisms in galaxies

(Hopkins *et al.* 2018). This leads to creating more possibilities within the geometry of galaxies to create holes for ionising photons to escape. The enhancement of ionising photons along with strong supernova feedback-driven changes in the ISM geometry may alter the contribution of massive galaxies to the reionisation of the Universe at $z > 6$ (Naidu *et al.* 2019). Thus, exploration of $z \sim 3-5$ massive quiescent galaxies may lead to newer challenges in cosmology.

Kinematical properties of $z \sim 3-5$ massive quiescent galaxies also provide clues for their formation scenarios. Deep rest-frame optical spectra have opened the door to obtaining velocity dispersion measurements of massive quiescent galaxies at $z > 3$ (Tanaka *et al.* 2019; Esdaile 2021b; Saracco *et al.* 2020). When the $z > 3$ quiescent galaxies are compared in the size-mass plane, they require a greater size evolution compared to what is expected from minor mergers and show evidence for dynamical masses to be lower than the stellar mass estimates (Esdaile 2021b). Additionally, despite sharing similar attributes as local elliptical galaxies, namely red colours and high stellar masses, there are indications that the morphologies of high-redshift massive quiescent galaxies are quite different. The modelled axis ratio from *HST* imaging shows indications of flattened ‘disc-like’ morphology in massive quiescent galaxies at $z > 3$ (Hill *et al.* 2019).

If these galaxies are formed in short durations, mass growth through mergers would be uncommon; thus, disc-like morphologies would be prominent at high redshift. However, morphological constraints at $z > 3$ are poor because existing high-resolution size measurements are based on rest-frame UV and are susceptible to underestimation from potential dust-reddening or overestimation from recent star formation events. Rest-frame optical size measurements, which trace older stellar mass, provide more robust measurements of size (Kubo *et al.* 2018). However, none of the current observations probe galaxies at sufficient depth and/or at the appropriate rest-frame wavelength windows to access features that could provide stronger constraints to the formation timescales (e.g. Schreiber *et al.* 2018b; Forrest *et al.* 2020a; Carnall *et al.* 2020) or morphologies.

The launch of the *James Webb Space Telescope (JWST)* in 2021 provides a unique opportunity for the detailed exploration of $z \sim 3-5$ massive quiescent galaxies. JWST Near-Infrared Spectrograph (NIRSpec) will obtain high-quality NIR spectroscopy of galaxies in the early Universe with high efficiency. This is due to its high sensitivity, multiplexing capabilities, continuous spectral coverage from 0.6 to 5.3 μ (Birkmann *et al.* 2011), low thermal background of the observatory, and the lack of atmospheric contamination in space. Thus, by obtaining spectroscopic confirmations for photometrically selected candidates, JWST will enable astronomers to provide tighter constraints to the formation histories, stellar populations, and kinematics of $z \sim 3-5$ quiescent galaxies.

JWST/NIRSpec will obtain rest-frame optical spectra of $z \sim 3-5$ quiescent galaxies, the gold standard required to reconstruct their formation histories. This wavelength regime covers a variety of important absorption features which can be compared with stellar evolutionary models to determine the star formation timescales and elemental abundances (Conroy 2013). Balmer absorption lines have a strong dependence on the age and the SFH of galaxies (Poggianti & Barbaro 1997). The variety of metal absorption lines such as Mg, Na, Ca, Ti, Na, and Fe observed in the rest-frame optical can be used with chemical evolution models to link to properties of previous star formation episodes (de La Rosa *et al.* 2011; Segers *et al.* 2016). Additionally, some of these features

are sensitive to different types of stars; thus, they can be used to infer stellar abundances and hence the IMF of these galaxies (e.g. La Barbera et al. 2013).

In this paper, we perform a JWST/NIRSpec observational case study based on a sample of galaxies presented by Schreiber et al. (2018b). The Schreiber et al. (2018b) sample is one of the most deepest and complete spectroscopic samples of $z \sim 3-5$ quiescent galaxies presented to date. They presented spectroscopy for 24 photometrically selected quiescent galaxy candidates with a $\sim 50\%$ spectroscopic completeness rate and a $\sim 90\%$ purity. The $z > 3$ spectroscopic sample reaches $K \sim 23.8$ with typical stellar masses of $\sim 0.3 - 2.0 \times 10^{11} M_{\odot}$. In Section 2, we discuss how JWST could be optimally used to confirm the redshift and quiescence of the photometrically selected quiescent galaxies to improve on the $\sim 50\%$ completeness rate. This is a crucial first step that is necessary before investing deeper observations to obtain stellar population and dynamical properties. Esdaile (2021b) presented deep ground-based KECK/MOSFIRE spectroscopy for the brightest galaxies in the Schreiber et al. (2018b) sample. This sample is biased towards the most massive $z \sim 3-4$ massive quiescent galaxies. Therefore, it probes an interesting range in mass, size, and velocity dispersion at $z > 3$ for JWST follow-up to obtain element abundances and resolved kinematics to constrain formation scenarios of massive quiescent galaxies. Thus, in Section 3 we discuss how JWST could enable detailed analysis of stellar population properties of the Esdaile (2021b) sample and present optimal observing strategies. Next, in Section 4 we discuss how JWST slit and integral field unit (IFU) spectroscopy could constrain the dynamics of the sample. Finally, in Section 5 we briefly discuss the expected advancements in the field with JWST and provide our conclusions. We assume a cosmology with $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_{\Lambda} = 0.7$, and $\Omega_M = 0.3$. We use AB magnitudes throughout the paper.

2. Probing the completeness with JWST

2.1 The need for spectroscopic confirmations

Deep multi-wavelength surveys utilising medium-bands (such as ZFOURGE; Straatman et al. 2016) have demonstrated that galaxy SED fitting with photometric data can obtain redshifts and evolutionary types of galaxies (i.e. blue/red star-forming or quiescent) with high accuracy (e.g. Spitler et al. 2014; Straatman et al. 2014). However, spectroscopy is crucial to constrain the number density of massive quiescent galaxies in the $z \sim 3-5$ epoch and rule out redshift and non-quiescent outliers (e.g. Schreiber et al. 2018b). Thus, in order to make robust comparisons with cosmological models, obtaining redshift confirmations and indications for quiescence from galaxy spectra is imperative.

The main limitation of purely characterising quiescent galaxies based on photometry is the redshift uncertainty. In SED fitting techniques, the shape of the multi-wavelength photometry is used to infer the redshift. Prominent breaks such as the Lyman and Balmer breaks (Alcalde-Pampliega et al. 2019) and the D4000 features have shown to constrain the redshift to $\sim 2\%$ accuracy (e.g. Whitaker et al. 2011; Straatman et al. 2016). As an example, Straatman et al. (2016) find that the inclusion of the FourStar (Persson et al. 2013) NIR J1, J2, J3, Hs, Hl filters reduces the photometric redshift uncertainty by up to $\sim 50\%$ at $z > 1.5$. Additionally, the redshift outlier fraction has also been shown to be dependent on the magnitude of a galaxy, with fainter sources showing

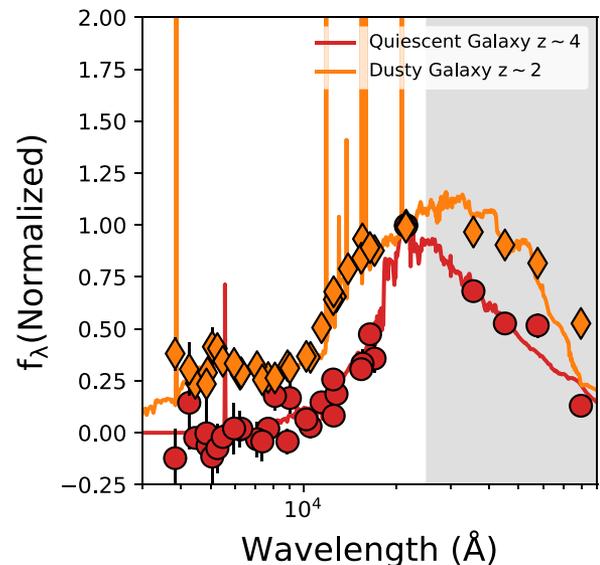


Figure 1. Two example SEDs from the ZFOURGE survey. Shown in red is a quiescent galaxy SED at $z \sim 4$ and in orange is a dusty star-forming galaxy at $z \sim 2$. The spectra are normalised in the K band. The spectral shape of the normalised SEDs of the $z \sim 4$ quiescent and $z \sim 2$ dusty star-forming at $\lesssim 2 \mu\text{m}$ looks largely similar; thus, spectroscopy is required to obtain emission lines to distinguish between the two types of galaxies. Alternatively, infrared/sub-mm observations can also be used to detect the dust continuum emission of dusty star-forming galaxies.

a systematic increase in the outlier fraction (Brinchmann et al. 2017). Thus, degeneracies between redshift, galaxy type, and S/N constraints add uncertainty to the determination of photometric redshifts. For example, as shown in Figure 1, the SED shape of $z \sim 2$ dusty (red) star-forming galaxies (e.g. Spitler et al. 2014) is largely similar to that of quiescent galaxies at $z \sim 4$. The Balmer break of $z \sim 4$ galaxies observed in the K -band could provide some constraints; however, the uncertainties in constraining the redshift and strong $H\alpha$ emission that contaminate $z \sim 2$ K -band photometry limit the diagnostic power.

The IR emission from the dusty star-forming galaxies can be used as an alternate indicator to differentiate them from $z \sim 4$ quiescent galaxies. In the pre-JWST era, this required extremely deep Spitzer or Herschel photometry and accurate source deblending techniques to account for the large PSF in such observations (e.g. Stefanon et al. 2021). JWST with $\sim 50\times$ collecting area than Spitzer will obtain exquisite spatial quality deep near and mid-IR observations in the *Hubble* legacy fields to revolutionise this space. Additionally, ALMA sub-mm observations can be used to detect molecular gas content (e.g. Boogaard et al. 2021) as an indirect tracer of quiescence, though the presence of gas around the galaxy is not necessarily an indicator for ongoing star formation (e.g. Kalita et al. 2021). However, ALMA is not a survey instrument and thus has a limited field of view (FoV). Therefore, covering an extragalactic field such as the *Hubble* Ultra Deep Field is very expensive (Walter et al. 2016).

Dusty star-forming galaxies can also be identified through rest-frame optical emission lines which can be obtained with relatively short integration times. Thus, spectroscopy can efficiently remove contaminants in photometrically selected galaxies such as dusty star-forming galaxies and bright active galactic nuclei. Spectroscopic confirmations provide tighter constraints to the evolution of the number density of massive quiescent galaxies

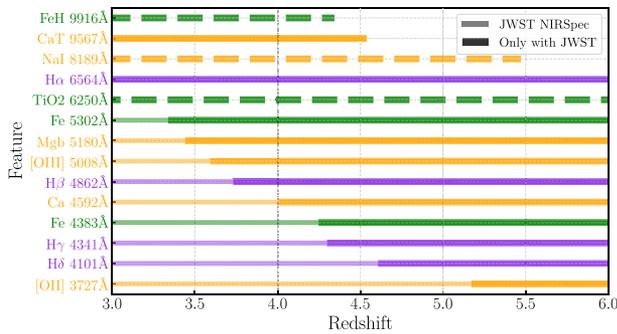


Figure 2. JWST NIRSpec wavelength coverage of rest-frame optical features that are crucial to determine the quiescence and constrain the stellar population properties of massive $z \sim 3-5$ quiescent galaxies. Redshifts only accessible via space-based spectroscopy of JWST are shown by the thicker/darker colour lines. Balmer emission/absorption features (sensitive to the SFH) are labelled in purple, forbidden emission lines, and α -element absorption lines are labelled in orange, while the other absorption features that constrain the overall stellar metallicity are labelled in green. IMF sensitive features are shown by dashed lines. It is evident that JWST NIRSpec spectroscopy is crucial to obtain a suite of spectral features that are necessary to analyse the $z \sim 3-5$ quiescent galaxy populations.

with cosmic time at an accuracy far greater than $\sim 2\%$ that can be obtained through photometric redshifts. Given JWST FoV, this can only be achieved through spectroscopic follow-up of quiescent galaxy candidates selected from deep JWST imaging surveys (e.g. Kauffmann *et al.* 2020) and/or deep ground-based imaging surveys (e.g. Weaver *et al.* 2021). Based on the ZFOURGE survey, Schreiber *et al.* (2018b) estimated the number density of quiescent galaxies between $3 < z < 4$ to be $\sim 2 \times 10^{-5} \text{ Mpc}^{-3}$. The NIRCам GTO JADES survey will cover $\sim 200 \text{ arcmin}^2$ FoV reaching up to ~ 30 th magnitude in the NIRCам filters (Williams *et al.* 2018). This is ~ 5 magnitudes deeper than the ZFOURGE survey but only half as small in terms of the FoV. Therefore, JADES alone would obtain a significantly higher number of older (fainter) quiescent galaxies than any ground-based survey has currently observed.

In addition to the spectroscopic confirmation of the redshift, spectroscopy provides crucial coverage of important spectral features that determines the quiescence of galaxies. In Figure 2, we show the wavelength coverage of spectral features that are necessary to determine the quiescence and stellar population properties of $z \sim 3-5$ quiescent galaxies. By obtaining coverage of Balmer lines or forbidden lines such as [O II] and [O III], strong constraints can be placed on the levels of star formation in these massive quiescent candidates. Absorption features of α -elements and other metals provide information on the stellar populations which is crucial to determine the SFH. When the redshift and quiescence are constrained through spectroscopy, multi-wavelength SED fitting would provide stronger constraints to stellar masses and SFHs, which are also important parameters to test cosmological evolution models.

2.2 NIRSpec PRISM/CLEAR spectroscopy as a redshift confirmation machine

In order to address the number density of massive quiescent galaxies in the $z \sim 3-5$ epoch, a systematic spectroscopic follow-up of all massive quiescent galaxy candidates is required. Rest-frame $U - V$ vs $V - J$ colour distributions have shown to be effective in identifying the star-forming and quiescent galaxies from each other in various cosmic epochs (e.g. Williams *et al.* 2009;

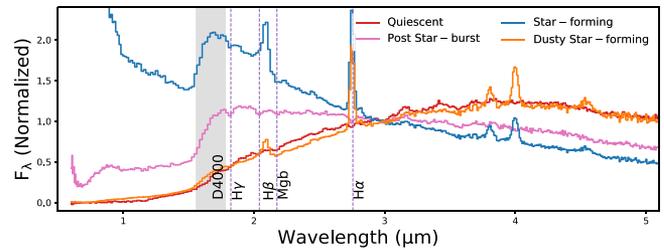


Figure 3. Simulated observations of JWST/NIRSpec PRISM/CLEAR spectroscopy for a quiescent ($E_{(B-V)} = 0.5$), a post-starburst ($E_{(B-V)} = 0.3$), dusty star-forming ($E_{(B-V)} = 1.0$), and dust-free star-forming ($E_{(B-V)} = 0.0$) galaxy at $z \sim 3$. Redder NIR colours due to age-dust degeneracies make it challenging to identify quiescent galaxies purely based on photometric observations. The high sensitivity coupled with the multiplexing capability makes NIRSpec an ideal instrument to obtain $0.6-5.3 \mu\text{m}$ spectroscopy to accurately distinguish between these types of galaxies and confirm their quiescence. The spectra shown here are all normalised to $K = 21.5$, and the typical continuum S/N obtained is $\sim 80 - 100$ in a $\sim 1500 \text{ s}$ exposure.

Straatman *et al.* 2014). Recent ground-based efforts have been successful in obtaining the rest-frame optical coverage of the brightest of these quiescent candidates (e.g. Schreiber *et al.* 2018b; Valentino *et al.* 2020). They have shown a high success rate in identifying quiescent galaxies based on rest-frame $U - V$ vs $V - J$ colours with a high purity but not necessarily high completeness. For example, Schreiber *et al.* (2018b) analysed 24 $U - V$ vs $V - J$ colour selected galaxies and obtained redshift confirmations for 12 galaxies out of which only two were $z \sim 2$ dusty star-forming interlopers (see discussion in Schreiber *et al.* 2018b).

However, studies have so far only targeted the brightest quiescent galaxies in the $z \sim 3-5$ epoch. This is driven by observational challenges in obtaining ground-based NIR spectroscopy. The limited wavelength coverage due to atmospheric absorption, strong skyline contamination, and limited multiplexing capabilities in ground-based NIR instruments have traditionally challenged astronomers to spectroscopically follow up mass/magnitude complete samples of quiescent galaxies. Additionally, targeting only the brightest galaxies biases the observed samples since most massive quiescent galaxies are in general older and fainter.

The multiplexing capability and the $0.6-5.3 \mu\text{m}$ wavelength coverage make JWST/NIRSpec an ideal instrument to obtain spectroscopic redshifts of photometrically selected massive quiescent galaxy candidates. The PRISM/CLEAR disperser/filter combination of NIRSpec provides continuous spectral coverage between 0.6 and $5.3 \mu\text{m}$ at a resolution of $R \sim 100$. Thus, by selecting quiescent candidates from deep ground and space-based imaging surveys, NIRSpec PRISM/CLEAR spectroscopy can be used to obtain the crucial rest-frame optical absorption features that are necessary to confirm the redshifts, quiescence, and rule out interlopers.

In Figure 3, we show that the NIRSpec PRISM/CLEAR spectroscopy is capable of accurately distinguishing between $z \sim 3$ quiescent, post-starburst, dusty (red) star-forming, and dust-free (blue) star-forming galaxies. Blue star-forming galaxies can be clearly identified based on the NIR colours and rest-UV continuum; however, prism spectra provide emission line measurements to further constrain the ISM and star formation properties of these galaxies. NIR colours will be similar between dusty star-forming and quiescent galaxies; however, the Balmer and forbidden emission features in the star-forming galaxies will be detectable by

$R \sim 100$ spectroscopy to rule out any star-forming contaminants. $H\beta$ and $H\gamma$ equivalent width can be used to determine the time since the last star formation episode (e.g. Glazebrook et al. 2017) and comparisons of the continuum around D4000 feature and the rest-frame near-UV continuum can distinguish between post-starburst and dusty star-forming/quiescent galaxies.

In addition to the simultaneous 0.6–5.3 μm wavelength coverage, the sensitivity of the NIRSpec instrument and the lower background level in space-based observations increase the efficiency of the JWST observations. Spectra shown in Figure 3 are generated using the Flexible Stellar Population Synthesis (FSPS) code (Conroy & Gunn 2010) and are normalised to $K = 21.5$ in the JWST exposure time calculator (ETC). With an exposure time of ~ 1500 s, ETC predicts that NIRSpec PRISM/CLEAR spectroscopy will obtain a typical continuum S/N of ~ 80 – 100 for these galaxies. Dusty star-forming, post-starburst, and quiescent galaxies are reddened with a $E_{(B-V)} = 1.0, 0.3, 0.5$, respectively, following the Cardelli, Clayton, & Mathis (1989) extinction law. No extinction is applied to the dust-free star-forming galaxies. For simplicity, we assume the sources to be pointlike. Glazebrook et al. (2017) obtained a S/N of 6 for a $K = 22.5 z = 3.7$ quiescent galaxy with 7 h of K band Keck/MOSFIRE spectroscopy at a resolution of 19 Å per pixel. For a $K = 21.5$, source Keck/MOSFIRE can reach a S/N of ~ 9 in 1500 s at similar resolution to NIRSpec PRISM/CLEAR spectroscopy.^a Therefore, even when only a limited wavelength coverage is considered, NIRSpec PRISM/CLEAR spectroscopy is $\sim \times 10$ more efficient.

As a potential caveat to the efficiency of quiescent galaxy surveys using NIRSpec PRISM/CLEAR, we also note that the NIRSpec micro-shutter assembly (MSA) adds extra observational challenges in obtaining uncontaminated spectra of galaxies. In addition to contamination from fixed open shutters, as stated in the JDox, MSA flux leakage due to small gaps in the MSA shutters can lead to $\sim 2\%$ contamination^b. Therefore, to obtain deep spectroscopy of faint spectral features, the NIRSpec fixed slits (S200A1 or S200A2 with 0.2" slit width) is ideally suited. Fixed slits provide the cleanest spectra in the NIRSpec detectors and yield a higher S/N compared to the MSA mode for the same exposure time. Additionally, it is possible to obtain simultaneous MSA and fixed slit spectroscopy by positioning the primary target of interest in the fixed slit location and populating the MSA shutters with secondary targets. This can be achieved by either configuring the MSA to only allow for rotational variation using the MSA planning design tool or by manually configuring the MSA shutters in open/close positions.

3. Stellar populations with JWST

3.1 Stellar population models used in this analysis

We use two different stellar population models in our analysis. To investigate how the spectral features change based on element abundances, we use the Villaume et al. (2017) empirical SSPs using the `alf` software (Conroy et al. 2018) to generate synthetic galaxy spectra. Villaume et al. (2017) models have the flexibility

^aSuch observations will be carried out by JWST Cycle 1 General Observers program ID 2565 'How Many Quiescent Galaxies Are There at $3 < z < 4$ Really?'.
^bhttps://jwst-docs.stsci.edu/near-infrared-spectrograph/nirspec-instrumentation/nirspec-instrumentation/nirspec-micro-shutter-assembly#NIRSpecMicro-ShutterAssembly-msa_leakageMSAfluxleakage.

Table 1. SFHs generated using FSPS models. The parameters related to the SFH used in FSPS are shown by the columns. Model A is an SSP, while Models B, C, and D have parametric SFHs. All models are generated at solar metallicity at an age of 1 Gyr.

Model name	sfh	tau	const	sf_start	sf_trunc	tburst	fburst
A	0	–	–	–	–	–	–
B	1	1.0	0.5	0.0	0.5	0.0	0.0
C	1	1.0	0.2	0.0	0.0	0.7	0.3
D	1	1.0	0.0	0.0	0.0	0.0	0.0

to change the abundance of individual elements which makes it ideally suited for this task. We use the same models to also investigate the accuracy in recovering element abundances from mock observed spectra as detailed in Section 3.2.

To investigate the role of SFH in recovering galaxy properties, we use the FSPS code (Conroy & Gunn 2010) to generate synthetic spectra for 4 mock galaxies as described in Table 1. Within FSPS, we select the MILES spectral library (Vazdekis et al. 2010) and MIST isochrones (Choi et al. 2016) to generate the synthetic models. All galaxies are generated at solar metallicity following a Calzetti et al. (2000) dust law with an optical depth of $\tau_{dust} = 0.5$. Unless specified all other parameters are kept at the `python-fsps` default values. Galaxy Model A is formed via a single burst event at $t = 0$. Model B has a delayed τ model in the form of $te^{-t/\tau}$ with $\tau = 1$ Gyr. A constant SFH is superimposed in the model to generate 50% of the total mass at 1 Gyr, and the star formation is truncated at 500 Myr. Model C is an exponentially declining SFH with a timescale of $\tau = 1$ Gyr. 20% of the total mass at 1 Gyr is formed through a constant SFH episode. Both of these are combined with an instantaneous burst at 700 Myr that generates 30% of the total stellar mass at 1 Gyr. Model D is an exponentially declining SFH with a timescale of $\tau = 1$ Gyr. Models C and D have declining but residual star formation, while Models A and B have been quenched for > 0.5 Gyr. All FSPS SFH parameters used are presented in Table 1 and are also shown in Figure 4.

The FSPS galaxies are normalised to a $10^{11} M_{\odot}$ to be similar to the Esdaile (2021b) sample. Observed galaxy spectra are generated at 1 Gyr and are fed through the JWST ETC to compute the observed spectrum from the G235M/FL170LP grism/filter combination with a continuum S/N ~ 30 per pixel. ETC output spectra are instrument throughput corrected and flux calibrated.

3.2 Element abundances

Stellar population properties of galaxies can be obtained via rest-UV and optical absorption line spectroscopy. O- and B-type stars which contribute to UV flux have short lifetimes. Therefore, rest-frame optical features from A and G type stars play the most important role in deciphering the underlying stellar population properties of quiescent galaxies. In Figure 5, we show prominent rest-frame optical absorption lines that are necessary to constrain the ages, metallicities, and element abundances of quiescent galaxies based on the Villaume et al. (2017) empirical SSPs. We choose SSPs purely for illustrative purposes because it simplifies the variation of the selected spectral features with time, α -abundance, and metallicity. Realistic SFHs with multiple generations of stellar populations require more advanced simultaneous treatment of multiple spectral features and are further explored later in the context of full spectral fitting.

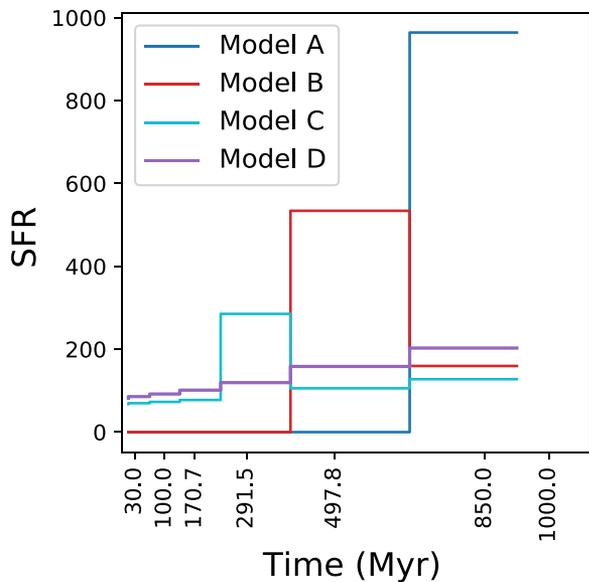


Figure 4. Visualisation of input SFHs generated using FSPS models as detailed in Table 1. The input parametric forms are averaged using seven time-bins which are fixed similar to the Prospector fits as described in Section 3.3. In lookback time, the first two bins are fixed to be between 0–30 and 30–100 Myr. The most distant bin is fixed to be between 850 and 1 000 Myr. The remaining four bins are split in logarithmic time evenly between 100 and 850 Myr.

It is clear from Figure 5 that not all features differ substantially between models such that they can be distinguished at a given sensitivity. As an example, the SFH (or rather the time since the last star formation episode) can be well constrained by the Balmer absorption lines with reasonable S/N limits (~ 100). However, in order to map changes in the lower mass IMF slope through the IMF sensitive features such as the Fe-H band and NaD absorption (e.g. Conroy & van Dokkum 2012) requires extremely high S/N sensitivity of $\gtrsim 350$. Achieving such levels of S/N in spectroscopy for populations of galaxies at $z > 3$, even with JWST requires upwards of hundreds of hours per target and is not practical. However, we note that these S/N levels are only presented in the context of simple SFHs and assume that diagnostics rely on a single feature. Much tighter constraints can be made by combined analysis of spectral features. Local galaxy analysis between individual features (i.e. Lick indices; Burstein *et al.* 1984) and full spectral fitting have shown that results from both methods are in broad agreement with each other (e.g. Conroy, Graves, & van Dokkum 2014).

Next, we go beyond individual spectral line analysis and perform full spectrum fits to identify optimal S/N levels and JWST/NIRSpec grism/filter combinations required to constrain the element abundances and the SFHs of $z \gtrsim 3$ quiescent galaxies similar to that of Esdaile (2021b) sample. The combined role of spectroscopy and multi-wavelength photometry in constraining the complex SFH of galaxies will be addressed in Section 3.3.

Rest-frame optical spectroscopy covers the spectral features that are necessary to constrain the formation history of quiescent galaxies. In addition to constraining the last star formation episode from the Balmer absorption features, the coverage of absorption features from elements such as Mg, Fe, Ti provides additional constraints to the durations of previous star formation episodes, metallicities, and chemical abundance patterns of galaxies (e.g.

Vazdekis *et al.* 2010). These are crucial to determine the SFHs of the early quiescent galaxies and can be compared with Λ CDM hierarchical galaxy formation models.

Here, we investigate what optimal JWST NIRSpec grism/filter combinations and S/N thresholds are required to accurately obtain element abundances of a simulated massive quiescent galaxy similar to that of Esdaile (2021b) sample. We use Villaume *et al.* (2017) empirical SSPs to generate a mock 1.5 Gyr old galaxy at $z = 3.2$ with velocity dispersion of 300 km s^{-1} and keep all elemental abundances at solar. We feed the mock galaxy to the JWST ETC to obtain a suite of mock observables at different S/N values for the JWST NIRSpec fixed slit S200A1 with G235H/FL170LP and G395H/FL290LP grism/filter combinations. We then use the full spectral fitting code *a1f* (Conroy *et al.* 2018) on the calibrated data to fit for velocity, velocity dispersion, age, $[Z/H]$, and element abundances of C, N, O, Mg, Si, Ca, Ti, and Na. Spectral fitting is performed to each individual grism/filter combination and to the combined spectrum from both grisms/filters to obtain 100 best-fit values for the input spectrum.

In Figure 6, we show the element abundance recovery of $[Mg/Fe]$, $[Ti/Fe]$, and $[Fe/H]$ for our mock observables. It is evident that at native resolutions of each grism/filter combinations, at a S/N $\gtrsim 30$ per pixel the recovery of the elemental abundances converges to the input values. We further find that there is no significant difference between the two grism/filter combinations. However, when the spectra from both grism/filter combinations are fit together, the uncertainty is slightly decreased. Even though spectral features such as Mg b are not covered by G395H/FL290LP for $z = 3.2$ galaxies, inherent relationships between various α -elements and other metals in the Villaume *et al.* (2017) templates allow Mg abundance to be converged albeit with slight systemic offset. For example, because Mg and Ca are both α elements, the coverage of the $\text{CaT}\lambda 8498\lambda 8542\lambda 8662$ features can be used to constrain the Mg abundance using theoretical response functions (Conroy *et al.* 2018). Thus, abundances of elements not included in the spectral range are inferred from predetermined element abundance ratios and chemical evolutionary models. However, direct measurements of element abundances should be preferred particularly for high-redshift galaxies where these assumptions may not hold.

Thus, for the $z \sim 3-4$ Esdaile (2021b) sample, most of the spectral features that are necessary to obtain α -element abundances can be obtained by the G235H/FL170LP grism/filter combination. For the fixed observed wavelength coverage of the FL170LP filter, we investigate whether the choice of the spectral resolution offered by the G235M and G235H grisms plays a significant role in the convergence of mock observable parameters. In Figure 7, we show the recovery of the parameters for the G235M and G235H grisms. It is evident that there is no significant difference between the two grisms for the recovery of the $[Mg/Fe]$, $[Ti/Fe]$, and $[Fe/H]$ element abundances. Therefore, it is advantageous to obtain G235M spectroscopy to increase the efficiency of $z > 3$ quiescent galaxy observing programmes. From our *a1f* simulations, we find that at a S/N of ~ 30 per pixel, an accuracy of $\sim 15\%$ can be obtained for element abundances which are comparable to the accuracy obtained for local globular clusters (Conroy *et al.* 2018).

Next, we go beyond SSPs to investigate the role of SFH in recovering element abundances from the G235H/FL170LP grism/filter combination. We use FSPS code (Conroy & Gunn 2010) to generate four different SFHs as detailed in Section 3. Simulated galaxy spectra are generated at 1 Gyr of age with a velocity dispersion

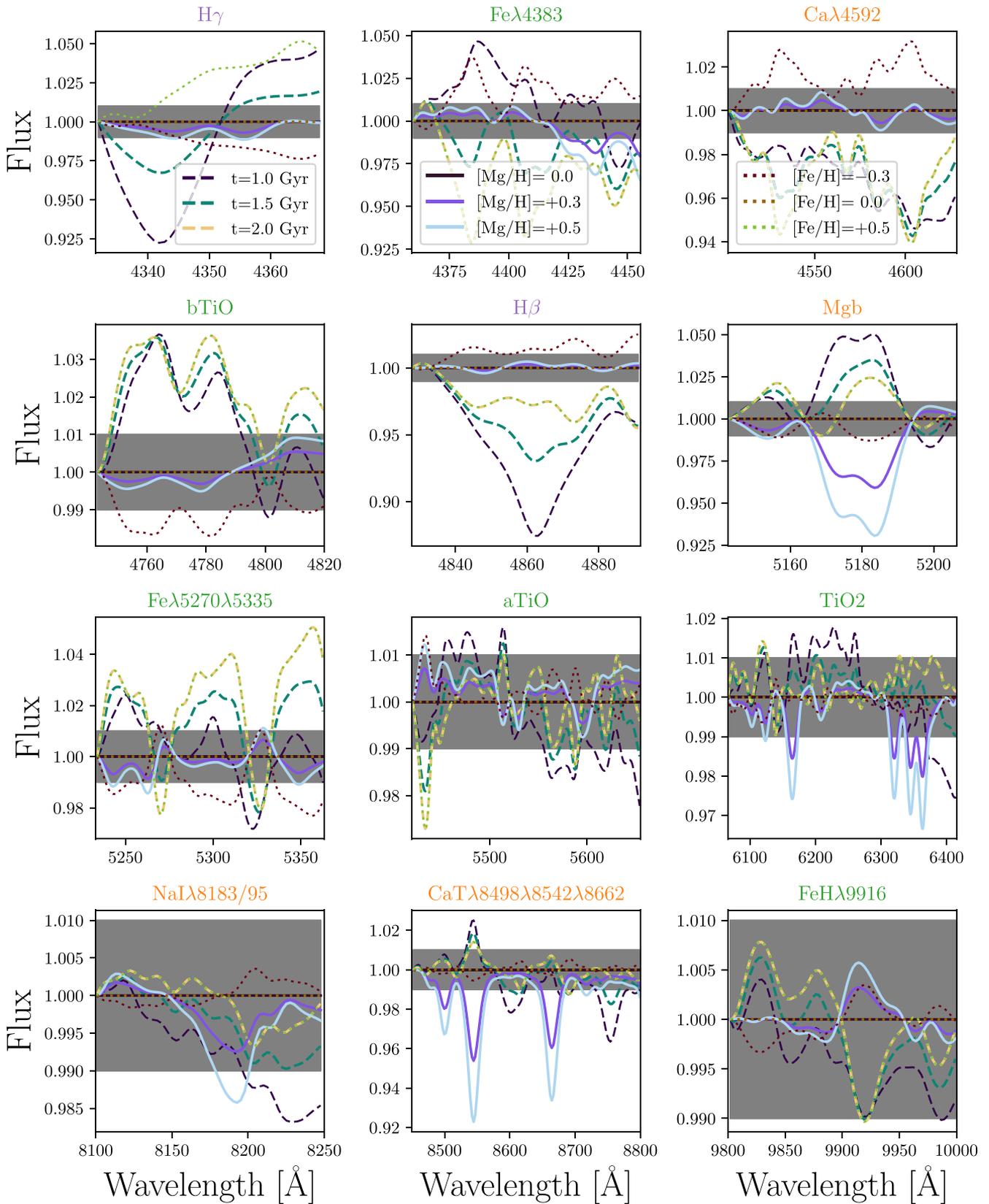


Figure 5. Here we show Villaume et al. (2017) empirical SSPs computed at different ages (dashed), α -abundances (solid), and metallicities (dotted) using `a1f`. Spectra are smoothed to a resolution of 100 km s^{-1} and are divided by a 1.5 Gyr old solar abundance spectrum, so relative changes in the spectra can be clearly identified. The grey-shaded region shows the relative accuracy that is obtained by a $S/N=100$ spectrum. It is evident that age through the Balmer absorption lines, α -abundances through Mgb, and metallicity ($[Fe/H]$) through Fe features can be recovered using these individual absorption features at this S/N level.

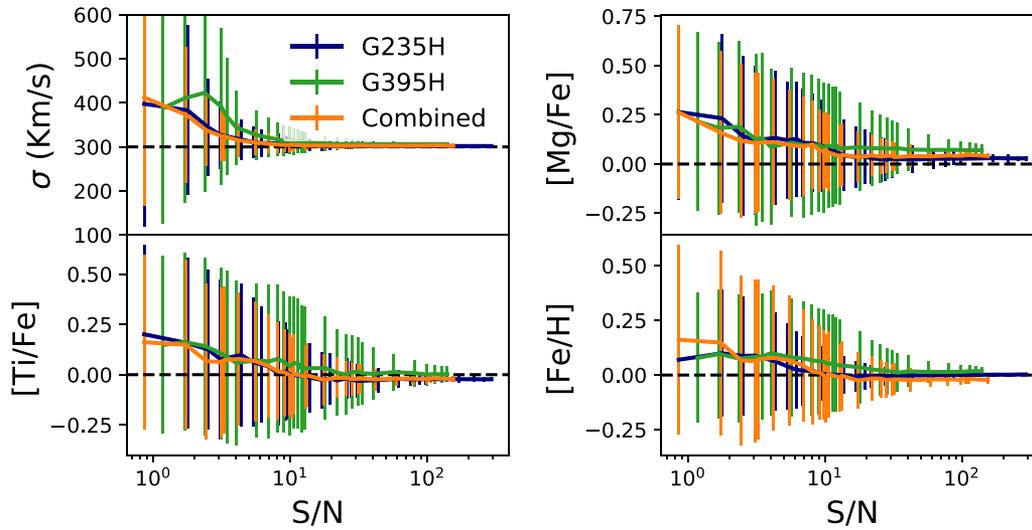


Figure 6. The recovery of top left: velocity dispersion and elemental abundances of top right: [Mg/Fe], lower left: [Ti/Fe], lower right: [Fe/H] of mock JWST NIRSpec S200A1 G235H/FL170LP and G395H/FL290LP observations. Full spectral fitting is performed using `a1f` for individual grism/filter combinations separately and together at their respective native grism resolutions. The input values to the model spectra are shown by the horizontal dashed lines.

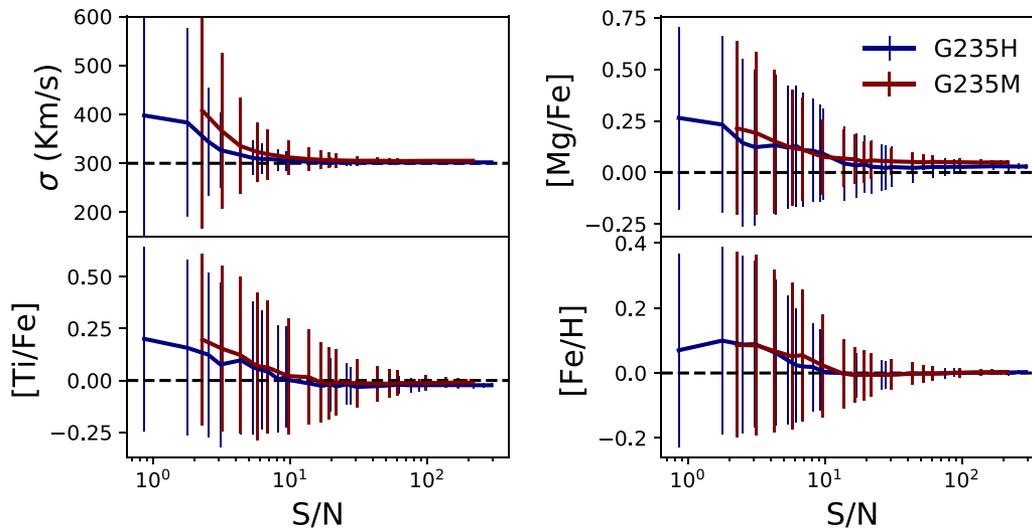


Figure 7. The recovery of top left: velocity dispersion and elemental abundances of top right: [Mg/Fe], lower left: [Ti/Fe], lower right: [Fe/H] from NIRSpec G235M/FL170LP and G235H/FL170LP grism/filter combinations using `a1f` at their respective native grism resolutions. The true value is shown by the horizontal dashed lines. It is evident that even with the lower resolution G235M grism, the input parameters can be accurately recovered at $S/N \gtrsim 30$. For the same exposure time, the S/N can increase by a factor of $\gtrsim 2$ between G235H and G235M grisms; therefore, obtaining absorption line spectroscopy using the G235M will be the most efficient to derive velocity dispersions and element abundances.

(σ) of 300 km s^{-1} . We then use the JWST ETC to generate mock observed spectra at $z = 3.2$ using the NIRSpec G235M/FL170LP grism/filter combination. The galaxy spectra are normalised to a stellar mass of $10^{11} M_{\odot}$ at 1 Gyr, and the exposure parameters are varied such that a continuum S/N of ~ 30 is achieved for the spectra.

In Figure 8, we show the recovery of element abundances for the 4 FSPS mock observations. The [Mg/Fe] abundance of all four galaxies are recovered well by the `a1f` fits to our simulations. Similarly, except for Model B, the [Ti/Fe] of the other SFHs is also well recovered. However, the [Fe/H] abundance cannot be recovered for any of the models. The [Fe/H] of the single burst

(Model A) is underestimated by ~ 0.08 dex, while the [Fe/H] of the other models is overestimated by ~ 0.1 dex. Therefore, despite the fact that Villaume *et al.* (2017) SSPs showed good convergence of [Fe/H] for NIRSpec G235M/FL170LP $S/N \sim 30$ observations, we find that FSPS models show a slight offset. We further run FSPS models up to $S/N \sim 100$ and find that the discrepancy cannot be resolved by increasing the S/N of the observations and/or by a different grism/filter combination offered by NIRSpec. The underlying stellar libraries and isochrones used by `a1f` to fit the mock galaxy spectra are largely similar to the ones used to create the mock observables at $\lesssim 0.7 \mu\text{m}$ ($\lesssim 2.9 \mu\text{m}$ in observed space at $z = 3.2$). Thus, we speculate that the offsets we observe in Figure 8

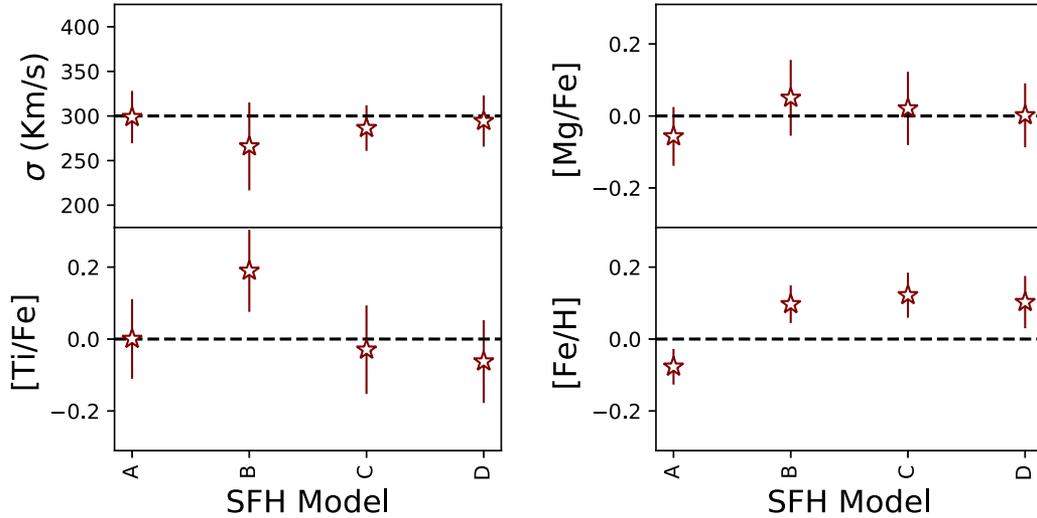


Figure 8. The recovery of top left: velocity dispersion and elemental abundances of top right: [Mg/Fe], lower left: [Ti/Fe], lower right: [Fe/H] from NIRSPEC G235M/FL170LP grism/filter combination using a1f. The models are generated using FSPS using different SFHs as detailed in Table 1. It is clear that with the exception of [Fe/H], other parameters are recovered accurately within the error limits for most SFHs.

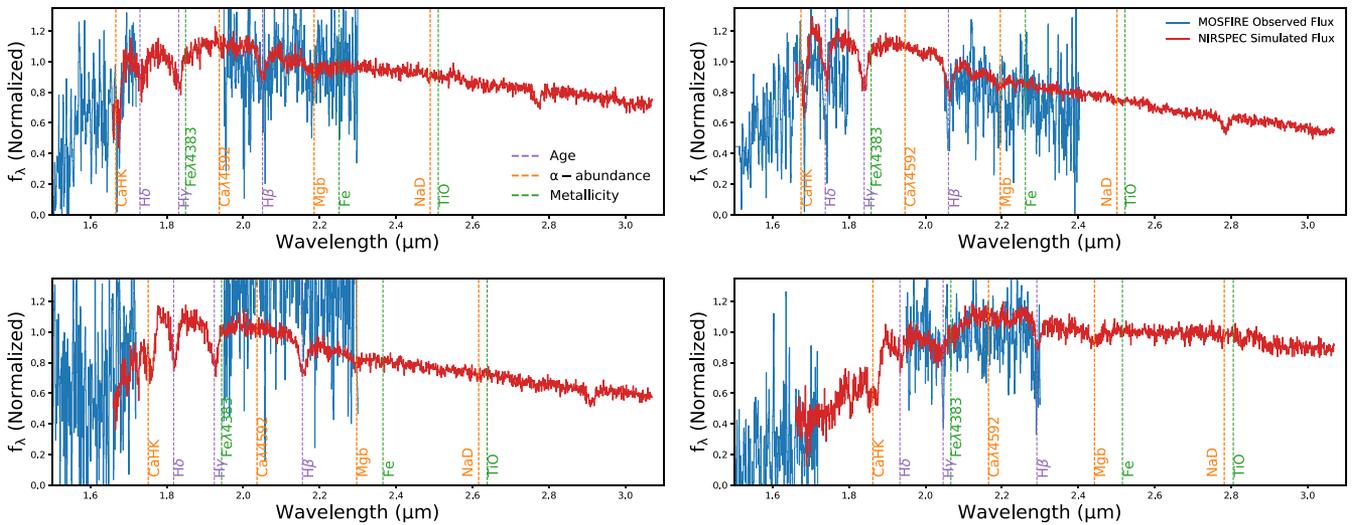


Figure 9. Simulated JWST NIRSPEC G235M/FL170LP observations of the four $z \sim 3-4$ quiescent galaxies presented by Esdaile 2021b). The best fit FAST++ templates to the galaxies from Schreiber et al. (2018b) are used to obtain the JWST mock observations. We show top left: 3D-EGS-40032 (~ 8 h), top right: 3D-EGS-18996 (~ 4 h), lower left: 3D-EGS-31322 (~ 4 h), and lower right: ZF-COSMOS-20115 (~ 5 h) where the time stated inside the brackets refers to the typical NIRSPEC G235M/FL170LP observing times necessary to obtain a continuum S/N of $\sim 30-40$. Ground-based H and K band Keck/MOSFIRE spectra from Schreiber et al. (2018b) are also shown for comparison. With < 10 h of exposure time, JWST/NIRSPEC can obtain > 7 times greater S/N quality at similar velocity resolutions compared to the current ground-based data for such targets and provides continuous coverage of features through atmospheric windows which are essential to analyse the stellar population properties of these galaxies. These features are colour coded according to their primary sensitivity to age, α -abundance, and metallicity.

are due to effects such as dust and complex SFHs (except for Model A) in FSPS which are not accounted for in a1f. Further detailed analysis on the role of galaxy properties in the recoveries is warranted but is out of the scope of this paper.

In Figure 9, we show the simulated JWST/NIRSPEC G235M/F170LP observations of the four $z \sim 3-4$ quiescent galaxies from the Esdaile 2021b) sample. These galaxies have per pixel S/N $\sim 5-10$ in ground-based spectra with exposure times up to ~ 15 h with Keck/MOSFIRE. Following morphological properties derived by Esdaile 2021b), we reconstruct each source

on the JWST ETC based on the HST WFC3/F160W imaging and use the best-fit FAST++^c spectra from Schreiber et al. (2018b) as the source spectra.

Following our aforementioned experiments with a1f, we alter the exposure times to obtain a continuum S/N $\sim 30-40$, to maximise the accuracy of the recovered spectral parameters. With exposure times ranging between 4 and 8 h, JWST/NIRSPEC

^c<https://github.com/cschreib/fastpp>.

G235M/F170LP provide $\gtrsim 7$ times better S/N at similar velocity resolutions compared to ground-based Keck/MOSFIRE spectra. It is also clear from the figure that space-based observations from JWST allow a continuous coverage of most crucial rest-frame optical spectral signatures for the $z \sim 3-4$ quiescent galaxy populations. These range from age, α -element abundance, and overall metallicity sensitive indicators as shown in Figure 5. Thus, rest-frame optical spectra of $z \sim 3-4$ quiescent galaxies with similar properties to the Esdaile (2021b) sample can be obtained from JWST NIRSpec with modest exposure times. Such high-quality observations of quiescent galaxies are crucial to constrain their formation and evolution properties with high confidence.

3.3 Star formation history

There is significant tension regarding the existence and abundance of massive quiescent galaxies in the $z > 3$ universe between observations and cosmological simulation models (e.g. Merlin *et al.* 2019). Thus, to understand the formation and subsequent quenching mechanisms of the early Universe, determining how massive quiescent galaxies built their mass efficiently is of great importance.

The SFH of massive quiescent galaxies at $z \sim 3-5$ could be addressed via two different methods. Firstly, through analysis of individual line strengths such as Balmer and D4000 features, the approximate duration of the most recent SFH can be constrained (Vazdekis *et al.* 2015). Additionally, if an enhancement of α -elements such as Mg, Ca, and Na is observed, this would provide indications that they were formed over a short starburst event due to the delay time of Type Ia supernovae (e.g. Kriek *et al.* 2016).

Secondly, more stringent constraints to the SFHs can be made by combining full-spectrum fitting techniques with multi-wavelength photometric fitting techniques. However, even full spectrophotometric fitting techniques have degeneracies that result in a factor of ~ 2 uncertainty in derived cosmological parameters, that is, the cosmic SFR density (e.g. Madau & Dickinson 2014; Yu & Wang 2016). This is a result of multiple levels of complex degeneracies and uncertainties inherent to stellar populations models (Conroy 2013) and SED fitting techniques (Walcher *et al.* 2011). If only photometry is used in SED fitting, inferences are based purely on the SED shape. This can result in systematic offsets in derived galaxy properties due to effects such as the age-dust-metallicity degeneracy that affects the shape of the SED (Bell & de Jong 2001).

When spectra are combined with photometry, stronger inferences can be made on galaxy properties. This is due to extra information that can be obtained from the spectral features. However, uncertainties in stellar evolution, limitations in empirical stellar libraries, and limitations in theoretical stellar atmosphere models can introduce systematic biases to these inferences (e.g. Bruzual 2007). The SFH of a galaxy will have imprints in the overall shape of the SED and on the strength of different elements observed in the spectra. Additionally, the presence of specific types of stars leaves wider spectral imprints which can be used to provide further constraints (e.g. Brinchmann, Kunth, & Durret 2008). However, simultaneous modelling of a variety of spectral features that consider variations in stellar types, stellar/ISM abundances, and IMFs is complicated (Gunawardhana *et al.* 2020). When this is combined with non-parametric SFHs to model the formation history,

high-quality data and modular SED fitting codes become crucial (e.g. Leja *et al.* 2017, 2019).

Next, we perform mock observations to investigate whether spectrophotometric fitting of galaxies could recover the SFHs in sufficient accuracy to distinguish between different mass growth scenarios. Given the Universe is ~ 2 billion years old at $z \sim 3.2$, SFH variations should be evaluated by the SED fitting codes within this short time frame. We use the FSPS-derived model galaxies with four different SFHs to generate the mock JWST NIRSpec spectra as detailed in Section 3. Observed photometry for the galaxies is computed using *sedpy* python package mirroring the full wavelength coverage of the ZFOURGE COSMOS field (Straatman *et al.* 2016). For simplicity, we assume that all photometric bands are detected at a S/N of 10.

Utilising the mock photometry and deep NIRSpec rest-frame optical spectra with full-spectrum fitting techniques of PROSPECTOR (Johnson *et al.* 2020) and its non-parametric approach to constraining the SFHs (e.g. Leja *et al.* 2019), we investigate whether precise constraints could be placed on the formation timescale of massive quiescent galaxies. In PROSPECTOR fitting, we allow stellar mass, stellar metallicity, and dust attenuation to vary as free parameters. Additionally, a *continuity_sfh* prior (see Section 2.2.3 of Leja *et al.* 2019) is used for the SFH. In lookback time, the first two bins are fixed to be between 0–30 Myr and 30–100 Myr. The most distant bin is fixed to be between $0.85 \times \text{age}$ of the Universe and the age of the Universe; thus, at $z = 3.2$, the final bin is between 1700 and 2000 Myr in lookback time. The remaining bins are split in logarithmic time evenly between 100 and 1700 Myr. In total, we use between 4 and 7 bins to investigate the SFH recovery for our four models.

In Figure 10, we show the recovered SFH for our four models. For Model A, the form of the SFH is recovered well by PROSPECTOR independent of the number of SFH bins investigated here. All recoveries suggest that the bulk of the star formation happened around 1 Gyr and that the current SFR is very minimal. This is further confirmed by the highest time resolution 7-bin SFH showing a clear peak at ~ 1 Gyr, with the SFR in the immediately preceding and following bins reducing by a $\gtrsim 10^{-3}$ and $\gtrsim 10^{-8}$, respectively.

For Model B, the increase in the SFR and the subsequent truncation of the SFR are well captured when the SFH is modelled using either 5 or 7 bins. The 4 and 6 bin SFHs fail to recognise the exponential increase in the SFR at earlier times. This is expected for the 4-bin SFH, because the 3rd time-bin by construct covers a large 100–1700 Myr time window. The galaxy only forms stars for 500 Myr in this window, which may result in the SFH sensitivity of this bin to reduce. However, the 6-bin SFH should have had sufficient sensitivity to recover the increase in SFR. The lack of star formation in the later stages of the galaxy is well recovered irrespective of the number of bins.

For Model C, the input SFH shown by the black dashed line is recovered well by all the fits. 6 and 7 bin SFHs provide the most accurate recovery of the burst; however, even with the limited time resolution of 4 and 5 SFH bin fits, the burst is still captured by the PROSPECTOR fits. The 7 and 5 bin SFHs bins also show that the SFR was declining before the galaxy underwent a burst. This, however, is not captured by the SFH with 6 bins even though it has finer time sampling compared to the SFH with 5 bins. It is interesting to note that at $z \sim 3.2$, even with 7 time-bins, the input SFH does not show the exponentially declining nature of the SFH in the post-burst phase (compared to Figure 4 Model C). Thus,

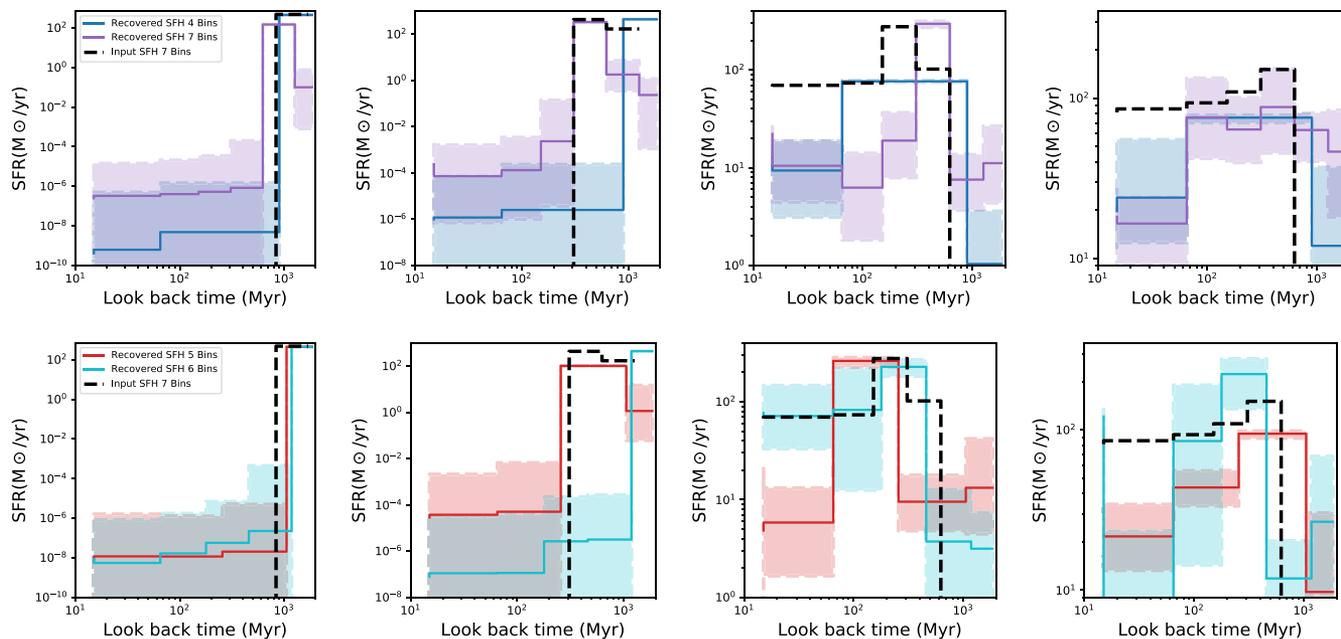


Figure 10. The recovery of SFHs based on full spectral + photometry fitting by the PROSPECTOR SED fitting code. Panels show the recovery of Models First Column: A Second Column: B Third Column: C Fourth Column: D as described in Table 1. All models are fit with varying numbers of non-parametric bins ranging from 4 to 7 bins. The top row shows the recovery for 4 and 7 bins, and the bottom row shows the recovery for 5 and 6 bins. The input SFH is shown as a black dashed line using 7 time-bins defined similarly to the 7 bin PROSPECTOR fit. The 16th and 84th percentiles are shown by the colour shading for each recovered SFH. For all models, PROSPECTOR is able to recover the overall shape of the input SFH. However, some fine tuning of the number of SFH bins is required to accurately recover the SFHs.

constraining the full form of the SFH would require a much larger number of time-bins. Additionally, only the 6-bin SFH is able to detect the ongoing star formation of this model at present time.

For Model D, the overall shape of the SFH is captured well by the fits. The 6-bin SFH suggests the galaxy to have a younger ~ 500 Myr component and an older ~ 2 Gyr component. The 7-bin SFH also suggests that a significant fraction of the galaxy has been formed at an older age compared to the input model. None of the recoveries are able to capture the residual star formation of this model at present time.

When comparing the recoveries between the models, it is evident that PROSPECTOR struggles to recover the ongoing star formation for Models C and D (with the exception of the 6-bin SFH recovery for Model C). Therefore, even with state-of-the-art SED fitting codes like PROSPECTOR, it is plausible that galaxies with ongoing residual star formation could be identified as quiescent when a full spectrophotometric fitting is performed. Analysis of optical rest-frame nebular optical emission lines from JWST spectroscopy could be used to provide constraints on the star formation limits of such galaxies. However, low emission lines such as [OII] have been observed in lower redshift quiescent galaxies (e.g. Maseda et al. 2021), which adds extra complexity. Dust continuum detections in the sub-mm would provide an independent constraint to the SFRs in such galaxies (e.g. Simpson et al. 2017; Schreiber et al. 2018a).

Albeit limitations in recovering residual star formation at later times, in general, our simulations with different SFHs and the different number of SFH bins demonstrate that the SED fitting techniques (investigated here in the context of PROSPECTOR) can recover the past form of the SFH to distinguish between different formation scenarios (extended formation vs short sharp

formation). We note that our simulations here are tied closely to investigate the recovery of parameters for galaxies that are of similar properties to the Esdaile 2021b sample. In this context, in order to recover the input SFHs accurately, some tuning of the number of SFH bins is required. The differences between bins are likely due to the sensitivity of certain photometric/spectral features to the SFH which are enhanced with finer/coarse time sampling. However, these should be well tested with simulated models for galaxies with short evolutionary times (age $\lesssim 3$ Gyr) and compared with other recovered galaxy properties such as stellar mass and metallicity. We defer this to a future analysis. We also note that the rest-frame optical includes both age and abundance-sensitive lines; thus, both the SFH and metallicity can be constrained to high accuracy using PROSPECTOR with fine-tuned binning. Additionally, when direct α -element constraints are lacking, metallicity along with strong constraints on the SFH can be used to infer the α -abundances.

3.4 The IMF

The stellar IMF is a fundamental parameter in galaxy evolution that plays a vital role in regulating galaxy formation and chemical evolution of the Universe. Even though traditionally the IMF has been considered to be universal (e.g. Salpeter 1955; Kroupa 2001; Baldry & Glazebrook 2003; Chabrier 2003), recent results have started to show evidence for systemic variations (e.g. Bastian, Covey, & Meyer 2010; Hopkins 2018). At $z \sim 0$, studies have shown that the lower mass of the IMF slope could vary as a function of velocity dispersion (e.g. Cappellari et al. 2012; Conroy et al. 2013), metallicity (e.g. Martín-Navarro et al. 2015b), and radial distance from the galactic centre (e.g. Martín-Navarro

et al. 2015a) in local early type galaxies (ETGs). However, mechanisms for such variations are still not well understood and overall agreement between different IMF indicators are still not clear (e.g. Smith 2014, 2020). Constraining the IMF in the early Universe is vital to determine how such variations may have been influenced by conditions prevalent in the Universe when the local ETGs were building up their stellar masses. Constraints on the early Universe IMF are also necessary to provide constraints to the cosmic SFH and to understand how the Universe transformed from hydrogen and helium to the current complex Universe.

The IMF of the $z \sim 3-5$ massive quiescent galaxies can be constrained using high-precision dynamical and stellar mass measurements. NIRSpc slit spectroscopy or NIRSpc IFU observations (see Section 4) can be used to constrain the dynamical masses of the galaxies. Given these early systems are compact and dense ($M_{\text{ast}} \sim 10^{10} - 10^{11}$, $r_e \sim 0.5 - 2$ kpc Straatman *et al.* 2014; Esdaile 2021b), the baryonic matter is expected to dominate the kinematics; thus, dynamical mass measurements are largely independent of the assumed dark matter halo profiles. This means that the dynamical mass is only contributed to by the stellar mass, and thus, the ratio of the dynamical mass to stellar mass (also known as the IMF mismatch parameter (e.g. Davis & McDermid 2017) can be used to understand the stellar populations via the IMF. JWST NIRSpc spectroscopy will also provide constraints to the rest-frame optical spectral shape of the $z \sim 3-5$ massive quiescent galaxies, which will provide strong constraints to stellar masses (Conroy 2013). Therefore, with JWST spectroscopy, stronger constraints on the IMF mismatch parameter can be obtained.

The IMF mismatch parameter can be compared with the distribution of the $z \sim 3-5$ galaxies in the stellar mass plane to see if there are differences in the inherent stellar populations of these galaxies, perhaps indicating different formation scenarios. Given quiescent galaxies by definition do not have active star formation, the ‘observed’ dynamical mass will only constrain the lower mass slope of the IMF (but is degenerated with high mass stellar remnants). This is due to the lack of massive O, B type stars, which have relatively short lifetimes of 10–100 Myr compared to the long-lived lower mass stars.

The IMF mismatch parameter is the only practical way to measure the IMF of unlensed individual massive quiescent galaxies at $z \sim 3-5$. As we showed in Figure 5, IMF-sensitive spectral features such as the Na I doublet and the Wing–Ford band resulting from K and M dwarf stars require $S/N > 300$ to distinguish between different IMF slopes (van Dokkum & Conroy 2010). Our NIRSpc simulations with a1f show that a $S/N \sim 350$ is required in the G395H/FL290LP grism/filter combination to recover the IMF slopes and reaching such levels even for the brightest $z \sim 3-5$ massive quiescent galaxies is not feasible due to the very high exposure times required.

Apart from the ‘observed’ IMF, the *relic* IMF of $z \sim 3-5$ massive quiescent galaxies can be reconstructed through spectrophotometric fitting of advanced SED fitting codes such as PROSPECTOR (private communication) and variable IMF chemical evolution codes such as galIMF (Yan *et al.* 2019). By providing IMF as an extra free parameter in SED fitting, the evolution of the IMF in the $z > 5$ Universe can be constrained. However, the time variation of the IMF, SFH, and α -element abundances is intricately connected (e.g. Martín-Navarro 2016); therefore, detailed stellar population modelling is required to understand degeneracies between them. $z > 5$ Universe is the time window when the $z \sim 3-5$ massive quiescent galaxies were forming the bulk of their

stars (e.g. Valentino *et al.* 2020). If these galaxies did indeed have top-heavy IMFs in their star formation phase (e.g. as formulated by Lacey *et al.* 2016, also see Gunawardhana *et al.* 2011; Nanayakkara *et al.* 2017, 2020), this would lead to differences in the number of ionising photons produced, the chemical enrichment, and stellar wind and supernovae feedback processes. Such changes would be crucial to the reionisation timescales of the Universe; thus, the reconstruction of the *relic* IMF in the $z > 6$ Universe is important to constrain the cosmological evolutionary models of the Universe.

4. Dynamical properties of galaxies with JWST

The dynamical properties of massive $z \sim 3-5$ quiescent galaxies are an important quantity to constrain the evolutionary properties of massive galaxies in the early Universe. Through JWST slit or IFU spectroscopy, velocity dispersions, dynamical masses, stellar masses, sizes, and the evolution of the mass–size–velocity dispersion plane of quiescent galaxies at $z \sim 3-5$ can be constrained. Analysing the $z \sim 3-5$ galaxies in this plane and comparing them with galaxies at lower redshifts are necessary to build up the cosmic evolutionary picture of quiescent galaxies and to determine how they evolve in the mass–size plane with cosmic time (e.g. Belli, Newman, & Ellis 2017a). Constraining kinematic properties is also important to determine whether these massive quiescent galaxies have undergone mergers in their evolutionary history (e.g. Belli *et al.* 2017b), which will shed further light into mass assembly and quenching processes in the early Universe.

4.1 Velocity dispersions through slit spectroscopy

JWST NIRSpc slit spectroscopy can be used to measure the velocity dispersions of $z \sim 3-5$ massive quiescent galaxies. In Figures 6 and 7, we show the recovery of the velocity dispersions of our mock NIRSpc S200A1 simulations. It is clear that with both the grisms and filters, the velocity dispersion is converged to the input value of the mock simulations. In Section 3.2, we established that G235M/FL170LP grism/filter combination provides the best wavelength coverage and the necessary resolution to obtain the element abundances of $z \sim 3-5$ galaxies.

In terms of velocity dispersion, the G235M medium-resolution disperser of NIRSpc can obtain dispersions of $\sim 130 \text{ km s}^{-1}$ for a 2-pixel resolution element while the high-resolution G235H disperser achieves a dispersion resolution of $\sim 50 \text{ km s}^{-1}$. Given the lowest velocity dispersion observed for $z > 3$ quiescent galaxies is $\sim 15 \text{ km s}^{-1}$ (Esdaile 2021b), the medium resolution grating is sufficient to obtain velocity dispersions of quiescent galaxies. We also perform mock JWST observations of a galaxy with $\sim 150 \text{ km s}^{-1}$ velocity dispersion and find that at a S/N of 30, the velocity dispersion can be recovered by the G235M/FL170LP grism/filter combination. Additionally, we find that the delivered S/N from G235M is higher than the S/N of G235H when binned to the same spectral resolution.

By obtaining a S/N of ~ 30 per pixel with JWST NIRSpc G235M/FL170LP observations, robust constraints can be obtained to the dynamical masses. For example, Esdaile (2021b) measured dynamical masses for four $z \sim 3-5$ massive quiescent galaxies to an accuracy of $\sim 20\%$. This was based on measurements of Balmer and Ca H & K absorption lines covering a total of $\sim 600 \text{ \AA}$ in rest-frame with a continuum S/N of $\sim 5-7$ achieved through spectral binning of 6 \AA in the rest-frame. With NIRSpc,

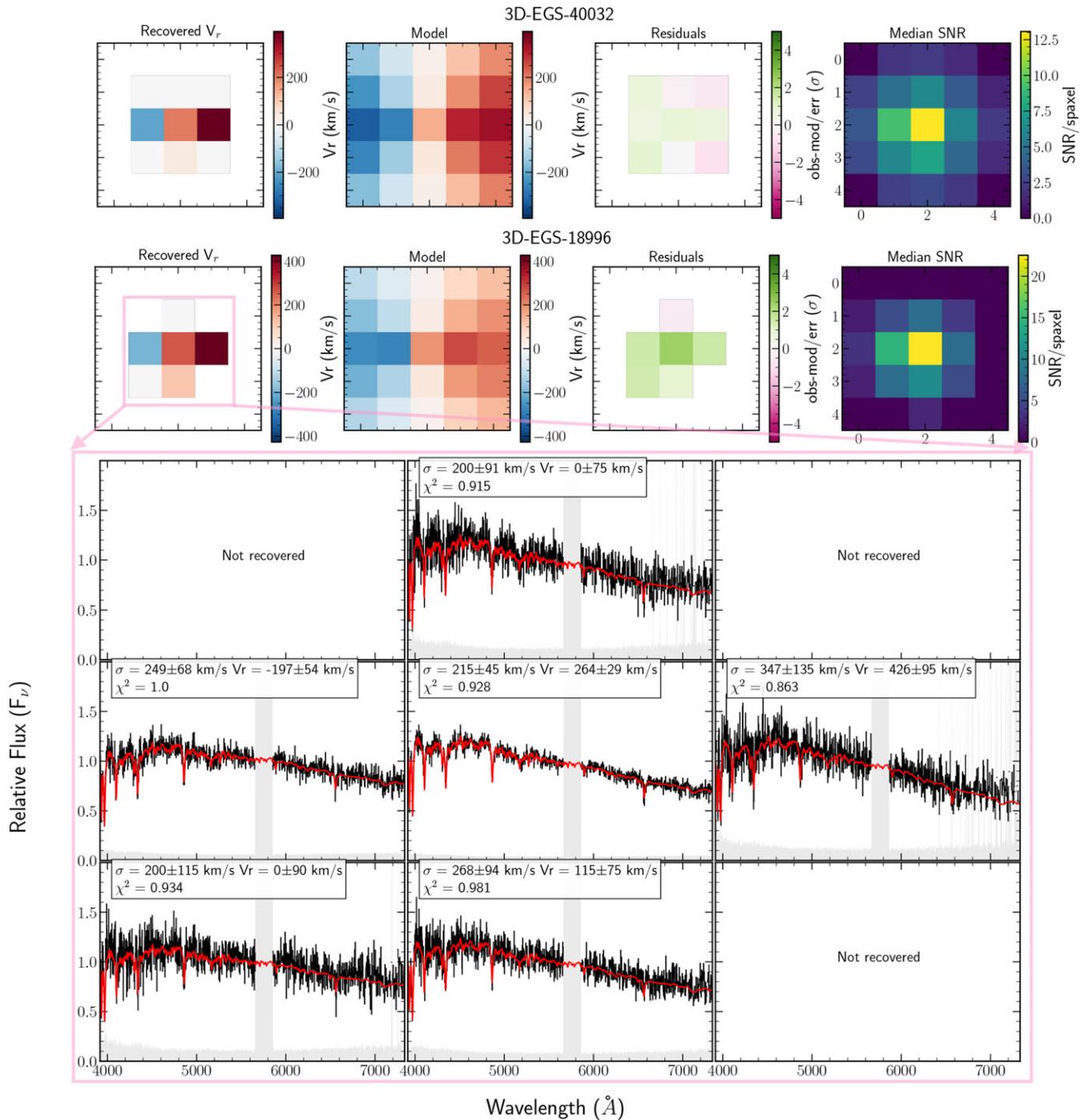


Figure 11. Simulated IFU observations for 2D kinematics recovery of two galaxies in the Esdaile 2021b) sample: 3D-EGS-40032 (G140M/F100LP grism/filter, 12 h exposure) and 3D-EGS-18996 (G235H/F170LP grism/filter, 8 h exposure). The chosen grism/filter combination optimises the spectral ranges (which includes [OII] emission from 3D-EGS-40032), sensitivity, and resolving power to measure existing velocity dispersions for each galaxy. Top: from left to right: recovered 2D kinematics from simulated IFU cube (left) for a Hernquist 2D model of a rotationally supported galaxy and with an inclination of 45° (middle-left), the normalised residuals (middle-right) and median S/N/spaxel from the IFU cube. The simulation for 3D-EGS-40032 is in the top row with a $V_r/\sigma = 1.5$ and 3D-EGS-18996 is below with a $V_r/\sigma = 1.8$. Bottom: Recovered radial velocities and velocity dispersions for 3D-EGS-18996 per inner set of spaxels in the simulated IFU observations cube. Black lines are the simulated spectra, red lines are the best-fit, and grey-shaded area is the corresponding noise spectra. The spectral fit is only included for $0.85 < \chi^2 < 1.5$

our simulations show that high-quality spectra of such galaxies can be obtained with typical S/N of $\sim 30\text{--}40$ per pixel at similar resolution with continuous coverage of $\sim 3\,000$ \AA in the rest-frame. They have no skyline contamination and reach an

accuracy of $\sim 2\text{--}3\%$ for dynamical mass measurements. This higher level of accuracy will enable robust comparisons with mass-matched control samples of galaxies at $z \sim 2$, which can also be obtained as filler targets using the JWST NIRSpec multi-object

spectroscopy (MOS) which utilises the NIRSpec micro-shutter assembly (MSA).

4.2 Galaxy kinematics through IFU spectroscopy

JWST NIRSpec IFU spectroscopy can be used to obtain resolved kinematics of $z \sim 3-5$ massive quiescent galaxies. Compared to ground-based IFU spectroscopy, the JWST NIRSpec IFU provides many advantages for kinematics analysis. Higher spatial resolution can be obtained due to diffraction-limited space-based observing at $2 \mu\text{m}$ and is not limited by the lack of bright stars necessary for adaptive optics corrections to increase the seeing. Generally, non-AO observations are preferred for the analysis of continuum sources with absorption line kinematics. This is because, in kinematics, the AO PSF imposes an interdependency between luminosity, velocity, and dispersion; thus, more advanced deconvolution is necessary for each spectral plane of the original data cube (Davies & Kasper 2012). This can be corrected by robustly using emission line features (after accounting for beam smearing); however, for continuum features a PSF convolved step-wise model for every pixel in the continuum and equivalent width would be required. This is hard to achieve even for bright sources (Thater *et al.* 2019). Additionally, the desired S/N is practically impossible to achieve even with the best ground-based IFUs such as Keck/OSIRIS, where the majority of the absorption features for $z \sim 3$ galaxies exist in the $H-$ band where the Strehl ratio is low ($\sim 10\%$). For example, for the two mock NIRSpec IFU simulations of two of the Esdaile 2021b) galaxies which we describe below, the total exposure time necessary to obtain resolved kinematics increases from ~ 8 to $> 40-100$ h (based on the surface brightness of the galaxy) in ground-based Keck/OSIRIS spectroscopy.

In order to investigate whether JWST NIRSpec IFU observations could recover the kinematics of $z \sim 3-5$ massive quiescent galaxies accurately, we perform full 2D mock IFU simulations to investigate the recovery of spatially resolved velocities with pPXF (Cappellari 2017). We use empirical SSPs from Villaume *et al.* (2017) to generate mock 1.5 Gyr old galaxies with velocity dispersions consistent with the measurements in Esdaile 2021b). These model spectra are input into the JWST ETC, including their physical size and properties determined from modelling using *HST* F160W band images. Several mock IFU cubes were then generated to obtain a suite of S/N values. Hernquist rotational models (Hernquist 1990) were then applied to each simulated IFU observation, assuming a 45° inclination, with various rotational-to-velocity dispersion ratios (V_r/σ) to simulate galaxies with rotational support. A Hernquist model is chosen as suitable for modelling a rotating spheroidal star system; the estimated mass and the scale length (taken to be the effective radius of the galaxy) produce the velocity profile. The Plummer model (Plummer 1911) produces similar rotational curves. Finally, we used pPXF on the instrument calibrated data to fit for radial velocity and velocity dispersion.

In Figure 11, we show the recovery of both 3D-EGS-18996 and 3D-EGS-40032 2D kinematics. We note that the JWST ETC outputs $0.1''$ spaxels which do not account for drizzling; however, we expect in practice that drizzling will improve the sampling compared to what is shown here. Given the compact sizes and sharp decrease in S/N per spaxel, binning to this IFU spaxel scale is likely optimal for signal, while still allowing sufficient spatial scale to

infer rotation. To assess the S/N requirements for the mock observations, this analysis was repeated for different exposure times. We found that a median S/N ~ 7 per spaxel is required in the adjacent to central pixels to recover a radial velocity and infer rotation with $> 2\sigma$ confidence. At this S/N, V_r/σ can be constrained to $V_r/\sigma \sim 1.5$ in 3D-EGS-18996 and $V_r/\sigma \sim 1.3$ in 3D-EGS-40032 and confirm rotational support in these galaxies. Additionally, this would limit the underestimation in dynamical mass from 0.4 dex (for a $V_r/\sigma \sim 3$) to $\sim 0.1-0.15$ dex (from Equation (5) Belli *et al.* 2017a) in both galaxies and allow for high-precision IMF constraints without potential large systematic errors.

5. Discussion and conclusions

Through a case study of the Esdaile 2021b) sample, we have shown here that the launch of JWST will address crucial observational challenges that currently limit our understanding of the abundance, formation, and evolutionary mechanisms of $z \sim 3-5$ quiescent galaxies. Given these galaxies provide insights into galaxy evolution in the $z > 6$ Universe, the detailed analysis would determine how the first generation of galaxies in the Universe formed and evolved. Current cosmological models find it challenging to match the observed abundance of massive quiescent galaxies in the $z \sim 3-5$ Universe (e.g. Merlin *et al.* 2019; Valentino *et al.* 2020); thus, the analysis of the stellar populations and galaxy kinematics is vital to reconstruct the star formation and merger history of these galaxies.

The development of medium-band imaging instruments has allowed surveys like ZFOURGE (Straatman *et al.* 2016) and FENIKS (Esdaile *et al.* 2021a) to be carried out where the photometric redshifts can be obtained with high accuracy. This is driven by the increased spectral sampling of the split H , J , and K bands (e.g. Nanayakkara *et al.* 2016). JWST NIRSpec CLEAR spectroscopy provides the most prominent avenue to probe the completeness of these photometrically selected massive quiescent galaxies in the early Universe. This is driven by the multiplexing capability of NIRSpec along with the simultaneous continuous wavelength coverage between 0.6 and $5.3 \mu\text{m}$. The large wavelength coverage is well suited to identify star-forming contaminants in $z \sim 3-5$ quiescent candidates. In Figure 3, we showed that the resolution and the sensitivity of NIRSpec PRISM/CLEAR spectroscopy were sufficient to accurately distinguish between quiescent, post-starburst, and dusty star-forming galaxies. However, the spectral resolution of PRISM/CLEAR spectroscopy is not sufficient to perform a detailed analysis of element abundances to constrain the stellar population properties of $z \sim 3-5$ massive quiescent galaxies.

The optimal strategy to obtain spectral features to perform a detailed analysis of stellar populations is to obtain deeper high-resolution JWST NIRSpec spectroscopy for samples that are already confirmed to be quiescent and are at a redshift range of interest. Using simulated spectra from Villaume *et al.* (2017), we generated mock NIRSpec observations for G235M/FL170LP, G235H/FL170LP, and G395H/FL290LP grism/filter combinations to identify the best grism/filter combinations and the S/N levels required to recover accurate element abundances and velocity dispersions. We found that a S/N of ~ 30 is required to reach the necessary accuracy levels for element recoveries and that the G235M/FL170LP grism/filter combination is optimally suited to carry out the necessary observations. We found that the lower resolution of the G235M enables more efficient observations by

achieving the necessary S/N level without compromising the accuracy, even when the higher resolution G235H grism is binned to the same velocity. Further analysis with parametric SFHs from FSPS (Conroy & Gunn 2010) showed that the G235M/FL170LP grism/filter combination could still largely recover element abundances and velocity dispersions. However, we found small offsets between input and recovered [Fe/H] abundances for these parametric SFH models.

The G235M/FL170LP grism/filter combination is ideal to obtain coverage of rest-frame optical features of the $z \sim 3-5$ massive quiescent candidates. As an example, we showed that Esdaile (2021b) sample will cover the crucial age, α -abundance, and metallicity sensitive indicators, all within < 8 h of exposure time. Compared to current best ground-based spectroscopy, the improvement on S/N from JWST NIRSpec was found to be $7\times$ at similar velocity resolutions. As we showed in Figure 5, the Balmer absorption features are most sensitive to the ages, Mgb is strongly sensitive to the α abundance, and the overall metallicity is well constrained from the various Fe features that fall within the G235M/FL170LP wavelength coverage. With combined full-spectrum fitting with software like `a1f`, the stellar population properties can be constrained to high accuracy.

In addition to the most recent star formation episode that can be recovered from the Balmer absorption features, novel advanced SED fitting codes such as PROSPECTOR (Johnson et al. 2020) allow the full SFH of galaxies to be recovered. Using mock JWST G235M/FL170LP observations and photometry mapping the ZFOURGE COSMOS field coverage, we showed that spectrophotometric fitting can accurately recover the SFHs of galaxies to distinguish between different formation mechanisms. However, we observed the accuracy of the outcome to be dependent on the number of SFH bins and that PROSPECTOR struggled to recover the ongoing star formation for some of the models. More work on understanding the role of the number and duration of SFH bins in accurately recovering the SFH of galaxies with short formation timescales is warranted.

Future advancements of SED fitting codes will allow IMF to be varied as an extra free parameter (i.e. PROSPECTOR, private communication). This will enable astronomers to perform robust statistical modelling of SFHs with variable IMFs in the early Universe. Recent studies have suggested that star-forming galaxies are likely to have a higher fraction of high mass stars compared to the traditional Salpeter (1955) IMF (e.g. Gunawardhana et al. 2011; Nanayakkara et al. 2017), and recent semi-analytical models have started to implement such changes to their evolutionary models (Lacey et al. 2016). However, adding extra layers of free parameters could lead to larger degeneracies between the parameters. Additionally, with the superior quality NIRSpec data and advanced spectrophotometric fitting codes these extra degeneracies will be further constrained by the confident detections of multiple α -elements, Fe, and the suite of Balmer absorption lines. They will provide stringent constraints to metallicity, element abundances, and SFH timescales.

Apart from the reconstruction of the SFHs, the integral field spectrographic capabilities of JWST/NIRSpec allow the kinematics of the $z \sim 3-5$ massive quiescent galaxies to be explored. In Figure 11, we showed through mock observations that the kinematics of galaxies similar to Esdaile (2021b) sample can be recovered. In addition to constraints to dynamical masses (and inferences about stellar population properties such as IMF), the kinematics of such galaxies provide vital clues to whether these

massive galaxies already show ordered rotation or if they show signatures of mergers. This can be used as an independent method to establish the formation pathways to early massive quiescent galaxies. Thus, future surveys with JWST will be able to determine whether the early massive quiescent galaxies could be formed within the first $\sim 1-2$ Gyr following Λ CDM hierarchical merger cosmology.

With the advancements made in understanding the abundances, properties of the stellar populations, and formation scenarios of $z \sim 3-5$ massive quiescent galaxies through JWST, the next generation of cosmological simulations will be able to address the cosmic puzzle of how such galaxies in the early Universe were so efficient in building up stellar mass within short time frames. This will allow additional constraints on galaxy evolution in the $z > 6$ Universe to be made and thus will provide more stringent constraints on reionisation pathways and timescales of the Universe and the relative contribution of massive galaxies to the epoch of reionisation.

Acknowledgements. We thank Michael Maseda, Marijn Franx, and Joel Leja for helpful discussions. We also thank the staff at the James Webb Space Telescope Help Desk. T. N., K. G., M. D., and C. J. acknowledge support from Australian Research Council Laureate Fellowship FL180100060. This project made use of `astropy` (Astropy Collaboration et al. 2018), `matplotlib` (Hunter 2007), and `pandas`.

References

- Alcalde-Pampliega, B., et al. 2019, *ApJ*, **876**, 135
 Astropy Collaboration, et al. 2018, *AJ*, **156**, 123
 Baldry, I. K., & Glazebrook, K. 2003, *ApJ*, **593**, 258
 Bastian, N., Covey, K. R., & Meyer, M. R. 2010, *ARA&A*, **48**, 339
 Bell, E. F., & de Jong, R. S. 2001, *ApJ*, **550**, 212
 Belli, S., Newman, A. B., & Ellis, R. S. 2017a, *ApJ*, **834**, 18
 Belli, S., et al. 2017b, *ApJ*, **841**, L6
 Birkmann, S. M., et al. 2011, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8150, Cryogenic Optical Systems and Instruments XIII, ed. J. B. Heaney, & E. T. Kvamme, 81500B, doi: 10.1117/12.893896
 Boogaard, L. A., et al. 2021, *ApJ*, **916**, 12
 Brinchmann, J., Kunth, D., & Durret, F. 2008, *A&A*, **485**, 657
 Brinchmann, J., et al. 2017, *A&A*, **608**, A3
 Bruzual, A. G. 2007, in Stellar Populations as Building Blocks of Galaxies, Vol. 241, A. Vazdekis, & R. Peletier, 125 (arXiv:astro-ph/0703052), doi: 10.1017/S1743921307007624
 Burstein, D., Faber, S. M., Gaskell, C. M., & Krumm, N. 1984, *ApJ*, **287**, 586
 Calzetti, D., Armus, L., Bohlin, R. C., Kinney, A. L., Koornneef, J., & Storchi-Bergmann, T. 2000, *ApJ*, **533**, 682
 Cappellari, M. 2017, *MNRAS*, **466**, 798
 Cappellari, M., et al. 2012, *Nat*, **484**, 485
 Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, *ApJ*, **345**, 245
 Carnall, A. C., et al. 2020, *MNRAS*, **496**, 695
 Casali, M., et al. 2007, *A&A*, **467**, 777
 Chabrier, G. 2003, *PASP*, **115**, pp. 763
 Choi, J., Dotter, A., Conroy, C., Cantiello, M., Paxton, B., & Johnson B. D. 2016, *ApJ*, **823**, 102
 Chon, S., Omukai, K., & Schneider, R. 2021, arXiv e-prints, p. arXiv:2103.04997
 Conroy, C. 2013, *ARA&A*, **51**, 393
 Conroy, C., & Gunn, J. E. 2010, *ApJ*, **712**, 833
 Conroy, C., Dutton, A. A., Graves, G. J., Mendel, J. T., & van Dokkum, P. G. 2013, *ApJ*, **776**, L26
 Conroy, C., Graves, G. J., & van Dokkum, P. G. 2014, *ApJ*, **780**, 33
 Conroy, C., & van Dokkum, P. 2012, *ApJ*, **747**, 69
 Conroy, C., Villaume, A., van Dokkum, P. G., & Lind, K. 2018, *ApJ*, **854**, 139

- Davé, R., Thompson, R., & Hopkins, P. F. 2016, *MNRAS*, **462**, 3265
- Davies, R., & Kasper, M. 2012, *ARA&A*, **50**, 305
- Davis, T. A., & McDermid, R. M. 2017, *MNRAS*, **464**, 453
- de La Rosa, I. G., La Barbera, F., Ferreras, I., & de Carvalho, R. R. 2011, *MNRAS*, **418**, L74
- Esdaile, J., et al. 2021a, *AJ*, **162**, 225
- Esdaile, J., et al. 2021b, *ApJ*, **908**, L35
- Forrest, B., et al. 2020a, *ApJ*, **890**, L1
- Forrest, B., et al. 2020b, *ApJ*, **903**, 47
- Glazebrook, K., et al. 2017, *Nat*, **544**, 71
- Gunawardhana, M. L. P., et al. 2011, *MNRAS*, **415**, 1647
- Gunawardhana, M. L. P., et al. 2020, *MNRAS*, **497**, 3860
- Hernquist, L. 1990, *ApJ*, **356**, 359
- Hill, A. R., van der Wel, A., Franx, M., Muzzin, A., Skelton, R. E., Momcheva, I., van Dokkum, P., & Whitaker, K. E. 2019, *ApJ*, **871**, 76
- Hopkins, A. M. 2018, *PASA*, **35**, 39
- Hopkins, P. F., et al. 2018, *MNRAS*, **480**, 800
- Hunter, J. D. 2007, *CSE*, **9**, 90
- Jafariyazani, M., Newman, A. B., Mobasher, B., Belli, S., Ellis, R. S., & Patel, S. G. 2020, *ApJ*, **897**, L42
- Johnson, B. D., Leja, J., Conroy, C., & Speagle, J. S. 2020, arXiv e-prints, p. arXiv:2012.01426
- Kalita, B. S., et al. 2021, *ApJ*, **917**, L17
- Kauffmann, O. B., et al. 2020, *A&A*, **640**, A67
- Kissler-Patig, M., et al. 2008, *A&A*, **491**, 941
- Kobayashi, C., & Nomoto, K. 2009, *ApJ*, **707**, 1466
- Kriek, M., et al. 2016, *nat*, **540**, 248
- Kroupa, P. 2001, *MNRAS*, **322**, 231
- Kubo, M., Tanaka, M., Yabe, K., Toft, S., Stockmann, M., & Gómez-Guijarro, C. 2018, *ApJ*, **867**, 1
- La Barbera, F., Ferreras, I., Vazdekis, A., de la Rosa, I. G., de Carvalho, R. R., Trevisan, M., Falcón-Barroso, J., & Ricciardelli, E. 2013, *MNRAS*, **433**, 3017
- Lacey, C. G., et al. 2016, *MNRAS*, **462**, 3854
- Leja, J., Carnall, A. C., Johnson, B. D., Conroy, C., & Speagle, J. S. 2019, *ApJ*, **876**, 3
- Leja, J., Johnson, B. D., Conroy, C., van Dokkum, P. G., & Byler, N. 2017, *ApJ*, **837**, 170
- Madau, P., & Dickinson, M. 2014, *ARA&A*, **52**, 415
- Marchesini, D., et al. 2010, *ApJ*, **725**, 1277
- Marsan, Z. C., Marchesini, D., Brammer, G. B., Geier, S., Kado-Fong, E., Labbé, I., Muzzin, A., & Stefanon, M. 2017, *ApJ*, **842**, 21
- Martín-Navarro, I. 2016, *MNRAS*, **656**, L104
- Martín-Navarro, I., Barbera, F. L., Vazdekis, A., Falcón-Barroso, J., & Ferreras, I. 2015a, *MNRAS*, **447**, 1033
- Martín-Navarro, I., et al. 2015b, *ApJ*, **806**, L31
- Maseda, M. V., et al. 2021, arXiv e-prints, p. arXiv:2110.00009
- Merlin, E., et al. 2019, *MNRAS*, **490**, 3309
- Naidu, R. P., Tacchella, S., Mason, C. A., Bose, S., Oesch, P. A., & Conroy, C. 2019, arXiv e-prints, p. arXiv:1907.13130
- Nanayakkara, T., et al. 2016, *ApJ*, **828**, 21
- Nanayakkara, T., et al. 2017, *MNRAS*, **468**, 3071
- Nanayakkara, T., et al. 2020, *ApJ*, **889**, 180
- Narayanan, D., & Davé, R. 2012, *MNRAS*, **423**, 3601
- Nomoto, K., Tominaga, N., Umeda, H., Kobayashi, C., & Maeda, K. 2006, *Nucl. Phys. A*, **777**, 424
- Patel, S. G., Hong, Y. X., Quadri, R. F., Holden, B. P., & Williams, R. J. 2017, *ApJ*, **839**, 127
- Pauldrach, A. W. A., Hoffmann, T. L., & Lennon, M. 2001, *A&A*, **375**, 161
- Persson, S. E., et al. 2013, *PASP*, **125**, 654
- Plummer, H. C. 1911, *MNRAS*, **71**, 460
- Poggianti, B. M., & Barbaro, G. 1997, *A&A*, **325**, 1025
- Salpeter, E. E. 1955, *ApJ*, **121**, 161
- Saracco, P., et al. 2020, *ApJ*, **905**, 40
- Schreiber, C., et al. 2018a, *A&A*, **611**, A22
- Schreiber, C., et al. 2018b, *A&A*, **618**, A85
- Segers, M. C., Schaye, J., Bower, R. G., Crain, R. A., Schaller, M., & Theuns, T. 2016, *MNRAS*, **461**, L102
- Sharda, P., & Krumholz, M. R. 2022, *MNRAS*, **509**, 1959
- Simpson, J. M., et al. 2017, *ApJ*, **844**, L10
- Smith, R. J. 2014, *MNRAS*, **443**, L69
- Smith, R. J. 2020, *ARA&A*, **58**, 577
- Sparre, M., et al. 2015, *MNRAS*, **447**, 3548
- Spitler, L. R., et al. 2014, *ApJ*, **787**, L36
- Stefanon, M., et al. 2021, *ApJS*, **257**, 68
- Steidel, C. C., Strom, A. L., Pettini, M., Rudie, G. C., Reddy, N. A., & Trainor, R. F. 2016, *ApJ*, **826**, 159
- Straatman, C. M. S., et al. 2014, *ApJ*, **783**, L14
- Straatman, C. M. S., et al. 2016, *ApJ*, **830**, 51
- Sutherland, W., et al. 2015, *A&A*, **575**, A25
- Tanaka, M., et al. 2019, *ApJ*, **885**, L34
- Thater, S., Krajnović, D., Cappellari, M., Davis, T. A., de Zeeuw, P. T., McDermid, R. M., & Sarzi, M. 2019, *A&A*, **625**, A62
- Valentino, F., et al. 2020, *ApJ*, **889**, 93
- van Dokkum, P. G., & Conroy, C. 2010, *Nature*, **468**, 940
- Vazdekis, A., Sánchez-Blázquez, P., Falcón-Barroso, J., Cenarro, A. J., Beasley, M. A., Cardiel, N., Gorgas, J., & Peletier, R. F. 2010, *MNRAS*, **404**, 1639
- Vazdekis, A., et al. 2015, *MNRAS*, **449**, 1177
- Villaume, A., Conroy, C., Johnson, B., Rayner, J., Mann, A. W., & van Dokkum, P. 2017, *ApJS*, **230**, 23
- Walcher, J., Groves, B., Budavári, T., & Dale, D. 2011, *Ap&SS*, **331**, 1
- Walter, F., et al. 2016, *ApJ*, **833**, 67
- Weaver, J. R., et al. 2021, arXiv e-prints, p. arXiv:2110.13923
- Wellons, S., et al. 2015, *MNRAS*, **449**, 361
- Whitaker, K. E., et al. 2011, *ApJ*, **735**, 86
- Williams, R. J., Quadri, R. F., Franx, M., van Dokkum, P., & Labbé, I. 2009, *ApJ*, **691**, 1879
- Williams, C. C., et al. 2018, *ApJS*, **236**, 33
- Yan, Z., Jerabkova, T., Kroupa, P., & Vazdekis, A. 2019, *A&A*, **629**, A93
- Yu H., & Wang F. Y., 2016, *ApJ*, **820**, 114