Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **498,** 2645–2661 (2020) doi:10.1093/mnras/staa2586
Advance Access publication 2020 August 26

Effects of self-consistent rest-ultraviolet colours in semi-empirical galaxy formation models

Jordan Mirocha [1]★† Charlotte Mason [2]‡ and Daniel P. Stark[3]

[1]*McGill University Department of Physics and McGill Space Institute, 3600 Rue University, Montréal, QC H3A 2T8, Canada*
[2]*Center for Astrophysics | Harvard and Smithsonian, 60 Garden St, Cambridge, MA 02138, USA*
[3]*Steward Observatory, University of Arizona, 933 N Cherry Ave, Tucson, AZ 85721, USA*

Accepted 2020 August 20. Received 2020 August 10; in original form 2020 May 14

ABSTRACT

Connecting the observed rest-ultraviolet (UV) luminosities of high-z galaxies to their intrinsic luminosities (and thus star formation rates, SFRs) requires correcting for the presence of dust. We bypass a common dust-correction approach that uses empirical relationships between infrared (IR) emission and UV colours, and instead augment a semi-empirical model for galaxy formation with a simple – but self-consistent – dust model and use it to jointly fit high-z rest-UV luminosity functions (LFs) and colour–magnitude relations ($M_{\rm UV}$–β). In doing so, we find that UV colours evolve with redshift (at fixed UV magnitude), as suggested by observations, even in cases without underlying evolution in dust production, destruction, absorption, or geometry. The observed evolution in our model arises due to the reduction in the mean stellar age and rise in specific SFRs with increasing z. The UV extinction, $A_{\rm UV}$, evolves similarly with redshift, though we find a systematically shallower relation between $A_{\rm UV}$ and $M_{\rm UV}$ than that predicted by IRX–β relationships derived from $z \sim 3$ galaxy samples. Finally, assuming that high 1600-Å transmission ($\gtrsim$0.6) is a reliable Ly α emitter (LAE) indicator, modest scatter in the effective dust surface density of galaxies can explain the evolution both in $M_{\rm UV}$–β and LAE fractions. These predictions are readily testable by deep surveys with the *James Webb Space Telescope*.

Key words: dust, extinction – galaxies: evolution – galaxies: high-redshift – galaxies: ISM.

1 INTRODUCTION

Current constraints on galaxy formation are based largely on the rest-ultraviolet (UV) properties of redshift $z \gtrsim 4$ galaxies, e.g. luminosity functions (LFs; Bouwens et al. 2015; Finkelstein et al. 2015) and UV colour–magnitude relations (CMDs, $M_{\rm UV}$–β; Bouwens et al. 2009; Finkelstein et al. 2012; Dunlop et al. 2013; Bouwens et al. 2014). Such observations probe the star formation rate (SFR) of high-z galaxies, given that the rest-UV emission is dominated by massive young stars, and have thus allowed astronomers to begin piecing together the cosmic star formation rate density (SFRD) in the early Universe (see Madau & Dickinson 2014, for a review). UV colours, generally quantified by a power-law spectral slope β (defined by $f_\lambda \propto \lambda^\beta$), are critical to this inference as they are modulated by dust extinction in a characteristic wavelength-dependent manner, allowing one to 'dust correct' UV magnitude measurements, so long as multiband photometry covering the rest UV continuum ($1300 \lesssim \lambda/\text{Å} \lesssim 2600$) is available.

Unfortunately, the link between β and UV extinction, $A_{\rm UV}$, is potentially complicated. One common approach is to assume that thermal radiation emitted from dust grains (in the infrared, IR) is a reliable tracer of the energy lost in the UV. In low-redshift star-forming galaxies, for which there is both rest-UV and rest-IR coverage, there is indeed a relationship between a galaxy's infrared 'excess' and its UV slope (the so-called IRX–β relationship; Meurer, Heckman & Calzetti 1999). For a known input stellar spectrum, assumed dust opacity (as a function of wavelength), and bolometric correction (to recover total dust emission from narrow-band IR observations), one can determine $A_{\rm UV}$ from β, and thus convert observed magnitudes into intrinsic magnitudes. Meurer et al. (1999, hereafter M99) found that $A_{1600} = 4.43 - 1.99\beta$, using a Calzetti, Kinney & Storchi-Bergmann (1994, hereafter C94) dust law and input stellar spectra from STARBURST99 (Leitherer et al. 1999).

In recent years, many groups have adopted some variant of the IRX–β-based procedure as a way to use observed UVLFs and CMDs to calibrate semi-analytic models (SAMs), which may not model dust explicitly. There are several ways this approach may break down. For example, the origin of the IRX–β relation is an active area of research, both observationally (e.g. Overzier et al. 2011; Casey et al. 2014; Reddy et al. 2018) and theoretically (e.g. Narayanan et al. 2018; Ma et al. 2019; Salim & Boquien 2019; Schulz et al. 2020), so it may be premature to apply it at arbitrarily high redshift, where the properties of stars and dust may differ from low-z samples. Inferences based on IRX–β arguments could be biased for a less interesting reason, which is that the assumptions underlying the M99 relation are often not made in SAMs. For example, adopting a different stellar population synthesis (SPS) model can change the input stellar spectrum, as can changes in the star formation histories (SFHs) of individual galaxies,

★ E-mail: jordan.mirocha@mcgill.ca
† CITA National Fellow.
‡ Hubble Fellow.

thus posing a self-consistency issue when invoking empirical IRX–β relations. For example, model galaxies generally have rapidly rising SFHs, and are thus intrinsically bluer than the assumed input β adopted in M99 (see e.g. Finlator, Oppenheimer & Davé 2011; Wilkins et al. 2013; Mancini et al. 2016), which is based on the assumption of a constant SFR.

In this work, we take a different approach that avoids self-consistency issues by using a simple dust model in lieu of an IRX–β assumption. This approach allows us to compute self-consistent solutions for the spectra of objects in our model, assess the circumstances in which evolution in the properties of dust is required by current measurements, and make physically motivated predictions for upcoming β measurements to be conducted with the *James Webb Space Telescope* (*JWST*).

In Section 2, we detail our dust model and the underlying assumptions about star formation in early galaxies. We detail our main results in Section 3 and discuss them in a broader context in Section 4. In Section 5, we summarize our findings.

We adopt AB magnitudes throughout (Oke & Gunn 1983), i.e.

$$M_\lambda = -2.5 \log_{10} \left(\frac{f_\lambda}{3631 \text{ Jy}} \right), \tag{1}$$

and adopt the following cosmology: $\Omega_\mathrm{m} = 0.3156$, $\Omega_\mathrm{b} = 0.0491$, $h = 0.6726$, and $\sigma_8 = 0.8159$, very similar to the recent Planck Collaboration VI (2018) constraints.

2 MODEL

Our model is similar to other semi-empirical models that have appeared in the literature in recent years. We outline our model for star formation in galaxies in Section 2.1, our approach to dust in Section 2.2, and our method for generating synthetic spectra and estimating UV colours in Section 2.3. Much of this has been described previously (Mirocha, Furlanetto & Sun 2017; Mirocha & Furlanetto 2019), and is publicly available within the ARES[1] code.

2.1 Star formation

We assume the SFR is proportional to the baryonic mass accretion (MAR) on to dark matter (DM) haloes (as in, e.g. Mason, Trenti & Treu 2015; Sun & Furlanetto 2016), i.e.

$$\dot{M}_*(M_\mathrm{h}, z) = f_*(M_\mathrm{h}, z) \dot{M}_\mathrm{b}(z, M_\mathrm{h}), \tag{2}$$

where f_* is the efficiency of star formation. The baryonic MAR is well approximated by a power law in mass and redshift (e.g. McBride, Fakhouri & Ma 2009; Dekel et al. 2013); however, rather than adopting a parametric form for the MAR calibrated by simulations, we derive it directly from the halo mass function (HMF; see appendix A of Furlanetto et al. 2017, for more details). We adopt the Tinker et al. (2010) HMF in this work, generated by the HMF code[2] (Murray, Power & Robotham 2013).

We assume that the star formation efficiency (SFE) is a double power law in M_h (Moster et al. 2010), i.e.

$$f_*(M_\mathrm{h}) = \frac{f_{*,10}\, \mathcal{C}_{10}}{\left(\frac{M_\mathrm{h}}{M_\mathrm{p}}\right)^{-\alpha_{*,\mathrm{lo}}} + \left(\frac{M_\mathrm{h}}{M_\mathrm{p}}\right)^{-\alpha_{*,\mathrm{hi}}}}, \tag{3}$$

where $f_{*,10}$ is the SFE at $10^{10}\,\mathrm{M}_\odot$, M_p is the mass at which f_* peaks, and $\alpha_{*,\mathrm{hi}}$ and $\alpha_{*,\mathrm{lo}}$ describe the power-law index at masses above and below the peak, respectively. The additional constant $\mathcal{C}_{10} \equiv (10^{10}/M_\mathrm{p})^{-\alpha_{*,\mathrm{lo}}} + (10^{10}/M_\mathrm{p})^{-\alpha_{*,\mathrm{hi}}}$ is introduced to re-normalize the standard double power-law formula to $10^{10}\,\mathrm{M}_\odot$, rather than the peak mass. This model predicts $z > 6$ UVLFs in good agreement with observations when calibrating only to measurements at $z \sim 6$ (Furlanetto et al. 2017; Mirocha et al. 2017), in agreement with the results of other similar models from recent studies (e.g. Trenti et al. 2010; Behroozi, Wechsler & Conroy 2013a; Mason et al. 2015; Mashian, Oesch & Loeb 2016; Sun & Furlanetto 2016; Tacchella et al. 2018; Behroozi et al. 2019).

In this work, we extend our models to $z \sim 4$ to more adequately address issues of time evolution in the M_UV–β relationship and UVLFs. This redshift range, though somewhat arbitrary, has become the de-facto interval for this kind of modelling in recent years, presumably due to the availability of homogenous data sets, though pushing to even lower redshifts would of course be advantageous and is a current work in progress. We include lognormal scatter in the SFR of haloes at fixed halo mass, $\sigma_{\log_{10}\mathrm{SFR}} = 0.3$, which we take to represent scatter in halo accretion rates, but such scatter will also resemble the dispersion in halo assembly times (e.g. Ren, Trenti & Mutch 2018). We then synthesize the spectra of all galaxies in the model, rather than assuming a constant steady-state value for the relationship between UV luminosity and SFR. In other words, for each halo in our model, with index i, we determine the intrinsic spectrum at redshift z_obs by integrating over the past SFH, i.e.

$$L_{\lambda,i}(z_\mathrm{obs}) = \int_{z_\mathrm{obs}}^{z_\mathrm{form}} \dot{M}_{*,i}(z') l_\lambda(\Delta t') \left| \frac{\mathrm{d}t}{\mathrm{d}z} \right| \mathrm{d}z, \tag{4}$$

where $l_\lambda(\Delta t')$ is the luminosity of a simple stellar population of age $\Delta t = t(z_\mathrm{obs}) - t(z')$. We include both the stellar continuum, provided by BPASS, and nebular continuum emission following standard procedures: For free–free and free–bound emission, we use the emission coefficients of Ferland (1980, his table 1), while for the two-photon emission probability, we take the fitting formula of Fernandez & Komatsu (2006, their equation 22), which was derived from the tabulated results of Brown & Mathews (1970). We adopt a temperature of 2×10^4 K in H II regions when generating the nebular continuum. We find that inclusion of the nebular continuum constitutes a small correction, equivalent to an $\sim$10 per cent reduction in dust content.

The synthesized luminosity is then reddened by an optical depth $\tau_{\lambda,i}$, yielding the observed luminosity,

$$L'_{\lambda,i} = L_{\lambda,i}(z_\mathrm{obs}) \mathrm{e}^{-\tau_{\lambda,i}}. \tag{5}$$

We adopt the BPASS version 1.0 (Eldridge & Stanway 2009) single-star models throughout when modeling $l_\lambda(\Delta t')$ with an intermediate stellar metallicity of $Z_\odot/5$. These choices largely affect the inferred normalization of the SFE and dust opacity. As a result, any change to the stellar model will largely be absorbed by normalization parameters, leaving constraints on the *shape* of the SFE and dust scale length relatively unaffected.

2.2 Dust obscuration

For each galaxy in our model we track the build-up of metals by assuming a fixed metal yield per unit SFR, i.e. $\dot{M}_Z = f_Z \dot{M}_*$, where the metal production efficiency f_Z is set to 0.1 in our fiducial case. We further assume that a fraction $f_\mathrm{d} = 0.4$ of these metals reside in dust grains (Dwek 1998). This 'instantaneous recyling' approximation is reasonable, at least in the $z \gtrsim 6$ limit, as the Universe is too young for asymptotic giant branch stars to have become a non-negligible source of dust production (e.g. Dwek, Galliano & Jones 2007).

[1] https://ares.readthedocs.io/en/latest/.
[2] https://hmf.readthedocs.io/en/latest/.

To redden galaxy spectra, we must also make an assumption about the geometry of the dust distribution and the opacity of dust (per unit mass). For simplicity, we adopt a simple spherically-symmetric dust screen model, where the dust optical depth along obscured lines of sight is given by

$$\tau_\lambda = \int_0^R \rho_{\rm dust}(r)\kappa_\lambda {\rm d}r. \quad (6)$$

We take the absorption coefficient κ to be a power law as we only explore a relatively narrow range in wavelength in this study,

$$\kappa_\lambda = \kappa_{1000}\left(\frac{\lambda}{10^3\,\text{Å}}\right)^{\gamma_\kappa}, \quad (7)$$

where $\kappa_{1000} = \kappa(\lambda = 10^3\,\text{Å}) \equiv 10^5 {\rm cm}^2\,{\rm g}^{-1} \simeq 20\,{\rm pc}^2{\rm M}_\odot^{-1}$ and $\gamma_\kappa \equiv -1$ in our fiducial model. These choices are consistent with an SMC-like dust law in the rest-UV (Weingartner & Draine 2001), though more complex models may be warranted, e.g. to accommodate the 2175 Å; 'bump' present in some galaxies (for a recent review of the dust attenuation law, see Salim & Narayanan 2020).

In this framework, sources at the centre of spherically-symmetric, uniform density dust clouds, are obscured by an optical depth given by

$$\tau_\lambda = \kappa_\lambda N_{\rm d} = \kappa_\lambda \frac{3M_{\rm d}}{4\pi R_{\rm d}^2}, \quad (8)$$

i.e. the characteristic scale $R_{\rm d}$ determines both the dust density and the length of sightlines passing through the dust envelope. To start, we model $R_{\rm d}$ generically as a power law in mass,

$$R_{\rm d} = R_0\left(\frac{M}{10^{10}{\rm M}_\odot}\right)^{\alpha_{\rm d}}\ {\rm kpc}, \quad (9)$$

where R_0 normalizes the scale length at $M_{\rm h} = 10^{10}\,{\rm M}_\odot$ and $\alpha_{\rm d}$ controls the dependence of $R_{\rm d}$ on $M_{\rm h}$. Similar approaches have been taken in previous work, e.g. Somerville et al. (2012) adopt $R_{\rm d} = R_{\rm gas}$, where the radius of the cold gas disc $R_{\rm gas}$ is assumed to be a constant fraction of the stellar scale length. Note that the virial radii of dark matter haloes evolve as $R_{\rm vir} \propto M_{\rm h}^{1/3}(1+z)^{-1}$, while $R_{\rm d} \propto M_{\rm h}^{1/2}$ implies dust column densities that are proportional to $M_{\rm d}/M_{\rm h}$ (see equation 8).

We will show in Section 3 that a more complicated function is likely warranted, at which point we will employ a double power law for $R_{\rm d}$ as well as f_*, i.e.

$$R_{\rm d}(M_{\rm h}) = \frac{R_{\rm d,10}\,\mathcal{D}_{10}}{\left(\frac{M_{\rm h}}{M_{\rm p}}\right)^{-\alpha_{\rm d,lo}} + \left(\frac{M_{\rm h}}{M_{\rm p}}\right)^{-\alpha_{\rm d,hi}}}, \quad (10)$$

where each parameter is analogous to those in equation (3). We discuss our choice of $R_{\rm d}$ parametrization further in Sections 3.3 and 4.2.

We make no effort to model dust *emission* in this work – any link to the IRX–β relation would require modelling of dust temperatures. Imara et al. (2018) present a very similar approach to ours, but focus instead on the implications at longer wavelengths. We expect that our predictions for dust emission would be similar to those of Imara et al. (2018) were we to make the same assumptions for how stellar radiation is reprocessed by dust. As we will discuss in Section 4, UV extinction is a prediction of our model, rather than an input, as is effectively the case for IRX–β-based models.

2.3 Synthetic observations

In order to fairly compare with constraints on the $M_{\rm UV}$–β relation at high z, we 'observe' our model galaxies using the same magnitude definition and photometric filters as in Bouwens et al. (2014, hereafter B14). The filters[3,4] employed vary with redshift as follows:

(i) $z \sim 4$: i_{775}, I_{814}, z_{850}, (Y_{105}), J_{125};
(ii) $z \sim 5$: z_{850}, Y_{105}, (J_{125}), H_{160};
(iii) $z \sim 6$: Y_{105}, (J_{125}), H_{160};
(iv) $z \sim 7$: J_{125}, H_{160}.

B14 employed the filters listed above in the ERS (Windhorst et al. 2011) and CANDELS (Grogin et al. 2011; Koekemoer et al. 2011) fields, but in the deeper XDF (Illingworth et al. 2013), and HUDF09 (Bouwens et al. 2011) fields, filters enclosed in parentheses were not used. The Y_{098} filter was used when available. The UV magnitude in B14 is defined as the geometric mean of the photometric measurements for each galaxy, which we indicate with angular brackets, $\langle M_{\rm UV}\rangle$.

We make no effort in this study to conduct mock surveys and perform sample selection self-consistently, and therefore have no basis on which to use different combinations of filters for objects at the same redshift. For consistency, we adopt the ERS/CANDELS filters for all objects, and only use the Y_{098} filter at $4 \lesssim z \lesssim 6$. We expect this to be a reasonable approach, given that most of the information about dust is in the brightest, reddest objects, which are captured best by the wider field surveys.

Our UV slope predictions are based on two slightly different approaches: (i) an empirically motivated approach, in which β is defined as a power-law fit through the available photometry, as in B14, and (ii) a more theoretical approach, in which β is defined as a power-law fit to the 'true' galaxy spectrum, i.e. the spectrum generated by our forward model (see equations 3–5) at its native resolution, sampled by the C94 windows, which we indicate as $\beta_{\rm c94}$, as in Finkelstein et al. (2012).

As pointed out in Finkelstein et al. (2012), photometric estimates of β can bias estimates of dust attenuation. This is largely for two reasons: (i) The spectrum of galaxies is not a pure power law, so the inferred UV slope can change because filters intersect different parts of the the rest UV spectrum for galaxies at slightly different redshifts, and (ii) absorption lines, particularly at the shortest wavelengths, are unavoidable with photometry, and thus suppress inferred magnitudes despite being unrelated to dust. High-resolution spectra can circumvent these problems, simply by only including 'clean' spectral windows in the fit (e.g. the C94 windows, which exclude wavelength ranges with strong absorption/emission features). Alternatively, one can perform SED fitting on photometric measurements, and estimate β from the best-fitting SED, rather than the photometry. For the duration of this paper, we adopt the purely empirical approach, both to remain consistent with B14 and because it is computationally more efficient. In select cases, we compare to the $\beta_{\rm c94}$ approach and Finkelstein et al. (2012) measurements, generally finding good agreement.

We compare β estimation techniques thoroughly in Appendix A, where we also provide a full listing of the *HST* filters and the redshifts at which they are used (Table A1; shown graphically in Fig. A3) along with the NIRCAM wide and medium filters that lie within the range of the Calzetti et al. (1994) windows. Unless indicated otherwise, these are the filters used for β estimation throughout this paper. We confirm that photometrically estimated UV slopes are biased relative to intrinsic, spectroscopically estimated (in C94 windows)

[3]WFC: http://www.stsci.edu/hst/acs/analysis/throughputs/tables.
[4]WFC3: http://www.stsci.edu/hst/wfc3/ins_performance/throughputs/Throughput_Tables.

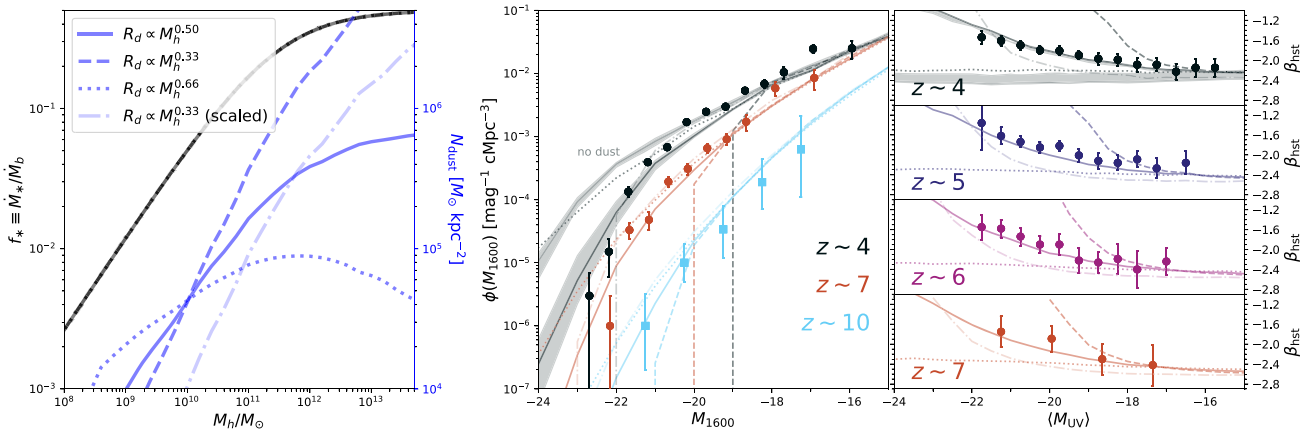

Figure 1. Effects of variations in the relationship between dust scale length, R_d, and halo mass, M_h. Left-hand panel: SFE (black) and dust column density (blue; right axis) as a function of M_h for four different $R_d(M_h)$ models. Note that the dot–dashed curve is systematically shifted, but keeps the same power law as the dashed curve. Right-hand panel: corresponding UVLFs (left-hand panel) and CMDs (right-hand panel) at a series of redshifts. Models in which $R_d \propto M_h^{1/2}$ (solid) are in good agreement with measurements. Shallower slopes $R_d \propto M_h^{1/3}$ result in much too steep M_{UV}–β relations (dashed), even if the normalization of R_d is adjusted to systematically reduce dust reddening (dot–dashed). The shaded regions for $z \sim 4$ models represent mild, $d\kappa/d\lambda = \pm 0.3$ changes in the wavelength dependence of the dust opacity (topmost bands), and differences caused by stellar metallicity for dust-free models (between $Z = 0.001$ and $Z_\odot = 0.02$; horizontal bands). Data shown include UVLFs from Bouwens et al. (2015) ($4 \lesssim z \lesssim 8$) and Oesch et al. (2018) ($z \sim 10$), and M_{UV}–β measurements from B14. A dust-free model is also shown for reference at $z \sim 4$. Note that these models are for illustrative purposes (i.e. they are not the result of fits; see Section 3.4 for MCMC results), and take $f_{*,10} = 0.05$, $M_{*,peak} = 2.8 \times 10^{11}$ M$_\odot$, $\alpha_{*,lo} = 2/3$, $\alpha_{*,hi} = 0$, and $R_{d,10} = 1.4$ kpc.

UV slopes. The effect is generally of order $\Delta\beta \simeq +0.1$ for objects with intrinsic slopes $\beta \sim -2.4$, particularly at $z \gtrsim 6$ (for both *HST* and *JWST*), i.e. measured slopes are biased red. Note that we neglect other possible causes of biases in β estimation, e.g. selection effects (see e.g. Dunlop et al. 2012).

3 RESULTS

3.1 Basic trends

In Fig. 1, we show how the relationship between dust scale length, R_d, and halo mass, M_h, influences UVLFs and the M_{UV}–β relation. We show scenarios spanning the range from $R_d \propto M_h^{1/3}$ to $R_d \propto M_h^{2/3}$.

For illustrative purposes, we fix $f_{*,10} = 0.05$, $M_{*,peak} = 2.8 \times 10^{11}$ M$_\odot$, $\alpha_{*,lo} = 2/3$, $\alpha_{*,hi} = 0$, and $R_{d,10} = 1.4$ kpc. In this case, the production of dust continues even as star formation slows at high mass, resulting in monotonically rising β. The B14 measurements prefer $\alpha_d \simeq 0.5$. The UV colours are extremely sensitive to α_d: The M_{UV}–β relation becomes too steep for $\alpha_d = 1/3$ (dashed), and much too shallow for $\alpha_d = 2/3$ (dotted). In the former case, while reducing the normalization length scale, $R_{d,10}$, can help (equivalent to decrease in dust yield), the shape of the UVLF and M_{UV}–β remain problematic (dash–dotted). The shaded region for $z \sim 4$ models shows how changing the wavelength dependence of the dust opacity (equation 7) between $\kappa_\lambda \propto \lambda^{-1.3}$ and $\kappa_\lambda \propto \lambda^{-0.7}$ affects the UVLFs and CMDs.

The sharp decline in the number counts of galaxies at the bright end of the UVLF is generally interpreted to be in part a sign of dust reddening, but also a byproduct of a decline in the efficiency of star formation in high-mass haloes (e.g. Behroozi, Conroy & Wechsler 2010; Moster et al. 2010). As a result, the assumption of a constant high-mass SFE used in Fig. 1 is likely unreasonable. This complicates the simple power-law R_d model used thus far because the decline in the SFE also causes a decline in dust production, which, if sharp enough, can cause UV colours to start becoming *bluer* as objects become brighter. Observations at $z \lesssim 4$ suggest that β continues to rise monotonically for increasingly bright galaxies (e.g. Lee et al. 2012). In our model, ensuring that β rises monotonically with decreasing M_{UV} requires a change in how R_d scales with halo mass.

We explore the impact of high-mass SFE variations in Fig. 2 for three different high-mass SFE slopes. For a strong decline, $\alpha_{*,hi} = -0.8$, predictions for the bright end of the $z \sim 4$ UVLF are in much better agreement with B14 measurements, though come at the cost of imparting a turnover in M_{UV}–β (dotted lines). This sharp decline in the SFE also has implications for galaxy stellar mass functions (SMFs), which we explore further in Section 3.3.

3.2 Effects of scatter in dust column density

The models shown thus far assume a 1:1 mapping between halo mass and dust scale length. Some scatter in the dust column density, N_d, at fixed M_h still arises due to scatter in the SFR of galaxies (and thus dust production rate), but this is likely overly conservative. To explore the impact of scatter further, we explore scenarios with lognormal scatter in N_d, $\sigma_{\log_{10} N_d}$, at fixed halo mass, deferring a discussion of intrinsic scatter in the M_{UV}–β relation in Section 3.4. In what follows, we also force the dust scale length to be a shallow function of halo mass, $R_d \propto M_h^{1/3}$, for reasons that will become apparent momentarily.

The introduction of N_d scatter has an interesting impact on M_{UV}–β. Consider a faint galaxy, $M_{UV} \sim -17$, with the average amount of dust attenuation, so that $\beta \sim -2.4$. Now, if we subject this galaxy to a strong negative fluctuation in N_d, it will become brighter and bluer, and thus enter an M_{UV} bin occupied (generally) by galaxies that reside in more massive, slightly more rare, haloes. The opposite case of a positive N_d spike will lead our galaxy to migrate in the opposite direction in the M_{UV}–β plane, where it will occupy an M_{UV} bin with galaxies that live in smaller, more common haloes. As a result, there will be a net blueward bias in β at fixed M_{UV}: galaxies scattering toward smaller M_{UV} will always be outnumbered by unscattered objects in the same magnitude bin, while galaxies scattering to brighter M_{UV} will always outnumber the 'typical' galaxy in that bin. Note that this effect is strongest in models with the steepest

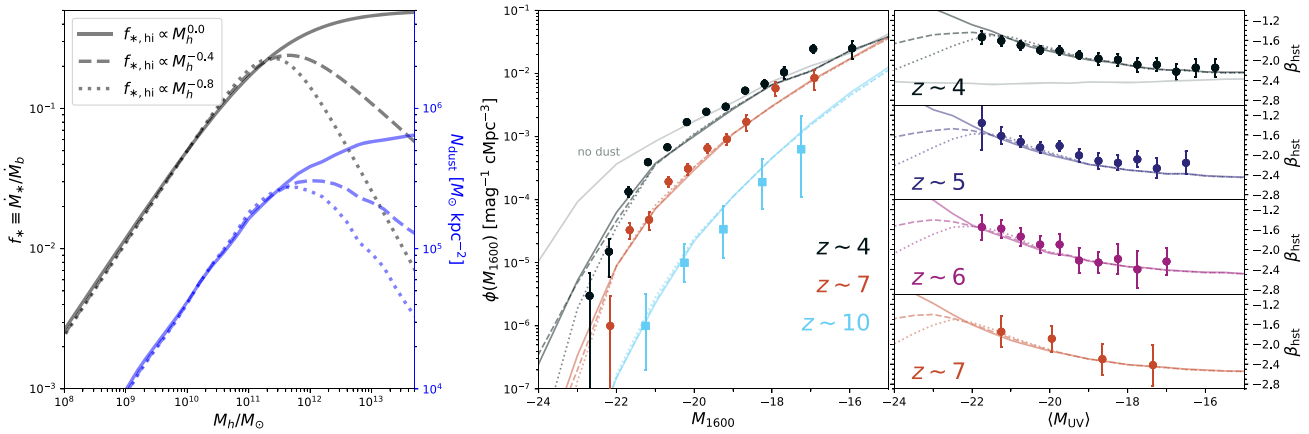

Figure 2. Effects of variations in the efficiency of star formation in high-mass haloes. Same as Fig. 1, except linestyles indicate a change in the slope of the SFE above the peak, $\alpha_{*,\rm hi}$. Because dust production is directly proportional to galaxy SFR in our model, a downturn in the SFE can cause galaxies to become bluer as they grow more massive (dotted lines). Solutions to this issue are discussed in Section 3.3. Each model here takes $f_{*,10} = 0.05$, $M_{*,\rm peak} = 2.8 \times 10^{11}\,{\rm M}_\odot$, $\alpha_{*,\rm lo} = 2/3$, $R_{\rm d,10} = 1.4$ kpc, and $R_{\rm d} \propto M_{\rm h}^{1/2}$.

UVLFs, and could thus be subject to revision if future observations find shallower UVLFs.

We show this effect in Fig. 3. Due to the net bias toward bluer colours, models with more $N_{\rm d}$ scatter can accommodate shallower relationships between $R_{\rm d}$ and $M_{\rm h}$, hence our adoption of the $R_{\rm d} \propto M_{\rm h}^{1/3}$ limit for each model in Fig. 3. Without scatter, $\beta(M_{\rm UV})$ is much too sharp, as shown also in the dashed lines of Fig. 1, but non-zero scatter curbs this behaviour. There is tension between UVLFs and $M_{\rm UV}$–β, which varies as a function of redshift, though this tension can be alleviated by slighly generalizing the $R_{\rm d}$ parametrization and calibrating the model properly via multidimensional fitting, as we describe in the next subsection.

3.3 Model calibration

In order to properly calibrate the model and quantify degeneracies between star formation and dust parameters, we perform a multidimensional Markov chain Monte Carlo (MCMC) fit to the $z \sim 4$, 6, and 8 UVLFs from Bouwens et al. (2015) and $z \sim 4$ and 6 $M_{\rm UV}$–β relations from Bouwens et al. (2014) using EMCEE[5] (Foreman-Mackey et al. 2013).

We note before moving on to the results of this calibration that fitting to the B14 empirical β fits is more efficient computationally than, e.g. fitting to the Finkelstein et al. (2012) UV slopes determined via SED fitting. In order to compare fairly with the Finkelstein et al. (2012) measurements, one needs higher wavelength resolution in order to adequately sample the spectra of objects within the Calzetti et al. (1994) spectral windows, of which there are 10, in contrast to the usual approximately two to five *HST* filters used in the B14 analysis. Our approach scales as the number of wavelengths over which to perform spectral synthesis, making the empirical approach a more efficient option. We compare our best-fitting models to the Finkelstein et al. (2012, hereafter F12) results shortly.

The simplest model we explore has a total of six free parameters: the typical four parameters needed to describe a double power-law SFE ($f_{*,10}$, $M_{*,\rm peak}$, $\alpha_{\rm lo}$, $\alpha_{\rm hi}$), and two parameters for the dust scale length ($R_{\rm d,10}$, $\alpha_{\rm d}$). We do not allow any of these parameters to evolve with cosmic time.

[5]https://emcee.readthedocs.io/en/stable/.

This simple, redshift-independent but halo mass-dependent model for star formation and dust obscuration agrees reasonably well with observations as shown in Fig. 4 (dotted lines). However, due to the tight link between star formation and dust production, the decline in the SFE at high-mass needed to match the steepness of the $z \sim 4$ UVLF has two unfortunate side effects: (i) a turnover in the $M_{\rm UV}$–β relation, and (ii) decline in the bright end of the SMF much steeper than suggested by constraints from Song et al. (2016) and Stefanon et al. (2017) (see dotted lines in bottom panels in Fig. 4).

We employ two strategies to remedy this problem in all that follows. First, we impose a prior requiring β to be a monontonic function of $M_{\rm UV}$ over the range of magnitudes probed by observations (including UVLFs and CMDs), which either eliminates a turnover in $M_{\rm UV}$–β entirely or pushes it to slightly brighter objects, helping to reduce the disagreement between the bright end of the UVLF and SMF. Secondly, we introduce an additional degree of freedom in our parametrization of $R_{\rm d}$, allowing it to be a double power law in $M_{\rm h}$ rather than a single unbroken power law.[6] With this parametrization, as the SFE declines at high mass to match the steepness of the UVLF, the dust scale length can become shallower to ensure that β continues to rise. This solution is amenable to shallower SFE curves at high mass, resulting in better agreement with SMFs at the bright end as well. Along with the standard four parameters for the SFE, this results in a total of nine free parameters, which we calibrate via fitting to the B14 $M_{\rm UV}$–β relation at $z = 4$ and 6, and UVLFs from B15 at $z \sim 4$, 6 and 8. Best-fitting values of the model parameters and their uncertainties are summarized in Table 1, with a subset of the posterior distributions shown in Fig. 6.

In Fig. 4, we show the rest UV calibration of this final model (top panels) and its predictions for the SMF as M_*–β relation (bottom panels) at all $4 \lesssim z \lesssim 10$ (solid lines). The top panels of Fig. 4 are not terribly surprising, as much of the data shown are used in the calibration. Most noteworthy in this context is the evolution in the $M_{\rm UV}$–β relation, which arises despite the assumption that the production rate, opacity, and scale length of dust are constant in

[6]Physically, this could be a signature of morphological changes occurring at high mass. Alternatively, because $R_{\rm d}$ is degenerate with $f_{\rm d}$ and κ, it could be an indicator of changes in how dust is produced and/or destroyed in high-mass galaxies.

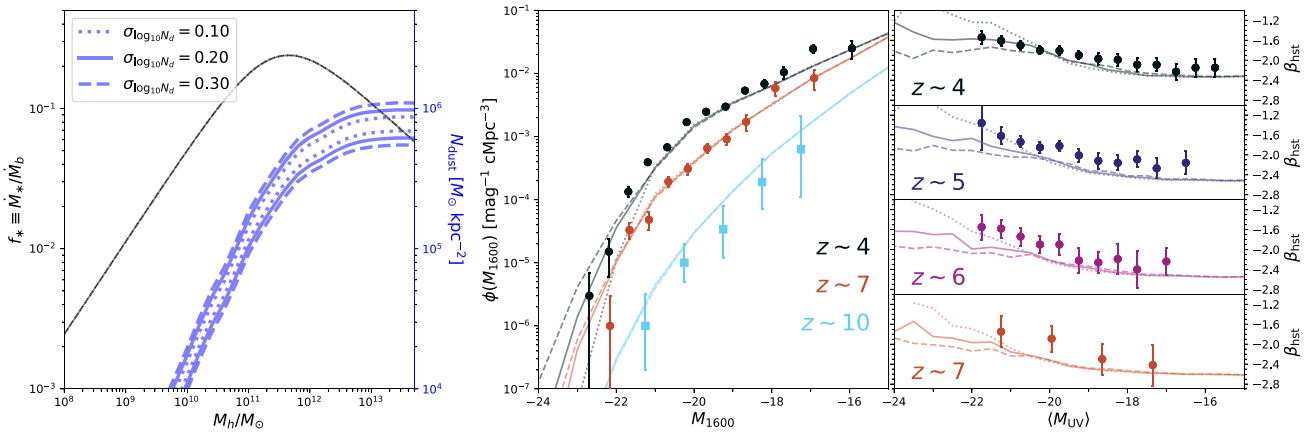

Figure 3. Effects of scatter in dust column density at fixed halo mass. Same as Fig. 1, except linestyles indicate the amount of lognormal scatter, with $\sigma_{\log_{10} N_{\rm d}} = 0.1$, 0.2, and 0.3. We assume an intermediate case for high-mass SFE, $f_{*,\rm hi} \propto M_{\rm h}^{-0.4}$, and a shallow limit for the dust scale length, $R_{\rm d} \propto M_{\rm h}^{1/3}$. The overall effect is a net blueward bias, as objects upscattered into brighter magnitude bins always outnumber the typical object in that bin. Each model here takes $f_{*,\,10} = 0.05$, $M_{*,\,\rm peak} = 2.8 \times 10^{11}$ M$_\odot$, $\alpha_{*,\,\rm lo} = 2/3$, $\alpha_{*,\,\rm hi} = -0.4$, $R_{\rm d,\,10} = 2.4$ kpc, and $R_{\rm d} \propto M_{\rm h}^{1/3}$.

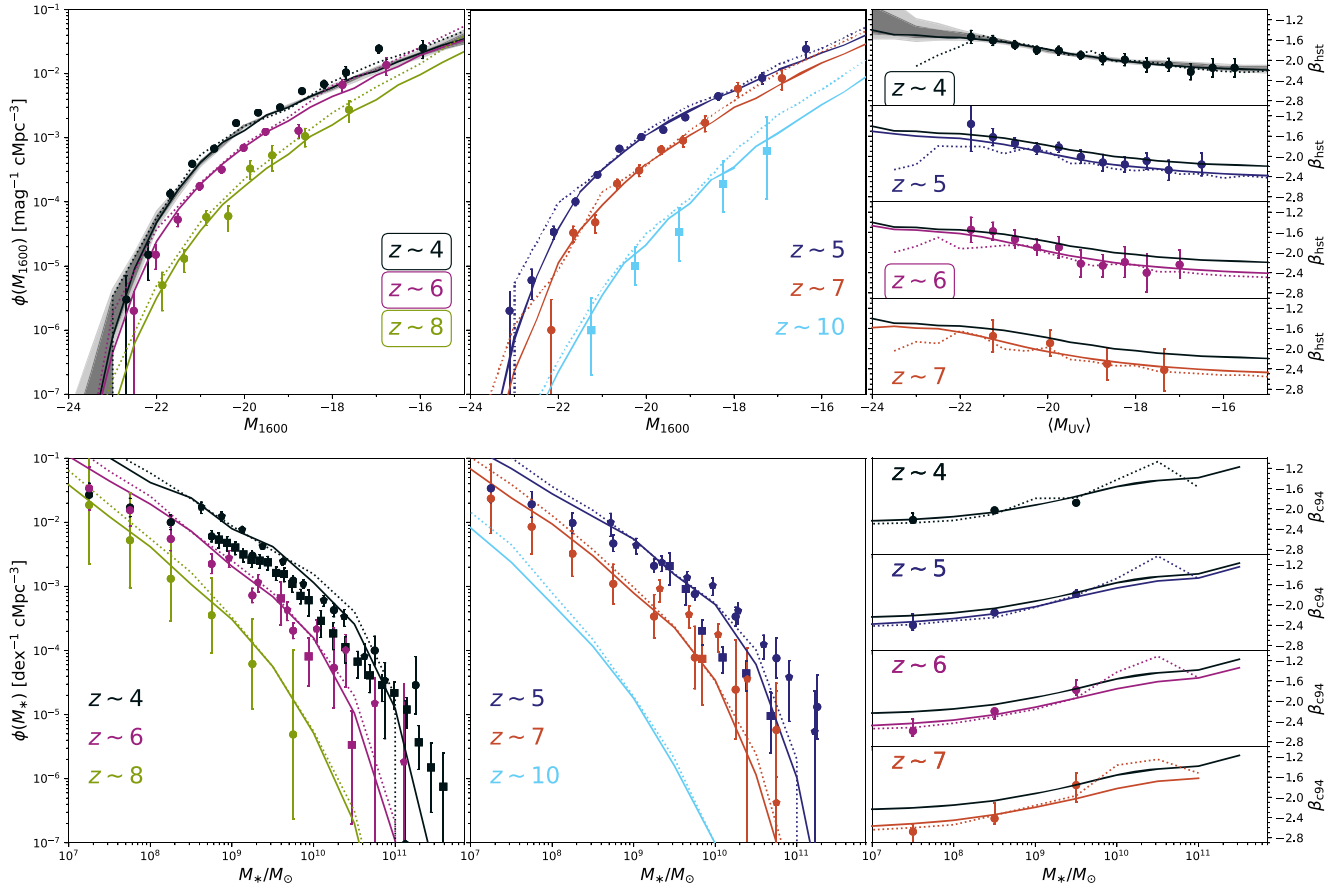

Figure 4. Evolution of UVLF, SMF, and CMDs. Top panels: rest-UV information only, including UVLFs at $z \sim 4$, 6, and 8 (left-hand panels), $z \sim 5$, 7, and 10 (centre panels), and $M_{\rm UV}$–β relation from $4 \leq z \leq 7$. Data shown include UVLFs from Bouwens et al. (2015) ($4 \lesssim z \lesssim 8$) and Oesch et al. (2018) ($z \sim 10$), and $M_{\rm UV}$–β from B14 (right-hand panels). Bottom panels: predictions in terms of stellar masses, rather than $M_{\rm UV}$, including SMFs at $z \sim 4$, 6, and 8 (left-hand panels), $z \sim 5$, 7, and 10 (centre panels), and M_*–β relations at $4 \leq z \leq 7$ (right-hand panels). Data shown include SMFs from Song et al. (2016) (circles), Stefanon et al. (2017) (squares), and Duncan et al. (2014) (pentagons), and M_*–β from Finkelstein et al. (2012). Note that the Stefanon et al. (2017) and Duncan et al. (2014) stellar masses have been shifted by 0.25 dex to convert from a Chabrier into Salpeter IMF, and the $z \sim 4$ UV colours are repeated in each panel. Measurements used in the calibration are indicated with boxes around the corresponding redshifts.

Table 1. Marginalized 68 per cent constraints on the parameters of our fiducial model.

Parameter	Recovery	Prior range
$\log_{10}(f_{*,\,\rm norm})$	$-1.26^{+0.061}_{-0.022}$	$(-3, 0)$
$\log_{10}(M_{*,\,\rm peak}/{\rm M}_\odot)$	$11.16^{+0.165}_{-0.186}$	$(9, 13)$
$\alpha_{*,\,\rm lo}$	$0.80^{+0.100}_{-0.143}$	$(0, 1.5)$
$\alpha_{*,\,\rm hi}$	$-0.53^{+0.244}_{-0.020}$	$(-1, 1.5)$
$R_{\rm d,\,10}$/kpc	$1.12^{+0.096}_{-0.083}$	$(0.01, 10)$
$\alpha_{\rm d,\,lo}$	$0.09^{+0.217}_{-0.068}$	$(0, 2)$
$\alpha_{\rm d,\,hi}$	$0.69^{+0.159}_{-0.023}$	$(-1, 2)$
$\log_{10}(M_{\rm d,\,peak}/{\rm M}_\odot)$	$12.01^{+0.314}_{-1.154}$	$(9, 13)$
$\sigma_{\log_{10} N_d}$	<0.039	$(0, 0.6)$

Notes. The first block of four parameters are those describing the SFE (see equation 3 and Section 2.1), while the next five parameters describe the dust scale length and scatter in dust column density (see equation 10 and Section 2.2). Fits were performed using broad uninformative priors on each parameter, as listed in the final column.

time. This evolution arises due to evolution in the typical stellar age, but also because specific SFRs rise rapidly with redshift, which is a generic prediction of most models (e.g. Dayal et al. 2013; Behroozi, Wechsler & Conroy 2013b). In other words, part of the evolution in $M_{\rm UV}$–β is due to evolution in $M_{\rm UV}$ (at fixed stellar or halo mass) alone, with the rest arising due to the bluer colours typical of increasingly young stellar populations at high redshift (see Section 4.2 for more discussion). From the bottom panels of Fig. 4, it is clear that none of our models predict an SMF as shallow as Song et al. (2016) at the faint end. While jointly fitting UVLFs and SMFs is one potential solution to this problem, we have opted for a pure rest-UV-based approach in order to avoid complicating the calibration procedure further. Setting a prior on ${\rm d}\beta/{\rm d}M_{\rm UV} < 0$ is a simple way to avoid tension between model and data at the bright end, while remaining agnostic about issues at the faint end of the SMF.

In Fig. 5, we show the key ingredients of our model as recovered via MCMC fitting. First, in the top panel, we show the SFE (filled grey contours) compared to a dust-free solution (dotted) and a solution obtained via the M99+B14 approach [7] (dashed). The M99+B14 approach converts intrinsic UV magnitudes into observed UV magnitudes by solving for the extinction $A_{\rm UV}$ required to simultaneously satisfy the link between $A_{\rm UV}$ and β put forth in M99 and the connection between observed $M_{\rm UV}$ and β reported in B14. As expected, the treatment of dust affects both the normalization and shape of the SFE as a function of $M_{\rm h}$, with offsets of a factor of ∼2–3 near the peak. The posterior distribution for the component parameters, as well as the reconstructed SFRD, are included in Appendix B. In the bottom panel, we show the recovered dust scale length with pure power laws included to guide the eye. The departure from a pure power law is subtle – at high-mass, $M_{\rm h} \gtrsim 10^{11}\,{\rm M}_\odot$, our solution roughly tracks the $R_{\rm d} \propto R_{\rm vir} \propto M_{\rm h}^{1/3}$ solution, while at lower mass a steeper relation is preferred. The blue line shows the corresponding dust column density for the best-fitting model only (right axis).

[7]Note that these functions are very similar to those presented in Mirocha et al. (2017) and Mirocha & Furlanetto (2019), except we have replaced the Sheth, Mo & Tormen (2001) mass function with the Tinker et al. (2010) mass function and re-run the fit to be consistent with this work.

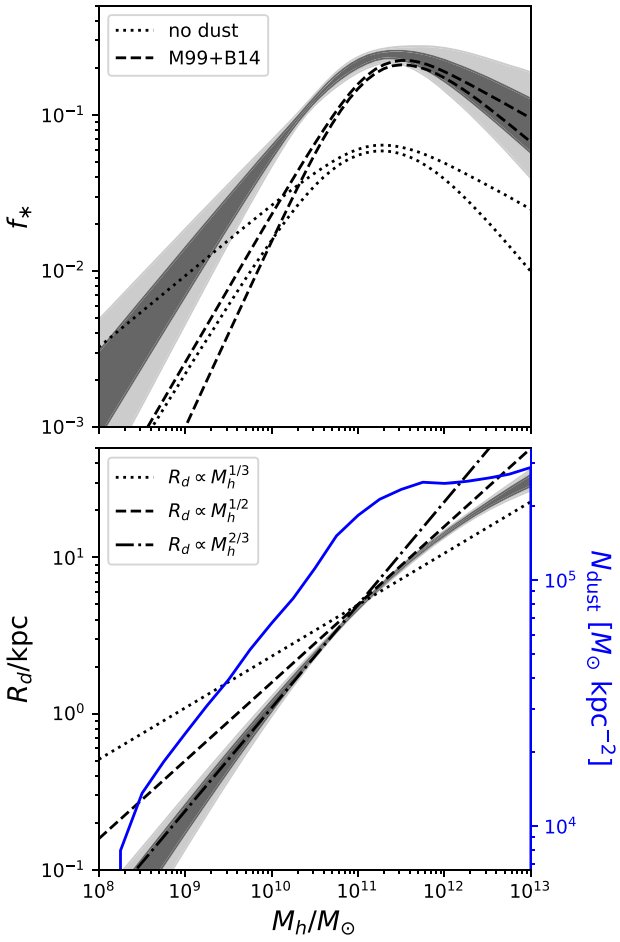

Figure 5. Reconstructed SFE and dust scale length. Top panel: Filled contours indicate 68 and 95 per cent confidence intervals on the SFE using the nine parameter models in this work, compared to a model with no dust calibrated to $z \sim 6$ (dotted), and a model using the common IRX–β-based approach (M99+B14; dashed) calibrated with UVLFs at $z \sim 4$ and 6. Note that our recovered low-mass slope, $\alpha_{*,\,\rm lo} \simeq 0.8$, is consistent both with the predictions of energy-regulated feedback models ($\alpha_{*,\,\rm lo} = 2/3$) and steeper $\alpha_{*,\,\rm lo} = 1$ scenarios (e.g. Tacchella et al. 2018) at the $\sim 2\sigma$ level. Bottom panel: recovered dust scale length (black) and corresponding dust column density for best-fitting model (blue; right axis).

Finally, in Fig. 6, we show a subset of the posterior distribution. From the leftmost panel, we see the degeneracy between components of the double power-law $R_{\rm d}$ model. Solutions favouring a single, unbroken power law would track the dotted line, but clearly such solutions are not preferred by our fits. The maximum likelihohod model has a steep slope at high-mass, $\alpha_{\rm d,\,hi} \sim 0.1$, and shallower slope at low-mass, $\alpha_{\rm d,\,lo} \sim 0.7$, with a change in slope occurring at $10^{11} \lesssim M_{\rm d,\,peak}/{\rm M}_\odot \lesssim 10^{12}$ (panel b). There is of course a mild degeneracy between the dust scale length and high-mass SFE slope (panel c), with $\alpha_{*,\,\rm hi} \simeq -0.5$. The dust scale length and scatter have no significant degeneracy (panel d). Finally, in panel (e), we see that the halo mass of the break in the dust scale length is poorly constrained, and though $M_{\rm d,\,peak} > M_{*,\,\rm peak}$ is preferred, the 68 per cent contours are consistent with occurring at $M_{\rm d,\,peak} = M_{*,\,\rm peak}$. A triangle plot of the SFE parameters is included in Appendix B compared to the results of simpler models published in previous studies.

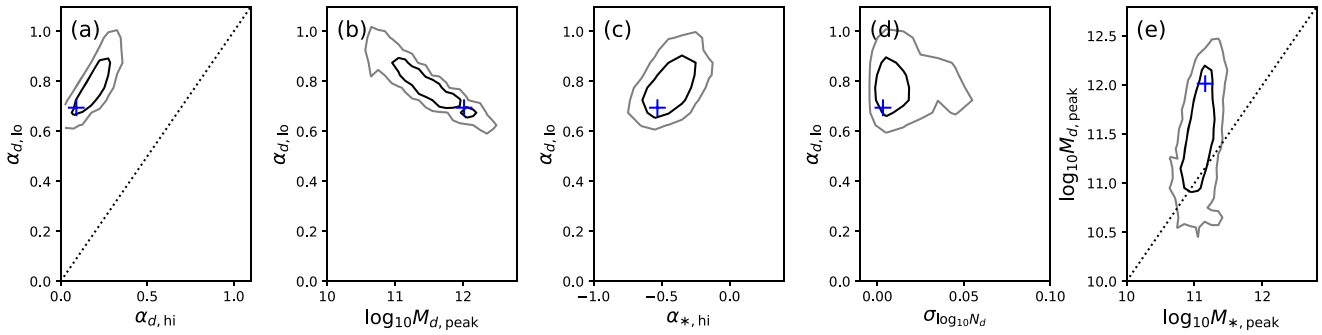

Figure 6. Subset of posterior distributions for our fiducial 9-parameter model. Black contours show 68 and 95 per cent confidence regions, while blue crosses indicate the maximimum likelihood point in each plane. Dotted 1:1 lines are included in first and last panels to assess the degeneracy between each component of R_d, and whether the peak mass in SFE and R_d are consistent. Best-fitting parameters are summarized in Table 1.

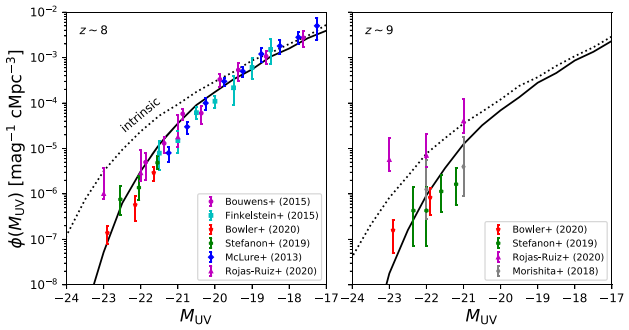

Figure 7. A closer look at $z \sim 8$–9 UVLFs. Dashed lines show our best-fitting model before applying dust reddening, while the solid line is our best-fitting model including dust reddening and dimming. The relatively shallow bright end reported recently by Bowler et al. (2020) is in reasonably good agreement with our predictions.

3.4 Model predictions

The bottom panels of Fig. 4 show our model's predictions for the galaxy SMF and relation between M_* and β. Agreement is reasonably good, though, as is the case for many models (e.g. Tacchella et al. 2018), the slope of the $z \sim 4$–5 SMF at low mass is in considerable tension with observational constraints. Our predictions are closer to the Duncan et al. (2014) SMF measurements than Song et al. (2016), with a steep slope continuing toward even low-mass, as suggested by the measurements of Bhatawdekar et al. (2019). At the bright end, our predictions agree well with the Stefanon et al. (2017) constraints. Predictions for M_*–β are in good agreement with the constraints from Finkelstein et al. (2012).

In Fig. 7, we zoom-in on our predictions for the $z \sim 8$ and 9 UVLFs, given the continued progress in finding bright galaxies at these redshsifts (e.g. McLeod, McLure & Dunlop 2016; Livermore et al. 2018; Morishita et al. 2018; Stefanon et al. 2019; Bowler et al. 2020; Rojas-Ruiz et al. 2020). We show our predictions both for the observed UVLF (solid) and the intrinsic UVLF (dotted), i.e. the UVLF uncorrected for dust. We find that $z \sim 8$–9 galaxies are still reddened considerably at the bright end, by $\sim$1 mag (at, e.g. $M_{\mathrm{UV,\,obs}} = -23$). It is difficult to compare fairly to the Bowler et al. (2020) measurements in full, as our models must match the systematically higher B15 measurements by construction. The agreement at $M_\mathrm{UV} \simeq -23$ may at least suggest that recent measurements do not necessarily indicate reduced dust content in bright galaxies at $z \sim 8$–9.

Because our model is fundamentally anchored to the evolution of dark matter haloes, it is straightforward to make predictions for future UV colour measurements with *JWST*. We show these predictions in Fig. 8, including evolution in β at various M_UV (left-hand panels) and M_* (right-hand panels). We expect the mild trend in $\beta(z; M_\mathrm{UV})$ observed thus far at $4 \lesssim z \lesssim 7$ to continue to higher redshift (top left-hand panel), as has been shown in other empirically calibrated models (e.g. Williams et al. 2018). However, photometric measurements of β generally do not recover the 'true' UV colour evolution (computed using C94 windows; dotted lines). For example, evolution in β as computed with NIRCAM wide filters (dash–dotted) exhibits sharp features at redshifts where two-filter coverage requires excursions outside the C94 range (see Fig. A3 and Table A1). The NIRCAM medium filters probe the underlying evolution more faithfully (dashed), at least at $z \gtrsim 7$, with only a slight redward bias, $\delta\beta \simeq 0.1$, as expected via photometry due to absorption lines in the stellar continuum. Further investigation into the evolution in the *shape* of the M_UV–β and M_*–β relation (bottom panels) seems potentially informative, as our models do not reflect the trends observed in B14 and F12 in detail – in fact, we predict little to no evolution in the β gradients with respect to M_UV (bottom left-hand panel) or M_* (bottom right-hand panel). However, measurement uncertainties are large, so we have not investigated potential sources of this disagreement in detail at this stage.

A key advantage of our approach is that we do not invoke an IRX–β relationship to correct for dust, instead self-consistently solving for the UV luminosity and colours of high-z galaxies with a semi-empirical dust model. As a result, the relationship between A_UV and M_UV is a prediction of our model, rather than an input. We show our recovered $A_\mathrm{UV}(M_\mathrm{UV})$ curves in Fig. 9 compared to the results obtained when assuming a Meurer et al. (1999) relation and the B14 fits to M_UV–β. Our predicted A_UV values are systematically lower than the M99+B14 approach for bright galaxies (dashed lines).

Finally, an interesting question in high-z galaxy evolution is whether or not redshift evolution in M_UV–β and Ly α emitter (LAE) fractions are related to the same underlying phenomenon. Evolution in both colours (e.g. Finkelstein et al. 2012; Bouwens et al. 2014) and Ly α emission (e.g. Shapley et al. 2003; Verhamme et al. 2008; Pentericci et al. 2009; Stark et al. 2010; Hayes et al. 2011; Oyarzún et al. 2017; Yang et al. 2017) has been attributed to evolution in dust, but, to our knowledge, there has been no effort to connect these phenomena explicitly in a physical model. To explore this potential link, we make the simplifying assumption that any object with sufficiently low dust opacity will be an LAE, with the critical opacity left as a free parameter to be determined.

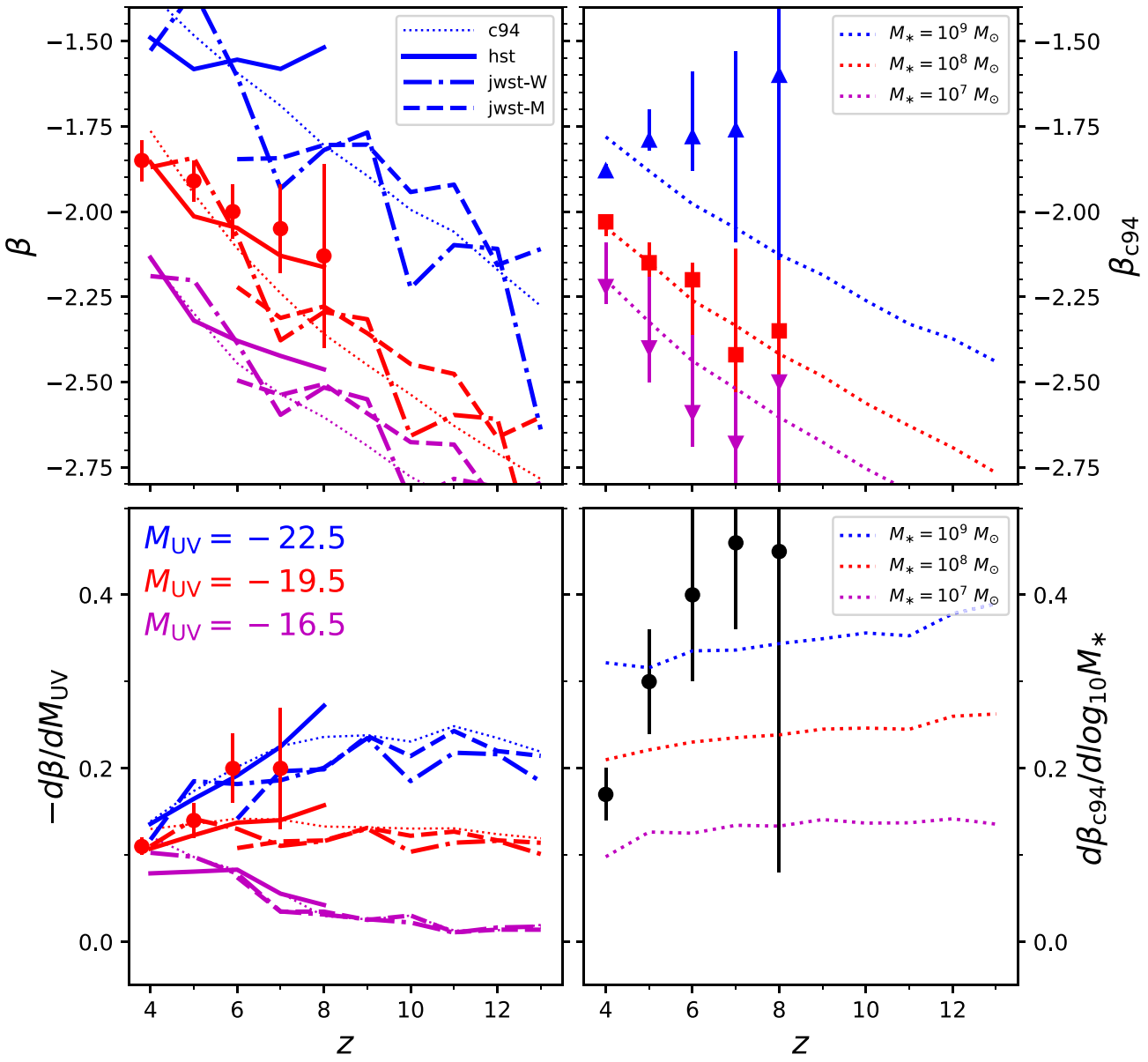

Figure 8. Predictions for redshift evolution in UV colours at fixed M_{UV} (left-hand panels; compared to B14) and fixed M_* (right-hand panels; compared to F12). The top panels show the evolution of the UV slope in three different M_{UV} (left-hand panels) and M_* (right-hand panels) bins, while the bottom panels illustrate evolution in the gradient of β with respect to M_{UV} (left-hand panels) and M_* (right-hand panels). Continued evolution in β at fixed M_{UV} should be detected by *JWST*, with the most accurate recovery enabled by coverage in the NIRCAM medium filters (dashed). Pure wide-band photometry requires sampling the rest-UV spectrum outside the C94 windows, resulting in a bias in β estimates (dash–dotted). Evolution in β at fixed M_* is expected (top right-hand panel), while little evolution in the gradient of the M_{UV}–β and M_*–β relations is expected at fixed M_{UV} (bottom left-hand panel) and M_* (bottom right-hand panel), respectively. Note that the F12 M_*–β measurements are fit over $7.5 \leq \log_{10}M_*/M_\odot \leq 9.5$ (black circles, bottom right-hand panel), whereas we report the slope at three different M_* values (see the legend in the top right-hand panel). Results in the right-hand panels use β as measured in the C94 spectral windows, while β values in the left-hand panels are computed using *HST* photometry (solid), NIRCAM wide (dash–dotted) and medium (dashed) filters, and C94 windows (dotted). Refer to Section 2.3 and Appendix A for more information about filter choices.

In a model with a 1:1 relationship between halo mass and dust column density, there will be a characteristic mass (or M_{UV}) at which galaxies become LAEs (assuming some equivalent width cut) – this mass is set simply by the dust column density for which $\tau_{dust} \sim 1$. Scatter in dust column density has an interesting side effect in this context: The transition from objects that are optically thick to dust at 1600 Å is no longer a sharp function of halo mass and/or M_{UV}. In our framework, scatter in dust column density is degenerate with the dust scale length: An intrinscally shallow $R_d(M_h)$ relationship (and thus steep M_{UV}–β relation) can be counteracted by scatter, and vice versa (see Section 3.2). So, though we cannot self-consistently predict the LAE fraction, we can explore different regions of the posterior distribution to see if the preferred values of $\sigma_{\log_{10}N_d}$ are preferred also by LAE measurements.

As shown in Fig. 10, the fraction of objects with high 1600-Å transmission ($e^{-\tau_{1600}} \geq 0.6 \pm 0.025$) looks remarkably similar to the LAE fraction, x_{LAE}, at $3.5 \lesssim z \lesssim 4.5$ as measured by Stark et al. (2010), at least for $\sigma_{\log_{10}N_d} \gtrsim 0.12$ (solid lines). With less scatter,

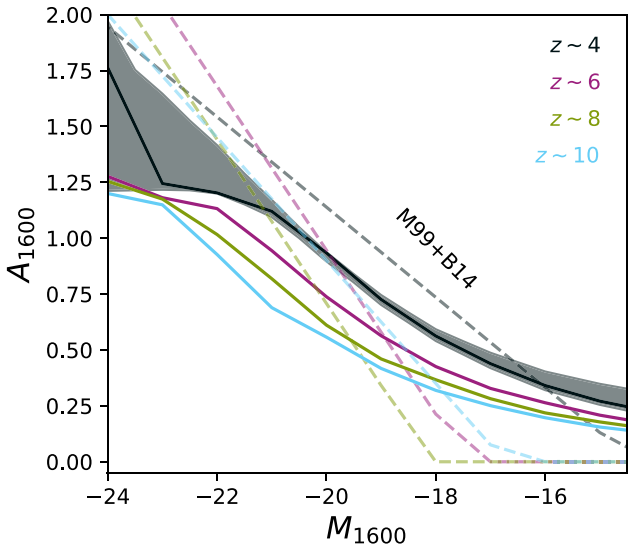

Figure 9. UV extinction as a function of UV magnitude at 1600 Å in the best-fitting model. Dashed lines show the predictions from the M99 IRX–β relation and B14 $M_{\rm UV}$–β relation from $z \sim 4$ to 10, while our model predictions are shown as solid lines. Note that the relationship between $A_{\rm UV}$ and M_* is redshift-*independent*, as we have assumed a constant dust yield per stellar mass and redshift-independent dust scale length, so we do not show it here. Gray shaded region indicates the 68 per cent confidence region for the reconstructed $z \sim 4$ $A_{\rm UV}(M_{\rm UV})$ relation.

the fraction of objects with high UV transmission transitions more abruptly between 0 and 1 (dotted curves). The redshift evolution of $f_{\rm haloes}(> e^{-\tau_{1600}})$ in coarse $M_{\rm UV}$ bins also agrees reasonably well with the redshift evolution measured by Stark et al. (2011). The more recent Kusakabe et al. (2020) measurements are perhaps more accommodating of models with little scatter, $\sigma_{\log_{10} N_{\rm d}} < 0.1$, as preferred in our fits, at least for $M_{\rm UV} \gtrsim -20$. Of course, the caveat here is that the critical value of $e^{-\tau_{1600}} \geq 0.6 \pm 0.025$ was tuned by-eye until the normalization of models and measurements matched. Despite this, it is at least intriguing that $\sigma_{\log_{10} N_{\rm d}}$ values permitted by $M_{\rm UV}$–β measurements generate reasonable LAE populations, and that relatively little scatter, $\sigma_{\log_{10} N_{\rm d}} \simeq 0.1$, is needed to do so. Further exploration of this effect, e.g. how to accommodate larger values of $\sigma_{\log_{10} N_{\rm d}}$, may thus be warranted.

Because the $M_{\rm UV}$–β–LAE connection is largely an issue of scatter in dust column in our framework, we show also our model's predictions for the intrinsic scatter in β at fixed $M_{\rm UV}$, $\Delta\beta$. A larger amount of scatter in $N_{\rm d}$ of course results in more scatter also in β, as we see in the top panel of Fig. 11. In order to ensure that variations in $\sigma_{\log_{10} N_{\rm d}}$ would not worsen agreement with measured UVLFs and CMDs, we draw points directly from the posterior (indicated in the bottom panels of Fig. 11), except for values $\sigma_{\log_{10} N_{\rm d}} > 0.1$, for which there are none. With $\sigma_{\log_{10} N_{\rm d}} \sim 0.12$, our predictions come close to the empirical findings of Rogers et al. (2014), who found evidence of steadily rising intrinsic scatter with increasing galaxy luminosity. The scatter in β in our models is well approximated as a Gaussian, in line with the assumptions of Rogers et al. (2014) and empirical findings of Castellano et al. (2012).

4 DISCUSSION

Our model, though simple, remedies potential inconsistencies in IRX–β approaches while making testable predictions for upcoming observations. In this section, we discuss the implications of the model, and assess the degree to which it is a useful conceptual framework for thinking about dust reddening in high-z galaxies.

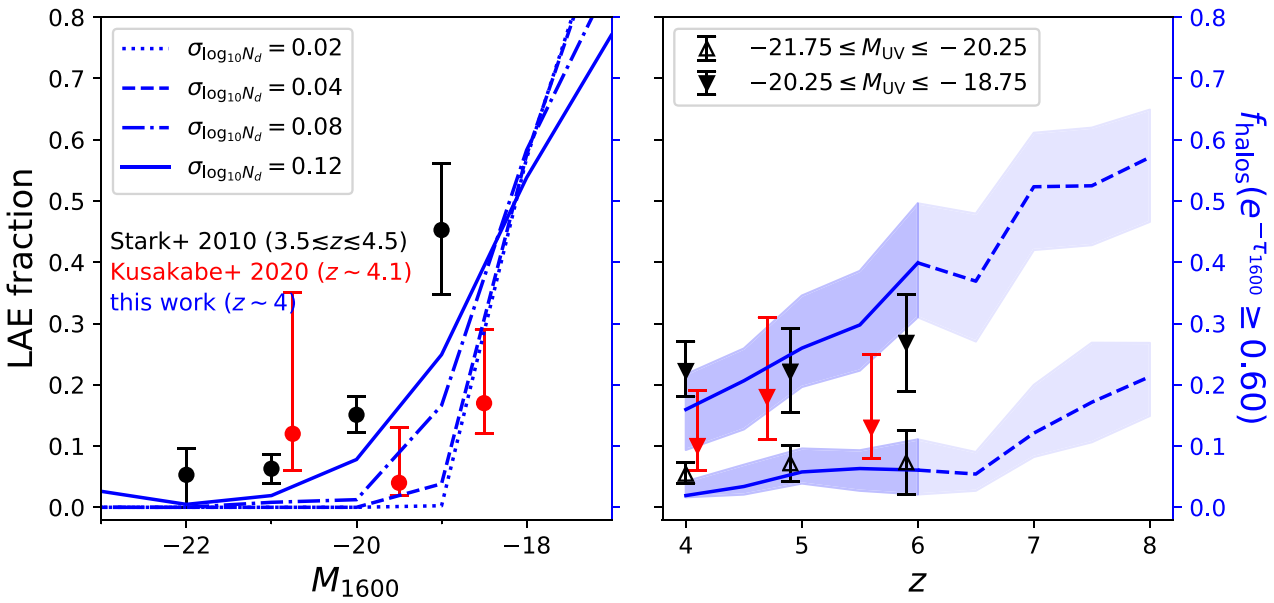

Figure 10. Potential connection between $M_{\rm UV}$–β and LAEs. Left-hand panel: LAE fractions (using equivalent width cut $\geq$ 55 Å) reported by Stark et al. (2010) from $3.5 \lesssim z \lesssim 4.5$ compared to fraction of objects with 1600-Å transmission in excess of 60 $\pm$ 2.5 per cent (right axis; blue) for four different levels of scatter in dust column density. Right-hand panel: evolution in LAE emitter fractions with redshift from Stark, Ellis & Ouchi (2011), separately for bright ($-21.75 \leq M_{\rm UV} \leq -20.25$) and faint ($-20.25 \leq M_{\rm UV} = -18.75$) objects. Shaded areas explore small deviations in the critical transmission ($\pm$0.025). Model predicions at $z \gtrsim 6$ would be subject to the effects of reionization, hence the change in linestyle/shading in the right-hand panel. In both panels, we also show the more recent measurements of Kusakabe et al. (2020) for comparison, which take EW $\geq$ 55 Å (both panels) and $-20.25 \leq M_{\rm UV} \leq -18.75$ (right-hand panel).

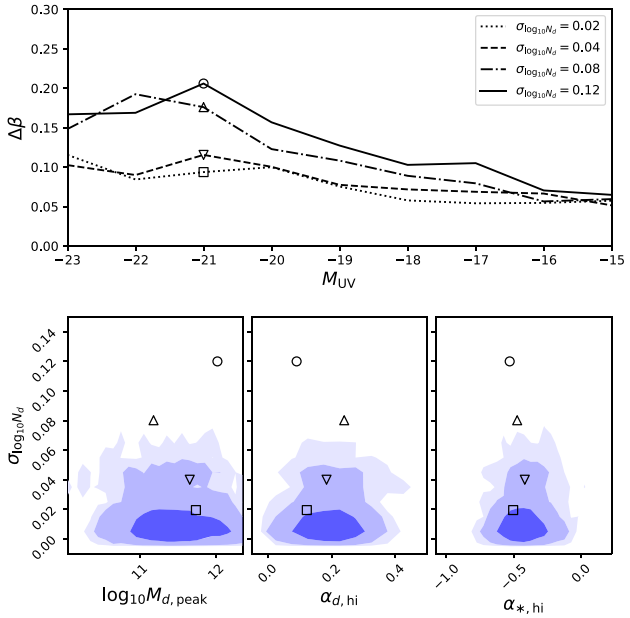

Figure 11. Intrinsic scatter in $\beta(M_{\rm UV})$, $\Delta\beta$, at $z \sim 5$ (top panel) at various points drawn from the posterior (bottom panels)**.** Models shown are the exact same as those shown in Fig. 10, from negligible scatter (dotted) to $\sigma_{\log_{10} N_{\rm d}} \simeq 0.12$ dex (solid). Recall from Fig. 10 that $\sigma_{\log_{10} N_{\rm d}} \gtrsim 0.1$ is generally preferred by LAE measurements. Scatter in SFR at fixed halo mass is responsible for the baseline $0.05 \lesssim \Delta\beta \lesssim 0.1$ scatter, in the limit of $\sigma_{\log_{10} N_{\rm d}} \approx 0$. Bottom panels: location of models in most relevant dimensions of posterior distribution.

4.1 Physical interpretation of the model

Taken at face value, our model predicts the UV luminosity and reddening of galaxies under the assumption that dust within galaxies is distributed uniformly in a sphere and is a source of attenuation only (i.e. no scattering), while star formation is centrally concentrated (see equation 8). Only in this limit does the dust scale length uniquely determine both the dust density and the path length through dust to stellar sources. In reality, the distribution of dust in galaxies is unlikely to be so ideal, so we do not adhere strongly to this geometrical interpretation. Instead, we think of this model as a simple way to connect dust reddening to halo properties. In other words, we caution against overinterpreting the dust scale lengths we infer, and instead emphasize bulk properties like the dust mass, column density, UV luminosity, and colours. Because of this, it is perhaps more reasonable to refer to an *effective* dust scale length or column density, i.e. that which is representative of the reddening over an entire galaxy, composed of many distinct star-forming regions and dust columns.

A key input to our model, aside from the dust scale length described above, is of course the dust production efficiency. We assume a constant dust yield $f_{\rm d} = 0.4$ (Dwek 1998), which fixes the dust-to-metal ratio (DTMR) in each of our model galaxies. Because we also assume a constant metal production efficiency, and instantaneous return model in which $\dot{M}_Z \propto \dot{M}_*$, the dust masses of galaxies in our model will always be a constant fraction of the stellar mass. In reality, the situation is likely more complicated. For example, the DTMR likely scales with halo mass in a non-trivial way depending on the interplay between dust production, destruction, and grain growth in the ISM. However, SAMs including simple treatments of these processes generally predict variations over $7 \lesssim \log_{10}(M_*/{\rm M}_\odot) \lesssim 12$ of only a factor of ∼2–3, and perhaps ∼10 in extreme cases (Popping, Somerville & Galametz 2017). Given this rather shallow modulation of dust content with stellar mass, we do not expect our results to dramatically change upon generalizing the model.

State of the art numerical simulations generally do not include a network of processes specific to the formation and evolution of dust (though see e.g. Li, Narayanan & Davé 2019; McKinnon et al. 2019), and instead link the dust content of grid zones (or gas particles) directly to the local hydrogen column, metallicity, and potentially temperature (Ma et al. 2019; Vogelsberger et al. 2020). Because the birth clouds of stars are generally unresolved, subgrid additions (following, e.g. Charlot & Fall 2000) are often necessary. Though this approach to dust production is still idealized, such simulations can of course investigate the effects of highly inhomogeneous and anisotropic dust distributions, which we cannot. Predictions for the effective dust column density as a function of halo (or stellar) mass could provide important guidance for simpler SAMs like ours. Given that our model allows flexibility in $R_{\rm d}$, rather than tying it to $R_{\rm vir}$, it would be particularly interesting to see the extent to which simulations predict strong mass or redshift evolution in $R_{\rm d}$, $N_{\rm d,\,eff}$, and/or the degree to which dust traces gas mass at high redshift.

4.2 Evolving dust?

Given that our model adopts a fixed dust yield and a time-independent dust opacity and scale length, all evolution in UV colours (at fixed $M_{\rm UV}$) arises from evolution in components of the model unrelated to dust. As a result, any observed evolution in β at fixed $M_{\rm UV}$ need not be due to evolution in the properties of dust, at least in the limit in which the SFE and $R_{\rm d}$ are universal. Evolution in UV colours occurs naturally in our model due to two independent effects: (i) Objects of fixed $M_{\rm UV}$ are hosted by smaller haloes at higher redshift and thus have less dust than objects of the same $M_{\rm UV}$ at lower redshift (sSFR grows rapidly with z; see Section 2.1), and (ii) the mean stellar age is simply younger at high redshift.

There is a significant body of work suggesting the need for evolution in the properties of dust with cosmic time, some of which are largely empirical[8] (Reddy et al. 2010; Capak et al. 2015), while others invoke evolving dust properties to reconcile galaxy formation models with observational constraints (Guo & White 2009; Somerville et al. 2012; Qiu et al. 2019; Yung et al. 2019). The latter admit to being ad hoc: Time-independent dust properties result in dramatic underprediction of bright galaxies at high redshifts (e.g. Somerville et al. 2012; Yung et al. 2019).

The results of numerical simulations are varied. Illustris prefers redshift evolution in the dust opacity (Vogelsberger et al. 2020), while the FIRE and CROC simulations introduce no such evolution (Khakhaleva-Li & Gnedin 2016; Ma et al. 2019), but still obtain luminosity functions that are roughly in agreement with observations. Ma et al. (2019) concluded that their simulations are consistent with evolution in dust properties, though such evolution would have to be geometrical in nature given their assumption of a constant DTMR.

In this work, we find that evolution in UV colours and UV extinction (at fixed $M_{\rm UV}$), which are often interpreted as signatures of evolving dust, arise naturally even for scenarios in which no dust (or even metallicity) evolution is allowed. The general need for evolving dust in some theoretical models is likely due to the shrinking sizes of

[8]These inferences still require assumptions about, e.g. the temperature of dust, evolution of which could masquerade as evolution in dust content. Furthermore, dust may be multiphase, complicating procedures based on a single temperature (Liang et al. 2019).

high-z galaxies at fixed mass. For example, many SAMs connect the dust scale length to the scale length of galaxy discs (Somerville et al. 2012; Qiu et al. 2019), which shrink rapidly in concert with the virial radii of their host dark matter haloes, $R_{\rm vir} \propto M_{\rm h}^{1/3}(1+z)^{-1}$. We are able to reproduce this effect: For example, if we force the relationship between dust scale length and halo masses to look like the relation between halo virial radii and mass ($R_{\rm d} \propto M_{\rm h}^{1/3}$, neglecting the redshift dependence), our $M_{\rm UV}$–β relations and UVLFs are much too steep (see dashed and dash–dotted curves in Fig. 1). However, because our model can adjust the scale length as a free parameter, and appeal to scatter in dust columns densities to shallow out intrinsically steep $M_{\rm UV}$–β relations, we are able to avoid redshift-dependent dust optical depths. The implicit prediction here is that if the scale length of dust does not perfectly track the scale length of gas in galaxy discs, the intrinsic properties of dust in galaxies need not evolve with time. This prediction is in principle testable, now that spatially resolved dust continuum maps can be obtained for high-z galaxies using ALMA (e.g. Gullberg et al. 2018).

We note before moving on that there are other potential sources of mild evolution in the UV colours of galaxies that we have not included, such as metallicity evolution (e.g. Wilkins et al. 2016), which will of course affect the intrinsic spectrum of galaxies and potentially amplify the effects of scatter in dust column density. Furthermore, below $z \sim 6$, asymptotic giant branch stars become a relevant source of dust production that we have effectively neglected by assuming an instantaneous dust return model. We defer a detailed discussion of these effects to future work.

4.3 Scatter in dust column

In simulations and some SAMs, scatter in the dust column density will inevitably arise due to viewing angle effects (Yung et al. 2019). It is unclear if our treatment of $\sigma_{\log_{10} N_{\rm d}}$ as a free parameter is acting to mimic these effects. Our model predicts scatter in β (at fixed $M_{\rm UV}$) at the level of $\Delta\beta \simeq 0.2$ at $M_{\rm UV} \gtrsim -19$ (see Fig. 11). This is comparable, though slightly lower, than the observed Rogers et al. (2014) trend and predictions from Yung et al. (2019). Given the importance of scatter in potentially setting both the shape of $M_{\rm UV}$–β and LAE fractions, future constraints on $\Delta\beta$ – including its distribution function – may serve as an important discriminator between models, and help determine if the $M_{\rm UV}$–β/LAE connection explored here is at work in real galaxies.

4.4 Connection with LAE fraction

We find that the $M_{\rm UV}$–β $M_{\rm UV}$–$x_{\rm LAE}$ relations can be explained by scatter in dust column density at fixed halo mass, $\sigma_{\log_{10} N_{\rm d}} \gtrsim 0.1$, assuming that 1600-Å transmission is a reliable predictor of whether or not galaxies are strong LAEs. In this case, 'strong' means equivalent widths of ≥ 55 Å, so as to compare directly with the measurements of Stark et al. (2010).

This binary model for the LAE fraction is much simpler than others in the literature. A common approach is to model the intrinsic Ly α equivalent width (EW), which requires assumptions about the SFR, escape fraction, and kinematics of galaxies, and then apply an EW cut in order to compare with observations (e.g. Dayal, Ferrara & Gallerani 2008). Our results suggest that, to zeroth order, the abundance of LAEs could be set by the dust column density PDF of galaxies, and that this PDF is also responsible (in part) for the shape of the $M_{\rm UV}$–β relation.

Though the potential connection between patchiness and Ly α emission, we explore in Fig. 10 has been considered before (e.g. Neufeld 1991; Hansen & Oh 2006; Finkelstein et al. 2008, 2009), to our knowledge, there have been no attempts to draw an explicit connection between the $M_{\rm UV}$–β and $M_{\rm UV}$–$x_{\rm LAE}$ in a forward model, despite the fact that both are often attributed to dust. The viability of scatter in $N_{\rm d}(M_{\rm h})$ as the mechanism responsible for each trend is, in principle, testable, perhaps most clearly via measurements of the intrinsic scatter in $M_{\rm UV}$–β (see Section 4.3).

4.5 Colour selection criteria

Given that the scatter we invoke is lognormal, it is possible that galaxies in our model will experience fluctuations in dust column substantial enough to migrate outside the typical high-z galaxy colour selection windows. However, for the small values $\sigma_{\log_{10} N_{\rm d}} \lesssim 0.1$ preferred by our fits, this is a small effect. For example, using the B14 colour selection criteria, and assuming the Madau (1995) model for IGM absorption, we find that only ∼1–3 per cent of objects would be falsely excluded from $z \gtrsim 4$ samples due to the effects of scatter. Falsely excluded galaxies are generally those with $\dot{M}_* \gtrsim 10^2$ M$_\odot$ yr^{-1}, with rest UV colours in *HST* bands of $i_{775} - J_{125} \gtrsim 1$ and $B_{435} - V_{606} \sim 3$, and can thus be confused with red elliptical galaxies at $z \sim 1$ (Coleman, Wu & Weedman 1980). This region of the colour–colour space is sparsely populated observationally (see e.g. fig. 3 of Bouwens et al. 2015), but future detections in this space may warrant further attention. Most of these objects ($\gtrsim 90$ per cent) have $M_{\rm UV} \lesssim -20$ and so are detectable for *HST*.

4.6 Implications for *JWST*

Our model will be readily testable with constraints on high-z galaxy counts and colours from *JWST*. The most noticeable tension in the model is at the low-mass end of the $4 \lesssim z \lesssim 6$ SMF, where current measurements diverge (e.g. Duncan et al. 2014; Song et al. 2016; Stefanon et al. 2017), We note that, while our predictions are consistent with steep low-mass slopes (e.g. Duncan et al. 2014; Bhatawdekar et al. 2019), such steep slopes are known to cause tension with Local Group constraints (Graus et al. 2016). Constraints on the SMF should improve considerably with *JWST*, given the substantial expansion of coverage in the infrared, which will much more fully probe the rest-optical emission of high-z galaxies, and thus potentially resolve the current disagreement. Improved colour constraints, particularly for massive galaxies, would also improve the model calibration. However, given the small field of view of NIRCAM, such constraints may only be possible for shallow, wide-area surveys.

The redshift evolution in $\beta(M_{\rm UV})$ predicted by our model is also readily testable with *JWST*. In principle, using the medium NIRCAM filters, rest-UV colours can be measured out to $z \sim 15$, provided there are galaxies bright enough to detect at such high redshifts. We find only minimal evolution in the *shape* of $\beta(M_{\rm UV})$ and $\beta(M_*)$. The medium filters are key to all future $z \gtrsim 8$ colour constraints, as the wide filters are cannot cleanly isolate the rest UV continuum generally used for estimating β (see Fig. A3).

Finally, testing the hypothesis that the scatter in dust column density drives both $M_{\rm UV}$–β and $M_{\rm UV}$–$x_{\rm LAE}$ will require improved constraints on the intrinsic scatter in $\beta(M_{\rm UV})$. We compare favourably, to the Rogers et al. (2014) estimates of scatter in $\beta(M_{\rm UV})$, in that the scatter rises monotonically for increasingly bright galaxies, and for $\sigma_{\log_{10} N_{\rm d}} \sim 0.1$ yields $\Delta\beta \simeq 0.2$ at $M_{\rm UV} \sim -21$, in agreement with Rogers et al. (2014, though slightly lower). Larger values of $\sigma_{\log_{10} N_{\rm d}}$, likely closer to $\sigma_{\log_{10} N_{\rm d}} \approx 0.2$, would improve agreement, though our fits clearly prefer $\sigma_{\log_{10} N_{\rm d}} \lesssim 0.1$. This could simply be

a limitation of the model – perhaps alternate parametrizations or additional flexibility would permit larger $\sigma_{\log_{10} N_d}$ values. We leave this as an avenue to pursue in future work.

4.7 Implications for cosmic SFRD and reionization

Assumptions about dust necessarily impact the inferred cosmic SFRD, and thus predictions for reionization. Naively, one might expect dust-free models to provide a floor in predictions for the SFRD, since any increase in the dust content of galaxies will require enhancements to the SFE in order to preserve agreement with observed UVLFs. However, the presence of dust can also modulate the inferred *shape* of the SFE (see Fig. 5). As a result, the introduction of dust will shift the SFRD upward at late times, $z \lesssim 10$, but potentially result in little overall change in the SFRD at earlier times, when fainter haloes (which form stars less efficiently) become the dominant source population. Indeed, we find a steeper slope $f_* \propto M_h^{2/3}$ (or steeper) here compared to the $f_* \propto M_h^{1/2}$ scaling reported in Mirocha et al. (2017), the latter of which neglected dust (see Appendix B). As a result, systematic uncertainties in modeling dust – at least in the framework presented here – are unlikely to cause a substantially earlier start to reionization or reheating. Any evidence for efficient star formation at $z \gtrsim 10$ from, e.g. kinetic Sunyaev–Z'eldovich constraints from the CMB (e.g. Miranda et al. 2017) or global 21-cm signal measurements (Bowman et al. 2018), is most likely indicative of star formation in haloes below the atomic cooling threshold (Mirocha & Furlanetto 2019; Mebane, Mirocha & Furlanetto 2020). This statement may be subject to change in models in which the SFE of galaxies grows with redshift or changes shape in non-trivial ways, a possibility that we defer to future work.

4.8 Implications for IRX–β

A common approach to dust-correcting rest-UV measurement of high-z galaxies is to invoke empirical correlations between infrared excesses and β (Meurer et al. 1999). The IR excess can be related to UV attenuation under the assumption of a known intrinsic UV slope and dust opacity, making it possible to convert the observed M_{UV} into an intrinsic UV magnitude, and thus SFR. The standard M99 approach assumes $\beta_0 = -2.23$, appropriate for constant star formation in the STARBURST99 models. Our models use the BPASS models (Eldridge & Stanway 2009) and assume galaxy SFHs are rising rapidly and have scatter, which modulates the input β_0 for each galaxy. The effects of breaking the assumptions made in M99 has been pointed out previously also by other authors (e.g. Wilkins et al. 2013).

Our forward model generally predicts *less* UV attenuation at fixed M_{UV} than the M99 relation, at least at the bright end, $M_{UV} \lesssim -18$, consistent with other studies (see e.g. Mancini et al. 2016). This is likely a byproduct of our joint inference approach, as the M_{UV}–β relation and UVLFs have competing requirements (see Section 3.1 and Fig. 2). For the faint, $M_{UV} \gtrsim -18$ dust-poor galaxies, our model predicts more attenuation than M99, which makes sense given that our model galaxies are intrinsically bluer than $\beta_{in} = -2.23$.

5 CONCLUSIONS

We have presented a simple, but self-consistent model for dust reddening in high-z galaxies that does not require assumptions about the IRX–β relationship. Instead, we flexibly parametrize the dust column density of galaxies as a function of halo mass, and link dust production directly to star formation. Upon calibrating the model parameters via joint fitting of high-z UV luminosity function and colour–magnitude relation constraints, we find the following:

(i) Models without redshift-dependent dust properties still predict evolution in M_{UV}–β, given that stellar ages are declining and specific SFRs are rising with redshift. In other words, much of the evolution in M_{UV}–β reflects evolution in the typical halo (or stellar) mass of galaxies in our model, and thus their integrated dust production. This result is conservative given our neglect of metallicity evolution, which is also expected to result in UV colour evolution at fixed stellar mass (see Section 4.2; Figs 4 and 8).

(ii) This lack of evolution is at odds with other models in the literature, which require ad hoc redshift evolution in the dust opacity in order to prevent excessive reddening in high-z galaxies. This need may be real: If the dust scale length is related to the scale length of galaxy discs, and thus dark matter halo virial radii, it will contract rapidly in both halo mass and redshift and cause reddening to increase as well. Our model assumes no redshift evolution in the dust scale length (at fixed M_h) and so avoids this effect. Observationally, constraining the effective dust scale length may be difficult, so guidance from simulations may offer important insights in this context (see Section 4.1).

(iii) Scatter in the relationship between dust column density and halo mass can help accommodate dust scale lengths that track halo virial radii (see Fig. 3). Furthermore, for values of the lognormal scatter in effective dust column density at fixed haloes mass, $\sigma_{\log_{10} N_d} \simeq 0.1$, the evolution in the abundance of galaxies with 1600 Å ≥ 0.6 resembles the evolution in the LAE population in the range $3 \lesssim z \lesssim 6$ (see Fig. 10). This could be an indicator that the shape of M_{UV}–β and M_{UV}–x_{LAE} are driven by the same phenomenon.

(iv) Measurements of the intrinsic scatter, $\Delta\beta$, provide important constraints on this aspect of the model. An increased scatter in the dust column density increases the scatter in β as well, with scatter continuing to grow in even brighter objects (see Fig. 11).

(v) UV colours are expected to continue to evolve smoothly with redshift at fixed M_{UV}, and are, in principle, measurable with *JWST* out to $z \sim 15$ at all $M_{UV} \lesssim -17$ (see Fig. 8). We predict little to no redshift evolution in the shape of the M_{UV}–β and M_{UV}–M_* relations.

The aforementioned results depend on the assumption of a time-independent SFE, which is common in the recent literature, but also on a universal $R_d(M_h)$ relation. Both assumptions are subject to change in simple models, which predict evolution in the SFE and $R_d(M_h)$, if indeed $R_d \propto R_{vir}$. We explore the consequences of these assumptions in a forthcoming paper (Mirocha 2020).

SOFTWARE

NUMPY (Van Der Walt, Colbert & Varoquaux 2011), SCIPY (Virtanen et al. 2020), MATPLOTLIB (Hunter 2007), H5PY,[9] and MPI4PY (Dalcın, Paz & Storti 2005).

ACKNOWLEDGEMENTS

JM acknowledges stimulating conversations with Chris Willott, Nissim Kanekar, Louis Abramson, Adrian Liu, Alan Heavens, James Rhoads, Tracy Webb, and Steve Finkelstein, and the anonymous referee for comments that helped improve this paper. JM also acknowledges support through a CITA National Fellowship. CM acknowledges support through the NASA Hubble Fellowship grant HST-HF2-51413.001-A awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for

[9] http://www.h5py.org/.

Research in Astronomy, Inc., for NASA, under contract NAS5-26555. Computations were made on the supercomputers Cedar (from Simon Fraser University and managed by Compute Canada) and Mammouth (from the Université de Sherbrooke and managed by Calcul Québec and Compute Canada). The operation of these supercomputers is funded by the Canada Foundation for Innovation (CFI), the ministère de l'Économie, de la science et de l'innovation du Québec (MESI), and the Fonds de recherche du Québec-Nature et technologies (FRQ-NT).

DATA AVAILABILITY

The data underlying this paper are available upon request, but can also be re-generated from scratch using the publicly available ARES code.

REFERENCES

Behroozi P. S., Conroy C., Wechsler R. H., 2010, ApJ, 717, 379
Behroozi P. S., Wechsler R. H., Conroy C., 2013a, ApJ, 762, L31
Behroozi P. S., Wechsler R. H., Conroy C., 2013b, ApJ, 770, 57
Behroozi P. et al., 2019, MNRAS, 488, 3143
Bhatawdekar R. et al., 2019, MNRAS, 486, 3805
Bouwens R. J. et al., 2009, ApJ, 705, 936
Bouwens R. J. et al., 2011, ApJ, 737, 90
Bouwens R. J. et al., 2014, ApJ, 793, 115 (B14)
Bouwens R. J. et al., 2015, ApJ, 803, 34
Bowler R. A. A., Jarvis M. J., Dunlop J. S., McLure R. J., McLeod D. J., Adams N. J., Milvang-Jensen B., McCracken H. J., 2020, MNRAS, 493, 2059
Bowman J. D. et al., 2018, Nature, 555, 67
Brown R. L., Mathews W. G., 1970, ApJ, 160, 939
Calzetti D., Kinney A. L., Storchi-Bergmann T., 1994, ApJ, 429, 582 (C94)
Capak P. L. et al., 2015, Nature, 522, 455
Casey C. M. et al., 2014, ApJ, 796, 95
Castellano M. et al., 2012, A&A, 540, A39
Charlot S., Fall S. M., 2000, ApJ, 539, 718
Coleman G. D., Wu C. C., Weedman D. W., 1980, ApJS, 43, 393
Dalc ın L., Paz R., Storti M., 2005, J. Parallel Distrib. Comput., 65, 1108
Dayal P., Ferrara A., Gallerani S., 2008, MNRAS, 389, 1683
Dayal P. et al., 2013, MNRAS, 434, 1486
Dekel A. et al., 2013, MNRAS, 435, 999
Duncan K. et al., 2014, MNRAS, 444, 2960
Dunlop J. S. et al., 2012, MNRAS, 420, 901
Dunlop J. S. et al., 2013, MNRAS, 432, 3520
Dwek E., 1998, ApJ, 501, 643
Dwek E., Galliano F., Jones A. P., 2007, ApJ, 662, 927
Eldridge J. J., Stanway E. R., 2009, MNRAS, 400, 1019
Ferland G. J., 1980, PASP, 92, 596
Fernandez E. R., Komatsu E., 2006, ApJ, 646, 703
Finkelstein S. L. et al., 2008, ApJ, 678, 655
Finkelstein S. L. et al., 2009, ApJ, 691, 465
Finkelstein S. L. et al., 2012, ApJ, 756, 164 (F12)
Finkelstein S. L. et al., 2015, ApJ, 810, 71
Finlator K., Oppenheimer B. D., Davé R., 2011, MNRAS, 410, 1703
Foreman-Mackey D. et al., 2013, PASP, 125, 306
Furlanetto S. R. et al., 2017, MNRAS, 472, 1576
Graus A. S. et al., 2016, MNRAS, 456, 477
Grogin N. A. et al., 2011, ApJS, 197, 35
Gullberg B. et al., 2018, ApJ, 859, 12
Guo Q., White S. D. M., 2009, MNRAS, 396, 39
Hansen M., Oh S. P., 2006, MNRAS, 367, 979
Hayes M. et al., 2011, ApJ, 730, 8
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Illingworth G. D. et al., 2013, ApJS, 209, 6
Imara N. et al., 2018, ApJ, 854, 36
Khakhaleva-Li Z., Gnedin N. Y., 2016, ApJ, 820, 133
Koekemoer A. M. et al., 2011, ApJS, 197, 36
Kusakabe H. et al., 2020, A&A, 638, A12
Lee K.-S. et al., 2012, ApJ, 758, L31
Leitherer C. et al., 1999, ApJS, 123, 3
Li Q., Narayanan D., Davé R., 2019, MNRAS, 490, 1425
Liang L. et al., 2019, MNRAS, 489, 1397
Livermore R. C. et al., 2018, ApJ, 861, L17
McBride J., Fakhouri O., Ma C.-P., 2009, MNRAS, 398, 1858
McKinnon R., Kannan R., Vogelsberger M., O'Neil S., Torrey P., Li H., 2019, preprint (arXiv:1912.02825)
McLeod D. J., McLure R. J., Dunlop J. S., 2016, MNRAS, 459, 3812
Ma X. et al., 2019, MNRAS, 487, 1844
Madau P., 1995, ApJ, 441, 18
Madau P., Dickinson M., 2014, ARA&A, 52, 415
Mancini M. et al., 2016, MNRAS, 462, 3130
Mashian N., Oesch P. A., Loeb A., 2016, MNRAS, 455, 2101
Mason C. A., Trenti M., Treu T., 2015, ApJ, 813, 21
Mebane R. H., Mirocha J., Furlanetto S. R., 2020, MNRAS, 493, 1217
Meurer G. R., Heckman T. M., Calzetti D., 1999, ApJ, 521, 64 (M99)
Miranda V. et al., 2017, MNRAS, 467, 4050
Mirocha J., 2020, MNRAS, preprint (arXiv:2008.04322)
Mirocha J., Furlanetto S. R., 2019, MNRAS, 483, 1980
Mirocha J., Furlanetto S. R., Sun G., 2017, MNRAS, 464, 1365
Morishita T. et al., 2018, ApJ, 867, 150
Moster B. P. et al., 2010, ApJ, 710, 903
Murray S. G., Power C., Robotham A. S. G., 2013, Astron. Comput., 3, 23
Narayanan D. et al., 2018, ApJ, 869, 70
Neufeld D. A., 1991, ApJ, 370, L85
Oesch P. A. et al., 2018, ApJ, 855, 105
Oke J. B., Gunn J. E., 1983, ApJ, 266, 713
Overzier R. A. et al., 2011, ApJ, 726, L7
Oyarzún G. A. et al., 2017, ApJ, 843, 133
Pentericci L. et al., 2009, A&A, 494, 553
Planck Collaboration VI, 2018, preprint (arXiv:1807.06209)
Popping G., Somerville R. S., Galametz M., 2017, MNRAS, 471, 3152
Qiu Y. et al., 2019, MNRAS, 489, 1357
Reddy N. A. et al., 2010, ApJ, 712, 1070
Reddy N. A. et al., 2018, ApJ, 853, 56
Ren K., Trenti M., Mutch S. J., 2018, ApJ, 856, 81
Rogers A. B. et al., 2014, MNRAS, 440, 3714
Rojas-Ruiz S. et al., 2020, ApJ, 891, 146
Salim S., Boquien M., 2019, ApJ, 872, 23
Salim S., Narayanan D., 2020, preprint (arXiv:2001.03181)
Schulz S. et al., 2020, MNRAS, 497, 4773
Shapley A. E. et al., 2003, ApJ, 588, 65
Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1
Somerville R. S. et al., 2012, MNRAS, 423, 1992
Song M. et al., 2016, ApJ, 825, 5
Stark D. P. et al., 2010, MNRAS, 408, 1628
Stark D. P., Ellis R. S., Ouchi M., 2011, ApJ, 728, L2
Stefanon M. et al., 2017, ApJ, 843, 36
Stefanon M. et al., 2019, ApJ, 883, 99
Sun G., Furlanetto S. R., 2016, MNRAS, 460, 417
Tacchella S. et al., 2018, ApJ, 868, 92
Tinker J. L. et al., 2010, ApJ, 724, 878
Trenti M. et al., 2010, ApJ, 714, L202
Van Der Walt S., Colbert S. C., Varoquaux G., 2011, Comput. Sci. Eng., 13, 22
Verhamme A. et al., 2008, A&A, 491, 89
VirtanenP. et al., 2020, Nature Methods, 17, 261
Vogelsberger M. et al., 2020, MNRAS, 492, 5167
Weingartner J. C., Draine B. T., 2001, ApJ, 548, 296
Wilkins S. M. et al., 2013, MNRAS, 430, 2885
Wilkins S. M. et al., 2016, MNRAS, 455, 659
Williams C. C. et al., 2018, ApJS, 236, 33
Windhorst R. A. et al., 2011, ApJS, 193, 27
Yang H. et al., 2017, ApJ, 844, 171
Yung L. Y. A. et al., 2019, MNRAS, 483, 2983

APPENDIX A: PHOTOMETRIC ESTIMATES OF UV SLOPES

It is important to extract $M_{\rm UV}$ and β from theoretical models in an observationally motivated way, given that biases (in β especially) comparable to observed trends can arise if using an idealized approach.

In Fig. A1, we illustrate the difference between spectroscopic and photometric estimates of UV magnitudes and colours. In the former case, we assume $M_{\rm UV} = M_{1600}$, and measure β as a power-law fit to intrinsic galaxy spectra through the C94 windows. Our photometric estimates follow B14, who compute $M_{\rm UV}$ as the geometric mean of all photometry, which we indicate with angular brackets, $\langle M_{\rm UV}\rangle$. Similarly, β is computed as a power-law fit through available photometry. In Fig. A2, we show $\Delta\beta \simeq 0.05$–0.2 biases in the UV colour estimated with *HST* photometry (open symbols), similar to what was found in Finkelstein et al. (2012, see their fig. 4), particularly at $z \gtrsim 6$. These biases will persist, even with *JWST* (see filled symbols in lower right-hand panel of Fig. A2).

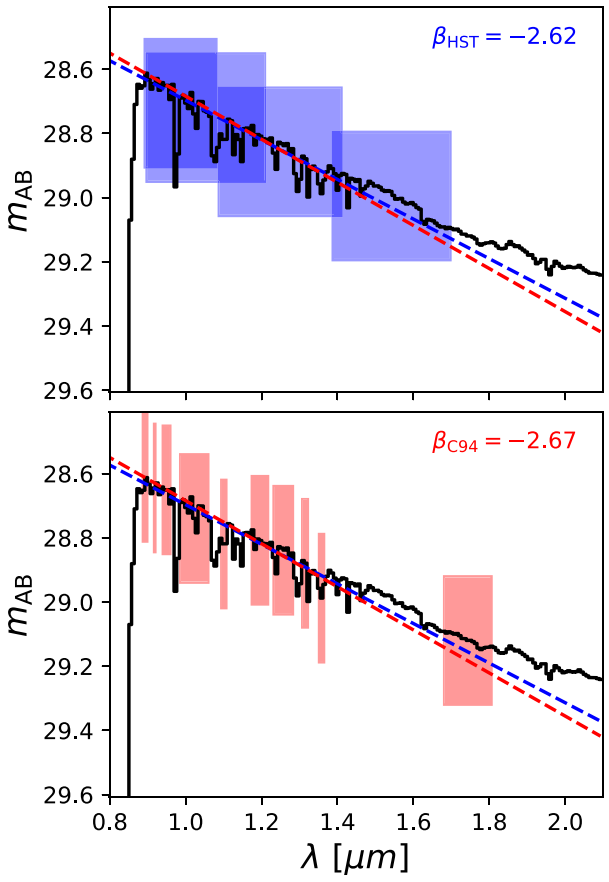

Figure A1. Example extraction of β from photometry. The top panel shows the four *HST* filters used for galaxies at $z \sim 6$, while the bottom panel shows the placement of the Calzetti et al. (1994) spectral windows. Dashed lines show best-fitting power laws using the respective photometric windows. Input spectrum corresponds to a dust-free galaxy from our model observed at $z = 6$, with $M_* = 10^9$ M$_\odot$, $\dot{M}_* \simeq 3$ M$_\odot$ yr^{-1}, and $M_{\rm h} \sim 10^{11}$ M$_\odot$.

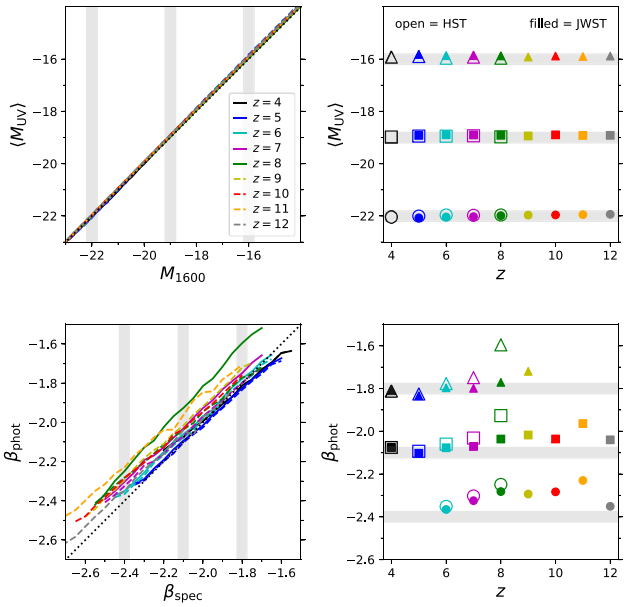

Figure A2. Biases in photometrically estimated UV magnitudes and colours. Top panel: UV magnitude, $\langle M_{\rm UV}\rangle$, defined as geometric mean of available photometry, as a function of monochrotmatic UV magnitude, M_{1600}, at a series of redshifts (left-hand panels), and as a function of redshift holding M_{1600} fixed at -22, -19, and -16 (right-hand panels). Bottom panels: UV colour estimated via power-law fits to available photometry, $\beta_{\rm phot}$, versus the 'true' UV colour, $\beta_{\rm spec}$, computed via power-law fits to the forward-modelled galaxy SEDs in the Calzetti et al. (1994) spectral windows. We show $\beta_{\rm phot}$ as a function of z in three fixed $\beta_{\rm spec}$ bins, $\beta_{\rm spec} = -2.4$, -2.1, and -1.9. Open symbols in right-hand panels indicate measurements performed only with *HST* WFC/WFC-3, while filled symbols use *HST* and *JWST* NIRCAM medium filters. See Table A1 for listing of filters used at each redshift.

From Fig. A3, it is clear that at $z \gtrsim 6$, the rest $1268 \lesssim \lambda \lesssim 2580$ range is sampled more sparsely than at $z \lesssim 6$. This is particularly noticeable at $z \sim 7$–8 for *HST*, with coverage heavily weighted to the bluest part of the rest UV spectrum. Coverage at $9 \lesssim 11$ for *JWST* is also weighted to the bluest part of the rest-UV spectrum, which has many absorption features, hence the redward bias (bottom right-hand panel of Fig. A2). Clearly, use of the NIRCAM medium filters will be required in order to get accurate colours at $z \gtrsim 8$. One could, in principle, use the wide filters at $z > 8$ also, though their increasing spectral width with wavelength corresponds to an increase in contamination from emission outside the C94 spectral range (see unfilled boxes).

We include a full listing of the filters used as a function of redshift both for *HST* and the *JWST* wide and medium filters in Table A1.

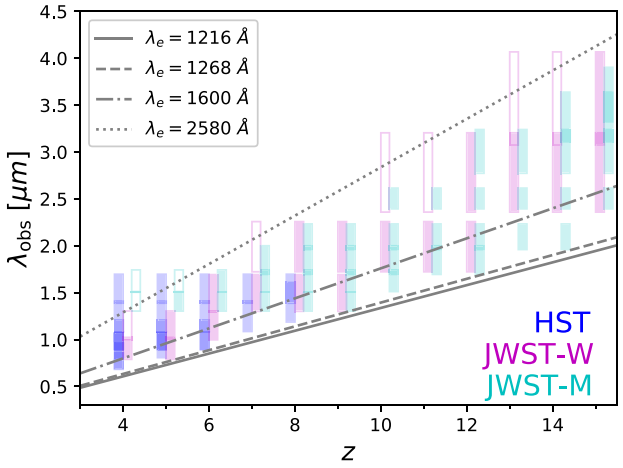

Figure A3. Filter placement versus galaxy redshift. Blue boxes show filters used in *HST* analyses of B14, while filled magenta and cyan boxes show the NIRCAM wide (W) and medium (M) filters that lie completely within the C94 windows. Open boxes are added to supplement the NIRCAM filter set at redshifts for which only one filter lies within C94 windows. For reference, we show the observed wavelength of Ly α (solid), the Calzetti et al. (1994) spectral range often used to compute β (dashed, dotted), and 1600 Å, where magnitudes are often reported. See Table A1 for full filter listing. Note that boxes are offset horizontally for clarity, but placement is determined using integer redshifts only.

APPENDIX B: UPDATED SFE CONSTRAINTS

For completeness, here we present our new constraints on the SFRD and SFE parameters, and compare directly to two other common approaches: (i) use of the IRX–β method of dust correction, or (ii) neglect of dust attenuation entirely.

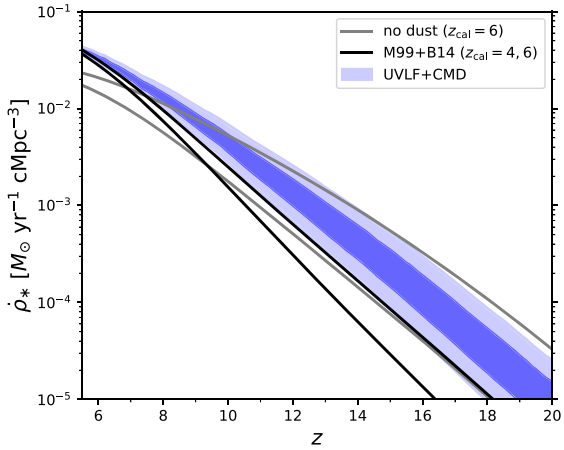

Figure B1. SFRD at high z predicted by our models. Blue contours show 68 and 95 per cent credibility regions for our new models presented in this work, while grey and black contours show the 68 per cent confidence regions obtained when neglecting dust and employing an IRX–β-based approach (described in the text), respectively.

First, in Fig. B1, we show the SFRD reconstructed from our fits compared to a dust-free model calibrated to $z \sim 6$ UVLFs from B15 (grey contours), and a model calibrated to $z \sim 4$ and 6 UVLFs from B15 using the common M99+B14 IRX–β-based approach (black contours). Our new models (blue) agree with the IRX–β approach at $z \sim 6$, predicting $\dot{\rho}_* \simeq 4 \times 10^{-2}$ M$_\odot$ yr^{-1} cMpc^{-3}. At higher redshifts, the new models tend toward the dust-free $z \sim 6$ calibration because, while the inclusion of dust biases the normalization of the SFE high, it also biases the shape of the SFE toward steeper slopes. As a result, our new estimates of the $z \gtrsim 10$ SFRD are largely unchanged compared to previous work (Mirocha et al. 2017).

Table A1. Filters used to estimate UV slope as a function of redshift.

Filter	Name	$z \sim 4$	$z \sim 5$	$z \sim 6$	$z \sim 7$	$z \sim 8$	$z \sim 9$	$z \sim 10$	$z \sim 11$	$z \sim 12$	$z \sim 13$
F775W	i_{775}	✓									
F814W	I_{814}	✓									
F850LP	z_{850}	✓	✓								
F098M	Y_{098}	✓	✓	✓	✓						
F105W	Y_{105}	✓	✓	✓							
F125W	J_{125}	✓	✓	✓	✓						
F140W	JH_{140}					✓					
F160W	H_{160}	✓	✓	✓	✓	✓					
F090W	n/a	✓	✓								
F115W	n/a	✗	✓	✓							
F150W	n/a			✓	✓	✓	✓				
F200W	n/a				✗	✓	✓	✓	✓	✓	
F277W	n/a							✗	✗	✓	✓
F356W	n/a										✗
F140M	n/a	✓	✓	✓	✓	✓	✓				
F162M	n/a	✗	✗	✓	✓	✓	✓	✓			
F182M	n/a				✓	✓	✓	✓	✓	✓	
F210M	n/a					✓	✓	✓	✓	✓	✓
F250M	n/a							✓	✓	✓	✓
F300M	n/a									✓	✓

Notes. The first eight rows indicate *HST* filters used by Bouwens et al. (2014) to estimate β, while the bottom two blocks of filters demarcated by horizontal lines are the *JWST* NIRCAM wide and medium filters used in our analysis. We use ✓ symbols to indicate filters that lie within the spectral range of the C94 filters, and ✗ symbols to indicate filters used out of necessity, when only one filter lies in the desired range. Use of undesirable filters at $z = 7$, 10, and 13 are responsible for the sharp features in Fig. 8 when adopting the NIRCAM wide filters.

In Fig. B2, we show the posterior distribution of the SFE parameters. Following the same colour scheme as in Fig. B1, we can quickly see that there are systematic differences when neglecting dust, particularly in the normalization of the SFE, which here we anchor to $10^{10}\,M_{\odot}$ haloes (first column; $f_{*,10}$). Our new approach, while qualitatively similar to the M99+B14 contours, does exhibit some important differences. For example, the new models prefer shallower SFE slopes at low mass (third column; $\alpha_{*,lo}$), which result in slightly higher SFRDs at $z \gtrsim 8$, as described above and shown in Fig. B1.

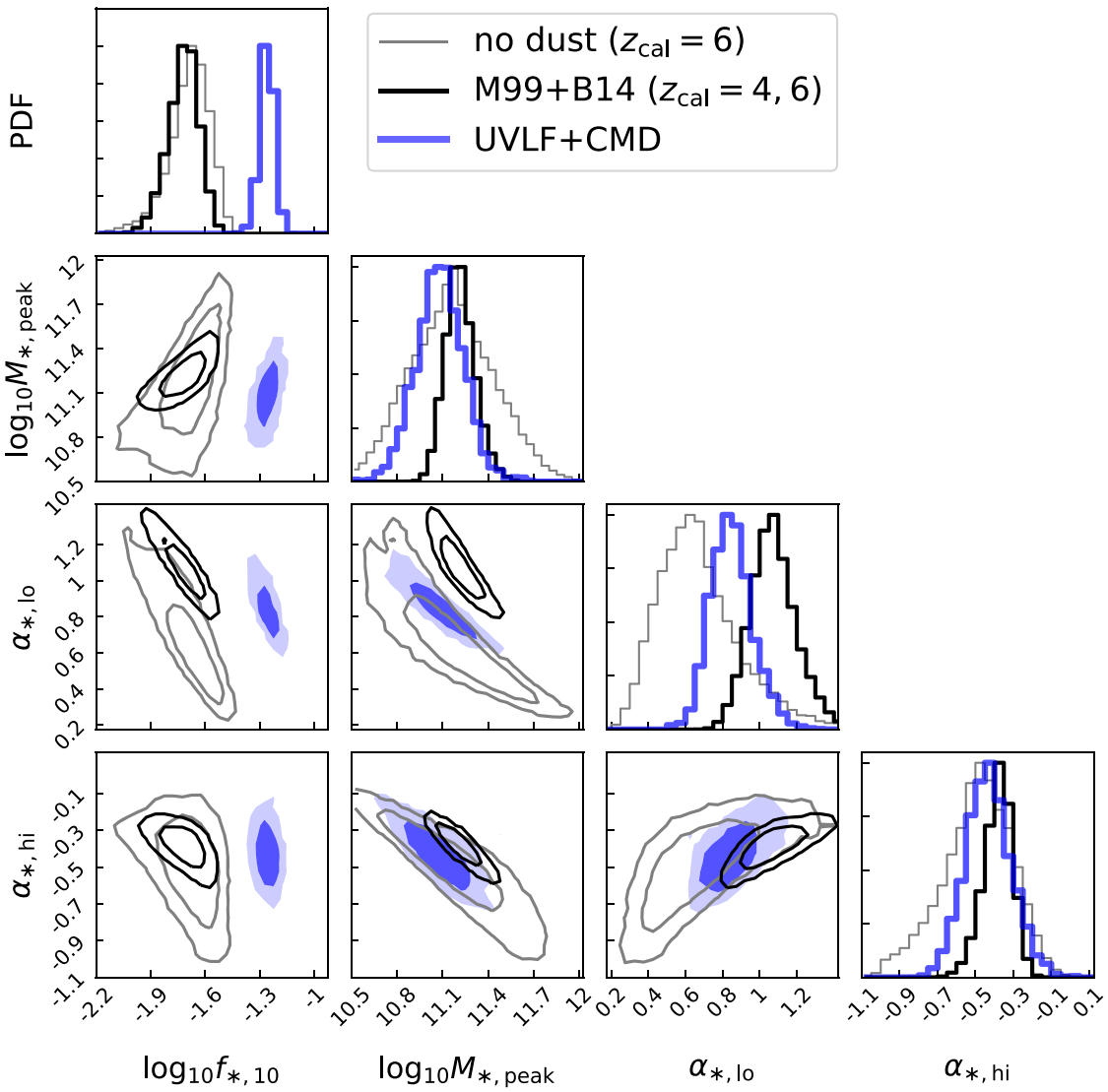

Figure B2. Posterior distribution of SFE parameters for three different approaches to rest-UV inference. Results using the new model presented in this work are shown in blue, while the dust-free and M99+B14 results are shown in grey and black, respectively. In each case, inner contours represent 68 per cent confidence regions, while outer contours indicate 95 per cent. Best fits and 1σ error bars are presented in Table 1.

This paper has been typeset from a TEX/LATEX file prepared by the author.