

The MOSDEF survey: probing resolved stellar populations at $z \sim 2$ Using a new bayesian-defined morphology metric called patchiness

Tara Fetherolf¹,¹★ Naveen A. Reddy,¹ Alice E. Shapley,² Mariska Kriek,³ Brian Siana,¹ Alison L. Coil,⁴ Bahram Mobasher,¹ William R. Freeman,¹ Sedona H. Price⁵, Ryan L. Sanders⁶,⁶† Irene Shivaie⁷,⁷† Mojegan Azadi,⁸ Laura de Groot,⁹ Gene C.K. Leung¹⁰ and Tom O. Zick³

¹Department of Physics & Astronomy, University of California, Riverside, 900 University Ave., Riverside, CA 92521, USA

²Department of Physics & Astronomy, University of California, Los Angeles, 430 Portola Plaza, Los Angeles, CA 90095, USA

³Astronomy Department, University of California at Berkeley, Berkeley, CA 94720, USA

⁴Center for Astrophysics and Space Sciences, Department of Physics, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA

⁵Max-Planck-Institut Für Extraterrestrische Physik, Postfach 1312, Garching, D-85741, Germany

⁶Department of Physics, University of California, Davis, 1 Shields Avenue, Davis, CA 95616, USA

⁷Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721, USA

⁸Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA

⁹Department of Physics, The College of Wooster, 1189 Beall Avenue, Wooster, OH 44691, USA

¹⁰Department of Astronomy, University of Texas at Austin, Austin, TX, 78712, USA

Accepted 2022 November 11. Received 2022 October 19; in original form 2022 January 7

ABSTRACT

We define a new morphology metric called ‘patchiness’ (P) that is sensitive to deviations from the average of a resolved distribution, does not require the galaxy centre to be defined, and can be used on the spatially resolved distribution of any galaxy property. While the patchiness metric has a broad range of applications, we demonstrate its utility by investigating the distribution of dust in the interstellar medium (ISM) of 310 star-forming galaxies at spectroscopic redshifts $1.36 < z < 2.66$ observed by the MOSFIRE Deep Evolution Field survey. The stellar continuum reddening distribution, derived from high-resolution multiwaveband CANDELS/3D-*HST* imaging, is quantified using the patchiness, Gini, and M_{20} coefficients. We find that the reddening maps of high-mass galaxies, which are dustier and more metal-rich on average, tend to exhibit patchier distributions (high P) with the reddest components concentrated within a single region (low M_{20}). Our results support a picture where dust is uniformly distributed in low-mass galaxies ($\lesssim 10^{10} M_{\odot}$), implying efficient mixing of dust throughout the ISM. On the other hand, the dust distribution is patchier in high-mass galaxies ($\gtrsim 10^{10} M_{\odot}$). Dust is concentrated near regions of active star formation and dust mixing time-scales are expected to be longer in high-mass galaxies, such that the outskirt regions of these physically larger galaxies remain relatively unenriched. This study presents direct evidence for patchy dust distributions on scales of a few kpc in high-redshift galaxies, which previously has only been suggested as a possible explanation for the observed differences between nebular and stellar continuum reddening, star formation rate indicators, and dust attenuation curves.

Key words: methods: data analysis – dust, extinction – galaxies: evolution – galaxies: high-redshift – galaxies: ISM – galaxies: structure.

1 INTRODUCTION

Galaxy morphology – or the observed structure of galaxies based on the distribution of their stars, gas, and dust – is an important tool for understanding how galaxies assemble across cosmic time. The morphology of a galaxy can most fundamentally be classified based on its visual structure, as is done when identifying galaxies on the well-known ‘Hubble sequence’ (Hubble 1926; de Vaucouleurs 1959). While the classification of galaxies has significantly advanced by crowd-sourcing volunteers for visual classification (e.g. Galaxy Zoo; Lintott et al. 2008, 2011; Kartaltepe et al. 2015) and developing

machine-learning algorithms to train computers (e.g. Banerji et al. 2010; Dieleman, Willett & Dambre 2015; Domínguez Sánchez et al. 2018), the disc-, bulge-, and bar-like structures used in visual classifications of local galaxies are not typically observed in high-redshift galaxies, which instead appear clumpy and irregular in shape (e.g. Griffiths et al. 1994; Dickinson 2000; van den Bergh 2002; Papovich et al. 2005; Shapley 2011; Law et al. 2012; Conselice 2014; Guo et al. 2015, 2018). However, quantifying the morphology of irregularly shaped high-redshift galaxies, especially at $z \sim 2$ when galaxies were rapidly assembling their stellar mass (see Madau & Dickinson 2014), is a critical step in understanding how they evolve into the ordered structures that are observed in the local Universe.

In this regard, there are also quantitative morphology metrics that are dependent on the distribution of flux from images at one

★ E-mail: tara.fetherolf@gmail.com

† Hubble Fellow

or two wavebands rather than visually defined patterned structures, such as the Sérsic index (Sérsic 1963), bulge-to-disc light ratio (i.e. GALFIT; Peng et al. 2002, 2010), CAS parameters (i.e. concentration, asymmetry, and clumpiness; Conselice 2003), Gini coefficient (Abraham, van den Bergh & Nair 2003; Lotz, Primack & Madau 2004), second-order moment of light (i.e. M_{20} ; Lotz et al. 2004, 2008), ICD (internal color dispersion; Papovich et al. 2005), and MID statistics (multimode, intensity, and deviation; Freeman et al. 2013). Quantitative metrics may be more appropriate for defining the morphology of high-redshift galaxies, but several of these metrics still require a well-defined centre for the galaxy (e.g. Sérsic index, concentration, M_{20}), which is not trivial to define for clumpy, irregularly shaped galaxies. Furthermore, since these metrics were originally designed to be used on resolved images at only one or two wavebands, they are not necessarily suitable for probing the distribution of resolved physical properties in galaxies that are inferred from their multiwavelength photometry.

The distribution of stellar mass and sites of recent star formation can be used to investigate the efficiency of star formation (e.g. Lang et al. 2014; Jung et al. 2017; Tacchella et al. 2018), the history of merging galaxies (e.g. Conselice 2003; Lotz et al. 2004, 2008; Cibinel et al. 2015), and the overall assembly of galaxies (e.g. Wuyts et al. 2012; Hemmati et al. 2014; Boada et al. 2015). Revealing the intrinsic properties of stellar populations within galaxies requires a correction for the obscuring effects of dust, which depends on the physical properties of the dust grains, the total amount of dust, and its distribution relative to the stars (e.g. Draine & Lee 1984; Fitzpatrick & Massa 1986; Calzetti, Kinney & Storchi-Bergmann 1994; Charlot & Fall 2000). Patchy or clumpy dust distributions have been theorized as the cause behind observed variations in reddening (e.g. Calzetti et al. 1994; Wild et al. 2011; Price et al. 2014; Reddy et al. 2015, 2018a, 2020) and star formation rates (SFRs; e.g. Boquien et al. 2009, 2015; Hao et al. 2011; Reddy et al. 2015; Katsianis et al. 2017; Fetherolf et al. 2021) that are deduced from different probes, such as UV and $H\alpha$. Furthermore, the shape and slope of the dust attenuation curve has been found to vary with galactic properties both in local and high-redshift galaxies (e.g. Reddy et al. 2006, 2010, 2018b; Leja et al. 2017; Salim, Boquien & Lee 2018; Shivaee et al. 2020), and these variations could also be explained by differences in their dust distribution.

In this study, we investigate the inferred distribution of dust in the interstellar medium (ISM) of high-redshift galaxies based on their resolved stellar continuum reddening maps $[E(B - V)_{\text{stars}}]$. To achieve this goal, we define a new general morphology metric, called ‘patchiness,’ which is sensitive to deviations that are both above and below the average of for a given resolved distribution and can be used to probe any resolved property, such as the flux distribution, stellar population and reddening maps, or spatially resolved emission line measurements. Resolved stellar population and robust reddening maps are constructed using the high-resolution multiwaveband *Hubble Space Telescope* (*HST*) imaging from the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS; Grogin et al. 2011; Koekemoer et al. 2011). We also use spectroscopic redshifts and emission line measurements from the MOSFIRE Deep Evolution Field (MOSDEF; Kriek et al. 2015) survey to mitigate degeneracies in the resolved SED fitting, and to derive globally averaged nebular reddening $[E(B - V)_{\text{gas}}]$ and gas-phase metallicities. The MOSDEF survey obtained rest-frame optical spectra for ~ 1500 star-forming and AGN galaxies that also have been observed through CANDELS and the 3D-*HST* grism survey (Brammer et al. 2012; Skelton et al. 2014). The legacy of these data is such that we can probe how the distribution of dust in

the ISM evolves with globally measured properties for a statistically large sample of 310 galaxies at $1.36 < z < 2.66$.

The data and sample selection are introduced, and the methodology for constructing resolved stellar population and robust reddening maps is outlined in Section 2. The new morphology metric, deemed ‘patchiness,’ is defined in Section 3 alongside the Gini and M_{20} coefficients that are additionally used to probe the distribution of resolved stellar continuum reddening. Our results are presented in Section 4 with a discussion of how the ISM of high-redshift galaxies evolves with stellar mass and gas-phase metallicity. Finally, we summarize our findings in Section 5.

A cosmology that assumes $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_{\Lambda} = 0.7$, and $\Omega_m = 0.3$ is used throughout this work. The vacuum wavelengths of emission lines are used and magnitudes are presented in the AB system (Oke & Gunn 1983).

2 DATA, SAMPLE SELECTION, AND STELLAR POPULATION AND REDDENING MAPS

In this section, we present the data, sample selection, and methodology for creating resolved stellar population and reddening maps. We introduce the CANDELS resolved imaging and 3D-*HST* photometric catalogues in Section 2.1. An overview of the MOSDEF survey is presented in Section 2.2 and the sample selection is defined in Section 2.3. In Section 2.4, we outline our methods for processing the resolved photometry and, finally, our SED-fitting assumptions are listed in Section 2.5.

2.1 CANDELS/3D-*HST* photometry

We use *HST*-resolved imaging that was obtained by CANDELS (Grogin et al. 2011; Koekemoer et al. 2011). Specifically, we use the *HST*/ACS and *HST*/WFC3 instruments to obtain imaging in the $F435W$, $F606W$, $F775W$, $F814W$, $F850LP$, $F125W$, $F140W$, and $F160W$ filters (B_{435} , V_{606} , i_{775} , I_{814} , z_{850} , J_{125} , JH_{140} , and H_{160}). CANDELS imaging covers $\sim 900 \text{ arcmin}^2$ of the well-studied AEGIS, COSMOS, GOODS-N, GOODS-S, and UDS extragalactic fields to a 90 percent completeness at $\sim 25 \text{ mag}$ in the H_{160} filter. We use the reprocessed CANDELS imaging that has been made publicly available¹ by the 3D-*HST* grism survey team (Brammer et al. 2012; Skelton et al. 2014; Momcheva et al. 2016). The reprocessed *HST* images have been drizzled to a $0''.06 \text{ pixel}^{-1}$ scale and spatially smoothed to the $0''.18$ resolution of the H_{160} images. We also utilize the 3D-*HST* broad-band catalogue that includes ancillary ground- and space-based photometry at 0.3 to $0.8 \mu\text{m}$ covering the CANDELS extragalactic fields. The 3D-*HST* v4.0 catalogue galaxy IDs are used throughout this paper.

2.2 MOSDEF spectroscopy

The MOSDEF survey (Kriek et al. 2015) obtained rest-frame optical spectra for ~ 1500 star-forming galaxies and AGN. Observations for the MOSDEF survey were taken using the 10-m Keck I telescope MOSFIRE multiobject spectrograph (McLean et al. 2010, 2012) with the $0''.7$ slit widths in the Y , J , H , and K bands ($R = 3400, 3000, 3650$, and 3600). The CANDELS imaging is used to select targets for the MOSDEF survey down to a stellar mass limit of $\sim 10^9 M_{\odot}$, corresponding to an H_{160} -band limit of $24.0, 24.5$, and 25.0 mag in three respective redshift bins: $1.37 < z < 1.70$, $2.09 < z < 2.61$,

¹<https://3dhst.research.yale.edu/>

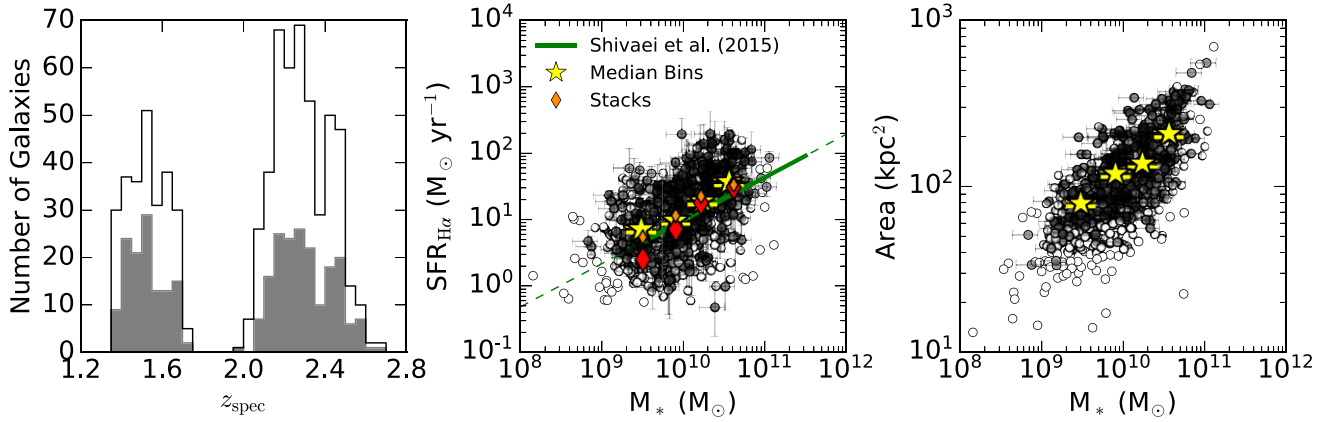


Figure 1. Left-hand panel: The spectroscopic redshift distribution of the star-forming galaxies in the sample at $1.24 < z < 2.66$. The white and grey histograms show the parent (735 galaxies) and final (310 galaxies) samples, respectively. Middle panel: the $H\alpha$ SFR versus stellar mass ‘main-sequence’ relationship. The white empty and grey filled points show the parent and final samples, respectively. $H\alpha$ emission is corrected for dust based on the Balmer decrement and the Cardelli, Clayton & Mathis (1989) extinction curve. $H\alpha$ SFRs are calculated assuming a stellar metallicity of $Z_* = 0.2 Z_\odot$, a Bruzual & Charlot (2003) stellar population synthesis model, and a Chabrier (2003) IMF. The (Shivaei et al. 2015) SFR– M_* ‘main-sequence’ linear relation that best fits the first two years of MOSDEF data is shown by the green line. Stacks of the spectra in four bins of stellar mass containing an equal number of objects are shown by the diamond symbols, where the stacks from the parent and final samples are shown by the large red and small orange diamonds, respectively. The yellow stars show the median $H\alpha$ SFR and M_* of the four bins in stellar mass for the final sample, with the yellow bars indicating the standard error in the median $H\alpha$ SFRs. Right-hand panel: The size versus mass relation for galaxies in the sample, where the area of each galaxy is measured from the detected regions in the 3D-HST segmentation maps. The colours and symbols are the same as in the middle panel, except the yellow stars show the median segmentation map areas in four bins of stellar mass and the yellow bars indicate the standard error in the median sizes. Compared to the parent sample that represents the broader population of galaxies observed as part of the MOSDEF survey, galaxies in the final sample tend to be marginally biased against low-mass and compact galaxies due to the S/N and resolution requirements enforced by the multifilter Voronoi binning technique (see Section 2.4 and Fetherolf et al. 2020).

and $2.95 < z < 3.80$ (hereafter, the $z \sim 1.5$, ~ 2.3 , and ~ 3.3 samples). These redshift ranges are selected so that several strong rest-frame optical emission lines could be observed in the near-IR windows of atmospheric transmission, including: $[O II]\lambda\lambda 3727, 3730$, $H\beta$, $[O III]\lambda\lambda 4960, 5008$, $H\alpha$, $[N II]\lambda\lambda 6550, 6585$, and $[S II]\lambda 6718, 6733$. At least one slit star is included on each slit mask for the absolute flux calibration of the spectra and to correct for slit loss. In this work, we utilize the spectroscopic redshifts and emission line measurements from the MOSDEF survey, and refer the reader to Kriek et al. (2015) for more information regarding the observing strategy and data-reduction procedures for the MOSDEF survey.

In order to measure line fluxes, a Gaussian function is fit to each emission line with an underlying linear fit to the continuum. Double and triple Gaussians are fit to the $[O II]\lambda\lambda 3727, 2730$ doublet and $H\alpha + [N II]\lambda\lambda 6550, 6585$ lines, respectively. $H\alpha$ and $H\beta$ emission line fluxes are corrected for Balmer absorption using the stellar population model that best fits the observed 3D-HST broad-band photometry (see Section 2.5). To obtain flux errors, the 1D spectra are perturbed 1000 times by their error spectra, line fluxes are remeasured, and the 1σ errors are assumed to be the 68th-percentile width of the distribution. The spectroscopic redshift is determined from the observed wavelength of the highest signal-to-noise ratio (S/N) line (typically $H\alpha$ or $[O III]\lambda 5008$). Refer to Kriek et al. (2015) and Reddy et al. (2015) for more details on the emission line flux measurements for the MOSDEF survey.

2.3 Sample selection

Galaxies with robust spectroscopic redshifts and coverage of $H\alpha$ and $H\beta$ emission lines are included in the sample. Robust spectroscopic measurements are based on at least two strong emission features with an $S/N \geq 2$. AGNs are identified and removed from the sample using their X-ray luminosities, optical emission line ratios ($\log([N II]/H\alpha)$

> -0.3), and/or mid-IR luminosities (Coil et al. 2015; Azadi et al. 2017, 2018; Leung et al. 2019). These criteria produce a parent sample of 735 galaxies at $1.24 < z < 2.66$. Additional S/N and resolution constraints are applied to the photometry using the multifilter Voronoi binning technique outlined by Fetherolf et al. (2020) in order to obtain robust dust reddening maps (also see Section 2.4), resulting in a final sample of 310 star-forming galaxies at $1.36 < z < 2.66$.

Fig. 1 shows the spectroscopic redshift distribution, SFR– M_* , and size– M_* relations for the galaxies in the parent (white histogram and empty points) and final samples (grey histogram and filled grey points). The solid green line shows the (Shivaei et al. 2015) main-sequence relationship derived from the first two years of data from the MOSDEF survey. Sizes of galaxies are defined by their segmentation map surface area (see Section 2.4). The objects in the samples are equally divided into four bins of stellar mass and the spectra are stacked using SPECLINE² (Shivaei et al. 2018). The red diamonds show the stacked $H\alpha$ SFR measurements for the binned parent (large red diamonds) and final (small orange diamonds) samples and the yellow stars show the median bins of the individual measurements for galaxies in the final sample, with the yellow bars indicating the standard error in the median $H\alpha$ SFRs. The distribution of the outlying empty white points and the 0.33 dex difference between the stacked points of the lowest mass bin between the two samples suggests that the final sample is marginally biased against low-mass and low-SFR galaxies relative to the galaxies in the parent sample. These galaxies tend to be compact and/or faint such that they do not have sufficient S/N that is necessary for reliably measuring spatially resolved fluxes across several filters, as is required by the multifilter Voronoi binning technique that is described in Section 2.4 (also see

²<https://github.com/IreneShivaei/specline/>

Fetherolf et al. 2020). Separating the centre and right-hand panels of Fig. 1 by the $z \sim 1.5$ and 2.3 samples reveals the evolution of the SFR– M_* and size– M_* relationships (e.g. van der Wel et al. 2014), but otherwise there are no significant differences between the two sub-samples.

2.4 Resolved photometry

The methodology for constructing stellar population and robust reddening maps is briefly outlined here, but we refer the reader to Fetherolf et al. (2020) for more detail. Individual galaxies in the sample are separated into sub-images that are 80×80 pixels in size (4.8×4.8 arcsec², or $\sim 40 \times 40$ kpc at $z \sim 2$). Pixels that are not associated with the central galaxy in each sub-image are masked using the noise-equalized $J_{125} + JH_{140} + H_{160}$ SOURCE-EXTRACTOR (Bertin & Arnouts 1996) segmentation map provided by the 3D-*HST* survey team (Skelton et al. 2014).

In order to group low S/N pixels and avoid correlated signal between individual pixels, the pixels associated with the galaxy are binned using an adaptive Voronoi binning technique (Cappellari & Copin 2003) that has been modified (see Fetherolf et al. 2020) to incorporate the S/N distribution of multiple filters with resolved imaging. Fetherolf et al. (2020) found that additionally considering the S/N distribution of filters at shorter wavelengths, which typically have lower S/N compared to the H_{160} filter, resulted in better constrained resolved SEDs compared to using the H_{160} S/N distribution alone. In particular, the estimated $E(B - V)_{\text{stars}}$ are better constrained when more than one filter is required to reach a certain S/N threshold for each Voronoi bin, which also helps reduce the degeneracy between the SED-measured stellar population ages and dust attenuation. To ensure robust resolved stellar population fits, Voronoi bins are required to satisfy an $S/N \geq 5$ in at least five filters with resolved imaging. Bins that do not satisfy this requirement are deemed as ‘outskirt’ bins and are not included in the analysis.

The counts from the pixels that make up an individual Voronoi bin are summed and converted to an AB magnitude. Magnitude errors are measured similarly by summing the noise in quadrature from the corresponding RMS maps. A minimum magnitude error of 0.05 mag is adopted to prevent any single photometric point from driving the best-fitting resolved SED. The resolved CANDELS photometry is additionally supplemented with unresolved *Spitzer*/IRAC photometry (3.6, 4.5, 5.8, and 8.0 μm ; Skelton et al. 2014) in order to constrain stellar masses and stellar population ages. In order to incorporate the *Spitzer*/IRAC photometry into the resolved photometry, a constant H_{160} –IRAC³ colour is assumed by dividing the total IRAC fluxes by the H_{160} flux within each Voronoi bin (see Fetherolf et al. 2020).

2.5 SED fitting

SED-derived $E(B - V)_{\text{stars}}$, stellar population ages, SFRs, and stellar masses are computed for both the galaxy as a whole from the unresolved 3D-*HST* broad-band photometry,⁴ and for individual Voronoi bins. The best-fitting (Bruzual & Charlot 2003) stellar population synthesis model is determined using χ^2 minimization relative to the photometry (see Reddy et al. 2012). A Chabrier

(2003) initial mass function (IMF) and constant SFHs are assumed. Stellar ages are permitted to vary between 50 Myr and the age of the Universe at the spectroscopic redshift of the galaxy.⁵ Sub-solar metallicities ($0.2 Z_{\odot}$) and an SMC extinction curve (Fitzpatrick & Massa 1990; Gordon et al. 2003) are assumed and reddening is allowed to range between $0.0 \leq E(B - V)_{\text{stars}} \leq 0.4$.⁶ Examples of the resultant stellar population and reddening maps are shown for UDS 25642 ($z = 1.42$) and GOODS-N 12345 ($z = 2.27$) in Fig. 2. The number in the top left-hand corner of each panel shows the average $E(B - V)_{\text{stars}}$, average stellar population age, total summed SFR, and total summed stellar mass obtained from the Voronoi bins. The low S/N ‘outskirt’ components that are not used throughout this analysis are indicated by the grey regions in each panel.

Measuring resolved SED parameter errors can be computationally expensive for a large number of Voronoi bins and galaxies. Therefore, a subset of 50 galaxies that have a range of stellar population parameters that are representative of the larger sample are selected to obtain typical SED parameter errors. The measured unresolved and resolved photometric fluxes are perturbed by their errors and refit 100 times. The 68 models with the lowest χ^2 are used to obtain the 1σ uncertainties in the SED properties derived from the unresolved photometry and individual Voronoi bins. The average SED parameter errors of the Voronoi bins for an individual galaxy are obtained by taking the average of the individual Voronoi bin parameter errors (summed in quadrature) within the galaxy. Finally, the typical resolved SED parameter errors are obtained by taking the average of the parameter errors determined from the 50 galaxies in the subsample. The typical unresolved SED parameter errors are 0.01 in $E(B - V)_{\text{stars}}$, 0.20 dex in log stellar population age, 0.05 dex in log SFR, and 0.16 dex in log stellar mass. The typical resolved SED parameter errors based on the individual Voronoi bins are 0.03 in $E(B - V)_{\text{stars}}$, 0.21 dex in log stellar population age, 0.16 dex in log SFR, and 0.10 dex in log stellar mass.

3 MORPHOLOGY METRICS

Studying the structure of galaxies at high redshifts is key towards understanding how galaxies build their stellar mass. Commonly used metrics for quantifying galactic structure include Sérsic surface brightness profiles (Sérsic 1963), light concentration, asymmetry,

⁵Reddy et al. (2012) found that if stellar ages are restricted to being older than the typical dynamical time-scale, then either constant SFHs or exponentially rising SFHs best reproduce the SFRs of $z \sim 2$ galaxies compared to other tracers such as IR+UV. Alternatively assuming exponentially rising SFHs results in stellar population ages that are on average ~ 30 per cent older than those determined when assuming constant SFHs (Reddy et al. 2012). For our sample, we find that the SED-derived SFRs measured when assuming exponentially rising and declining SFHs are typically within 0.03 dex (higher and lower, respectively) of those measured when assuming constant SFHs, which is within the typical uncertainty of the SED-derived SFRs (0.05 dex).

⁶Assuming an SMC extinction curve Fitzpatrick & Massa (1990), Gordon et al. (2003) and sub-solar metallicities, opposed to a Calzetti et al. (2000) attenuation curve and solar metallicities, does not affect the relative order of mass measurements (Reddy et al. 2018a). The reddening of the observed rest-frame $B - V$ colours can be reproduced by a combination of the $E(B - V)$ and attenuation curve. The same range of observed rest-frame $B - V$ colours can be reproduced using either a steep attenuation curve and an $E(B - V)$ with a small dynamic range [such as SMC and $0.0 \leq E(B - V)_{\text{stars}} \leq 0.4$] or a shallow attenuation curve and an $E(B - V)$ with a larger dynamic range [such as Calzetti and $0.0 \leq E(B - V)_{\text{stars}} \leq 0.6$]. Therefore, the results presented here would not significantly change if a Calzetti et al. (2000) attenuation curve and solar metallicities were alternatively assumed.

³Fetherolf et al. (2020) showed that constraining the H_{160} –IRAC does not unduly influence the degeneracy between stellar population ages and attenuation within galaxies.

⁴The results do not significantly change when alternatively using SED parameters derived from the integrated Voronoi bin fluxes.

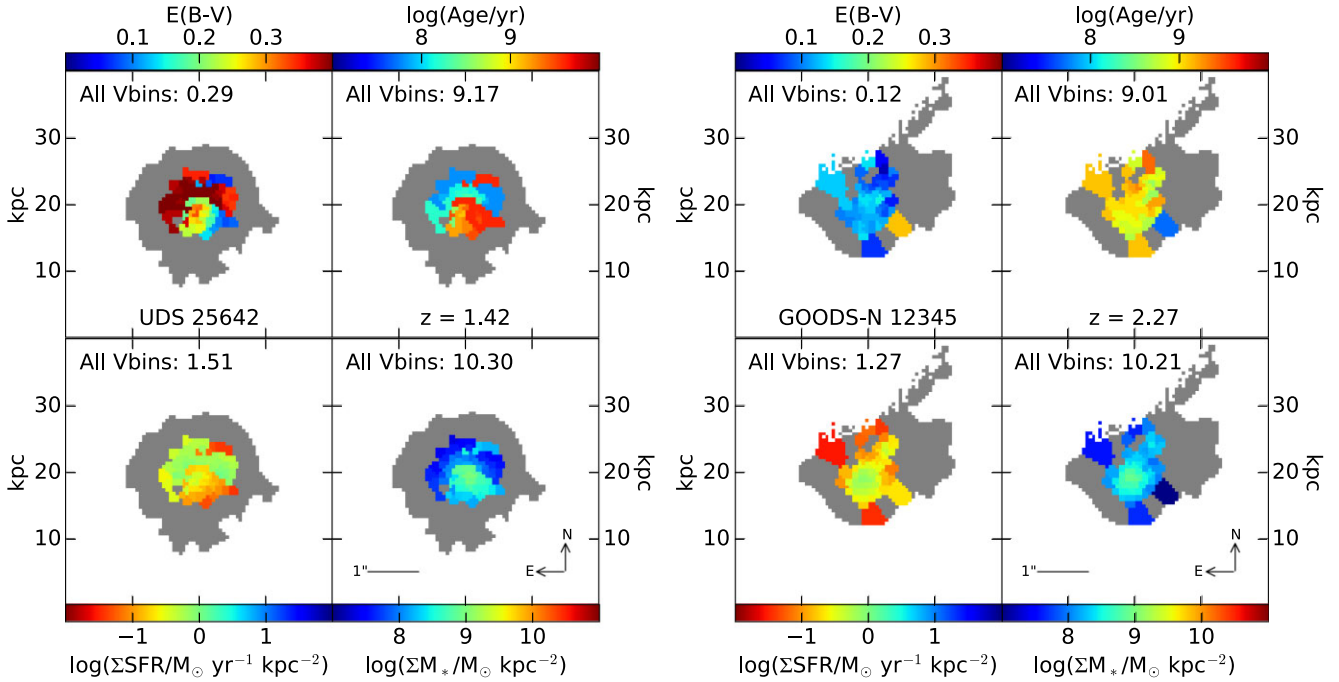


Figure 2. Examples of the stellar population and reddening maps for two example galaxies in our sample from each redshift bin. The left-hand panels show UDS 25642 ($z = 1.42$) and the right-hand panels show GOODS-N 12345 ($z = 2.27$). The individual panels for a single galaxy show the distribution of $E(B - V)_{\text{stars}}$ (top left), stellar population age (top right-hand panel), SFR surface density (bottom left-hand panel), and stellar mass surface density (bottom right-hand panel). The average $E(B - V)_{\text{stars}}$ and stellar ages weighted by the Voronoi bin areas are shown by the values inside the top left and top right-hand panels, respectively. The total log SFR and total log stellar mass from all of the Voronoi bins are shown by the values inside the bottom left- and bottom right-hand panels, respectively. The grey regions indicate low S/N ‘outskirt’ regions that are not used in the analysis.

clumpiness (CAS; Conselice 2003), Gini coefficient, second-order moment of light (Gini- M_{20} ; Abraham et al. 2003; Lotz et al. 2004), and internal colour dispersion (ICD; Papovich et al. 2003). All of these measures either require a well-defined centroid, are based on one- to two-filter photometry, or focus exclusively on the brightest parts of the light distribution. Alternatively, we aim to quantify structure in a way that is not dependent on a centroid – especially considering that galaxies at progressively higher redshifts have increasingly irregular structure such that a ‘centre’ may not be clearly defined (e.g. Griffiths et al. 1994; Dickinson 2000; van den Bergh 2002; Papovich et al. 2005; Shapley 2011; Law et al. 2012; Conselice 2014; Guo et al. 2015, 2018).

In Section 3.1, we introduce ‘patchiness’ (P) as a new morphology metric, which quantifies the dispersion of light (or any resolved property) and is sensitive to both faint and bright regions of light without requiring a defined centre. The definitions and choice of the Gini (G) and M_{20} coefficients to be used alongside patchiness are explained in Section 3.2. Finally, while these morphology metrics could be used on any SED parameters, we have chosen in this paper to focus on how P , G , and M_{20} is applied to the $E(B - V)_{\text{stars}}$ distribution, which is specifically described in Section 3.3.

3.1 The patchiness metric

Patchiness, P , is defined using a Bayesian maximum-likelihood technique (for a recent review of Bayes’ Theorem, see Sharma 2017) that measures the Gaussian probability that every individual resolved element, X_i , equals the weighted average of the distribution, $\bar{X}_w$. Since the stellar population and dust maps that we are using do not have equally sized bins, we choose to weight each resolved element

(i.e. Voronoi bin) by their area.⁷ The weighted average for a Voronoi bin distribution is defined by

$$\bar{X}_w = \frac{\sum_{i=1}^{N_{\text{Vbins}}} n_{\text{pix},i} X_i}{\sum_{i=1}^{N_{\text{Vbins}}} n_{\text{pix},i}}, \quad (1)$$

where N_{Vbins} is the number of Voronoi bins within a single galaxy and n_{pix} is the number of pixels contained within each Voronoi bin. Note that the denominator of equation (1) is simply the total area of (or number of pixels contained within) the galaxy. Patchiness is used to quantify the resolved distribution of some parameter within a single galaxy and is defined by

$$P = -\ln \left\{ \prod_{i=1}^{N_{\text{Vbins}}} \frac{1}{\sqrt{2\pi}\sigma_i} \exp \left[-\frac{(X_i - \bar{X}_w)^2}{2\sigma_i^2} \right] \right\}, \quad (2)$$

where X_i is the value of a single resolved element (Voronoi bins, in this case), $\bar{X}_w$ is the weighted (or non-weighted) average of the distribution, and σ_i is the parameter uncertainty within each resolved element. The patchiness metric is defined such that the measured P value will be low when the resolved elements equal the weighted average (high likelihood) and P will be high when there are deviations from the average (low likelihood). For the resolved distributions probed in this study, the range of P spans several orders of magnitude. In order to show $\log P$ in the figures, all calculated P values are offset by a constant such that $P_{\text{min}} = 1$.

The patchiness metric is unique compared to other traditional morphology metrics in that it does not require a defined centroid

⁷Alternatively weighting by either the H_{160} flux or log stellar mass of individual Voronoi bins does not significantly affect our results.

or radius, it is sensitive to below-average outliers in addition to deviations above the mean, and it is insensitive to properties with large dynamic ranges (e.g. stellar mass) or values that approach zero [e.g. $E(B - V)$]. Patchiness is most comparable to the ICD (Papovich et al. 2003), which measures the dispersion of light between the resolved imaging obtained in two different filters for a given galaxy. However, patchiness can be directly applied to any resolved property – including those derived from resolved SED fitting, which incorporates the flux distribution in several photometric bands.

It is important to note that the patchiness metric (and other quantitative metrics) should only be used to compare the relative patchiness of resolved distributions for galaxies that make up a uniformly defined and analyzed sample. The calculated values of patchiness will vary depending on how the resolved distribution is defined, such as the choice of binning and the method for defining which resolved elements are included in the morphology analysis for a single galaxy. An in-depth analysis of the patchiness metric is presented in Appendix A using the resolved SED-derived $E(B - V)_{\text{stars}}$, stellar population ages, SFRs, and stellar masses. Our conclusions from Appendix A are summarized briefly as follows. (1) Galaxies that are patchier in one resolved SED-derived property are more likely to be patchier in other resolved SED-derived properties, possibly due to dependencies between SED-derived parameters. (2) The ICD is most significantly correlated with the patchiness of the $E(B - V)_{\text{stars}}$ distribution compared to other SED-derived properties. (3) Patchiness is significantly correlated with the Gini coefficient (also see Section 3.2). (4) Caution should be taken when measuring the patchiness on a pixel-by-pixel scale due to correlated signal between neighbouring pixels and stochastic measurements from pixels with low S/N. (5) Patchiness values should only be compared between galaxies at similar redshifts, as the patchiness values are not preserved due to surface brightness dimming. Inconsistencies between quantified galaxy structure when measured across a range of redshifts is a common issue with non-parametric morphology indicators (e.g. Giavalisco et al. 1996; Conselice 2003, 2014; Lisker 2008) and, thus, is not unique to the performance of the patchiness metric. For this reason, the subsequent analyses are separated by the $z \sim 1.5$ and 2.3 redshift bins. (6) Patchiness is not biased towards higher values for galaxies with more resolved elements, but galaxies with more resolved elements may be intrinsically patchier for some resolved properties (such as stellar mass).

Finally, it is important to emphasize that the morphology of galaxies can be best understood when several morphology metrics are used together. In the present study, we choose to pair patchiness with the Gini and M_{20} coefficients, which are sensitive to the resolved elements with the highest measured values or inferred quantities (e.g. brightest or highest mass regions). Morphology metrics that are sensitive to outliers with specific characteristics in a resolved distribution (i.e. high flux or high mass), such as Gini and M_{20} , complement the patchiness metric, which is sensitive to any type of outlier (high or low) in a resolved distribution.

3.2 The Gini coefficient and M_{20}

The Gini coefficient is most well known from economics, where it is used to measure the distribution of wealth in a population. Abraham et al. (2003) adapted the Gini coefficient to measure the concentration of light within galaxies based on their resolved imaging. After sorting the resolved measurements from lowest to highest flux, the Gini

coefficient is calculated by

$$G = \frac{1}{|\bar{X}_w|n(n-1)} \sum_{i=1}^n (2i - n - 1)|X_i|, \quad (3)$$

where X_i is the flux within an individual resolved element (i.e. Voronoi bin), n is the total number of resolved elements, and $\bar{X}_w$ is the weighted average of the resolved distribution. If all n resolved elements have an equal amount of flux, then $G = 0$. Conversely, if a single resolved element has all of the flux, then $G = 1$. While a high Gini value implies a high concentration of flux, the flux is not necessarily concentrated in a single region since there is no spatial information in the metric.

Lotz et al. (2004) introduced the M_{20} parameter, which is the normalized second-order moment of the 20 per cent brightest regions in a galaxy. The centre of the brightest region is defined at (x_c, y_c) , which is determined by minimizing the total second-order moment. The total second-order moment, M_{tot} is defined by

$$M_{\text{tot}} = \sum_i M_i = \sum_{i=1}^n X_i [(x_i - x_c)^2 + (y_i - y_c)^2], \quad (4)$$

where (x_i, y_i) is the location of each resolved element and X_i is the flux of each resolved element. The X_i fluxes are then ordered from highest to lowest flux and M_{20} is defined as

$$M_{20} = \log \left(\frac{\sum M_i}{M_{\text{tot}}} \right), \text{ while } \sum_i X_i < 0.2 X_{\text{tot}}, \quad (5)$$

where only the brightest 20 per cent of resolved elements are included. Like the Gini coefficient, M_{20} is also a measure of the concentration of light, but includes spatial information of the brightest regions. Low M_{20} values (typically negative) indicate that the brightest regions are more significantly grouped within a single region of the galaxy. As M_{20} increases (less negative and towards zero), the brightest regions tend to be more spatially spread across several clusters of light.

The G and M_{20} coefficients complement each other in that G identifies galaxies where there exists regions of high brightness compared to the average, and M_{20} identifies the spatial distribution of the brightest regions. Together, the G and M_{20} coefficients have been used to characterize the morphologies of galaxies (Förster Schreiber et al. 2011; Wuyts et al. 2012; Lee et al. 2018), classify galaxies into traditional Hubble types (Lotz et al. 2004, 2008; Boada et al. 2015), and identify merging galaxies (Lotz et al. 2008; Boada et al. 2015). However, a critical assessment of the Gini coefficient by Lisker (2008) found that the Gini coefficient significantly depends on the S/N of resolved elements and the radius used to determine the inclusion of pixels when calculating G . Therefore, G measurements and Gini-defined classification schemes (e.g. Lotz et al. 2008) cannot be directly compared between studies with differing methodologies or sample characteristics. The same can be said for P , which is why we emphasize in Section 3.1 that P should only be used to compare the patchiness of resolved distributions for galaxies in uniformly defined and analyzed samples at similar redshifts (also see Appendix A). An in-depth comparison between P and the G and M_{20} coefficients is discussed in Appendix A, where we find that P is significantly correlated with G in all resolved SED-derived properties for galaxies in our sample.

3.3 Quantifying the spatial distribution of $E(B - V)_{\text{stars}}$

Typically G and M_{20} are used to quantify the concentration of light at a single wavelength (e.g. Lotz et al. 2004, 2008; Förster

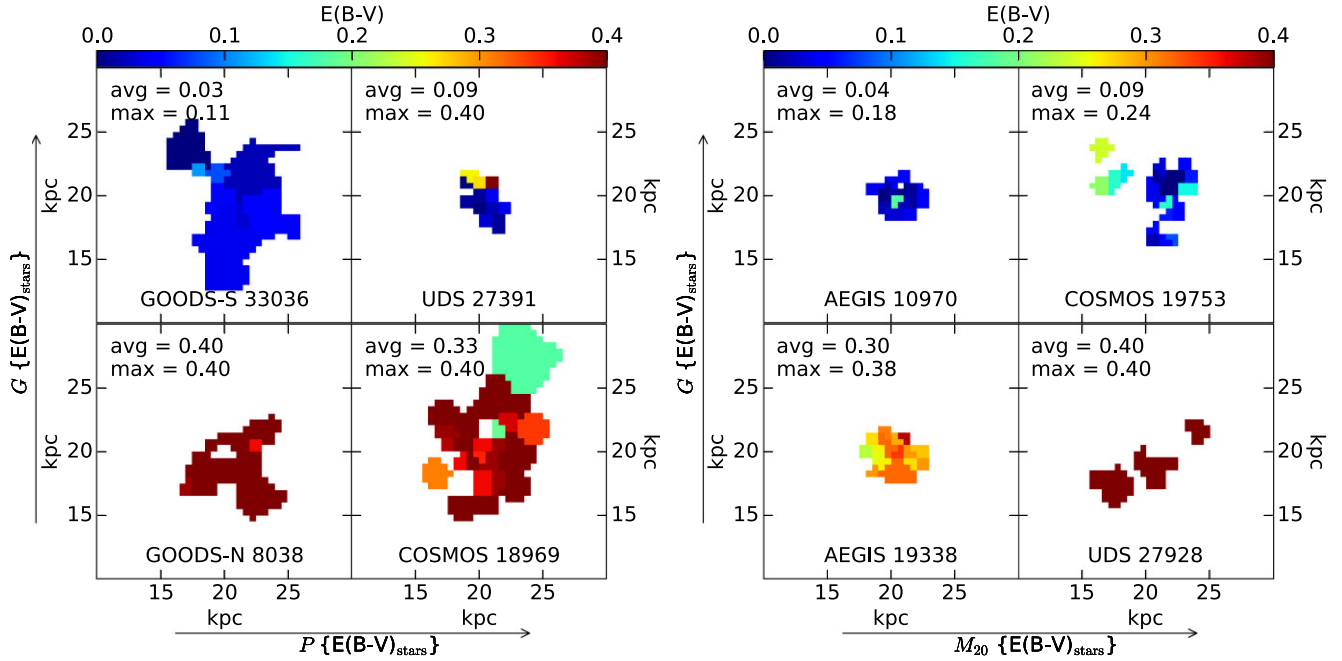


Figure 3. Reddening maps of eight example galaxies that span the P , G , and M_{20} parameter space. The average $E(B - V)_{\text{stars}}$ across each galaxy and the maximum $E(B - V)_{\text{stars}}$ of the resolved components are listed in the top left-hand corner of each panel. Left-hand panel: Patchiness generally probes the dispersion of the resolved $E(B - V)$ maps such that P is sensitive to both the bluest [low $E(B - V)$] and reddest [high $E(B - V)$] regions relative to the average reddening. Gini, on the other hand, is only sensitive to the resolved components that are relatively high in reddening compared to the global average. Right-hand panel: Similar to Gini, M_{20} also probes the reddest regions in the $E(B - V)$ maps, but M_{20} is also sensitive to the spatial distribution of the reddest regions. Lower M_{20} suggests that the reddest regions are contained within a single region of the galaxy, whereas higher M_{20} indicates that the reddest areas are distributed across several spatially separated regions.

Schreiber et al. 2011). However, Wuyts et al. (2012) emphasized the importance of studying multicolour morphologies using G and M_{20} by calculating these metrics on the resolved stellar mass distribution for the first time. We further extend the usage of G and M_{20} by quantifying the concentration of the reddest Voronoi bins [$E(B - V)_{\text{stars,bin}} > \langle E(B - V)_{\text{stars}} \rangle$] in the resolved reddening maps. In the analysis that follows, we pair the concentration of the reddest regions probed by G and M_{20} with the patchiness, P , of the $E(B - V)_{\text{stars}}$ distribution. The G , M_{20} , and P values are calculated using the equations defined in Section 3.1 and Section 3.2, where X_i is the $E(B - V)_{\text{stars}}$ of an individual Voronoi bin, $\bar{X}_w$ is the average $E(B - V)_{\text{stars}}$ [$\langle E(B - V)_{\text{stars}} \rangle$] of the Voronoi bins weighted by their sizes, and σ_i is the uncertainty in the resolved $E(B - V)_{\text{stars}}$ (≈ 0.03 mag; see Section 2.5). Uncertainties in P , G , and M_{20} are determined by perturbing the entire resolved distribution by the uncertainties of the resolved elements and remeasuring P , G , and M_{20} 100 times. The 1σ uncertainty is set to be the range of the 68 perturbations that are closest to the originally measured morphology value. Examples of the $E(B - V)_{\text{stars}}$ maps ordered by their P values are shown in Appendix B.

The revised interpretation for measuring P , G , and M_{20} on the $E(B - V)_{\text{stars}}$ distribution requires redefining ‘bright’ regions as the reddest regions in the distribution, or the Voronoi bins with the highest $E(B - V)_{\text{stars}}$. A high G then indicates the presence of regions with red (high) inferred $E(B - V)_{\text{stars}}$ relative to $\langle E(B - V)_{\text{stars}} \rangle$, and a low G implies a more uniform distribution of reddening between regions that are close to the average value. The spatial continuity of the reddest [highest 20 percent $E(B - V)_{\text{stars}}$] regions is quantified by M_{20} , where a low M_{20} indicates that the reddest regions tend to be concentrated within a single clump and a high M_{20} indicates that the

reddest regions occur in several clumps that are spatially separated across the galaxy. Fig. 3 shows examples of the $E(B - V)_{\text{stars}}$ maps for eight galaxies that span the parameter space between P , G , and M_{20} . It can be seen that the galaxies in the top panels (high G) exhibit larger differences between $\langle E(B - V)_{\text{stars}} \rangle$ and the reddest region [maximum resolved $E(B - V)_{\text{stars}}$] compared to the galaxies shown in the bottom panels (low G). Patchiness, on the other hand, more generally probes the dispersion of a resolved distribution and, thus, is sensitive to both the reddest and bluest regions in the $E(B - V)_{\text{stars}}$ maps [see examples of $E(B - V)_{\text{stars}}$ ordered by P in Appendix B]. Furthermore, from the right set of panels in Fig. 3, it can be seen that the reddest regions (highest resolved $E(B - V)_{\text{stars}}$) are distributed generally within a single region for the galaxies shown in the left sub-panels (low M_{20}), but are spatially spread across the galaxies shown in the right sub-panels (high M_{20}).

A complication when calculating G on the $E(B - V)_{\text{stars}}$ distribution is that G may be artificially high or unphysical ($G > 1$) as $\langle E(B - V)_{\text{stars}} \rangle$ approaches zero (consistent with no dust reddening inferred from the best-fitting SEDs). Therefore, we raise caution when using G on any resolved distribution where the average could be near zero, such as the $E(B - V)_{\text{stars}}$ distribution. The relationship between the $\langle E(B - V)_{\text{stars}} \rangle$ and G calculated on the $E(B - V)_{\text{stars}}$ distribution for galaxies in the sample is shown in Fig. 4 and is separated by the two redshift bins. The regions where $\langle E(B - V)_{\text{stars}} \rangle \leq 0.05$ and ≤ 0.10 mag are shaded in green. While galaxies with low $\langle E(B - V)_{\text{stars}} \rangle$ typically tend to have high G , there are also individual galaxies within the green shaded regions that have relatively low $\langle E(B - V)_{\text{stars}} \rangle$ and low or average G such that G is not necessarily artificially boosted when $\langle E(B - V)_{\text{stars}} \rangle$ is nearly zero. On the high $\langle E(B - V)_{\text{stars}} \rangle$ end, there is a similar effect where low G values

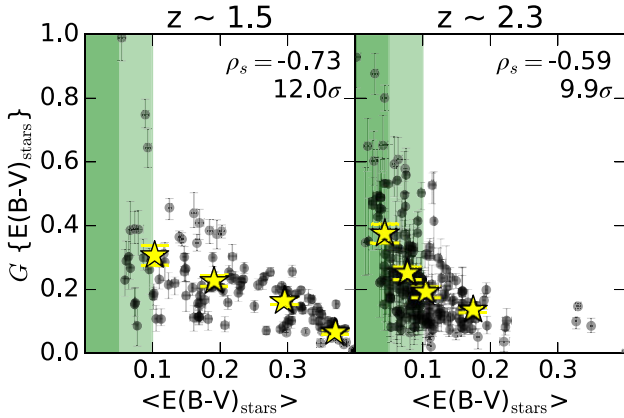


Figure 4. Gini of the $E(B - V)_{\text{stars}}$ distributions versus $\langle E(B - V)_{\text{stars}} \rangle$ for the 126 galaxies in the $z \sim 1.5$ sample (left-hand panel) and 184 galaxies in the $z \sim 2.3$ sample (right-hand panel). The yellow stars show the average G and $\langle E(B - V)_{\text{stars}} \rangle$ in four bins of $\langle E(B - V)_{\text{stars}} \rangle$ for the galaxies in each redshift sample, with the yellow bars indicating the standard error in the average G values. The dark and light green shaded regions highlight where the average $E(B - V)_{\text{stars}}$ are less than 0.05 and 0.10 mag, respectively. The Spearman rank correlation coefficient and its significance is listed in the top right-hand corner of each panel. We observe a wide range of Gini values at low average $E(B - V)_{\text{stars}}$, suggesting that the observed correlation is not purely driven by an artificially boosted Gini when $\langle E(B - V)_{\text{stars}} \rangle$ is nearly zero.

are caused by the maximum $E(B - V)_{\text{stars}}$ of 0.40 mag allowed in the SED models (see Section 2.5). If the galaxy is very red on average [high $\langle E(B - V)_{\text{stars}} \rangle$], then the ‘reddest’ regions will not appear as significant outliers in the resolved distribution. Referring to Fig. 3, the example reddening maps show that galaxies with low G (bottom panels) tend to be redder on average [higher $\langle E(B - V)_{\text{stars}} \rangle$] than those with higher G (top panels). These limitations of G on the resolved $E(B - V)_{\text{stars}}$ distribution highlight the importance of pairing G with a morphology metric that is not exclusively sensitive towards the highest or lowest values in a resolved distribution – such as the patchiness metric, P . Furthermore, while P and G are generally positively correlated (see Appendix A), the examples shown in the left set of sub-panels of Fig. 3 demonstrate how P and G can be used together to further quantify differences in the morphology of galaxies. For example, GOODS-S 33036 (top left-hand sub-panel) shows how a single outlier can drive G higher while the global distribution remains generally smooth with low P . Therefore, G can still be used to probe the concentration of dust reddening as long as the contribution of $\langle E(B - V)_{\text{stars}} \rangle$ is considered when interpreting trends with G .

4 DUST REDDENING DISTRIBUTION

Our analysis showing the morphology metrics calculated on the $E(B - V)_{\text{stars}}$ distribution compared to the total stellar mass and globally averaged gas-phase metallicity are discussed in Sections 4.1 and 4.2, respectively. We then place these results in the context of a physical interpretation for how the dust distribution evolves throughout the ISM of high-redshift galaxies in Section 4.3.

4.1 $E(B - V)$ versus stellar mass

Mass is a fundamental parameter of galaxies. Connections between the stellar mass of a galaxy and its other physical properties, such as their SFR (e.g. Daddi et al. 2007; Noeske et al. 2007; Pannella et al.

2009; Wuyts et al. 2011; Reddy et al. 2012; Whitaker et al. 2012, 2014; Shivaie et al. 2015), gas-phase oxygen abundance (Tremonti et al. 2004; Erb et al. 2006; Kewley & Ellison 2008; Mannucci et al. 2010; Andrews & Martini 2013; Sanders et al. 2015, 2018, 2021), and dust content (e.g. Reddy et al. 2006, 2010, 2015; Pannella et al. 2009; Yoshikawa et al. 2010; Price et al. 2014; Hemmati et al. 2015; Nelson et al. 2016; Tacchella et al. 2018; Shivaie et al. 2020), give clues to the processes that drive the evolution of galaxies. Understanding the distribution of dust is also critical for inferring the bolometric output from the underlying stellar populations. Therefore, in this section, we investigate how the dust distribution, as inferred by the P , G , and M_{20} morphology metrics measured on the resolved $E(B - V)_{\text{stars}}$ maps, varies as a function of the total stellar mass for galaxies at $z \sim 2$. Total stellar masses are derived from the SEDs that best fit the integrated 3D-*HST* photometry (see Section 2.5). Examples of the reddening maps for the lowest and highest mass galaxies in our $z \sim 1.5$ and 2.3 samples are shown in Fig. 5, with the log stellar mass of each galaxy listed in the bottom right-hand corner of each panel. The highest mass galaxies in both redshift samples tend to exhibit more complicated $E(B - V)_{\text{stars}}$ distributions, whereas the low-mass galaxies exhibit more uniform and generally bluer $E(B - V)_{\text{stars}}$ distributions (see further discussion in Section 4.3).

In addition to using the morphology metrics on the $E(B - V)_{\text{stars}}$ distribution, the global $E(B - V)_{\text{stars}}$ derived from the integrated 3D-*HST* photometry (see Section 2.5) and the global $E(B - V)_{\text{gas}}$ are also incorporated into this analysis. Balmer decrements ($H\alpha/H\beta$) measured for each galaxy by the MOSDEF survey (see Section 2.2) are used to calculate the globally averaged $E(B - V)_{\text{gas}}$, where a Cardelli et al. (1989) galactic extinction curve is assumed.⁸ Typical ISM conditions with $T = 10\,000$ K, $n_e = 10^2$ cm⁻³, and Case B recombination are assumed. Zero dust extinction is assumed for galaxies with $H\alpha/H\beta < 2.86$ (Osterbrock 1989). If $H\alpha$ or $H\beta$ is not detected with an $S/N \geq 3$, then individual measurements for $E(B - V)_{\text{gas}}$ are represented by their 3σ upper limit.

Fig. 6 shows all measures of $E(B - V)$ versus stellar mass for the $z \sim 1.5$ (top panels) and $z \sim 2.3$ (bottom panels) samples. Each top row of panels shows the unresolved globally averaged $E(B - V)_{\text{gas}}$ (panel a), $E(B - V)_{\text{stars}}$ (panel b), and difference between $E(B - V)_{\text{gas}}$ and $E(B - V)_{\text{stars}}$ (panel c) versus stellar mass. Each bottom row of panels shows the P (panel d), G (panel e), and M_{20} (panel f) morphology metrics calculated on the $E(B - V)_{\text{stars}}$ distribution versus stellar mass. Stacked measurements of $E(B - V)_{\text{gas}}$ in four bins of stellar mass containing an equal number of objects from the parent (large red diamonds) and final (small orange diamonds) samples are also shown (see Section 2.3). The Spearman rank correlation coefficient (ρ_s) is calculated from the individual measurements – not including those with upper or lower limits – and its significance is shown in the top left-hand corner of each panel. Throughout our analysis, we consider $>2\sigma$ significance as statistically significant given that there is a >95 percent confidence of rejecting the null hypothesis.

In agreement with previous studies, we find that $E(B - V)_{\text{gas}}$ (panel a), $E(B - V)_{\text{stars}}$ (panel b), and the difference between $E(B - V)_{\text{gas}}$ and $E(B - V)_{\text{stars}}$ (panel c) are significantly correlated with stellar mass (e.g. Yoshikawa et al. 2010; Price et al. 2014; Hemmati et al. 2015; Reddy et al. 2015; Nelson et al. 2016; Tacchella et al. 2018). The relationship between P and stellar mass (panel d) shows that

⁸Reddy et al. (2020) found that the shape of the nebular dust attenuation curve derived directly from the MOSDEF sample is similar in shape at rest-frame optical wavelengths to the Cardelli et al. (1989) Galactic extinction curve.

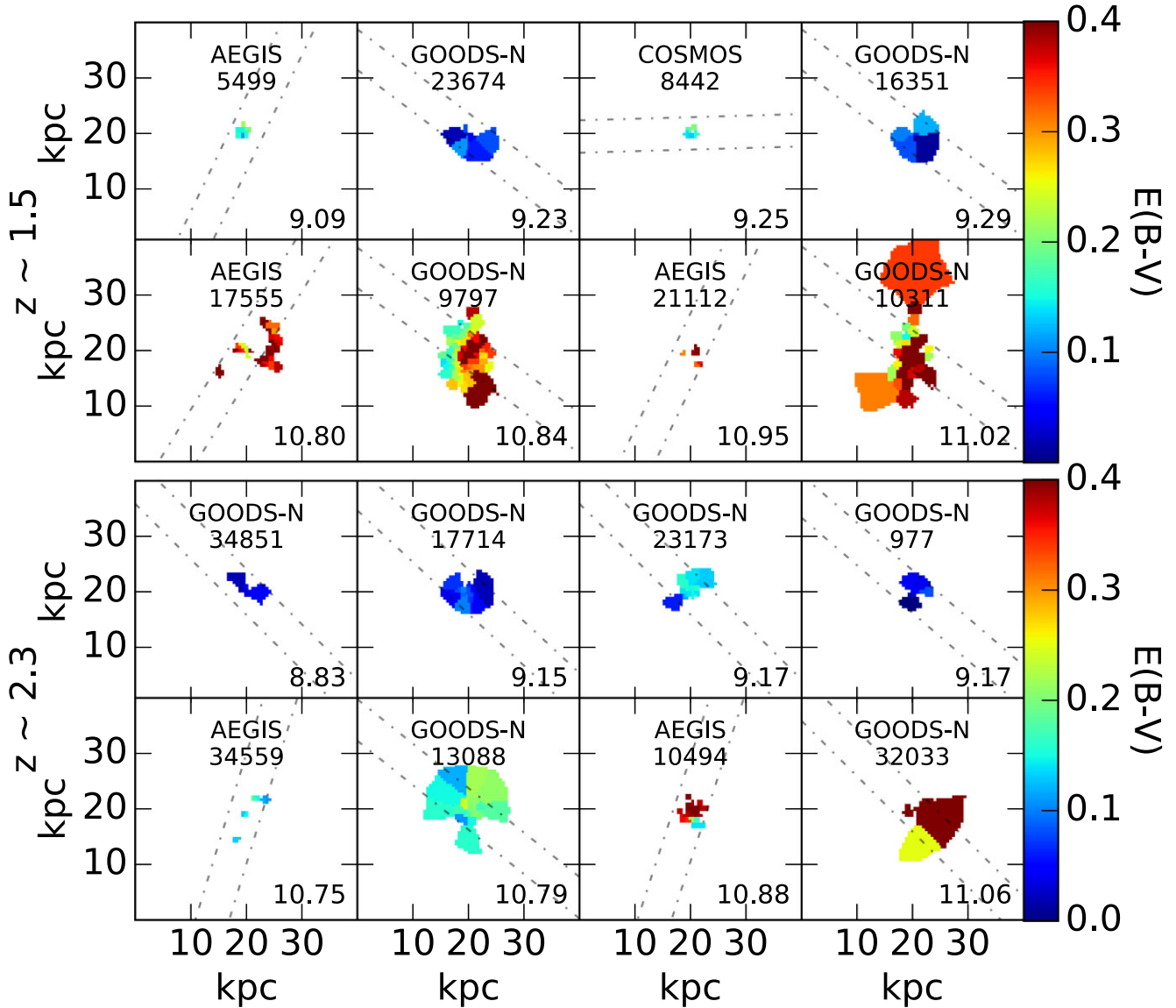


Figure 5. Examples of the $E(B - V)_{\text{stars}}$ maps for galaxies in the $z \sim 1.5$ (top panel) and $z \sim 2.3$ (bottom panel) samples. The top and bottom rows of each panel section show the galaxies with the lowest and highest masses, respectively, from each redshift sample. The $\log(M_*/M_\odot)$ of each galaxy is shown in the bottom right-hand corner of each panel. The dash-dotted grey lines show the placement of the MOSFIRE spectroscopic slit.

low-mass galaxies tend to have smoother $E(B - V)_{\text{stars}}$ distributions while high-mass galaxies exhibit generally patchier dust reddening distributions, but with a higher dispersion in P between galaxies of similar mass (see the left-hand panel of Fig. 7). The G of the $E(B - V)_{\text{stars}}$ distribution anticorrelates with stellar mass (panel e). An anticorrelation between G and stellar mass implies that the reddest regions in low-mass galaxies are significant outliers relative to $\langle E(B - V)_{\text{stars}} \rangle$ and the reddest regions in high-mass galaxies are comparable to $\langle E(B - V)_{\text{stars}} \rangle$ (see Section 3.3). Since high-mass galaxies are known to be typically dustier than low-mass galaxies, their reddest regions are not significant outliers compared to the average reddening and, thus, it is expected that they would exhibit lower Gini $E(B - V)_{\text{stars}}$ distributions (see Section 3.3). Finally, the high-mass galaxies in the $z \sim 1.5$ sample tend to have lower M_{20} values than low-mass $z \sim 1.5$ galaxies (top panel f). Low M_{20} values indicate that the reddest regions are generally concentrated within a

single region in high-mass $z \sim 1.5$ galaxies. There is no correlation observed between M_{20} and stellar mass for the $z \sim 2.3$ sample (bottom panel f).

We further investigate the significance of the relationship between P measured on the $E(B - V)_{\text{stars}}$ distribution versus stellar mass (Fig. 6d) by performing an independent t -test. The $z \sim 1.5$ and 2.3 samples are each split into a low- and high-mass bin, each with an equal number of galaxies where the threshold between low- and high-mass galaxies is $\sim 10^{10} M_\odot$. The independent t -test on the $z \sim 1.5$ sample shows that there is a significant difference in the P measured between low- and high-mass galaxies ($p < 0.05$), whereas the $z \sim 2.3$ sample shows a marginal difference in the P measured between low- and high-mass galaxies ($p < 0.1$). The significant difference shown by the independent t -test could be caused by the large difference in the dispersion of P between each redshift bin (see left-hand panel of Fig. 7), therefore we also perform bootstrap resampling

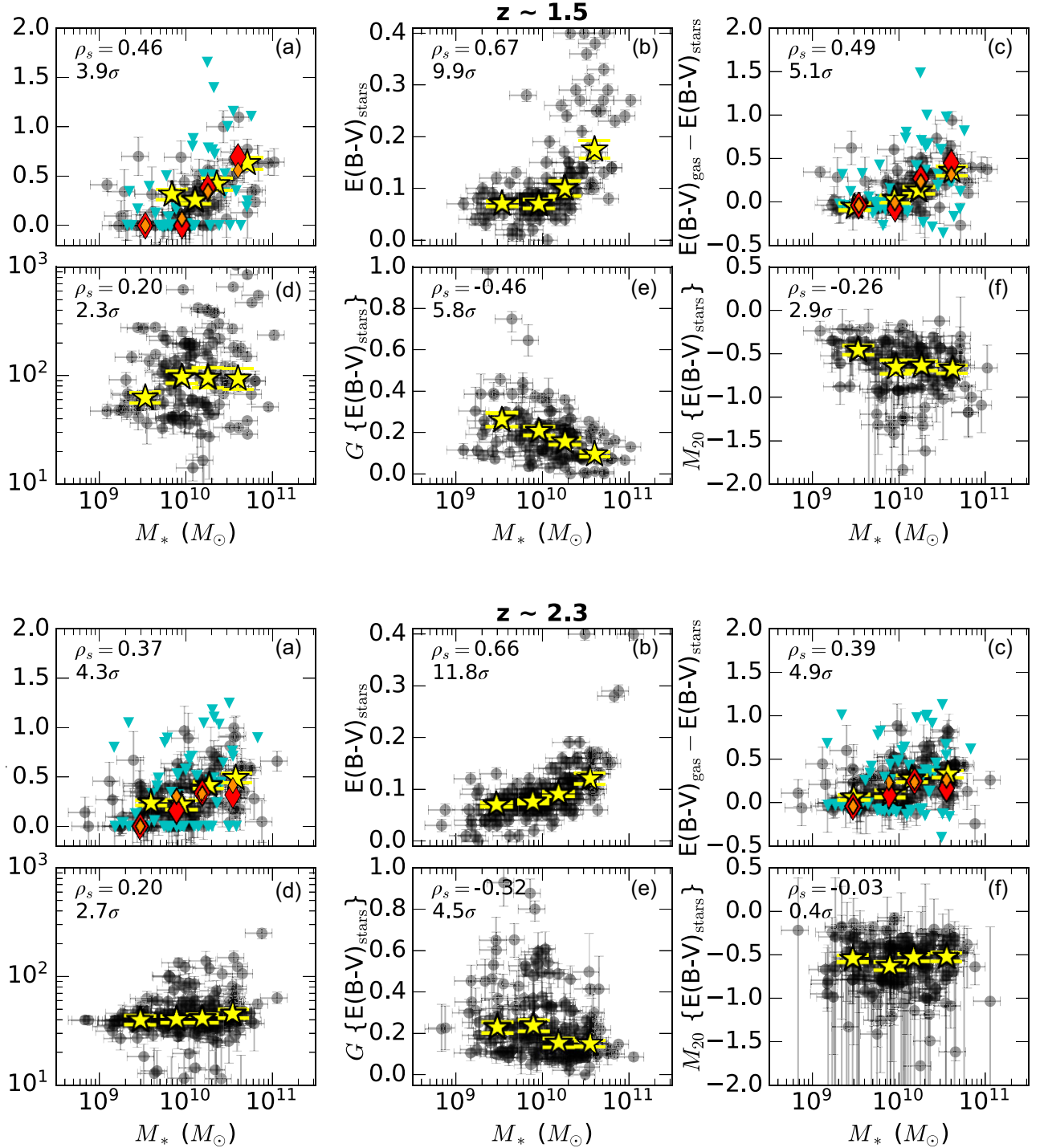


Figure 6. Relationships between stellar mass and several probes of dust reddening for the 126 galaxies in the $z \sim 1.5$ sample (top panel) and 184 galaxies in the $z \sim 2.3$ sample (bottom panel). Top row of sub-panels: Globally averaged $E(B-V)$ derived from the unresolved data. Balmer decrements ($H\alpha/H\beta$) measured from the MOSDEF survey are used to calculate $E(B-V)_{\text{gas}}$ assuming a Cardelli et al. (1989) galactic extinction curve. The light blue triangles show upper limits on $E(B-V)_{\text{gas}}$. Stacks of the spectra in four bins of stellar mass containing an equal number of objects are shown by the diamond symbols, where the stacks from the parent and final samples are shown by the large red and small orange diamonds, respectively. The globally averaged $E(B-V)_{\text{stars}}$ are inferred from the integrated 3D-*HST* photometry. The yellow stars show the median y-axis values and median M_* in four bins of stellar mass containing an equal number of objects, with the yellow bars indicating the standard error in the median y-axis values. The Spearman rank correlation coefficient and its significance is listed in the top left-hand corner of each panel. Bottom row of sub-panels: The P , G , and M_{20} morphology metrics calculated on the $E(B-V)_{\text{stars}}$ distribution derived from the resolved SED fitting.

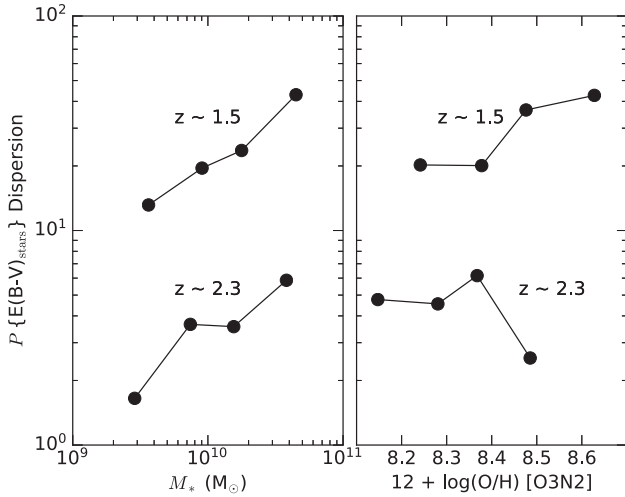


Figure 7. Dispersion in the P calculated on the $E(B - V)_{\text{stars}}$ distribution versus stellar mass (left-hand panel) and O3N2 metallicity (right-hand panel). The points indicate four bins in stellar mass or metallicity, each containing an equal number of objects. The systematic offset in P dispersion between the $z \sim 1.5$ and ~ 2.3 samples is either due to the $z \sim 1.5$ galaxies having overall more variation in their $E(B - V)_{\text{stars}}$ maps at similar masses (see Fig. 5) or their higher resolution (i.e. more Voronoi bins) compared to galaxies in the $z \sim 2.3$ sample (see Fig. A6).

over 1000 iterations to determine the 68 per cent confidence interval on the distribution of P values in each sub-sample. We find that the confidence intervals from the bootstrap resampling of the low- and high-mass bins do not overlap with each other for either the $z \sim 1.5$ or 2.3 galaxies. After testing higher confidence intervals, we find that they begin to overlap at the 90 and 87 per cent confidence levels for the $z \sim 1.5$ and 2.3 samples, respectively. Therefore, we consider the trend between P measured on the stellar continuum reddening distribution and stellar mass to be significant to at least 1.5σ confidence.

In extension to the discussion throughout Appendix A, we also further investigate how P of the $E(B - V)_{\text{stars}}$ distribution may be correlated with the number of Voronoi bins and the sizes of the galaxies in our sample. We find that the P of the $E(B - V)_{\text{stars}}$ distribution for the $z \sim 1.5$ sample is positively correlated with both the number of Voronoi bins and the size of the galaxies at 3.5σ and 2.4σ , respectively. On the other hand, the $z \sim 2.3$ sample shows that P on the $E(B - V)_{\text{stars}}$ distribution is inversely correlated with the number of Voronoi bins to 4.0σ significance, such that galaxies with fewer Voronoi bins tend to have patchier dust distributions, and P measured on the $E(B - V)_{\text{stars}}$ distribution is not significantly correlated with the size of the galaxy (1.6σ). When combining the $z \sim 1.5$ and 2.3 samples, P is not significantly correlated with the number of Voronoi bins (1.2σ) but is significantly correlated with the size of the galaxy (5.3σ). The significant correlation between P and the size of the galaxies is not surprising given that there is a very strong size- M_* relationship (see Fig. 1) and Fig. 6 d shows how P correlates with stellar mass in both redshift bins. While P correlates significantly with the number of Voronoi bins in the $z \sim 1.5$ sample, this trend is inverse in the $z \sim 2.3$ sample and does not persist in the combined sample. Therefore, we suggest that the patchiness metric is not biased by the number of Voronoi bins in a galaxy, and instead that the trend between P and the number of Voronoi bins is more directly related to the clumpy light distribution observed from the

imaging of the $z \sim 1.5$ galaxies (see also Appendix B). Based on our investigations on the significance in the difference between the P measured from the low- and high-mass bins and the insignificant correlation between P and the number of Voronoi bins, we overall suggest that the trend between P and M_* is driven by the physical conditions in the ISM.

It is expected that galaxies with centrally peaked radial dust profiles would exhibit lower M_{20} since the reddest regions are defined as being located within a single region. Several studies have observed $E(B - V)_{\text{stars}}$ and $E(B - V)_{\text{gas}}$ to be centrally peaked in $z \sim 2$ galaxies, with some evidence for steeper radial gradients as stellar mass increases (e.g. Nelson et al. 2013, 2016; Hemmati et al. 2015; Tacchella et al. 2018). These observations are in agreement with our findings of lower M_{20} in the $z \sim 1.5$ high-mass galaxies (Fig. 6f top panel). Tacchella et al. (2018) also found that several galaxies in their sample of 10 star-forming galaxies at $z \sim 2$ ($M_* \sim 10^{10} - 10^{11.4} M_\odot$) exhibited secondary local maxima in their radial attenuation profiles while the stacked profile showed a mostly smooth gradient. Secondary non-central peaks in the $E(B - V)_{\text{stars}}$ distribution would cause generally higher P and, if the secondary peaks are within the 20 per cent reddest regions, higher M_{20} values. We find examples of galaxies with both types of $E(B - V)_{\text{stars}}$ distributions with high P on the high-mass end of the $z \sim 2.3$ sample. For example, COSMOS 19753 (top right-hand sub-panel of Fig. 3) has high M_{20} (several red clumps) and AEGIS 10494 (left of the bottom right-hand panel of Fig. 5) has low M_{20} (single red clump). Therefore, a mix of galaxies with a single central peak and those with additional off-centre peaks in their radial dust attenuation profiles could explain both patchier dust distributions in high-mass galaxies and the flat correlation between stellar mass and M_{20} observed in the $z \sim 2.3$ sample (bottom panels d and f in Fig. 6f). Alternatively, it is possible that galaxies in the $z \sim 2.3$ sample are not sufficiently resolved (see Appendix A) to robustly probe the shape of the radial dust attenuation profile. While high Gini could indicate steeper radial gradients, the reddest regions probed by the Gini coefficient do not necessarily need to be grouped within a single region (e.g. the centre). Furthermore, when Gini is measured on the dust reddening, high Gini values can be attributed to $\langle E(B - V)_{\text{stars}} \rangle$ being nearly zero (see Fig. 4), which could drive the significance of the relationship between Gini and stellar mass.

While patchiness is a newly derived morphology metric, we show in Appendix A that P calculated on the $E(B - V)_{\text{stars}}$ distribution is significantly correlated with ICD (Papovich et al. 2003). Boada et al. (2015) found that higher mass galaxies (up to $M_* = 10^{11} M_\odot$) at $z \sim 2$ tend to have higher ICD, which is comparable with our findings that the $E(B - V)_{\text{stars}}$ distribution tends to be patchier in higher mass galaxies (Fig. 6d). Several studies have suggested that patchy dust distributions could explain observed differences between nebular and stellar continuum reddening (e.g. Calzetti et al. 1994; Wild et al. 2011; Price et al. 2014; Reddy et al. 2015, 2018a, 2020) and SFR indicators (e.g. Boquien et al. 2009, 2015; Hao et al. 2011; Reddy et al. 2015; Katsianis et al. 2017; Fetherolf et al. 2021). Our observation of high-mass galaxies exhibiting patchier stellar continuum dust reddening distributions (Fig. 6d) and larger differences between their globally averaged nebular and stellar continuum reddening (Fig. 6c) further suggests a connection between patchy dust distributions and the differences between the stellar continuum and nebular reddening. However, we find no correlation between the P of the $E(B - V)_{\text{stars}}$ distribution and the difference between $E(B - V)_{\text{gas}}$ and $E(B - V)_{\text{stars}}$, which suggests that other processes may also be contributing to the observed differences between stellar continuum and nebular reddening.

4.2 $E(B - V)$ versus metallicity

There exists a well-known relationship between stellar mass and gas-phase oxygen abundance, hereafter referred to as the ‘metallicity’ (e.g. Tremonti et al. 2004). There is also a known correlation between dust and metals (e.g. Dwek 1998; Jenkins 2009; Reddy et al. 2010; Mattsson et al. 2014; Rémy-Ruyer et al. 2014; Shivaie et al. 2017; De Vis et al. 2019; Shapley et al. 2020; Galliano et al. 2021). Therefore, it is reasonable to expect that the morphology metrics that probe the $E(B - V)_{\text{stars}}$ distribution may correlate with metallicity. In this section, we show how P , G , and M_{20} calculated on the $E(B - V)_{\text{stars}}$ distribution vary as a function of metallicity.

Gas-phase metallicities are based on the $[\text{N II}]\lambda 6585$, $\text{H}\alpha$, $[\text{O III}]\lambda 5008$, and $\text{H}\beta$ emission line measurements from the MOSDEF survey (see Section 2.2). The empirical calibrations from Pettini & Pagel (2004) are assumed to obtain gas-phase oxygen abundances $[12 + \log(\text{O}/\text{H})]$ based on the $\log \{([\text{O III}]\lambda 5008/\text{H}\beta)/([\text{N II}]\lambda 6585/\text{H}\alpha)\}$ (O3N2) indicator.⁹ Since these ratios include lines that are close in wavelength, the lines are not corrected for dust obscuration. If $[\text{N II}]\lambda 6585$ and/or $\text{H}\beta$ is not detected with an $\text{S/N} \geq 3$, then an upper limit on the metallicity is assumed. If $[\text{O III}]\lambda 5008$ and/or $\text{H}\alpha$ is not detected with an $\text{S/N} \geq 3$, then a lower limit on the metallicity is assumed. Any other combination of lines that are not detected with an $\text{S/N} \geq 3$ is not useful for obtaining gas-phase metallicities. Based on these restrictions, we obtained gas-phase metallicities for 262 galaxies, with 128 of those galaxies having either upper or lower limits on their metallicities. Metallicities with upper or lower limits on their measurements are not included when determining the Spearman rank correlation coefficient and its significance and the results do not change significantly when all measurements are included.

Fig. 8 shows all measures of $E(B - V)$ (similar to Fig. 6) versus O3N2 metallicity, split by the $z \sim 1.5$ (top panels) and $z \sim 2.3$ (bottom panels) samples. The orange diamonds in panels (d)–(f) show the metallicities measured from the spectra that were stacked in two bins of each morphology metric (P , G , and M_{20}). Most significantly, we find that higher metallicity galaxies in the $z \sim 1.5$ sample tend to be overall dustier with higher $E(B - V)_{\text{gas}}$ and $E(B - V)_{\text{stars}}$ (top panels a and b). The correlation between metallicity and globally averaged dust reddening is consistent with the known connection between dust and metals (e.g. Dwek 1998; Jenkins 2009; Reddy et al. 2010; Mattsson et al. 2014; Rémy-Ruyer et al. 2014; Shivaie et al. 2017; De Vis et al. 2019; Shapley et al. 2020; Galliano et al. 2021). The correlation between metallicity and M_{20} (top panel f) suggests that high-metallicity galaxies at $z \sim 1.5$ exhibit centrally peaked radial dust profiles (see Section 4.3), similar to galaxies with higher masses (Nelson et al. 2016). There is no significant correlation between metallicity and P (panel d), but the dispersion in P is higher in $z \sim 1.5$ high-metallicity galaxies (right-hand panel of Fig. 7). This suggests that low-metallicity galaxies have comparable $E(B - V)_{\text{stars}}$ distributions while there is more variety in the $E(B - V)_{\text{stars}}$ distributions of high-metallicity galaxies, in that the distribution can be either smooth or patchy.

For the $z \sim 2.3$ sample, we find that high-metallicity galaxies tend to have higher $E(B - V)_{\text{stars}}$ relative to low-metallicity galaxies (bottom panel b). On the other hand, unlike the $z \sim 1.5$ sample, we find no significant trend between $E(B - V)_{\text{gas}}$ and metallicity for the $z \sim 2.3$ sample (bottom panel a). The lack of correlation between

metallicity and $E(B - V)_{\text{gas}}$ in the $z \sim 2.3$ sample, such that $E(B - V)_{\text{gas}}$ is typically the average of the sample $[\langle E(B - V)_{\text{gas}} \rangle = 0.31 \text{ mag}]$ at any metallicity, may be related to why the difference between $E(B - V)_{\text{stars}}$ and $E(B - V)_{\text{gas}}$ is uncorrelated with metallicity (bottom panel c). Shivaie et al. (2020) suggested that a larger difference between $E(B - V)_{\text{stars}}$ and $E(B - V)_{\text{gas}}$ observed in low-metallicity galaxies may be attributed to patchier dust distributions. However, we do not observe a correlation between P measured on the $E(B - V)_{\text{stars}}$ distribution and metallicity for galaxies in either the $z \sim 1.5$ or 2.3 samples (panel d). It is possible that the metal-rich gas may follow a similar distribution to the dust that reddens the stellar continuum [as probed by $E(B - V)_{\text{stars}}$] across the ISM, but the global enrichment of the galaxy is not directly connected to the dust reddening towards the most massive stars that have recently formed [as probed by $E(B - V)_{\text{gas}}$]. Due to the significant relationship between $E(B - V)_{\text{stars}}$ and metallicity, lower metallicity $z \sim 2.3$ galaxies also tend to show higher G on the $E(B - V)_{\text{stars}}$ distribution (bottom panel e; also see Fig. 4). Finally, there is no significant correlation between the concentration of the reddest regions, as probed by M_{20} , and the global metallicity for the $z \sim 2.3$ sample, which is similar to what is found between M_{20} and stellar mass (Fig. 6f).

4.3 Physical interpretation of the observed trends

Here, we present a possible physical interpretation for how the dust distribution evolves in $z \sim 2$ galaxies based on the relationships observed among nebular reddening, stellar continuum reddening, stellar mass, and metallicity. Specifically, our results suggest that the dust distribution in high-redshift galaxies tends to transition from smooth to patchy with increasing stellar mass. While there exists a well-known relationship between the stellar mass and metallicity of galaxies (e.g. Tremonti et al. 2004; Erb et al. 2006; Kewley & Ellison 2008; Mannucci et al. 2010; Andrews & Martini 2013; Sanders et al. 2015, 2018, 2021), we showed in Section 4.2 that the distribution of dust as probed by P , G , and M_{20} in the resolved $E(B - V)_{\text{stars}}$ maps is not significantly connected with the global metal enrichment of the ISM compared to the total stellar mass. We observed a significant correlation between metallicity and the average $E(B - V)_{\text{stars}}$ (panel b in Fig. 8), but there are marginal to null relationships between metallicity and P , G , and M_{20} (panels d–f in Fig. 8). These findings are consistent with previous studies that have found radial metallicity profiles in $z \sim 2$ galaxies that are generally flat for a wide range of stellar masses ($M_* \sim 10^7\text{--}10^{11} M_\odot$; Jones et al. 2015; Leethochawalit et al. 2016; Wuyts et al. 2016; Wang et al. 2019; Simons et al. 2021). Radial dust profiles, on the other hand, tend to be centrally peaked in massive $z \sim 2$ galaxies ($M_* \sim 10^9\text{--}10^{11} M_\odot$; Nelson et al. 2013, 2016; Hemmati et al. 2015; Tacchella et al. 2018). Therefore, while higher metallicity galaxies are dustier on average, the actual spatial distribution of dust is not necessarily correlated with the global metallicity of the galaxy. Instead, we propose that the mechanisms responsible for distributing dust throughout the ISM of high-redshift galaxies are more fundamentally connected to the stellar mass of these galaxies. In the following discussion, we focus on describing how the distribution of dust tends to change with stellar mass by referencing Fig. 9. The corresponding interpretation with metallicity is mentioned when relevant. The characteristics of the dust distribution in low- and high-mass galaxies are described separately, then we summarize how galaxies may transition from a smooth to patchy dust distribution as they increase in stellar mass. In the subsequent discussion, we use a threshold of $10^{10} M_\odot$ to define low- and high-mass galaxies.

⁹Using the N2 indicator produces results that are statistically consistent with those using the O3N2 indicator.

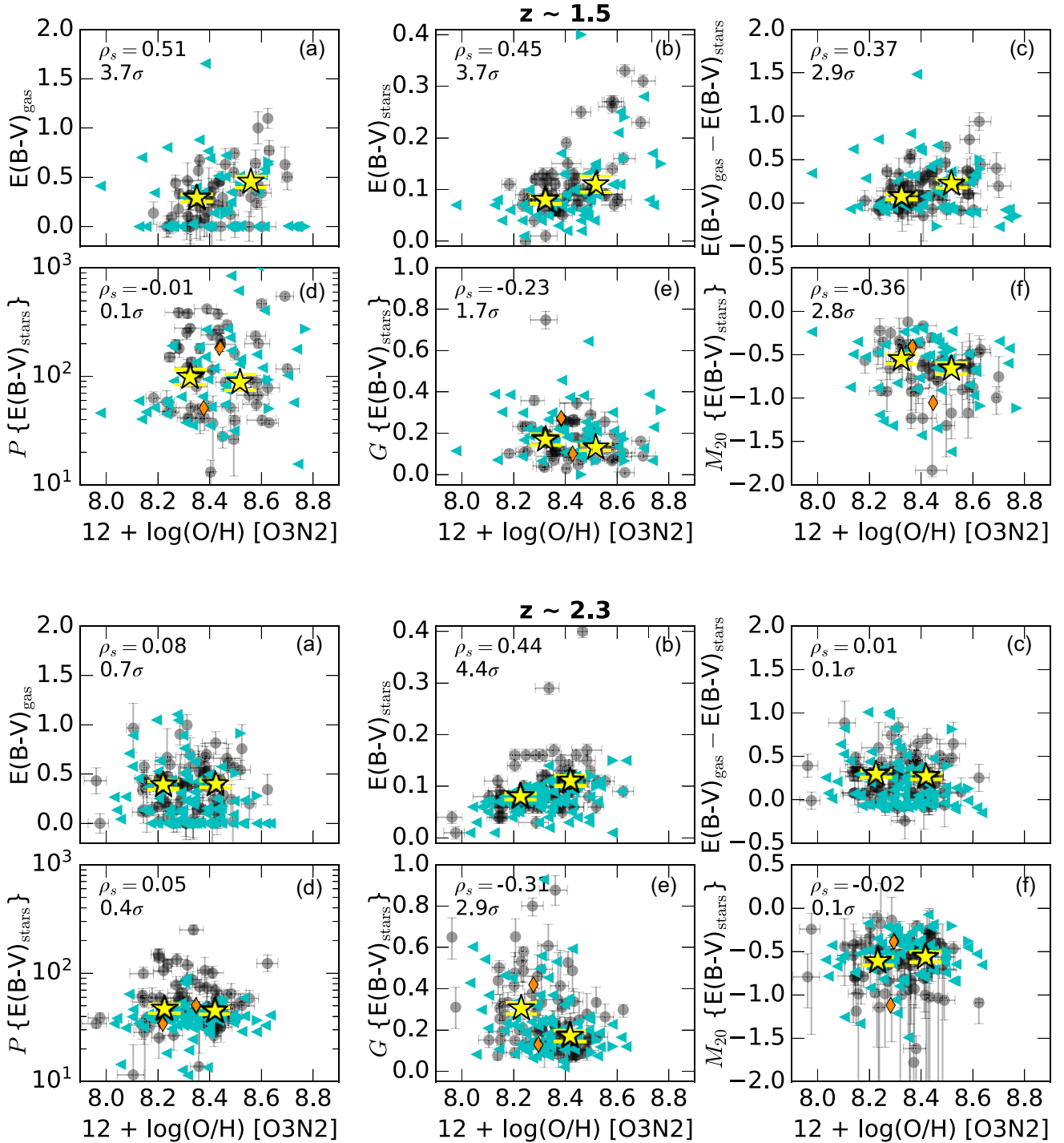


Figure 8. Same as Fig. 6, except showing the relationship between metallicity and various probes of dust reddening for a subset of 104 galaxies from the $z \sim 1.5$ sample (top panel) and 158 galaxies from the $z \sim 2.3$ sample (bottom panel). Gas-phase oxygen abundances are obtained by assuming the Pettini & Pagel (2004) calibrations for the O3N2 indicator. The left- and right-facing light blue triangles show lower and upper limits on the metallicities, respectively. Alternatively using the N2 metallicity indicator produces results that are statistically consistent with those from the O3N2 indicator. Stacks of the spectra in bins of P , G , or M_{20} containing an equal number of objects are shown by the small orange diamond symbols in panels (d)–(f).

A schematic example for low-mass galaxies is shown, on the left-hand side of Fig. 9. Low-mass galaxies are observed to exhibit lower globally averaged $E(B - V)_{\text{gas}}$ and $E(B - V)_{\text{stars}}$ that are approximately equal to each other (panels a–c in Fig. 6), indicating

that low-mass galaxies (and similarly low-metallicity galaxies) have low dust content. Low-mass galaxies tend to have lower measures of P on their $E(B - V)_{\text{stars}}$ distributions (panel d in Fig. 6), which implies that the distribution of dust is typically smooth in these galaxies (as

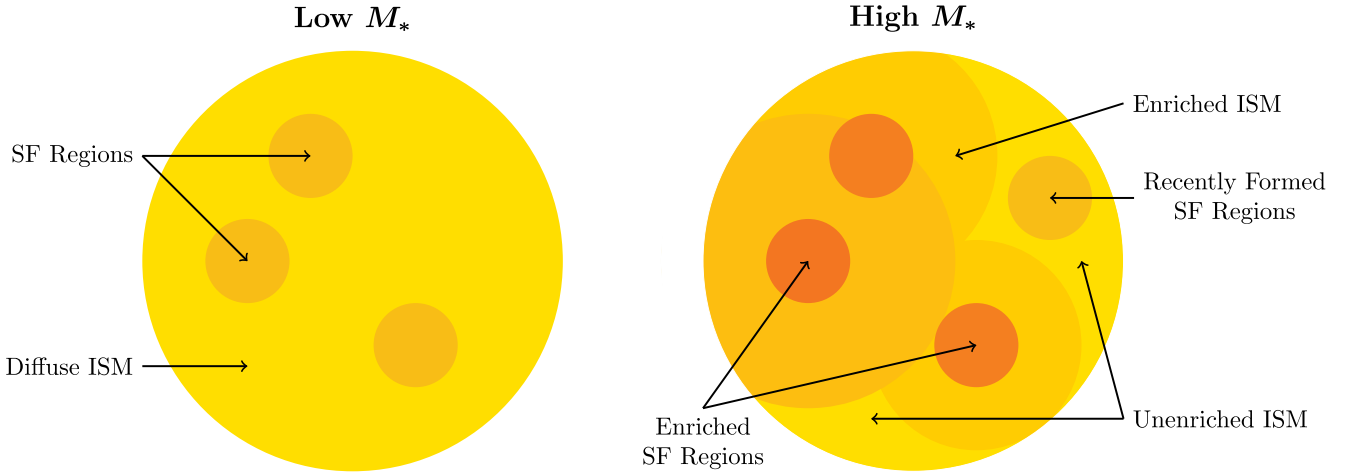


Figure 9. Illustration demonstrating dust production within star-forming regions that spreads radially outwards from each star-forming region and enriches the ISM as galaxies increase in stellar mass. Yellow regions represent the unenriched ISM, small orange circles designate star-forming regions that become redder as dust is produced through stellar evolutionary processes (AGB stars, supernovae, etc.), and large orange circles show the enriched ISM caused by dust radially spreading from star-forming regions to the nearby ISM. The radially expanding dust (large orange circles) may overlap such that the dust would appear concentrated within a single continuous clump. The dustiest region is illustrated as being off-centre to emphasize that the observer-defined centre of a given galaxy is not necessarily the geometric centre. Dust mixing time-scales may be faster in low-mass galaxies due to their physically smaller sizes compared to high-mass galaxies, but note that the size- M_* relation is not explicitly shown in this illustration.

depicted by the large yellow circle in Fig. 9; also see Fig. 5).¹⁰ Furthermore, both low-mass and low-metallicity galaxies exhibit similar degrees of patchiness in their $E(B - V)_{\text{stars}}$ distributions (i.e., low dispersion; see Fig. 7). The higher G in low-mass galaxies is likely boosted by there being very little dust overall [$\langle E(B - V)_{\text{stars}} \rangle \approx 0$; see Fig. 4], but localized regions with higher $E(B - V)_{\text{stars}}$ compared to the global average may exist (Fig. 6e). The higher M_{20} values (Fig. 6f) indicate that the highest $E(B - V)_{\text{stars}}$ regions are not concentrated within a single area in low-mass galaxies (as depicted by the small orange circles in Fig. 9). The localized regions with higher $E(B - V)_{\text{stars}}$ likely correspond to sites of recent star formation,¹¹ but the distribution of star formation as probed by the SED-derived SFR maps is still observed to be generally smooth in low-mass galaxies (see Fig. 10). These results are consistent with a picture in which dust distributes throughout low-mass galaxies on a short time-scale (< 100 Myr), such that there is a generally negligible difference between the observed $E(B - V)_{\text{stars}}$ and $E(B - V)_{\text{gas}}$ measured globally across these galaxies (Fig. 6c).

The right side of Fig. 9 shows a schematic for high-mass galaxies. While $E(B - V)_{\text{gas}}$ is observed to be higher than $E(B - V)_{\text{stars}}$ in high-mass galaxies, both $E(B - V)_{\text{gas}}$ and $E(B - V)_{\text{stars}}$ are high when compared to values typical of low-mass galaxies (panels a–c in Fig. 6;

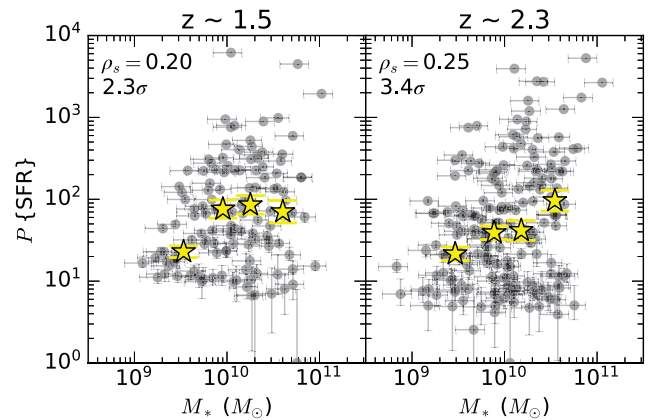


Figure 10. Patchiness of the SED-derived SFR distributions versus stellar mass for the 126 galaxies in the $z \sim 1.5$ sample (left-hand panel) and 184 galaxies in the $z \sim 2.3$ sample (right-hand panel). The yellow stars show the average P and stellar mass in four bins of stellar mass for the galaxies in each redshift sample, with the yellow bars indicating the standard error in the average P values. The Spearman rank correlation coefficient and its significance is listed in the top left-hand corner of each panel.

e.g. Yoshikawa et al. 2010; Price et al. 2014; Hemmati et al. 2015; Reddy et al. 2015; Nelson et al. 2016; Tacchella et al. 2018; Shivaei et al. 2020). These trends suggest that while high-mass galaxies (and similarly high-metallicity galaxies) are generally dustier than low-mass galaxies (medium-sized orange circles), the regions that formed stars most recently are dustier than that which is typical of the ISM across the galaxy (small dark orange circles; Fig. 6c). There is a wide dispersion in the patchiness of high-mass galaxies and high-metallicity $z \sim 1.5$ galaxies (Fig. 7), but, on average, the $E(B - V)_{\text{stars}}$ distributions of high-mass galaxies tend to be patchier than those of low-mass galaxies and they have lower G (panels d–e in Fig. 6). Since $\langle E(B - V)_{\text{stars}} \rangle$ is redder in high-mass and high-metallicity galaxies, the reddest regions are not sufficiently significant outliers to drive

¹⁰Low P is not caused by having fewer Voronoi bins (see Appendix A) or low S/N components. Low S/N Voronoi bins are not included in the analysis (see Section 2.4).

¹¹Confirmation that the regions with higher $E(B - V)_{\text{stars}}$ (small orange circles in Fig. 9) correspond to sites of recent star formation would require resolved nebular emission line maps [i.e. resolved $E(B - V)_{\text{gas}}$ maps], which are not available for individual galaxies in our sample. The 3D-*HST* resolved emission line maps have been used to study the distribution of recent star formation and dust in stacks of several galaxies (e.g. Wuyts et al. 2013; Nelson et al. 2013, 2016), but the S/N of the emission-line map elements are not typically sufficient for performing a comparative analysis on the distribution of the stellar continuum and nebular emission in individual galaxies at high redshifts.

G to higher values (see Fig. 4). G is also not sensitive to the bluest regions in the $E(B - V)_{\text{stars}}$ distribution, by definition (equation 3). Therefore, a high P and low G in high-mass galaxies indicate that there are several low $E(B - V)_{\text{stars}}$ outliers that are driving P to higher values (yellow regions near edges). Lastly, the dustiest regions are generally concentrated within a single clump (overlapping orange regions) given the lower M_{20} values in the $E(B - V)_{\text{stars}}$ distributions of high-mass and high-metallicity galaxies in the $z \sim 1.5$ sample. The centres of the reddest regions, defined by (x_c, y_c) from Section 3.2, are typically within 1 kpc of the centre of the sub-images (i.e. observed centre; see Section 2.4) and are at most 4.5 kpc from the centre. It then follows that the interpretation of low M_{20} , which indicates a single dusty region in the ISM of $z \sim 1.5$ high-mass and high-metallicity galaxies, is consistent with observations of centrally peaked radial dust profiles in these galaxies (e.g. Nelson et al. 2013, 2016; Hemmati et al. 2015; Tacchella et al. 2018).

The lack of centrally concentrated dust reddening (probed by M_{20}) in $z \sim 2.3$ high-mass galaxies could be explained by spatially independent young star-forming regions or there being insufficient time for the dust to diffuse throughout the ISM in these physically large galaxies. The possibility of spatially independent young star-forming regions in high-mass galaxies is further supported by a correlation between P of the SED-derived resolved SFR and stellar mass (Fig. 10), which is more significant for galaxies in the $z \sim 2.3$ sample (3.4σ) than the $z \sim 1.5$ sample (2.3σ). Alternatively, the total amount of dust increases throughout the diffuse ISM in high-mass galaxies, but mixes throughout the ISM on longer time-scales than that of low-mass galaxies due to the physically larger sizes of high-mass galaxies (see right-hand panel of Figs 1 and 6d). Otherwise, the $z \sim 2.3$ galaxies may not be sufficiently resolved to robustly probe the dynamic range of P and M_{20} since the $z \sim 2.3$ galaxies have fewer resolved components (i.e. Voronoi bins) than the $z \sim 1.5$ galaxies in this study (see Appendix A).

The distribution of dust has traditionally been approximated by either a uniform dust screen or a two-component dust model (Calzetti et al. 1994; Charlot & Fall 2000) with higher optical depths towards the youngest star-forming regions, which causes $E(B - V)_{\text{gas}}$ to be systematically larger than $E(B - V)_{\text{stars}}$ (Calzetti et al. 2000). Patchy dust distributions have also been invoked in order to explain both the observed differences between stellar continuum and nebular reddening, and UV and $H\alpha$ SFR indicators (e.g. Wild et al. 2011; Price et al. 2014; Boquien et al. 2015; Reddy et al. 2015). Our results suggest that low-mass galaxies have a generally diffuse and uniform distribution of dust throughout their ISM, with only minor deviations towards higher opacities in localized regions. Recent star formation activity may be distributed uniformly throughout low-mass galaxies, thus resulting in a smooth distribution of dust through stellar evolutionary processes. Furthermore, several simulation studies have suggested that dust mixes nearly instantaneously ($\sim$ few tens of Myr) between the cold and warm ISM (e.g. McKee 1989; Tielens 1998; Peters et al. 2017). Therefore, efficient dust mixing in spatially small (i.e. low mass) galaxies could explain their relatively smooth distributions as indicated in the dust reddening maps. As galaxies increase in stellar mass through *in situ* star formation, they will produce significantly more dust and have physically larger sizes than their low-mass counterparts. We suggest that regions closest to sites of dust production (star-forming regions) within high-mass galaxies will become enriched first, while the remainder of the ISM becomes enriched on longer time-scales. Meanwhile, a centrally peaked concentration of dust can form in high-mass galaxies ($z \sim 1.5$ sample) if star-forming regions are closely spaced or if star formation activity is generally higher in the centres of high-mass galaxies.

The observed variations in the dust distribution have implications for the selection of an appropriate dust attenuation curve, especially considering that the shape of the dust attenuation curve has been observed to vary with galaxy properties (e.g. Reddy et al. 2006, 2010, 2018b; Leja et al. 2017; Salim et al. 2018; Shivaee et al. 2020). Therefore, assuming a single attenuation curve shape for all high-redshift galaxies may not be appropriate. Since the highest mass galaxies tend to exhibit patchier dust distributions, their starlight may more appropriately be corrected for dust obscuration by assuming a greyer dust attenuation curve (e.g. Calzetti et al. 2000). On the other hand, a steeper SMC-like dust attenuation curve (Fitzpatrick & Massa 1990; Gordon et al. 2003) may be more appropriate for low-mass galaxies with smoother dust distributions. Investigating the distribution of dust in galaxies using a sample that is split between different attenuation curve assumptions would require a further in-depth investigation with a larger sample size given that comparative morphology metrics – such as P , G , and M_{20} – are best applied to uniformly defined samples (see Section 3 and Appendix A).

To summarize, as depicted by Fig. 9, we suggest that the evolution of the dust distribution in the ISM is correlated with stellar mass and possibly the physical size of the galaxy. Low-mass galaxies exhibit smooth dust distributions, possibly due to uniformly distributed star formation activity and short dust mixing time-scales in these generally compact galaxies. Meanwhile, high-mass galaxies tend to show more complex dust distributions on average (but with higher dispersion) compared to low-mass galaxies, perhaps due to their physically larger sizes and longer dust mixing time-scales. Galaxy simulations that include the effects of dust propagation through the ISM should help to quantify the dust mixing time-scales for galaxies of different masses. Finally, we defer a discussion about how the ISM dust distribution evolves across cosmic time to a future work with a larger sample across redshift bins given that the trends observed in Section 4.1 and Section 4.2 are generally statistically consistent between our $z \sim 1.5$ and 2.3 redshift samples.

5 SUMMARY

We investigated the inferred distribution of dust for a sample of 310 star-forming galaxies from the MOSDEF survey at spectroscopic redshifts $1.36 < z < 2.66$. Using a new morphology metric called patchiness (P), the Gini coefficient (G), and second-order moment of light (M_{20}), we quantified robust dust reddening maps that were constructed from CANDELS/3D-*HST* high-resolution imaging. Globally averaged $E(B - V)_{\text{gas}}$, $E(B - V)_{\text{stars}}$, and gas-phase metallicity (O3N2) were also used to help build a physical interpretation for the evolution of the distribution of dust throughout the ISM of high-redshift galaxies. We found that the total amount of dust is correlated both with stellar mass and metallicity, but the distribution of dust (as probed by P , G , and M_{20}) is more significantly connected to the stellar mass (Fig. 6) of a galaxy than its globally averaged gas-phase metallicity (Fig. 8).

The patchiness metric (P ; equation 2) is sensitive to both the high and low outliers of a distribution. Low-mass galaxies tend to exhibit low P in their $E(B - V)_{\text{stars}}$ distributions, indicating that the dust is uniformly distributed throughout their ISM. Meanwhile, high-mass galaxies are more likely to exhibit patchy $E(B - V)_{\text{stars}}$ distributions (with a large dispersion in P). High-mass galaxies have both extremely dusty star-forming regions and regions that are sufficiently far from sites of young star formation that are not enriched to the same extent, possibly due to the physically larger size of high-mass galaxies. The progression of higher to lower M_{20} values (equation 5) in the dust distribution of the $z \sim 1.5$ sample suggests

that the dustiest regions in low-mass galaxies exist in several isolated clumps, which then become more centrally concentrated as galaxies increase in stellar mass and metallicity. The resolved $E(B - V)_{\text{stars}}$ distribution also exhibits high G in low-mass and low-metallicity galaxies, which is attributed to sites of star formation that enhance the dust reddening in a few isolated regions relative to the generally low $\langle E(B - V)_{\text{stars}} \rangle$ of the ISM. However, G is systematically higher at $\langle E(B - V)_{\text{stars}} \rangle$ by definition (equation 3). Therefore, these localized regions of higher dust opacities are, in fact, insignificant compared to the global $E(B - V)_{\text{stars}}$ distribution such that the dust distribution remains smooth overall. Since the Gini coefficient is only sensitive to values that are above the average of a distribution, G on the $E(B - V)_{\text{stars}}$ distribution becomes less sensitive as galaxies become dustier at higher stellar masses and metallicities (Section 3.3 and Fig. 4).

Overall, we propose that the dust formed through stellar evolutionary processes is mixed efficiently throughout the ISM of low-mass galaxies due to either uniformly distributed star formation activity or their compact sizes. High-mass galaxies, on the other hand, have longer dust mixing time-scales possibly due to their physically larger sizes. High-mass galaxies also produce significantly more dust than low-mass galaxies such that only the regions closest to the sites of dust production (star-forming regions) become enriched on short time-scales. An illustration of the transition between a smooth to a more complex dust distribution as galaxies increase in stellar mass is shown in Fig. 9.

As sample sizes increase with the advent of statistically large resolved imaging surveys, galaxy morphology measurements must move away from visual classification and more towards quantitative metrics. We showed that using patchiness, Gini, and M_{20} together was critical towards painting a complete picture of the resolved structures within galaxies. In particular, higher patchiness values were driven by the bluest regions in the dust reddening distribution of high-mass galaxies, thus emphasizing that void-like or darker regions within galaxies could reveal valuable information about galaxy structure. Resolved observations of high-redshift galaxies have also revealed that galaxy structure is increasingly irregular, which brings into question how the centre of these galaxies should be defined when quantifying structure radially. Increasing the sample of high-redshift galaxies with resolved imaging and spectroscopy will further build upon the results of this paper, such as those that will be obtained using NIRC2 and NIRS2 on the *James Webb Space Telescope* for galaxies out to $z \sim 6$. These observations will further enable the construction of high-resolution stellar population, reddening, and emission line maps for high-redshift galaxies – such as spatially resolved Balmer decrements – and will reveal how galaxies assemble and evolve across cosmic time.

DATA AVAILABILITY

Resolved CANDELS/3D-*HST* photometry is available at <https://3d.hst.research.yale.edu/Data.php>. Spectroscopic redshifts, 1D spectra, and 2D spectra from the MOSDEF survey are available at <http://mosdef.astro.berkeley.edu/for-scientists/data-releases/>.

ACKNOWLEDGEMENTS

This work is based on observations taken by the CANDELS Multi-Cycle Treasury Program and the 3D-*HST* Treasury Program (GO 12177 and 12328) with the NASA/ESA *HST*, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. The MOSDEF team acknowledges support from an NSF AAG collaborative grant (AST-1312780,

1312547, 1312764, and 1313171) and grant AR-13907 from the Space Telescope Science Institute. Some of the data presented herein were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California, and the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation. The authors wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Maunakea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain.

Facilities: *HST* (WFC3, ACS), Keck:I (MOSFIRE), and *Spitzer* (IRAC)

Software: ASTROPY (Astropy Collaboration et al. 2013, 2018), MATPLOTLIB (Hunter 2007), NUMPY (Oliphant 2007), SCIPY (Oliphant 2007), SPECLINE (Shivaei et al. 2018), Voronoi Binning Method (Cappellari & Copin 2003)

REFERENCES

- Abraham R. G., van den Bergh S., Nair P., 2003, *ApJ*, 588, 218
- Andrews B. H., Martini P., 2013, *ApJ*, 765, 140
- Astropy Collaboration et al. 2013, *A&A*, 558, A33
- Astropy Collaboration et al. 2018, *AJ*, 156, 123
- Azadi M. et al. 2017, *ApJ*, 835, 27
- Azadi M. et al. 2018, *ApJ*, 866, 63
- Banerji M. et al. 2010, *MNRAS*, 406, 342
- Barden M., Jahnke K., Häußler B., 2008, *ApJS*, 175, 105
- Bertin E., Arnouts S., 1996, *A&AS*, 117, 393
- Boada S. et al. 2015, *ApJ*, 803, 104
- Boquien M. et al. 2009, *ApJ*, 706, 553
- Boquien M. et al. 2015, *A&A*, 578, A8
- Brammer G. B. et al. 2012, *ApJS*, 200, 13
- Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000
- Calzetti D., Kinney A. L., Storchi-Bergmann T., 1994, *ApJ*, 429, 582
- Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, *ApJ*, 533, 682
- Cappellari M., Copin Y., 2003, *MNRAS*, 342, 345
- Cardelli J. A., Clayton G. C., Mathis J. S., 1989, *ApJ*, 345, 245
- Chabrier G., 2003, *Publ. Astron. Soc. Pac.*, 115, 763
- Charlot S., Fall S. M., 2000, *ApJ*, 539, 718
- Cibinel A. et al. 2015, *ApJ*, 805, 181
- Coil A. L. et al. 2015, *ApJ*, 801, 35
- Conselice C. J., 2003, *ApJS*, 147, 1
- Conselice C. J., 2014, *ARA&A*, 52, 291
- Daddi E. et al. 2007, *ApJ*, 670, 156
- de Vaucouleurs G., 1959, *Handbuch der Physik*, 53, 275
- De Vis P. et al. 2019, *A&A*, 623, A5
- Dickinson M., 2000, *Phil. Trans. R. Soc.*, 358, 2001
- Dieleman S., Willett K. W., Dambre J., 2015, *MNRAS*, 450, 1441
- Domínguez Sánchez H., Huertas-Company M., Bernardi M., Tuccillo D., Fischer J. L., 2018, *MNRAS*, 476, 3661
- Draine B. T., Lee H. M., 1984, *ApJ*, 285, 89
- Dwek E., 1998, *ApJ*, 501, 643
- Erb D. K., Steidel C. C., Shapley A. E., Pettini M., Reddy N. A., Adelberger K. L., 2006, *ApJ*, 646, 107
- Fetherolf T. et al. 2020, *MNRAS*, 498, 5009
- Fetherolf T. et al. 2021, *MNRAS*, 508, 1431
- Fitzpatrick E. L., Massa D., 1986, *ApJ*, 307, 286
- Fitzpatrick E. L., Massa D., 1990, *ApJS*, 72, 163
- Förster Schreiber N. M., Shapley A. E., Erb D. K., Genzel R., Steidel C. C., Bouché N., Cresci G., Davies R., 2011, *ApJ*, 731, 65
- Freeman P. E., Izbicki R., Lee A. B., Newman J. A., Conselice C. J., Koekemoer A. M., Lotz J. M., Mozena M., 2013, *MNRAS*, 434, 282
- Galliano F. et al. 2021, *A&A*, 649, A18

- Giavalisco M., Livio M., Bohlin R. C., Macchetto F. D., Stecher T. P., 1996, *AJ*, 112, 369
- Gordon K. D., Clayton G. C., Misselt K. A., Landolt A. U., Wolff M. J., 2003, *ApJ*, 594, 279
- Griffiths R. E. et al. 1994, *ApJ*, 435, L19
- Grogin N. A. et al. 2011, *ApJS*, 197, 35
- Guo Y. et al. 2015, *ApJ*, 800, 39
- Guo Y. et al. 2018, *ApJ*, 853, 108
- Hao C.-N., Kennicutt R. C., Johnson B. D., Calzetti D., Dale D. A., Moustakas J., 2011, *ApJ*, 741, 124
- Hemmati S. et al. 2014, *ApJ*, 797, 108
- Hemmati S., Mobasher B., Darvish B., Nayyeri H., Sobral D., Miller S., 2015, *ApJ*, 814, 46
- Hubble E. P., 1926, *ApJ*, 64, 321
- Hunter J. D., 2007, *Comput. Sci. Eng.*, 9, 90
- Jenkins E. B., 2009, *ApJ*, 700, 1299
- Jones T. et al. 2015, *AJ*, 149, 107
- Jung I. et al. 2017, *ApJ*, 834, 81
- Kartalteppe J. S. et al. 2015, *ApJS*, 221, 11
- Katsianis A. et al. 2017, *MNRAS*, 472, 919
- Kewley L. J., Ellison S. L., 2008, *ApJ*, 681, 1183
- Koekemoer A. M. et al. 2011, *ApJS*, 197, 36
- Kriek M. et al. 2015, *ApJS*, 218, 15
- Lang P. et al. 2014, *ApJ*, 788, 11
- Law D. R., Steidel C. C., Shapley A. E., Nagy S. R., Reddy N. A., Erb D. K., 2012, *ApJ*, 745, 85
- Lee B. et al. 2018, *ApJ*, 853, 131
- Leethochawalit N., Jones T. A., Ellis R. S., Stark D. P., Richard J., Zitrin A., Auger M., 2016, *ApJ*, 820, 84
- Leja J., Johnson B. D., Conroy C., van Dokkum P. G., Byler N., 2017, *ApJ*, 837, 170
- Leung G. C. K. et al. 2019, *ApJ*, 886, 11
- Lintott C. J. et al. 2008, *MNRAS*, 389, 1179
- Lintott C. et al. 2011, *MNRAS*, 410, 166
- Lisker T., 2008, *ApJS*, 179, 319
- Lotz J. M., Primack J., Madau P., 2004, *AJ*, 128, 163
- Lotz J. M. et al. 2008, *ApJ*, 672, 177
- McLean I. S. et al., 2010, in McLean I. S., Ramsay S. K., Takami H., eds, Proc. SPIE Conf. Ser., Vol. 7735, Ground-based and Airborne Instrumentation for Astronomy III. SPIE, p. 77351E
- McKee C., 1989, in Allamandola L. J., Tielens A. G. G. M., eds, Proc. IAU Symp. 135, Interstellar Dust. Kluwer Academic Publishers, Dordrecht, p. 431
- Mannucci F., Cresci G., Maiolino R., Marconi A., Gnerucci A., 2010, *MNRAS*, 408, 2115
- Madau P., Dickinson M., 2014, *ARA&A*, 52, 415
- Mattsson L., De Cia A., Andersen A. C., Zafar T., 2014, *MNRAS*, 440, 1562
- McLean I. S. et al., 2012, in McLean I. S., Ramsay S. K., Takami H., eds, Proc. SPIE Conf. Ser., Vol. 8446, Ground-Based and Airborne Instrumentation for Astronomy IV. SPIE, p. 84460J
- Momcheva I. G. et al., 2016, *ApJS*, 225, 27
- Nelson E. J. et al. 2013, *ApJ*, 763, L16
- Nelson E. J. et al. 2016, *ApJ*, 817, L9
- Noeske K. G. et al. 2007, *ApJ*, 660, L43
- Oke J. B., Gunn J. E. 1983, *ApJ*, 266, 713
- Oliphant T. E., 2007, *Comput. Sci. Engin.*, 9, 10
- Osterbrock D. E., 1989, *Astrophysics of Gaseous Nebulae and Active Galactic Nuclei*. University Science Books, Mill Valley, CA
- Pannella M. et al. 2009, *ApJ*, 698, L116
- Papovich C., Giavalisco M., Dickinson M., Conselice C. J., Ferguson H. C., 2003, *ApJ*, 598, 827
- Papovich C., Dickinson M., Giavalisco M., Conselice C. J., Ferguson H. C., 2005, *ApJ*, 631, 101
- Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2002, *AJ*, 124, 266
- Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2010, *AJ*, 139, 2097
- Peters T. et al. 2017, *MNRAS*, 467, 4322
- Pettini M., Pagel B. E. J., 2004, *MNRAS*, 348, L59
- Price S. H. et al. 2014, *ApJ*, 788, 86
- Reddy N. A., Steidel C. C., Fadda D., Yan L., Pettini M., Shapley A. E., Erb D. K., Adelberger K. L., 2006, *ApJ*, 644, 792
- Reddy N. A., Erb D. K., Pettini M., Steidel C. C., Shapley A. E., 2010, *ApJ*, 712, 1070
- Reddy N. A., Pettini M., Steidel C. C., Shapley A. E., Erb D. K., Law D. R., 2012, *ApJ*, 754, 25
- Reddy N. A. et al. 2015, *ApJ*, 806, 259
- Reddy N. A. et al. 2018a, *ApJ*, 853, 56
- Reddy N. A. et al. 2018b, *ApJ*, 869, 92
- Reddy N. A. et al. 2020, *ApJ*, 902, 123
- R  my-Ruyer A. et al. 2014, *A&A*, 563, A31
- Salim S., Boquien M., Lee J. C., 2018, *ApJ*, 859, 11
- Sanders R. L. et al. 2015, *ApJ*, 799, 138
- Sanders R. L. et al. 2018, *ApJ*, 858, 99
- Sanders R. L. et al., 2021, *ApJ*, 914, 19
- S  rsic J., 1963, *BAAA*, 6, 41
- Shapley A. E., 2011, *ARA&A*, 49, 525
- Shapley A. E., Steidel C. C., Adelberger K. L., Dickinson M., Giavalisco M., Pettini M., 2001, *ApJ*, 562, 95
- Shapley A. E., Cullen F., Dunlop J. S., McLure R. J., Kriek M., Reddy N. A., Sanders R. L., 2020, *ApJ*, 903, L16
- Sharma S., 2017, *ARA&A*, 55, 213
- Shivaei I. et al. 2015, *ApJ*, 815, 98
- Shivaei I. et al. 2017, *ApJ*, 837, 157
- Shivaei I. et al. 2018, *ApJ*, 855, 42
- Shivaei I. et al. 2020, *ApJ*, 899, 117
- Simons R. C. et al. 2021, *ApJ*, 923, 203
- Skelton R. E. et al. 2014, *ApJS*, 214, 24
- Tacchella S. et al. 2018, *ApJ*, 859, 56
- Tielens A. G. G. M., 1998, *ApJ*, 499, 267
- Tremonti C. A. et al. 2004, *ApJ*, 613, 898
- van den Bergh S., 2002, *Publ. Astron. Soc. Pac.*, 114, 797
- van der Wel A. et al. 2014, *ApJ*, 788, 28
- Wang X. et al. 2019, *ApJ*, 882, 94
- Whitaker K. E., van Dokkum P. G., Brammer G., Franx M., 2012, *ApJ*, 754, L29
- Whitaker K. E. et al. 2014, *ApJ*, 795, 104
- Wild V., Charlot S., Brinchmann J., Heckman T., Vince O., Pacifici C., Chevallard J., 2011, *MNRAS*, 417, 1760
- Worthey G. 1994, *ApJS*, 95, 107
- Wuyts S. et al. 2011, *ApJ*, 738, 106
- Wuyts S. et al. 2012, *ApJ*, 753, 114
- Wuyts S. et al. 2013, *ApJ*, 779, 135
- Wuyts E. et al. 2016, *ApJ*, 827, 74
- Yoshikawa T. et al. 2010, *ApJ*, 718, 112

APPENDIX A: ANALYSIS OF THE PATCHINESS METRIC

In Section 3.1, we defined a morphology measure that is sensitive to both faint and bright regions within galaxies, called ‘patchiness’ (equation 2). In this appendix, an in-depth analysis of the patchiness metric, P , is presented using the patchiness calculated from the resolved $E(B - V)_{\text{stars}}$, stellar population age, SFR, and stellar mass maps. This analysis of the patchiness metric includes observed correlations between the calculated P of all resolved SED-derived properties; a comparison between P and the ICD (Papovich et al. 2003), Gini (Abraham et al. 2003; Lotz et al. 2004), and M_{20} (Lotz et al. 2004, 2008) morphology metrics; P calculated from the pixel-scale distribution; the behaviour of P when galaxies are artificially redshifted; and the potential dependence of P on the number of resolved elements.

First, the calculated P on the resolved SED-derived $E(B - V)_{\text{stars}}$, stellar population ages, SFRs, and stellar masses are directly compared in Fig. A1. If a given galaxy exhibits a patchy distribution in any SED-derived property, it is likely to exhibit a patchy distribution

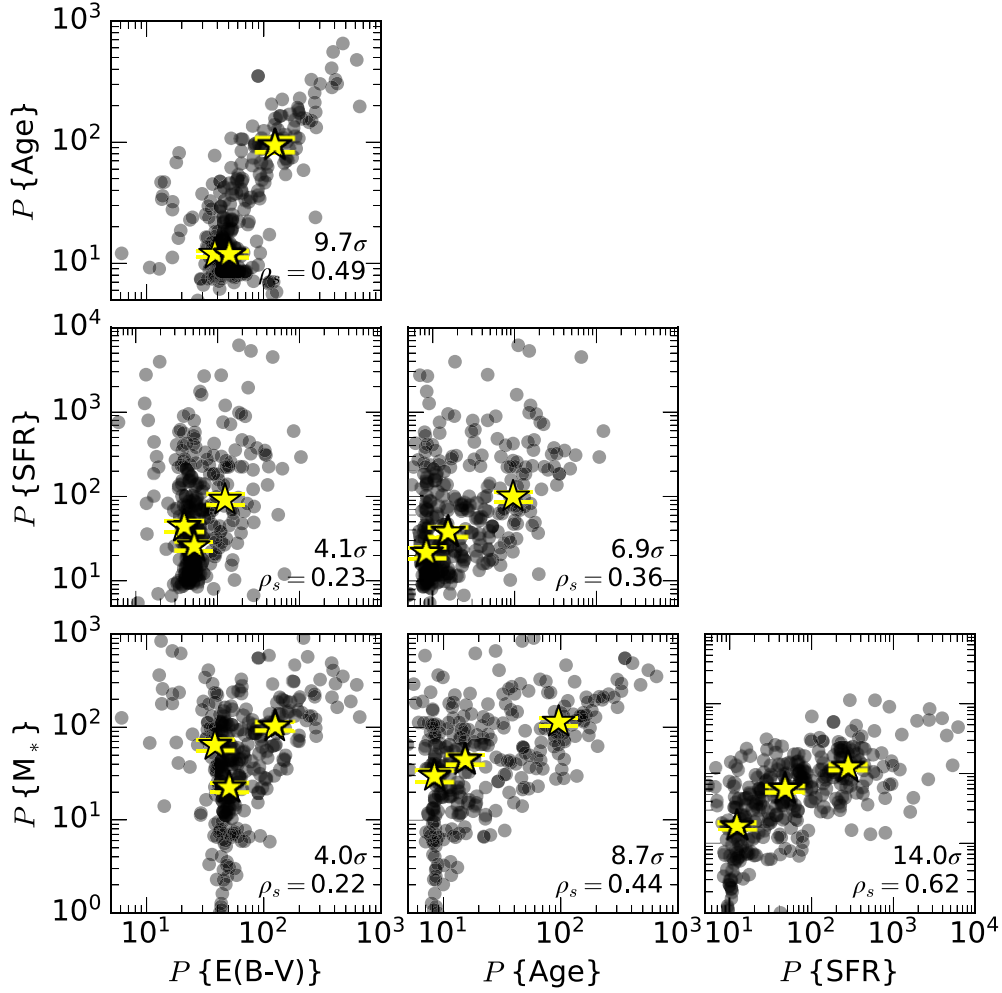


Figure A1. Relationships between P calculated from the resolved SED-derived $E(B - V)_{\text{stars}}$, stellar population ages, SFRs, and stellar masses for the 310 galaxies in our sample. The yellow stars show the median P values in three equally sized bins that are based on the P measurement of the x-axis, with the yellow bars indicating the standard error in the median P values. The Spearman rank correlation coefficient and its significance is listed in the bottom right-hand corner of each panel.

in the other resolved stellar population properties. Most significantly, galaxies with patchier SFRs tend to have patchier stellar masses and galaxies with patchier $E(B - V)_{\text{stars}}$ distributions tend to have patchier stellar ages. While the relationship between the P of the resolved SFRs and stellar masses could provide clues as to how galaxies build their stellar mass in the context of the resolved SFR- M_* relationship, the significance of this relationship could alternatively be explained by the covariance between SFR and stellar mass as both are derived from the normalization of the SED model to the photometry. Similarly, the relationship between patchy reddening and stellar ages could be related to the age-extinction degeneracy (e.g. Worthey 1994; Shapley et al. 2001), where the SED of an old, dust-free stellar population has a similar shape compared to the SED of a young, dusty stellar population.

Of the preexisting morphology metrics, patchiness is most closely related to the ICD (Papovich et al. 2003). The ICD measures deviations in colour and, thus, is sensitive to the general dispersion in flux rather than exclusively the brightest regions. The primary advantage of patchiness compared to the ICD is that patchiness can be applied to any resolved property, whereas the ICD is limited to a single colour (i.e. resolved imaging for two filters). Based on the methodology presented in Section 2.4, P calculated on the resolved

stellar population and reddening maps incorporates information from at least five filters with resolved imaging. We measure the ICD using the $V_{606}-H_{160}$ colour,¹² which spans the Balmer/4000 Å breaks and thus is expected to have some dependence on the dispersion in stellar population age and stellar mass. The ICD measured from the $V_{606}-H_{160}$ colour is compared to the P calculated from the resolved stellar population properties in Fig. A2 and is found to correlate most significantly with the P of the $E(B - V)_{\text{stars}}$ and stellar mass distribution, with a marginal correlation towards higher P measured from the stellar population age distribution.

Throughout this paper, we use patchiness alongside the previously defined Gini and M_{20} coefficients (Abraham et al. 2003; Lotz et al. 2004, 2008). Fig. A3 shows how patchiness is related to G and M_{20} when each of these morphology metrics are measured on the resolved SED-derived $E(B - V)_{\text{stars}}$, stellar population ages, SFRs, and stellar masses. In all cases, P and G are significantly positively correlated, whereas there is no significant correlation between P and M_{20} . The correlation between P and G is attributed to both metrics

¹²The V_{606} and H_{160} filters are the bluest and reddest resolved filters, respectively, that were used to observe all of the galaxies in our sample.

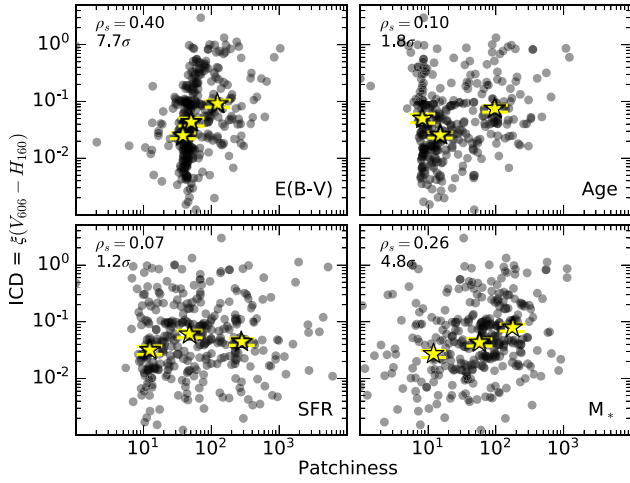


Figure A2. The ICD (Papovich et al. 2003) measured from the $V_{606} - H_{160}$ colour-resolved imaging versus the P measured from the resolved stellar population and reddening maps for the 310 galaxies in our sample. The yellow stars show the median ICD and P values in three equally sized bins that are based on the P measurement of the x -axis, with the yellow bars indicating the standard error in the median ICD and P values. The Spearman rank correlation coefficient and its significance is listed in the top left-hand corner of each panel.

being sensitive to the resolved components with the highest values relative to the average of the distribution, but P is additionally driven by regions with the lowest values relative to the average. While M_{20} is similarly based on the highest values of a resolved distribution, M_{20} is primarily sensitive to the spatial distribution of the highest valued regions and, thus, is not expected to be directly correlated with P or G .

To examine the behaviour of the patchiness metric on other resolved scales, patchiness is calculated on the pixel-to-pixel scale for a subset of 50 galaxies that have stellar population parameters that are representative of the larger sample. In the pixel-to-pixel patchiness calculation, all pixels identified by the 3D-*HST* segmentation map are included. Fig. A4 shows how P measured on the pixel-to-pixel

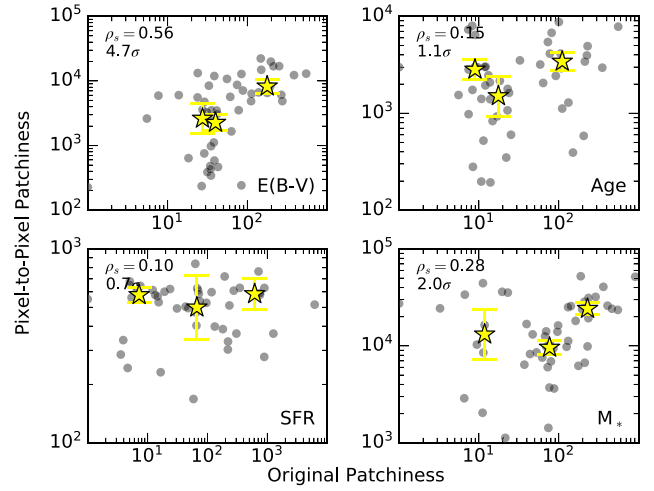


Figure A4. The P calculated from the pixel-to-pixel $E(B - V)_{\text{stars}}$, stellar population age, SFR, and stellar mass distributions compared to the P calculated from the Voronoi bin stellar population properties for a subset of 50 galaxies in the sample. The yellow stars show the median P values in three equally sized bins that are based on the P measurement of the x -axis, with the yellow bars indicating the standard error in the median P values. The Spearman rank correlation coefficient and its significance is listed in the top left-hand corner of each panel.

distribution compares to P measured on the Voronoi bin distribution. The relative P calculated on the $E(B - V)_{\text{stars}}$ distribution is generally preserved between a pixel and Voronoi bin scale and marginally preserved for P calculated on the stellar mass distribution. However, caution should be taken when measuring P on a pixel-to-pixel scale due to correlated signal between neighbouring pixels and low S/N pixels that cause unphysical outliers in the resolved distribution that artificially drive P to higher values. Therefore, we discourage the use of the patchiness parameter on pixel-to-pixel scales. Alternatively, we recommend using a minimum bin size that is at least as large as the point spread function and that low S/N components that can cause unreliable resolved measurements are removed from the analysis (also see Fetherolf et al. 2020).

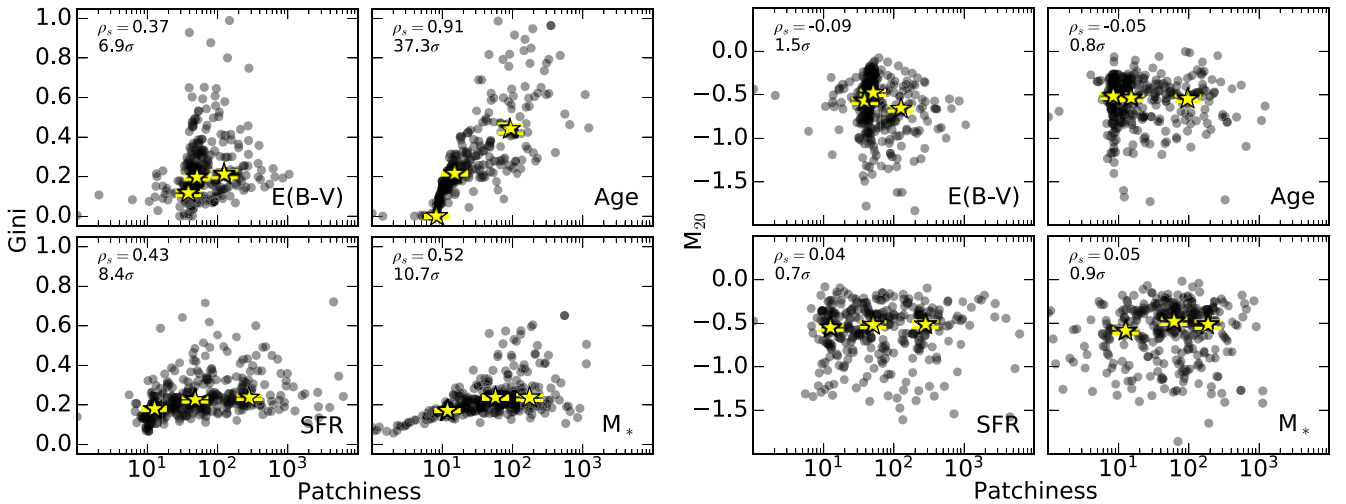


Figure A3. The Gini (left-hand panel) and M_{20} (right-hand panel) coefficients compared to the patchiness metric, all calculated from the $E(B - V)_{\text{stars}}$, stellar population age, SFR, and stellar mass distributions for the 310 galaxies in our sample. The yellow stars show the median G and M_{20} values in three equally sized bins that are based on the P measurement of the x -axis, with the yellow bars indicating the standard error in the median G and M_{20} values. The Spearman rank correlation coefficient and its significance is listed in the top left-hand corner of each panel.

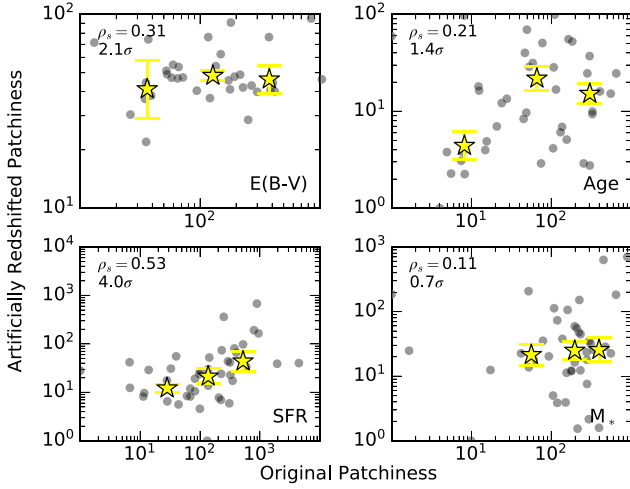


Figure A5. The P calculated from stellar population and reddening maps for a subset of 43 galaxies that have been artificially redshifted from $z \sim 1.5$ to 2.3 compared to their original P values. The yellow stars show the median P values in three equally sized bins that are based on the P measurement of the x -axis, with the yellow bars indicating the standard error in the median P values. The Spearman rank correlation coefficient and its significance is listed in the top left-hand corner of each panel.

The effect of redshift on the patchiness metric is explored by artificially redshifting each of the galaxies in the $z \sim 1.5$ sample (126 galaxies) to a redshift that is randomly selected from a Gaussian distribution with the mean and standard deviation of the $z \sim 2.3$ sample ($z = 2.30 \pm 0.14$). For simplicity, the change in pixel scale between these redshifts is considered negligible and no correction is applied for the shift in the bandpass due to the large widths of the *HST* filters. Cosmological dimming is applied by multiplying the observed flux in each filter by the square of the ratio of luminosity distances

at the two redshifts (Barden, Jahnke & Häußler 2008). The Voronoi binning and resolved SED-fitting procedures are then repeated using the dimmed flux distributions. The artificially redshifted sample is restricted to galaxies with at least five Voronoi bins that have an $S/N \geq 5$ in at least five filters (see Section 2.4), resulting in a subsample of 43 galaxies that have been artificially redshifted to $z \sim 2.3$. The P measured from the artificially redshifted stellar population and reddening maps is compared to the original P measurements in Fig. A5. Except in the case of the resolved SFRs, the ordering of P is generally not preserved when the effects of cosmological dimming are considered. Furthermore, there are intrinsic differences in the physical properties of galaxies between redshifts $z \sim 1.5$ and 2.3 (e.g. see recent reviews by Shapley 2011; Madau & Dickinson 2014). Therefore, we recommend using P only for relative comparisons of galaxies at similar redshifts and in uniformly defined sample selections (see Section 3.1).

Finally, a possible dependence between P and the number of Voronoi bins in the galaxy is investigated through Fig. A6. The $z \sim 1.5$ (left-hand panels) and $z \sim 2.3$ samples (right-hand panels) are each split into two equally sized bins based on the number of Voronoi bins. Their distribution in P (blue and red histograms) and average P (vertical dashed lines) are shown for each stellar population property. While galaxies with more Voronoi bins (red histogram) tend to have higher P on average, the range of the distribution in P is either consistent or broader than the distribution of P for galaxies with fewer Voronoi bins (blue histogram). One exception is the distribution of P calculated from the resolved stellar mass (bottom right sub-panels of Fig. A6), where galaxies with more Voronoi bins tend to have patchier stellar mass distributions while galaxies with fewer Voronoi bins have smoother stellar mass distributions. However, since this offset is not equivalently seen for all resolved stellar population properties, we suggest that the relationship between the number of resolved elements and the calculated P is physically driven – especially regarding stellar mass.

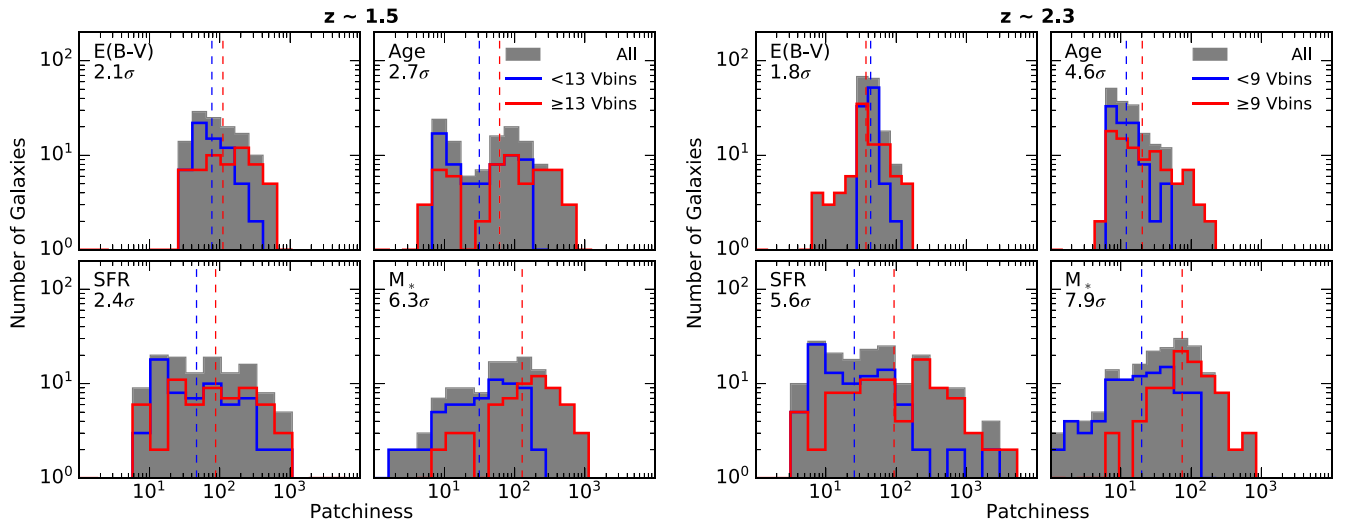


Figure A6. The distribution of P measured from the resolved $E(B - V)_{\text{stars}}$, stellar population ages, SFRs, and stellar masses for the 126 galaxies in the $z \sim 1.5$ sample (left-hand panels) and 184 galaxies in the $z \sim 2.3$ sample (right-hand panels). The distribution of the full sample (grey histogram) is shown alongside two equally sized bins based on the number of Voronoi bins, where galaxies with few Voronoi bins are outlined by the blue histogram and galaxies with many Voronoi bins are highlighted by the red histogram. The vertical dashed lines indicate the average P of the two bins. The significance in the difference between the average of each distribution derived using a two-sided t -test is listed in the top left-hand corner of each sub-panel. While galaxies with more Voronoi bins tend to have higher P than galaxies with fewer Voronoi bins, the range in P is consistent between the two distributions for most stellar population properties (except stellar mass) such that P is determined to not be systematically increased for galaxies with more resolved elements, but a relationship between P and the number of Voronoi bins may be physically driven.

Furthermore, the number and size of Voronoi bins are defined by grouping regions of similar brightness (Cappellari & Copin 2003) and low S/N components that could artificially drive P towards higher values are not included in the analysis (see Section 2.4). Therefore, we conclude that the calculation of P is not systematically biased towards higher P when a galaxy contains more resolved elements.

The most important conclusions from this in-depth analysis on the patchiness metric are summarized as follows (1) calculating P on the pixel-to-pixel distribution is not recommended due to unphysical low S/N outliers that could drive P towards higher values, (2) the P (and other morphology metrics) of galaxies at different redshifts should be analyzed separately, and (3) P is not dependent on the number of resolved elements within a given galaxy. These conclusions align with our recommendation to compare the relative P values of resolved distributions for galaxies within uniformly defined samples. The resolved distributions for the sample of galaxies used throughout this paper are uniformly defined by separating the $z \sim 1.5$ and 2.3 samples

and applying an strict criteria for including resolved elements in the analysis (see Section 2.4). The findings in this appendix also highlight how the patchiness metric can be used on other resolved stellar population properties [besides $E(B - V)_{\text{stars}}$, which is the focus of this work].

APPENDIX B: $E(B - V)$ MAPS ORDERED BY PATCHINESS

In Fig. B1, we show examples of the $E(B - V)_{\text{stars}}$ maps ordered from the highest to lowest P values measured from the $E(B - V)_{\text{stars}}$ maps. It can be seen that the galaxies with the highest P values tend to have more complex $E(B - V)_{\text{stars}}$ distributions, whereas the galaxies with the lowest P values tend to have much more uniform $E(B - V)_{\text{stars}}$ distributions. Furthermore, the lowest P values can exhibit either generally low or high $E(B - V)_{\text{stars}}$ and can be made of a high number of resolved elements (see Appendix A).

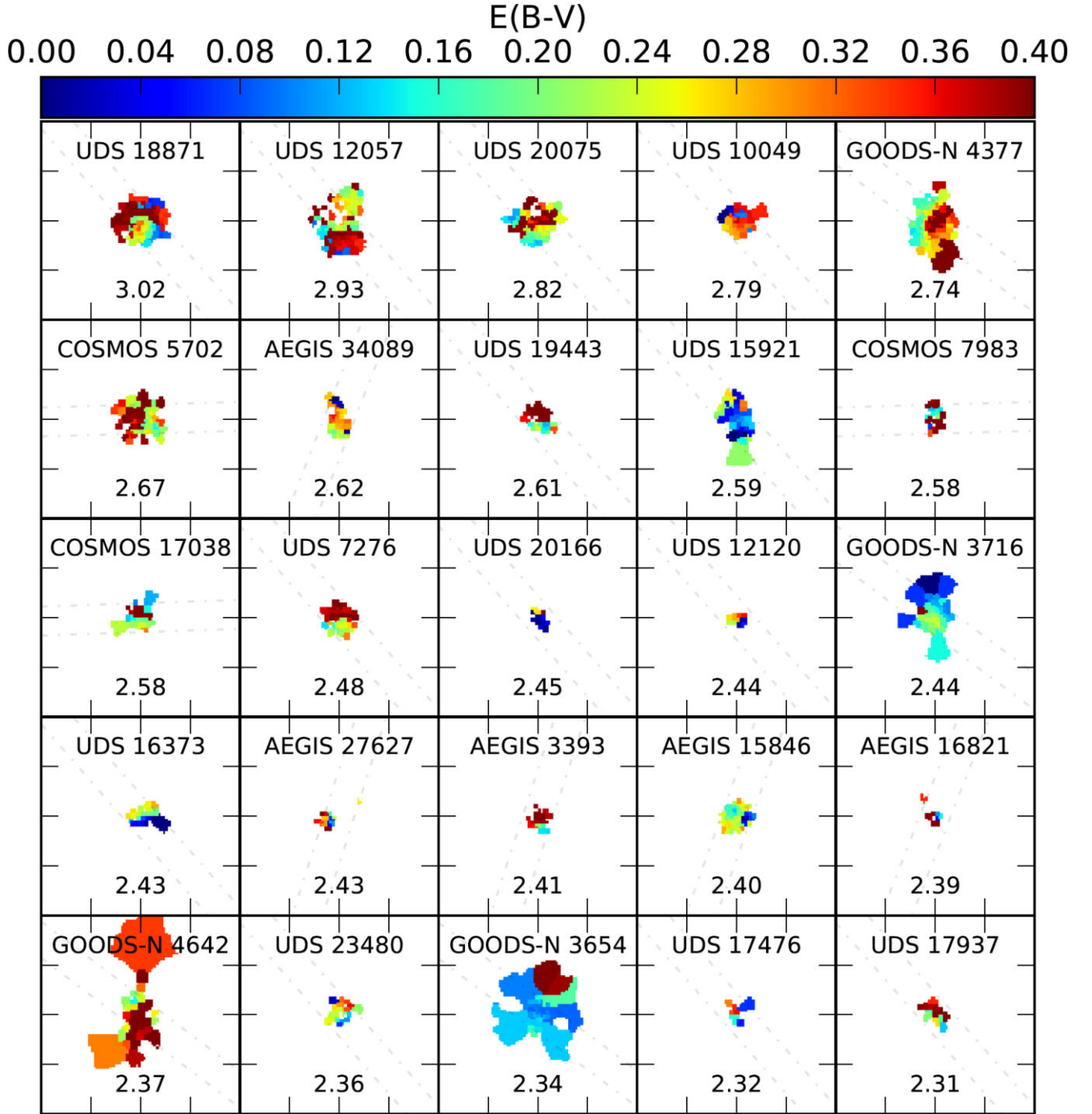


Figure B1. Examples of the $E(B - V)_{\text{stars}}$ maps for galaxies with the highest $\log(P)$ values, which are listed at the bottom of each panel. The dash-dotted grey lines show the placement of the MOSFIRE spectroscopic slit. Examples of the $E(B - V)_{\text{stars}}$ maps for galaxies with intermediate $\log(P)$ values, which are listed at the bottom of each panel. Examples of the $E(B - V)_{\text{stars}}$ maps for galaxies with the lowest $\log(P)$ values, which are listed at the bottom of each panel.

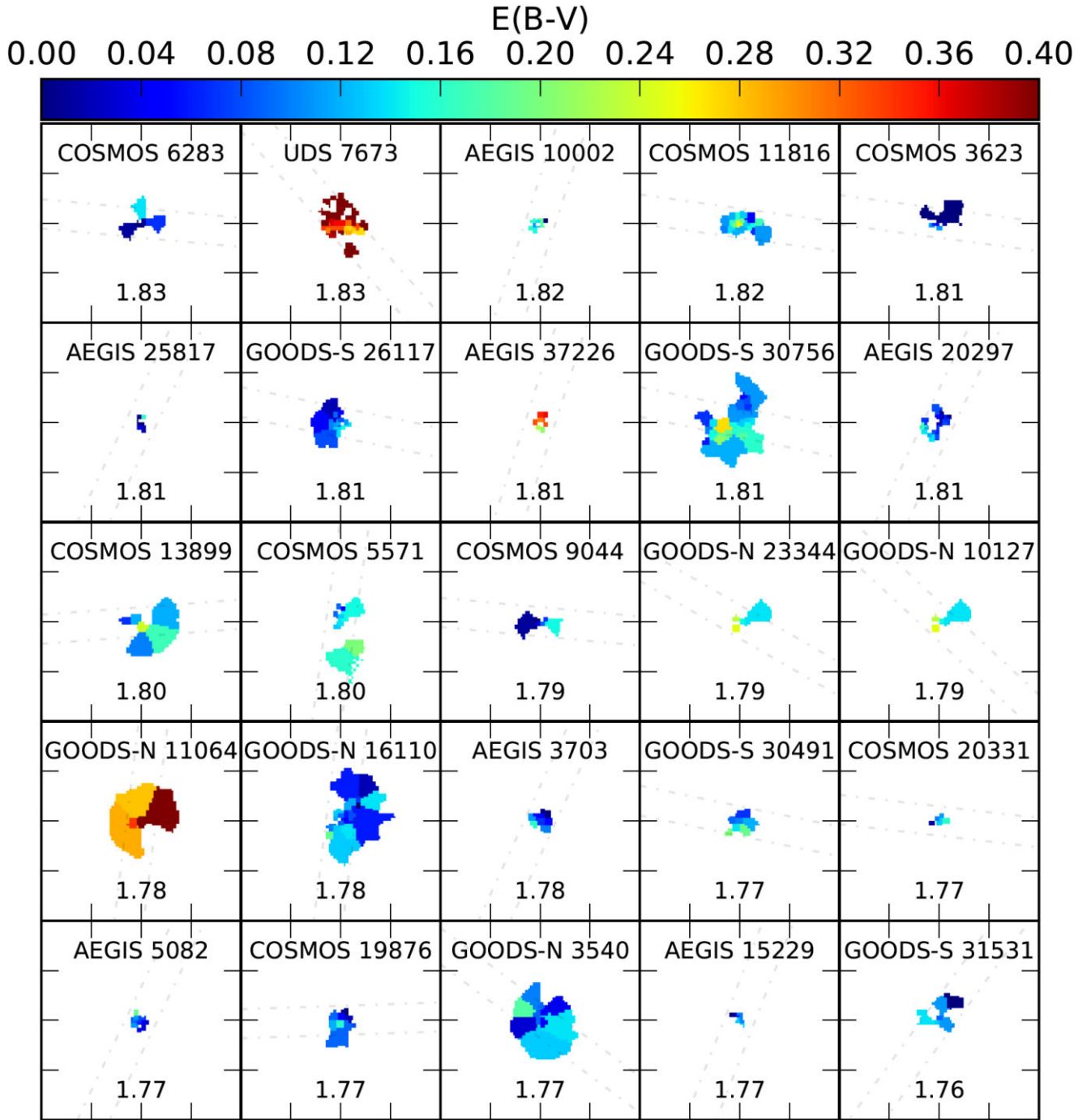


Figure B1 – continued

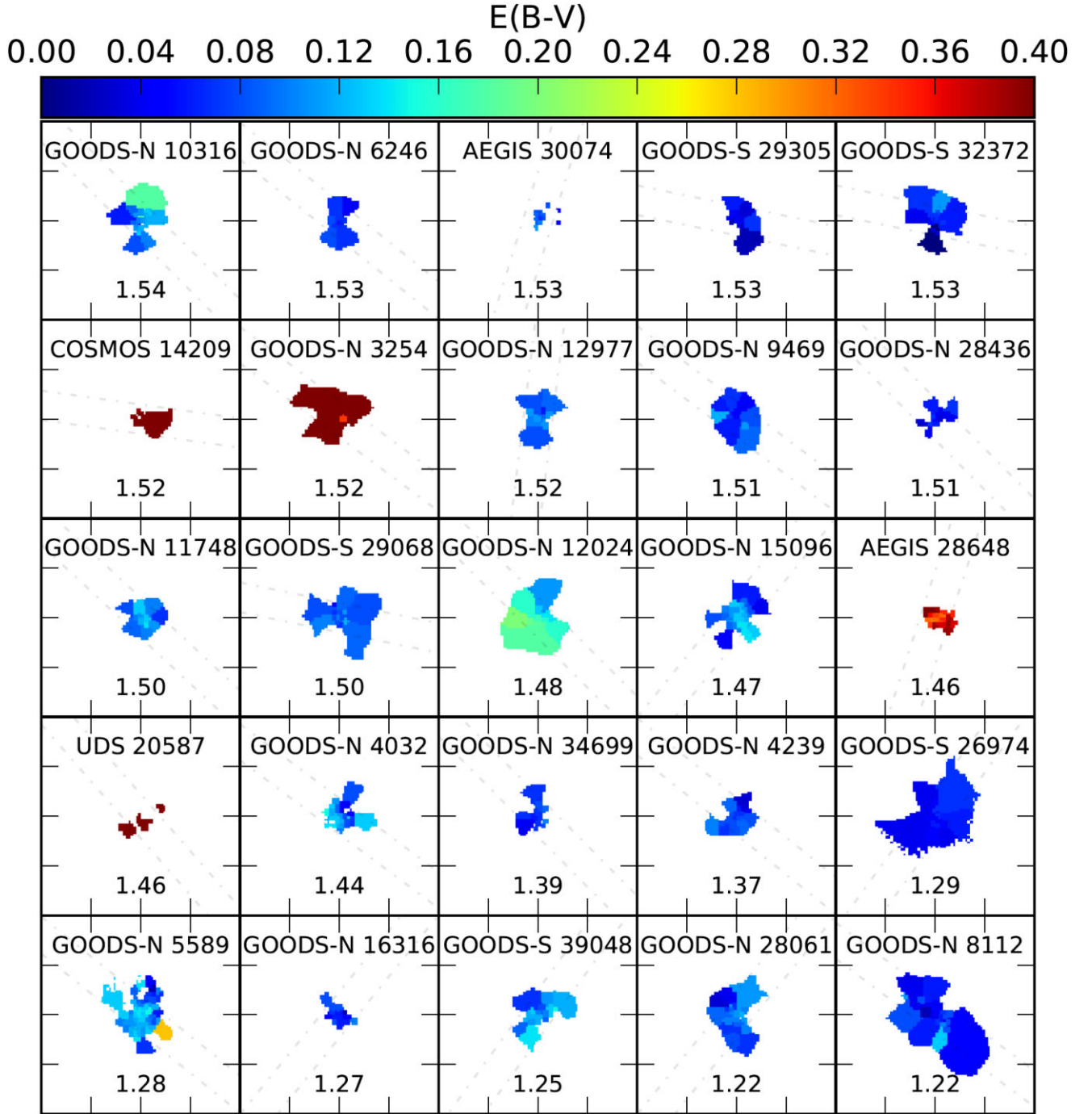


Figure B1 – continued

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