

Beyond UVJ: Color Selection of Galaxies in the JWST Era

JACQUELINE ANTWI-DANSO  ^{1,2}, CASEY PAPOVICH  ^{1,2}, JOEL LEJA  ^{3,4,5}, DANILO MARCHESINI  ⁶, Z. CEMILE MARSAN  ⁷, NICHOLAS S. MARTIS  ^{8,9}, IVO LABBÉ  ¹⁰, ADAM MUZZIN  ⁷, KARL GLAZEBROOK  ¹⁰, CAROLINE M. S. STRAATMAN  ^{11,12} AND KIM-VY H. TRAN  ^{13,14}

¹George P. and Cynthia Woods Mitchell Institute for Fundamental Physics and Astronomy, Texas A&M University, College Station, TX, USA

²Department of Physics and Astronomy, Texas A&M University, 4242 TAMU, College Station, TX, USA

³Department of Astronomy & Astrophysics, The Pennsylvania State University, University Park, PA 16802, USA

⁴Institute for Computational & Data Sciences, The Pennsylvania State University, University Park, PA, USA

⁵Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802, USA

⁶Department of Physics and Astronomy, Tufts University, Medford, MA, USA

⁷Department of Physics and Astronomy, York University, Toronto, Ontario, Canada

⁸Department of Astronomy and Physics, St. Mary's University, Halifax, Nova Scotia, Canada

⁹National Research Council of Canada, Herzberg Astronomy & Astrophysics Research Centre, 5071 West Saanich Road, Victoria, BC, Canada, V9E 2E7

¹⁰Center for Astrophysics & Supercomputing, Swinburne University, Hawthorn, Victoria, Australia

¹¹Department of Physics and Astronomy, Ghent University, Krijgslaan 281 S9, B-9000 Gent, Belgium

¹²Max-Planck Institut für Astronomie, Königstuhl 17, D-69117, Heidelberg, Germany

¹³School of Physics, University of New South Wales, Kensington, Australia

¹⁴ARC Centre for Excellence in All-Sky Astrophysics in 3D

Submitted to ApJ

ABSTRACT

We present a new rest-frame color-color selection method using “synthetic $u_s - g_s$ and $g_s - i_s$ ”, $(ugi)_s$ colors to identify star-forming and quiescent galaxies. Our method is similar to the widely-used $U - V$ versus $V - J$ (UVJ) diagram. However, UVJ suffers known systematics. Spectroscopic campaigns have shown that UVJ -selected quiescent samples at $z \gtrsim 3$ include $\sim 10 - 30\%$ contamination from galaxies with dust-obscured star formation and strong emission lines. Moreover, at $z > 3$, UVJ colors are extrapolated because the rest-frame J-band shifts beyond the coverage of the deepest bandpasses at $< 5 \mu\text{m}$ (typically *Spitzer*/IRAC $4.5 \mu\text{m}$ or future *JWST*/NIRCam observations). We demonstrate that $(ugi)_s$ offers improvements to UVJ at $z > 3$, and can be applied to galaxies in the *JWST* era. We apply $(ugi)_s$ selection to galaxies at $0.5 < z < 6$ from the (observed) 3D-HST and UltraVISTA catalogs, and to the (simulated) JAGUAR catalogs. We show that extrapolation can affect $(V - J)_0$ color by up to 1 magnitude, but changes $(g_s - i_s)_0$ color by ≤ 0.2 mag, even at $z \simeq 6$. While $(ugi)_s$ -selected quiescent samples are comparable to UVJ in completeness (both achieve $\sim 85\text{-}90\%$ at $z = 3 - 3.5$), $(ugi)_s$ reduces contamination in quiescent samples by nearly a factor of two, from $\simeq 35\%$ to $\simeq 17\%$ at $z = 3$, and from $\simeq 60\%$ to $\simeq 33\%$ at $z = 6$. This leads to improvements in the true-to-false-positive ratio (TP/FP), where we find $\text{TP}/\text{FP} \gtrsim 2.2$ for $(ugi)_s$ at $z \simeq 3.5 - 6$, compared to $\text{TP}/\text{FP} < 1$ for UVJ -selected samples. This indicates that contaminants will outnumber true quiescent galaxies in UVJ at these redshifts, while $(ugi)_s$ will provide higher-fidelity samples.

Keywords: techniques: photometric — catalogs — surveys — galaxies: high-redshift

1. INTRODUCTION

Over the past two decades, deep extragalactic surveys have yielded impressive numbers of galaxies in the high redshift universe, thereby improving our understanding of galaxy evolution. The development of sensitive near-infrared

(NIR) imagers 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) and their incorporation into multiwavelength surveys have yielded thousands of galaxy candidates at $z = 3 - 6$. The addition of medium-band filters with deep NIR surveys such as NMBS (Whitaker et al. 2011) and ZFOURGE (Straatman et al. 2016) have additionally improved the color selection of galaxies at these redshifts, enabling the detection and characterization of red galaxies over cosmic time (e.g. Marchesini

Corresponding author: Jacqueline Antwi-Danso
jadanso@tamu.edu

et al. 2010; Straatman et al. 2014, Spitzer et al. 2014; Patel et al. 2017). Using these data, we have demonstrated that we can accurately determine redshifts and distinguish between star-forming and quiescent galaxies with photometric data alone. Similarly, surveys of deep fields and lensing clusters using the *Hubble Space Telescope* (*HST*) have revealed hundreds of galaxies and galaxy candidates at $z = 7 - 10$, enabling us to identify and characterize some of the very first sources in the Universe (Grogan et al. 2011, Koekemoer et al. 2011, Bradley et al. 2012, 2014, Ellis et al. 2013, Schmidt et al. 2014, Lotz et al. 2017, Salmon et al. 2018, Coe et al. 2019).

The recently-launched *James Webb Space Telescope* (*JWST*) has the potential to do the same, detecting galaxies well out to $z \sim 12$ (Behroozi et al. 2020, Robertson 2022) due to its unparalleled NIR imaging and spectroscopic capabilities. The Near-InfraRed Camera (NIRCam) and the Mid-InfraRed Instrument (MIRI) on *JWST* are designed for broadband photometry from 0.7 – 25.5 microns with unprecedented resolution and sensitivity. Accepted ERS, GTO, and GO¹ programs will carry out the first observations with *JWST*, test each instrument’s capabilities, and set the tone for future science with Cycle 1 and beyond. These will allow us to extend observations of the galaxy rest-frame UV luminosity function to higher redshifts ($z > 10$) and fainter luminosity as well as detect rest-frame optical and NIR spectral energy distributions (SEDs) of high redshift galaxies. Even with these amazing capabilities, spectroscopy of a large number of sources with *JWST* is unfeasible due to its short operational lifespan, and as MIRI has the smallest field of view of all the imaging instruments, we will obtain mid-IR data for less than a third of the survey area of our largest upcoming programs. As a result, most of the early surveys will carry out broadband imaging spanning $\sim 0.9\text{--}4.4$ microns, with the NIR observations being instrumental for characterizing the rest-frame optical of $z \sim 7$ galaxies, similar to what was done with *Spitzer*/IRAC photometry (e.g., Labb   et al. 2013, Oesch et al. 2014, Smit et al. 2015, Roberts-Borsani et al. 2016, Castellano et al. 2017).

For this reason, it is incumbent that we come up with an efficient and accurate method to select galaxies based on photometric colors alone, a method that retains its efficiency and accuracy in the *JWST* era. The most widely-used method for $z \leq 4$ galaxies is the *UVJ* diagram. Star-forming and quiescent galaxies have been shown to exhibit a bimodality in the rest-frame $U - V$ versus $V - J$ color-color space out to $z \sim 3$ (Labb   et al. 2005, Williams et al. 2009, Muzzin et al. 2013, Whitaker et al. 2011, Straatman et al. 2014, 2016). This bimodality correlates with specific star formation rate, MIPS 24 micron flux down to implied SFRs of $\sim 40 M_{\odot}\text{yr}^{-1}$ (Brammer et al. 2011), and morphology (Papovich et al. 2012, 2015; Patel et al. 2012). Additionally, dusty star-forming and quiescent galaxies are difficult to distinguish in single-color diagrams (e.g. $U\text{-}V$ alone; Wolf et al. 2009, Brammer et al.

2009, Leja et al. 2019b), as both SED types exhibit red colors and peak at $1\,\mu\text{m}$ in the rest-frame. The physical motivation then for including a second color (e.g. $V\text{-}J$) is to add a longer wavelength baseline in order to trace out the rest-frame NIR, where dusty star-forming galaxies tend to have redder colors and quiescent galaxies have bluer colors. Consequently, the *UVJ* diagram is a useful and efficient way of selecting different galaxy populations in large photometric datasets.

Despite its utility, the *UVJ* diagram is fraught with problems at $z > 3$. Due to the growing number of spectroscopically-confirmed quiescent samples at these redshifts, we now know that *UVJ*-selected quiescent samples have 21–30% contamination from galaxies with significant levels of ongoing star-formation (e.g. Schreiber et al. 2018, Forrest et al. 2020a). There are several reasons for this. One is that the $\text{H}\beta + [\text{O III}]$ lines can boost the V -band flux, causing galaxies with strong emission lines to appear redder in $U - V$ than they actually are. Additionally, J -band fluxes at $z > 3$ are less accurate because they require extrapolation. This occurs because the J -band corresponds to the *Spitzer*/IRAC bands at these redshifts, for which the data is often shallow (typically by ≈ 2.5 mag, if it is available at all (Sanders et al. 2007, Ashby et al. 2015).

In this *Paper*, we present a new set of synthetic top-hat filters to better separate galaxies in different stages (quiescent versus star-forming) using $(g_s - i_s)$ versus $(u_s - g_s)$ colors (a “ $(ugi)_s$ ” diagram). The outline for this *Paper* is as follows. In Section 2 we describe the datasets we use to test the color-selection methods. In Section 3 we introduce the filters (u_s , g_s , and i_s), which lie at 2900 Å, 4500 Å, and 7500 Å, respectively, and were designed to address the aforementioned issues for use with current ground-based photometric catalogs and in the *JWST* era (Figure 1). In Section 4, we highlight the necessity for a new set of filters for use at $z > 3$ by demonstrating the effects of extrapolation on *UVJ*. Finally, we evaluate the sample selection efficiency and limitations of each method in Section 5. Throughout, we assume a ΛCDM cosmology with $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$ and $H_0 = 70\,\text{km s}^{-1}\,\text{Mpc}^{-1}$. Rest-frame colors are quoted in AB magnitudes (Oke & Gunn 1983), and stellar population parameters were computed using a Chabrier IMF (Chabrier 2003).

2. DATA AND METHODS

2.1. Observational Data ($0.5 < z < 4$)

To test and calibrate the color-color methods, we employ four galaxy catalogs that provide both observed and rest-frame colors (we derive these ourselves where not provided), high-quality measures of photometric (or spectroscopic) redshifts, stellar masses, and star formation rates (SFR). These are important as we wish to test the ability of selection methods to identify galaxies in active and quiescent stages of their evolution. We therefore opt to employ both theoretical catalogs constructed from known details about stellar population models (Section 2.2) and real catalogs that contain physical high-redshift galaxies with imperfect knowledge of their stellar populations, dust attenuation, and emission line contributions (Sections 2.1.1 and 2.1.2). It is important to understand

¹ Early Release Science, Guaranteed Time Observations, and General Observer

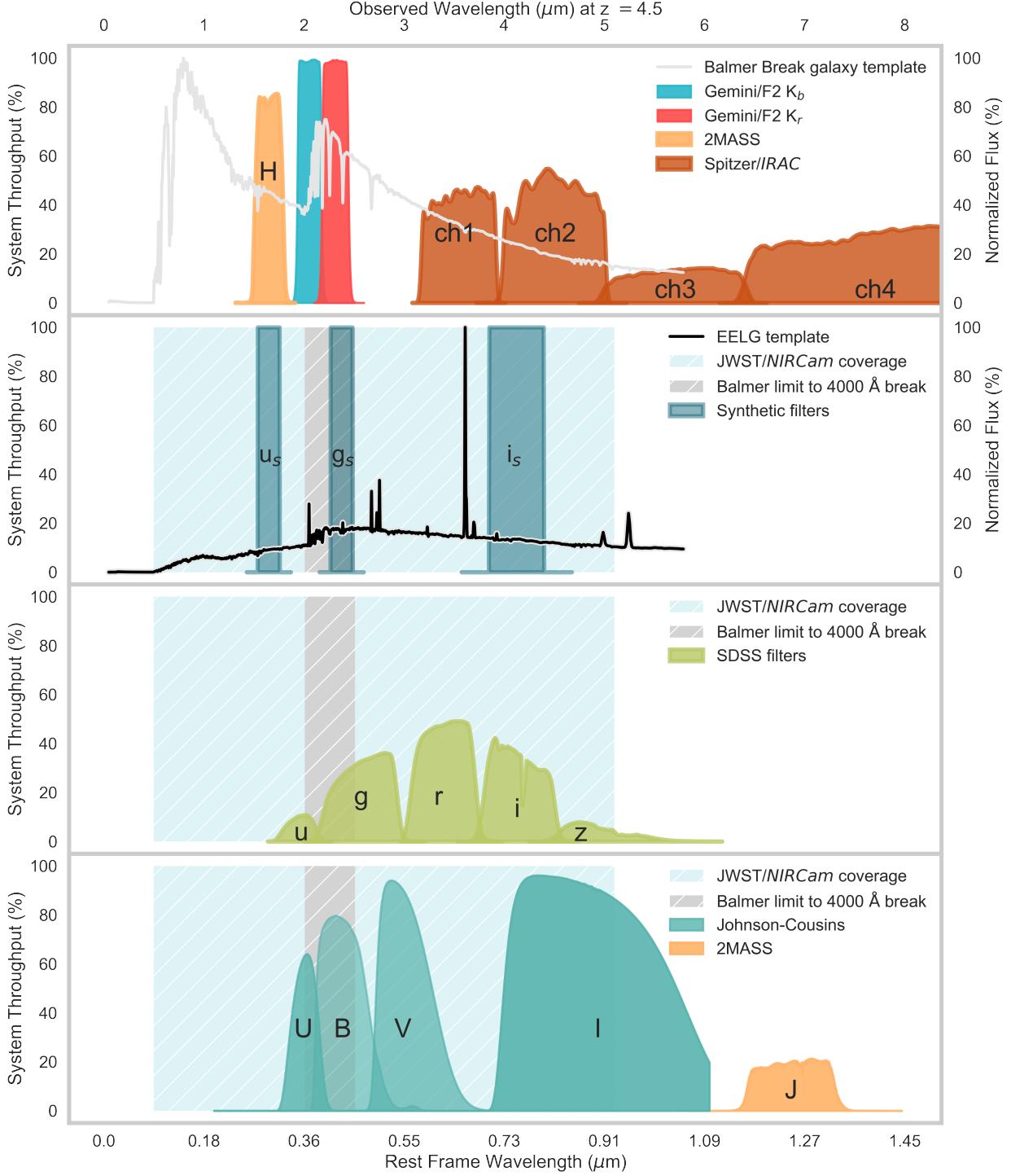


Figure 1. “Classic” photometric bands available in extragalactic catalogs and rest-frame color schemes. Rest-frame UVJ colors at $z > 3$ require extrapolation because the J band is redward of $5\mu\text{m}$ (and it overlaps with *Spitzer*/IRAC channels 3 and 4, for which there is often limited or no data). This will be a problem even with *JWST*/NIRCam observations, because they will also only extend out to $5\mu\text{m}$. Additionally, the V -band filter will be susceptible to artificial boosting from galaxies with strong $\text{H}\beta + [\text{O III}]$ emission lines (at $0.49 - 0.51\mu\text{m}$, illustrated by the SED of the extreme-emission-line galaxy in the second panel), which comprise a significant fraction of star-forming galaxies at $z > 3$, reaching $\sim 50\%$ by $z = 6$. We introduce a set of synthetic (ugr_i)_s filters (second panel) designed to mitigate these problems by (1) being more sensitive to the 4000\AA /Balmer break; (2) avoiding common emission lines; and (3) maintaining overlap with the *Spitzer*/IRAC and *JWST*/NIRCam filters out to $z = 6$.

color-color selection efficacy from both types of catalogs as we move into the even more distant universe.

2.1.1. 3D-HST

One of the empirical catalogs we use are those from 3D-HST (Skelton et al. 2014), as these provide some of the most robust results from spectral-energy-distribution (SED) fitting, but are limited to $z < 2.5$ (because the catalogs are selected in the *HST*/F160W band). The data were taken in five well-studied extragalactic fields of CANDELS (Grogin et al. 2011; Koekemoer et al. 2011) imaged in 19–45 photometric bands. We use results of Prospector fits to the 3D-HST photometric catalogs from Leja et al. (2019c). The 3D-HST catalogs provide observed-frame 0.3–24 μm photometry and redshifts for $\sim 200,000$ galaxies, complete down to $10^9 M_\odot$ at $z = 2$ (Tal et al. 2014). The catalog redshifts comprise, in order of accuracy, (1) ground-based spectroscopic redshifts, (2) near-infrared grism redshifts (Momcheva et al. 2016), and (3) photometric redshifts from EAZY (Brammer et al. 2008).

We use stellar population parameters inferred with the 14-parameter Bayesian SED-fitting Prospector- α framework described in (Johnson & Leja 2017) and (Leja et al. 2017), fit to a 90% stellar mass-complete sample of 58461 of galaxies at $0.5 < z < 2.5$ (Leja et al. 2019c). The lower redshift limit ($z = 0.5$) is set by the redshift at which the aperture photometry starts to become unreliable. Prospector- α utilizes the wide variety of physical implementations available in FSPS (Conroy & Gunn 2010), including a flexible six-parameter nonparametric star formation history (SFH), state-of-the-art MIST stellar isochrones, a wide range of stellar metallicities, a two-component dust attenuation model with a flexible dust attenuation curve, dust emission via energy balance, nebular line and continuum emission, and a model for the mid-infrared emission of dusty AGN tori. All of these enable Prospector- α to mitigate the persistent factor-of-two uncertainty in stellar masses and SFRs derived by SED-fitting codes (Papovich et al. 2001, Marchesini et al. 2009, Muzzin et al. 2009, Wuyts et al. 2009, Behroozi et al. 2010, Pforr et al. 2012, Conroy 2013, Mitchell et al. 2013, Leja et al. 2015, 2019c, Mobasher et al. 2015, Santini et al. 2015, Tomczak et al. 2016, Carnall et al. 2018) by making it possible to model systematics on stellar masses and SFRs galaxy by galaxy.

We note that although *UV + IR* SFRs, which are intended to be more reliable than those from SED-fitting codes, as they account for reprocessed emission from dust-obscured star formation, are available for the 3D-HST catalogs (e.g. Barro et al. 2019, Whitaker et al. 2014), they tend to underestimate the flux from old ($t > 100$ Myr) stellar populations. This becomes particularly important for low sSFR objects, with the UV and IR luminosity from old stars comprising $\sim 50\%$ of the total flux at $\log(\text{sSFR} < -10.5)$, and dominating the total light output at $\log(\text{sSFR} = -13)$ (Martis et al. 2019, Leja et al. 2019c). Because our study is primarily concerned with selecting highly pure samples of quiescent galaxies (which generally have low sSFRs) we chose the SFRs inferred by this modified Prospector- α model, which self-consistently

estimates the effect of dust heating by old stars, over catalogs based on *UV + IR*-derived SFRs.

Additionally, we compare our findings with Prospector- α to those using SFRs, stellar masses, and dust extinction from FAST (Kriek et al. 2009) (provided with the 3D-HST catalogs) because this has been the traditional choice in the literature. These are based on single-parameter, τ -model star-formation histories, where the SFR scales as $\exp(-t/\tau)$ (for an e -folding timescale, τ). Although FAST stellar masses are well-constrained because they mostly depend on rest-frame optical photometry, which is well covered by the 3D-HST data, star formation rates and dust are not (Wuyts et al. 2012). We discuss this further in Section 5.1. In this work we use a subsample of galaxies with reasonable photometry (i.e. 3D-HST `use_phot` = 1, with $K_s \text{ SNR} \geq 7$) and stellar mass $M_* > 10^{10} M_\odot$. Reported star formation rates from Prospector- α are averaged over the most recent 100 Myr (rather than instantaneous values).

2.1.2. UltraVISTA

Catalogs such as CANDELS and 3D-HST lack large samples of galaxies at the upper end of the stellar mass function, particularly at the highest redshifts (e.g., Merlin et al. 2018) because they have a small survey area (~ 0.2 sq degrees). Old and dusty galaxies tend to be better sampled by K-selected surveys. For this reason we also use results from the UltraVISTA survey (McCracken et al. 2012), which covers 1.5 deg 2 in the COSMOS field with multiwavelength data, and better samples massive galaxies at high redshifts. For galaxies at $2.5 < z < 4$, we use MAGPHYS fits from Martis et al. (2019) to the UltraVISTA DR3 catalog which is mass-complete at $\log(M_*/M_\odot) > 10.5$ out to $z \simeq 4$. This deep catalog (5σ limiting magnitude in $K_s = 25.2$ mag) includes UV-NIR photometry spanning 49 bands. These were supplemented with *Spitzer*/MIPS 24 μm observations and 100 μm and 160 μm observations from the *Herschel* PACS Evolutionary Probe (PEP; Lutz et al. 2011), and 250 μm , 350 μm , and 500 μm observations from the *Herschel* Multi-Tiered Extragalactic Survey (HerMES; Oliver et al. 2012). Redshifts in the catalog are either from the EAZY photometric redshift code or spectroscopic redshifts where available.

Martis et al. (2019) used the high-z extension of MAGPHYS (da Cunha et al. 2008, 2015) to model the UV-FIR photometry of the UltraVISTA sources. The star formation history is parameterized as a continuous delayed exponential of the form $\gamma^2 t e^{(-\gamma t)}$, where γ is the inverse of the star formation timescale, τ . This form allows high redshift galaxies to have SFHs that rise with time (negative t), suggested by many observations at high redshifts (Lee et al. 2010, Papovich et al. 2011, Reddy et al. 2012, Carnall et al. 2019). Additionally, it helps reduce systematic effects introduced by too simplistic functional forms of the SFH (e.g., exponentially declining, see subsection 2.3). Moreover, the inclusion of the far-IR data from *Herschel* ensures stronger constraints on the estimated SFRs compared to those estimated from modeling the UV-to-NIR photometry alone. MAGPHYS SFRs are reported

by averaging the most recent SFH over a 10 Myr period. For our sample, we select galaxies with K_s mag ≤ 23.5 AB.

2.2. Simulation Data at $4 < z < 6$

Our goal is to develop a simple color-color selection method for distinguishing samples of quiescent and star-forming galaxies that works well for current ground- and space-based photometric data, and also in the *JWST* era. To achieve this, we use the *JWST* extragalactic mock catalog, publicly available as the JAdes extraGalactic Ultradeep Artificial Realizations (JAGUAR) package (Williams et al. 2018). Unlike semi-analytic models and hydrodynamical simulations, JAGUAR uses an empirically-driven approach: modeling observed galaxy distributions and scaling relations and matching realizations to our deepest extragalactic surveys.

The JAGUAR catalog is briefly described as follows: number counts for star-forming and quiescent galaxies are determined using the Tomczak et al. (2014) mass functions and UV luminosity functions at $0 < z < 10$ (Bouwens et al. 2015, Oesch et al. 2018; for star-forming galaxies at $z > 4$). UV luminosity functions are used to estimate star-forming galaxy counts at $z > 4$, rather than stellar mass functions because the Balmer/4000 Å break shifts into near-infrared at these redshifts, where most facilities have low sensitivity, which could make stellar mass estimates at these redshifts uncertain. The downside to using UV luminosity functions, however, is that dusty star-forming galaxies, which are the primary contaminants of quiescent galaxies in color-selected samples, are underrepresented in UV-selected samples. We artificially infuse dusty star-forming galaxies into the JAGUAR catalog by reddening a subsample of star-forming galaxies using the Calzetti et al. (2000) dust law and the mass-dependent Panella et al. (2009) relation for the extinction of star-forming galaxies in the ultraviolet (see Appendix A). Even with this addition, we see a dearth of dusty star-forming galaxies in the JAGUAR color-color plots (discussed further in Section 5.1.2).

JAGUAR generates mock SEDs and stellar population parameters for each galaxy using BEAGLE (Chevallard & Charlot 2016) and matches them to 3D-HST objects based on redshift and stellar mass (for $\log(M_*/M_\odot) > 8.7 + 0.4z$, which corresponds to $\log(M_*/M_\odot) = 9.89$ at $z = 3$). Outside the parameter space covered by 3D-HST, they adopt the properties of the mock SEDs from BEAGLE. To derive stellar population parameters, they use a delayed exponential SFH and model stellar emission with the Bruzual & Charlot (2003) stellar population synthesis code. The JAGUAR catalogs incorporate line and continuum emission from gas photoionized by young, massive stars using the Gutkin et al. (2016) models and dust attenuation by the two-component model of Charlot & Fall (2000). The catalog contains broadband photometry in *HST* and *JWST*/NIRCam filters from $0.4 - 4.8\mu\text{m}$, emission line measurements, and galaxy morphological properties. As with the datasets above, we use JAGUAR SFRs averaged over the most recent 100 Myr.

2.3. Rest-Frame Colors

Using the aforementioned datasets (3D-HST, UltraVISTA, and JAGUAR), we derive $(ugi)_s$ rest-frame colors using eazy-py (Brammer 2021), a Python-based SED-fitting code based on the photometric redshift fitting code EAZY (Brammer et al. 2008). EAZY was built to handle faint galaxy samples with limited spectroscopic redshifts, as we often have with deep NIR photometric surveys. It fits a non-negative, linear combination of empirically-derived templates (in a user-defined list) to the observed photometry. Two features that distinguish EAZY from other photometric redshift fitting codes and make it ideal for fitting high redshift galaxies are (1) a template error function, which seeks to account for wavelength-dependent corrections of the templates, such as variations in the dust extinction law and missing spectral features; and (2) an apparent magnitude prior, which assigns low probabilities to low-redshift solutions for extremely bright galaxies at high redshift.

One of the important modifications eazy-py makes to EAZY is that it determines rest-frame colors by doing a “weighted interpolation.” It refits the templates to the data, weighting more strongly the observed photometry that is nearest the rest-frame band (in wavelength) and down-weights photometry that is farther away. The rest-frame colors are then interpolated from the model fluxes flanking the rest-frame band of interest.

Generally, the way the best-fit SED is determined impacts the derived rest-frame colors. When the best fit is an arbitrary linear combination of templates (as it is in EAZY and eazy-py), this method produces similar results as interpolating from the observed photometry. This is important to note, as rest-frame colors that are based on the best-fit SED are heavily influenced by the choice of template set and the assumed star-formation history (SFH). Merlin et al. (2018) found that rest-frame *UVJ* colors can differ by up to 0.3 mag when using an exponentially-declining (τ) SFH versus a top-hat SFH, and this change occurs mostly in the $V - J$ direction. Rest-frame colors based on the observed photometry are not subject to these systematics, however they are more sensitive to sharp effects (e.g. emission lines and spectral breaks) and low S/N. eazy-py adopts the best features of each method by using empirical template sets (which do not assume a SFH) and using all the available photometry, weighting more strongly bands closest to the rest-frame band of interest. This way, the estimated uncertainties on the rest-frame colors (e.g. Figures 4, 5, and 6) are purely statistical and not prone to the limitations and biases of the chosen template set and SFH (we return to this issue in Section 5.3.2).

For catalogs that provide *UVJ* and *ugi* colors, we use those, as they are tied to the stellar masses and star formation rates that we use for the subsequent analysis in this *Paper*. This applies for instance to the *UVJ* colors from Prospector- α from $0.5 < z < 2.5$ (Figure 2), which were determined by marginalizing over the likelihoods of the SED parameters. For catalogs that do not provide rest-frame colors (e.g., JAGUAR), we use a set of 10 templates which model the following galaxy

Table 1. Synthetic filter details. Central wavelengths (λ_c) and widths ($\Delta\lambda$) of our proposed $(ugi)_s$ filters.

Filter Name	λ_c (Å)	$\Delta\lambda$ (Å)
Synthetic u (u_s)	2900	400
Synthetic g (g_s)	4500	400
Synthetic i (i_s)	7500	1000

populations: emission line galaxies, galaxies that are both old and dusty, old quiescent galaxies, and post-starbursts. We used the v.1.0. template error function and fixed the redshifts to the provided photometric (or spectroscopic redshifts).

3. FILTER DESIGN AND CALIBRATION

3.1. The Ideal Filter Combination

The ideal filter set for distinguishing star-forming galaxies from quiescent galaxies at $z > 3$ should match UVJ in simplicity and convenience while reducing the biases inherent in UVJ selection such that it will be useful even in the *JWST* era. We designed a set of top-hat filters with central wavelengths at 2900 Å, 4500 Å, and 7500 Å, corresponding to the SDSS u , g , and i filters, respectively, shown in Figure 1.

The synthetic filters were designed to accomplish three main goals. First, they robustly straddle the Balmer/4000 Å break, similar to the synthetic U_m and B_m filters in Kriek et al. (2010). The wider wavelength spacing of the u_s and g_s filters combined with their narrower widths means that the Balmer Break produces a stronger color signature in $u_s - g_s$ than in $u - g$ and $U - V$. We explore this idea further in Section 5.3.1. For this reason, the $(ugi)_s$ filters complement data from NIR medium-band surveys, which were designed to achieve higher resolution sampling of the Balmer/4000 Å break. This includes new imaging using K_b and K_r filters ($R \sim 10$) from the FENIKS survey, which split the K_s band into a bluer and redder filter, similar to the strategy employed by the NMBS and ZFOURGE surveys, which split the J and H bands for increased wavelength sampling of the Balmer break. These K_b and K_r filters have been shown to reduce photometric redshift uncertainties by factors of 2 and 4 at $z < 3$ and $z > 4$, respectively, and reduce contamination from low-redshift interlopers, usually dusty star-forming galaxies, by a factor of 2 at $4 \lesssim z \lesssim 5$ (Esdaile et al. 2021). We plan to apply our $(ugi)_s$ selection to data from FENIKS in a forthcoming paper.

Second, the synthetic $(ugi)_s$ filters avoid regions with strong emission lines. These can boost photometry in a particular bandpass, mimicking red $U - V$ or blue $V - J$ colors, hence quiescence (Section 5.2). A similar problem occurs with SDSS $u - g$ colors. We can quantify the magnitude increase (Δm) in a filter of width $\Delta\lambda$ using the following, where W_0 is the equivalent width of the emission line (Papovich et al. 2001):

$$\Delta m \approx -2.5 \log \left[1 + \frac{W_0(1+z)}{\Delta\lambda} \right] \quad (1)$$

A star-forming galaxy at $z = 3$, for instance, with a rest-frame $H\beta+[O\text{ III}]$ emission line equivalent width (EW) of 500 Å will boost the $U - V$ color by $\simeq 1.2$ mag. This problem is exacerbated at higher redshifts, because high EW objects comprise an increasingly larger fraction of the star forming population in those regimes (Endsley et al. 2021, Tang et al. 2021). The synthetic filters solve this problem by avoiding $[O\text{ II}]$ (3727 Å), $H\beta+[O\text{ III}]$ (in the region 4863–5007 Å), and $H\alpha+[N\text{ II}]$ (in the region 6548–6563 Å). However the FWHM of the $(ugi)_s$ filters is small (400 Å compared to 991 Å for V and 1390 Å for SDSS g). This means that although g_s should not be contaminated, if it is (e.g., due to photometric redshift uncertainties), a stronger magnitude change is introduced. We discuss this further in Section 5.3.2.

Third, the synthetic $(ugi)_s$ filters improve on uncertainties from the age-dust degeneracy while avoiding extrapolation. Galaxies with blue $U - V$ colors generally exhibit unobscured star formation. Those that are red in $U - V$, however, could either be galaxies that have evolved stellar populations, or star-forming galaxies that are obscured by dust. However, because quiescent galaxies generally have relatively lower dust attenuation, they tend to have bluer $V - J$ colors. Thus, the longer the wavelength baseline, the higher the sensitivity to dust. It is for this reason that FUV/mid-infrared colors are more efficient than UVJ at selecting galaxies with low sSFRs (Leja et al. 2019b).

Unfortunately though, a longer wavelength baseline is more likely to require the extrapolation of observed-frame colors beyond where there is data in order to determine rest-frame colors. At $z \gtrsim 4$, the rest-frame J -band corresponds to $\gtrsim 6 \mu\text{m}$ observed, generally only available from IRAC 5.8 μm and redder bands (e.g. forthcoming *JWST*/MIRI data). At longer wavelengths, the flux sensitivity of the IRAC data is substantially lower (by factors of 4 and 8 in IRAC channels 3 and 4 compared to IRAC channels 1 and 2), hence those data are often unavailable or too shallow (Figure 3)². Rest-frame UVJ colors at $z > 4$ therefore require extrapolation and hence are less reliable. Figures 4 and 5 illustrate the effects of extrapolation on the J -band. The new synthetic filters proposed here cover the longest wavelength baseline that is possible while maintaining overlap with *Spitzer*/IRAC channels 1 and 2 and *JWST*/NIRCam coverage, to avoid extrapolation at redshifts out to $z \sim 6$ (the end of the epoch of reionization). We demonstrate the effects of extrapolation on colors using the rest-frame J -band further in Section 4.

Finally, we discuss the choice of a synthetic u_s filter over an ultraviolet one (e.g., rest-frame NUV). $NUV - r$ or $FUV - V$ color selection, for instance, is a better discriminator of current versus past star formation activity than $U - V$ (Ilbert et al. 2010, Martin et al. 2007, Arnouts et al. 2007, Ilbert et al. 2013, Man et al. 2016, Hwang et al. 2021, Leja et al. 2019b). Similarly, the $FUV - V$ color has also been shown to have a much stronger correlation with specific star formation rate (Leja et al. 2019b). $NUVrJ$ -selected quiescent samples have

² IRAC Instrument Handbook: <https://doi.org/10.26131/irsa486>

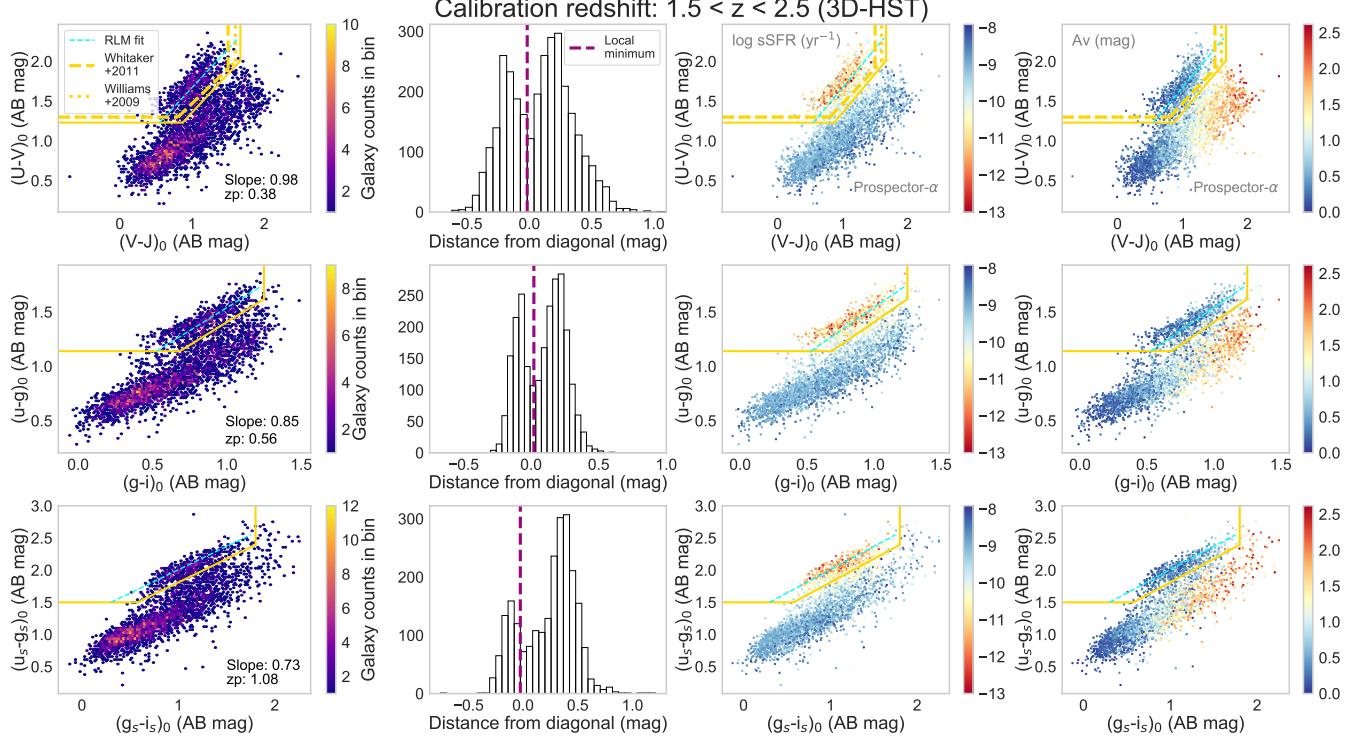


Figure 2. Calibration and validation of color-color diagrams. From top to bottom, the rows show different color-color schema applied to samples of galaxies from the 3D-HST catalog with stellar mass, $\log(M_*/M_\odot) \geq 10$, and redshift range $z = 1.5 - 2.5$. **Top:** UVJ ; **Middle:** ugi ; and **Bottom:** $(ugi)_s$. The **Left** panel of each row shows the color-color distributions. We determine the slope of the diagonal line by fitting the “red sequence” (of galaxies in the upper left of each color-color space) using robust linear methods (RLM), which are less sensitive to outliers than ordinary least squares. We then measure the y-intercept (zeropoint [zp]) by identifying a local minimum between the red and blue sequences (illustrated in the histogram in the second panel of each row). Similar to UVJ , the rest-frame ugi and synthetic $(ugi)_s$ colors correlate with specific star formation rate and dust estimates from **Prospector- α** , shown in the third and fourth panels in each row, respectively (dots show different galaxies color-coded by the aforementioned properties).

also been shown to have low contamination ($\leq 10\text{--}15\%$) from dusty star-forming galaxies at $z \leq 3$ (Man et al. 2016, Hwang et al. 2021). These are all reasons to choose an ultraviolet filter over an optical u filter.

However, using a rest-frame color selection involving a UV filter would miss recently-quenched or “post-starburst” galaxies, particularly at high redshifts. Post-starbursts have intermediate-age stellar populations, which means that while they have strong Balmer/4000 Å breaks, they also have a higher UV continuum than that of a typical quenched galaxy due to residual star formation (see the template in the first panel of Figure 1). As a result of their elevated UV flux, their $NUV - r$ colors are *blue*, similar to those of star-forming galaxies. This has been demonstrated in the literature. For example, Valentino et al. (2020) showed two post-starburst galaxies at $z \sim 4$ that would be classified as star-forming using $NUVrJ$ -selection and quiescent using UVJ .

By extension, post-starburst galaxies may be missed using rest-frame $NUV - g$ colors. These galaxies are observed (and expected) to be more common at higher redshifts ($z > 3$ e.g. Forrest et al. 2020b, D’Eugenio et al. 2020). It is therefore

prudent to use a color-selection that will include them. For this reason, the rest-frame $u_s - g_s$ color is better suited to identify galaxies with colors typical of post-starburst galaxies, which will be an important class of objects at $z > 3$.

3.2. Color-Color Selection Criteria

In this section, we describe how we define empirical UVJ , ugi , and $(ugi)_s$ selection criteria that separate quiescent and star-forming galaxies in our data. The basic quantities that we are interested in are the slope and y-intercept of the diagonal line, as well as the positions of the horizontal and vertical color cuts. These are based on a fit to an initial quiescent selection, which we vary until we obtain a clear bimodal distribution centered at zero. This process is a modification of the method in Kawinwanichakij et al. (2016). It is summarized as follows, and illustrated in Figure 2:

1. Find quiescent galaxies using literature criteria for the specified redshift where available (e.g. Williams et al. (2009), Whitaker et al. (2011), Muzzin et al. (2013) for UVJ) or by making an initial guess based on the

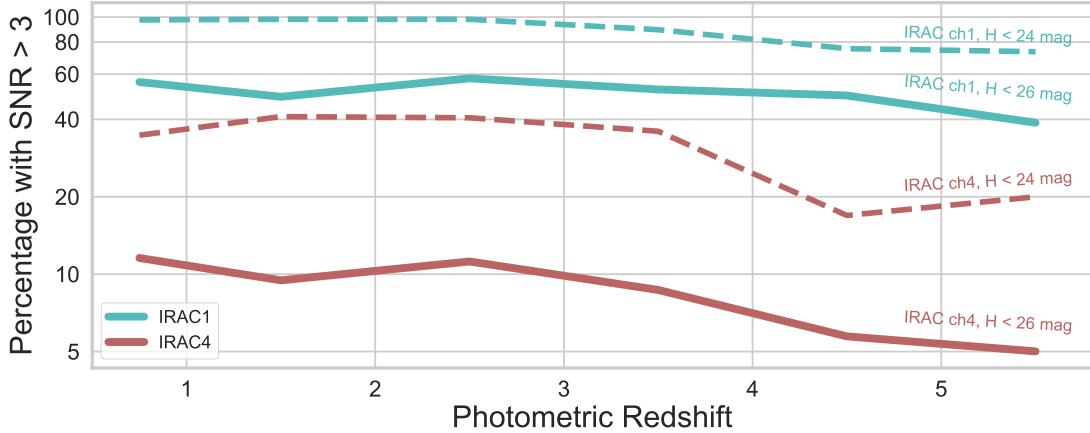


Figure 3. Percentage of detected galaxies with signal-to-noise ratio greater than 3 as a function of redshift for two of the reddest bands in our deepest photometric catalogs, Spitzer/IRAC channels 1 and 4 (with central wavelengths $3.6\ \mu\text{m}$ and $7.9\ \mu\text{m}$). We use a sample of 32528 galaxies in 3D-HST at $0.5 < z < 6$ with reasonable photometry and at two different detection limits (H mag < 24 and H mag < 26), with an additional requirement that the galaxy is detected in the K -band ($\text{SNR} > 5$). Channels 1 and 4 correspond to the rest-frame J -band at $z \sim 2$ and $z \sim 4.5$. This means that rest frame UVJ colors will require extrapolation for up to 60% of galaxies at $z \sim 2$ and up to 95% of galaxies at $z \sim 4.5$. This problem will persist even with *JWST* data, as the NIRCam filter coverage does not extend beyond $5\mu\text{m}$ and MIRI data will be obtained for $< 30\%$ of the survey area of upcoming GO, GTO, and ERS programs. The Spitzer/IRAC channels 2 and 3 data have similar detection fractions as channels 1 and 4, respectively.

distribution of data in color-color space (for ugi and $(ugi)_s$).

2. Update the slope of the diagonal line using the slope from an iteratively re-weighted robust linear methods (RLM) fit to a sigma-clipped subsample of the galaxies in the “red sequence” region of the color-color space (these are candidate quiescent galaxies). We opt to use RLM fits as they are less sensitive to outliers and do not assume homoscedastic errors (e.g. Croux & Rousseeuw 1992). This is illustrated in the first panel of each row of Figure 2.
3. Find the distance of each point in the data set from the diagonal line to delineate the red and blue sequences. This is illustrated in the second panel of each row of Figure 2.
4. Bin distances to determine the local minimum between the two sequences. The optimal number of bins is determined using the Freedman-Diaconis rule (Freedman & Diaconis 1981), which minimizes the area under the probability distribution function of the data and that of the function that best matches the data. Finally, we use the local minimum to update the y-intercept of the diagonal line.

We evaluate our method first by applying it to a sample of galaxies from the ZFOURGE survey (Straatman et al. 2016) and the third data release of the UltraVISTA survey at $0.3 < z < 1.6$, following Kawinwanichakij et al. (2016). Our zeropoints and slopes for the aforementioned are within 0.02 and 0.17 mag of Kawinwanichakij et al. (2016) for the two datasets, respectively.

We then determine the color-color selection criteria for each dataset used in this paper (3D-HST, UltraVISTA, and JAGUAR) using the method outlined above. This is important, as the color lines are not portable from one survey to another due to systematic biases (discussed further in Section 5.1). Because the observed data at $z > 3$ are sparse and generally of poorer quality, we calibrate the color-color lines on a broad redshift range ($z = 1.5 - 2.5$) and apply these calibrations to the entire dataset, up to $z = 6$ in the discussion that follows. Calibrating the lines on a broad redshift range enables us to account for any evolution in the zeropoint as a function of redshift. We apply a zeropoint adjustment at $z = 4 - 6$ based on the UVJ colors of spectroscopic samples at $z = 3 - 4$ in the literature (Section 5.1.1).

Figure 2 shows a bimodality in each panel, indicating the presence of a distinct star-forming and quiescent population in all three color-selection methods: UVJ , ugi , and $(ugi)_s$. We define the selection of quiescent galaxies for each of the color-color selections in Figure 2 with the following criteria. For UVJ :

$$(U - V) > 1.23 \wedge (V - J) < 1.67 \wedge (U - V) > (V - J) \times 0.98 + 0.38 \quad (2)$$

For ugi :

$$(u - g) > 1.14 \wedge (g - i) < 1.25 \wedge (u - g) > (g - i) \times 0.85 + 0.56 \quad (3)$$

For $(ugi)_s$:

$$(u_s - g_s) > 1.5 \wedge (g_s - i_s) < 1.8 \wedge (u_s - g_s) > (g_s - i_s) \times 0.73 + 1.08 \quad (4)$$

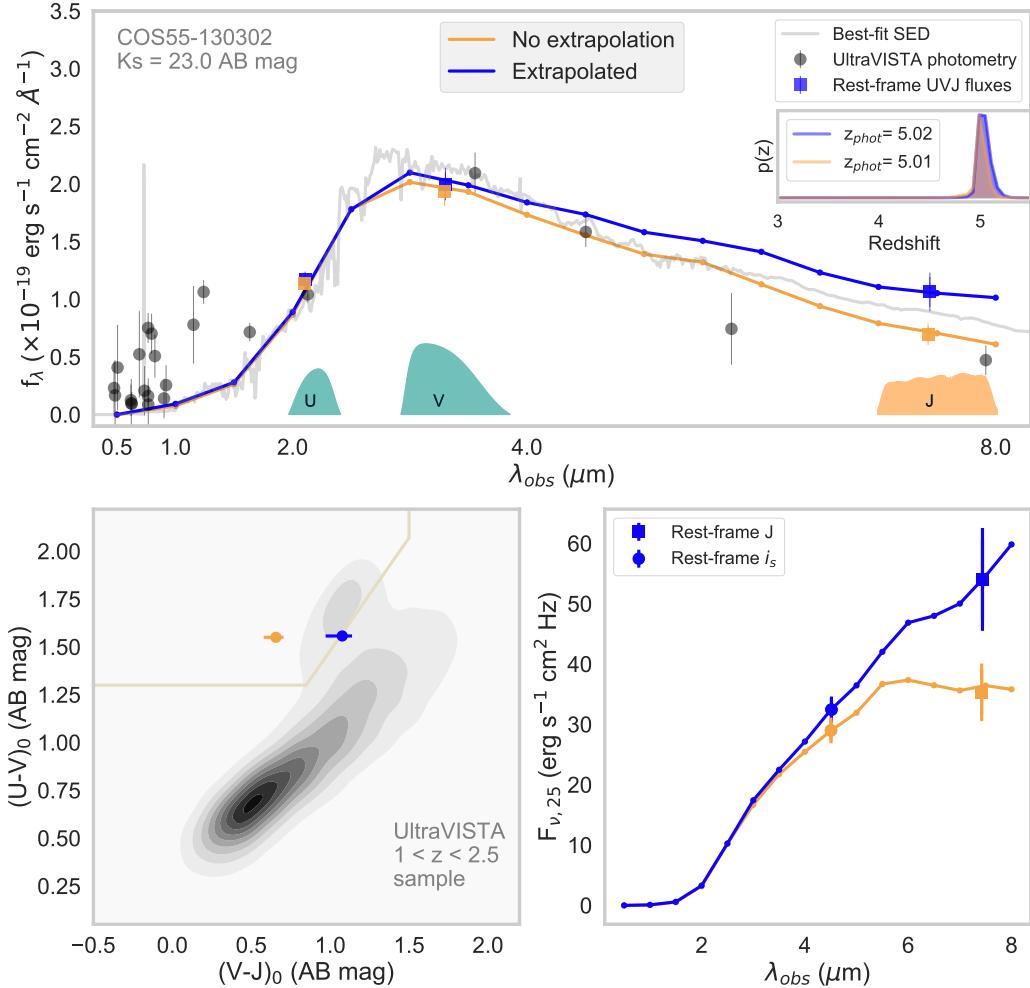


Figure 4. Examining the impact of extrapolation on the rest-frame *UVJ* colors of a $z = 5$ quiescent galaxy candidate. **Top:** Photometry, best-fit SED, and photometric redshift solution of COS55-130302 with extrapolation (blue curve/points) and without extrapolation (orange curve/points). For each case, we derive rest-frame fluxes in synthetic ($R = 10$) tophat filters in the observed frame (from $0.5 - 8 \mu\text{m}$). We simulate extrapolation by removing the observed-frame photometry corresponding to the *J*-band and redder. At $z = 5$, this corresponds to excluding data from *Spitzer*/IRAC channels 3 and 4, which is either unavailable or too shallow for at least 70% of galaxies at this redshift (Figure 3). **Bottom left:** Although extrapolation has little impact on the best-fit SED and photometric redshift, it results in an almost factor of 2 difference in the estimated rest-frame *J*-band flux, resulting in a 0.46 mag change in $(V - J)_0$ color. Due to this extrapolation, galaxies like like COS55-130302 can be erroneously scattered out of *UVJ*-selected samples. **Bottom right:** We repeat this experiment with the $(ugis)_s$ filters and find that our $(g_s - i_s)_0$ color, in comparison, is only affected by 0.08 mag, indicating that the $(ugis)_s$ diagram is better suited for selecting quiescent galaxies at $z > 4$.

In all cases, \wedge is the logical AND operator.

4. EFFECTS OF EXTRAPOLATION

UVJ colors at $z > 2.5$ can be uncertain because the rest-frame *J*-band is subject to extrapolation. Generally, rest-frame colors are determined by interpolating between the two filters flanking the rest-frame band of interest (see Section 2.3 for a more detailed description). At $z = 3$, for instance, the *J*-band overlaps with *Spitzer*/IRAC channels 1 and 2 at

$3.6 \mu\text{m}$ and $4.5 \mu\text{m}$, respectively. These data are often available (at $\text{SNR} > 3$) for only $\sim 55\%$ of galaxies in our deepest photometric catalogs (Figure 3). This means that for $\sim 45\%$ of galaxies at $z \sim 3$, their rest-frame *J* fluxes are determined by extrapolating from the reddest observed data point. This problem is exacerbated at higher redshifts ($z \gtrsim 4$) because longer wavelength data (channels 3 and 4) are available for an even smaller fraction of galaxies (< 7%).

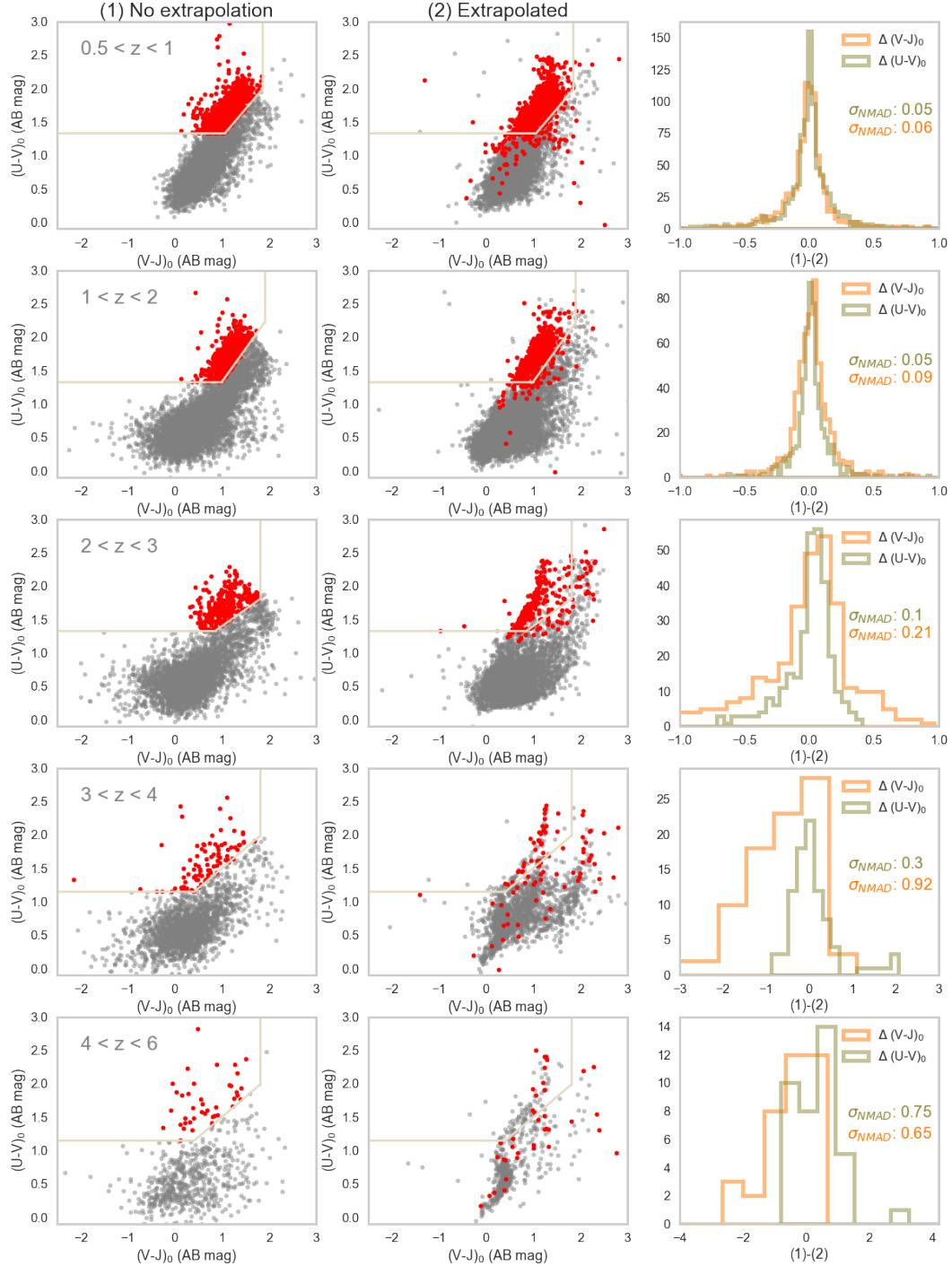


Figure 5. Impact of extrapolation on UVJ -selected quiescent galaxies with rest-frame J -band coverage from *Spitzer* at $0.5 < z < 6$. Each row shows the galaxies in bins of increasing redshift (top to bottom). The **Left** panel in each row shows the distribution using rest-frame colors derived with all the available photometry (no extrapolation). The red dots show galaxies selected as UVJ -quiescent. The **Middle** panel of each row shows the colors of the same galaxies on the left when the rest-frame colors are extrapolated (when the observed fluxes corresponding to the J -band and redder are removed). The **Right** panel of each row shows the distributions of the difference in $U - V$ and $V - J$ colors between the non-extrapolated and extrapolated cases for the quiescent galaxies. We quantify this change using the normalized median absolute deviation (σ_{NMAD}). The effects of extrapolation are seen as an increase in σ_{NMAD} with increasing redshift, with $\sigma(\Delta[V - J])$ increasing from 0.2 mag at $z \sim 2$ to >0.65 mag at $z > 3$. NB: the abscissa ranges in the top three right-hand plots are the same, however we adjust it for the bottom two ($z > 3$) to better show the large change extrapolation introduces to rest-frame colors at these redshifts.

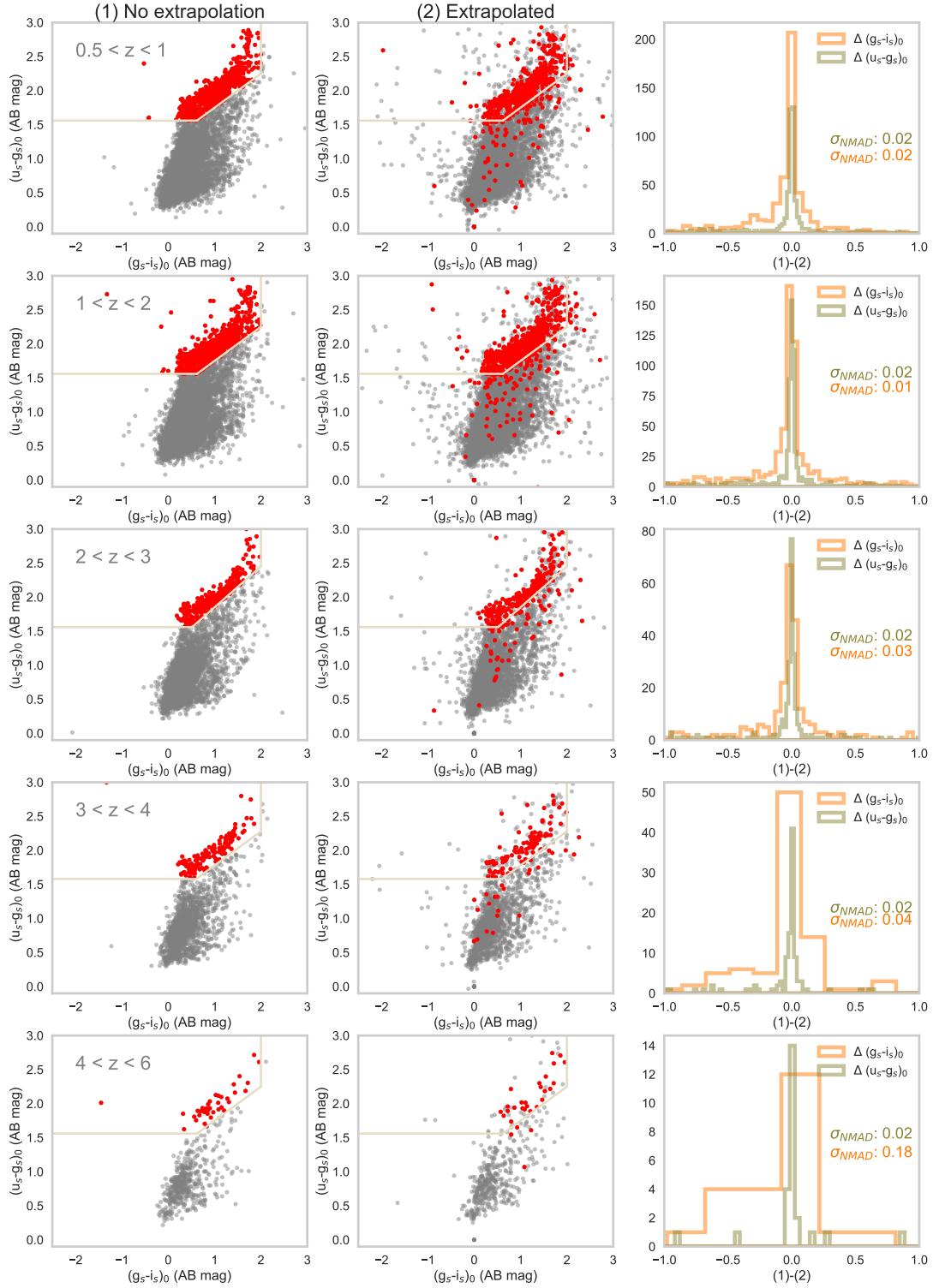


Figure 6. Similar to Figure 5, but showing the impact of extrapolation on synthetic $(ugi)_s$ -selected quiescent galaxies at $0.5 < z < 6$. The $(ugi)_s$ filters are more resistant than UVJ to the effects of extrapolation, particularly at $z > 4$. The scatter in the distribution of the change in $(g_s - i_s)_0$ (Right panels in each row) increases only slightly with redshift and is never higher than $\sigma(\Delta[g_s - i_s]) = 0.2$ mag. This is significantly improved compared to the effects of extrapolation on the UVJ colors. NB: the limits of the abscissa of the right-most panels are the same in all rows and the same as that of the first three rows of Fig 5.

The availability of mid-IR data with the *James Webb Space Telescope* will not solve this problem. At the beginning of its mission, *JWST* will carry out a number of General Observer (GO), Guaranteed Time Observations (GTO), and Early Release Science (ERS) surveys that are aimed at studying galaxies at high redshift. The largest programs (in order of decreasing survey area) are COSMOS-Web (Kartaltepe et al. 2021), Public Release IMaging for Extragalactic Research (PRIMER; Dunlop et al. 2021), the *JWST* Advanced Deep Extragalactic Survey (JADES; Williams et al. 2018), and the Cosmic Evolution Early Release Science survey (CEERS; Finkelstein et al. 2017). Together, these programs cover the five standard extragalactic fields (COSMOS, UDS, GOODS-N, GOODS-S, and EGS). Although they all include MIRI imaging, mid-IR data will be available for only a small portion of the surveyed area, due to the instrument's relatively small field of view. COSMOS-Web and PRIMER, which together survey COSMOS and UDS, will have MIRI coverage for up to $\simeq 30\%$ and $\simeq 60\%$ of their planned survey areas, respectively (noting that the survey area of PRIMER is only 1/6 that of COSMOS-Web). Additionally, the majority of planned MIRI observations will be shallow, i.e. 5σ depth $\lesssim 24$ AB mag. For reference, the candidate $z = 5$ quiescent galaxy in Figure 4 has $m_{IRAC4} \approx 25$ AB mag. The latter two surveys, which cover GOODS-N, GOODS-S, and EGS, will have MIRI coverage for 7.4% and 25% of their total area, respectively. This means that we will have adequate MIRI coverage for only two out of the five extragalactic fields, and consequently a small fraction of galaxies surveyed. As such, extrapolation will continue to be a problem for the majority of galaxies at $z > 3$, even with *JWST* data.

We illustrate this with a candidate quiescent galaxy in the UltraVISTA catalog with a photometric redshift at $z_{ph} \approx 5$ (*JWST* GO Program #2362, Marsan et al. 2021). This galaxy offers an excellent example because it includes many high S/N photometric points from the UV-to mid-IR, such that the true SED is well-represented by the data. We simulate extrapolation by removing photometric points corresponding to the rest-frame *J*-band (and redder) at that redshift. We then re-derive the photometric redshift and rest-frame colors from the data using eazy-py (using the same method described in Section 2.3). We determine the rest-frame *UVJ* fluxes and corresponding uncertainties in the “non-extrapolated” case (where we use all available photometric bands) and the “extrapolated” case (where we have excluded bands corresponding to the rest-frame *J*-band or redder, specifically IRAC ch3 and 4 in this case). This latter case mimics the effects of extrapolation on the derived rest-frame *UVJ* fluxes, shown in Figure 4. We also show the rest-frame colors derived in synthetic top-hat ($R = 10$) filters spanning $0.5 - 8\mu\text{m}$ for both cases.

Extrapolation (removing data points covering the rest-frame *J*-band and redder) has a large impact on the derived rest-frame *J*-band flux (though it does not significantly change the redshift, which is constrained by the data near the Balmer/4000 Å break). In this case, the rest-frame *J*-band flux is changed by a factor of 2, causing the $V - J$ color to become redder by ≈ 0.5 mag. This moves the galaxy out of the quies-

cent region of the *UVJ* diagram. We repeat this experiment with the $(ugi)_s$ filters (removing observed data corresponding to the i_s -band and redder) and find that the $(g_s - i_s)_0$ color, in comparison, is only affected by 0.08 mag, indicating that the $(ugi)_s$ diagram is more robust against extrapolation and hence better suited to select quiescent galaxies at $z > 4$ with incomplete data.

Our findings with COS-130302 beg the question: how often do studies misclassify galaxies as quiescent using the *UVJ* diagram due to extrapolation? We investigate this by repeating the experiment above on a SNR-limited sample (K_s SNR ≥ 5) of galaxies at $0 < z < 6$ from the 3D-HST survey. This is illustrated in Figure 5. We quantify performance by measuring the normalized median absolute deviation (σ_{NMAD} , Beers et al. 1990, Brammer et al. 2008) of the change in rest-frame colors ($\Delta(U - V)_0$ and $\Delta(V - J_0)$) of quiescent galaxies in the “non-extrapolated” (using all data) and “extrapolated” cases (removing any observed data corresponding to and redward of the rest-frame *J*-band).

UVJ selection performs very well out to $z \sim 2$. For $z \lesssim 2$ the difference between non-extrapolated and extrapolated colors of quiescent galaxies is small, with $\sigma[\Delta(V - J_0)] \lesssim 0.1$ mag. At $z > 2$, the scatter increases to ~ 0.3 mag, and it becomes much worse at $z > 4$, reaching $\sigma[\Delta(V - J_0)] \sim 1$ mag at $4 < z < 6$. ($U - V$)₀ colors also start to become extrapolated at $z \sim 2$, where the rest-frame *V*-band enters the observed *H*-band, which is only detected at SNR > 3 for up to $\sim 45\%$ of galaxies at $z \gtrsim 2$ (estimated for Figure 3, not shown). This likely explains the factor of 2 increase in σ_{NMAD} (0.05 – 0.1) from $1 < z < 2$ to $2 < z < 3$ in ($U - V$)₀.

In contrast, the effects of extrapolation on the synthetic rest-frame $(ugi)_s$ colors are much lower, even at high redshift. Figure 6 shows the results. The change in $\Delta(ugi)_s$ colors remains small between the extrapolated and fiducial cases, with $\sigma[\Delta(ugi)_s] \leq 0.2$ mag at all redshifts, even at $4 < z < 6$.³ Therefore, we conclude that extrapolation of the rest-frame *J*-band can have a serious impact on the selection of quiescent galaxy candidates at high redshifts ($z > 4$). This is especially important as the quiescent fraction is expected to decline sharply at this epoch (e.g., Merlin et al. 2019) such that these types of systematics of selection may dominate searches for such objects (see discussion in Section 5 below).

We note that these results are an illustration of the effects of extrapolation on the selection of high redshift quiescent galaxies in the specific case where rest-frame colors are derived semi-empirically. When they are estimated solely from the best-fit SED, as is typically done in the literature, the effects of extrapolation on a single galaxy are varied and dependent on many assumptions, including its intrinsic SED shape, SNR, redshift, the assumed star formation history, and the template sets and parameters used for SED-fitting. Our experiments

³The effects of extrapolation on $U - V$ color may be reduced with *JWST*/NIRCam data because unlike with ground-based telescopes, a number of narrow, medium, and broad-band filters exist between the *K*-band and IRAC channel 1. Nevertheless, these effects will likely persist for $V - J$ colors.

demonstrate the need for an updated color-selection method for high redshift galaxies. This will continue to be necessary in the *JWST* era. We have demonstrated that the effects of extrapolation on our proposed synthetic $(ugi)_s$ filters will be minimal out to $z \sim 6$.

In the remaining sections, we show that the synthetic $(ugi)_s$ filters match *UVJ* in simplicity and discrimination efficiency, making $(ugi)_s$ a viable replacement.

5. PURITY TESTS AND DISCUSSION

5.1. Tests of Completeness and Contamination using Color-Color Selection

Color-color selections involve a balance between sample purity and completeness. Generally, the higher the completeness, the higher the number of contaminants and vice versa. This tradeoff becomes increasingly important for quenched galaxies at higher redshift ($z > 3$), where features such as the Balmer/4000 Å break, which are key for determining photometric redshifts, are shifted into the near-infrared, a regime prone to numerous systematics. Additionally, galaxies at those redshifts generally have lower signal-to-noise photometry, which can decrease the accuracy of rest-frame colors. As a result, dusty star forming galaxies⁴, which are the primary contaminants of quiescent galaxies, are more likely to scatter into the quiescent region. Furthermore, galaxies with strong nebular line emission (e.g., strong [O III] emission) are another source of contamination, and such sources comprise a greater fraction of the star-forming population at these redshifts (Endsley et al. 2021). Efficient color-color selection of galaxies at these redshifts therefore requires careful culling of such objects in order to maintain high completeness while lowering the contamination rate. In this section, we contrast the performance of the $(ugi)_s$ diagram with that of ugi and *UVJ* as it pertains to identifying and removing such interlopers.

One major challenge is evaluating “truth” in the purity and contamination of quiescent-galaxy searches. Several approaches have been taken in the literature, including using estimates of galaxy star formation rates (and specific SFRs) (e.g. Heinis et al. 2014, Chang et al. 2015, Tasca et al. 2015, Tomczak et al. 2016, Pacifici et al. 2016). Others identify active galaxies based on emission line strength (e.g. Rodighiero et al. 2014, Belfiore et al. 2018) or far-IR emission (e.g. Whitaker et al. 2014, Lee et al. 2015, Ilbert et al. 2015, Schreiber et al. 2015, Erfanianfar et al. 2016). Speagle et al. (2014) and Popesso et al. (2019) showed that the slope of the SFR-M* plane, commonly referred to as the star-forming main sequence (SFMS), changes by up to 1 dex depending on which of these methods is used. Similar results were found with the Illustris-TNG simulations, with quies-

cent fractions on the higher mass end ($\log(M_*/M_\odot) > 10.5$) varying between 20 – 40% (Donnari et al. 2019). Moreover, these thresholds only correspond to a “relative” quiescence because they fail to account for the evolution in the intrinsic stellar populations of quiescent galaxies over cosmic time, making them applicable only to a specific redshift or stellar mass bracket. Additionally, while *UV + IR* SFRs, which some regard as the “gold standard”, are available for our dataset at $z < 4$, these can be problematic for a number of reasons. This is because IR luminosities (typically estimated from the 24 μm flux) tend to overestimate SFRs by up to a factor of 4, because they do not account for heating from intermediate age and old stars, which tends to dominate the IR budget for $\log(\text{sSFR}) < -13$ (Martis et al. 2019, Leja et al. 2019c, Belli et al. 2017, Fang et al. 2018). Simulations also report similar findings, i.e. that UV luminosities can be overestimated by 0.2 – 0.5 dex at $z = 1 - 4$ for galaxies with $\log(M_*/M_\odot) \approx 8.5 - 10.5$ (Katsianis et al. 2020) and that the total IR luminosity is not a good proxy for dust-obscured star formation for rapidly-quenched or “post-starburst” galaxies (Roebuck et al. 2019). In summary, the definition of what constitutes “truth” in quiescent-galaxy selection matters.

To mitigate this effect, we choose a data-driven approach. In order to determine which galaxies are truly quiescent, we take SFR and stellar mass estimates for galaxies from two catalogs that account for these systematics by self-consistently modeling the entire SED: the Leja et al. (2019c) measurements from Prospector- α for 3D-HST galaxies at $0.5 < z < 2.5$, and the Martis et al. (2019) measurements from MAGPHYS for galaxies in UltraVISTA at $z = 3 - 4$. Leja et al. (2019c) use Bayesian modeling methods with flexible star-formation histories (SFHs) that allow for more accurate estimates of galaxy SFRs (derived from panchromatic data, i.e. UV – mid-IR). These estimates are appropriate for galaxies out to $z \lesssim 3$ as they are based on galaxies selected in the *H*-band (and at higher redshifts, the Balmer/4000 Å break shifts beyond the *HST/WFC3 H*-band coverage). The catalogs of Martis et al. (2019) are based on UV – far-IR SED modeling using a rising SFH that accounts for stochasticity. For massive galaxies ($\log M_*/M_\odot > 10.5$) these provide useful constraints on distinguishing star-forming and quiescent galaxies at $z > 3$. Both catalogs include the effects of emission lines on the rest-frame colors, which allows us to also test contamination from sources with strong emission lines. We described these catalogs in detail in Section 2.

We then define regions in the specific SFR (sSFR)–stellar mass plane that separates star-forming galaxies from those that are quiescent. We identify a “ridgeline” i.e. the density peak in the SFR-mass plane, which corresponds to star-forming galaxies (Renzini & Peng 2015). At lower sSFRs lies the “green valley”, below which lie the quiescent galaxies (Schawinski et al. 2014, Sherman et al. 2021). The advantage of this approach is that it does not depend on any pre-separation of quiescent and star-forming galaxies, because the star-forming sequence is directly measured from the data.

We then divide the data into five redshift bins from $0.5 < z < 4$ and determine its 3D distribution by dividing the data

⁴ These can either be dusty galaxies *at the same redshift* as the quiescent galaxies or at lower redshift whose colors mimic the colors of high-redshift quiescent galaxies. Most of the dusty contaminants, however, will fall into the former category. The placement of the $(ugi)_s$ filters along the Balmer/4000Åbreak will help reduce the effect of the latter (see Section 3).

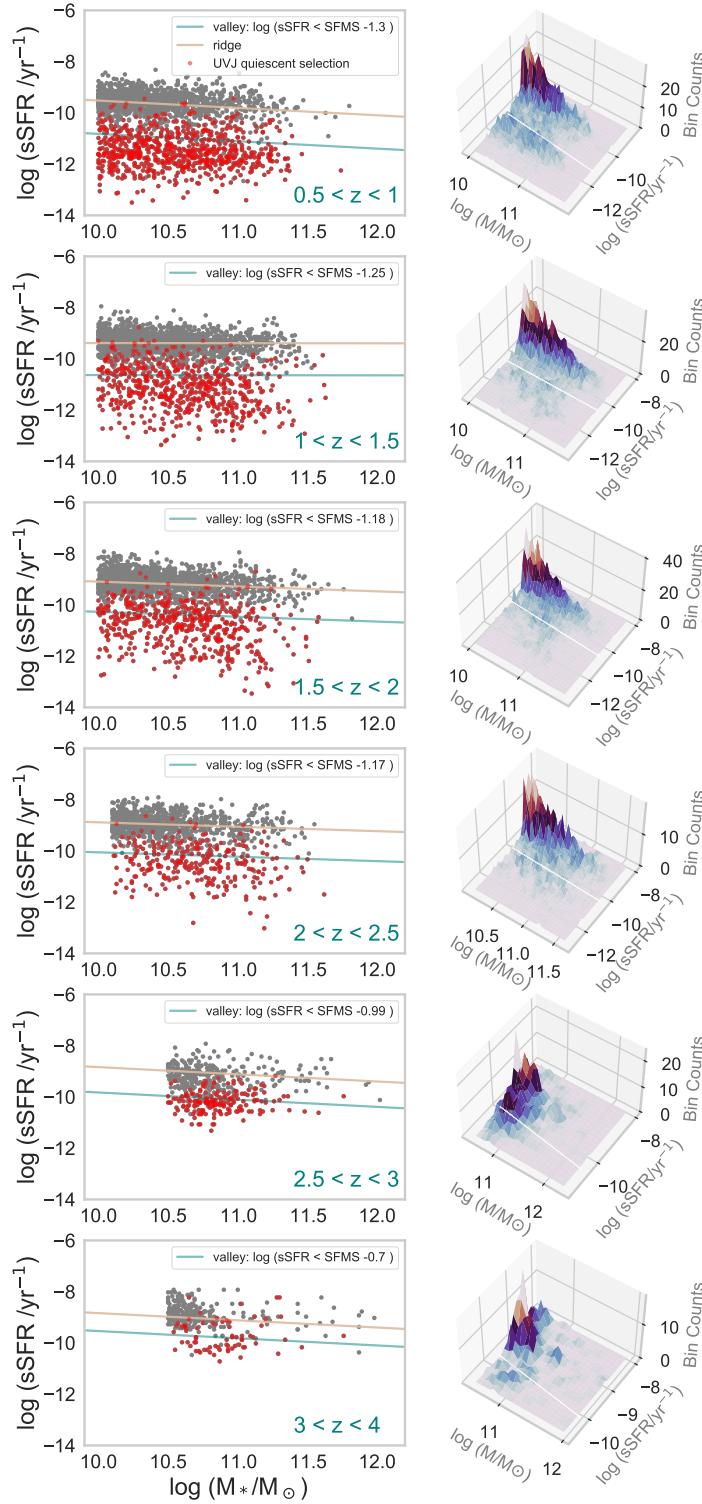


Figure 7. Selection of quiescent galaxies based on specific star formation rate (sSFR). Each row shows the distribution of sSFR as a function of stellar mass in bins of redshift (low-to-high, top-to-bottom). In each row, the **Left** panel shows the selection of quiescent galaxies by identifying the density peak of the star forming main sequence (SFMS) at each redshift and the green valley between the red and blue sequences, which we define as where the data falls below 10% of the peak. Because our galaxy samples at $z > 2.5$ are small, we use the slope of the SFMS at $2 < z < 2.5$ for the two highest redshift bins. The mass range at each redshift reflects the 90% stellar mass completeness limit of the 3D-HST ($0.5 < z < 2.5$) and UltraVISTA ($2.5 < z < 4$) surveys. The **Right** panel in each row then shows a 3D projection of the SFMS at each redshift. The white line shows the estimated location of the green valley between the red and blue sequences.

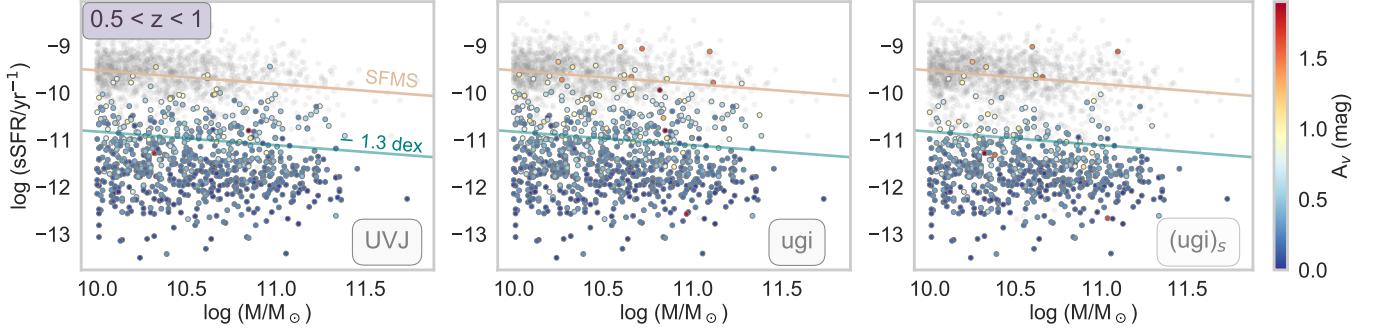


Figure 8. Demographics of quiescent galaxies selected in each method at $0.5 < z < 1$ in 3D-HST. The colored points are the galaxies identified as quiescent in UVJ , ugi , and $(ugi)_s$. We color code each sample by dust extinction from *Prospector- α* . Gray points mark galaxies outside of the quiescent region. The “green valley” (teal line) and peak of the star forming main sequence (brown line) are the same as those from Figure 7 at this redshift. Due to a slightly smaller baseline than UVJ , ugi and $(ugi)_s$ select a few galaxies with high dust obscuration ($A_V \geq 1$), as these tend to lie closer to the diagonal line in these color diagrams than in UVJ . The majority of false positives (galaxies above the teal line) have moderate dust extinction values (median $A_V \approx 0.5$) and low specific star formation rates ($< -10 \text{ yr}^{-1}$). $(ugi)_s$ -selected quiescent samples have less contamination from these dusty galaxies than UVJ and ugi .

into 0.3 dex square bins. In each redshift bin, we define the green valley by determining where the data falls below 10% of the star-forming density peak. In Figure 7, we show the two- and three-dimensional SFR-M* plane as well as the green valley, denoted by the white line. The mass range covered by the data reflects the 90% completeness limit at each redshift (taken from Skelton et al. 2014 for 3D-HST and Martis et al. 2019 and Marsan et al. 2022 for UltraVISTA). UVJ -quiescent galaxies are shown in red. We define “true” quiescent galaxies as those with a sSFR below the estimated location of the “green valley” using the definition above.

5.1.1. Results at $z \leq 4$

In Figures 8 and 9, we contrast the performance of the $(ugi)_s$ diagram with those of ugi and UVJ . Figure 8 shows that our definition of quiescence correctly selects galaxies with low sSFRs and low dust extinction (as a function of redshift), and that majority of the contaminants are dusty galaxies with low sSFRs. In Figure 9, we quantify this performance, also showing results using UVJ -selection criteria from Williams et al. (2009) and Whitaker et al. (2011). Completeness here is defined as the number of true positives (TP; galaxies with specific SFRs below the threshold of the green valley in Figure 7) in the quiescent region divided by the total number of true positives in the sample. By this definition, the false negative fraction (i.e. the number of true quiescent galaxies missed) is $(1 - \text{Completeness})$. The contamination rate is the number of selected false positives (FP; galaxies that are selected to be quiescent, but have specific SFRs above the green-valley threshold in Figure 7) divided by the total number of selected galaxies. Thus, the sample purity is $(1 - \text{Contamination})$. We also evaluate the ratio of true-to-false positives (TP/FP) as an additional criterion of efficiency. Note that our definitions of completeness and contamination have different denominators. This is why the contamination rate in Figure 9 is not 50% at $z = 3$ even though the UVJ TP/FP ratio is 2.

We see that the $(ugi)_s$ diagram, although comparable in completeness to UVJ , has significantly less contamination, hence a much higher TP/FP ratio as a function of redshift. A slightly smaller wavelength baseline in $(u-g)$ and $(g-i)$ means less quiescent galaxies overall are selected in $(ugi)_s$ and ugi (hence $(ugi)_s$ is $\sim 5\%$ less complete) at lower redshifts. At higher redshifts ($z > 2.5$), we start to see the effects of extrapolation and increased contamination from strong [O III] emitters on UVJ and ugi , problems which the $(ugi)_s$ filters are designed to fix. By $z = 3.5$, the UVJ selection has a TP/FP ratio of ~ 1 , which means that *the contaminants in the UVJ-selected quiescent sample are just as numerous as the true quiescent galaxies*.

SFRs and stellar masses from SED-fitting codes have a persistent 0.5 and 0.2 dex scatter (see Section 2.1.1), respectively. While the flexible SFHs and self-consistent SED modeling employed in *Prospector- α* and MAGPHYS reduce systematics and hence provide an improved ability to mitigate this effect, it is still worth noting. In order to investigate the effect of this offset on our derived completeness and contamination fractions, we repeat our analysis above using SFRs and stellar masses in the 3D-HST catalogs derived using FAST, which uses more a simplistic parameterization of the SFH (usually exponential or “delayed exponential” models) and tends to be the traditional choice for estimating stellar population parameters in the literature. Figure 9 (bottom panel) shows the results from these tests. All three methods (UVJ , ugi , $(ugi)_s$) using FAST have lower completeness rates. We attribute this to the adopted star formation history, which has been shown to produce highly biased stellar population parameters (e.g. Carnall et al. 2019), favoring young stellar populations, hence higher sSFRs. This particularly impacts red galaxies (which could be either dusty star-forming galaxies or quiescent galaxies) at high-redshift, for which this SFH has been shown to be inadequate (see Section 2.1.2). Nevertheless, the effect

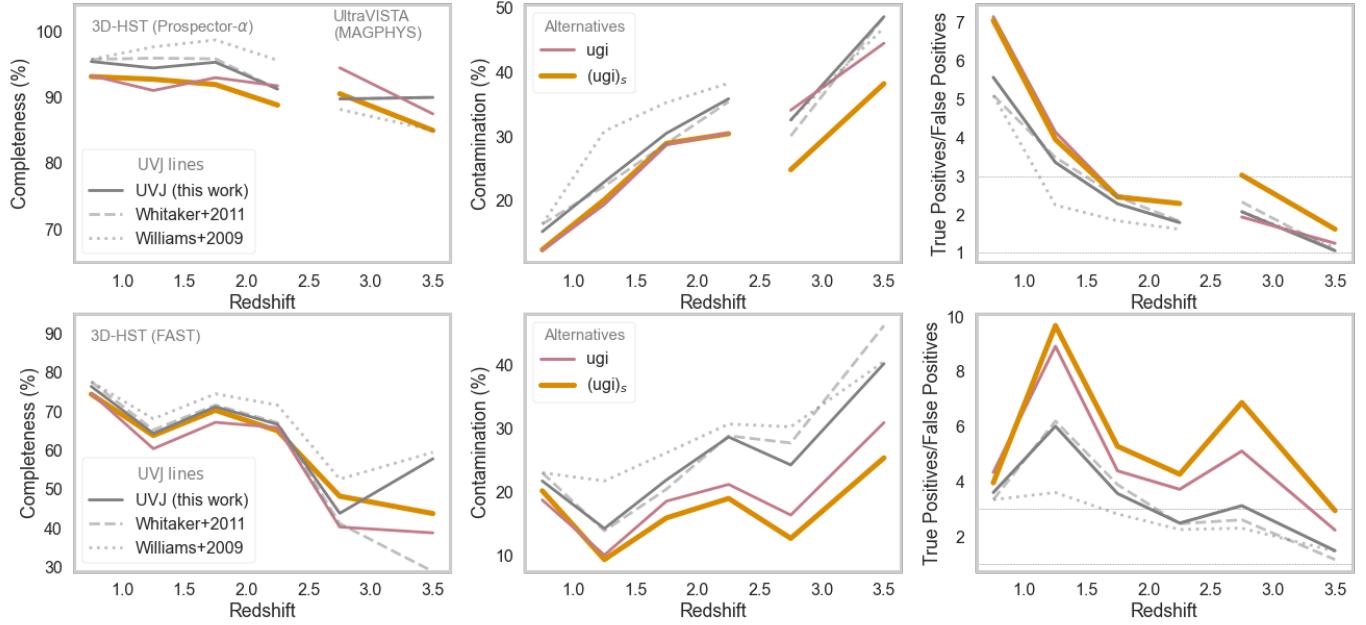


Figure 9. Purity tests of color-color selected quiescent galaxies. The **Top** row shows results using *Prospector- α* fits to the 3D-HST catalogs (Leja et al. 2019c) for galaxies at $z < 2.5$ and the UltraVISTA/MAGPHYS catalog (Martis et al. 2019) for galaxies at $z > 2.5$. The **Bottom** row shows results using SFRs and stellar masses from FAST, which is the traditional choice in the literature. In each row, the columns show the completeness (**Left**), contamination (**Middle**), and true-to-false positive ratio (TP/FP) for quiescent galaxies selected by three color-color methods (*UVJ*, *ugi*, and $(ugi)_s$, as labeled). In all cases we define true quiescent galaxies based on specific SFRs (i.e. by defining the star-forming main sequence in each redshift bin; Section 5). In all cases we find that the different methods produce similar completeness. However, the contamination of $(ugi)_s$ is superior to the other color-color selections. For this reason the TP/FP ratio is highest for $(ugi)_s$. For example, the TP/FP rate for *UVJ* selection dwindles to ~ 1 by $z = 3.5$ (where there is an equal number of false positives for each true positive detected). In contrast, $(ugi)_s$ -selected samples have a TP/FP ratio that is nearly a factor of 2 higher.

appears to be systematic, affecting all color-color methods similarly.

Our results with FAST mirror those derived from *Prospector- α* and MAGPHYS above: the completeness of quiescent galaxies is similar for all three color-color methods (left column of Figure 9). However, the contamination in the quiescent selection is consistently lower using the $(ugi)_s$ colors at $z > 3$. Finally, the $(ugi)_s$ color-selection of quiescent galaxies has the highest TP/FP ratios (right column of Fig 9). This is primarily a result of the lower contamination from the $(ugi)_s$ selection, and for this reason we see the largest improvement in TP/FP at $z > 3$. This fact will be crucial at these redshifts as contaminants (false positives) can otherwise exceed true positives, as in *UVJ*.

Our results for the completeness and contamination of *UVJ*-selected quiescent galaxies at $z \leq 4$ are broadly consistent with similar results in the literature using photometric data from large area surveys, spectroscopic campaigns, and simulations. Spectroscopic studies at $z < 1$ have shown that *UVJ*-selected quiescent galaxies at various mass cutoffs show $\sim 10 - 65\%$ contamination from galaxies with signs of ongoing star formation (e.g. Tacchella et al. 2022). Using a sample of \log

$(M_*/M_\odot) > 10$ galaxies at $1.5 < z < 2.5$ in 3D-HST, Leja et al. (2019b) find that galaxies with the lowest specific-SFRs ($\log(\text{sSFR}) < -10.5$) show up to $\sim 30\%$ contamination. Using CANDELS, Shahidi et al. (2020) find 80% completeness for $\log(M_*/M_\odot) > 10$ galaxies at $2.8 < z < 4$ and 50% completeness at $z = 4 - 5.4$. Using SHELA (Papovich et al. 2016, Wold et al. 2019), Sherman et al. (2021) find up to 38% contamination from dusty galaxies for $\log(M_*/M_\odot) > 11$ quiescent galaxies at $1.5 < z < 3$. Finally, Díaz-García et al. (2019) find a 20% contamination rate at $z \sim 1$ from dusty galaxies using data from the ALHAMBRA survey (Moles et al. 2008). Similarly, for the simulations, Lustig et al. (2022) find 67-75% completeness and 50-60% contamination for $\log(M_*/M_\odot) > 11$ quiescent galaxies at $z = 2.7$ in Magneticum⁵ (Steinborn et al. 2016) and Illustris TNG⁶ (e.g. Springel et al. 2018).

Lastly, we would like to caution the reader on the importance of calibrating color selection methods on the dataset which they choose to use for their study. A number of stud-

⁵ www.magneticum.org

⁶ www.tng-project.org

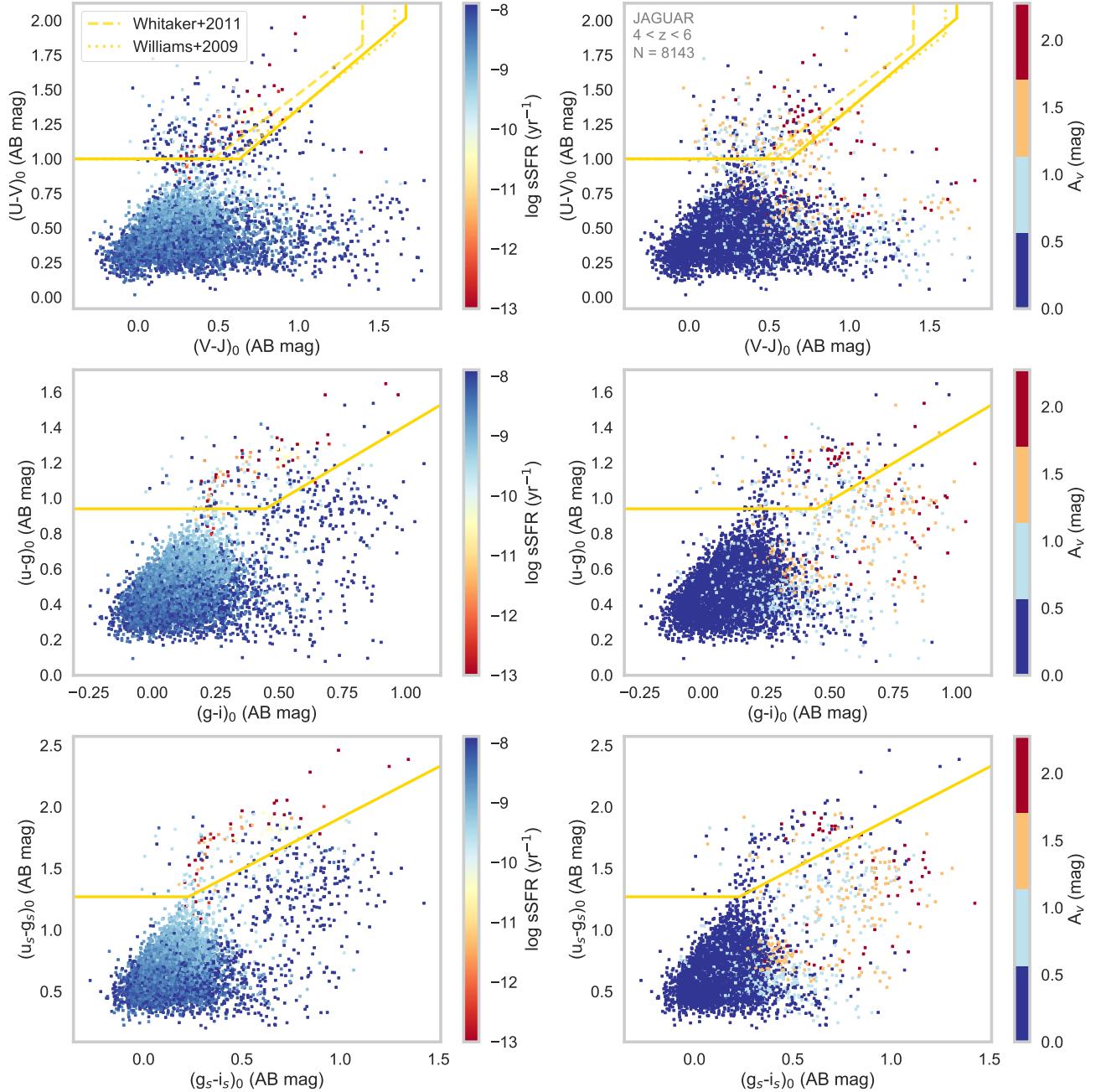


Figure 10. Color-color diagrams for simulated data with added noise at $4 < z < 6$. Each row shows color-color diagrams for galaxies with $\log(M_*/M_\odot) \geq 8.9$ in the JAGUAR catalog, color-coded by specific star formation rate (left panel in each row) and dust extinction (right panel in each row). The **Top** row shows the UVJ colors. The **Middle** row shows the ugi colors. The **Bottom** row shows the $(ugi)_s$ colors. In each case, the rest-frame colors were derived using *HST*, *JWST/NIRCam* and *Spitzer/IRAC* photometry from $0.4-5\mu\text{m}$. The $(U - V)_0$, $(u - g)_0$, and $(u_s - g_s)_0$ horizontal lines were each lowered by 0.23 mag to accommodate the increasingly younger population of quiescent galaxies at these redshifts, although we maintained the slopes and zeropoints from Figure 2.

ies, including Kawinwanichakij et al. (2016) have pointed out that there are systematic offsets in the rest-frame colors of galaxies at fixed mass and redshift in different surveys. This means that quiescent selection lines defined on different datasets will yield different numbers of quiescent galaxies, resulting in different completeness and contamination rates. We demonstrate this by applying the Whitaker et al. (2011) and Williams et al. (2009) *UVJ* lines to our dataset and reporting their completeness and contamination rates in Figure 9. Although the Whitaker et al. (2011) line appears to be quite similar to ours, the results for the Williams et al. (2009) lines are very different, yielding up to 10% more quiescent galaxies and contaminants. We also determined during some earlier tests that the Muzzin et al. (2013) lines (not shown in Figure 9) yield up to 20% less completeness and 5% less contamination than the other *UVJ* lines applied to our dataset. For these reasons, although we provide $(ugi)_s$ lines, we urge the reader to treat them as initial selections and calibrate the color selection lines on their own dataset in order for the method to be most accurate.

5.1.2. Results at $z \geq 4$: What to expect with JWST

Using the same methodology and definitions detailed above, we determine the completeness and contamination of the quiescent selections of *UVJ*, *ugi*, and $(ugi)_s$ in simulated JWST data using the JAGUAR catalog (Williams et al. 2018). We do this using a sample of 10,200 galaxies split into two redshift bins: $3 < z < 4$ and $4 < z < 6$. The latter bin is large because there are only a handful of quiescent galaxies in the catalog at these redshifts (TP = 139; Figure 10). For the $z > 4$ bin, we lower the $(U - V)_0$ line by 0.23 mag to include more post-starburst galaxies in the quiescent region, because there is evidence from both photometric and spectroscopic samples at high- z that these galaxies tend to have much lower $U - V$ colors than the quenched population (Marsan et al. 2015, Carnall et al. 2020, Marsan et al. 2022, Forrest et al. 2020b, Schreiber et al. 2018, discussed further in Section 5.3.2). We lowered the horizontal cut by 0.23 mag to match the $U - V$ colors of the young quiescent galaxies in these studies (i.e. $(U - V)_0 \sim 1$). In the same vein, we lower the horizontal lines of the other color selection methods (*ugi*, $(ugi)_s$) by the same amount.

The completeness, contamination, and true-to-false positive rates of the selection methods, *UVJ*, *ugi*, $(ugi)_s$ are shown as a function of redshift from $z = 3.5$ and $z = 6$ in Figure 11. As in Figure 9, although the color selection methods are comparable in completeness at these redshifts, $(ugi)_s$ has a much lower contamination rate (up to $\sim 30\%$ less) than *UVJ* and *ugi*, resulting in a higher true-to-false positive rate ($\gtrsim 2.5$). Conversely, *ugi* and *UVJ* have true-to-false positive rates of $\lesssim 1$ by $z = 6$, indicating that the true quiescent galaxies are outnumbered by contaminants, making these methods unfeasible for selecting quiescent galaxies at these redshifts. For this reason, the $(ugi)_s$ -selection method provides improved fidelity in color-selected samples of quiescent galaxies at these high redshifts. It is worth noting that these numbers represent the best case scenario, as the number of contaminant dusty star forming galaxies may be higher in reality, since they tend

to be underrepresented in UV-selected samples like JAGUAR (Appendix A).

5.2. Contamination from Extreme Emission-Line Galaxies in Quiescent Galaxy Selection

Here, we test the ability of each color selection method to minimize the number of galaxies with strong emission lines in the quiescent region. Extreme emission line galaxies (EELGs) tend to have $EW_0 \sim 100 \text{ \AA} - 1000 \text{ \AA}$ in $[\text{O II}] \lambda\lambda 3726, 3729, \text{H}\beta + [\text{O III}]$ and $\text{H}\alpha + [\text{N II}]$ although a handful of massive galaxies have been discovered at $z > 7$ with $EW(\text{H}\beta + [\text{O III}]) \approx 1000 - 2000 \text{ \AA}$ (Smit et al. 2014, 2015, Roberts-Borsani et al. 2016, Castellano et al. 2017). At $z \sim 0.1$, EELGs correspond to rare, starbursting, low-mass galaxies ($\log(M_*/M_\odot) \sim 8 - 9$) where the mass doubling time can be 100 Myr (e.g., Izotov et al. 2016). These galaxies are 10–100× more common at $z \gtrsim 1$ (van der Wel et al. 2011, Maseda et al. 2018) than at $z \leq 0.5$ and may be the dominant star-formation mode at $z = 3 - 4$ (Kurczynski et al. 2016, Cohn et al. 2018, Tran et al. 2020).

Galaxies with strong emission lines can mimic the rest-frame *UVJ* colors of galaxies with old stellar populations by boosting their photometry and making them appear redder in $(U - V)_0$ than they actually are (e.g. Labb   et al. 2005). Strong emission lines can be a result of either ongoing star formation or AGN activity, both of which are observed to increase at higher redshifts (e.g. Sobral et al. 2014, D'Eugenio et al. 2020, Marsan et al. 2022). $\text{H}\beta + [\text{O III}]$ emission with $EW \leq 200 \text{ \AA}$ has been shown to boost fluxes in the *K*-band⁷ by 5–30% (Schreiber et al. 2018, Forrest et al. 2020a) and in the IRAC bands by up to 65% (Labb   et al. 2013). Based on their central wavelengths and widths, the synthetic u_s and i_s filters require extremely strong lines (rest-frame $EW > 800 \text{ \AA}$ in $[\text{O II}]$ and $\text{H}\alpha + [\text{N II}]$ respectively) to be contaminated. The synthetic g_s filter, however, has only a 163 \AA buffer from the central wavelength of $\text{H}\beta \lambda 4861$. This implies that it can be contaminated by galaxies with $\text{H}\beta + [\text{O III}] > 500 \text{ \AA}$. While the fraction of star-forming galaxies with $\text{H}\beta + [\text{O III}] EW > 500 \text{ \AA}$ is low at $z = 1.7 - 2.3$ (5.4 – 8.2%, Boyett et al. 2022), it increases to $\sim 40\%$ at $z = 5$ (Rasappu et al. 2016⁸) and could reach $\sim 50\%$ at $z = 6$, assuming that the rest-frame optical emission evolves as a power law, $(1+z)^P$, and given that fractions at $z \sim 7$ are measured to be $\sim 65\%$ (Endsley et al. 2021). Consequently, we can expect an increasingly larger number of EELGs to contaminate the quiescent region at higher redshifts by boosting the fluxes in the reddest observed photometric bands (*K*-band – IRAC channel 4).

We quantify this contamination for *UVJ*, *ugi* and $(ugi)_s$ using mass-complete samples from 3D-HST and JAGUAR at $1.2 < z < 2.3$ and $4 < z < 6$, respectively. The 3D-HST line measurements are from *HST/WFC3 G141* grism observations (Momcheva et al. 2016) and those from the simulated JAGUAR catalog are determined using the BEAGLE SED fit

⁷ Originally shown by Shapley et al. 2005.

⁸ converted from $EW(\text{H}\alpha)$ by Boyett et al. 2022

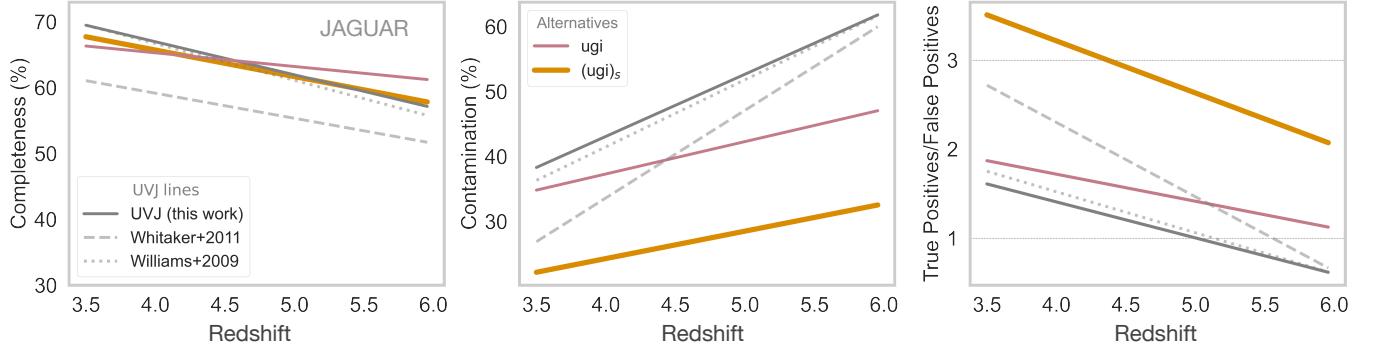


Figure 11. Sample selection efficiency of UVJ , ugi , and $(ugi)_s$ at $3 < z < 6$. We show the completeness and contamination of the quiescent selections of UVJ , ugi , and $(ugi)_s$ using specific star formation rates from the JAGUAR catalog, which were generated using BEAGLE. As in Figure 9, although the color selection methods are comparable in completeness at these redshifts, $(ugi)_s$ has a much lower contamination rate (up to $\sim 30\%$ less) than UVJ and ugi , resulting in a higher true-to-false positive rate at $z = 6$ (≥ 2.3). Conversely ugi and UVJ have true-to-false positive rates of $\lesssim 1$ by $z = 6$, indicating that the true quiescent galaxies are outnumbered by contaminants, making these methods unfeasible for selecting quiescent galaxies at these redshifts.

ting code (Williams et al. 2018). In Figure 12, we show the fraction of galaxies in the quiescent region with $H\beta+[O\ III]$ equivalent widths $> 500 \text{ \AA}$ for all three color selection methods. For the $EW > 500 \text{ \AA}$ population at both redshift ranges, although the methods have similar fractions of EELGs in their quiescent regions, the $(ugi)_s$ diagram has a smaller number of them (a factor of 1.6 lower). However, for the $EW > 300 \text{ \AA}$ EELG population, the contamination in $(ugi)_s$ is lower by a factor of 3. For comparison, $EW > 200 \text{ \AA}$ contamination fractions at $z > 4$ are: 0.38 (UVJ , N=92), 0.22 (ugi , N = 37), 0.11 ($(ugi)_s$, N = 14). Photometric redshift uncertainties could also be a factor (as an erroneous redshift could shift an emission line into the rest-frame band, though this effect will be lessened for galaxies with stronger emission lines as these tend to have lower redshift uncertainties, see Momcheva et al. 2016). Nevertheless, in Appendix B we show that photometric redshift errors of $\sigma_z/(1+z) < 9.1\%$ correspond to fewer than 16% of sources having their rest-frame colors impacted by emission lines (and most modern surveys have better photometric-redshift accuracy than this). Therefore we do not expect this to significantly impact the rest-frame colors. Our results are consistent with spectroscopic surveys, which report a contamination rate of 21–30% from $EW(H\beta+[O\ III]) > 100 \text{ \AA}$ galaxies in UVJ -selected quiescent samples at $z = 3 – 4$, with at least a third of those having $f[O\ III]/f_{H\beta} > 6$ (Schreiber et al. 2018, Forrest et al. 2020a), indicative of AGN activity, particularly at high mass (e.g. Strom et al. 2017, Belli et al. 2017, Reddy et al. 2018).

5.3. Implications for Quiescent Galaxy Selection at $z > 4$: Current and Future

5.3.1. The Beginning of the End

We are approaching the end of the era of discovery of quiescent galaxies. We have spent the past two decades attempting

to obtain a complete census of different populations of galaxies as a function of redshift, and color selection methods have played a pivotal role in this (e.g. Daddi et al. 2004, Kriek et al. 2009, Quadri et al. 2007, Guo et al. 2013, Nayyeri et al. 2014). Color selection methods are useful because they can quickly isolate different galaxy populations in large datasets over a range of redshifts using only 3 bands. Although star-forming galaxies may well exist out to $z \sim 13$ (Harikane et al. 2022), we may be reaching the era where there are no more quiescent galaxies to be selected ($z \sim 6$). We have spectroscopically confirmed quiescent galaxies out to $z \sim 4$ (McCarthy et al. 2004, Cimatti et al. 2004, Marsan et al. 2015, Valentino et al. 2020, Forrest et al. 2020b, Saracco et al. 2020, Tanaka et al. 2019, Glazebrook et al. 2017) and we have a handful of photometrically-selected candidates at $z = 5 – 6$ (e.g. Marsan et al. 2021, Merlin et al. 2019, Mawatari et al. 2016) with searches underway using ground-based surveys that leverage the detection capabilities of medium-band filters (Esdaile et al. 2021). However, given that passive fractions are $< 10\%$ at $z > 4$ (e.g. Santini et al. 2021), by $z \sim 6$ we will be witnessing the emergence of the first quiescent galaxies. These quiescent galaxies will be some of the first to have started forming stars and quench and will be important for answering questions about structure formation in the early universe and settling debates about quenching mechanisms (see Nanayakkara et al. 2022 for a review). With this in mind, it is ever so important that we select pure and complete samples of quiescent galaxies for spectroscopic follow-up so that we do not miss any of this interesting population.

In this Paper, we have presented three synthetic filters: u_s , g_s , and i_s which accomplish this goal most efficiently, using data from our largest and deepest ground-based photometric surveys and simulated *JWST* data. We consider one of our intended goals in designing our filter set: achieving increased

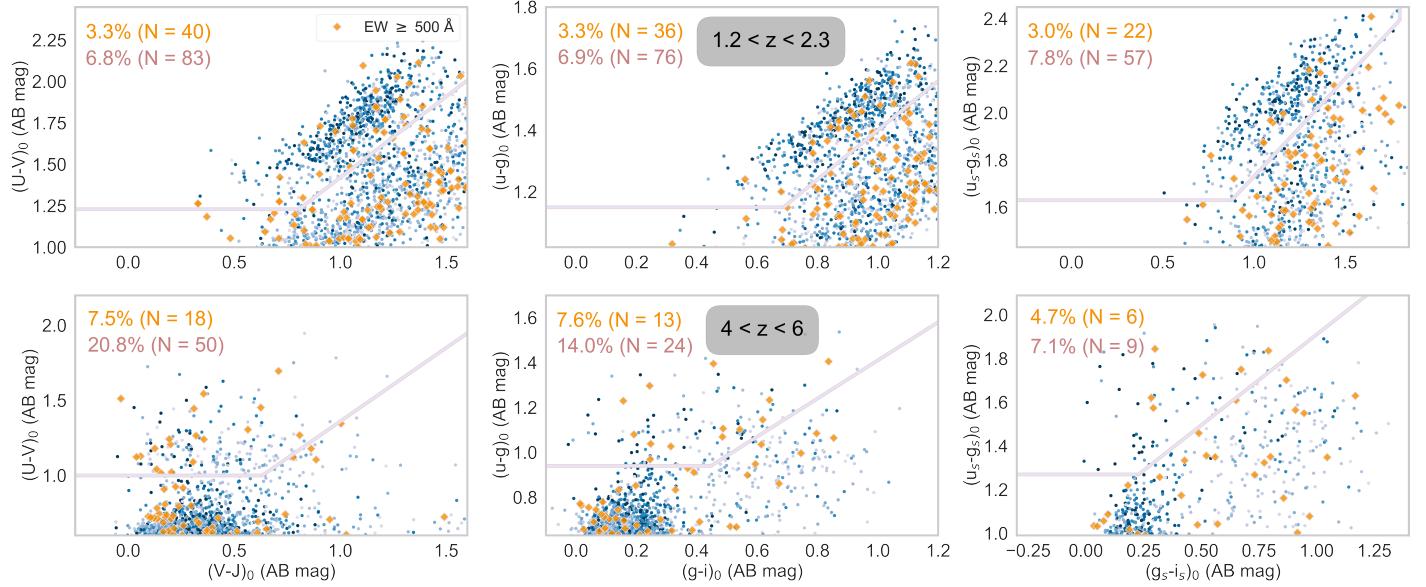


Figure 12. Fraction of extreme emission line galaxies in UVJ , ugi , and $(ugi)_s$ using mass-complete samples from 3D-HST and JAGUAR at $1.2 < z < 2.3$ (**Top**) and $4 < z < 6$ (**Bottom**), respectively. For each redshift bin, each color diagram shares the same line parameters as Figure 10. As in Figure 10, the ordinate limits and horizontal color lines shown are different for each diagram in order to capture the evolution of galaxies in color space as a function of redshift. The fractions and numbers of rest-frame $EW > 500 \text{ \AA}$ galaxies selected as quiescent in each color diagram are shown in orange. For the $EW > 500 \text{ \AA}$ population in both redshift bins, although the methods have similar fractions of EELGs in their quiescent regions, the $(ugi)_s$ diagram has a smaller number of them (a factor of 1.6 lower). We also show the numbers and fractions of $EW > 300 \text{ \AA}$ in brown (not plotted) for comparison. For the $EW > 300 \text{ \AA}$ EELG population, the contamination in $(ugi)_s$ is lower by a factor of 3.

sensitivity to the Balmer/4000 Å break. In Figure 13, we show the distribution of the vertical colors of galaxies in the quiescent regions of UVJ , ugi and $(ugi)_s$ at $1.2 < z < 2.3$ and $4 < z < 6$. The strength of the Balmer break is predicted to be lower at fixed sSFR for galaxies at $z = 2 - 10$ than for their lower redshift counterparts (Shen et al. 2020). If so, we expect the vertical colors in the high-z bin to be bluer than those in the low-z bin and that the most efficient color selection method would have redder vertical colors overall, signaling that the Balmer/4000 Å break produces a stronger color signature in that color space. This is precisely what we see in Figure 13, with median values of $(u_s-g_s)_0 = 1.73$, $(U-V)_0 = 1.19$, and $(u-g)_0 = 1.13$ at $z > 4$. A wider range in $(u_s-g_s)_0$ ($\sigma_{NMAD} = 0.19$) over $(U-V)_0$ (0.15) and $(u-g)_0$ (0.13) at those redshifts may also indicate that $(ugi)_s$ is better at capturing the diversity of quiescent SEDs at higher redshift than UVJ and ugi . If true, this property will be extremely useful for studying the age gradients of the first quiescent galaxies (e.g. Whitaker et al. 2013).

This is particularly important because, so far, most of the quiescent galaxies that have been confirmed spectroscopically at $z > 3$ are young, recently-quenched galaxies, with a few exceptions (e.g. Galaxy D in Kalita et al. 2021). Older quiescent galaxies are predicted to be 1 – 2 magnitudes fainter in the K -band than their post-starburst counterparts at the same redshift, with stronger Balmer/4000 Å breaks and weaker absorption lines due to their evolved stellar populations (Forrest et al.

2020b). This suggests that their absence from current surveys is simply a selection effect, therefore this population should be revealed with deep ongoing NIR surveys e.g. FENIKS (Esdaile et al. 2021) and the upcoming *JWST* ERS, GTO, and GO programs, which will enable us to detect some of the faintest galaxies in the distant universe.

5.3.2. Possible Limitations of the $(ugi)_s$ Diagram

Next, we consider an apparent limitation of $(ugi)_s$, and color-color-selection methods in general at high-z. If indeed $(ugi)_s$ is more efficient at selecting high-z quiescent galaxies, why then is it $\lesssim 70\%$ complete at $z > 4$ (Figure 11)? First, the simplest explanation is that they were scattered out of the quiescent region due to the addition of noise, since that would preferentially affect faint and low mass galaxies. In fact, the completeness of quiescent galaxies is reduced by $\sim 20\%$ for all three methods after we add noise (Appendix A).

A second possibility is that we are missing some quiescent galaxies because our adopted slope for the $(ugi)_s$ diagonal line excludes them. For the sake of consistency, we assumed that the slope of the diagonal line remains constant as a function of redshift, similar to what others have done (Williams et al. 2009, Muzzin et al. 2013, Whitaker et al. 2011). Like those studies, we lowered the horizontal cut to account for the evolution in color due to quiescent galaxies having increasingly younger ages at higher redshifts (see discussion in paragraph below). However, it is not entirely clear if the slope needs

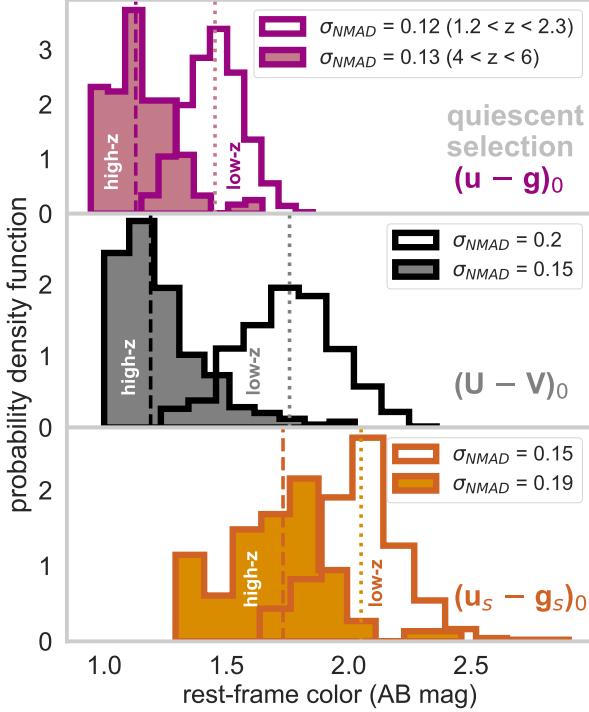


Figure 13. Vertical color of the red sequence in each color selection method at low and high redshift ($1.2 < z < 2.3$ and $4 < z < 6$). The **Top** row shows the results for $(ug_i)_0$. The **Middle** row shows the results for UVJ . The **Bottom** row shows the results for $(ug_i)_s$. We calculate the median and associated normalized median absolute deviation for each histogram to compare the methods with each other. Galaxies at $z < 2.3$ (low-z) tend to exhibit redder colors than those at $z > 4$ (high-z) because the strength of the Balmer/4000 Å break decreases as a function of redshift. The $(u_s-g_s)_0$ color, however, is reddest in both bins, with a median of 1.73 mag (2.05) at $z > 4$ ($z < 2.3$), compared with 1.19 mag (1.76) in $(U-V)_0$ and 0.13 (1.45) mag in $(u-g)_0$. $(ug_i)_s$ also has a larger spread in colors. Both of these indicate that $(ug_i)_s$ is better at isolating quiescent galaxies and capturing the diversity of quiescent SEDs in the high-z Universe.

to evolve as well, particularly at $z > 3$. This can be resolved with larger unbiased samples of quiescent galaxies with high signal-to-noise photometry.

A third possibility is that lowering the horizontal lines of each color diagram by 0.23 mag is not enough to capture all the post-starbursts in the sample. This would be the case for dust-free star-forming galaxies that quenched abruptly (< 175 Myr) prior to observation, as they can have $(U-V)_0$ colors as low as 0.5 mag (Merlin et al. 2018), which suggests that genuinely quiescent galaxies can be found outside of the fiducial UVJ boundaries, and by extension, the $(ug_i)_s$ boundaries as well. Metal-poor galaxies ($Z = 0.02 Z_{\odot}$), which take ~ 3 Gyr to become UVJ -quiescent, also meet this fate, as the time needed to enter the quiescent region exceeds the age of the Universe at that redshift (Tacchella et al. 2018). Post-starburst galaxies are predicted to comprise an increasingly larger fraction of the quiescent population at $z > 3$ because galaxies will

have less time to evolve passively after quenching. One possible solution is to use a slightly redder u_s filter ($\lambda_c = 3200 \text{ \AA}$, as opposed to 2900 \AA) at $z > 4$. This would produce a stronger $(u_s-g_s)_0$ color signature for the entire quiescent population, potentially moving more post-starbursts into the quiescent region. This, however, may be at the risk of increasing the contamination from dusty star-forming galaxies, as this modification would reduce the wavelength baseline needed to accurately measure dust reddening in the near-infrared. Others have resolved this by eliminating the horizontal ($U - V$) cut altogether and instead extending the diagonal line to bluer colors (e.g. Marsan et al. 2022, Forrest et al. 2020a). While this would increase the number of post-starbursts in the quiescent region, it would also unfortunately, increase the number of non-dusty star-forming contaminants. It is also important to note that the way the slope of the diagonal line is determined becomes even more critical in this case (see our cautionary note in Section 5.1.1).

The fourth and final possible reason for $< 70\%$ completeness at $z > 4$ is that we may be missing old galaxies that are significantly obscured by dust ($E(B-V) \gtrsim 0.4$, i.e. $A_V \gtrsim 1$), causing them to fail the $V - J$ (hence the $(g_s - i_s)_0$) cut (Merlin et al. 2018). Although local ($z \sim 0$) massive, quiescent galaxies have little-to-no dust (Smith et al. 2012), recent observations have shown that their quiescent counterparts at higher redshifts ($z \gtrsim 1.5$) could be dusty, with up to 12% showing MIPS and *Herschel* detections at $z = 3 - 4$ (Martis et al. 2019). This estimate does not take into consideration gas depletion timescales, which are reported to be quite short (100–600 Myr) for quiescent galaxies at $z \sim 1.5 - 3$ (Williams et al. 2021, Caliendo et al. 2021, Whitaker et al. 2021a). It may still be possible to catch a quiescent galaxy in its dusty phase if the quenching timescale is longer than the gas depletion timescale. A few galaxies with some measure of dust obscuration (verified via ALMA detections) have been reported in the literature (Schreiber et al. 2018, Whitaker et al. 2021b) and found in simulations (Akins et al. 2022). While it is possible that these galaxies may indeed represent an evolutionary phase for quiescent galaxies, it is equally likely that they are dusty star-forming galaxies masquerading as quiescent galaxies due to their low sSFRs (heavily-dependent on the choice of SFH, e.g. Marchesini et al. 2010).

We can resolve this through targeted spectroscopic campaigns searching for the presence (or lack of) absorption features, which would unambiguously confirm their quiescent nature. Additionally, flexible SFHs and self-consistent SED modeling are recommended because they have been shown to better approximate stellar masses as they account for emission from old stars (Martis et al. 2019, Leja et al. 2019c), thereby resulting in more accurate sSFRs for dusty star-forming galaxies and lower dust masses for quiescent galaxies (Whitaker et al. 2021b). In summary, if this population of high-z dusty quiescent galaxies is real, removing the vertical $(g_s - i_s)_0$ color cut (e.g. Martis et al. 2019, Marsan et al. 2022) would increase completeness, but would also increase contamination from dusty star-forming galaxies.

To conclude, we consider the utility and accuracy of color selection methods compared to others in the literature. Are there better ones out there? The short answer is yes. The longer answer is no. Other methods such as template fitting (e.g. Shahidi et al. 2020) and unsupervised machine learning algorithms, e.g. t-distributed stochastic neighbor embedding (t-SNE, Steinhardt et al. 2020), are able to use a larger portion of the SED information to classify galaxies, as opposed to only a select few bands. They have been shown to achieve this task quite well, boasting $\sim 80\%$ completeness and $< 30\%$ contamination at $z < 2$. One major draw-back of these methods is that they are computationally-intensive and relatively complex (particularly those that employ non-parametric (flexible) models, e.g., Leja et al. 2019a). Color selection, on the other hand, is convenient and accessible, especially as many catalogs now come with pre-computed rest-frame colors. Additionally, their performance hinges upon the availability of several photometric points, which is a luxury that many surveys with the next generation of telescopes (e.g. *Euclid*, the *Vera Rubin Observatory*, and the *Nancy Grace Roman Space Telescope*) will not have. This will also be the case for *JWST* observations in new fields, as these will have no existing ancillary data from other catalogs. Lastly, these alternative methods are model-dependent. SED-fitting depends on using galaxy templates that capture the variety of galaxies in the Universe and machine learning methods must be trained on large, representative samples of galaxies. Both of these are typically available at lower redshifts ($z < 3$), and for several reasons (outlined above), we cannot assume that quiescent galaxies at low- z resemble those at high- z enough for these methods to be effective. When rest-frame colors are derived either directly from the observed photometry or using empirically-derived templates, as we have done in this paper (Sections 2 and 4), they are not subject to the underlying assumptions and biases associated with these model-dependent methods, hence are more representative of reality. We argue that this will remain important through the *JWST* era.

Going forward, progress in the study of galaxies over cosmic time will likely continue to rely heavily on color selection techniques, both to distinguish between different galaxy populations and to identify targets for spectroscopy. Over the next few months, *JWST* will provide a rapid and extremely efficient spectroscopic test⁹ of quiescence and hence of the $(ugi)_s$ color selection method. In the meantime, the next step is to standardize these independent observational results and bring them to a common scale by using a single color-color selection method over a range of redshifts, and we have demonstrated that the $(ugi)_s$ diagram can do that efficiently. In a follow-up paper, we will explore the distribution of galaxy properties at a range of redshifts in color-color space, comparing the global properties of star-forming and quiescent galaxies (e.g. stellar

mass, star formation rate, number density, IR luminosity, stellar age, and morphology) of $(ugi)_s$ as has been done for *UVJ* (e.g. Williams et al. 2010, Leja et al. 2015, Patel et al. 2012, Price et al. 2014, Straatman et al. 2014, Spitler et al. 2014, Forrest et al. 2016, Martis et al. 2016, Fang et al. 2018, Belli et al. 2019).

1 This work benefited from the generous support of the George
2 P. and Cynthia Woods Mitchell Institute for Fundamental
3 Physics and Astronomy at Texas A&M University. This ma-
4 terial is based upon work supported by the National Science
5 Foundation under Grants AST-2009632 and AST-2009442.
6 CMSS acknowledges support from Research Foundation—
7 Flanders (FWO) through Fellowship 12ZC120N.

8 Based on observations taken by the 3D-HST Treasury
9 Program (HST-GO-12177 and HST-GO-12328) with the
10 NASA/ESA Hubble Space Telescope, which is operated by
11 the Association of Universities for Research in Astronomy,
12 Inc., under NASA contract NAS5-26555 and data products
13 from observations made with ESO Telescopes at the La Silla
14 Paranal Observatory under ESO program ID 179.A-2005 and
15 on data products produced by TERAPIX and the Cambridge
16 Astronomy Survey Unit on behalf of the UltraVISTA con-
17 sortium. This research made extensive use of NASA's Astro-
18 physics Data System for bibliographic information.

19 JAD would like to thank Gabe Brammer for help with trou-
20 bleshooting and running his SED-fitting code `eazy-py`, Justin
21 Spilker and Rob Kennicutt for insightful discussions, and Tay-
22 lor Hutchison and Jonathan Cohn for their help and support
23 throughout the preparation of this manuscript.

DATA AVAILABILITY

The synthetic $(ugi)_s$ filters are available on Github¹⁰. Other data products can be provided upon request to the corresponding author.

Software: Astropy (Astropy Collaboration et al. 2018), `eazy-py` (Brammer 2021), Jupyter (Kluyver et al. 2016), `matplotlib` (Hunter 2007), `numpy` (Harris et al. 2020), `pandas` (The Pandas Development Team 2020), `seaborn` (Waskom 2021), `scipy` (Virtanen et al. 2020), `statsmodels` (Seabold & Perktold 2010)

⁹ The cost of confirming quiescence at $z > 3$ comes out to $\sim \$250$ per object, assuming 15 hours of integration with Keck/MOSFIRE. With *JWST*/NIRSPEC, this will be reduced to ~ 1 hour and $\sim \$10$ per object (assuming a 5-year mission).

¹⁰ <https://github.com/jacqdanso/synthetic-ugi-filters>

APPENDIX

A. ESTIMATING NOISE AND CREATING A SAMPLE OF DUSTY STAR-FORMING GALAXIES IN JAGUAR

Williams et al. (2018) generated number counts of star-forming galaxies at $z > 4$ in JAGUAR using UV luminosity functions, rather than stellar mass functions (Section 2.2). Dusty star-forming galaxies, however, tend to be underrepresented in UV-selected samples. Observations suggest that they comprise $\sim 50 - 60\%$ of the $\log_{10}(M_*/M_\odot) > 10.5$ population at $z = 2 - 3$ (Martis et al. 2016), and $\sim 40 - 50\%$ at $z = 3 - 4$ (Spitler et al. 2014, Martis et al. 2019). JAGUAR features $\sim 20\%$ at these redshifts, a factor of at least 2 lower than what we expect from observations. This becomes problematic for evaluating completeness and contamination fractions of color-selected quiescent samples, as dusty star-forming galaxies are their primary contaminants.

To resolve this, we artificially introduce more dusty star-forming galaxies into the JAGUAR catalogs by doing the following. For a subsample of galaxies with $A_V \geq 1$ and $\log(sSFR/\text{yr}) \geq -9.5$, we determined the UV extinction at 1500 Å as a function of stellar mass using the following relation from Pannella et al. (2009) for $\log(M_*/M_\odot) \geq 10.1$ star-forming galaxies at $z \approx 2$, assuming that it does not evolve from $z = 2 - 6$.

$$A_{1500} = 4.07 \times \log(M_*) - 39.32 \quad (\text{A1})$$

For the low mass ($\log(M_*/M_\odot) \leq 10.1$), we assume $A_{1500} = 1$ mag (Salim et al. 2005). We applied 1 mag of scatter to A_{1500} , using a lognormal distribution to prevent the perturbed extinction values from going negative. Using the UV extinction, we estimated the color excess for each galaxy using the reddening curve from (Calzetti et al. 2000):

$$E(B - V) = \frac{A_{1500}}{k(1500)} \quad (\text{A2})$$

The extinction as a function of wavelength, $A(\lambda)$, can then be obtained from the color excess, and the reddened flux in each bandpass is:

$$f_{v,\text{reddened}} = 10^{\frac{2.5 \log(f_v) + A(\lambda)}{2.5}} \quad (\text{A3})$$

We add uncertainties to the *HST* and *JWST* photometry using publicly-available code made for the JAGUAR catalogs (Hainline et al. 2020). For a user-specified survey depth, the code estimates noise directly from the photometry by determining the total flux in the smallest fixed aperture between $r=0.16''$ and $r=0.64''$ that matches the galaxy's half-light radius, estimated using the Sersic index of the galaxy. The uncertainty on the flux in each band is then given as the Poisson noise and the instrument read noise summed in quadrature. We determine the final noise estimate per band (for each galaxy) as a random draw from a Gaussian with a width set to the uncertainty on the flux. We show the *UVJ*, *ugi*, and *(ugi)_s* diagrams for $\log(M_*/M_\odot) \geq 8.9$ galaxies at $4 < z < 6$ using the perturbed data in Figure 10. The mass cutoff is based on the expected mass completeness limit at $z = 6$ for the Cosmic Evolution Early Release Science (CEERS; P.I.: S. L. Finkelstein) survey (Kauffmann et al. 2020).

The addition of noise preferentially affects low mass ($\log(M_*/M_\odot) < 10$) and faint galaxies, causing them to scatter in and out of the quiescent region. The completeness of the perturbed sample is lower by $\sim 10\%$ from the original at $z = 3 - 4$ and $\sim 20\%$ at $z = 4 - 6$. The contamination rates after the addition of noise are similar to that of the unperturbed sample.

B. EFFECT OF PHOTOMETRIC REDSHIFT UNCERTAINTY ON CONTAMINATION FROM STRONG EMISSION LINES

In Section 5.2 and Figure 12, we presented an estimate of the fraction of EELG contaminants in the quiescent selection of each color selection using a sample of galaxies at low- z ($1.2 < z < 2.3$) and high- z ($4 < z < 6$). We demonstrated that their fractions in the *(ugi)_s*-selected quiescent samples are lower by a factor of $1.6 - 3.5$ than that of *UVJ* at these redshifts, where the fraction of star-forming galaxies with rest-frame $\text{EW} > 500$ Å is expected to be $\simeq 10 - 50\%$ (Boyett et al. 2022). Underlying these estimates, however, is the assumption that we have precise and accurate photometric redshifts. For the 3D-HST sample, the maximum photo- z scatter (based on spectroscopic redshifts up to $z \sim 6$) is 2.6% (Skelton et al. 2014), and that for JAGUAR is effectively 0%, since it is simulated data. In reality, redshift uncertainties for ground- and space-based photometric surveys may be higher.

In this section, we consider the effect of photometric redshift uncertainty. We do this by estimating the photometric redshift uncertainty required for the strongest emission lines to shift into the synthetic filters (that is, we translate the wavelength shift required for a given emission line to contaminate a specified synthetic filter into a $1-\sigma$ uncertainty on the redshift). Photometric redshift scatter is typically reported as the normalized median absolute deviation:

$$\sigma_{NMAD} = 1.48 \times \text{median}\left(\frac{\Delta z}{1 + z_{true}}\right) \quad (\text{B4})$$

where z_{true} is often determined using spectroscopy. Given an emission line that has shifted into a photometric bandpass

$$\lambda_{new} = \lambda_{orig}(1 + z + \Delta z) \quad (\text{B5})$$

Table 2. Photometric redshift uncertainties required for the strongest emission lines to contaminate the synthetic filters.

Synthetic Filter	Filter Edge (Å)	Emission Line	Rest Wavelength ^a (Å)	Buffer ^b (Å)	σ_z ^c	Δz (at $z = 3$)
u_s	3100 (R)	[O II]	3729.875	627	0.249	-0.996
g_s	4300 (L)	[O II]	3729.875	573	0.228	0.912
g_s	4700 (R)	H β	4862.68	163	0.048	-0.192
g_s	4700 (R)	[O III]	4960.295, 5008.24	307	0.091	-0.364
i_s	7000 (L)	H β +[O III]	4862.68, 4960.295, 5008.24	2993 ^d	0.885	3.54
i_s	7000 (L)	H α +[N II]	6564.61, 6585.27	452	0.102	0.408
i_s	8000 (R)	Si III	9068.6, 9530.6	1069	0.174	-0.698

^aVacuum rest-frame wavelengths obtained from Momcheva et al. (2016) and the SDSS database ^a.

^bWavelength separation between the central wavelength of the line and filter edge, used to determine λ_{new} in Equation B6.

^cPhotometric redshift scatter, given by Equation B6.

^dConservative estimate using H β λ4863.

^a <http://classic.sdss.org/dr6/algorithms/linestable.html>

the photometric redshift scatter required for the line to contaminate a given filter is

$$\sigma_z = 1.48 \times \left(\frac{\lambda_{new} - 1}{\lambda_{orig}(1+z)} \right) \quad (\text{B6})$$

We report our results for the four strongest emission lines in Table 2. For blended lines and doublets, we make conservative estimates by using the central wavelength of the emission line closest to the synthetic filter in question. Based on our calculations, the filter most likely to get contaminated is g_s with H β +[O III]. Due to the proximity of these emission lines to the filter, small photo-z uncertainties ($\simeq 5 - 9\%$) can cause the lines to scatter into the g_s filter and artificially boost the $(u_s - g_s)_0$ color. Is this enough to warrant concern? Medium-band photometric surveys such as ZFOURGE (Straatman et al. 2016) and NMBS (Whitaker et al. 2011) boast 1 – 2% photo-z uncertainties due to finer sampling of the SED. Deep NIR surveys such as UltraVISTA (Muzzin et al. 2013) that include medium-band photometry from the aforementioned also report similar uncertainties up to $z \sim 2$. Typical photo-z uncertainties are on the order of 5% (e.g. COSMOS2020, Weaver et al. 2022). Similarly, photo-z uncertainties of 5 – 8% are expected for the faintest objects ($m_{F200W} > 24$ AB) in upcoming photometric surveys with JWST such as CEERS (Kauffmann et al. 2020). Since these estimates are well within our redshift uncertainty tolerance for the g_s filter, contamination from galaxies with strong emission lines will likely be small compared to UVJ as we have demonstrated in Section 5.2.

A reasonable critique of this conclusion is that photo-z uncertainties from ground-based surveys are underestimated, as they are usually evaluated using galaxies with spectroscopic redshifts. These are typically highly biased samples, comprising only $\sim 1\%$ of the catalog and tend to be 2–5 magnitudes brighter than the rest of the sample. Our counter is that for precisely this reason, EELGs will have some of the most precise redshifts in the sample, as they tend to be bright and the presence of emission lines makes it easier to obtain spectroscopic redshifts for them. Additionally, many SED fitting codes (including eazy-py) now include extreme emission line galaxy templates in order to account for the effects of emission lines on the observed photometry (see Section 2.3). One way to test the accuracy of the photometric redshift uncertainties from our sample in 3D-HST is to look at the predicted change in magnitude caused by a contaminating emission line. Because g_s is much narrower than V ($\Delta\lambda = 400$ Å versus 991 Å), a stronger magnitude change is introduced in g_s than in V when contaminated by an emission line (Equation 1). This means that if the photo-z uncertainty exceeds our estimated threshold, there should be more galaxies with moderate H β +[O III] line strengths (EW > 200 Å) in the $(ugi)_s$ quiescent region than that for UVJ since they will produce a strong magnitude change of (1.93 mag) in the former and a relatively weaker one in the latter (0.64 mag). What we see is the opposite, with the number of EW > 200 Å galaxies in $(ugi)_s$ being a factor of $\simeq 1.4$ less than that in UVJ (92 as opposed to 130). This suggests that the photometric redshift estimates in 3D-HST have an uncertainty of < 5% and lends credibility to using spectroscopic redshifts for estimating photometric redshift uncertainties.

REFERENCES

- Akins, H. B., Narayanan, D., Whitaker, K. E., et al. 2022, ApJ, 929, 94, doi: [10.3847/1538-4357/ac5d3a](https://doi.org/10.3847/1538-4357/ac5d3a) Arnouts, S., Walcher, C. J., Le Fèvre, O., et al. 2007, A&A, 476, 137, doi: [10.1051/0004-6361:20077632](https://doi.org/10.1051/0004-6361:20077632)

- Ashby, M. L. N., Willner, S. P., Fazio, G. G., et al. 2015, *ApJS*, 218, 33, doi: [10.1088/0067-0049/218/2/33](https://doi.org/10.1088/0067-0049/218/2/33)
- Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, *AJ*, 156, 123, doi: [10.3847/1538-3881/aabc4f](https://doi.org/10.3847/1538-3881/aabc4f)
- Barro, G., Pérez-González, P. G., Cava, A., et al. 2019, *ApJS*, 243, 22, doi: [10.3847/1538-4365/ab23f2](https://doi.org/10.3847/1538-4365/ab23f2)
- Beers, T. C., Flynn, K., & Gebhardt, K. 1990, *AJ*, 100, 32, doi: [10.1086/115487](https://doi.org/10.1086/115487)
- Behroozi, P., Conroy, C., Wechsler, R. H., et al. 2020, *MNRAS*, 499, 5702, doi: [10.1093/mnras/staa3164](https://doi.org/10.1093/mnras/staa3164)
- Behroozi, P. S., Conroy, C., & Wechsler, R. H. 2010, *ApJ*, 717, 379, doi: [10.1088/0004-637X/717/1/379](https://doi.org/10.1088/0004-637X/717/1/379)
- Belfiore, F., Maiolino, R., Bundy, K., et al. 2018, *MNRAS*, 477, 3014, doi: [10.1093/mnras/sty768](https://doi.org/10.1093/mnras/sty768)
- Belli, S., Newman, A. B., & Ellis, R. S. 2019, *ApJ*, 874, 17, doi: [10.3847/1538-4357/ab07af](https://doi.org/10.3847/1538-4357/ab07af)
- Belli, S., Genzel, R., Förster Schreiber, N. M., et al. 2017, *ApJL*, 841, L6, doi: [10.3847/2041-8213/aa70e5](https://doi.org/10.3847/2041-8213/aa70e5)
- Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015, *ApJ*, 803, 34, doi: [10.1088/0004-637X/803/1/34](https://doi.org/10.1088/0004-637X/803/1/34)
- Boyett, K. N. K., Stark, D. P., Bunker, A. J., Tang, M., & Maseda, M. V. 2022, *MNRAS*, doi: [10.1093/mnras/stac1109](https://doi.org/10.1093/mnras/stac1109)
- Bradley, L. D., Bouwens, R. J., Zitrin, A., et al. 2012, *ApJ*, 747, 3, doi: [10.1088/0004-637X/747/1/3](https://doi.org/10.1088/0004-637X/747/1/3)
- Bradley, L. D., Zitrin, A., Coe, D., et al. 2014, *ApJ*, 792, 76, doi: [10.1088/0004-637X/792/1/76](https://doi.org/10.1088/0004-637X/792/1/76)
- Brammer, G. 2021, gbrammer/eazy-py: Tagged release 2021, 0.5.2, Zenodo, doi: [10.5281/zenodo.5012705](https://doi.org/10.5281/zenodo.5012705)
- Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, *ApJ*, 686, 1503, doi: [10.1086/591786](https://doi.org/10.1086/591786)
- Brammer, G. B., Whitaker, K. E., van Dokkum, P. G., et al. 2009, *ApJL*, 706, L173, doi: [10.1088/0004-637X/706/1/L173](https://doi.org/10.1088/0004-637X/706/1/L173)
- . 2011, *ApJ*, 739, 24, doi: [10.1088/0004-637X/739/1/24](https://doi.org/10.1088/0004-637X/739/1/24)
- Bruzual, G., & Charlot, S. 2003, *MNRAS*, 344, 1000, doi: [10.1046/j.1365-8711.2003.06897.x](https://doi.org/10.1046/j.1365-8711.2003.06897.x)
- Caliendo, J. N., Whitaker, K. E., Akhshik, M., et al. 2021, *ApJL*, 910, L7, doi: [10.3847/2041-8213/abe132](https://doi.org/10.3847/2041-8213/abe132)
- Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, *ApJ*, 533, 682, doi: [10.1086/308692](https://doi.org/10.1086/308692)
- Carnall, A. C., McLure, R. J., Dunlop, J. S., & Davé, R. 2018, *MNRAS*, 480, 4379, doi: [10.1093/mnras/sty2169](https://doi.org/10.1093/mnras/sty2169)
- Carnall, A. C., McLure, R. J., Dunlop, J. S., et al. 2019, *MNRAS*, 490, 417, doi: [10.1093/mnras/stz2544](https://doi.org/10.1093/mnras/stz2544)
- Carnall, A. C., Walker, S., McLure, R. J., et al. 2020, *MNRAS*, 496, 695, doi: [10.1093/mnras/staa1535](https://doi.org/10.1093/mnras/staa1535)
- Casali, M., Adamson, A., Alves de Oliveira, C., et al. 2007, *A&A*, 467, 777, doi: [10.1051/0004-6361:20066514](https://doi.org/10.1051/0004-6361:20066514)
- Castellano, M., Pentericci, L., Fontana, A., et al. 2017, *ApJ*, 839, 73, doi: [10.3847/1538-4357/aa696e](https://doi.org/10.3847/1538-4357/aa696e)
- Chabrier, G. 2003, *PASP*, 115, 763, doi: [10.1086/376392](https://doi.org/10.1086/376392)
- Chang, Y.-Y., van der Wel, A., da Cunha, E., & Rix, H.-W. 2015, *ApJS*, 219, 8, doi: [10.1088/0067-0049/219/1/8](https://doi.org/10.1088/0067-0049/219/1/8)
- Charlot, S., & Fall, S. M. 2000, *ApJ*, 539, 718, doi: [10.1086/309250](https://doi.org/10.1086/309250)
- Chevallard, J., & Charlot, S. 2016, *MNRAS*, 462, 1415, doi: [10.1093/mnras/stw1756](https://doi.org/10.1093/mnras/stw1756)
- Cimatti, A., Daddi, E., Renzini, A., et al. 2004, *Nature*, 430, 184, doi: [10.1038/nature02668](https://doi.org/10.1038/nature02668)
- Coe, D., Salmon, B., Bradač, M., et al. 2019, *ApJ*, 884, 85, doi: [10.3847/1538-4357/ab412b](https://doi.org/10.3847/1538-4357/ab412b)
- Cohn, J. H., Leja, J., Tran, K.-V. H., et al. 2018, *ApJ*, 869, 141, doi: [10.3847/1538-4357/aaed3d](https://doi.org/10.3847/1538-4357/aaed3d)
- Conroy, C. 2013, *ARA&A*, 51, 393, doi: [10.1146/annurev-astro-082812-141017](https://doi.org/10.1146/annurev-astro-082812-141017)
- Conroy, C., & Gunn, J. E. 2010, FSPS: Flexible Stellar Population Synthesis, Astrophysics Source Code Library, record ascl:1010.043. <http://ascl.net/1010.043>
- Croux, C., & Rousseeuw, P. J. 1992, in Computational Statistics, ed. Y. Dodge & J. Whittaker (Heidelberg: Physica-Verlag HD), 411–428
- da Cunha, E., Charlot, S., & Elbaz, D. 2008, *MNRAS*, 388, 1595, doi: [10.1111/j.1365-2966.2008.13535.x](https://doi.org/10.1111/j.1365-2966.2008.13535.x)
- da Cunha, E., Walter, F., Smail, I. R., et al. 2015, *ApJ*, 806, 110, doi: [10.1088/0004-637X/806/1/110](https://doi.org/10.1088/0004-637X/806/1/110)
- Daddi, E., Cimatti, A., Renzini, A., et al. 2004, *ApJ*, 617, 746, doi: [10.1086/425569](https://doi.org/10.1086/425569)
- D'Eugenio, C., Daddi, E., Gobat, R., et al. 2020, *ApJL*, 892, L2, doi: [10.3847/2041-8213/ab7a96](https://doi.org/10.3847/2041-8213/ab7a96)
- Díaz-García, L. A., Cenarro, A. J., López-Sanjuan, C., et al. 2019, *A&A*, 631, A156, doi: [10.1051/0004-6361/201832788](https://doi.org/10.1051/0004-6361/201832788)
- Donnari, M., Pillepich, A., Nelson, D., et al. 2019, *MNRAS*, 485, 4817, doi: [10.1093/mnras/stz712](https://doi.org/10.1093/mnras/stz712)
- Dunlop, J. S., Abraham, R. G., Ashby, M. L. N., et al. 2021, PRIMER: Public Release IMaging for Extragalactic Research, JWST Proposal. Cycle 1, ID. #1837
- Ellis, R. S., McLure, R. J., Dunlop, J. S., et al. 2013, *ApJL*, 763, L7, doi: [10.1088/2041-8205/763/1/L7](https://doi.org/10.1088/2041-8205/763/1/L7)
- Endsley, R., Stark, D. P., Chevallard, J., & Charlot, S. 2021, *MNRAS*, 500, 5229, doi: [10.1093/mnras/staa3370](https://doi.org/10.1093/mnras/staa3370)
- Erfanianfar, G., Popesso, P., Finoguenov, A., et al. 2016, *MNRAS*, 455, 2839, doi: [10.1093/mnras/stv2485](https://doi.org/10.1093/mnras/stv2485)
- Esdaile, J., Labbé, I., Glazebrook, K., et al. 2021, *AJ*, 162, 225, doi: [10.3847/1538-3881/ac2148](https://doi.org/10.3847/1538-3881/ac2148)
- Fang, J. J., Faber, S. M., Koo, D. C., et al. 2018, *ApJ*, 858, 100, doi: [10.3847/1538-4357/aabcba](https://doi.org/10.3847/1538-4357/aabcba)
- Finkelstein, S. L., Dickinson, M., Ferguson, H. C., et al. 2017, The Cosmic Evolution Early Release Science (CEERS) Survey, JWST Proposal ID 1345. Cycle 0 Early Release Science
- Forrest, B., Tran, K.-V. H., Tomczak, A. R., et al. 2016, *ApJL*, 818, L26, doi: [10.3847/2041-8205/818/2/L26](https://doi.org/10.3847/2041-8205/818/2/L26)

- Forrest, B., Marsan, Z. C., Annunziatella, M., et al. 2020a, *ApJ*, 903, 47, doi: [10.3847/1538-4357/abb819](https://doi.org/10.3847/1538-4357/abb819)
- Forrest, B., Annunziatella, M., Wilson, G., et al. 2020b, *ApJL*, 890, L1, doi: [10.3847/2041-8213/ab5b9f](https://doi.org/10.3847/2041-8213/ab5b9f)
- Freedman, D., & Diaconis, P. 1981, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 57, 453
- Glazebrook, K., Schreiber, C., Labb  , I., et al. 2017, *Nature*, 544, 71, doi: [10.1038/nature21680](https://doi.org/10.1038/nature21680)
- Grogan, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, *ApJS*, 197, 35, doi: [10.1088/0067-0049/197/2/35](https://doi.org/10.1088/0067-0049/197/2/35)
- Guo, Y., Ferguson, H. C., Giavalisco, M., et al. 2013, *ApJS*, 207, 24, doi: [10.1088/0067-0049/207/2/24](https://doi.org/10.1088/0067-0049/207/2/24)
- Gutkin, J., Charlot, S., & Bruzual, G. 2016, *MNRAS*, 462, 1757, doi: [10.1093/mnras/stw1716](https://doi.org/10.1093/mnras/stw1716)
- Hainline, K. N., Hviding, R. E., Rieke, M., et al. 2020, *ApJ*, 892, 125, doi: [10.3847/1538-4357/ab7dc3](https://doi.org/10.3847/1538-4357/ab7dc3)
- Harikane, Y., Inoue, A. K., Mawatari, K., et al. 2022, *ApJ*, 929, 1, doi: [10.3847/1538-4357/ac53a9](https://doi.org/10.3847/1538-4357/ac53a9)
- Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, *Nature*, 585, 357, doi: [10.1038/s41586-020-2649-2](https://doi.org/10.1038/s41586-020-2649-2)
- Heinis, S., Buat, V., B  thermin, M., et al. 2014, *MNRAS*, 437, 1268, doi: [10.1093/mnras/stt1960](https://doi.org/10.1093/mnras/stt1960)
- Hunter, J. D. 2007, *Computing in Science & Engineering*, 9, 90, doi: [10.1109/MCSE.2007.55](https://doi.org/10.1109/MCSE.2007.55)
- Hwang, Y.-H., Wang, W.-H., Chang, Y.-Y., et al. 2021, *ApJ*, 913, 6, doi: [10.3847/1538-4357/abf11a](https://doi.org/10.3847/1538-4357/abf11a)
- Ilbert, O., Salvato, M., Le Floc'h, E., et al. 2010, *ApJ*, 709, 644, doi: [10.1088/0004-637X/709/2/644](https://doi.org/10.1088/0004-637X/709/2/644)
- Ilbert, O., McCracken, H. J., Le F  vre, O., et al. 2013, *A&A*, 556, A55, doi: [10.1051/0004-6361/201321100](https://doi.org/10.1051/0004-6361/201321100)
- Ilbert, O., Arnouts, S., Le Floc'h, E., et al. 2015, *A&A*, 579, A2, doi: [10.1051/0004-6361/201425176](https://doi.org/10.1051/0004-6361/201425176)
- Izotov, Y. I., Guseva, N. G., Fricke, K. J., & Henkel, C. 2016, *MNRAS*, 462, 4427, doi: [10.1093/mnras/stw1973](https://doi.org/10.1093/mnras/stw1973)
- Johnson, B., & Leja, J. 2017, Bd-J/Prospector: Initial Release, v0.1, Zenodo, Zenodo, doi: [10.5281/zenodo.1116491](https://doi.org/10.5281/zenodo.1116491)
- Kalita, B. S., Daddi, E., D'Eugenio, C., et al. 2021, *ApJL*, 917, L17, doi: [10.3847/2041-8213/ac16dc](https://doi.org/10.3847/2041-8213/ac16dc)
- Kartaltepe, J., Casey, C. M., Bagley, M., et al. 2021, COSMOS-Webb: The Webb Cosmic Origins Survey, JWST Proposal. Cycle 1, ID. #1727
- Katsianis, A., Gonzalez, V., Barrientos, D., et al. 2020, *MNRAS*, 492, 5592, doi: [10.1093/mnras/staa157](https://doi.org/10.1093/mnras/staa157)
- Kauffmann, O. B., Le F  vre, O., Ilbert, O., et al. 2020, *A&A*, 640, A67, doi: [10.1051/0004-6361/202037450](https://doi.org/10.1051/0004-6361/202037450)
- Kawinwanichakij, L., Quadri, R. F., Papovich, C., et al. 2016, *ApJ*, 817, 9, doi: [10.3847/0004-637X/817/1/9](https://doi.org/10.3847/0004-637X/817/1/9)
- Kissler-Patig, M., Pirard, J. F., Casali, M., et al. 2008, *A&A*, 491, 941, doi: [10.1051/0004-6361:200809910](https://doi.org/10.1051/0004-6361:200809910)
- Kluyver, T., Ragan-Kelley, B., P  rez, F., et al. 2016, in *Positioning and Power in Academic Publishing: Players, Agents and Agendas*, ed. F. Loizides & B. Schmidt, IOS Press, 87 – 90
- Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, *ApJS*, 197, 36, doi: [10.1088/0067-0049/197/2/36](https://doi.org/10.1088/0067-0049/197/2/36)
- Kriek, M., van Dokkum, P. G., Labb  , I., et al. 2009, *ApJ*, 700, 221, doi: [10.1088/0004-637X/700/1/221](https://doi.org/10.1088/0004-637X/700/1/221)
- Kriek, M., Labb  , I., Conroy, C., et al. 2010, *ApJL*, 722, L64, doi: [10.1088/2041-8205/722/1/L64](https://doi.org/10.1088/2041-8205/722/1/L64)
- Kurczynski, P., Gawiser, E., Acquaviva, V., et al. 2016, *ApJL*, 820, L1, doi: [10.3847/2041-8205/820/1/L1](https://doi.org/10.3847/2041-8205/820/1/L1)
- Labb  , I., Huang, J., Franx, M., et al. 2005, *ApJL*, 624, L81, doi: [10.1086/430700](https://doi.org/10.1086/430700)
- Labb  , I., Oesch, P. A., Bouwens, R. J., et al. 2013, *ApJL*, 777, L19, doi: [10.1088/2041-8205/777/2/L19](https://doi.org/10.1088/2041-8205/777/2/L19)
- Lee, N., Sanders, D. B., Casey, C. M., et al. 2015, *ApJ*, 801, 80, doi: [10.1088/0004-637X/801/2/80](https://doi.org/10.1088/0004-637X/801/2/80)
- Lee, S.-K., Ferguson, H. C., Somerville, R. S., Wiklind, T., & Giavalisco, M. 2010, *ApJ*, 725, 1644, doi: [10.1088/0004-637X/725/2/1644](https://doi.org/10.1088/0004-637X/725/2/1644)
- Leja, J., Carnall, A. C., Johnson, B. D., Conroy, C., & Speagle, J. S. 2019a, *ApJ*, 876, 3, doi: [10.3847/1538-4357/ab133c](https://doi.org/10.3847/1538-4357/ab133c)
- Leja, J., Johnson, B. D., Conroy, C., van Dokkum, P. G., & Byler, N. 2017, *ApJ*, 837, 170, doi: [10.3847/1538-4357/aa5ff](https://doi.org/10.3847/1538-4357/aa5ff)
- Leja, J., Tacchella, S., & Conroy, C. 2019b, *ApJL*, 880, L9, doi: [10.3847/2041-8213/ab2f8c](https://doi.org/10.3847/2041-8213/ab2f8c)
- Leja, J., van Dokkum, P. G., Franx, M., & Whitaker, K. E. 2015, *ApJ*, 798, 115, doi: [10.1088/0004-637X/798/2/115](https://doi.org/10.1088/0004-637X/798/2/115)
- Leja, J., Johnson, B. D., Conroy, C., et al. 2019c, *ApJ*, 877, 140, doi: [10.3847/1538-4357/ab1d5a](https://doi.org/10.3847/1538-4357/ab1d5a)
- Lotz, J. M., Koekemoer, A., Coe, D., et al. 2017, *ApJ*, 837, 97, doi: [10.3847/1538-4357/837/1/97](https://doi.org/10.3847/1538-4357/837/1/97)
- Lustig, P., Strazzullo, V., Remus, R.-S., et al. 2022, arXiv e-prints, arXiv:2201.09068. <https://arxiv.org/abs/2201.09068>
- Lutz, D., Poglitsch, A., Altieri, B., et al. 2011, *A&A*, 532, A90, doi: [10.1051/0004-6361/201117107](https://doi.org/10.1051/0004-6361/201117107)
- Man, A. W. S., Greve, T. R., Toft, S., et al. 2016, *ApJ*, 820, 11, doi: [10.3847/0004-637X/820/1/11](https://doi.org/10.3847/0004-637X/820/1/11)
- Marchesini, D., van Dokkum, P. G., F  rster Schreiber, N. M., et al. 2009, *ApJ*, 701, 1765, doi: [10.1088/0004-637X/701/2/1765](https://doi.org/10.1088/0004-637X/701/2/1765)
- Marchesini, D., Whitaker, K. E., Brammer, G., et al. 2010, *ApJ*, 725, 1277, doi: [10.1088/0004-637X/725/1/1277](https://doi.org/10.1088/0004-637X/725/1/1277)
- Marsan, C., Muzzin, A., Alcorn, L. Y., et al. 2021, Dawn of the Monsters: JWST Characterization of Extremely Massive Galaxies at z 5, JWST Proposal. Cycle 1, ID. #2362
- Marsan, Z. C., Marchesini, D., Brammer, G. B., et al. 2015, *ApJ*, 801, 133, doi: [10.1088/0004-637X/801/2/133](https://doi.org/10.1088/0004-637X/801/2/133)
- Marsan, Z. C., Muzzin, A., Marchesini, D., et al. 2022, *ApJ*, 924, 25, doi: [10.3847/1538-4357/ac312a](https://doi.org/10.3847/1538-4357/ac312a)

- Martin, D. C., Wyder, T. K., Schiminovich, D., et al. 2007, ApJS, 173, 342, doi: [10.1086/516639](https://doi.org/10.1086/516639)
- Martis, N. S., Marchesini, D. M., Muzzin, A., et al. 2019, ApJ, 882, 65, doi: [10.3847/1538-4357/ab32f1](https://doi.org/10.3847/1538-4357/ab32f1)
- Martis, N. S., Marchesini, D., Brammer, G. B., et al. 2016, ApJL, 827, L25, doi: [10.3847/2041-8205/827/2/L25](https://doi.org/10.3847/2041-8205/827/2/L25)
- Maseda, M. V., van der Wel, A., Rix, H.-W., et al. 2018, ApJ, 854, 29, doi: [10.3847/1538-4357/aaa76e](https://doi.org/10.3847/1538-4357/aaa76e)
- Mawatari, K., Yamada, T., Fazio, G. G., Huang, J.-S., & Ashby, M. L. N. 2016, PASJ, 68, 46, doi: [10.1093/pasj/psw041](https://doi.org/10.1093/pasj/psw041)
- McCarthy, P. J., Le Borgne, D., Crampton, D., et al. 2004, ApJL, 614, L9, doi: [10.1086/425306](https://doi.org/10.1086/425306)
- McCracken, H. J., Milvang-Jensen, B., Dunlop, J., et al. 2012, A&A, 544, A156, doi: [10.1051/0004-6361/201219507](https://doi.org/10.1051/0004-6361/201219507)
- Merlin, E., Fontana, A., Castellano, M., et al. 2018, MNRAS, 473, 2098, doi: [10.1093/mnras/stx2385](https://doi.org/10.1093/mnras/stx2385)
- Merlin, E., Fortuni, F., Torelli, M., et al. 2019, MNRAS, 490, 3309, doi: [10.1093/mnras/stz2615](https://doi.org/10.1093/mnras/stz2615)
- Mitchell, P. D., Lacey, C. G., Baugh, C. M., & Cole, S. 2013, MNRAS, 435, 87, doi: [10.1093/mnras/stt1280](https://doi.org/10.1093/mnras/stt1280)
- Mobasher, B., Dahlen, T., Ferguson, H. C., et al. 2015, ApJ, 808, 101, doi: [10.1088/0004-637X/808/1/101](https://doi.org/10.1088/0004-637X/808/1/101)
- Moles, M., Benítez, N., Aguerri, J. A. L., et al. 2008, AJ, 136, 1325, doi: [10.1088/0004-6256/136/3/1325](https://doi.org/10.1088/0004-6256/136/3/1325)
- Momcheva, I. G., Brammer, G. B., van Dokkum, P. G., et al. 2016, ApJS, 225, 27, doi: [10.3847/0067-0049/225/2/27](https://doi.org/10.3847/0067-0049/225/2/27)
- Muzzin, A., Marchesini, D., van Dokkum, P. G., et al. 2009, ApJ, 701, 1839, doi: [10.1088/0004-637X/701/2/1839](https://doi.org/10.1088/0004-637X/701/2/1839)
- Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013, ApJ, 777, 18, doi: [10.1088/0004-637X/777/1/18](https://doi.org/10.1088/0004-637X/777/1/18)
- Nanayakkara, T., Esdaille, J., Glazebrook, K., et al. 2022, PASA, 39, e002, doi: [10.1017/pasa.2021.61](https://doi.org/10.1017/pasa.2021.61)
- Nayyeri, H., Mobasher, B., Hemmati, S., et al. 2014, ApJ, 794, 68, doi: [10.1088/0004-637X/794/1/68](https://doi.org/10.1088/0004-637X/794/1/68)
- Oesch, P. A., Bouwens, R. J., Illingworth, G. D., Labb  , I., & Stefanon, M. 2018, ApJ, 855, 105, doi: [10.3847/1538-4357/aab03f](https://doi.org/10.3847/1538-4357/aab03f)
- Oesch, P. A., Bouwens, R. J., Illingworth, G. D., et al. 2014, ApJ, 786, 108, doi: [10.1088/0004-637X/786/2/108](https://doi.org/10.1088/0004-637X/786/2/108)
- Oke, J. B., & Gunn, J. E. 1983, ApJ, 266, 713, doi: [10.1086/160817](https://doi.org/10.1086/160817)
- Oliver, S. J., Bock, J., Altieri, B., et al. 2012, MNRAS, 424, 1614, doi: [10.1111/j.1365-2966.2012.20912.x](https://doi.org/10.1111/j.1365-2966.2012.20912.x)
- Pacifici, C., Kassin, S. A., Weiner, B. J., et al. 2016, ApJ, 832, 79, doi: [10.3847/0004-637X/832/1/79](https://doi.org/10.3847/0004-637X/832/1/79)
- Pannella, M., Carilli, C. L., Daddi, E., et al. 2009, ApJL, 698, L116, doi: [10.1088/0004-637X/698/2/L116](https://doi.org/10.1088/0004-637X/698/2/L116)
- Papovich, C., Dickinson, M., & Ferguson, H. C. 2001, ApJ, 559, 620, doi: [10.1086/322412](https://doi.org/10.1086/322412)
- Papovich, C., Finkelstein, S. L., Ferguson, H. C., Lotz, J. M., & Giavalisco, M. 2011, MNRAS, 412, 1123, doi: [10.1111/j.1365-2966.2010.17965.x](https://doi.org/10.1111/j.1365-2966.2010.17965.x)
- Papovich, C., Bassett, R., Lotz, J. M., et al. 2012, ApJ, 750, 93, doi: [10.1088/0004-637X/750/2/93](https://doi.org/10.1088/0004-637X/750/2/93)
- Papovich, C., Labb  , I., Quadri, R., et al. 2015, ApJ, 803, 26, doi: [10.1088/0004-637X/803/1/26](https://doi.org/10.1088/0004-637X/803/1/26)
- Papovich, C., Shipley, H. V., Mehrtens, N., et al. 2016, ApJS, 224, 28, doi: [10.3847/0067-0049/224/2/28](https://doi.org/10.3847/0067-0049/224/2/28)
- Patel, S. G., Holden, B. P., Kelson, D. D., et al. 2012, ApJL, 748, L27, doi: [10.1088/2041-8205/748/2/L27](https://doi.org/10.1088/2041-8205/748/2/L27)
- Patel, S. G., Hong, Y. X., Quadri, R. F., Holden, B. P., & Williams, R. J. 2017, ApJ, 839, 127, doi: [10.3847/1538-4357/aa6bf4](https://doi.org/10.3847/1538-4357/aa6bf4)
- Persson, S. E., Murphy, D. C., Smee, S., et al. 2013, PASP, 125, 654, doi: [10.1086/671164](https://doi.org/10.1086/671164)
- Pforr, J., Maraston, C., & Tonini, C. 2012, MNRAS, 422, 3285, doi: [10.1111/j.1365-2966.2012.20848.x](https://doi.org/10.1111/j.1365-2966.2012.20848.x)
- Popesso, P., Morselli, L., Concas, A., et al. 2019, MNRAS, 490, 5285, doi: [10.1093/mnras/stz2635](https://doi.org/10.1093/mnras/stz2635)
- Price, S. H., Kriek, M., Brammer, G. B., et al. 2014, ApJ, 788, 86, doi: [10.1088/0004-637X/788/1/86](https://doi.org/10.1088/0004-637X/788/1/86)
- Quadri, R., Marchesini, D., van Dokkum, P., et al. 2007, AJ, 134, 1103, doi: [10.1086/520330](https://doi.org/10.1086/520330)
- Rasappu, N., Smit, R., Labb  , I., et al. 2016, MNRAS, 461, 3886, doi: [10.1093/mnras/stw1484](https://doi.org/10.1093/mnras/stw1484)
- Reddy, N. A., Pettini, M., Steidel, C. C., et al. 2012, ApJ, 754, 25, doi: [10.1088/0004-637X/754/1/25](https://doi.org/10.1088/0004-637X/754/1/25)
- Reddy, N. A., Shapley, A. E., Sanders, R. L., et al. 2018, ApJ, 869, 92, doi: [10.3847/1538-4357/aaed1e](https://doi.org/10.3847/1538-4357/aaed1e)
- Renzini, A., & Peng, Y.-j. 2015, ApJL, 801, L29, doi: [10.1088/2041-8205/801/2/L29](https://doi.org/10.1088/2041-8205/801/2/L29)
- Roberts-Borsani, G. W., Bouwens, R. J., Oesch, P. A., et al. 2016, ApJ, 823, 143, doi: [10.3847/0004-637X/823/2/143](https://doi.org/10.3847/0004-637X/823/2/143)
- Robertson, B. E. 2022, Annual Review of Astronomy and Astrophysics, 60, null, doi: [10.1146/annurev-astro-120221-044656](https://doi.org/10.1146/annurev-astro-120221-044656)
- Rodighiero, G., Renzini, A., Daddi, E., et al. 2014, MNRAS, 443, 19, doi: [10.1093/mnras/stu1110](https://doi.org/10.1093/mnras/stu1110)
- Roebuck, E., Sajina, A., Hayward, C. C., et al. 2019, ApJ, 881, 18, doi: [10.3847/1538-4357/ab2bf5](https://doi.org/10.3847/1538-4357/ab2bf5)
- Salim, S., Charlot, S., Rich, R. M., et al. 2005, ApJL, 619, L39, doi: [10.1086/424800](https://doi.org/10.1086/424800)
- Salmon, B., Coe, D., Bradley, L., et al. 2018, ApJL, 864, L22, doi: [10.3847/2041-8213/aadc10](https://doi.org/10.3847/2041-8213/aadc10)
- Sanders, D. B., Salvato, M., Aussel, H., et al. 2007, ApJS, 172, 86, doi: [10.1086/517885](https://doi.org/10.1086/517885)
- Santini, P., Ferguson, H. C., Fontana, A., et al. 2015, ApJ, 801, 97, doi: [10.1088/0004-637X/801/2/97](https://doi.org/10.1088/0004-637X/801/2/97)
- Santini, P., Castellano, M., Merlin, E., et al. 2021, A&A, 652, A30, doi: [10.1051/0004-6361/202039738](https://doi.org/10.1051/0004-6361/202039738)

- Saracco, P., Marchesini, D., La Barbera, F., et al. 2020, *ApJ*, 905, 40, doi: [10.3847/1538-4357/abc7c4](https://doi.org/10.3847/1538-4357/abc7c4)
- Schawinski, K., Urry, C. M., Simmons, B. D., et al. 2014, *MNRAS*, 440, 889, doi: [10.1093/mnras/stu327](https://doi.org/10.1093/mnras/stu327)
- Schmidt, K. B., Treu, T., Trenti, M., et al. 2014, *ApJ*, 786, 57, doi: [10.1088/0004-637X/786/1/57](https://doi.org/10.1088/0004-637X/786/1/57)
- Schreiber, C., Pannella, M., Elbaz, D., et al. 2015, *A&A*, 575, A74, doi: [10.1051/0004-6361/201425017](https://doi.org/10.1051/0004-6361/201425017)
- Schreiber, C., Glazebrook, K., Nanayakkara, T., et al. 2018, *A&A*, 618, A85, doi: [10.1051/0004-6361/201833070](https://doi.org/10.1051/0004-6361/201833070)
- Seabold, S., & Perktold, J. 2010, in 9th Python in Science Conference
- Shahidi, A., Mobasher, B., Nayyeri, H., et al. 2020, *ApJ*, 897, 44, doi: [10.3847/1538-4357/ab96c5](https://doi.org/10.3847/1538-4357/ab96c5)
- Shapley, A. E., Steidel, C. C., Erb, D. K., et al. 2005, *ApJ*, 626, 698, doi: [10.1086/429990](https://doi.org/10.1086/429990)
- Shen, X., Vogelsberger, M., Nelson, D., et al. 2020, *MNRAS*, 495, 4747, doi: [10.1093/mnras/staa1423](https://doi.org/10.1093/mnras/staa1423)
- Sherman, S., Jogee, S., Florez, J., et al. 2021, *MNRAS*, 505, 947, doi: [10.1093/mnras/stab1350](https://doi.org/10.1093/mnras/stab1350)
- Skelton, R. E., Whitaker, K. E., Momcheva, I. G., et al. 2014, *ApJS*, 214, 24, doi: [10.1088/0067-0049/214/2/24](https://doi.org/10.1088/0067-0049/214/2/24)
- Smit, R., Bouwens, R. J., Labb  , I., et al. 2014, *ApJ*, 784, 58, doi: [10.1088/0004-637X/784/1/58](https://doi.org/10.1088/0004-637X/784/1/58)
- Smit, R., Bouwens, R. J., Franx, M., et al. 2015, *ApJ*, 801, 122, doi: [10.1088/0004-637X/801/2/122](https://doi.org/10.1088/0004-637X/801/2/122)
- Smith, M. W. L., Eales, S. A., Gomez, H. L., et al. 2012, *ApJ*, 756, 40, doi: [10.1088/0004-637X/756/1/40](https://doi.org/10.1088/0004-637X/756/1/40)
- Sobral, D., Best, P. N., Smail, I., et al. 2014, *MNRAS*, 437, 3516, doi: [10.1093/mnras/stt2159](https://doi.org/10.1093/mnras/stt2159)
- Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, *ApJS*, 214, 15, doi: [10.1088/0067-0049/214/2/15](https://doi.org/10.1088/0067-0049/214/2/15)
- Spitler, L. R., Straatman, C. M. S., Labb  , I., et al. 2014, *ApJL*, 787, L36, doi: [10.1088/2041-8205/787/2/L36](https://doi.org/10.1088/2041-8205/787/2/L36)
- Springel, V., Pakmor, R., Pillepich, A., et al. 2018, *MNRAS*, 475, 676, doi: [10.1093/mnras/stx3304](https://doi.org/10.1093/mnras/stx3304)
- Steinborn, L. K., Dolag, K., Comerford, J. M., et al. 2016, *MNRAS*, 458, 1013, doi: [10.1093/mnras/stw316](https://doi.org/10.1093/mnras/stw316)
- Steinhardt, C. L., Weaver, J. R., Maxfield, J., et al. 2020, *ApJ*, 891, 136, doi: [10.3847/1538-4357/ab76be](https://doi.org/10.3847/1538-4357/ab76be)
- Straatman, C. M. S., Labb  , I., Spitler, L. R., et al. 2014, *ApJL*, 783, L14, doi: [10.1088/2041-8205/783/1/L14](https://doi.org/10.1088/2041-8205/783/1/L14)
- Straatman, C. M. S., Spitler, L. R., Quadri, R. F., et al. 2016, *ApJ*, 830, 51, doi: [10.3847/0004-637X/830/1/51](https://doi.org/10.3847/0004-637X/830/1/51)
- Strom, A. L., Steidel, C. C., Rudie, G. C., et al. 2017, *ApJ*, 836, 164, doi: [10.3847/1538-4357/836/2/164](https://doi.org/10.3847/1538-4357/836/2/164)
- Sutherland, W., Emerson, J., Dalton, G., et al. 2015, *A&A*, 575, A25, doi: [10.1051/0004-6361/201424973](https://doi.org/10.1051/0004-6361/201424973)
- Tacchella, S., Carollo, C. M., F  rster Schreiber, N. M., et al. 2018, *ApJ*, 859, 56, doi: [10.3847/1538-4357/aabf8b](https://doi.org/10.3847/1538-4357/aabf8b)
- Tacchella, S., Conroy, C., Faber, S. M., et al. 2022, *ApJ*, 926, 134, doi: [10.3847/1538-4357/ac449b](https://doi.org/10.3847/1538-4357/ac449b)
- Tal, T., Dekel, A., Oesch, P., et al. 2014, *ApJ*, 789, 164, doi: [10.1088/0004-637X/789/2/164](https://doi.org/10.1088/0004-637X/789/2/164)
- Tanaka, M., Valentino, F., Toft, S., et al. 2019, *ApJL*, 885, L34, doi: [10.3847/2041-8213/ab4ff3](https://doi.org/10.3847/2041-8213/ab4ff3)
- Tang, M., Stark, D. P., Chevallard, J., et al. 2021, *MNRAS*, 501, 3238, doi: [10.1093/mnras/staa3454](https://doi.org/10.1093/mnras/staa3454)
- Tasca, L. A. M., Le F  vre, O., Hathi, N. P., et al. 2015, *A&A*, 581, A54, doi: [10.1051/0004-6361/201425379](https://doi.org/10.1051/0004-6361/201425379)
- The Pandas Development Team. 2020, *pandas-dev/pandas: Pandas*, latest, Zenodo, doi: [10.5281/zenodo.3509134](https://doi.org/10.5281/zenodo.3509134)
- Tomczak, A. R., Quadri, R. F., Tran, K.-V. H., et al. 2014, *ApJ*, 783, 85, doi: [10.1088/0004-637X/783/2/85](https://doi.org/10.1088/0004-637X/783/2/85)
- . 2016, *ApJ*, 817, 118, doi: [10.3847/0004-637X/817/2/118](https://doi.org/10.3847/0004-637X/817/2/118)
- Tran, K.-V. H., Forrest, B., Alcorn, L. Y., et al. 2020, *ApJ*, 898, 45, doi: [10.3847/1538-4357/ab8cba](https://doi.org/10.3847/1538-4357/ab8cba)
- Valentino, F., Tanaka, M., Davidzon, I., et al. 2020, *ApJ*, 889, 93, doi: [10.3847/1538-4357/ab64dc](https://doi.org/10.3847/1538-4357/ab64dc)
- van der Wel, A., Straughn, A. N., Rix, H. W., et al. 2011, *ApJ*, 742, 111, doi: [10.1088/0004-637X/742/2/111](https://doi.org/10.1088/0004-637X/742/2/111)
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *Nature Methods*, 17, 261, doi: [10.1038/s41592-019-0686-2](https://doi.org/10.1038/s41592-019-0686-2)
- Waskom, M. L. 2021, *Journal of Open Source Software*, 6, 3021, doi: [10.21105/joss.03021](https://doi.org/10.21105/joss.03021)
- Weaver, J. R., Kauffmann, O. B., Ilbert, O., et al. 2022, *ApJS*, 258, 11, doi: [10.3847/1538-4365/ac3078](https://doi.org/10.3847/1538-4365/ac3078)
- Whitaker, K. E., Labb  , I., van Dokkum, P. G., et al. 2011, *ApJ*, 735, 86, doi: [10.1088/0004-637X/735/2/86](https://doi.org/10.1088/0004-637X/735/2/86)
- Whitaker, K. E., van Dokkum, P. G., Brammer, G., et al. 2013, *ApJL*, 770, L39, doi: [10.1088/2041-8205/770/2/L39](https://doi.org/10.1088/2041-8205/770/2/L39)
- Whitaker, K. E., Franx, M., Leja, J., et al. 2014, *ApJ*, 795, 104, doi: [10.1088/0004-637X/795/2/104](https://doi.org/10.1088/0004-637X/795/2/104)
- Whitaker, K. E., Narayanan, D., Williams, C. C., et al. 2021a, *ApJL*, 922, L30, doi: [10.3847/2041-8213/ac399f](https://doi.org/10.3847/2041-8213/ac399f)
- Whitaker, K. E., Williams, C. C., Mowla, L., et al. 2021b, *Nature*, 597, 485, doi: [10.1038/s41586-021-03806-7](https://doi.org/10.1038/s41586-021-03806-7)
- Williams, C. C., Curtis-Lake, E., Hainline, K. N., et al. 2018, *ApJS*, 236, 33, doi: [10.3847/1538-4365/aabccb](https://doi.org/10.3847/1538-4365/aabccb)
- Williams, C. C., Spilker, J. S., Whitaker, K. E., et al. 2021, *ApJ*, 908, 54, doi: [10.3847/1538-4357/abcf6](https://doi.org/10.3847/1538-4357/abcf6)
- Williams, R. J., Quadri, R. F., Franx, M., van Dokkum, P., & Labb  , I. 2009, *ApJ*, 691, 1879, doi: [10.1088/0004-637X/691/2/1879](https://doi.org/10.1088/0004-637X/691/2/1879)
- Williams, R. J., Quadri, R. F., Franx, M., et al. 2010, *ApJ*, 713, 738, doi: [10.1088/0004-637X/713/2/738](https://doi.org/10.1088/0004-637X/713/2/738)
- Wold, I. G. B., Kawinwanichakij, L., Stevans, M. L., et al. 2019, *ApJS*, 240, 5, doi: [10.3847/1538-4365/aaee85](https://doi.org/10.3847/1538-4365/aaee85)
- Wolf, C., Arag  n-Salamanca, A., Balogh, M., et al. 2009, *Monthly Notices of the Royal Astronomical Society*, 393, 1302, doi: [10.1111/j.1365-2966.2008.14204.x](https://doi.org/10.1111/j.1365-2966.2008.14204.x)

Wuyts, S., Franx, M., Cox, T. J., et al. 2009, ApJ, 696, 348,
doi: [10.1088/0004-637X/696/1/348](https://doi.org/10.1088/0004-637X/696/1/348)

Wuyts, S., Förster Schreiber, N. M., Genzel, R., et al. 2012, ApJ,
753, 114, doi: [10.1088/0004-637X/753/2/114](https://doi.org/10.1088/0004-637X/753/2/114)