

High redshift *JWST* predictions from IllustrisTNG: Dust modelling and galaxy luminosity functions

Mark Vogelsberger^{1*}, Dylan Nelson², Annalisa Pillepich³, Xuejian Shen¹, Federico Marinacci^{5,4,1}, Volker Springel², Rüdiger Pakmor², Sandro Tacchella⁴, Rainer Weinberger⁴, Paul Torrey^{6,1}, Lars Hernquist⁴

¹ Department of Physics, Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

² Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, 85741 Garching, Germany

³ Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany

⁴ Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA, 02138, USA

⁵ Department of Physics & Astronomy, University of Bologna, via Gobetti 93/2, 40129 Bologna, Italy

⁶ Department of Astronomy, University of Florida, 316 Bryant Space Sciences Center, Gainesville, FL 32611 USA

17 April 2019

ABSTRACT

The *James Webb Space Telescope* (*JWST*) promises to revolutionise our understanding of the early Universe, and contrasting its upcoming observations with predictions of the Λ CDM model requires detailed theoretical forecasts. Here, we exploit the large dynamic range of the IllustrisTNG simulation suite, TNG50, TNG100, and TNG300, to derive multi-band galaxy luminosity functions from $z = 2$ to $z = 10$. We put particular emphasis on the exploration of different dust attenuation models to determine galaxy luminosity functions for the rest-frame ultraviolet (UV), and apparent wide NIRC*am* bands. Our most detailed dust model is based on continuum Monte Carlo radiative transfer calculations employing observationally calibrated dust properties. This calibration results in constraints on the redshift evolution of the dust attenuation normalisation and dust-to-metal ratios yielding a stronger redshift evolution of the attenuation normalisation compared to most previous theoretical studies. Overall we find good agreement between the rest-frame UV luminosity functions and observational data for all redshifts, also beyond the regimes used for the dust model calibrations. Furthermore, we also recover the observed high redshift ($z = 4 - 6$) UV luminosity versus stellar mass relation, the $H\alpha$ versus star formation rate relation, and the $H\alpha$ luminosity function at $z = 2$. The bright end ($M_{UV} > -19.5$) cumulative galaxy number densities are consistent with observational data. For the F200W NIRC*am* band, we predict that *JWST* will detect ~ 80 (~ 200) galaxies with a signal-to-noise ratio of 10 (5) within the NIRC*am* field of view, 2.2×2.2 arcmin², for a total exposure time of 10^5 s in the redshift range $z = 8 \pm 0.5$. These numbers drop to ~ 10 (~ 40) for an exposure time of 10^4 s.

Key words: methods: numerical, galaxies: evolution, galaxies: formation

1 INTRODUCTION

The concordance paradigm of structure formation, the Λ CDM model (e.g., Planck Collaboration 2016), provides testable predictions for the growth of dark matter halos and is the basis of our current theory for galaxy formation (White & Rees 1978; Blumenthal et al. 1984). Confronting these predictions with observations at multiple epochs of cosmic evolution is crucial to confirm or falsify this theoretical framework. One of the most basic quantities of the galaxy population is the galaxy luminosity function that measures

the comoving number density of galaxies as a function of their luminosities at different redshifts.

JWST will open a new window into the high redshift Universe to study faint and distant galaxies during the early phases of cosmic evolution. Specifically, *JWST* will quantify the galaxy population and galaxy luminosity functions at higher redshifts than ever before. Most importantly it will also dramatically increase the statistical sample sizes of high redshift galaxies. Two *JWST* instruments, the Near InfraRed Camera (NIRC*am*) and the Mid InfraRed Instrument (MIRI), are designed to obtain broadband photometry over the wavelength range from 0.7 to 25.5 μ m with unprecedented sensitivity and angular resolution. This wavelength coverage al-

* email:mvogelsb@mit.edu

lows *JWST* to probe the rest-frame UV, optical, and near-infrared (IR) spectral energy distributions (SEDs) of high-redshift galaxies, which will lead to novel insights into the evolution of the earliest galaxy populations. These observations are crucial to study the physics that shapes the galaxy population in the early Universe. Furthermore, quantifying the high redshift galaxy population is also important to understand the sources of reionisation in the early Universe.

These upcoming *JWST* observations will extend the successes of previous studies to explore the high redshift Universe. For example, the Lyman-break technique (e.g., Steidel & Hamilton 1993; Steidel et al. 1996) allowed the detection of galaxies at $z \sim 3$. These early limits were pushed further with the Advanced Camera on the Hubble Space Telescope (*HST*), which extended the Lyman-break technique to $z \sim 6$ (Bouwens et al. 2003; Stanway et al. 2003). The Wide-Field Camera 3 with near-IR filters then further increased the number of galaxies that could be identified at $z \sim 7$ (e.g., Bouwens et al. 2010; Wilkins et al. 2010; Oesch et al. 2010a), enlarging the sample sizes of galaxies at these redshifts substantially, with a few examples at $z \sim 10$. Specifically, *HST* has detected about 2000 galaxy candidates at high redshifts ($z \sim 4-10$) from both blank and gravitationally lensed fields (e.g., Koekemoer et al. 2013; Lotz et al. 2017). Brighter objects have also been discovered with ground-based facilities, such as the United Kingdom Infra-Red Telescope (UKIRT) and the Visible and Infrared Survey Telescope for Astronomy (VISTA) (McLure et al. 2009; Bowler et al. 2015). These combined observations have provided constraints on the space density of relatively bright galaxy populations up to $z \sim 10$ (e.g., McLure et al. 2009, 2013; Castellano et al. 2010; van der Burg et al. 2010; Oesch et al. 2013, 2014, 2018; Schenker et al. 2013; Tilvi et al. 2013; Bowler et al. 2014, 2015; Bouwens et al. 2014b, 2015, 2016, 2017; Schmidt et al. 2014; Atek et al. 2015a; McLeod et al. 2015, 2016; Finkelstein et al. 2015; Livermore et al. 2017; Ishigaki et al. 2018).

Despite these successes, it still remains difficult to observationally quantify evolutionary details of the high redshift galaxy luminosity functions. For example, probing the high redshift faint-end galaxy population and luminosity functions with existing observational facilities is challenging. Most of the currently existing constraints rely on fields that are lensed by massive foreground galaxy clusters with uncertain magnification corrections (e.g., Kawamata et al. 2016; Bouwens et al. 2017; Prieue et al. 2017). It is therefore difficult to derive robust results for the evolution of the faint-end of the galaxy luminosity function. *JWST* will probe the luminosity function in unlensed fields much deeper, and importantly also provide more robust redshift measurements for high redshift candidates that are currently selected by the Lyman-break technique. This will lead to a better understanding of the faint-end population of the galaxy luminosity function and its evolution with redshift with important consequences for the source of reionisation.

Given the prospects of *JWST* to quantify the galaxy population of the high redshift Universe, theoretical predictions are required to provide templates for comparisons with upcoming *JWST* observational data. Various hydrodynamical simulations have therefore been employed to explore the high redshift Universe and to derive *JWST* predictions, for example: the First Billion Years simulation suites (e.g., Paardekooper et al. 2013), the BlueTides simulation (e.g., Wilkins et al. 2016, 2017), the Renaissance simulations suite (e.g., Xu et al. 2016; Barrow et al. 2017), the FIRE-2 simulations (e.g., Ma et al. 2018a), the Sphinx simulation (e.g., Rosdahl et al. 2018) and others (e.g., Dayal et al. 2013; Shimizu et al. 2014). In addition, semi-analytic models of galaxy formation (e.g.,

Cowley et al. 2018; Yung et al. 2018), and empirical models (e.g., Tacchella et al. 2018) have also been studied to derive *JWST* predictions and high redshift luminosity functions. We note that most of the studied hydrodynamical simulations have only been evolved down to relatively high redshifts due to the demanding computational requirements. It is therefore often unclear how reliable some of the high redshift predictions are, given that some of these models could not be tested towards lower redshifts, where observational data for comparisons is available. These uncertainties then directly propagate into the derived *JWST* predictions for the high redshift Universe. We therefore stress that reliable high redshift galaxy population forecasts require a galaxy formation model that has been calibrated towards lower redshifts. Such predictions based on well-tested hydrodynamical cosmological simulations and galaxy formation models are currently largely missing.

The goal of this paper is to fill this gap by analysing the recently finished IllustrisTNG simulation suite to predict the high redshift galaxy population as it will be observed by *JWST*. The focus of this paper is a study of high redshift galaxy luminosity functions to derive galaxy number densities as a function of galaxy luminosities at different epochs ranging from $z = 2$ to $z = 10$. We maximise the dynamic range of our analysis by combining all three IllustrisTNG simulations, TNG50, TNG100 and TNG300. Dust plays a crucial role in shaping the bright end of the galaxy luminosity function, and we therefore explore multiple dust models in our study. One of our dust models is based on full Monte Carlo dust radiative transfer calculations where the dust-to-metal ratio is calibrated against current constraints on the rest-frame UV luminosity functions at $z = 2 - 10$. This dust calibration procedure leads to novel constraints on the dust-to-metal ratios, and dust attenuation normalisations. Based on the luminosity functions for various *JWST* NIRCам bands, we can also derive predictions for the expected number of detected galaxies given specific NIRCам configurations.

The structure of this paper is as follows. In Section 2 we briefly describe the IllustrisTNG simulation suite and state the numerical parameters of TNG50, TNG100 and TNG300 in detail. Our dust attenuation models are then described in Section 3. The main results are presented in Section 4, where we discuss various galaxy luminosity functions and *JWST* survey forecasts. Specifically, we discuss rest-frame UV luminosity functions, and then explore *JWST* luminosity functions for all NIRCам wide filter bands. Towards the end of Section 4 we also briefly discuss the predicted mass-to-light ratio relations and $H\alpha$ luminosity star formation rate relation that provide another important verification of our simulation predictions and radiative transfer based dust modelling in particular. Our summary and conclusions are presented in Section 5.

2 THE ILLUSTRIS TNG SIMULATION SUITE

Our analysis is based on the IllustrisTNG simulation suite (Marnacci et al. 2018; Naiman et al. 2018; Nelson et al. 2018; Pillepich et al. 2018b; Springel et al. 2018), including the newest addition of TNG50 (Nelson et al. 2019; Pillepich et al. 2019), which consists of three primary simulations: TNG50, TNG100 and TNG300. IllustrisTNG is the follow-up project of the Illustris simulations (Vogelsberger et al. 2014a,b; Genel et al. 2014; Nelson et al. 2015; Sijacki et al. 2015). The IllustrisTNG simulation suite employs the following cosmological parameters (Planck Collaboration 2016): $\Omega_m = 0.3089$, $\Omega_b = 0.0486$, $\Omega_\Lambda = 0.6911$, $H_0 = 100 h \text{ km s}^{-1} \text{ Mpc}^{-1} = 67.74 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\sigma_8 = 0.8159$, and

IllustrisTNG Simulation	run	volume side length [$h^{-1}\text{Mpc}$]	N_{gas}	N_{dm}	m_{b} [$h^{-1}M_{\odot}$]	m_{dm} [$h^{-1}M_{\odot}$]	$\epsilon_{\text{dm,stars}}$ [$h^{-1}\text{kpc}$]	$\epsilon_{\text{gas}}^{\text{min}}$ [$h^{-1}\text{kpc}$]
TNG300	TNG300(-1)	205	2500^3	2500^3	7.4×10^6	4.0×10^7	1.0	0.25
TNG100	TNG100(-1)	75	1820^3	1820^3	9.4×10^5	5.1×10^6	0.5	0.125
TNG50	TNG50(-1)	35	2160^3	2160^3	5.7×10^4	3.1×10^5	0.2	0.05

Table 1. IllustrisTNG simulation suite. The table presents the basic numerical parameters of all IllustrisTNG simulations studied in this paper: simulation volume side length, number of gas cells (N_{gas}), number of dark matter particles (N_{dm}), baryon mass resolution (m_{b}), dark matter mass resolution (m_{DM}), Plummer-equivalent maximum physical softening length of dark matter and stellar particles ($\epsilon_{\text{dm,stars}}$) and the minimal comoving cell softening length $\epsilon_{\text{gas}}^{\text{min}}$. In the following we will identify TNG50 with TNG50-1, TNG100 with TNG100-1 and TNG300 with TNG300-1.

$n_{\text{s}} = 0.9667$. The three major simulations cover three different periodic, uniformly sampled volumes, roughly $\sim 50^3, 100^3, 300^3 \text{ Mpc}^3$ for TNG50, TNG100 and TNG300, respectively. The numerical parameters of the different simulations are summarised in Table 1. All simulations were carried out with the moving-mesh code AREPO (Springel 2010; Pakmor et al. 2016) combined with the IllustrisTNG galaxy formation model (Weinberger et al. 2017; Pillepich et al. 2018a) which is an updated version of the Illustris galaxy formation model (Vogelsberger et al. 2013; Torrey et al. 2014). We note that TNG50, TNG100, and TNG300 differ in their highest numerical resolution as listed in Table 1. In the following we will identify TNG50 with TNG50-1, TNG100 with TNG100-1 and TNG300 with TNG300-1.

3 GALAXY LUMINOSITIES AND DUST MODELS

The central goal of our paper is to provide high redshift galaxy luminosity functions. We therefore have to assign luminosities and multiple band magnitudes to each galaxy of all different simulations of IllustrisTNG. We will describe this assignment process in this section. In this work we define a galaxy as being either a central or satellite galaxy as identified by the SUBFIND algorithm (Springel et al. 2001; Dolag et al. 2009). For the following analysis, we impose a stellar mass cut for galaxies. We only consider galaxies with a stellar mass larger than 100 times the baryonic mass resolution, $100 \times m_{\text{b}}$, within twice the stellar half mass radius. Galaxies resolved with a lower number of resolution elements will not be considered, since we assume that their structure is not reliably modelled. Furthermore, galaxy luminosities, magnitudes, stellar masses and star formation rates are calculated within a fixed physical aperture of 30 pkpc and based on gravitationally bound particles and cells.

3.1 Dust-free galaxy emission

We first describe our method to derive dust-free or intrinsic magnitudes, i.e. assuming that no dust is absorbing or scattering photons within the galaxy. Consequently these dust-free magnitudes also do not include dust emission. We note that dust emission also has no impact on the derived magnitudes for the given redshift and filter ranges studied in this paper. To calculate dust-free magnitudes we treat each stellar particle in the simulation as a simple stellar population (SSP) using a stellar population synthesis method. Here we employ the Flexible Stellar Population Synthesis (FSPS) code (Conroy et al. 2009; Conroy & Gunn 2010) with MIST isochrones (Paxton et al. 2011, 2013, 2015; Choi et al. 2016; Dotter 2016) and the MILES stellar library (Sánchez-Blázquez

et al. 2006a,b), assuming a Chabrier initial mass function (Chabrier 2003) consistent with the IllustrisTNG galaxy formation model. To accelerate the luminosity and magnitude calculations, we first generate a two-dimensional magnitude grid with one axis containing initial stellar metallicity values and the other axis containing stellar age values. The initial metallicity, Z_{i} , axis contains 13 values uniformly logarithmically spaced from $10^{-4.35}$ to $10^{-1.35}$. The stellar ages, t_{ssp} , axis contains 107 values uniformly logarithmically spaced from 10^{-4} to $10^{1.3}$ Gyr. For each point on this grid, we construct the corresponding SSP and calculate the rest-frame SED normalised to an initial stellar mass of $1 M_{\odot}$. We then use the SEDPY¹ code to convolve the resulting fluxes, either absolute or apparent, with transmission curves of arbitrary bands and calculate rest-frame magnitudes. We note that all the magnitudes in this work are AB magnitudes, i.e. assuming a reference flux of 3631 Jy. To extract rest-frame UV magnitudes, M_{UV} , we convolve the fluxes with a tophat filter of width 10 nm centred on 150 nm; i.e. $M_{\text{UV}} = M_{150\text{nm}}$. This definition of UV magnitude is consistent with most other observational and theoretical studies (e.g., Duncan et al. 2014; Bowler et al. 2014, 2015; Finkelstein et al. 2015; Liu et al. 2016; Cullen et al. 2017).

To calculate apparent dust-free magnitudes, we redshift the rest-frame SED, $L_{\nu}(v_{\text{em}})$, and calculate the apparent flux at a given frequency as:

$$f_{\nu}(v_{\text{obs}}) = \frac{(1+z)L_{\nu}(v_{\text{em}})}{4\pi d_L^2} e^{-\tau_{\text{IGM}}^{\text{eff}}(v_{\text{obs}})}, \quad (1)$$

where $v_{\text{obs}} = v_{\text{em}}/(1+z)$, z is the redshift of the source, d_L is the luminosity distance to the source and $\tau_{\text{IGM}}^{\text{eff}}(\nu)$ is the effective optical depth of the clumpy intergalactic medium (IGM). This optical depth, $\tau_{\text{IGM}}^{\text{eff}}$, is calculated based on the IGM absorption model of Madau (1995) and Madau et al. (1996). We note that we do not consider IGM absorption corrections for rest-frame magnitudes since this absorption occurs at $\lesssim 120 \text{ nm}$ (Lyman- α), which is shorter than the wavelength ranges of all bands studied in our work. Based on this apparent flux we then derive band magnitudes through a convolution as described above.

HII regions surrounding young stellar populations reprocess the Lyman continuum photons into nebular continuum and line emissions, which represents an important contribution to galaxy SEDs. This emission depends on both the spectra of ionising photons and the properties of the surrounding medium, including geometry, gas density, chemical content and the covering fraction of Lyman continuum photons. Here we adopt the nebular emission

¹ <https://github.com/bd-j/sedpy>

model of [Byler et al. \(2017\)](#), which is based on photoionization calculations using the CLOUDY code ([Ferland et al. 2013](#)). The model assumes that the fraction of escaping Lyman continuum photons is zero and nebular emission is purely determined by the gas phase metallicity and the ionisation parameter. Following [Wilkins et al. \(2016\)](#) and [Byler et al. \(2017\)](#), we choose the gas phase metallicity to be the same as the initial metallicity of the stellar particle, which is inherited from the gas cell from which the stellar particle has been created. The ionisation parameter, as defined in [Byler et al. \(2017\)](#), encodes the intensity of the ionising source and the geometry of the gas cloud. We choose the ionisation parameter to be 0.01 as suggested in [Byler et al. \(2018\)](#). The time span that a SSP is surrounded by its birth cloud is limited, so nebular emission is only active for SSPs with $t_{\text{ssp}} < t_{\text{esc}} = 10$ Myr. Stars older than t_{esc} can escape the birth clouds.

Based on the magnitude grid, we assign rest-frame and apparent band magnitudes to stellar particles of galaxies. For each stellar particle we employ a bilinear interpolation of the grid values to calculate the band magnitudes corresponding to the Z_i and t_{ssp} values of the stellar particle. In a last step, we sum up the luminosity contribution from all stellar particles in a galaxy within the physical aperture of 30 pkpc. We mass-weight each particle by its initial mass m_i to derive the dust-free luminosities and the corresponding band magnitudes of the whole galaxy. This model is essentially the same as the dust-free model A of [Nelson et al. \(2018\)](#).

3.2 Dust attenuation models

Dust attenuation can significantly alter the SED of a galaxy, and correspondingly its luminosity and magnitude in certain bands. Consequently, dust attenuation also changes the galaxy luminosity function, especially towards the bright end, where large amounts of dust are available within galaxies. It is therefore necessary to model the effects of dust attenuation to turn the dust-free band magnitudes of the previous section into dust-corrected band magnitudes. In the following we explore three different dust attenuation models of increasing complexity:

- Model A – empirical dust model
- Model B – resolved dust optical depth model
- Model C – resolved dust radiative transfer model

For Model A, we employ observationally derived empirical scaling relations to link dust-free rest-frame UV magnitudes with observed dust-attenuated rest-frame UV magnitudes. This model is solely based on the dust-free rest-frame UV magnitudes, and this approach does not require any further input from the simulations. For Model B, we consider in addition also the simulated gas distribution and spatial configuration of stars within each simulated galaxy to derive dust column densities to attenuate the light of stellar particles without explicit radiative transfer. This second model therefore takes into account additional information from the simulated galaxies to derive dust-attenuated luminosities and magnitudes. For Model C, we also consider the local gas distribution in galaxies and perform full dust Monte Carlo radiative transfer calculations of photons propagating through the dusty interstellar medium of simulated galaxies. Most theoretical high redshift luminosity function studies have employed Model A (e.g., [Tacchella et al. 2013](#); [Mason et al. 2015](#); [Liu et al. 2016](#); [Tacchella et al. 2018](#)) or Model B (e.g., [Kitzbichler & White 2007](#); [Somerville et al. 2012](#); [Clay et al. 2015](#); [Trayford et al. 2015](#); [Wilkins et al. 2017](#); [Yung et al. 2018](#)) to correct for the attenuation of dust. Some works have also explored Monte Carlo radiative transfer calculations (e.g., [Trayford](#)

[et al. 2017](#)), however not at high redshift employing a calibrated dust model. Here we go beyond those approaches by also performing full dust radiative transfer for galaxies in TNG50, TNG100, and TNG300 using a calibrated dust model from $z = 2$ to $z = 10$ to derive dust-attenuated galaxy luminosity functions. In the following we describe the details of each dust model.

3.2.1 Model A – empirical dust model

Instead of directly modelling dust attenuation based on the local gas and dust properties around stars in simulated galaxies, we can employ observationally motivated empirical scaling relations to derive dust-attenuated rest-frame UV magnitudes of the simulated galaxies. This is the basic idea of our dust Model A. We note that this approach is limited to UV magnitudes since the empirical scaling relations have only been established for this band. It is therefore not possible to derive dust-attenuated luminosities or magnitudes for other bands. Predictions for those bands will be derived based on dust Models B and C as discussed below.

To implement this dust Model A, we first assume that the observed, and therefore dust-attenuated, UV spectrum of a galaxy can be characterised by a simple power law $f_\lambda \sim \lambda^\beta$, where β is the slope of the UV continuum. The attenuation of UV light and the corresponding reprocessed light into IR wavelengths correlates with this slope. Many works have studied this $\text{IRX} = (L_{\text{IR}}/L_{\text{UV}})$ versus β relation in detail finding that (e.g., [Meurer et al. 1999](#); [Siana et al. 2009](#); [Casey et al. 2014](#)):

$$\text{IRX} = B [10^{0.4(C_0 + C_1\beta)} - 1], \quad (2)$$

where B , C_0 and C_1 are constants. Since $\text{IRX} = B[10^{0.4A_{\text{UV}}} - 1]$, the attenuation of the galaxy UV light, A_{UV} , depends then linearly on β : $A_{\text{UV}} = C_0 + C_1\beta$. For the constants, [Meurer et al. \(1999\)](#) found $C_0 = 4.43$, $C_1 = 1.99$ for the originally proposed relation. For the Small Magellanic Cloud $\text{IRX} - \beta$ relation ([Lequeux et al. 1982](#); [Prevot et al. 1984](#); [Bouchet et al. 1985](#); [Bouwens et al. 2016](#)), these constants are $C_0 = 2.45$, $C_1 = 1.1$. More recent measurements find $C_0 = 3.36$, $C_1 = 2.04$ ([Casey et al. 2014](#)). Here, we adopt the original Meurer relation for our Model A.

Using the $A_{\text{UV}} - \beta$ relation we can calculate the dust attenuation strength based on the β slope. The slope, β , itself furthermore correlates with the observed dust-attenuated rest-frame UV magnitude. This empirical relation can, for example, be parametrised as (e.g. [Bouwens et al. 2012, 2014a](#)):

$$\langle \beta \rangle (M_{\text{UV}}^{\text{dust}}) = \frac{d\beta}{dM_{\text{UV}}^{\text{dust}}}(z) [M_{\text{UV}}^{\text{dust}} - M_0] + \beta_{M_0}(z), \quad (3)$$

where $M_{\text{UV}}^{\text{dust}}$ is the rest-frame UV magnitude observed after dust attenuation. Some observationally derived values for $\beta_{M_0}(z)$, $d\beta/dM_{\text{UV}}^{\text{dust}}(z)$ and M_0 are summarised in Table 2.

To derive the mean attenuation at a given $M_{\text{UV}}^{\text{dust}}$ and ultimately link $M_{\text{UV}}^{\text{dust}}$ with $M_{\text{UV}}^{\text{dust-free}}$, the dust-free rest-frame UV magnitude, we briefly introduce two commonly employed approaches: (i) taking averages of the inferred dust-attenuated and dust-free luminosities or (ii) taking averages of the inferred dust-attenuated and dust-free magnitudes. These two approaches are used to derive dust-attenuated rest-frame UV magnitudes based on the dust-free rest-frame UV magnitudes, or for the reverse operation. For each approach, the mean dust attenuation can furthermore be derived in two ways either analytically or numerically. We note that these various methods of establishing the link between $M_{\text{UV}}^{\text{dust}}$ and $M_{\text{UV}}^{\text{dust-free}}$ are not fully equivalent once scatter in β at a

redshift z	M_0^a [mag]	$\beta_{M_0^a}$	$d\beta/dM_{UV}^a$ [mag $^{-1}$]	M_0^b [mag]	$\beta_{M_0^b}$	$d\beta/dM_{UV}^b$ [mag $^{-1}$]	M_0^c [mag]	$\beta_{M_0^c}$	$d\beta/dM_{UV}^c$ [mag $^{-1}$]
4	-19.5	$-1.85 \pm 0.01 \pm 0.06$	-0.11 ± 0.01	-20.31	$-1.86^{+0.03}_{-0.02}$	$+0.01 \pm 0.03$	-19.5	$-2.00 \pm 0.02 \pm 0.10$	-0.11 ± 0.01
5	-19.5	$-1.91 \pm 0.02 \pm 0.06$	-0.14 ± 0.02	-19.94	$-1.97^{+0.07}_{-0.04}$	$+0.00 \pm 0.06$	-19.5	$-2.08 \pm 0.03 \pm 0.10$	-0.16 ± 0.03
6	-19.5	$-2.00 \pm 0.05 \pm 0.08$	-0.20 ± 0.04	-19.54	$-2.02^{+0.13}_{-0.04}$	-0.10 ± 0.07	-19.5	$-2.20 \pm 0.05 \pm 0.14$	-0.15 ± 0.04
7	-19.5	$-2.05 \pm 0.09 \pm 0.13$	-0.20 ± 0.07	-19.39	$-2.42^{+0.31}_{-0.13}$	-0.20 ± 0.11	-19.5	$-2.27 \pm 0.07 \pm 0.28$	-0.21 ± 0.07
8	-19.5	$-2.13 \pm 0.44 \pm 0.27$	-0.15 (fixed)	-19.35	$-2.03^{+0.46}_{-0.38}$	-0.03 ± 0.26			

^a Bouwens et al. (2014a), ^b Finkelstein et al. (2012) (M_0 is calculated based on $0.5L^*$, where the luminosity bin is $0.25L^* - 0.75L^*$ and L^* is taken from the reference), ^c Bouwens et al. (2012)

Table 2. Observationally derived parameters for the $M_{UV}^{\text{dust}} - \langle\beta\rangle$ relation. We also state random and systematic errors, respectively. The quoted values differ quite significantly between the different observational studies.

given M_{UV}^{dust} is considered. However, this scatter only causes minor differences in the actual mapping between M_{UV}^{dust} and $M_{UV}^{\text{dust-free}}$ as we will demonstrate below.

Analytical approach: We first discuss an analytic approach to link dust-attenuated and dust-free magnitudes.

(i) *Luminosity based average:* Here we derive an analytic relation between the dust-free rest-frame UV magnitude, $M_{UV}^{\text{dust-free}}$, and the dust-corrected rest-frame UV magnitude, M_{UV}^{dust} , based on luminosity averages. This method has often been adopted in previous works (e.g., Tacchella et al. 2013; Mason et al. 2015; Tacchella et al. 2018). Most of these studies also commonly assume that there is some scatter in β for a given M_{UV}^{dust} . This scatter is assumed to follow a Gaussian distribution at each M_{UV}^{dust} with a constant dispersion $\sigma_\beta = 0.34$ (e.g., Smit et al. 2012; Tacchella et al. 2013, 2018). We can then calculate the average attenuation as a function of $\langle\beta\rangle$. To derive this attenuation, we start from the IRX– β relation and derive $\langle\text{IRX}\rangle$ for the Gaussian β distribution. IRX correlates linearly with the ratio between dust-free UV luminosity and dust-attenuated UV luminosity as:

$$1 + \text{IRX}/B = 10^{0.4A_{UV}} = L_{UV}^{\text{dust-free}}/L_{UV}^{\text{dust}}. \quad (4)$$

We can then derive the dust-free UV luminosity for a given dust-attenuated luminosity as:

$$\ln\left(\frac{L_{UV}^{\text{dust-free}}}{L_{UV}^{\text{dust}}}\right) = \ln(1 + \text{IRX}/B) = 0.4 \ln 10 (C_0 + C_1\beta). \quad (5)$$

The $L_{UV}^{\text{dust-free}}/L_{UV}^{\text{dust}}$ values follow a log-normal distribution since β has a Gaussian distribution. Since this equation is evaluated at a given L_{UV}^{dust} , we have $\langle L_{UV}^{\text{dust-free}}/L_{UV}^{\text{dust}} \rangle = \langle L_{UV}^{\text{dust-free}} \rangle / L_{UV}^{\text{dust}}$. Then the mean attenuation at a given L_{UV}^{dust} or M_{UV}^{dust} is given by:

$$\begin{aligned} \langle A_{UV} \rangle (M_{UV}^{\text{dust}}) &= -2.5 \log\left(\langle L_{UV}^{\text{dust-free}} \rangle / L_{UV}^{\text{dust}}\right) \\ &= C_0 + 0.2 \ln 10 C_1^2 \sigma_\beta^2 + C_1 \langle\beta\rangle (M_{UV}^{\text{dust}}). \end{aligned} \quad (6)$$

We can now use this relation of averages to link $M_{UV}^{\text{dust-free}}$ with M_{UV}^{dust} based on $M_{UV}^{\text{dust}} = M_{UV}^{\text{dust-free}} + \langle A_{UV} \rangle (M_{UV}^{\text{dust}})$, where we set negative $\langle A_{UV} \rangle (M_{UV}^{\text{dust}})$ values to zero to avoid an unphysical attenuation. Replacing $\langle A_{UV} \rangle (M_{UV}^{\text{dust}})$ with Equation 6, and then substituting $\langle\beta\rangle$ with Equation 3 finally yields:

tuting $\langle\beta\rangle$ with Equation 3 finally yields:

$$M_{UV}^{\text{dust}} = \begin{cases} \frac{1}{1 - C_1 \frac{d\beta}{dM_{UV}^{\text{dust}}}(z)} \left[M_{UV}^{\text{dust-free}} + C_0 + C_1 \beta_{M_0}(z) - C_1 \frac{d\beta}{dM_{UV}^{\text{dust}}}(z) M_0 + 0.2 \ln 10 C_1^2 \sigma_\beta^2 \right], & \text{for } \langle A_{UV} \rangle (M_{UV}^{\text{dust}}) > 0 \\ M_{UV}^{\text{dust-free}}, & \text{for } \langle A_{UV} \rangle (M_{UV}^{\text{dust}}) \leq 0. \end{cases} \quad (7)$$

This relation links $M_{UV}^{\text{dust-free}}$ with M_{UV}^{dust} and allows us to apply dust attenuation corrections to the dust-free galaxy rest-frame UV magnitudes calculated in Section 3.1.

(ii) *Magnitude based average:* Instead of calculating average luminosities, we can also directly calculate average dust-free rest-frame UV magnitudes for a given dust-attenuated rest-frame UV magnitude. Taking magnitudes instead of luminosity averages then leads to:

$$\langle A_{UV} \rangle (M_{UV}^{\text{dust}}) = C_0 + C_1 \langle\beta\rangle (M_{UV}^{\text{dust}}). \quad (8)$$

Replacing $\langle A_{UV} \rangle (M_{UV}^{\text{dust}})$ in $M_{UV}^{\text{dust}} = M_{UV}^{\text{dust-free}} + \langle A_{UV} \rangle (M_{UV}^{\text{dust}})$ with Equation 8, and then replacing $\langle\beta\rangle$ with Equation 3 we find:

$$M_{UV}^{\text{dust}} = \begin{cases} \frac{1}{1 - C_1 \frac{d\beta}{dM_{UV}^{\text{dust}}}(z)} \left[M_{UV}^{\text{dust-free}} + C_0 + C_1 \beta_{M_0}(z) - C_1 \frac{d\beta}{dM_{UV}^{\text{dust}}}(z) M_0 \right], & \text{for } \langle A_{UV} \rangle (M_{UV}^{\text{dust}}) > 0 \\ M_{UV}^{\text{dust-free}}, & \text{for } \langle A_{UV} \rangle (M_{UV}^{\text{dust}}) \leq 0. \end{cases} \quad (9)$$

Again this relation links $M_{UV}^{\text{dust-free}}$ with M_{UV}^{dust} and allows us to apply dust attenuation corrections to the dust-free rest-frame UV magnitudes of galaxies derived in Section 3.1. However, the σ_β term does not appear in this relation since the average has been performed over magnitudes rather than luminosities. We will demonstrate below that this does not affect the resulting dust correction in any significant way.

Numerical sampling approach: We next discuss a numerical approach to link dust-attenuated and dust-free magnitudes that is based on a numerical sampling of the underlying distributions. This method can also be applied to luminosities or magnitudes directly. For this approach the attenuation of numerical samples with negative A_{UV} will be set to zero, which leads to a rise in $\langle A_{UV} \rangle$. This

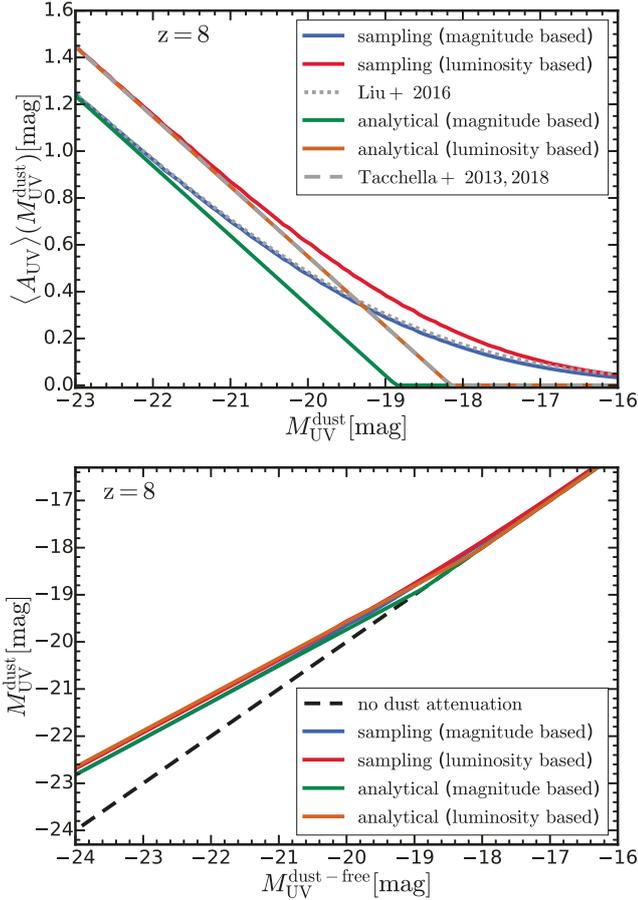


Figure 1. Dust attenuation relations of dust Model A. *Top panel:* Derived $\langle A_{UV} \rangle (M_{UV}^{dust}) - M_{UV}^{dust}$ relations at $z = 8$ for the different approaches described in the text. For the parameters of the $M_{UV}^{dust} - \langle \beta \rangle$ relation we employ the values of Bouwens et al. (2014a). The various methods lead to visible differences for the derived attenuation relations. *Bottom panel:* Derived $M_{UV}^{dust-free}$ to M_{UV}^{dust} mapping at $z = 8$ for the different approaches. Despite the differences in the $\langle A_{UV} \rangle (M_{UV}^{dust}) - M_{UV}^{dust}$ relations, we find that the mappings between dust-free rest-frame UV magnitudes and dust-attenuated magnitudes differ only slightly for the different approaches. We also compare the relations to results from Tacchella et al. (2013); Liu et al. (2016); Tacchella et al. (2018). Despite the small differences between the various mappings all approaches lead to the same effective dust attenuation correction of dust-free magnitudes.

is the main difference compared to the analytical approach. This method has, for example, been adopted in Smit et al. (2012); Liu et al. (2016).

(i) *Luminosity based average:* We first select 200 linearly spaced M_{UV}^{dust} values ranging from -27 mag to -13 mag. At each M_{UV}^{dust} or correspondingly L_{UV}^{dust} , we sample 10^5 β values drawn from a normal distribution centred around $\langle \beta \rangle (M_{UV}^{dust})$ with $\sigma_\beta = 0.34$. We then calculate the $L_{UV}^{dust-free}$ value for each individual sample using the IRX $-\beta$ relation. If the resulting $L_{UV}^{dust-free}$ is smaller than L_{UV}^{dust} , we set $L_{UV}^{dust-free} = L_{UV}^{dust}$; i.e. we set the attenuation to zero for these negative cases. Finally, we combine all the samples and divide them based on their $M_{UV}^{dust-free}$ values into 30 linearly spaced bins ranging from -25 mag to -15 mag. In each bin, the mean L_{UV}^{dust} of the samples within each bin is then calculated. The

M_{UV}^{dust} value corresponding to this mean L_{UV}^{dust} is then derived. This then finally results in an average mapping from $M_{UV}^{dust-free}$ to M_{UV}^{dust} .

(ii) *Magnitude based average:* Similar to the luminosity average approach, we first select 200 linearly spaced M_{UV}^{dust} values ranging from -27 mag to -13 mag. At each M_{UV}^{dust} , we sample 10^5 β values with a normal distribution centred at $\langle \beta \rangle (M_{UV}^{dust})$ with $\sigma_\beta = 0.34$. We then calculate the A_{UV} value for each individual sample with the relation $A_{UV} = C_0 + C_1 \beta$. Negative A_{UV} values are set to zero. We then calculate $M_{UV}^{dust-free}$ for each individual sample with $M_{UV}^{dust-free} = M_{UV}^{dust} - A_{UV}$. Finally, we combine all the samples and split them based on their $M_{UV}^{dust-free}$ values into 30 linearly spaced bins ranging from -25 mag to -15 mag. In each bin, the mean M_{UV}^{dust} of the samples within the bin is then calculated. This then results in an averaged mapping from $M_{UV}^{dust-free}$ to M_{UV}^{dust} .

In Figure 1, we compare the $\langle A_{UV} \rangle (M_{UV}^{dust}) - M_{UV}^{dust}$ relations (top panel) and the resulting mappings from $M_{UV}^{dust-free}$ to M_{UV}^{dust} (bottom panel) derived by the different approaches evaluated at $z = 8$. For the parameters of the $M_{UV}^{dust} - \langle \beta \rangle$ relation we employ the values of Bouwens et al. (2014a). The dust attenuation curves differ between the analytical and numerical sampling approach towards the faint end of the galaxy population. For the numerical sampling approach, samples with negative A_{UV} are set to zero attenuation which leads to a rise in $\langle A_{UV} \rangle$, especially towards the faint end where A_{UV} can easily drop below zero. The magnitude based and luminosity based averages also lead to differences mainly at the bright end. Luminosity based averages lead to stronger attenuation due to the extra term caused by the assumed scatter in β . Although these different approaches and methods result in differences in the derived $\langle A_{UV} \rangle (M_{UV}^{dust}) - M_{UV}^{dust}$ relations, the differences in the final magnitude mappings are negligible as demonstrated in the lower panel of Figure 1. In the following we will therefore adopt the analytic approach with luminosity based averages to correct our dust-free magnitudes for dust in Model A.

The advantage of this empirical dust model is that it follows from observational scaling relations, and only depends on the dust-free galaxy magnitudes as predicted by the simulation. At the same time however, this is also a weakness. Specifically, both the observed $M_{UV}^{dust} - \langle \beta \rangle$ relation and the observed IRX $-\beta$ relation still contain significant uncertainties. Bouwens et al. (2014a) provides currently the most complete observational dataset to derive the $M_{UV}^{dust} - \langle \beta \rangle$ relation with more than 4000 galaxies at $z = 4 - 8$. However, there have been several other studies of the $M_{UV}^{dust} - \langle \beta \rangle$ relation (Wilkins et al. 2011; Finkelstein et al. 2012; Bouwens et al. 2012; Dunlop et al. 2013; Rogers et al. 2013). Besides different sample selections, also the measurement of the β slopes differ between these studies, which causes significant differences of up to $\Delta\beta = 0.5$. After correcting for different systematic biases, Bouwens et al. (2014a) demonstrated that these different observational studies are roughly in agreement ($\Delta\beta < 0.3$) with each other. Furthermore, the underlying IRX $-\beta$ relation is also uncertain (Meurer et al. 1999). Specifically, no consensus has yet been reached regarding the detailed IRX $-\beta$ relation at $z \gtrsim 3$. Several studies (e.g., Reddy et al. 2010, 2012; Nordon et al. 2013; McLure et al. 2013; Koprowski et al. 2018; McLure et al. 2018; Reddy et al. 2018) have explored the IRX $-\beta$ relationship at high redshifts. For typical but little or modestly obscured systems, the results are broadly consistent with the Meurer et al. (1999) IRX $-\beta$ relation, but deviate from it for more IR luminous, highly obscured galax-

ies. However, there is still a debate on whether they support a gray attenuation curve (e.g., Calzetti et al. 2000; McLure et al. 2018; Koprowski et al. 2018) or are more Small Magellanic Cloud-like (e.g., Capak et al. 2015; Reddy et al. 2018). All these caveats have to be taken into account when using Model A. Given these observational inconsistencies, we will in the following treat the two parameters of the $M_{\text{UV}}^{\text{dust}} - \langle \beta \rangle$ relation, the intercept β_{M_0} and the slope $d\beta/dM_{\text{UV}}$, as adjustable parameters that can be derived through a calibration procedure described below. We have also tested explicitly that the observationally derived $M_{\text{UV}}^{\text{dust}} - \langle \beta \rangle$ relations lead to IllustrisTNG dust-attenuated galaxy luminosity functions that are inconsistent with observed rest-frame UV luminosity functions. This also motivates a calibration based approach for dust Model A.

3.2.2 Model B – resolved dust optical depth model

For our second dust model we explicitly consider the resolved gas properties around every stellar particle to derive the dust attenuation based on an estimate of the dust optical depth of the surrounding gas. We note that this model does not include dust emission, which mainly affects the rest-frame SED of galaxies in far-IR and therefore has no influence on rest-frame UV or other band magnitudes that we study in this paper. In addition to the resolved component, this second dust model also includes an unresolved dust component for dust attenuation. This unresolved component is required due to the limited numerical resolution of our simulations. Below we describe the implementation of both dust components in detail. This method is similar to the fiducial dust-model C of Nelson et al. (2018) which was used to compare IllustrisTNG colour distributions versus SDSS at $z \sim 0$.

Resolved dust: To calculate the resolved dust component, we first map the gas distribution around each galaxy onto a cubic grid. This cube has a side length of 60 pkpc centred on the most bound particle of each subhalo hosting the galaxy. This physical side length is sufficiently large to cover most of the gas that contributes to dust attenuation at all redshifts. In the following we observe each galaxy along the z axis. The cubic grid has a pixel size of 1 pkpc in the x - y plane and samples the z direction with a spacing of 0.1 pkpc. We verified that the resulting luminosities and magnitudes do not change if we, for example, increase the x - y plane pixel size to 2 pkpc. We assume that dust is only forming in cold phases of the interstellar medium (ISM) and gas cells therefore need to be either star-forming or cold enough with temperature < 8000 K to contribute to the dust grid values. The hydrogen mass and gas metallicity within each grid cell are then calculated based on this criterion. For each stellar particle we then find its z -location on the grid and calculate accumulated cold gas values of all the grid cells in front of the stellar particle along the line-of-sight direction. Specifically, we calculate the hydrogen column density N_{H} and hydrogen mass-weighted gas metallicity Z_{g} . Finally, we compute the V band (547.7 nm) resolved optical depth for this stellar particle as (Guideroni & Rocca-Volmerange 1987; Devriendt et al. 1999; Devriendt & Guideroni 2000; De Lucia & Blaizot 2007; Kitzbichler & White 2007; Fontanot et al. 2009; Guo & White 2009; Clay et al. 2015; Nelson et al. 2018):

$$\tau_{\text{V}}^{\text{res}} = \tau_{\text{dust}}(z) \left(\frac{Z_{\text{g}}}{Z_{\odot}} \right)^{\gamma} \left(\frac{N_{\text{H}}}{N_{\text{H},0}} \right), \quad (10)$$

where $\gamma = 1$, $Z_{\odot} = 0.0127$ and $N_{\text{H},0} = 2.1 \times 10^{21} \text{ cm}^{-2}$ are normalisation values for gas metallicity and hydrogen column density, respectively. In Guideroni & Rocca-Volmerange (1987), γ is typically wavelength dependent with two regimes interpolating

extinction curves between the solar neighbourhood and the Magellanic clouds. Here, we use this relation to derive the V band optical depth in systems with different metallicity. Guideroni & Rocca-Volmerange (1987) and some later works set $\gamma = 1.6$ for $\lambda > 200$ nm. However, in Somerville et al. (2012); Yung et al. (2018), the optical depth for the V band is assumed to linearly depend on the metal mass in the galaxy disc, which is equivalent to $\gamma = 1$. Wilkins et al. (2017) also assumed that the dust optical depth depends linearly on metal column density. Ma et al. (2018b), following the approach in Hopkins et al. (2005), also scaled dust attenuation linearly with metallicity. We have tested explicitly that different γ values do not lead to significant differences in the resulting luminosity functions. We therefore adopt $\gamma = 1$ to derive the V band optical depth in the following. This then also implies that the redshift dependent scale factor $\tau_{\text{dust}}(z)$ scales like the average dust-to-metal ratio. We note that there are also some variations in the literature in terms of the employed gas density and metallicity values entering the optical depth relation above. For example, Nelson et al. (2018) used the neutral hydrogen density and neutral hydrogen mass weighted metallicity instead of hydrogen density and gas metallicity in Equation 10. However, they studied only the low redshift galaxy population. Other studies (e.g., Guideroni & Rocca-Volmerange 1987; Devriendt et al. 1999; Devriendt & Guideroni 2000; Clay et al. 2015), referred to N_{H} as hydrogen column density, while Guo & White (2009) referred to it as neutral hydrogen column density. Kitzbichler & White (2007) referred to the term used in Devriendt et al. (1999) as neutral hydrogen column density, while they referred to their N_{H} as hydrogen column density. As stated above, we employ here the hydrogen column density for calculating the V band optical depth.

Next, we have to translate the V band optical depth to an actual dust attenuation. The attenuation depends on the geometry of dust and stars. In the following, we employ the dust geometry model described in Calzetti et al. (1994), where the ionised gas and dust are co-spatial and uniformly mixed. We note that we calculate here dust attenuation and not extinction. Extinction only considers the removal of photons from the line of sight, both absorption and scattering. The light comes from a background point source and the dust is entirely foreground to the source. Extinction therefore is insensitive to the detailed spatial distribution of the foreground dust. However, attenuation refers to the situation where the radiation sources are distributed within the dust with a range of depths. Both radiation sources and dust have an extended and complex spatial distribution. The photons will not only be removed from a given line of sight, they can also be scattered into the line of sight from other points in the extended source. To calculate the dust attenuation, the radiative transfer equations have to be solved to derive the relation between observed luminosity and intrinsic luminosity. Using the solution of the radiative transfer equation for a homogeneous mixture of stars and dust, the V band attenuation can be written as (Calzetti et al. 1994):

$$A_{\text{V}}^{\text{res}} = -2.5 \log \left(\frac{L_{\text{V}}^{\text{dust}}}{L_{\text{V}}^{\text{dust-free}}} \right) = -2.5 \log \left(\frac{1 - e^{-\tau_{\text{V}}^{\text{res}}}}{\tau_{\text{V}}^{\text{res}}} \right). \quad (11)$$

Besides specifying the V band optical depth and attenuation, we also need to assume a certain functional form for the attenuation curve to calculate the dust attenuation at other wavelengths. This curve depends on details of the underlying dust distribution, like density, grain size distribution and chemical composition. In the absence of detailed knowledge of these quantities, we have to assume a certain attenuation curve. The Calzetti et al. (2000) attenuation curve is commonly adopted for the study of high redshift galaxies.

This curve is shown in Figure 2 along with other extinction and attenuation curves. The Calzetti relation was first proposed for local starburst galaxies and can be described as (Calzetti et al. 2000):

$$A_{\text{V}}^{\text{res}}(\lambda) = A_{\text{V}}^{\text{res}} \left(\frac{k'(\lambda)}{4.05} \right), \quad (12)$$

where $k'(\lambda)$ is the attenuation curve normalised to $A_{\text{V}} = 4.05$:

$$k'(\lambda) = 4.05 + 2.659 \begin{cases} \left(-1.857 + \frac{1.040}{\lambda} \right), & \text{for } 0.63 \mu\text{m} < \lambda < 2.20 \mu\text{m} \\ \left(-2.156 + \frac{1.509}{\lambda} - \frac{0.198}{\lambda^2} + \frac{0.011}{\lambda^3} \right), & \text{for } 0.12 \mu\text{m} < \lambda < 0.63 \mu\text{m}. \end{cases} \quad (13)$$

Attenuation outside of these wavelength ranges can be derived through extrapolation. We note that the Calzetti attenuation law does not include the 217.5 nm UV bump feature in dust attenuation, which is non-negligible for high redshift galaxies (e.g., Kriek & Conroy 2013; Scoville et al. 2015). In addition, the slope of the attenuation curve is observed to vary among different galaxies or sight lines, which imprints non-universal features in dust attenuation (e.g., Noll et al. 2009; Kriek & Conroy 2013; Salmon et al. 2016). Therefore Kriek & Conroy (2013) proposed a modified version of the Calzetti law to account for the UV bump feature and the change in slope:

$$A_{\text{V}}^{\text{res}}(\lambda) = \frac{A_{\text{V}}^{\text{res}}}{4.05} [k'(\lambda) + D(\lambda)] \left(\frac{\lambda}{\lambda_{\text{V}}} \right)^{\delta}, \quad (14)$$

where δ is the correction of the steepness of the attenuation curve and $D(\lambda)$ is a Lorentzian-like Drude profile to parameterise the UV bump feature defined as:

$$D(\lambda) = \frac{E_{\text{b}}(\lambda \Delta\lambda)^2}{(\lambda^2 - \lambda_0^2)^2 + (\lambda \Delta\lambda)^2}, \quad (15)$$

where E_{b} is the parameter defining the strength of the UV bump, $\lambda_0 = 217.5$ nm is the central wavelength of the UV bump, and $\Delta\lambda = 35$ nm is the FWHM of the bump (Seaton 1979; Noll et al. 2009). Kriek & Conroy (2013) found a relation between E_{b} and δ :

$$E_{\text{b}} = (0.85 \pm 0.09) - (1.9 \pm 0.4) \delta. \quad (16)$$

Therefore the shape of the attenuation curve is purely characterised by δ . In Buat et al. (2011), 30 galaxies at $1 < z < 2$ were studied and they derived an averaged attenuation curve with $\delta = -0.13$, $E_{\text{b}} = 1.26$ and $\Delta\lambda = 35.6$ nm. For simplicity, we assume $\delta = 0$ in this work, which implies no correction to the slope of the Calzetti attenuation curve while introducing the UV bump feature.

We then finally get $A^{\text{res}}(\lambda)$ for each stellar particle in the galaxy as:

$$A_{\text{V}}^{\text{res}}(\lambda) = \frac{A_{\text{V}}^{\text{res}}}{4.05} [k'(\lambda) + D(\lambda)]. \quad (17)$$

To calculate the dust attenuated magnitudes of an arbitrary band, we determine the attenuation at the effective wavelength of the band, $A(\lambda_{\text{eff}})$, to represent the attenuation in the entire wavelength range covered by a given filter:

$$M^{\text{dust}} = M^{\text{dust-free}} + A^{\text{res}}(\lambda_{\text{eff}}), \quad (18)$$

where M^{dust} and $M^{\text{dust-free}}$ are the dust-attenuated and dust-free magnitudes of an arbitrary band, respectively. We have verified that Equation 18 is a good approximation of the detailed procedure applying the attenuation to each wavelength and then convolving with

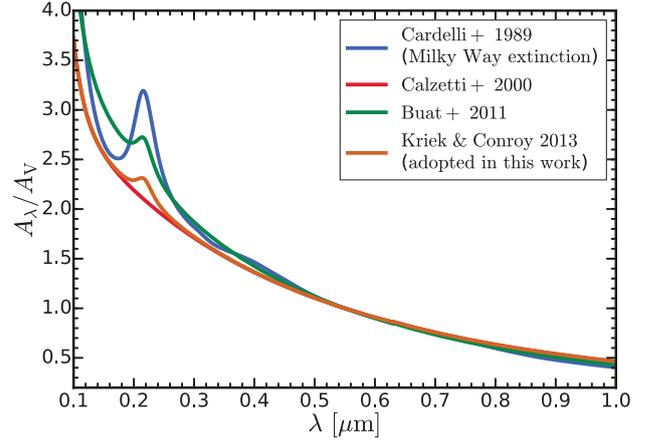


Figure 2. Different dust attenuation and extinction curves. The Milky Way extinction curve (Cardelli et al. 1989) shows a strong UV feature. The Calzetti attenuation curve (Calzetti et al. 2000) is based on observations of local starburst galaxies. The attenuation curve we adopt for our work (Kriek & Conroy 2013) adds a parametrised slope correction and UV feature on top of the Calzetti attenuation curve (Noll et al. 2009). We set the slope correction δ to zero while keeping the UV feature. Based on observations, Buat et al. (2011) derived the sample averaged parameters for this attenuation curve. They followed the same attenuation parametrisation as Kriek & Conroy (2013); Noll et al. (2009) but determined the parameters through average values of 30 galaxies at in the redshift range $1 < z < 2$.

the filter transmission function. For example, for a SSP with initial stellar mass of $10^{10} M_{\odot}$ and solar metallicity and applying the Calzetti attenuation curve with $A_{\text{V}} = 0.2$, we find that the rest-frame magnitude differences between these two approaches are ≤ 0.01 in all SDSS optical bands and the GALEX FUV band.

Unresolved dust: Apart from the resolved dust attenuation of the ISM, there are also dust components which cannot be resolved by our simulations due to the limited numerical resolution. One major component that cannot be resolved by our simulations are the birth clouds surrounding young stars, which can cause significant dust attenuation. We also have to account for this dust attenuation contribution in addition to the resolved dust component. It is common to include this contribution using the power law extinction model presented in Charlot & Fall (2000). This model specifically considers the attenuation contribution from birth clouds around young stars. Many semi-analytic and hydrodynamical works have adopted this model to extend their resolved ISM attenuation model to smaller scales to add the attenuation contribution from birth clouds. The original Charlot & Fall (2000) model includes the attenuation by finite lifetime birth clouds surrounding young stellar populations as well as the ambient diffuse ISM. According to this model the V band optical depth can then be described as a combination of these two contributions:

$$\tau_{\text{V}} = \begin{cases} (\tau_1 + \tau_2), & \text{for } t_{\text{ssp}} \leq t_{\text{esc}} \\ \tau_2, & \text{for } t_{\text{ssp}} > t_{\text{esc}}, \end{cases} \quad (19)$$

where τ_1 is the optical depth of unresolved birth clouds around young stars, τ_2 is the optical depth of the ambient ISM gas and t_{esc} is an age criterion defining young stars associated with birth clouds. Stars older than t_{esc} can escape the birth clouds and their light is not subject to attenuation due to the birth cloud environment anymore.

Our dust model already includes the resolved ISM compo-

nent as described above. However, it does not include an attenuation contribution from the birth cloud environment. We therefore have to add this unresolved birth cloud component. This approach has also been commonly employed in other studies (e.g., Guo & White 2009; Somerville et al. 2012; Clay et al. 2015). We note that sometimes only the resolved ISM contribution to dust is considered neglecting the birth cloud attenuation (Kitzbichler & White 2007; Yung et al. 2018). For example, Yung et al. (2018) states that the addition of attenuation due to birth clouds leads to an overprediction of attenuation at high redshifts. Some models also include a resolved ISM component along with the birth cloud and ambient ISM combination (Nelson et al. 2018).

To apply the model for the birth cloud dust attenuation contribution, we have to know the optical depths values entering Equation 19. At $z = 0$, commonly adopted values are: $\tau_1 = 0.7$, $\tau_2 = 0.3$, $t_{\text{esc}} = 10$ Myr (e.g., Trayford et al. 2015; Byler et al. 2017; Nelson et al. 2018). At higher redshifts some works have assumed certain redshift dependencies of these parameters. For example, Somerville et al. (2012) assumed a $(1+z)^{-1}$ ($1/[(1+z)z]$ for $z > 1$) redshift dependence for the optical depth τ_1 and $(1+z)^{-1}$ for the optical depth τ_2 . These redshift dependencies were chosen to match the observed $L_{\text{IR}}/L_{\text{bol}}$ relation and the bright end of the observed UV luminosity functions at $z = 0.5-5$. They also assume a redshift dependence for the age criteria for young stellar populations, t_{esc} , while most other studies set this to a constant value of 10 Myr at all redshifts. Guo & White (2009), on the other hand, assumed that the optical depth of young stellar populations is three times that of the old stellar population, which is suggested by Charlot & Fall (2000). They then chose $(1+z)^{-0.4}$ for the redshift dependence of the optical depths. Trayford et al. (2015) adopted the same ratio between the optical depth of ISM and birth cloud. Cullen et al. (2017) found in their best-fit model that the dust optical depth of old stellar populations is 0.3 times that of young stellar population. They also found t_{esc} to be around 12 Myr. De Lucia & Blaizot (2007) and Clay et al. (2015) assumed that the additional optical depth of dust in birth clouds scales linearly with the resolved dust optical depth. The linear scale factor is assumed to have some random variation around the median value of 2.33 for different galaxies. They have employed a $(1+z)^{-1}$ redshift scaling for the optical depths values.

After reviewing existing unresolved dust models, we now describe our specific implementation of unresolved dust. First, we discard the attenuation of diffuse ISM from the Charlot & Fall (2000) model since this contribution is already included in our resolved ISM component. This leaves us with an additional attenuation due to the birth cloud environment that we have to implement for the unresolved dust component. In the absence of a detailed physical model for the dust abundance in birth clouds, we assume that the additional optical depth of birth clouds is two times (Charlot & Fall 2000; Kong et al. 2004; De Lucia & Blaizot 2007; Guo & White 2009; Henriques et al. 2015; Clay et al. 2015) the mean optical depth of resolved dust averaged over all the stellar particles in the galaxy:

$$\tau_{\text{V}}^{\text{unres}} = \begin{cases} 2 \langle \tau_{\text{V}}^{\text{res}} \rangle, & \text{for } t_{\text{spp}} \leq t_{\text{esc}} \\ 0, & \text{for } t_{\text{spp}} > t_{\text{esc}}, \end{cases} \quad (20)$$

where $t_{\text{esc}} = 10$ Myr, and $\langle \tau_{\text{V}}^{\text{res}} \rangle$ is the average optical depth of the resolved dust component. This mean value is derived by a simple average of $\tau_{\text{V}}^{\text{res}}$ over all stellar particles that contribute to the stellar light of the galaxy. Our complete dust model therefore has only

one free redshift-scaling parameter, which is the redshift-dependent normalisation of the optical depth $\tau_{\text{dust}}(z)$.

Under a simple uniform dust screen model the solution of the radiative transfer equation can then be used to derive the V band attenuation based on this unresolved optical depth Charlot & Fall (2000):

$$A_{\text{V}}^{\text{unres}} = -2.5 \log \left(\frac{L_{\text{dust}}}{L_{\text{dust-free}}} \right) = -2.5 \log \left(e^{-\tau_{\text{V}}^{\text{unres}}} \right). \quad (21)$$

Here the dust screen lies between the observer and source and the source is treated as point source. We note that the equation above simply evaluates to $A_{\text{V}}^{\text{unres}} = 1.086 \tau_{\text{V}}^{\text{unres}}$. We adopt a power law attenuation curve for unresolved dust with power law index -0.7 , the best-fit value in Charlot & Fall (2000), such that the attenuation at other wavelengths is given by:

$$A^{\text{unres}}(\lambda) = A_{\text{V}}^{\text{unres}} \left(\frac{\lambda}{\lambda_{\text{V}}} \right)^{-0.7}. \quad (22)$$

Similar to the resolved dust model, we then have:

$$M^{\text{dust}} = M^{\text{dust-free}} + A^{\text{unres}}(\lambda_{\text{eff}}). \quad (23)$$

This then fully specifies the unresolved dust component.

Combined resolved and unresolved dust attenuation: Given the dust-free $M^{\text{dust-free}}$ magnitude of a galaxy, we combine the attenuation from the resolved and unresolved dust component to calculate the final observed dust-corrected band magnitude as:

$$M^{\text{dust}} = M^{\text{dust-free}} + A^{\text{unres}}(\lambda_{\text{eff}}) + A^{\text{res}}(\lambda_{\text{eff}}), \quad (24)$$

where λ_{eff} is the effective wavelength of the band filter. This combined dust attenuation fully specifies our dust Model B.

3.2.3 Model C – dust radiative transfer

For the final dust model we consider the distribution of stars and dust in the simulated galaxies to perform full Monte Carlo continuum dust radiative transfer calculations using a modified version of the publicly available SKIRT code (Baes et al. 2011; Camps et al. 2013; Camps & Baes 2015; Saftly et al. 2014). SKIRT traces dust scattering and absorption for randomly sampled photon packets until they reach a detector. This results in integrated fluxes for each simulated galaxy. Based on these fluxes we then calculate the rest-frame magnitudes for each band through a convolution with the corresponding SEDPY transmission functions. In the following we describe the different components and steps that are required for the radiative transfer calculation. This method of using SKIRT post-processing radiative transfer is similar to that recently developed in Rodriguez-Gomez et al. (2018) to compare IllustrisTNG and PanSTARRS optical galaxy morphologies.

Stellar SEDs: Stellar particle SEDs are modelled with FSPS to be consistent with the SEDs employed in our other two dust models. To this end we have modified the SKIRT code in the following way. We first use FSPS to generate a series of dust-free SED templates as described in the previous sections. These SED templates have 5994 wavelength points covering a wavelength range from 9.1 nm to 1 cm for the adopted MILES stellar library. Then we modify the “FSPSVariableSEDFamily” class in SKIRT to implement this family of template SEDs. Based on the initial mass, metallicity and age of a given stellar particle, SKIRT then generates the SED of the particle through interpolation over those SED templates. Doppler shifts of spectra due to the motion of stellar particles are not included here.

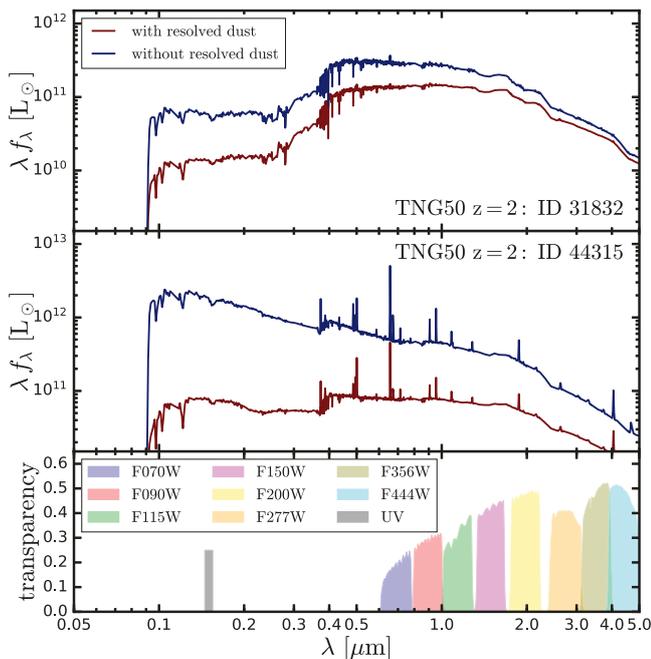


Figure 3. Spectral energy distribution for two example galaxies from TNG50 at $z = 2$. We show the intrinsic spectral energy distribution (blue) and the dust attenuated distribution (red) using SKIRT assuming a dust-to-metal ratio of 0.9, which represents the best-fit value of our dust parameter calibration procedure described in the text. The bottom panel shows the transmission functions of the different band filters employed in this work: UV band filter, and the eight wide *JWST* NIRCcam filters ($\lambda/\Delta\lambda \sim 4 - 5$).

Wavelength grid construction: SKIRT employs a wavelength grid that stores the wavelength values for emitted photons. At each of those wavelengths SKIRT then emits a prescribed number of photon packets. The output flux is also sampled by this wavelength grid. For the purpose of our high redshift luminosity function predictions between $z = 2$ and $z = 10$, it is sufficient to have a wavelength grid between $0.05 \mu\text{m}$ and $5 \mu\text{m}$. This covers all relevant *JWST* NIRCcam wide filters at all redshifts, and the rest-frame UV band. The SEDs with nebular emission have very narrow and sharp emission line features. We have therefore designed our wavelength grid in a specific way to account for this. We start from the wavelength grid of the input FSPS SED templates. This high resolution grid has 5164 sampled points between wavelengths of $0.05 \mu\text{m}$ and $5 \mu\text{m}$. In order to efficiently run SKIRT, we need to reduce the number of wavelength points without losing important spectral features like emission lines. We therefore first select every 12th grid point on this grid resulting in a 12 times lower resolution grid. In a next step, we select additional 11 points around each emission line to ensure that those are not lost during the grid down-sampling process. The exact wavelengths of these emission lines are taken from the FSPS datafile. Finally, in order to make accurate predictions for UV magnitudes, we keep all the wavelength grid points between 140 nm to 180 nm , which includes the coverage of our UV filter from the original fine grid. This procedure results in a wavelength grid with 1168 grid points covering wavelengths from $0.05 \mu\text{m}$ to $5 \mu\text{m}$, which suits our need to efficiently and accurately calculate galaxy fluxes and photometric properties.

Source distribution: Next we have to provide SKIRT with smoothing length values for the stellar particles representing the ra-

diation sources. This is required to calculate internally a smoothed photon source distribution function. We employ a K-D tree algorithm to calculate the smoothing length enclosing 64 ± 1 nearest stellar particles for all stellar particles within the galaxy. Given the spatial location and smoothing length values of stellar particles, SKIRT then creates a photon source distribution and emissivity profile through the entire space by interpolating over these kernels. For each wavelength on the wavelength grid, N_p photon packets are then launched isotropically, for which the launch positions are determined by the luminosity distribution given by the smoothed stellar particle distribution at that wavelength. The photon packets then propagate through the resolved ISM and interact with the dust cells randomly before they are finally collected by the photon detector. We set N_p equal to the number of all the stellar particles bound to a galaxy. This number is by construction larger than the number of stellar particles within the 30 pkpc aperture. In addition we apply $N_p = 10^5$ as an upper and $N_p = 10^2$ as a lower limit for the number of generated photon packets. We have tested the convergence of radiative transfer calculations on selected galaxies in TNG50 with M_* ranging from 10^7 to $10^{11} M_\odot$. Doubling the number of photon packets leads only to minor, $< 0.01 \text{ mag}$, changes in all rest-frame UV and optical bands.

Dust properties: The position, gas density and metallicity of the Voronoi gas cells from the simulation are used to calculate the distribution of dust in galaxies. We assume a spatially constant dust-to-metal ratio at each redshift to link the cold gas density and metallicity with the dust density:

$$\rho_{\text{dust}} = f_z(z) Z_g \rho_g, \quad (25)$$

where Z_g and ρ_g are the metallicity and density of the gas cell, and $f_z(z)$ is a redshift dependent dust-to-metal ratio, which has to be calibrated in the absence of a dust model to predict self-consistently the dust distribution within galaxies. Following Model B, only star-forming gas cells or gas cells colder than $< 8000 \text{ K}$ contribute to the ISM dust budget. Gas cells that do not meet any of these criteria have their dust-to-metal ratio set to zero. Dust emission is self-consistently included in this model and we assume a *Draine & Li (2007)* dust mixture of amorphous silicate and graphitic grains, including varying amounts of Polycyclic Aromatic Hydrocarbons (PAHs) particles. This model reproduces the average Milky Way extinction curve and is widely used (e.g., *Jonsson et al. 2010; Baes et al. 2011; Kimm & Cen 2013; Rémy-Ruyer et al. 2014*). The model is an update of the *Li & Draine (2001)* model to adjust PAHs to match early Spitzer results. The dust mass distribution is then mapped to an adaptively refined grid on which the optical depth at each wavelength is calculated. Specifically, the space is discretized by an octree algorithm and the root cell is a cube with sidelength of 60 pkpc centred on the position of the most bound particle of the subhalo. The octree is then refined between a minimum refinement level of 3 and a maximum refinement level of k . The minimum octree cell size should be smaller than the physical spatial resolution of the simulation, $60 \text{ pkpc}/2^k \leq \epsilon_{\text{gas,phy}}^{\text{min}}$. For TNG50, the minimum comoving gravitational softening length is $\epsilon_{\text{gas}}^{\text{min}} = 0.074 \text{ kpc}$ comoving, which correspondingly requires a maximum refinement level $k = 14$ at $z = 10$. For TNG100, $\epsilon_{\text{gas}}^{\text{min}} = 0.185 \text{ kpc}$ comoving and requires $k = 12$ at $z = 10$. The cell splitting criterion is set to 2×10^{-6} . This sets the maximum fraction of total dust mass that can be contained in a single dust cell. If the dust mass fraction within a dust cell is larger than this value and the maximum refinement level has not been reached, the cell will be subdivided further.

Magnitude Calculation: To finally calculate galaxy magnitudes, we place a detector 1 Mpc away from the simulated galaxy

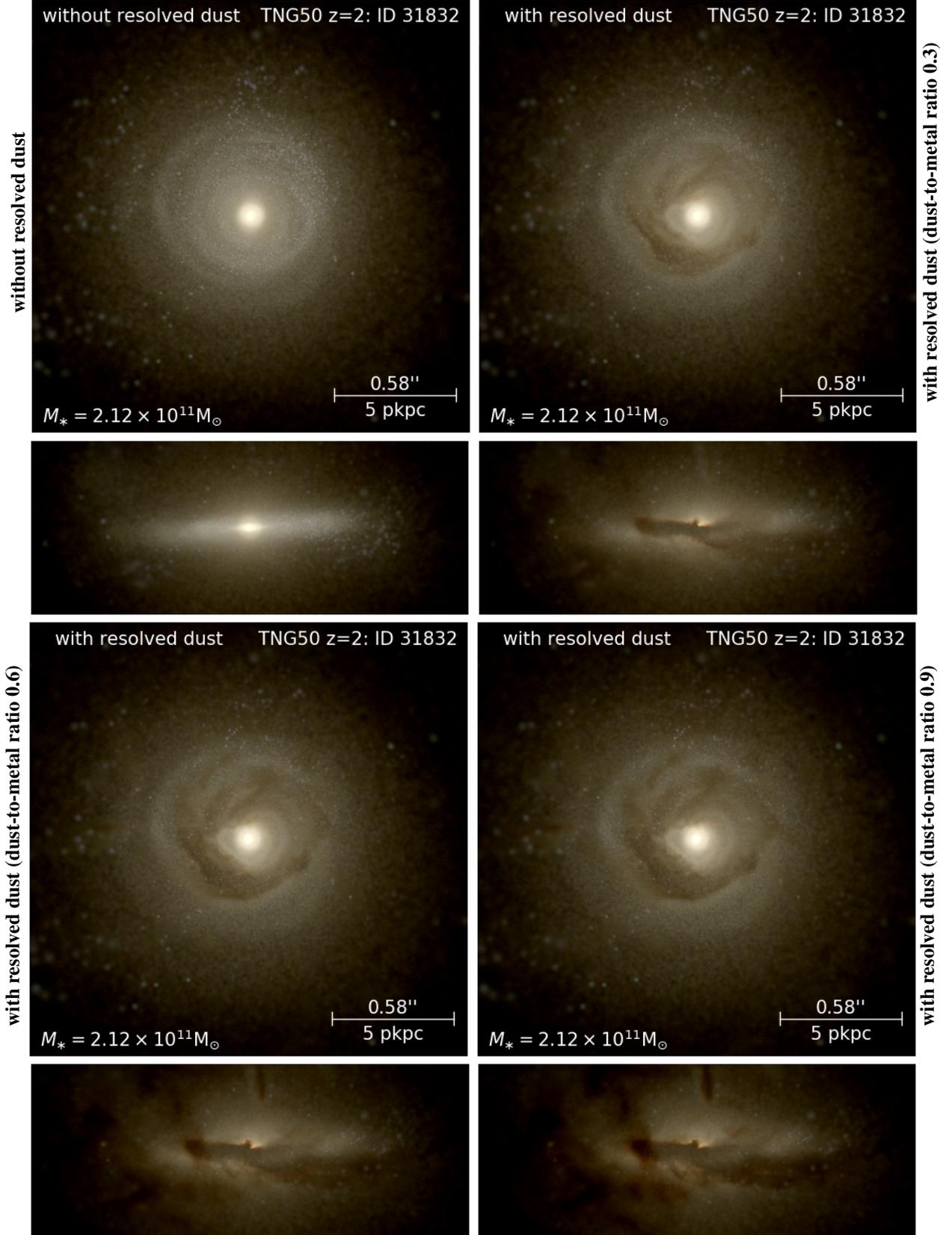


Figure 4. *JWST* NIRCam face-on and edge-on images of a TNG50 galaxy at $z = 2$. *Top left panel:* Galaxy images without resolved dust. *Other panels:* Images with resolved dust for three dust-to-metal ratios: 0.3 (upper right), 0.6 (lower left) and 0.9 (lower right). The images are based on apparent F115W, F150W, F200W filter fluxes of the galaxy ID 31832. The SED of this galaxy is shown in Figure 3. The images cover a $18 \text{ pkpc} \times 18 \text{ pkpc}$ field of view with 500×500 pixels. The radiative transfer employs 10^7 photon packets per wavelength on the reduced wavelength grid with 234 points covering $0.05 \mu\text{m}$ to $5 \mu\text{m}$. Our fiducial SED grid employs a finer grid with 1168 points. The central bright bulge region is covered by dust absorption when observed in the edge-on view.

along the positive z -direction to collect emitted photons. The integrated galaxy flux is then generated by SKIRT and then convolved with the transmission curve of the band transmission curves using the SEDPY code. For the calculation of apparent magnitudes, the rest-frame flux is redshifted and converted to observed spectra. In addition, we also correct for IGM absorption (Madau 1995; Madau et al. 1996) as described in Section 3.1.

Unresolved components: Similar to Model B, Model C also suffers from the coarse numerical resolution of the simulation. Specifically, this limited resolution is problematic for the modelling of the young stellar component. The treatment of the young stellar component is affected by two limitations of the simulation: (i) the relatively coarse sampling of star formation due to the limited mass resolution, and (ii) the inability of resolving birth cloud absorption associated with recent star formation. In principle, a single newly formed stellar particle or a star-forming gas cell should represent an ensemble of giant HII regions with stars formed inside of them with different formation times. Young and massive stars are then surrounded by dense birth clouds. Consequently, their dust-free stellar light emission will be reprocessed by a variety of physical processes. HII regions surrounding stellar populations reprocess the Lyman continuum photons into nebular continuum and line emissions. Furthermore, the unresolved dust in HII and photodissociation regions absorbs the UV/optical light and reprocesses it to IR.

The coarse sampling of star formation can be alleviated through a resampling technique (Trayford et al. 2017). However, this introduces a few additional free parameters for the modelling. We therefore follow another approach here by exploiting the high mass resolution of TNG50, which is about a factor of 15 and 120 higher than TNG100 and TNG300, respectively. Therefore, the TNG50 results are expected to be less affected by the problem of a coarse sampling of star formation given its high mass resolution. Using TNG50 we then apply a resolution correction to TNG100 and TNG300 to match the mass resolution of TNG50. This resolution correction procedure, as described in Section 3.3, alleviates the problem of a too coarse sampling of star formation in TNG100 and TNG300 without introducing additional free parameters.

The second problem, i.e. the inclusion of unresolved dust attenuation associated with young stellar populations, requires a modification of the stellar particles SEDs. Stellar particles with ages greater than 10 Myr are still modelled with FSPS spectra templates while stellar particles younger than 10 Myr are modelled with the MAPPINGS-III spectra library (Groves et al. 2008). These are based on 1D photoionization and radiative transfer calculations to model the radiative transfer of radiation from a newly formed massive stellar cluster with a Kroupa (2001) initial stellar mass function, using STARBURST99 (Leitherer et al. 1999) spectra for the dust-free stellar light emission. The radiation propagates through the surrounding spherically symmetric HII and photodissociation regions, whose covering fractions decrease as a function of time as it is cleared away by the strong winds from the massive cluster. We note that the use of the STARBURST99 population synthesis model and Kroupa initial stellar mass functions (Kroupa 2001) causes some inconsistency in our approach, since old particles are modelled with FSPS and a Chabrier initial stellar mass function. However, the resulting differences are small in the UV and optical ranges. Gonzalez-Perez et al. (2014) found that the differences in NIR are significant between different population synthesis models due to variations in the treatment of the thermally pulsating asymptotic giant branch stars, which should in principle only be important for older stellar population. We have explicitly

compared the SEDs of galaxies with young stellar population modelled by FSPS versus MAPPINGS-III and found no difference in the NIR range. We are therefore confident that the inconsistency in the population synthesis models does not have a significant impact on our results for all bands and all redshifts. The final output SEDs then include the stellar light attenuated by dust in the HII and photodissociation regions, and the corresponding dust, nebular continuum and line emissions. We note that the FSPS nebular emission model adopted in Model B is no longer employed here since those are already modelled through the MAPPINGS-III calculation.

The MAPPINGS-III SED of a specific stellar particle is controlled by five parameters that we specify in the following: *Star Formation Rate:* The star formation rate is required because the MAPPINGS-III SED templates are normalised to $\text{SFR} = 1 M_{\odot}/\text{yr}$. We assume $\text{SFR} = m_i/10 \text{ Myr}$ to maintain conservation of stellar mass, where m_i is the initial mass of the stellar particle. *Metallicity of the birth cloud:* The metallicity of the birth cloud is assumed to be the initial metallicity of the stellar particle, representing the metallicity of its birth environment. *Pressure of the surrounding ISM:* This pressure can be calculated based on the density of the star-forming gas. However, as demonstrated in Groves et al. (2008), the impact of pressure on UV to NIR spectra is very small. We therefore simply employ a fixed typical value of $\log_{10}(P_0/k_B)/\text{cm}^{-3}\text{K} = 5$ (Groves et al. 2008; Rodriguez-Gomez et al. 2018). *The compactness of the birth cloud:* This describes the density of the HII region and can be calculated as $\log_{10} C = 3/5 \log_{10} M_{\text{cl}}/M_{\odot} + 2/5 \log_{10} P_0/k_B/(\text{cm}^{-3}\text{K})$ (Groves et al. 2008), where M_{cl} is the mass of the stellar cluster and P_0 is the pressure of HII region. The compactness parameter mainly affects the temperature of the dust and its emission in FIR. It has almost no impact on optical/UV/NIR photometric properties (Groves et al. 2008). We therefore also fix this to a typical value of $\log_{10} C = 5$ (Groves et al. 2008; Rodriguez-Gomez et al. 2018). *Photodissociation regions covering fraction:* This defines the time-averaged fraction of the stellar cluster solid angle that is covered by the photodissociation regions. The MAPPINGS-III spectra library includes SEDs for the case of a fully covered or uncovered HII region. The final output SED is a linear combination of these two extreme cases controlled by the parameter f_{PDR} . We adopt here the fiducial value 0.2 for f_{PDR} (Groves et al. 2008; Jonsson et al. 2010; Rodriguez-Gomez et al. 2018). These parameters then fully specify the MAPPINGS-III SEDs and therefore our unresolved dust attenuation model for Model C.

In Fig 3, we demonstrate the impact of dust on two example rest-frame galaxy SEDs. The galaxies are selected at $z = 2$ from the TNG50 simulation, which has the highest baryonic mass resolution of all IllustrisTNG simulations. The dust-to-metal ratio is set to 0.9 for the generation of the SEDs for these two galaxies. This dust-to-metal ratio is the best-fit value of our dust calibration procedure described below. The SED of galaxy ID 44315 (middle panel) has a much stronger UV flux than galaxy ID 31832 (top panel) and the impact of dust attenuation is also stronger. We present face-on and edge-on images of galaxy ID 31832 in Figure 4. In the top left panel we present the galaxy without resolved dust attenuation. The remaining three panels present images with resolved dust attenuation. For the upper right panel we employ a dust-to-metal ratio of 0.3, for the lower left panel we use 0.6, and finally for the lower right panel we employ a dust-to-metal ratio of 0.9, which is the best-fit value of our calibration procedure at this redshift as we will describe below. We note that we include the unresolved dust component for all cases since this is intrinsic to the employed MAPPINGS-III spectral templates. These images are based on apparent F115W, F150W,

redshift z	Model A			Model B	Model C
	M_0 [mag]	β_{M_0}	$d\beta/dM_{UV}[\text{mag}^{-1}]$	τ_{dust}	dust-to-metal ratio
2	-19.5	-1.38	-0.28	0.46	0.90
3	-19.5	-1.63	-0.26	0.20	0.41
4	-19.5	-1.80	-0.30	0.13	0.24
5	-19.5	-2.03	-0.29	0.08	0.15
6	-19.5	-2.32	-0.47	0.06	0.11
7	-19.5	-2.53	-0.40	0.04	0.08
8	-19.5	-2.66	-0.34	0.03	0.06

Table 3. Best-fit dust parameters for dust Models A, B, and C. For Model A we quote the parameters for the $M_{UV}^{\text{dust}} - \langle\beta\rangle$ relation: $M_0, \beta_{M_0}, d\beta/dM_{UV}$. We note that we keep $M_0 = -19.5$ mag constant and assume an underlying Meurer et al. (1999) relation for the Model A fits. For Model B we quote the optical depth normalisation: τ_{dust} . For Model C we quote the dust-to-metal ratios. These best-fit parameters are all redshift dependent and then uniformly applied to all galaxies at these redshifts.

F200W NIRCcam wide filter fluxes. They cover a $18 \text{ pkpc} \times 18 \text{ pkpc}$ field of view with 500×500 pixels. The SKIRT radiative transfer calculation employed 10^7 photon packets per wavelength on a reduced wavelength grid with 234 grid points in the wavelength range from $0.05 \mu\text{m}$ to $5 \mu\text{m}$. We stress that this grid is still constructed to cover all nebular emission lines reliably. The inclusion and sufficient sampling of these emission lines is crucial for a correct modelling of the appearance of galaxies (e.g., Camps et al. 2016; Trayford et al. 2017). Figure 4 clearly demonstrates the important role of dust. In fact, the central bright bulge region of the galaxy is nearly completely covered by dust absorption when observed from the edge-on perspective. Even for a low dust-to-metal ratio of 0.3 the appearance of the galaxy is dramatically different from the image without resolved dust attenuation. Any analysis or predictions based on the stellar light emission of simulated galaxies, e.g. UV, optical or lines like $H\alpha$, have to take dust attenuation into account. Neglecting this effect, can significantly alter the results, interpretation and lead to wrong conclusions.

3.3 Numerical resolution corrections

The highest mass resolutions of the different IllustrisTNG simulations, TNG50, TNG100, and TNG300, differ from each other, see Table 1. Combining results of TNG50, TNG100, and TNG300 therefore requires a resolution correction for the luminosities, magnitudes, stellar masses and star formation rates. More specifically, this requires a resolution correction for TNG100 and TNG300 to match the numerical resolution of TNG50, which provides the highest numerical resolution of all simulations of the IllustrisTNG simulation suite. Once such a numerical resolution correction has been applied, the results of the three simulations can be combined. We achieve this resolution correction through a band magnitude resolution correction factor. A similar procedure has been applied for stellar masses and metal abundances of IllustrisTNG simulations (e.g., Pillepich et al. 2018b; Vogelsberger et al. 2018b). We stress that the resolution corrections differ among the different bands since the SED resolution dependence is wavelength dependent. Furthermore, these corrections also have to be derived for dust-free and dust-attenuated cases separately, and all correction factors are redshift dependent.

In the following we describe briefly the procedure to derive the resolution corrections for the different band magnitudes. To this end we study the resolution dependence of the three simulations TNG50, TNG100 and TNG300 at their various resolution levels. Specifically, we explore the dependence of galaxy magnitudes on

the baryonic mass resolution for the different bands, simulations, resolutions and redshifts. For each simulation, we first divide the galaxies in 20 logarithmically spaced halo mass bins ranging from $10^9 M_\odot$ to $10^{14} M_\odot$. We assume that the impact of resolution on the halo mass is relatively mild and that the halo mass is a stable quantity to divide galaxies into groups from simulations with different mass resolutions. We then calculate for each simulation the median band magnitude of all galaxies in a given halo mass bin for every band. By comparing the median magnitude of TNG100 (TNG300) with the median magnitude of TNG50 for every halo mass bin, we can derive a magnitude correction for TNG100 (TNG300) to the resolution level of TNG50 at a given halo mass. The simulation volume of TNG50 is too small to provide a large number of galaxies towards larger halo masses. We therefore set the magnitude correction factor to zero once the number of galaxies in a halo mass bin drops below 10 for TNG50. This assumption is supported by the fact that TNG100 and TNG300 have nearly identical luminosity functions and stellar mass functions towards the bright and massive end. We can therefore safely assume a vanishing correction factor in this regime. We apply this whole procedure of deriving correction factors to all redshifts of interest. The resulting magnitude correction factors for each dust model as a function of halo mass and redshift are then stored such that we can easily correct all band magnitudes for all simulations at all redshifts. Resolution corrected band magnitudes can then be used to construct resolution corrected galaxy luminosity functions. We note that the described resolution correction procedure can also be applied to other quantities like galaxy stellar masses and star formation rates. Those resolution corrected values build the basis to eventually derive the combined galaxy luminosity functions at all redshifts, for different bands, and for the three different dust models. Before describing how to derive this combined luminosity function based on the resolution corrected magnitudes, we first briefly discuss the dust model calibration step to determine the free dust parameters for the three dust models.

3.4 Dust model calibration

All three dust models presented above depend on some free parameters. Specifically, Model A depends on two dust parameters; Models B and C depend each on one free dust parameter. Ultimately, these free parameters reflect the uncertain knowledge of the amount and distribution of dust in the simulated galaxies. The exact amount of dust and its distribution is not directly tracked in IllustrisTNG. Only recently, some galaxy formation simulations

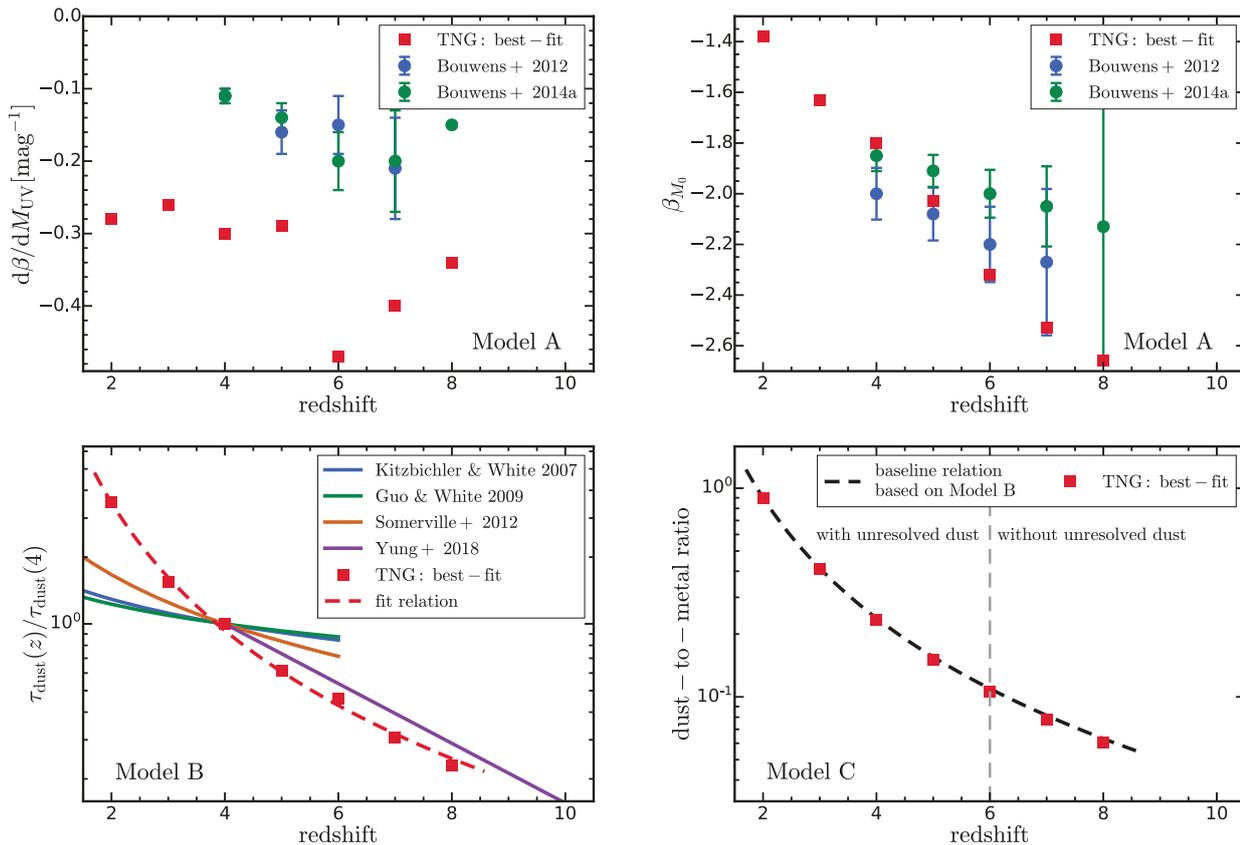


Figure 5. Redshift dependence of best-fit dust parameters for dust Models A, B, and C. *Upper left and upper right panels:* Best-fit values for the $M_{\text{UV}}^{\text{dust}} - \langle \beta \rangle$ relation assuming an underlying Meurer et al. (1999) relation and fixed $M_0 = -19.5$ mag for dust Model A. We also present the observationally inferred relations of Bouwens et al. (2012, 2014a). These deviate from the best-fit values for the slope, $d\beta/dM_{\text{UV}}$, redshift evolution but are consistent for the intercept β_{M_0} . *Lower left panel:* Best-fit $\tau_{\text{dust}}(z)$ values of Model B. We also include the redshift evolution from other theoretical works, mostly semi-analytic models, for the evolution of the dust optical depth normalisation employing similar dust models. *Lower right panel:* Best-fit dust-to-metal ratio for Model C. The redshift evolution of the dust-to-metal ratio is based on the optical depth redshift evolution of Model B as baseline relation.

and semi-analytical models have been able to track this dust component in detail through the direct modelling of the relevant dust physics (McKinnon et al. 2016; Aoyama et al. 2017; McKinnon et al. 2017; Popping et al. 2017; McKinnon et al. 2018; Vogelsberger et al. 2018a; Davé et al. 2019; Vijayan et al. 2019). We therefore have to rely on the gas phase metallicity as a proxy for the dust component in the galaxy. It is common to assume here a linear scaling and express the amount of dust in terms of a dust-to-metal ratio. This parameter is however non-universal, and also redshift dependent. We therefore have to perform a dust model calibration for all dust models and all redshifts.

For the calibration process we compare the predicted dust-attenuated rest-frame UV luminosity functions to existing observational data of rest-frame UV luminosity functions. Our calibration is based on the observational data compilation of Finkelstein (2016) at $z = 4 - 10$, which consists of both ground- and space-based observations from McLure et al. (2009); Castellano et al. (2010); van der Burg et al. (2010); McLure et al. (2013); Oesch et al. (2013, 2014); Schenker et al. (2013); Tilvi et al. (2013); Bowler et al. (2014, 2015); Bouwens et al. (2015, 2016); Finkelstein et al. (2015); Schmidt et al. (2014); McLeod et al. (2015, 2016). Additional observational constraints are taken from Alavi et al. (2014) at $z = 2$, Mehta et al. (2017) at $z = 2, 3$, Parsa et al. (2016) at $z = 2-4$, Reddy et al. (2008) and Reddy & Steidel (2009) at $z = 3$, van der

Burg et al. (2010) at $z = 3 - 5$, Bouwens et al. (2017); Atek et al. (2018) at $z = 6$, Atek et al. (2015b); Ouchi et al. (2009) at $z = 7$ and Oesch et al. (2018) at $z = 10$. We perform the calibration of dust parameters based on the resolution corrected TNG100 simulation result, which represents the best compromise of volume to sample the bright end, which is most sensitive to dust attenuation, and numerical resolution. We note that results of the dust calibration are sensitive to the radial cut applied for the stars contributing to the luminosity calculation of a given galaxy. Throughout this work, we use 24 points linearly spaced from -16 mag to -24 mag to create 23 bins for the UV luminosity function calculation.

Model A: In principle, Model A contains no free parameters, since all the relations and parameters of this model are directly tied to observational scaling relations. However, as discussed above, there are significant observational uncertainties in the parameters of these relations. Most importantly the $M_{\text{UV}}^{\text{dust}} - \langle \beta \rangle$ relation and its redshift evolution are still quite uncertain. Instead of relying on observational relations, we will therefore find the parameters for the $M_{\text{UV}}^{\text{dust}} - \langle \beta \rangle$ relation through a best-fit to the rest-frame UV luminosity function. We note that we have explicitly verified that directly applying existing observationally inferred $M_{\text{UV}}^{\text{dust}} - \langle \beta \rangle$ relations to our simulation results leads to predicted galaxy luminosity functions that are inconsistent with observed rest-frame UV luminosity functions. For the calibration of Model A, we set up a

two dimensional grid for each redshift, with $d\beta/dM_{\text{UV}}$ and β_{M_0} as the two axes. The $d\beta/dM_{\text{UV}}$ axis contains 81 points ranging from -0.8 mag^{-1} to 0 mag^{-1} . The β_{M_0} axis contains 301 points ranging from -4 to -1 . For each grid point we apply the dust attenuation of Model A and compare the resulting resolution corrected TNG100 rest-frame UV luminosity function with our sample of observed rest-frame UV luminosity functions and calculate χ^2 as:

$$\frac{\chi^2}{n_{\text{eff}} - \nu} = \frac{1}{n_{\text{eff}} - \nu} \left[\sum_{i=1}^{n_{\text{eff}}} \left(\frac{\log \phi(M_{\text{ob}}^i) - \log \phi_{\text{ob}}^i}{\sigma_{\text{ob}}^i} \right)^2 \right], \quad (26)$$

where $\nu = 2$ is the number of free parameters in the model and $\log \phi()$ is the interpolated resolution-corrected luminosity function predicted by our model. For the i^{th} observation data point, $\log \phi_{\text{ob}}^i$ is the logarithm of the observed galaxy number density at M_{ob}^i , σ_{ob}^i is the uncertainty of the observational measurement. Only observational data points with observed $M_{\text{UV}} \in [M_{\text{limit}}, -24 \text{ mag}]$ and within the magnitude coverage of the interpolated luminosity function predicted by the model are taken as effective points. M_{limit} is set to -18 mag at $z = 3$, -19 mag at $z = 2, 4, 5$ and -20 mag at $z \geq 6$. The number of these effective points is n_{eff} . The parameter choice resulting in the smallest $\chi^2/(n_{\text{eff}} - \nu)$ then represents the best-fit Model A dust parameters at a given redshift. The result of this calibration procedure is presented in the top row panels of Figure 5, where we also compare our best-fit values to observationally derived parameters. We note that at $z = 9, 10$, the number of luminosity function data points of TNG100 is not sufficient to derive reliably best-fit values for Model A. We therefore neglect dust attenuation for $z = 9$ and $z = 10$. This assumption is justified by the calibration of Model B, and also by the results of Model C as we will discuss below. These calibrations demonstrate that dust attenuation is indeed negligible at $z = 9$ and $z = 10$.

Model B: For Model B, the only free dust parameter is $\tau_{\text{dust}}(z)$. This optical depth normalisation parameter is a free scaling factor, which reflects the unknown dust-to-metal ratio in each simulated galaxy. Many previous works have found that the relation between dust attenuation or dust optical depth and gas metallicity and column density evolves through cosmic time reflecting a redshift-dependent dust-to-metal ratio. Also a fixed relation tends to overpredict the effect of dust attenuation for high redshift galaxies significantly (e.g., Somerville et al. 2012; Wilkins et al. 2013; Reddy et al. 2015; Whitaker et al. 2017; Reddy et al. 2018; Yung et al. 2018). Therefore, $\tau_{\text{dust}}(z)$ is expected to be redshift dependent. Similar to Somerville et al. (2012) and Yung et al. (2018), we calibrate the value of $\tau_{\text{dust}}(z)$ to match the rest-frame UV luminosity functions observed at $z = 2 - 10$. We explore 81 values of τ_{dust} linearly spaced from 0 to 0.8 at each redshift. We compare the resulting UV luminosity functions with observations and calculate χ^2 values using Equation 26 with $\nu = 1$. Since the Model B parameter variation only leads to significant differences at the bright end, we only consider observational data points with observed $M_{\text{UV}} \in [M_{\text{limit}}, -24 \text{ mag}]$ for the calibration process. M_{limit} is set to be -20 mag , except for -21 mag at $z \geq 8$. At each redshift, we then calculate the $\chi^2/(n_{\text{eff}} - \nu)$ for each parameter choice. The parameter choice resulting in the smallest $\chi^2/(n_{\text{eff}} - \nu)$ then represents the best-fit dust parameter for that redshift. The result of this procedure is presented in the lower left panel of Figure 5. At $z = 9, 10$ the bright end of TNG100 is not sampled well enough to provide reliable best-fit dust parameters for Model B. Extrapolating the $\tau_{\text{dust}}(z)$ scaling to these redshifts results in $\tau_{\text{dust}}(z) \lesssim 10^{-2}$, which implies a negligible dust correction. We therefore set $\tau_{\text{dust}}(z) = 0$ for $z = 9, 10$. Overall we find that the redshift evolution of $\tau_{\text{dust}}(z)$ is

well described by a power law relation with $z^{-1.92}$ as indicated in the lower left panel of Figure 5. The redshift dependence at $z \geq 4$ is similar to the dust attenuation scaling recently presented in Yung et al. (2018) for a similar dust model.

Model C: For Model C, the only free dust parameter is the dust-to-metal ratio $f_z(z)$. For the calibration of Model C we assume the power law fit of $\tau_{\text{dust}}(z)$ from Model B as a baseline since it is computationally expensive to perform radiative transfer calculations for all simulations for a large number of different dust-to-metal ratios. We therefore first perform a dust parameter exploration for $z = 2$ to $z = 8$ following the redshift dependence of $\tau_{\text{dust}}(z)$ found for Model B. We explore $f_z(z) = [0.3, 0.5, 0.9, 1.3, 1.7] \times (z/2)^{-1.92}$ for the dust-to-metal ratio and compare the resulting luminosity functions with observations. To accelerate this process, we require a less stringent cell splitting criterion for SKIRT: $1/N_{\text{gas}}$ with 10^{-3} as upper limit and 2×10^{-6} as lower limit instead of using a fixed value of 2×10^{-6} . We find that $f_z(z) = 0.9 \times (z/2)^{-1.92}$ results in the best-fit UV luminosity functions. This best-fit dust parameter choice is presented in the lower right panel of Figure 5. Based on this exploration result, we execute the SKIRT pipeline for this redshift dependent dust-to-metal parameter evolution with the fiducial cell splitting criterion, 2×10^{-6} , for all redshifts $z = 2 - 10$ to produce the final dust-attenuated luminosities for Model C. During the calibration processes we have also noticed that the unresolved dust attenuation of the MAPPINGS-III spectra templates results in too much dust attenuation for $z \geq 6$. This results in a suppression of the abundance of faint and intermediate brightness galaxies. We therefore do not consider the unresolved dust component for $z \geq 6$, i.e. we do not employ the MAPPINGS-III spectra templates for young stars for these high redshifts. The low dust-to-metal ratios at these high redshifts furthermore implies a vanishing attenuation due to resolved dust. Therefore, the high redshift luminosities are essentially identical to the intrinsic dust-free magnitudes. This is similar to the high redshift behaviour of Model B. In this case the resolved component also vanishes due to the small normalisation factor of the optical depths. Similarly, also the unresolved component vanishes towards higher redshifts since it scales linearly with the resolved dust attenuation by construction.

We can compare our findings for the best-fit dust-to-metal ratio with the rather limited amount of observational and theoretical findings. Observationally, Zafar & Watson (2013) found that the dust-to-metal ratio in a sample of foreground absorbers to Gamma-ray bursts (GRBs) and quasars does not vary significantly over a wide range of redshifts 0.1 – 6.3, metallicities, and hydrogen column densities, favouring a universally constant dust-to-metal ratio. Chen et al. (2013) found a weak redshift dependence of dust-to-metal ratios in lensed galaxies. De Cia et al. (2013) studied dust-to-metal ratios of GRB damped Lyman-alpha absorbers. They found a large scatter in the dust-to-metal ratio distribution at $z > 1.7$ and also found several systems with low dust-to-metal ratio at high redshift. Rémy-Ruyer et al. (2014) found that dust-to-metal ratios can be lower in low metallicity systems with a large dispersion. Wiseman et al. (2017) studied the evolution of dust-to-metal ratios through cosmic time also via GRB damped Lyman-alpha absorbers. They found that dust-to-metal ratios increase with metallicity (see also De Cia et al. 2016) but did not observe an obvious redshift dependence of dust-to-metal ratio. Theoretically, the analytical model of Inoue (2003) demonstrated that dust-to-metal ratios are lower in younger galaxies and have a significant redshift dependence. They suggested that the Lyman break galaxies at $z \sim 3$ have a low dust-to-metal ratio, < 0.1 . Kimm & Cen (2013) found that the dust-

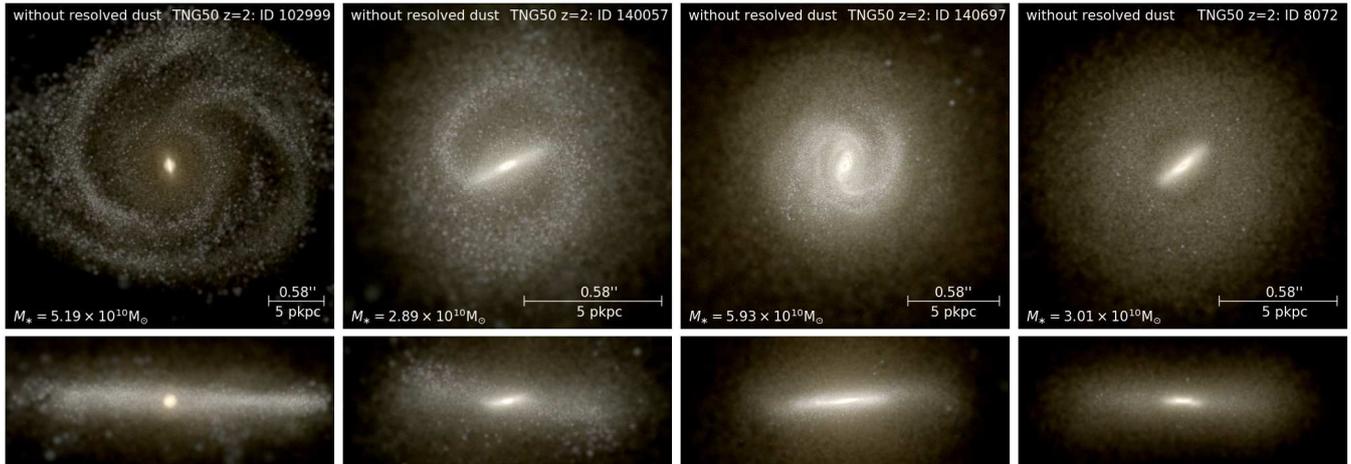
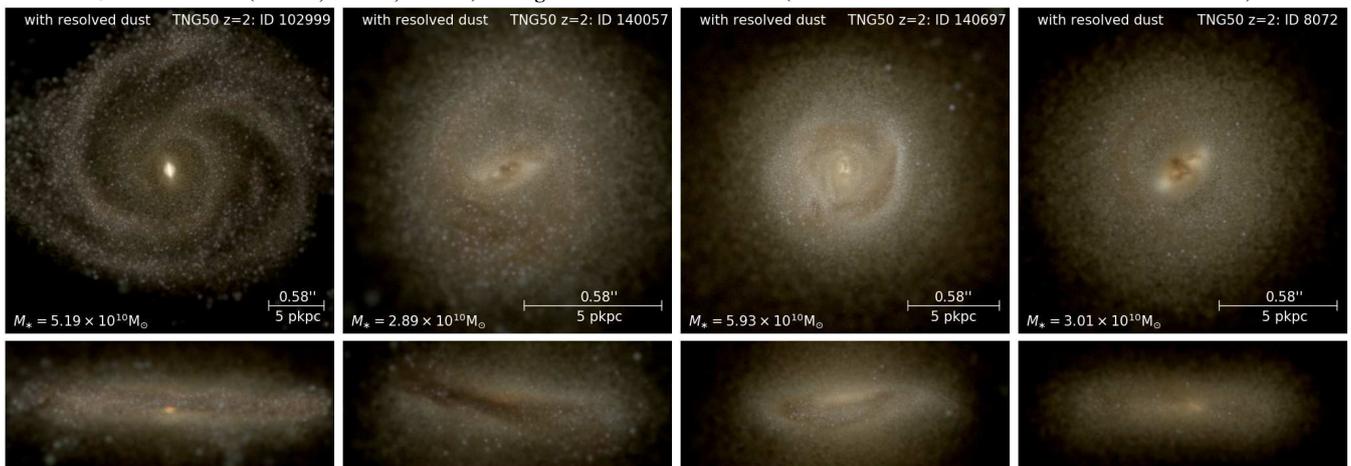
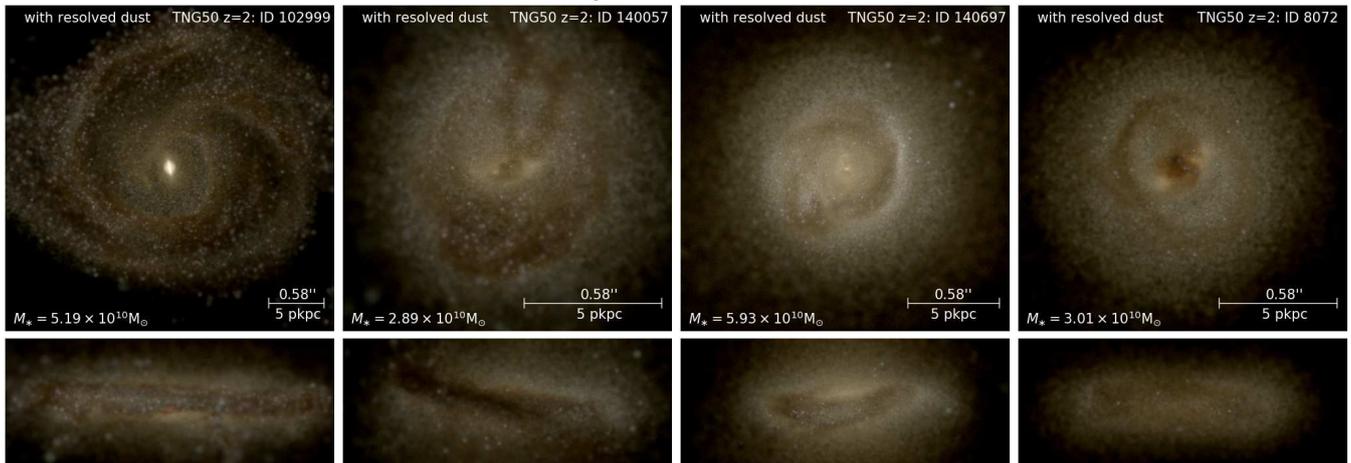
JWST NIRCcam (F115W, F150W, F200W) $z = 2$ galaxies without resolved dust**JWST NIRCcam (F115W, F150W, F200W) $z = 2$ galaxies with resolved dust (dust-to-metal ratio 0.3 = one third of best-fit value)****JWST NIRCcam (F115W, F150W, F200W) $z = 2$ galaxies with resolved dust (dust-to-metal ratio 0.9 = best-fit value)**

Figure 6. JWST NIRCcam face-on and edge-on images of TNG50 galaxies at $z = 2$. *Top row panels:* Galaxy images without resolved dust. *Middle and bottom row panels:* Galaxy images with resolved dust assuming different dust-to-metal ratios as indicated. The images are based on apparent F115W, F150W, F200W filter fluxes. The radiative transfer calculation employed 10^7 photon packets per wavelength on our reduced wavelength grid with 234 points in the wavelength range from $0.05\mu\text{m}$ to $5\mu\text{m}$. This grid is specifically constructed to sample emission lines. The inclusion and sufficient sampling of these emission lines is important for a correct modelling of the appearance of galaxies. We note that our fiducial wavelength grid to derive galaxy SEDs and magnitudes employs a finer grid with 1168 points within the same wavelength range. For the dust cases we assume dust-to-metal ratios of 0.9, which is the best-fit value based on the dust calibration procedure, and 0.3. In general the appearance of galaxies is severely affected by dust. Most importantly, light from the central bright regions of the various galaxies is strongly absorbed and scattered due to dust.

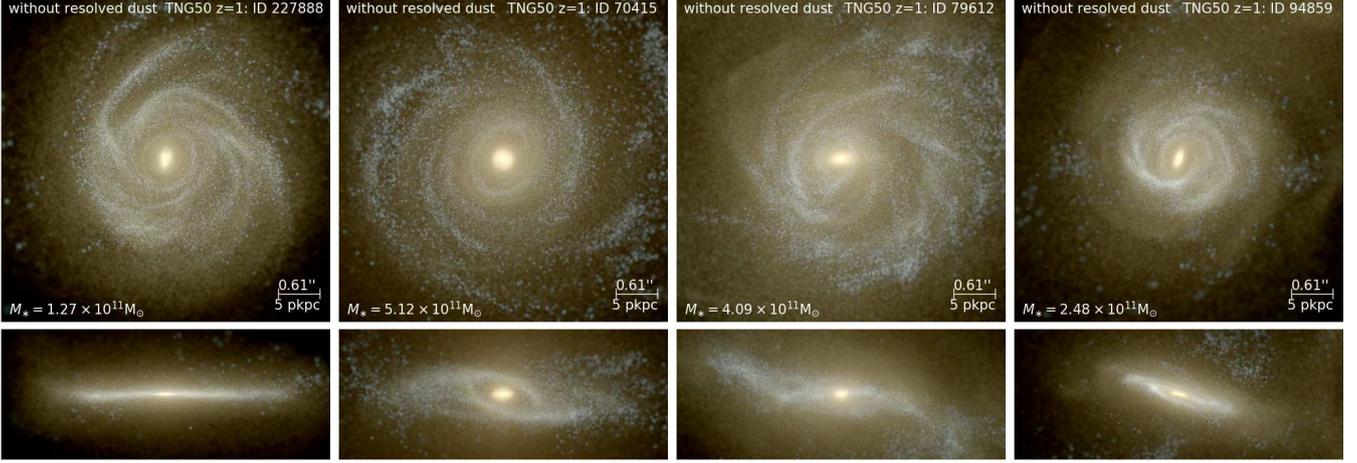
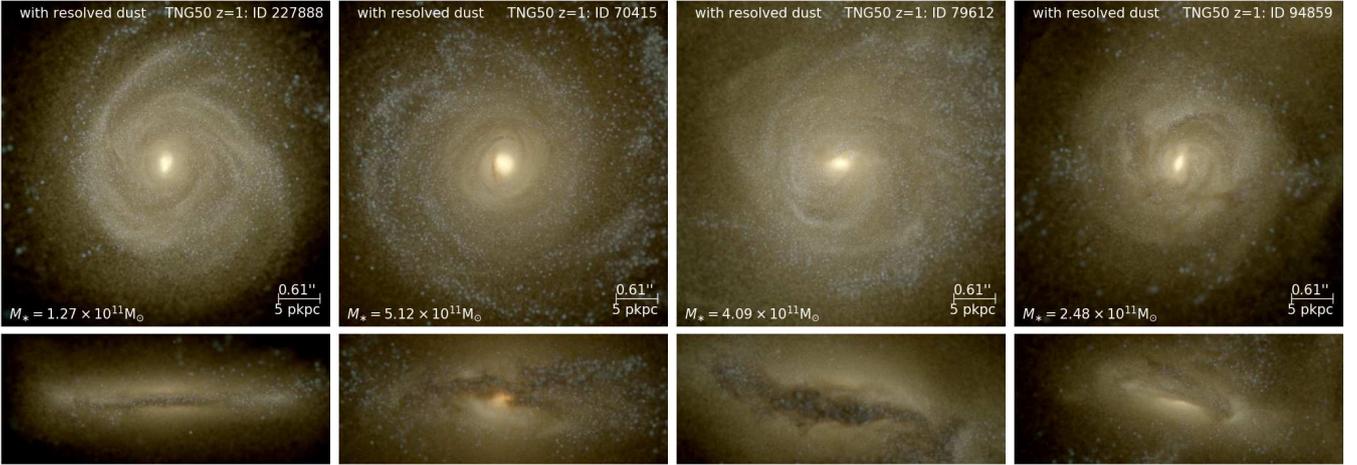
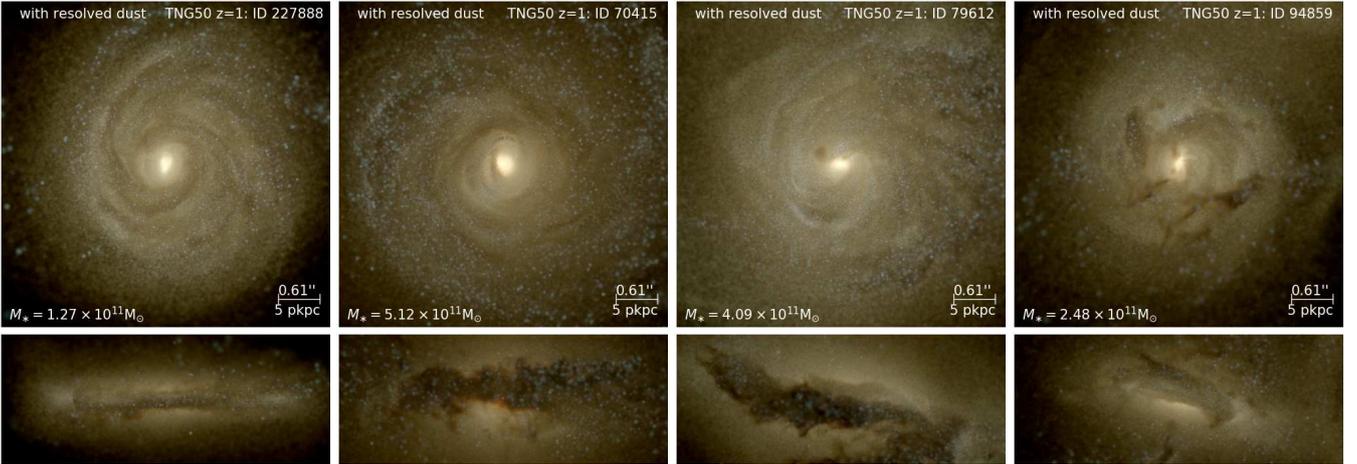
JWST NIRCam (F070W, F090W, F115W) $z = 1$ galaxies without resolved dust

JWST NIRCam (F070W, F090W, F115W) $z = 1$ galaxies with resolved dust (dust-to-metal ratio 0.3)

JWST NIRCam (F070W, F090W, F115W) $z = 1$ galaxies with resolved dust (dust-to-metal ratio 0.9)


Figure 7. JWST NIRCam face-on and edge-on images of TNG50 galaxies at $z = 1$. *Top row panels:* Galaxy images without resolved dust. *Middle and bottom row panels:* Galaxy images with resolved dust assuming different dust-to-metal ratios as indicated. The images are based on apparent F070W, F090W, F115W filter fluxes. The radiative transfer calculation employed 10^7 photon packets per wavelength on our reduced wavelength grid with 234 points in the wavelength range from $0.05\mu\text{m}$ to $5\mu\text{m}$. This grid is specifically constructed to sample emission lines. The inclusion and sufficient sampling of these emission lines is crucial for a correct modelling of the appearance of galaxies. We note that our fiducial wavelength grid to derive galaxy SEDs and magnitudes employs a finer grid with 1168 points within the same wavelength range. For the dust cases we assume dust-to-metal ratios of 0.9, which is the best-fit value based on the dust calibration procedure, and 0.3. In general the appearance of galaxies is severely affected by dust. Most importantly, light from the central bright regions of the various galaxies is strongly absorbed and scattered due to dust.

to-metal ratio is 0.06 (0.4) at $z = 7$ assuming an escape fraction of light of 0% (10%), which represent a lower (upper) limit, respectively. Yajima et al. (2014) explored dust-to-metal ratios in the range from 0.05 to 0.5, with 0.5 as their fiducial model. They found that a dust-to-metal ratio > 0.4 is required at $z = 3$ to match observational data. They also suggest that the dust-to-metal ratio at $z = 3$ is close to that of the local Universe. Bekki (2015) found that dust-to-metal ratios can be different between different galaxies at different redshifts and even varies within a single galaxy. McKinnon et al. (2016) used hydrodynamical simulations of Milky Way-like galaxies to study the evolution of dust and also found no evolution in dust-to-metal ratio at $z > 1$. However, they predicted a much lower dust-to-metal ratio at $z > 2$ compared to the value at $z = 0$. Popping et al. (2017) studied the dust content of galaxies from $z = 0$ to $z = 9$ and found that dust-to-metal ratios can vary in systems with different metallicity and stellar mass. One of their models shows a redshift dependence for the dust-to-metal ratio. Aoyama et al. (2017) found that the assumption of fixed dust-to-metal ratio is no longer valid once galaxies are older than 0.2 Gyr. Behrens et al. (2018) found in simulations that the dust-to-metal ratios for galaxies at $z \sim 8$ are only 0.08, i.e. far lower than the value in the local Universe. These various theoretical and observational results demonstrate the difficulty of measuring and modelling the dust content of galaxies. We note that it is expected that the dust-to-metal ratios vary with redshift since the metal content of galaxies also changes with redshift. Our dust calibration procedure is therefore necessary since definite predictions and measurements of dust abundances at high redshifts are currently not available.

We summarise the results of our dust calibration for the different dust models and redshifts in Table 3. We demonstrate the impact of different amounts of dust in Figure 6 for a few more TNG50 $z = 2$ galaxies. All galaxies are presented in face-on and edge-on projections. Galaxy images in the top row do not consider the impact of dust. The other two rows include the effects of dust for two different dust-to-metal ratio: 0.3 for the middle row, and 0.9 for the bottom row. The latter represents the best-fit $z = 2$ dust-to-metal ratio derived based on the calibration procedure discussed above. In Figure 7 we present also four $z = 1$ galaxies of TNG50 for the dust-free case and for two different dust-to-metal ratios. We note that we employed here a different set of NIRCcam wide filters: F070W, F090W, F115W instead of the F115, F150W, and F200W filters that we used at $z = 2$. For all galaxy images, it is obvious that the inclusion of dust attenuation severely affects the stellar light morphology of galaxies.

3.5 Combining TNG50, TNG100 and TNG300 predictions

So far we have described how to resolution correct the different simulations to the TNG50 resolution level, and how we calibrate our three dust models based on the resolution-corrected TNG100 simulation results. Next we describe how we can derive a combined galaxy luminosity function based on the resolution-corrected and dust-corrected magnitudes of all three IllustrisTNG simulations: TNG50, TNG100 and TNG300. This will then lead to predicted galaxy luminosity functions that probe a wide range of galaxy luminosities.

The combination procedure is based on the following steps. For each redshift, we first study the resolution-corrected luminosity functions of all three simulations. An example of those is presented in Figure 8. The panels in this figure demonstrate that the resolution-corrected luminosity functions of the different simulations overlap within some luminosity range but deviate from each

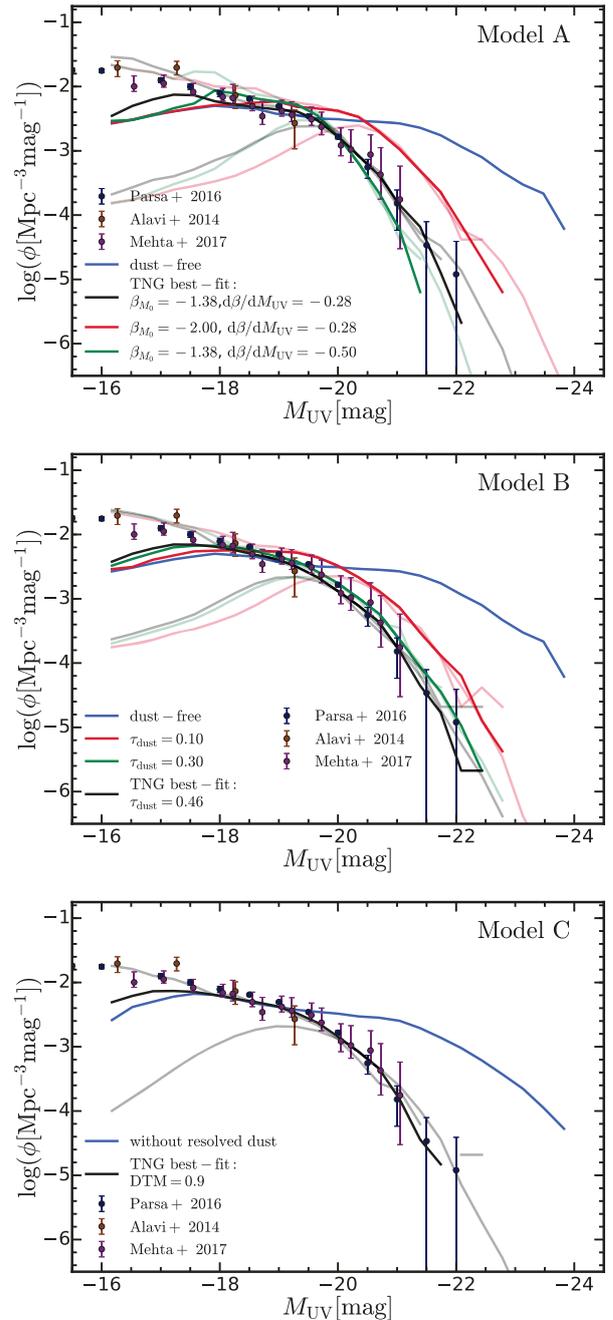


Figure 8. Comparison between rest-frame UV luminosity functions for different parameter choices for the three different dust models at $z = 2$. The resolution corrected luminosity functions of TNG100 are shown with solid lines. The luminosity functions of TNG50 and resolution corrected TNG300 are shown with transparent lines. Besides the best-fit values, we also show for Models A and B dust parameter variations around these best-fit values.

other towards the faint end once the corresponding simulation fails to resolve small galaxies. The magnitude where the simulations start to deviate can be seen as a sharp drop in the predicted abundance for simulations with a lower resolution. This marks a lower magnitude limit over which a given simulation is considered to be valid. Higher resolution simulations probe the faint end down to

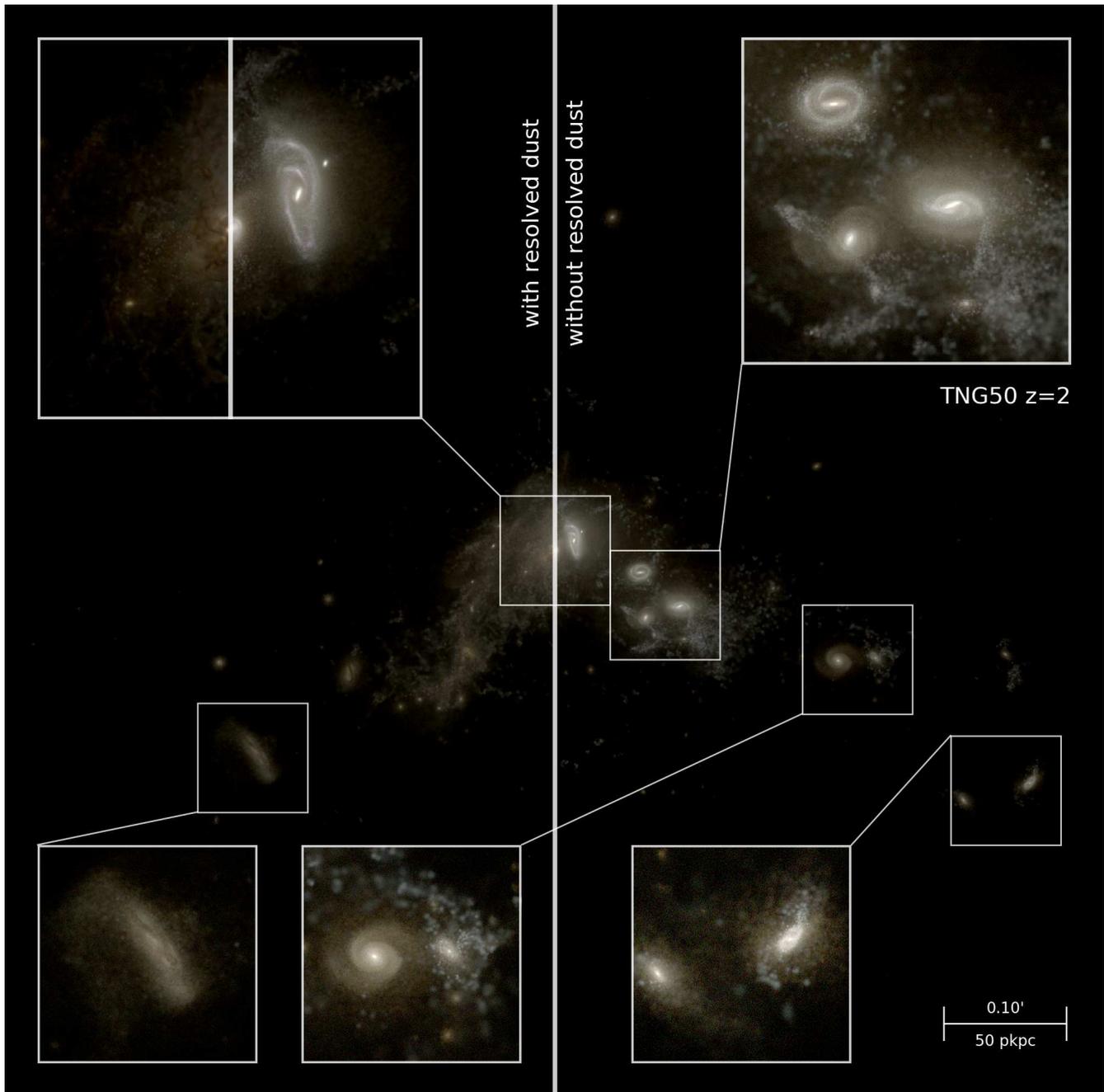


Figure 9. TNG50 protocluster JWST NIRCcam image at $z = 2$ with and without dust attenuation. The image has a physical extent of 450 pkpc and is based on apparent F115W, F150W, F200W filter fluxes. The field of view is sampled by 2000×2000 pixels and 10^7 photon packets on the reduced wavelength grid with 234 points in the range $0.05 \mu\text{m}$ to $5 \mu\text{m}$, which is five times coarser than our fiducial wavelength grid employed for the integrated galaxy fluxes. In addition to the large field of view, we place higher resolution detectors, each with 500×500 pixels, on the zoom-in regions as indicated. These additional detectors show the zoom-in images of individual galaxies marked in the larger field of view. The vertical line divides the image in a left half that includes dust attenuation, and a right half does not include dust instead shows stellar light without resolved dust attenuation.

lower magnitudes. Towards the bright end, the galaxy number densities fluctuate due to Poisson sampling noise set by the simulation volume. Larger simulation volumes lead to less noise towards the bright end. This sets for a given simulation an upper magnitude limit above which a simulation is not considered valid anymore. In order to reliably combine the luminosity functions, we therefore first determine at each redshift and for each simulation these lower and upper magnitude limits within which the luminosity function

is reliably sampled. These points are marked as those magnitudes where the different resolution corrected simulations predict galaxy abundances that differ more than 0.2 dex. Beyond these boundaries, the luminosity function does not contribute to the construction of the combined luminosity function.

For each galaxy magnitude bin within these limits we then calculate an average galaxy abundance, which represents the combined galaxy luminosity function at that specific magnitude. The

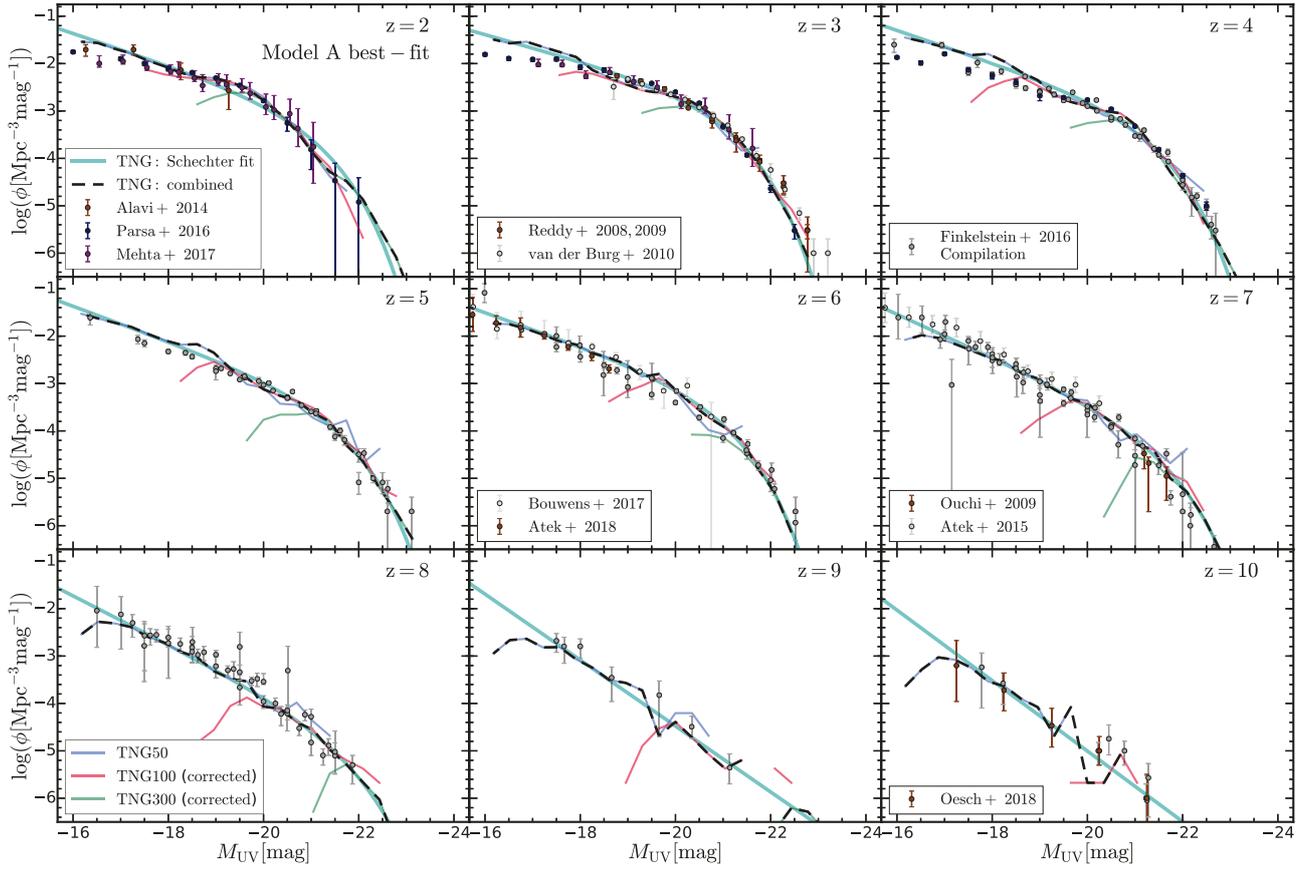


Figure 10. Best-fit Model A UV luminosity functions at $z = 2 - 10$. We show resolution corrected luminosity functions of TNG100 in solid red. Comparing those with the luminosity functions of TNG50 (blue) and the resolution corrected luminosity function of TNG300 (green) demonstrates that we achieve a good agreement between the different simulations over their common luminosity range. Combining the three simulations yields the combined luminosity functions (black dashed lines). Schechter fits for the combined luminosity functions are presented with cyan lines. In general the combined luminosity functions agree with the observational data very well. At $z \lesssim 4$, the combined luminosity function predicts slightly higher number densities and steeper faint-end slopes compared to observational data. Towards higher redshifts we find that IllustrisTNG predicts slightly lower abundances of galaxies. This is most noticeable at $z = 8$, where we predict lower abundances over a quite wide range of rest-frame UV luminosities from -18 mag to -21 mag.

average is performed with weights based on the number of galaxies in each magnitude bin. This gives more weight to simulations that sample a given bin with more galaxies. Specifically, we calculate the average for the combined luminosity function with:

$$\phi_{\text{combined}} = \frac{\phi_{50} N_{50}^2 + \phi_{100} N_{100}^2 + \phi_{300} N_{300}^2}{N_{50}^2 + N_{100}^2 + N_{300}^2}, \quad (27)$$

where N_X and ϕ_X refer to the number of galaxies of TNGX in each bin and the resolution-corrected luminosity function of TNGX, respectively. This procedure then leads to one combined galaxy luminosity function for each redshift and for all bands. In the following we will study those combined luminosity functions at different redshifts, for different bands, and for different dust models. We note that this combination procedure can also be applied to other quantities like the galaxy stellar mass functions.

4 RESULTS

After discussing in detail our methods, we will present our results in this section. We start with a first image of the high redshift uni-

verse as predicted by the TNG50 simulation on larger scales. To this end, we show in Figure 9 a *JWST* mock image of the most massive protocluster within the TNG50 volume at a redshift of $z = 2$. This image covers a field of view of 450×450 pkpc². In addition we also zoom in onto some galaxies by placing high resolution detectors on some regions of the field of view. This protocluster turns into a $z = 0$ cluster with a mass of $M_{200,\text{crit}} \sim 2 \times 10^{14} M_{\odot}$. The vertical line divides the image in a left half that includes dust attenuation based on dust Model C, and a right half does not include resolved dust attenuation.

4.1 Rest-frame UV luminosity functions

More quantitatively, we next discuss our result for the rest-frame UV luminosity functions. The UV luminosity functions at $z = 2$ with different dust models and different dust parameters are presented in Figure 8 that we briefly discussed already in the previous section. We present the best-fit resolution-corrected luminosity function of TNG100 in black lines along with the corresponding luminosity functions of TNG50 and the resolution-corrected lumi-

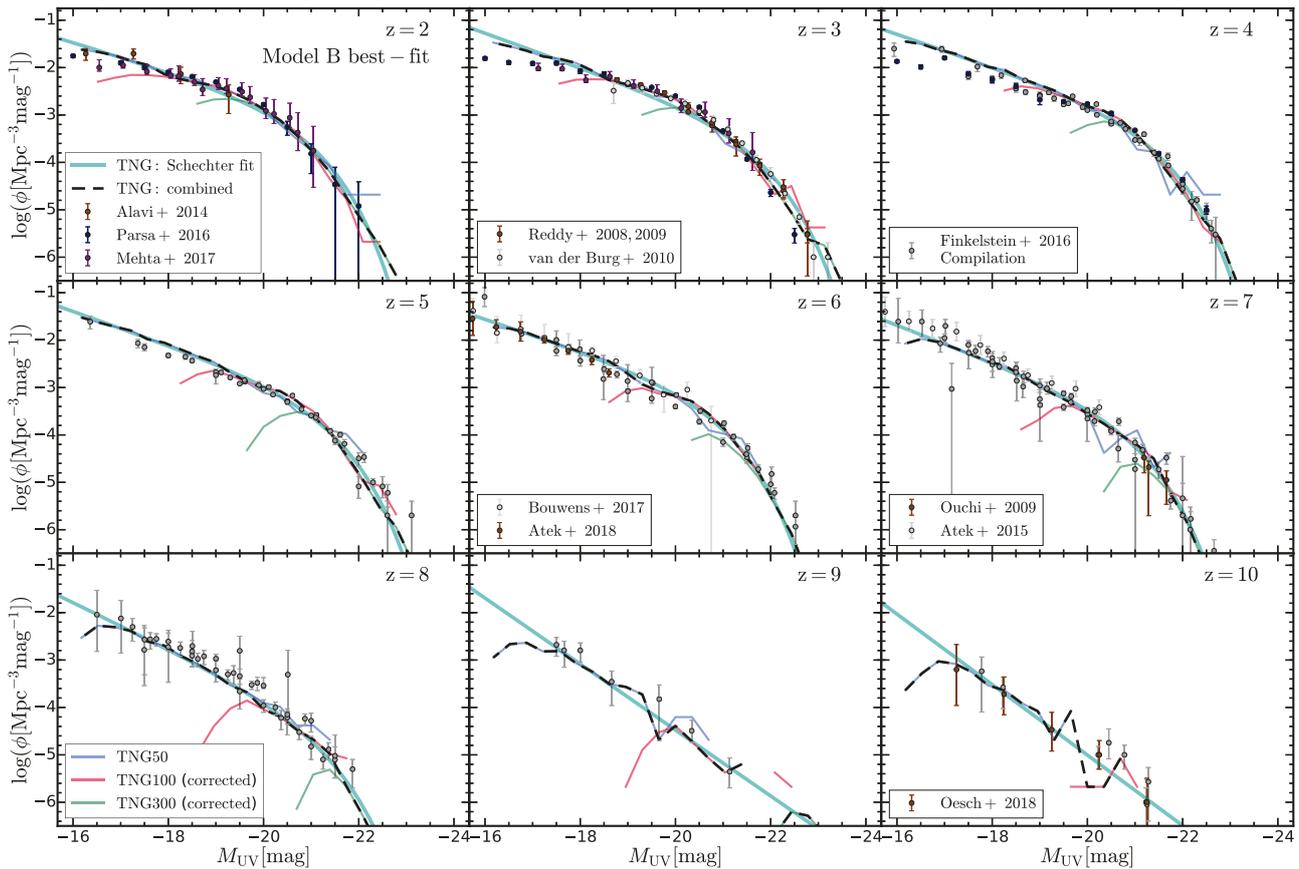


Figure 11. Best-fit Model B UV luminosity functions at $z = 2 - 10$. The combined luminosity functions agree better with observational data than the combined luminosity functions of Model A. However, at $z \lesssim 4$, the combined luminosity function still predicts slightly higher number densities and steeper slopes towards the faint end compared with observations. This is an inherent feature of the galaxy population predicted by TNG50. Similarly, towards higher redshifts we still find that IllustrisTNG predicts slightly lower abundances of galaxies for dust Model B.

nosity function of TNG300 with transparent lines. Figure 8 demonstrates that the resolution correction works reliably over the magnitude ranges where the individual simulations sample the galaxy population sufficiently in terms of resolution and simulation volume. In addition to the luminosity functions based on the best-fit dust parameters, we also present the luminosity functions for slightly different dust parameters (coloured lines) for Model A and Model B to highlight their impact on the shape of the UV luminosity functions. The variation of the Model A intercept value β_{M_0} dust parameter mainly affects the offset of the luminosity function while the slope $d\beta/dM_{UV}$ mostly affects the shape of the bright end. For Model B, as expected, increasing the dust parameter $\tau_{dust}(z = 2)$ shifts the luminosity function away from the dust-free luminosity function at the bright end while preserving the faint end shape. This behaviour is generic also at other redshifts, i.e. as expected dust attenuation mostly affects the bright end of the galaxy population but not the faint end. Towards the bright end the impact of dust is quite severe leading to large corrections on top of the dust-free luminosity function, which is also presented in the different panels. We also note that given the fact that the faint end is not significantly affected by dust attenuation, any predictions for the faint end, like the faint-end slope, are not affected by our dust modelling in any significant way. Furthermore, the faint end

is mostly sampled by TNG50, which also does not require any resolution corrections. This implies that the faint-end slope is a quite robust prediction of our analysis given that it does not depend on dust- or resolution-corrections. Of course, TNG50 itself might still not be fully converged, but we expect that resolution impacts are comparatively small for TNG50. We stress here however that only the faint-end prediction of Model B and C are not sensitive to dust parameter choices since those directly model the dust based on the metal content of galaxies. However, this is not the case for Model A, which is only an empirical relation providing a mapping between the dust-free and dust-attenuated rest-frame UV magnitudes. Therefore parameter variations of Model A can in principle also shift the galaxy number densities at the faint-end of the luminosity function.

Next we explore the best-fit dust-attenuated galaxy luminosity functions at all redshifts of interest beyond $z = 2$. In Figures 10, 11 and 12, we present the rest-frame UV luminosity functions from $z = 2$ to $z = 10$ for the best-fit dust parameters of Models A, B, and C, respectively. Here we include all simulation volumes, TNG50, TNG100, and TNG300, with the latter two being resolution-corrected. We stress again that the dust parameter calibration for all dust models is solely based on the resolution-corrected TNG100 simulation to avoid any mis-calibration due to

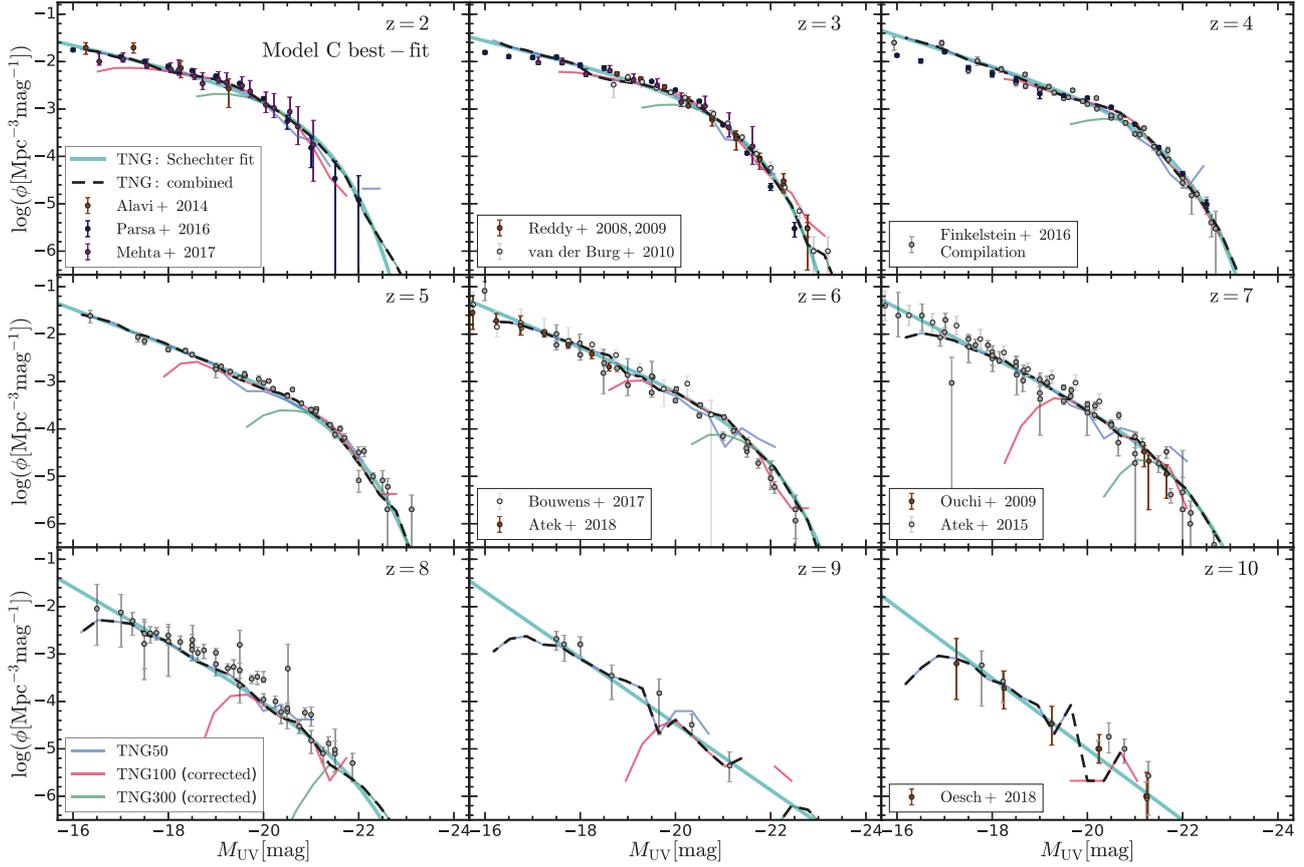


Figure 12. Best-fit Model C UV luminosity functions at $z = 2 - 10$. The combined luminosity functions agree with observational data as well as Model B although Model C leads to a slightly better agreement with observational data towards the faint end.

redshift z	Model A			Model B			Model C		
	$\log \phi^*$ [Mpc $^{-3}$ mag $^{-1}$]	M^* [mag]	α^*	$\log \phi^*$ [Mpc $^{-3}$ mag $^{-1}$]	M^* [mag]	α^*	$\log \phi^*$ [Mpc $^{-3}$ mag $^{-1}$]	M^* [mag]	α^*
2	-2.898	-20.71	-1.835	-2.805	-20.52	-1.759	-2.652	-20.45	-1.581
3	-2.654	-20.70	-1.695	-3.156	-21.32	-1.901	-2.735	-20.84	-1.627
4	-2.926	-20.93	-1.844	-2.983	-21.02	-1.859	-3.035	-21.10	-1.796
5	-3.244	-21.17	-1.924	-3.107	-20.95	-1.884	-3.398	-21.21	-1.941
6	-3.079	-20.61	-1.876	-3.025	-20.52	-1.833	-3.608	-21.31	-2.042
7	-3.846	-21.18	-2.133	-3.418	-20.58	-1.967	-4.209	-21.47	-2.279
8	-4.445	-21.38	-2.280	-4.111	-20.86	-2.216	-4.714	-21.44	-2.455

Table 4. Best-fit UV luminosity function Schechter parameters for the three dust models (A, B and C). We list for each dust model the three Schechter parameters: the normalisation Φ^* , the transition magnitude M^* , and the faint-end slope parameter α^* .

imperfections in combining the three simulation volumes. In addition to the individual resolution-corrected luminosity functions, we also present the combined luminosity function of the best-fit dust model for each redshift. The combined luminosity functions are shown with black dashed line and have been derived as discussed above. These combined luminosity functions can be well described by Schechter functions (Schechter 1976)

$$\phi(M) = \frac{0.4 \ln(10) \phi^*}{10^{0.4(M-M^*)(\alpha^*+1)}} e^{-10^{-0.4(M-M^*)}}, \quad (28)$$

shown in cyan transparent lines in each panel. We perform a χ^2 fit to derive the best-fit Schechter parameters. To accelerate this process we employ the observationally derived Schechter parameters of Finkelstein (2016) as initial guesses for the fitting procedure. We note that even with TNG50 resolution, it is challenging to sample the faint end of the luminosity function ($M_{UV} \gtrsim -17$ mag) at higher redshifts. Specifically, we observe a drop of the faint end luminosity function for redshifts above $z = 6$. This drop can affect the derived Schechter fits, especially in terms of the faint-end slope parameter α^* . We therefore exclude these magnitude ranges of the lu-

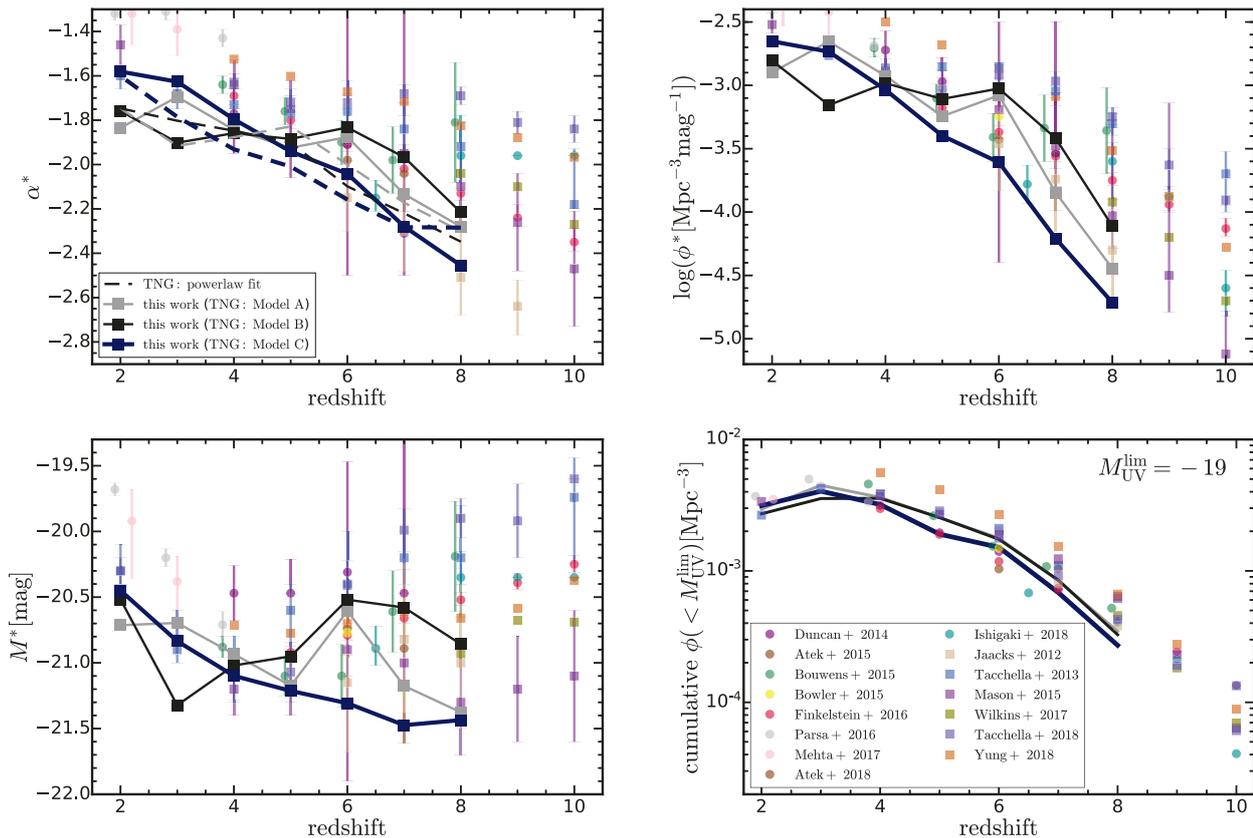


Figure 13. Best-fit Schechter function parameters and cumulative number densities of galaxies above a certain UV luminosity threshold. *Upper left, upper right, lower left panels:* Redshift evolution of the three derived Schechter fit parameters. We compare those with previous results from observations (Duncan et al. 2014; Atek et al. 2015b; Bowler et al. 2015; Bouwens et al. 2015; Finkelstein 2016; Parsa et al. 2016; Mehta et al. 2017; Atek et al. 2018; Ishigaki et al. 2018) (circles), and analytical models and simulations predictions (Jaacks et al. 2012; Tacchella et al. 2013; Mason et al. 2015; Wilkins et al. 2017; Tacchella et al. 2018; Yung et al. 2018) (squares). The Schechter fit parameters of Model A, B, C are shown in gray, black, dark blue squares respectively. The dashed lines in the upper left panel show the faint-end slopes based on single power law fits for $M_{UV} > -19.5$ mag, which yield consistent results with the Schechter slopes. *Lower right panel:* Cumulative number densities analytically calculated based on best-fit Schechter fits of our work and fits presented in previous work. Our results tend to agree slightly better with existing observational data compared to most previous theoretical studies.

luminosity function from the fitting procedure to avoid contamination of the derived best-fit parameters. We only fit the luminosity function in the range $M_{UV} < -17$ mag for $z = 6$ and $M_{UV} < -17.5$ mag for $z > 6$. Furthermore, we ignore bins with less than 10 galaxies at the bright end of TNG300, roughly corresponding to $\log(\phi) = -6$, for the fitting procedure. Our derived best-fit Schechter parameters for the rest-frame UV luminosity functions at different redshifts are listed in Table 4.

We compare our model predictions for the rest-frame UV luminosity function also with observational data. Specifically, we have added at each redshift the compilation of observational datasets described in Section 3.4. Most of these observational points are taken from the Finkelstein (2016) compilation with the original sources listed in Section 3.4. This compilation contains observations mainly from the CANDELS, HUDF, UltraVISTA and UKIDSS UDS surveys. In addition we also include a few more observations that are not included in this large compilation. Specifically, Atek et al. (2018) used gravitational lensing data of the *HST* Frontier Fields program to constrain the extreme faint end UV luminosity function at $z = 6$. Oesch et al. (2018) presented the UV luminosity function at $z = 10$ observed in all *HST* Legacy fields. Mehta et al. (2017) provided constraints on the UV luminosity

functions at $z = 2, 3$ from the *HST* Ultra-Violet Ultra Deep Field program. We note that some of the observations at $z = 2$ (e.g., Reddy et al. 2008; Reddy & Steidel 2009; Oesch et al. 2010b; Hathi et al. 2010) are not fully consistent with the most recent observations (Parsa et al. 2016; Alavi et al. 2014; Mehta et al. 2017). We therefore do not consider these constraints in our analysis. Generally our model predictions agree well with the different observational datasets, which is largely a consequence of the dust calibration procedure. At higher redshifts, $z \sim 9-10$, the predictive power of the IllustrisTNG simulation suite is limited due to the limited simulation resolution of TNG100 and TNG300. We are therefore not able to derive reliable best-fit Schechter parameters for these two high redshifts.

Inspecting our predictions in more detail, we find however also some discrepancies between observations and our theoretical predictions. For example, towards lower redshifts ($z < 4$), the predicted faint end galaxy number densities are slightly higher than observationally found. This also goes along with a slightly steeper faint-end slope. As discussed above, we note that the faint-end behaviour of the luminosity function is expected to be more reliably described by Models B and C, where the dust-free and dust-attenuated results should agree nearly perfectly. This is not neces-

sarily the case for Model A due to its empirical nature. The higher galaxy number densities of faint galaxies for these two dust models are therefore a direct consequence of the galaxy formation model employed in IllustrisTNG and are independent of resolution and dust corrections. Towards the bright end we find that our luminosity function predictions agree well with observational data. We stress here that the dust corrections for the bright end are crucial and obviously our predictions here are sensitive to the dust parameter calibration. We also note that at each redshift the dust attenuation is described only by one normalisation constant for Models B and C, which is the same for all galaxies at the given redshift. The luminosity functions for the three dust models also reveal that for $z \sim 8$ the abundance of galaxies around $M_{UV} \sim -19$ is slightly underpredicted compared to observations. To some degree this is also the case for $z \sim 7$. Despite these minor discrepancies, we conclude that the rest-frame UV luminosity functions of IllustrisTNG agree well with observational data over a wide range of luminosities and redshifts.

In Figure 13, we compare the derived best-fit UV luminosity function Schechter parameters with other values taken from the literature. Observational constraints (Duncan et al. 2014; Bowler et al. 2015; Bouwens et al. 2015; Parsa et al. 2016; Finkelstein 2016; Mehta et al. 2017; Atek et al. 2015b, 2018; Ishigaki et al. 2018) are presented with circles, theoretical predictions (Jaacks et al. 2012; Tacchella et al. 2013; Mason et al. 2015; Wilkins et al. 2017; Tacchella et al. 2018; Yung et al. 2018) are shown as squares. In general, we find slightly steeper faint-end slopes, α^* , and lower ϕ^* values while the predictions for M^* are broadly consistent with previous theoretical results and observational constraints. We have also performed a pure power law fit to the faint-end of the luminosity functions to find the faint-end slope. These power law fits have been performed for $M_{UV} > -19.5$ mag. The result for these slopes are presented as dashed lines in the upper left panel of Figure 13. Those are broadly consistent with the Schechter fit faint-end slopes. Based on the best-fit Schechter functions we can also estimate the cumulative number density of galaxies above a certain magnitude limit. In the bottom right panel of Figure 13, we show this cumulative number density of galaxies above a certain UV luminosity limit $M_{UV}^{\text{lim}} = -19$ mag. This cumulative number density is calculated by integrating the best-fit Schechter functions as:

$$\begin{aligned} \phi_{\text{cum}}(< M_{UV}^{\text{lim}}) &= \int_{L^{\text{lim}}}^{\infty} \phi^* \left(\frac{L}{L^*} \right)^{\alpha^*} \exp\left(-\frac{L}{L^*}\right) \frac{dL}{L^*} \\ &= \phi^* \Gamma_{\text{inc}}(\alpha^* + 1, 10^{-0.4(M_{UV}^{\text{lim}} - M^*)}), \end{aligned} \quad (29)$$

where $\Gamma_{\text{inc}}(a, z) = \int_z^{\infty} t^{a-1} e^{-t} dt$ is the incomplete upper gamma function. We process the best-fit Schechter parameters from the literature similarly to derive cumulative number densities for comparison. Over the range of redshifts where we have valid Schechter fits, we find consistent cumulative number densities compared to previous studies. Our results tend to be in slightly better agreement with observational data than the predictions of some analytical models. We note that this cumulative number density above $M_{UV}^{\text{lim}} = -19$ mag mostly probes the bright end of the rest-frame UV luminosity function. As discussed above, our theoretical predictions are in good agreement with observational data at the bright end. It is therefore not surprising that we find such a good agreement for this cumulative number density given our calibration procedure.

4.2 Apparent *JWST* band luminosity functions

After studying the rest-frame UV band luminosity functions, we will now explore the *JWST* apparent band magnitude luminosity functions. Hereafter, we will only use predictions of our dust Model C in all subsequent analysis. The *JWST* band luminosity functions are resolution corrected and combined in the same way as the UV luminosity functions presented above. The only difference is that we employ here 47 linearly spaced bins from apparent magnitude 16 mag to 32 mag for the analysis. Schechter fits for the combined luminosity functions of four selected *JWST* NIRCcam wide filter bands at each redshift are presented in Figure 14. For the Schechter fitting procedure we set the apparent magnitude range to < 29 mag at $z = 2, 3$ and < 29.5 mag at $z \geq 4$. Also similar to the procedure used to derive the UV luminosity function fits, we ignore bins with less than 10 galaxies, roughly corresponding to $\log \phi = -6$ for TNG300. Towards lower redshifts, the luminosity functions separate clearly. Towards higher redshifts, the luminosity functions of different bands begin however to overlap due to the closer rest-frame wavelength range that is probed. Each panel of Figure 14 also contains rough and typical detection limits in apparent magnitude and number density: ~ 29 mag is the rough detection limit for a 10^4 s exposure time with signal-to-noise ratio 10, and ~ 31 mag is the rough detection limit for a 10^5 s exposure time with signal-to-noise ratio 5. We note that the exact detection limit depends on the actual band filter, and ~ 29 mag, ~ 31 mag are here only approximate values. The precise detection limits for each of the bands are listed in Table 5. The luminosity functions of F070W, which is not shown here, and F090W quickly drop out of the detectable regime towards higher redshifts due to IGM absorption. The horizontal detection line corresponds to one object in the NIRCcam field of view per unit redshift and per magnitude calculated as described below.

The exact *JWST* detection limits of Table 5 are derived based on calculations using the *JWST* Exposure Time Calculator (ETC)² with the following configuration details. Sources are treated as point sources, and the exposure time is set to either 10^4 s or 10^5 s as indicated in the table. Furthermore, the readout pattern is set to DEEP8, which yields a high signal-to-noise ratio and can efficiently reach a maximum survey depth. For the 10^4 s exposure time we employ the full sub-array with 13 groups per integration, 1 integration per exposure, 4 exposures per specification. For the 10^5 s exposure time we employ 20 groups per integration, 1 integration per exposure, 24 exposures per specification. The aperture radius is set to $0.04''$ for the short-wavelength filters (F070W, F090W, F115W, F150W, F200W) and $0.08''$ for the long-wavelength filter (F277W, F356W, F444W). The aperture used for imaging aperture photometry is circular, which specifies the radius of the circular aperture used for extraction. The background subtraction is performed with a sky annulus with inner and outer radii of $0.6''$ and $0.99''$ for short-wavelength bands, $1.2''$ and $1.98''$ for long-wavelength bands. The ETC background model includes celestial sources (zodiacal light, interstellar medium, and cosmic infrared background) along with telescope thermal and scattered light. This background model varies with the target coordinates (RA, Dec) and time of year. Here, we choose the fiducial background at RA = 17:26:44, Dec = -73:19:56 on June 19, 2019³. We note that some authors employ background configurations with a pre-calculated percentile of the range of backgrounds predicted at

² <https://jwst.etc.stsci.edu/>

³ <https://jwst-docs.stsci.edu/display/JTI/NIRCcam+Imaging+Sensitivity>

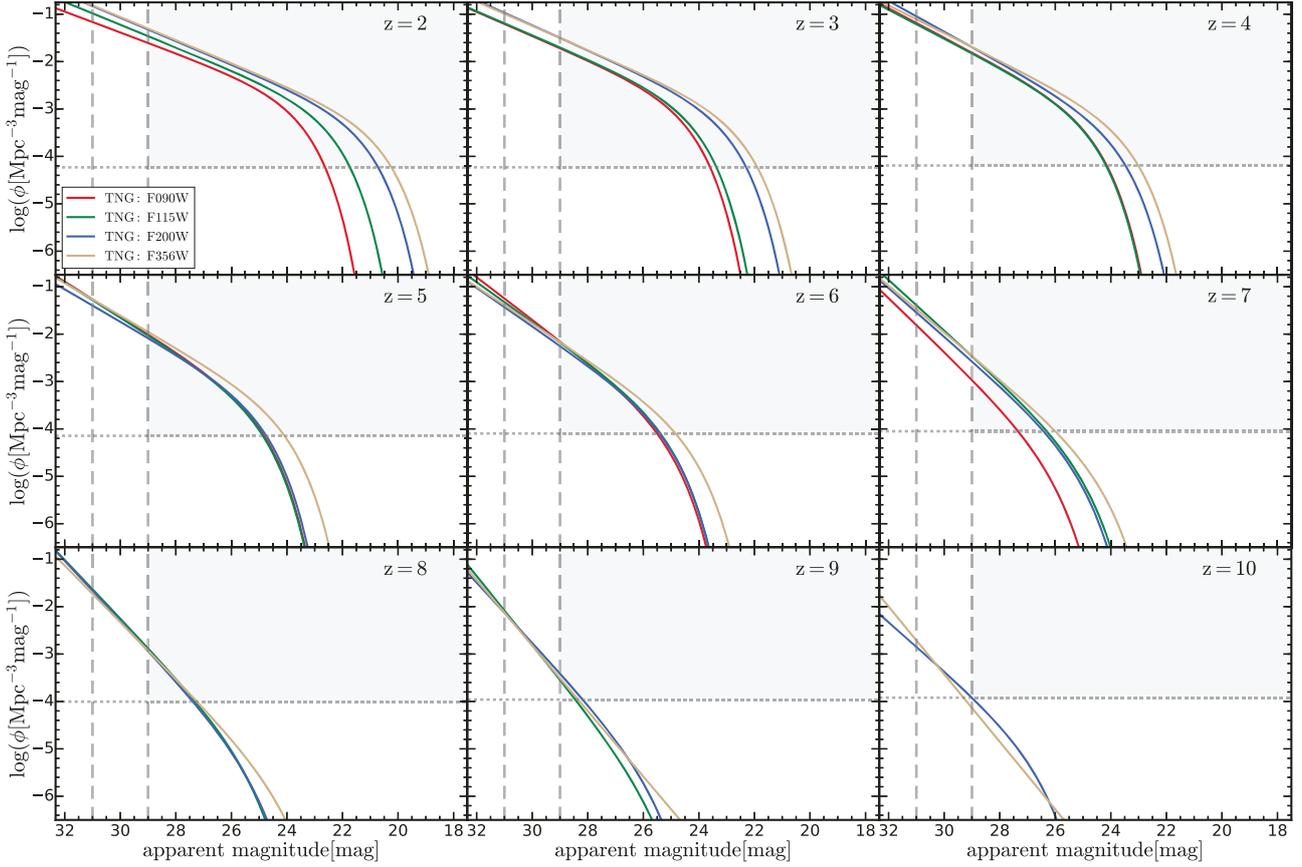


Figure 14. Schechter fits from the TNG apparent luminosity functions of four selected NIRCcam wide filters (F090W, F115W, F200W, F356W) at $z = 2 - 10$. Vertical dashed line show two approximate JWST detection limits: ~ 29 mag for an exposure time of 10^4 s and a target signal-to-noise ratio of 10; ~ 31 mag for an exposure time of 10^5 s and a target signal-to-noise ratio of 5. Horizontal dashed lines represent the number density required for the detection of one object per magnitude per unit redshift in the JWST NIRCcam field of view.

NIRCcam wide filter	wavelength [μm]	bandwidth [μm]	$m_{\text{SNR}=10, T_{\text{exp}}=10^4\text{s}}^{\text{lim}}$ [mag]	$m_{\text{SNR}=10, T_{\text{exp}}=10^5\text{s}}^{\text{lim}}$ [mag]	$m_{\text{SNR}=5, T_{\text{exp}}=10^4\text{s}}^{\text{lim}}$ [mag]	$m_{\text{SNR}=5, T_{\text{exp}}=10^5\text{s}}^{\text{lim}}$ [mag]
F070W	0.704	0.132	28.46	29.69	29.24	30.46
F090W	0.902	0.194	28.79	30.02	29.57	30.78
F115W	1.154	0.225	29.02	30.24	29.80	31.00
F150W	1.501	0.318	29.27	30.49	30.05	31.25
F200W	1.989	0.457	29.34	30.55	30.12	31.31
F277W	2.762	0.683	29.06	30.25	29.83	31.01
F356W	3.568	0.781	29.06	30.25	29.83	31.01
F444W	4.408	1.029	28.61	29.78	29.37	30.55

Table 5. JWST NIRCcam wide filter characteristics and detection limits. The table contains the pivot wavelengths, bandwidths and detection limits of the JWST NIRCcam wide filters adopted in this work. The detection limits are calculated assuming 10^4 s and 10^5 s exposure time with target signal-to-noise ratios of 5 and 10 based on the JWST Exposure Time Calculator.

the specified sky position over the window of visibility. For example, Yung et al. (2018) chose the percentile to be 10% referred to as “low” level configuration. For each band, we then set up all the ETC parameters as described and then vary the apparent magnitude of the source until a target signal-to-noise ratio is reached. This then sets the corresponding apparent magnitude detection limit for this band for various exposure times and target signal-to-noise ratios.

The final result of this procedure are the detection limits listed in Table 5.

In the upper left panel of Figure 15, we compare the JWST F200W band luminosity functions at all redshifts with previous theoretical predictions. The results from Yung et al. (2018) are based on a semi-analytical galaxy formation model along with a dust model similar to our Model B to provide JWST predictions

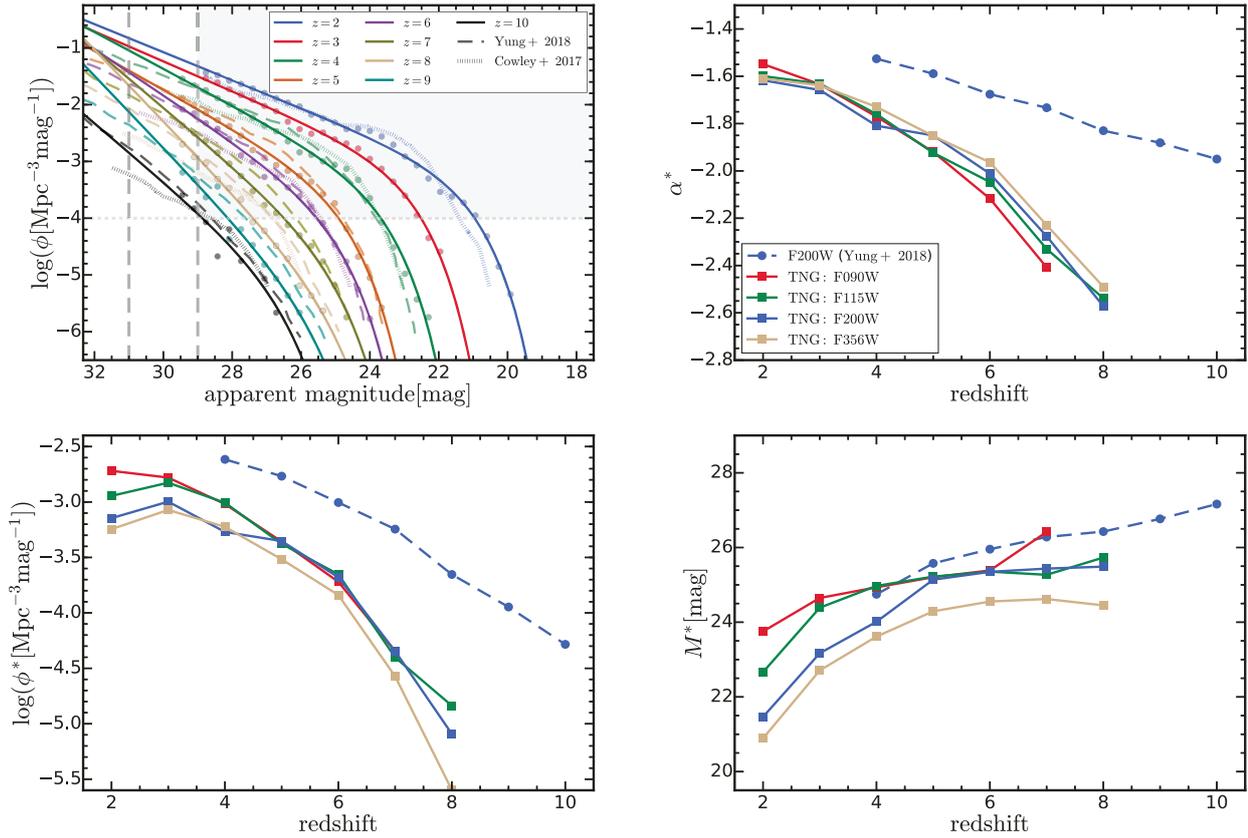


Figure 15. F200W band luminosity functions and best-fit Schechter function parameters for the predicted NIRCcam wide filter apparent luminosity functions. *Upper left:* *JWST* NIRCcam F200W band apparent luminosity functions at different redshifts. Symbols show the actual simulation data that determines the fits. We compare IllustrisTNG Schechter fits of apparent *JWST* luminosity function predictions with Schechter fits of Yung et al. (2018) and predictions from Cowley et al. (2018). Vertical and horizontal dashed lines have the same meaning as in Figure 14 but the detection limit in number density is a rough value ~ -4 for $z = 2 - 10$. *Upper right, lower left, lower right panels:* Best-fit Schechter function parameters at each redshift for selected wide filters. The best-fit Schechter parameters of Yung et al. (2018) for the F200W band are presented for comparison.

for *JWST* NIRCcam wide filters in the redshift range from $z = 4$ to $z = 10$. Cowley et al. (2018) combined the spectrophotometric radiative transfer code GRASIL (Silva et al. 1998) with the semi-analytic GALFORM code to model the impact of dust emission and absorption. They provide predictions for deep galaxy surveys with the *JWST* NIRCcam and MRI instruments. Our results are roughly consistent with those of Yung et al. (2018) but we predict lower number densities at the intermediate luminosity range (-28 mag to -26 mag) at $z \geq 5$. Our results are consistent with those of Cowley et al. (2018) towards the bright and intermediate luminosity range but differ at the faint end, where we predict much higher number densities. We note however that the predictions of Cowley et al. (2018) have a strong dependence on their feedback model. For example, their evolving feedback model predicts much higher galaxy number densities at the faint end. In the upper left panel of Figure 15, we also plot the approximate detection limits: ~ 31 mag for 10^3 s and signal-to-noise ratio 5, ~ 29 mag for 10^4 s and signal-to-noise ratio 10. We note that the numerical resolution of TNG50 is not sufficient to probe the luminosity function down to ~ 31 or beyond.

The Schechter fits of the apparent *JWST* luminosity functions are reliable up to $z = 8$. At $z = 9$ and $z = 10$, we are not able to derive reliable Schechter fit parameters due to the limited statistical sample similar to the limitation when deriving UV luminosity

function Schechter fits. The redshift evolution of the Schechter fit parameters are presented in Figure 15. We note that once the luminosity function of a given band drops out of the detectable regime, the shaded regions in Figure 14, we do not show the parameters anymore. We predict a steeper faint-end slope than the theoretical predictions of Yung et al. (2018), and we also find lower ϕ^* normalisation values. We note here however that the UV faint-end slope predicted by Yung et al. (2018) is also shallower than most observational data and other theoretical predictions.

4.3 *JWST* predictions

Based on the best-fit Schechter functions, we can integrate the apparent *JWST* luminosity functions down to the detection limit to derive the cumulative number density above that detection limit following Equation 29. We can then calculate the expected number of galaxies per unit redshift in the NIRCcam field of view as:

$$\frac{dN_{\text{exp}}}{dz} = \phi_{\text{cum}}(< m^{\text{lim}}) \frac{dV_{\text{com}}}{d\Omega dz}(z) \Delta\Omega, \quad (30)$$

where $\phi_{\text{cum}}(< m^{\text{lim}})$ is the cumulative number density calculated with Equation 29, $\Delta\Omega$ is the solid angle corresponding to the field of view of the *JWST* NIRCcam instrument (2.2×2.2 arcmin 2) and

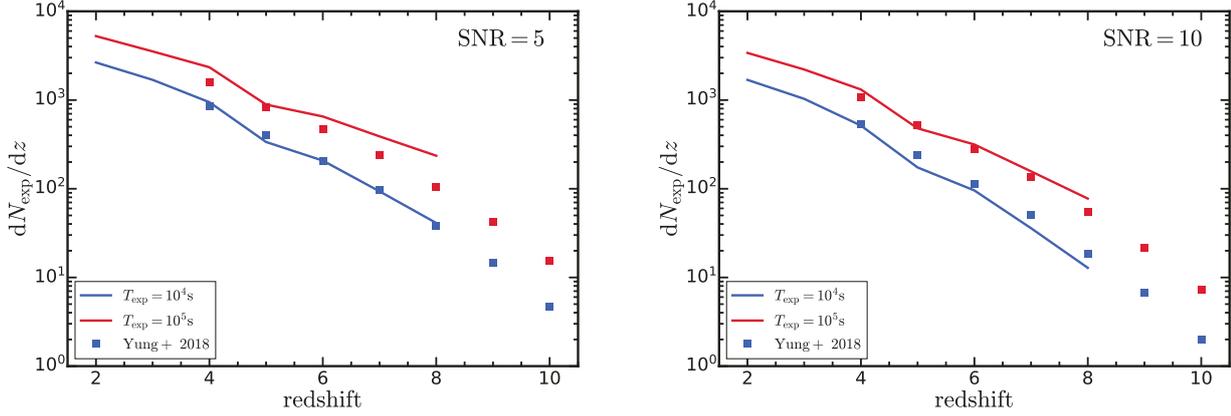


Figure 16. Expected number of galaxies per unit redshift observed in the NIRCcam F200W band in the NIRCcam field of view ($2.2 \times 2.2 \text{ arcmin}^2$). We present the results as a function of redshift with exposure times of 10^4 s and 10^5 s . The left and the right panel show the expected numbers assuming two different target signal-to-noise ratios: 5 (*left panel*) and 10 (*right panel*). Our predictions are roughly consistent with those from Yung et al. (2018). However, for a signal-to-noise ratio of 5 and $T_{\text{exp}} = 10^5 \text{ s}$ we predict a larger number of galaxies for the different redshifts. We note however that this setup has the faintest detection limit, $\sim 31 \text{ mag}$. Since our luminosity function cannot reach that low detection limit, our predictions here largely rely on an extrapolation. The steeper faint end slope we find then leads to the slight over-prediction compared with Yung et al. (2018).

$dV_{\text{com}}/d\Omega dz$ is the differential comoving volume element calculated as:

$$\frac{dV_{\text{com}}}{d\Omega dz}(z) = \frac{c(1+z)^2 d_A(z)^2}{H_0 E(z)}, \quad (31)$$

where $d_A(z)$ is the angular diameter distance and $H(z) = H_0 E(z)$ is the Hubble parameter at redshift z . We note that we can also use this equation to inversely derive the number density detection limit corresponding to only one object per magnitude per unit redshift in the field of view of the *JWST* NIRCcam instrument. This limit is plotted in Figure 14 and the upper left panel of Figure 15 as horizontal dashed lines. The predicted number counts of galaxies for the NIRCcam F200W band in the *JWST* NIRCcam field of view are presented in Figure 16. Here we also compare those values with results derived from the best-fit Schechter functions of Yung et al. (2018). For the left (right) panel, we assume a target signal-to-noise ratio of 5 (10). For a short exposure time, 10^4 s , we predict consistent numbers with Yung et al. (2018), despite some deficit at $z > 7$, where our luminosity function is lower. For a longer exposure time, 10^5 s , we predict slightly higher numbers since we find a steeper faint-end slope in the derived Schechter fits. However, here we also start approaching a regime where the detection limit reaches or is already beyond our numerical resolution limit. Due to this limitation in numerical resolution, we do not provide predictions for longer exposure times, e.g., 10^6 s , which would result in even lower apparent magnitudes as detection limits. We note that the expected number of detected galaxies in the field of view rises monotonically towards lower redshifts.

We can also use the *JWST* galaxy luminosities to make concrete predictions for the number counts of *JWST* NIRCcam high redshift dropout surveys. Following earlier theoretical work by Tacchella et al. (2018), we focus here on two extragalactic surveys that are currently planned for *JWST*. The first survey is a large (720 hours) observational program, the *JWST* Advanced Deep Extragalactic Survey (JADES). The second is the Cosmic Evolution Early Release Science Survey (CEERS). We use our apparent galaxy UV magnitudes, m_{UV} , to make predictions for galaxy number counts for these future *JWST* surveys. Using the estimates in Tacchella et al. (2018), the deep component of JADES (JADES-

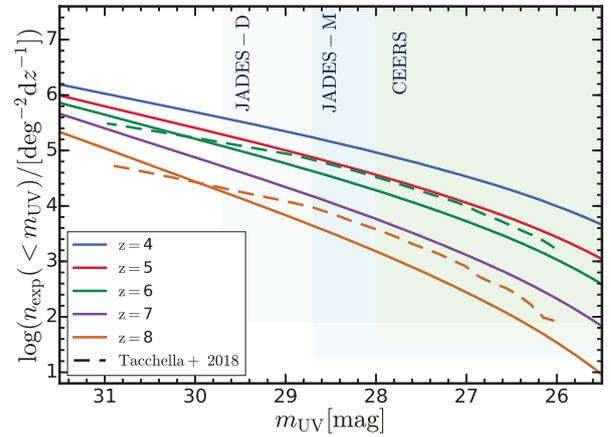


Figure 17. Predictions for future *JWST* NIRCcam high redshift dropout surveys. We present the predicted number of galaxies per square degree per unit redshift at $z = 4 - 8$ for galaxies brighter than a given apparent magnitude limit m_{UV} , which is the apparent magnitude evaluated at the wavelength where the rest-frame UV light is redshifted to. The shaded regions indicate the coverage of planned *JWST* surveys. The dashed lines show the prediction from Tacchella et al. (2018) for comparison.

D) will have an averaged survey depth of $\sim 29.7 \text{ mag}$ for all bands assuming a 10σ point source limit with a sky coverage of 46 arcmin^2 . The medium component JADES-M will have an averaged depth of 28.7 mag with a sky coverage of 190 arcmin^2 . The CEERS program will have an averaged depth of 28.0 mag with a sky coverage of 100 arcmin^2 . In Figure 17, we compare our predicted galaxy number counts with the coverage of these *JWST* surveys. The horizontal lower boundaries of the survey coverage regions are calculated assuming that there is only one object per unit redshift in the entire sky coverage of the surveys. We find that JADES-D can detect ~ 3 galaxies per unit redshift at the bright end up to $m_{\text{UV}} \sim 27 \text{ mag}$ at $z = 8$, while having a detection limit $m_{\text{UV}} \sim 29.7 \text{ mag}$ at the faint end. The total number of detectable galaxies can reach 235 per unit redshift at $z = 8$. JADES-M has a

larger sky coverage and can detect a ~ 4 times larger number of galaxies at the bright end. The total number of detectable objects is 238 comparable to that of JADES-D at $z = 8$. Our results are roughly consistent with the predictions of Tacchella et al. (2018) but we find a slightly different slope of the expected number of detected galaxies as a function of apparent magnitude. We conclude that the IllustrisTNG simulation suite can provide important and insightful predictions for these kind of upcoming *JWST* surveys.

4.4 UV luminosity versus stellar mass relation

So far we have only discussed galaxy luminosity functions, but the galaxy stellar mass function is actually the more fundamental property to quantify the galaxy population, and also a direct prediction of any galaxy formation model. However, deriving galaxy stellar mass functions from observed luminosity functions is a challenging task since the conversion from stellar light to stellar mass depends on multiple assumptions for photometry, redshift determination, stellar population synthesis modelling, star formation histories, metallicities, dust characteristics, and initial stellar mass functions. This introduces significant systematic uncertainties in the whole conversion process (e.g., Conroy & Gunn 2010). Especially, measuring the stellar masses of galaxies at high redshifts is challenging, because the rest-frame wavelengths probed by optical and near-infrared surveys cover only to the UV continuum requiring mid-IR observations to extend past a rest-frame wavelength of 4000\AA . In addition, selection effects, contamination and incompleteness of the galaxy sample make it much more difficult to constrain the galaxy stellar mass function at higher redshifts. An alternative approach is to convert well-constrained galaxy UV luminosity functions to a stellar mass function assuming a certain mass-to-light ratio relation. The advantage of UV luminosity functions is that they can typically be corrected for selection effects and provide good statistics at both the bright and faint end. Many observations at high redshifts revealed that the galaxy UV luminosity correlates well with the stellar mass of the galaxy although the relation contains also quite significant scatter (e.g., Stark et al. 2009, 2013; Duncan et al. 2014; Song et al. 2016). Once the mass-to-light ratio relation has been derived, the relation can be used to map observed UV luminosity functions to galaxy stellar mass functions without going through the demanding and uncertain SED modelling task. Specifically, one can construct galaxy stellar mass functions conveniently by convolving the $M_* - M_{\text{UV}}$ distribution with the observed rest-frame UV luminosity function. For example, González et al. (2011) derived the stellar masses for 401 star forming galaxies by fitting the observed optical to UV SEDs. With these samples, they constructed a stellar mass versus UV luminosity relation, which they then convolved with well-constrained UV luminosity functions to derive galaxy stellar mass functions for redshifts $z = 4 - 7$.

Based on our simulation predictions, we can also construct a $M_* - M_{\text{UV}}$ relation by combining the IllustrisTNG simulations, TNG50, TNG100 and TNG300. We can then compare this predicted relation to the observed relation as a further test of our galaxy formation and dust model. In Figure 18, we present the predicted relation between galaxy stellar masses and their UV magnitudes at $z = 4, 5, 6$. To derive this relation, we employ the same binning in UV magnitudes as for the rest-frame UV luminosity function analysis presented above. For each bin, we then calculate the median $\log(M_*/M_\odot)$ and the standard deviation of $\sigma_{\log M_*}$ for the TNG50, TNG100 and TNG300 simulations respectively. Here, the stellar masses have been resolution-corrected following the same proce-

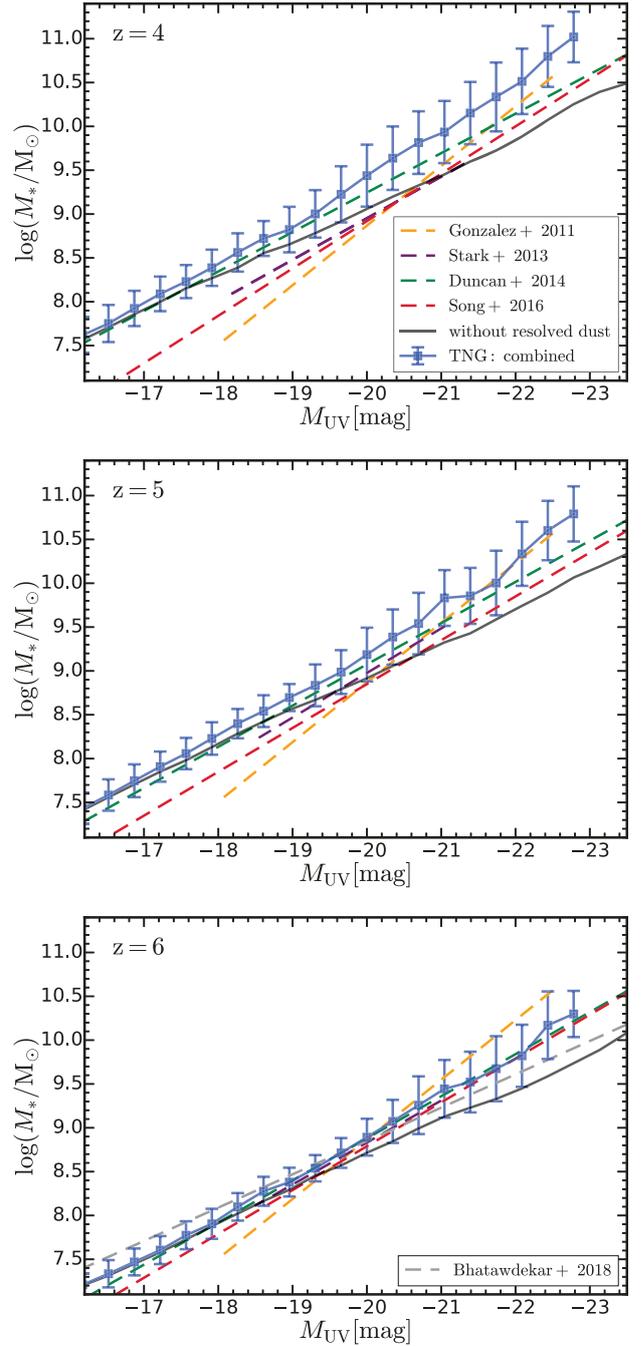


Figure 18. UV luminosity versus stellar mass relations at different redshifts. The galaxy stellar mass is defined here within a fixed physical aperture of 30 pkpc. The blue line shows the median $\log(M_*/M_\odot)$ within each dust-attenuated UV magnitude bin. The error bars represent the $\log(M_*/M_\odot)$ standard deviation of the samples in each bin. The black line indicates the relation when no resolved dust attenuation is taken into account. Observational constraints are taken from González et al. (2011); Stark et al. (2013); Duncan et al. (2014); Song et al. (2016); Bhatwdekar et al. (2018). For González et al. (2011); Stark et al. (2013); Song et al. (2016), the initial stellar mass function is converted from a Salpeter to a Chabrier initial stellar mass function by decreasing $\log(M_*/M_\odot)$ by 0.24.

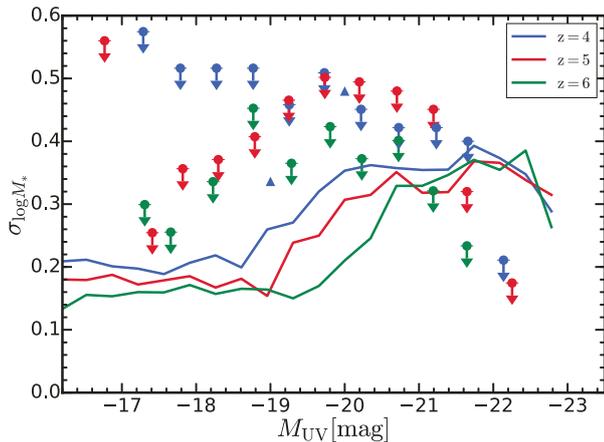


Figure 19. Scatter in the UV luminosity versus stellar mass relation at $z = 4, 5, 6$. The observational upper limits are taken from Song et al. (2016) where the scatter in each UV magnitude bin is measured and includes both intrinsic scatter and photometric and model uncertainties. The solid triangles are taken from González et al. (2011) where the photometric uncertainties at $M_{UV} = -20$ mag and at $M_{UV} = -19$ mag are quantified. The intrinsic scatter is then calculated as $\sigma_{\text{int}} = (\sigma_{\text{tot}}^2 - \sigma_{\text{pho}}^2)^{1/2}$. The scatter is larger towards lower redshifts.

ture that has also been used for the galaxy luminosities discussed above. Then similar to the combination step for the luminosity functions, we combine the median $\log(M_*/M_\odot)$ or $\sigma_{\log M_*}$ using Equation 27 in the luminosity range covered by two or three simulations, replacing ϕ with the median of $\log(M_*/M_\odot)$ or $\sigma_{\log M_*}$. At the faint end, where TNG100 (TNG300) does not provide sufficient resolution and therefore deviates from TNG50 (TNG100) predictions for the relation, we only use TNG50 (TNG50 and TNG100) to construct the combined $M_* - M_{UV}$ relation. In the combination step we do not consider bins with less than 10 galaxies for TNG300. The combined median $\log(M_*/M_\odot)$ versus M_{UV} relation is shown as the solid line in Figure 18. The error bars represent the standard deviation in $\log(M_*/M_\odot)$ of the samples in each bin. We note that stellar mass is here defined as the total mass of stellar particles within a radial aperture of 30 pkpc to be consistent with the UV luminosity calculation discussed above. Figure 18 reveals a strong correlation between stellar mass and UV luminosity. At the faint end, our results are consistent with Duncan et al. (2014), but not with González et al. (2011), Stark et al. (2013) and Song et al. (2016) at $z = 4, 5$. However, the number of galaxies at the faint end $M_{UV} > -18$ in these observations are very limited. The observational relations presented here have also some significant scatter, which we do not show for clarity of the figure. The discrepancies at the bright end are likely related to the fact that observations are based on UV selected galaxies and have a potential bias missing dusty star-forming galaxies and quiescent galaxies. This can lead to a shift in the relation towards UV brighter galaxies at a given stellar mass. At the bright end, the relation is no longer linear and deviates towards lower UV luminosity due to the influence of dust. This can be appreciated in the figure through a comparison with the dust-free lines. The relation is tighter towards higher redshifts due to younger stellar populations with dominant UV emission. Towards lower redshifts, the scatter increases since the stellar populations become older, which weakens the correlation between stellar mass and UV luminosity causing a visible larger scatter.

In Figure 19 we quantify this scatter in the mass-to-light ratio

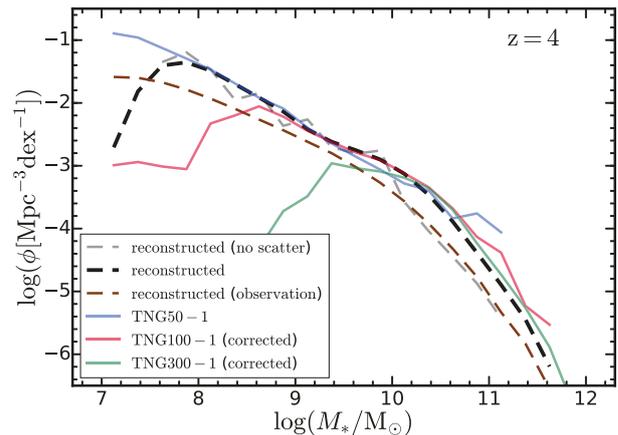


Figure 20. Reconstructed galaxy stellar mass function through a convolution of the rest-frame UV luminosity function with the UV luminosity versus stellar mass relation at $z = 4$. Solid curves show the stellar mass function directly from resolution corrected simulation outputs. The black dashed line represents the reconstructed stellar mass function using the $M_* - M_{UV}$ relation including its scatter. We also show the reconstructed stellar mass function using same relation without this scatter as a gray dashed line. The brown dashed line shows the reconstructed stellar mass function using an observationally determined $M_* - M_{UV}$ relation and scatter taken from Song et al. (2016).

relation in more detail by presenting the standard deviation in the stellar mass to UV magnitude relation at $z = 4, 5, 6$. This scatter increases from the faint end towards the bright end. Observations typically find that the scatter is around 0.3–0.5 dex at $z = 4–7$, which is generally larger than our values. For example, Song et al. (2016) found a stellar mass standard deviation of about 0.4 dex. Specifically, they found standard deviations of 0.52, 0.42, 0.36, 0.40, and 0.30 dex at $z = 4, 5, 6, 7$, and 8, respectively. Those were measured as the mean standard deviations in stellar mass bins with more than five galaxies. They found no noticeable correlation of the scatter with redshift or UV luminosity. The scatter at the bright end (measured at $-21.5 \text{ mag} < M_{UV} < -20.5 \text{ mag}$), where the effect of observational uncertainty should be minimal, is 0.43, 0.47, 0.36, 0.52, and 0.40 dex at $z = 4, 5, 6, 7$, and 8, respectively, similar to the quoted value above and the scatter at the faint end (at $-19.0 \text{ mag} < M_{UV} < -18.0 \text{ mag}$) of 0.51, 0.39, 0.39, and 0.41 dex at $z = 4, 5, 6$, and 7, respectively. We note that these scatter measurements represent upper limits since there are significant photometric and model uncertainties on top of the intrinsic scatter. González et al. (2011) found that the scatter in this relation at $z = 4$ is about 0.5 dex. They also estimated that the photometric uncertainty is about 0.14 dex at $M_{UV} = -20$ mag and 0.37 dex at $M_{UV} = -19$ mag. Considering these photometric uncertainties, the intrinsic scatter is close to what we find. Stark et al. (2013) estimated the intrinsic scatter to be in the range 0.2–0.5 dex at $z = 4$ due to photometric and model uncertainties.

Based on the mass-to-light ratio relation, we can also reconstruct the galaxy stellar mass function by convolving the derived UV luminosity function with the derived stellar mass versus UV magnitude relation. This follows the common approach to convert observed rest-frame UV luminosity functions into stellar mass functions. To this end, we take the binned $\log(M_*/M_\odot)$ versus M_{UV} relation with its scatter and convolve it with the combined luminosity functions derived above. Specifically, in each UV mag-

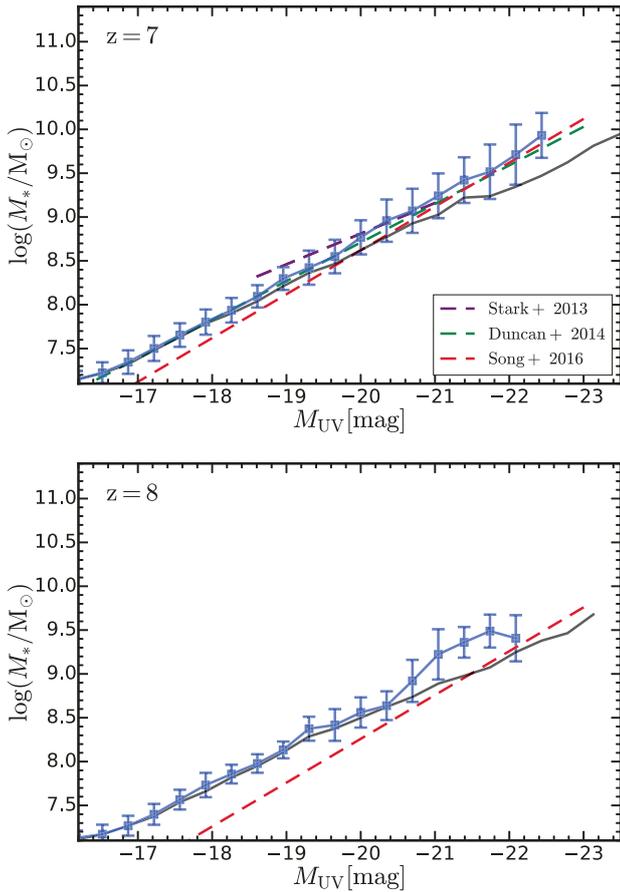


Figure 21. UV luminosity versus stellar mass relations at high redshifts. The galaxy stellar mass is defined here within a fixed physical aperture of 30 pkpc. The blue line shows the median $\log(M_*/M_\odot)$ within each dust-attenuated UV magnitude bin. The error bars represent the $\log(M_*/M_\odot)$ standard deviation of the samples in each bin. The black line indicates the relation when no dust attenuation is taken into account. Observational constraints are taken from [Stark et al. \(2013\)](#); [Duncan et al. \(2014\)](#); [Song et al. \(2016\)](#).

nitude bin, we generate stellar mass samples with a normal distribution in $\log(M_*/M_\odot)$ centred around the median value of the $\log(M_*/M_\odot)$ versus M_{UV} relation with its scatter. The total number of the samples is proportional to the galaxy number density at the given UV magnitude. By combining the samples derived from each UV magnitude bin, we generate a reconstructed galaxy stellar mass function, with 21 bin edges linearly spaced from 7 to 12 for $\log(M_*/M_\odot)$. In [Figure 20](#), we present the reconstructed galaxy stellar mass function and compare it with the directly derived resolution-corrected stellar function. We find good agreement between those two galaxy stellar mass functions over the entire dynamic range. We also plot the reconstructed function without including the scatter in the $\log(M_*/M_\odot)$ versus M_{UV} relation. In this case we find that neglecting the scatter significantly underestimates the galaxy number densities at the massive end. Using the observationally determined $\log(M_*/M_\odot)$ versus M_{UV} relation along with its scatter ([Song et al. 2016](#)) to reconstruct the stellar mass function, we find that the stellar mass is underestimated in the entire dynamic range consistent with the discrepancy we find in $\log(M_*/M_\odot)$ versus M_{UV} relation discussed above.

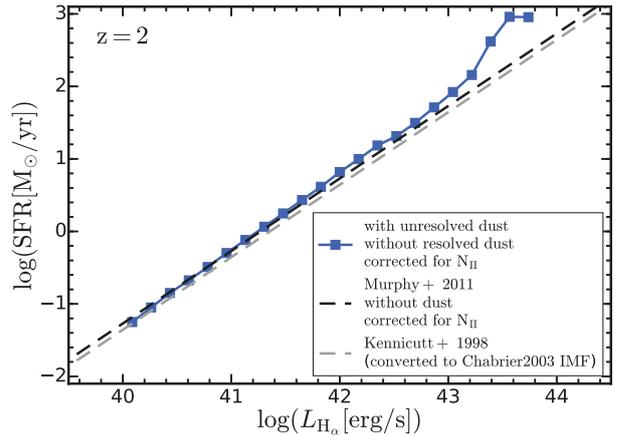


Figure 22. $H\alpha$ luminosity versus star formation rate relation at $z = 2$. We compare the relation to the observational results from [Murphy et al. \(2011\)](#), which is based on the observation of local star-forming regions in nearby galaxies. The simulation results are shown at $z = 2$. In addition we also show the [Kennicutt \(1998\)](#) observational relation converted to a Chabrier initial stellar mass function.

Finally, we can also use the simulation results to make predictions for the mass-to-light ratio relation towards higher redshifts, where observational data becomes more scarce and uncertain. [Figure 21](#) presents the $z = 7$ and $z = 8$ mass-to-light ratio relations and compares those with some observational data. At $z = 7$ we find good agreement between the simulation predictions and observational data. At $z = 8$ observations have not yet fully explored the mass-to-light ratio relation, and only limited data is available. This data is, as can be seen in [Figure 21](#), not consistent with our predictions at this high redshift.

4.5 $H\alpha$ luminosity versus star formation rate relation

The UV luminosity of a galaxy is an indicator for the stellar mass of the galaxy through the mass-to-light ratio relation as demonstrated above. Furthermore, the UV luminosity can also be employed as a star formation rate indicator. Spectral indices like the $H\alpha$ emission line luminosity, on the other hand, provide further indicators for star formation rates of galaxies ([Kennicutt 1998](#)). Young, massive stars produce large amounts of ionising photons that ionise the surrounding gas. Hydrogen recombination cascades then produce line emission, including the well-known Balmer series lines of $H\alpha$ (656.3 nm) and $H\beta$ (486.1 nm). These lines are strong and represent the most traditional star formation rate indicators. Above we have explored the correlation between UV luminosities and stellar masses. Here we now explore how $H\alpha$ luminosities correlate with star formation rates of galaxies (e.g., [Murphy et al. 2011](#)). Probing these UV and $H\alpha$ luminosity relations and comparing them to the observed relations provides important tests of our galaxy formation model and the employed luminosity and dust modelling. Specifically, they probe relations of the continuum, UV, and emission lines, $H\alpha$.

To derive the $H\alpha$ emission line luminosity of a given galaxy, we follow an approach commonly adopted in observations (e.g. [Lee et al. 2012](#); [Sobral et al. 2013, 2015](#); [Matthee et al. 2017](#)). We first construct broad band and narrow band tophat filters around the $H\alpha$ emission line. The broad band filter is centred at 656.3 nm with

a band width of $\Delta\lambda_{\text{BB}} = 80$ nm. The narrow band filter is also centred at 656.3 nm but with a band width of $\Delta\lambda_{\text{NB}} = 20$ nm. Both of these filters cover the entire rest-frame $\text{H}\alpha$ emission line. We then convolve the galaxy SED with these two filters and get the filter averaged fluxes f_{BB} (broad band) and f_{NB} (narrow band). We can then derive the line luminosity and equivalent width (EW) as:

$$L_{\text{H}\alpha} = \Delta\lambda_{\text{NB}} \frac{f_{\text{NB}} - f_{\text{BB}}}{1 - \frac{\Delta\lambda_{\text{NB}}}{\Delta\lambda_{\text{BB}}}}, \quad \text{EW} = \Delta\lambda_{\text{NB}} \frac{f_{\text{NB}} - f_{\text{BB}}}{f_{\text{BB}} - f_{\text{NB}} \frac{\Delta\lambda_{\text{NB}}}{\Delta\lambda_{\text{BB}}}}. \quad (32)$$

We note that the two N_{II} emission lines, located at 654.8 nm and 658.3 nm, contaminate the $\text{H}\alpha$ luminosity and equivalent width. To correct for this, we assume that the ratio between the N_{II} luminosity and the $\text{H}\alpha$ luminosity depends linearly on the total measured equivalent width with contributions from the lines of N_{II} and $\text{H}\alpha$ (Sobral et al. 2015):

$$f(\text{N}_{\text{II}}/\text{H}\alpha) = -0.296 \times \log(\text{EW}_{\text{N}_{\text{II}}+\text{H}\alpha}) + 0.8, \quad (33)$$

where the equivalent width is measured in units of \AA . Negative f are manually set to 0. We then calculate the $\text{H}\alpha$ line luminosity as Sobral et al. (2015):

$$L_{\text{H}\alpha} = L_{\text{N}_{\text{II}}+\text{H}\alpha} \times \frac{1}{1 + f(\text{N}_{\text{II}}/\text{H}\alpha)}. \quad (34)$$

To probe the $\text{H}\alpha$ luminosity versus star formation rate relation, we compute the instantaneous star formation rate of each galaxy within a physical aperture of 30 kpc consistent with the aperture applied above for galaxy luminosities and stellar masses. We present the result of this calculation in Figure 22 for $z = 2$ along with the observationally inferred linear relation taken from Murphy et al. (2011):

$$\left(\frac{\text{SFR}_{\text{H}\alpha}}{M_{\odot} \text{ yr}^{-1}} \right) = 5.37 \times 10^{-42} \left(\frac{L_{\text{H}\alpha}}{\text{erg s}^{-1}} \right) \quad (35)$$

This relation is a recalibration of the relation from Kennicutt (1998) and is dust-corrected. We therefore do not consider the impact of resolved dust, when deriving $\text{H}\alpha$ luminosities. We note that the emission lines, and so the $\text{H}\alpha$ line, are modelled through MAPPINGS-III assuming a Kroupa initial stellar mass function. This is consistent with the Murphy et al. (2011) relation that assumes the same initial stellar mass function. Figure 22 demonstrates that our results reproduce the observed relation. We also include in this figure the Kennicutt (1998) relation converted to a Chabrier initial stellar mass function, which has a marginally different normalisation compared to the Murphy et al. (2011) relation.

To explore the impact of dust on the $\text{H}\alpha$ emission, we next present the $\text{H}\alpha$ luminosity function at $z = 2$ along with a comparison to observational results in Figure 23. The solid blue line shows the luminosity function without resolved dust attenuation. The solid red line shows the luminosity function with resolved dust attenuation assuming dust-to-metal ratios of 0.9, i.e. the best-fit value based on the UV calibration. Not including the N_{II} line correction results in the dashed lines. Neglecting the resolved dust component clearly leads to an inconsistent $\text{H}\alpha$ luminosity function. Including the resolved dust on the other hand, moves the luminosity function towards significantly lower luminosities. We note that a dust-to-metal ratio of 0.9 leads to slightly too much attenuation of the $\text{H}\alpha$ fluxes, which can be seen in Figure 23 towards the bright end, where the simulation predictions start deviating from the observational data. Nevertheless, our results are still marginally consistent within the observational uncertainties. We also stress that a dust-to-metal ratio of 0.9 represents the best-fit to the UV luminosity

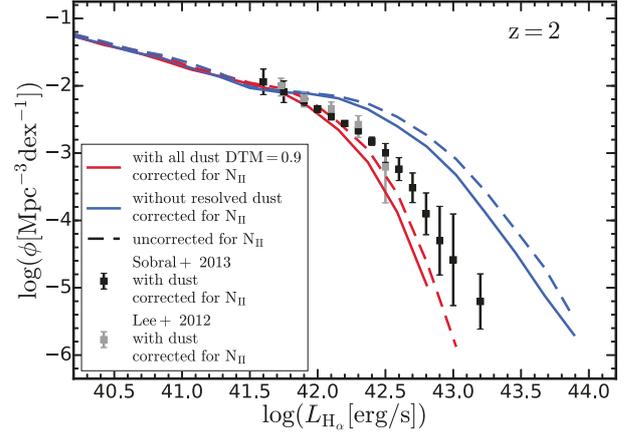


Figure 23. The $\text{H}\alpha$ luminosity function at $z = 2$. We present the luminosity functions of the $\text{H}\alpha$ emission evaluated at $z = 2$ and compare to observational data from Lee et al. (2012); Sobral et al. (2013). The solid blue line shows the luminosity function without resolved dust attenuation. The solid red line shows the luminosity function with resolved dust attenuation assuming a dust-to-metal ratios of 0.9. Not including the N_{II} line correction results in the dashed lines.

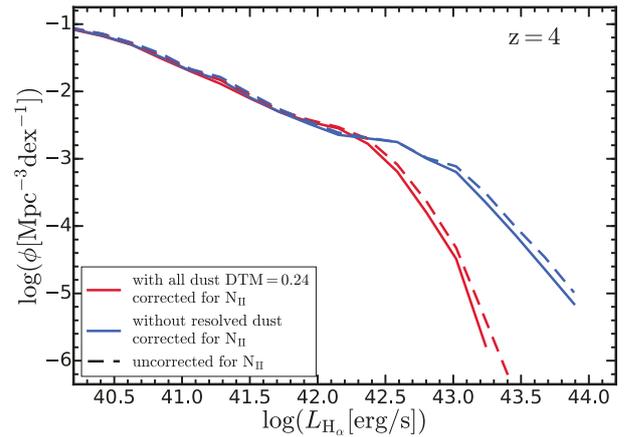


Figure 24. The $\text{H}\alpha$ luminosity function at $z = 4$. The solid blue line shows the luminosity function without resolved dust attenuation. The solid red lines shows the luminosity function with resolved dust attenuation assuming dust-to-metal ratios of 0.24. Not including the N_{II} line correction results in the dashed lines.

function at $z = 2$ as discussed above, and lower dust-to-metal ratios would lead to an inconsistent UV luminosity function. We note that the unresolved attenuation through the MAPPINGS-III model is calibrated only for the local universe and uncertainties arise when applying this towards higher redshifts. This can possibly lead to a too high attenuation as is shown in Figure 23. Measuring the $\text{H}\alpha$ luminosity function towards higher redshifts is observationally challenging. We therefore present in Figure 24 the predicted $\text{H}\alpha$ luminosity function at $z = 4$. We show here the luminosity function again with and without resolved dust assuming the best-fit dust-to-metal ratio as derived from the dust calibration step.

We conclude that our modelling of emission lines leads to con-

sistent $H\alpha$ relations. This provides another important test of our luminosity and dust attenuation modelling.

5 SUMMARY AND CONCLUSIONS

JWST is going to transform our understanding of the high redshift Universe. However, exploiting the unprecedented capabilities of this novel telescope depends on reliable theoretical predictions. Such predictions for the properties of high redshift galaxies are therefore of uttermost importance, and detailed cosmological simulations of the early Universe are required to provide those forecasts. Ideally, these cosmological simulations should be based on galaxy formation models that produce a realistic low redshift galaxy population and cover a wide dynamic range. The low redshift verification of galaxy formation models is necessary for the fidelity and consistency of the simulated high redshift Universe.

In this paper we have studied the high redshift galaxy population of the IllustrisTNG simulation suite consisting of TNG50, TNG100, and TNG300. Most importantly the IllustrisTNG simulations employ a galaxy formation model that produces a realistic low redshift Universe that is consistent with a wide range of observational data. The main focus of our work has been a detailed study of high redshift galaxy luminosity functions in various bands. To this end, we have combined the results of all three IllustrisTNG simulations, TNG50, TNG100, and TNG300, to maximise the dynamic range of these luminosity function predictions. We have devised a resolution correction and luminosity function combination procedure to predict combined luminosity functions over a large dynamic range. The resolution correction scheme extrapolates the TNG100 and TNG300 results to TNG50 resolution resulting in three simulations with the same effective numerical mass resolution. The luminosity function predictions of these simulations are then combined to achieve a large dynamic range in galaxy luminosities (e.g., $M_{UV} = -17$ to -22 at $z = 8$).

Dust attenuation severely affects luminosities of dust rich galaxies. We have therefore developed a variety of dust attenuation models for the IllustrisTNG galaxies to derive these luminosity functions in different bands and at different redshifts. Neglecting these corrections results in rest-frame UV luminosity functions that deviate significantly from observational data for redshifts $z \lesssim 5$ towards the bright end. We have therefore presented in this paper a detailed discussion of dust modelling for high redshift galaxies. Specifically, we have introduced three distinct dust attenuation models of increasing complexity and applied them in post-processing to the simulation output: an empirical dust model (Model A), a resolved dust optical depth model (Model B), and a resolved dust radiative transfer model (Model C). IllustrisTNG, similar to most other existing large volume galaxy formation simulations, does not follow the evolution of dust in galaxies. Each of the applied dust models therefore depends on some redshift-dependent free dust parameters that have to be determined through a calibration procedure:

- Dust Model A (empirical model): The two free parameters of Model A are the intercept β_{M_0} and the slope $d\beta/dM_{UV}$ values of the employed $M_{UV}^{\text{dust}} - \langle\beta\rangle$ relation.
- Dust Model B (resolved optical depth model): Model B contains an overall optical depth normalisation parameter for resolved dust.
- Dust Model C (resolved dust radiative transfer model): Similarly to Model B, the dust attenuation of Model C depends on one

parameter, the dust-to-metal ratio that sets at each redshift the conversion factor from the metal distribution in the simulated galaxies to the actual dust distribution required for the Monte Carlo radiative transfer calculation.

For Models B and C we consider both the impact of resolved and unresolved dust. The resolved dust component is modelled through resolved gas properties around each stellar particles. For the unresolved dust component we model the impact of birth clouds around young stars. For Model B this is handled through different optical depth normalisations and different extinction curves. The additional optical depth due to unresolved dust is applied to attenuate light from stellar particles younger than 10 Myr. For Model C, we instead use different stellar spectral energy distributions depending on whether stellar particles are surrounded by birth clouds or not. Specifically, the spectral energy distributions are altered for stellar particles younger than 10 Myr.

We have calibrated the free dust parameters of the three dust attenuation models based on a best-fit procedure against observed rest-frame UV luminosity functions. This dust model calibration procedure is based on the resolution-corrected TNG100 simulation. Specifically, we have calibrated the free dust parameters of all three dust models to match the observed rest-frame UV galaxy luminosity functions over the redshift range $z = 2$ to $z = 10$. The calibration process revealed for each dust model the redshift evolution of the free dust parameters constraining the dust models. We note that this calibration procedure results in redshift dependent dust parameters that are then applied to all galaxies at the corresponding redshift. For example, for Model C, we convert the metal abundances in each galaxy to dust abundances using a fixed dust-to-metal ratio at a given redshift. Based on this dust calibration we find:

- For dust Model A the inferred redshift evolution of the best-fit values for the intercept β_{M_0} agree reasonably well with the observationally derived redshift evolution. However, the redshift evolution of the slope $d\beta/dM_{UV}$ does not agree with observational estimates. Here our calibration prefers steeper slopes than what is observationally found.
- For dust Model B the calibration yields the redshift evolution of the optical depth normalisation. The redshift dependence of this normalisation can be described by a rather steep power law $z^{-1.92}$. This redshift evolution differs from other typically employed redshift dependencies of similar dust attenuation models especially towards lower redshifts ($z \lesssim 4$), where we predict a much stronger redshift dependence.
- For dust Model C the best-fit redshift evolution of the dust-to-metal ratios follows the same power law as the optical depth normalisation of Model B. For Model C we have also found that the additional unresolved birth cloud dust attenuation results in too much dust attenuation towards higher redshifts, $z \geq 6$. We therefore do not consider unresolved dust for $z \geq 6$ and only include the attenuation due to resolved dust. We also note that the resolved dust attenuation is negligible at higher redshifts due to the strong redshift dependence of the dust-to-metal ratio. For example, at $z = 8$ we predict a dust-to-metal ratio of about $\sim 5\%$. This results in high redshift luminosity functions that are essentially identical to the intrinsic dust-free luminosity functions. We note that our dust Model B also predicts that the contribution of unresolved dust is reduced at higher redshifts. For Model B this is a direct consequence of the employed proportionality between the resolved and unresolved optical depth in the model. Attenuation due to both components quickly vanishes towards higher redshifts due to the rather steep power law dependence of the re-

solved dust attenuation normalisation. Both dust models, Model B and Model C, therefore predict a vanishing unresolved dust attenuation contribution towards higher redshifts.

We stress that the redshift dependence of these dust models is crucial. Assuming, for example, a non-evolving optical depth normalisation or dust-to-metal ratio, results in inconsistent predictions for the rest-frame UV luminosity functions. Any high redshift galaxy luminosity predictions assuming constant dust-to-metal ratios are therefore questionable.

The dust calibration is the basis for our exploration of galaxy luminosity functions for a variety of bands. We first discussed the rest-frame UV luminosity functions, which agree for all our models well with currently existing UV luminosity function data in the redshift range $z = 2$ to $z = 10$. This result is largely a consequence of the dust calibration process, which is based on rest-frame UV luminosity functions. Upon closer inspection we find however also some important discrepancies.

- For dust Models A and B, towards lower redshifts ($z \lesssim 4$), the faint end of the rest-frame UV luminosity function deviates slightly from observational data, where IllustrisTNG, more specifically, TNG50 overpredicts the abundance of faint galaxies. Model C leads to a better agreement with the data, but the overprediction still occurs at $z = 3, 4$. For instance, at $z = 3, 4$ for Model C IllustrisTNG predicts about a ~ 0.2 dex too high galaxy number density at $M_{UV} = -16$ mag. These results are not affected by the employed resolution correction since this portion of the rest-frame UV luminosity function is probed solely by TNG50, where we do not apply any resolution correction.
- Towards higher redshifts we find that IllustrisTNG predicts slightly lower abundances of galaxies. This is most noticeable at $z = 8$, where we find lower abundances over a quite wide range of rest-frame UV luminosities from -18 mag to -21 mag. Dust attenuation is negligible at these redshifts. Therefore, these results are robust against dust modelling uncertainties.
- We fit the combined and dust-attenuated rest-frame UV luminosity functions with Schechter functions to quantify their shapes in more detail. This analysis revealed a rest-frame UV luminosity function faint-end slope, which is rather steep compared to observationally inferred values especially towards lower redshifts. For instance, at $z \sim 2$ we predict a best-fit slope around ~ -1.6 to ~ -1.8 for all three dust models. This value is steeper than typical observationally inferred values and also steeper than most other theoretical predictions. The observations usually suggest a shallower faint end slope around ~ -1.3 . This result is consistent with the slight overproduction of faint galaxies discussed above.

In addition to the rest-frame UV luminosity functions, we have also determined apparent galaxy luminosity functions for the NIRCam wide filters F070W, F090W, F115W, F150W, F200W, F277W, F356W and F444W based on dust Model C. We described all these apparent luminosity functions through Schechter fits and discuss the best-fit Schechter parameters redshift evolution.

- Similarly to the rest-frame UV luminosity function, our predictions for the faint-end slope of *JWST* NIRCam bands are also steeper than those of other theoretical predictions.
- Based on the derived best-fit Schechter functions, we then calculate concrete observational estimates for *JWST* NIRCam observations. Specifically, we provide predictions for the expected number of galaxies per unit redshift observed in the NIRCam F200W band within the 2.2×2.2 arcmin² field of view of the *JWST* NIRCam instrument. Our predictions are based on two

exposure times $T_{\text{exp}} = 10^4$ s and $T_{\text{exp}} = 10^5$ s and two different signal-to-noise ratios of 5 and 10. For example, at $z = 8 \pm 0.5$ we predict about ~ 200 galaxies to be detected with a signal-to-noise ratio of 5 for an exposure time of 10^5 s. This number drops to ~ 40 for an exposure time of 10^4 s. For $z = 4 \pm 0.5$ we predict more than ~ 1000 galaxies to be detected with a signal-to-noise ratio of 10 and the longer exposure time. These results are consistent with previous estimates.

- In addition we have also calculated number counts for *JWST* NIRCam high- z dropout surveys. Specifically, we have provided predictions for two extragalactic surveys that are currently planned with *JWST*. The first is a large (720 hours) observational program, the *JWST* Advanced Deep Extragalactic Survey (JADES). The second is the Cosmic Evolution Early Release Science Survey (CEERS). We have used the galaxy apparent UV luminosities to make predictions for galaxy number counts for these future *JWST* surveys. For example, at $z = 8$, we find that JADES-D can detect ~ 3 galaxies per unit redshift at the bright end up to $m_{UV} \sim 27$ mag, while having a decent detection limit $m_{UV} \sim 29.7$ mag at faint end. 235 galaxies per unit redshift can be detected in total. JADES-M has wider sky coverage and can detect ~ 5 times larger number of galaxies at the bright end. The total number of galaxies detected is 238 comparable to that of JADES-D.

Theoretically, the galaxy luminosity function has to be derived through stellar population synthesis and dust modelling based on simulation predictions. The galaxy stellar mass function, on the other hand, is the fundamental property that is directly predicted by the galaxy formation model. Observationally, galaxy stellar mass functions can be derived through a convolution of the rest-frame UV galaxy luminosity with mass-to-light ratio relations. We have therefore explored the stellar mass to UV magnitude relation, i.e. the mass-to-light ratio relation. This relation provides another important test of our galaxy formation, and dust attenuation model.

- We find that our mass-to-light ratio relations are consistent with observationally inferred ratios except for brighter galaxies, where we tend to predict more massive galaxies at a given rest-frame UV luminosity. This discrepancy towards the bright end is potentially related to the fact that observations are based on UV selected galaxies and have a potential bias missing dusty star-forming galaxies and quiescent galaxies. This can lead to a shift in the relation towards UV brighter galaxies at a given stellar mass. We have studied in detail the scatter in the M_* M_{UV} relation. This scatter increases from the faint end towards the bright end.
- The mass-to-light ratio relation is tighter towards higher redshifts due to younger stellar populations with dominant UV emission.
- Towards lower redshifts, the scatter in the mass-to-light ratio relation increases since the stellar populations become older, which weakens the correlation between stellar mass and UV luminosity causing a visibly larger scatter.
- We have employed the derived mass-to-light ratio relation to reconstruct the galaxy stellar mass function from the IllustrisTNG rest-frame UV luminosity functions. As expected, we find that this reconstruction of the galaxy stellar mass function only yields reliable results if the scatter in the mass-to-light ratio relation is taken into account to populate the massive end of the galaxy stellar mass function.

Finally, we have explored how well star formation rates of IllustrisTNG correlate with observational star formation rate indica-

tors. Spectral indices like the $H\alpha$ line luminosity are among the most important star formation rate indicators. We have therefore also extracted $H\alpha$ line luminosities for TNG50 galaxies.

- Our predictions correctly recover the observed relation between $H\alpha$ emission line luminosities and galaxy star formation rates at $z = 2$ demonstrating a tight relation between star formation rates and $H\alpha$ luminosity.
- We also find a marginally consistent $H\alpha$ luminosity function at $z = 2$ in reasonable agreement with observational data. This provides another verification of our dust modelling, and the dust calibration procedure.
- In addition we also provided a prediction for the $H\alpha$ luminosity function at $z = 4$ that can be confronted against upcoming observations.

We conclude that the combination of the three IllustrisTNG simulations, TNG50, TNG100, and TNG300, allows detailed predictions for the high redshift galaxy population up to $z \sim 10$. The derived galaxy number densities enable the derivation of the expected number of detectable galaxies with *JWST* in multiple bands assuming different survey characteristics leading to important forecasts for the first observations with *JWST*. A significant uncertainty stems from the unknown dust abundances in galaxies, which have to be inferred through a dust calibration procedure. Here novel more self-consistent dust models (e.g. [McKinnon et al. 2016, 2017, 2018](#)) would be required to overcome these limitations.

ACKNOWLEDGEMENTS

MV acknowledges support through an MIT RSC award, a Kavli Research Investment Fund, NASA ATP grant NNX17AG29G, and NSF grants AST-1814053 and AST-1814259. FM is supported by the Program “Rita Levi Montalcini” of the Italian MIUR. The primary IllustrisTNG simulations were realised with compute time granted by the Gauss Centre for Supercomputing (GCS): TNG50 under GCS Large-Scale Project GCS-DWAR, and TNG100 and TNG300 under GCS-ILLU on the GCS share of the supercomputer Hazel Hen at the High Performance Computing Center Stuttgart (HLRS). GCS is the alliance of the three national supercomputing centres HLRS (Universität Stuttgart), JSC (Forschungszentrum Jülich), and LRZ (Bayerische Akademie der Wissenschaften), funded by the German Federal Ministry of Education and Research (BMBF) and the German State Ministries for Research of Baden-Württemberg (MWK), Bayern (StMWFK) and Nordrhein-Westfalen (MIWF). Additional simulations were carried out on the Draco and Cobra supercomputers at the Max Planck Computing and Data Facility (MPCDF). Ancillary and test runs of the project were also run on the Stampede supercomputer at TACC/XSEDE (allocation AST140063), at the Hydra and Draco supercomputers at the Max Planck Computing and Data Facility, and on the MIT/Harvard computing facilities supported by FAS and MIT MKI.

REFERENCES

Alavi A., et al., 2014, *ApJ*, **780**, 143
 Aoyama S., Hou K.-C., Shimizu I., Hirashita H., Todoroki K., Choi J.-H., Nagamine K., 2017, *MNRAS*, **466**, 105
 Atek H., et al., 2015a, *ApJ*, **800**, 18
 Atek H., et al., 2015b, *ApJ*, **814**, 69
 Atek H., Richard J., Kneib J.-P., Schaerer D., 2018, *MNRAS*, **479**, 5184

Baes M., Verstappen J., De Looze I., Fritz J., Saftly W., Vidal Pérez E., Stalevski M., Valcke S., 2011, *ApJS*, **196**, 22
 Barrow K. S. S., Wise J. H., Norman M. L., O’Shea B. W., Xu H., 2017, *MNRAS*, **469**, 4863
 Behrens C., Pallottini A., Ferrara A., Gallerani S., Vallini L., 2018, *MNRAS*, **477**, 552
 Bekki K., 2015, *ApJ*, **799**, 166
 Bhatawdekar R., Conselice C., Margalef-Bentabol B., Duncan K., 2018, preprint, ([arXiv:1807.07580](#))
 Blumenthal G. R., Faber S. M., Primack J. R., Rees M. J., 1984, *Nature*, **311**, 517
 Bouchet P., Lequeux J., Maurice E., Prevot L., Prevot-Burnichon M. L., 1985, *A&A*, **149**, 330
 Bouwens R. J., et al., 2003, *ApJ*, **595**, 589
 Bouwens R. J., et al., 2010, *ApJ*, **709**, L133
 Bouwens R. J., et al., 2012, *ApJ*, **754**, 83
 Bouwens R. J., et al., 2014a, *ApJ*, **793**, 115
 Bouwens R. J., et al., 2014b, *ApJ*, **795**, 126
 Bouwens R. J., et al., 2015, *ApJ*, **803**, 34
 Bouwens R. J., et al., 2016, *ApJ*, **833**, 72
 Bouwens R. J., Oesch P. A., Illingworth G. D., Ellis R. S., Stefanon M., 2017, *ApJ*, **843**, 129
 Bowler R. A. A., et al., 2014, *MNRAS*, **440**, 2810
 Bowler R. A. A., et al., 2015, *MNRAS*, **452**, 1817
 Buat V., et al., 2011, *A&A*, **533**, A93
 Byler N., Dalcanton J. J., Conroy C., Johnson B. D., 2017, *ApJ*, **840**, 44
 Byler N., Dalcanton J., Conroy C., Johnson B., Levesque E., Berg D., 2018, preprint, ([arXiv:1803.04425](#))
 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
 Camps P., Baes M., 2015, *Astronomy and Computing*, **9**, 20
 Camps P., Baes M., Saftly W., 2013, *A&A*, **560**, A35
 Camps P., Trayford J. W., Baes M., Theuns T., Schaller M., Schaye J., 2016, *MNRAS*, **462**, 1057
 Capak P. L., et al., 2015, *Nature*, **522**, 455
 Cardelli J. A., Clayton G. C., Mathis J. S., 1989, *ApJ*, **345**, 245
 Casey C. M., et al., 2014, *ApJ*, **796**, 95
 Castellano M., et al., 2010, *A&A*, **511**, A20
 Chabrier G., 2003, *PASP*, **115**, 763
 Charlot S., Fall S. M., 2000, *ApJ*, **539**, 718
 Chen B., Dai X., Kochanek C. S., Chartas G., 2013, preprint, ([arXiv:1306.0008](#))
 Choi J., Dotter A., Conroy C., Cantiello M., Paxton B., Johnson B. D., 2016, *ApJ*, **823**, 102
 Clay S. J., Thomas P. A., Wilkins S. M., Henriques B. M. B., 2015, *MNRAS*, **451**, 2692
 Conroy C., Gunn J. E., 2010, *ApJ*, **712**, 833
 Conroy C., Gunn J. E., White M., 2009, *ApJ*, **699**, 486
 Cowley W. I., Baugh C. M., Cole S., Frenk C. S., Lacey C. G., 2018, *MNRAS*, **474**, 2352
 Cullen F., McLure R. J., Khochfar S., Dunlop J. S., Dalla Vecchia C., 2017, *MNRAS*, **470**, 3006
 Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, arXiv e-prints,
 Dayal P., Dunlop J. S., Maio U., Ciardi B., 2013, *MNRAS*, **434**, 1486
 De Cia A., Ledoux C., Savaglio S., Schady P., Vreeswijk P. M., 2013, *A&A*, **560**, A88
 De Cia A., Ledoux C., Mattsson L., Petitjean P., Srianand R., Gavignaud I., Jenkins E. B., 2016, *A&A*, **596**, A97
 De Lucia G., Blaizot J., 2007, *MNRAS*, **375**, 2
 Devriendt J. E. G., Guiderdoni B., 2000, *A&A*, **363**, 851
 Devriendt J. E. G., Guiderdoni B., Sadat R., 1999, *A&A*, **350**, 381
 Dolag K., Borgani S., Murante G., Springel V., 2009, *MNRAS*, **399**, 497
 Dotter A., 2016, *ApJS*, **222**, 8
 Draine B. T., Li A., 2007, *ApJ*, **657**, 810
 Duncan K., et al., 2014, *MNRAS*, **444**, 2960
 Dunlop J. S., et al., 2013, *MNRAS*, **432**, 3520

- Ferland G. J., et al., 2013, *Rev. Mex. Astron. Astrofis.*, **49**, 137
- Finkelstein S. L., 2016, *Publ. Astron. Soc. Australia*, **33**, e037
- Finkelstein S. L., et al., 2012, *ApJ*, **756**, 164
- Finkelstein S. L., et al., 2015, *ApJ*, **810**, 71
- Fontanot F., Somerville R. S., Silva L., Monaco P., Skibba R., 2009, *MNRAS*, **392**, 553
- Genel S., et al., 2014, *MNRAS*, **445**, 175
- Gonzalez-Perez V., Lacey C. G., Baugh C. M., Lagos C. D. P., Helly J., Campbell D. J. R., Mitchell P. D., 2014, *MNRAS*, **439**, 264
- González V., Labbé I., Bouwens R. J., Illingworth G., Franx M., Kriek M., 2011, *ApJ*, **735**, L34
- Groves B., Dopita M. A., Sutherland R. S., Kewley L. J., Fischera J., Leitherer C., Brandl B., van Breugel W., 2008, *ApJS*, **176**, 438
- Guiderdoni B., Rocca-Volmerange B., 1987, *A&A*, **186**, 1
- Guo Q., White S. D. M., 2009, *MNRAS*, **396**, 39
- Hathi N. P., et al., 2010, *ApJ*, **720**, 1708
- Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R., Guo Q., Lemson G., Springel V., Overzier R., 2015, *MNRAS*, **451**, 2663
- Hopkins P. F., Hernquist L., Martini P., Cox T. J., Robertson B., Di Matteo T., Springel V., 2005, *ApJ*, **625**, L71
- Inoue A. K., 2003, *PASJ*, **55**, 901
- Ishigaki M., Kawamata R., Ouchi M., Oguri M., Shimasaku K., Ono Y., 2018, *ApJ*, **854**, 73
- Jaacks J., Choi J.-H., Nagamine K., Thompson R., Varghese S., 2012, *MNRAS*, **420**, 1606
- Jonsson P., Groves B. A., Cox T. J., 2010, *MNRAS*, **403**, 17
- Kawamata R., Oguri M., Ishigaki M., Shimasaku K., Ouchi M., 2016, *ApJ*, **819**, 114
- Kennicutt Jr. R. C., 1998, *ARA&A*, **36**, 189
- Kimm T., Cen R., 2013, *ApJ*, **776**, 35
- Kitzbichler M. G., White S. D. M., 2007, *MNRAS*, **376**, 2
- Koekemoer A. M., et al., 2013, *ApJS*, **209**, 3
- Kong X., Charlot S., Brinchmann J., Fall S. M., 2004, *MNRAS*, **349**, 769
- Koprowski M. P., et al., 2018, *MNRAS*,
- Kriek M., Conroy C., 2013, *ApJ*, **775**, L16
- Kroupa P., 2001, *MNRAS*, **322**, 231
- Lee J. C., et al., 2012, *PASP*, **124**, 782
- Leitherer C., et al., 1999, *ApJS*, **123**, 3
- Lequeux J., Maurice E., Prevot-Burnichon M.-L., Prevot L., Rocca-Volmerange B., 1982, *A&A*, **113**, L15
- Li A., Draine B. T., 2001, *ApJ*, **554**, 778
- Liu C., Mutch S. J., Angel P. W., Duffy A. R., Geil P. M., Poole G. B., Mesinger A., Wyithe J. S. B., 2016, *MNRAS*, **462**, 235
- Livemore R. C., Finkelstein S. L., Lotz J. M., 2017, *ApJ*, **835**, 113
- Lotz J. M., et al., 2017, *ApJ*, **837**, 97
- Ma X., et al., 2018a, *MNRAS*, **477**, 219
- Ma X., et al., 2018b, *MNRAS*, **478**, 1694
- Madau P., 1995, *ApJ*, **441**, 18
- Madau P., Ferguson H. C., Dickinson M. E., Giavalisco M., Steidel C. C., Fruchter A., 1996, *MNRAS*, **283**, 1388
- Marinacci F., et al., 2018, *MNRAS*, **480**, 5113
- Mason C. A., Trenti M., Treu T., 2015, *ApJ*, **813**, 21
- Matthee J., Sobral D., Best P., Smail I., Bian F., Darvish B., Röttgering H., Fan X., 2017, *MNRAS*, **471**, 629
- McKinnon R., Torrey P., Vogelsberger M., 2016, *MNRAS*, **457**, 3775
- McKinnon R., Torrey P., Vogelsberger M., Hayward C. C., Marinacci F., 2017, *MNRAS*, **468**, 1505
- McKinnon R., Vogelsberger M., Torrey P., Marinacci F., Kannan R., 2018, *MNRAS*, **478**, 2851
- McLeod D. J., McLure R. J., Dunlop J. S., Robertson B. E., Ellis R. S., Targett T. A., 2015, *MNRAS*, **450**, 3032
- McLeod D. J., McLure R. J., Dunlop J. S., 2016, *MNRAS*, **459**, 3812
- McLure R. J., Cirasuolo M., Dunlop J. S., Foucaud S., Almaini O., 2009, *MNRAS*, **395**, 2196
- McLure R. J., et al., 2013, *MNRAS*, **432**, 2696
- McLure R. J., et al., 2018, *MNRAS*, **476**, 3991
- Mehta V., et al., 2017, *ApJ*, **838**, 29
- Meurer G. R., Heckman T. M., Calzetti D., 1999, *ApJ*, **521**, 64
- Murphy E. J., et al., 2011, *ApJ*, **737**, 67
- Naiman J. P., et al., 2018, *MNRAS*, **477**, 1206
- Nelson D., et al., 2015, *Astronomy and Computing*, **13**, 12
- Nelson D., et al., 2018, *MNRAS*, **475**, 624
- Nelson D., et al., 2019, arXiv e-prints,
- Noll S., et al., 2009, *A&A*, **499**, 69
- Nordon R., et al., 2013, *ApJ*, **762**, 125
- Oesch P. A., et al., 2010a, *ApJ*, **709**, L16
- Oesch P. A., et al., 2010b, *ApJ*, **725**, L150
- Oesch P. A., et al., 2013, *ApJ*, **773**, 75
- Oesch P. A., et al., 2014, *ApJ*, **786**, 108
- Oesch P. A., Bouwens R. J., Illingworth G. D., Labbé I., Stefanon M., 2018, *ApJ*, **855**, 105
- Ouchi M., et al., 2009, *ApJ*, **706**, 1136
- Paardekooper J.-P., Khochfar S., Dalla Vecchia C., 2013, *MNRAS*, **429**, L94
- Pakmor R., Springel V., Bauer A., Mocz P., Munoz D. J., Ohlmann S. T., Schaaf K., Zhu C., 2016, *MNRAS*, **455**, 1134
- Parisa S., Dunlop J. S., McLure R. J., Mortlock A., 2016, *MNRAS*, **456**, 3194
- Paxton B., Bildsten L., Dotter A., Herwig F., Lesaffre P., Timmes F., 2011, *ApJS*, **192**, 3
- Paxton B., et al., 2013, *ApJS*, **208**, 4
- Paxton B., et al., 2015, *ApJS*, **220**, 15
- Pillepich A., et al., 2018a, *MNRAS*, **473**, 4077
- Pillepich A., et al., 2018b, *MNRAS*, **475**, 648
- Pillepich A., et al., 2019, arXiv e-prints,
- Planck Collaboration 2016, *A&A*, **594**, A13
- Popping G., Somerville R. S., Galametz M., 2017, *MNRAS*, **471**, 3152
- Prevot M. L., Lequeux J., Maurice E., Prevot L., Rocca-Volmerange B., 1984, *A&A*, **132**, 389
- Priewe J., Williams L. L. R., Liesenborgs J., Coe D., Rodney S. A., 2017, *MNRAS*, **465**, 1030
- Reddy N. A., Steidel C. C., 2009, *ApJ*, **692**, 778
- Reddy N. A., Steidel C. C., Pettini M., Adelberger K. L., Shapley A. E., Erb D. K., Dickinson M., 2008, *ApJS*, **175**, 48
- Reddy N. A., Erb D. K., Pettini M., Steidel C. C., Shapley A. E., 2010, *ApJ*, **712**, 1070
- Reddy N., et al., 2012, *ApJ*, **744**, 154
- Reddy N. A., et al., 2015, *ApJ*, **806**, 259
- Reddy N. A., et al., 2018, *ApJ*, **853**, 56
- Rémy-Ruyer A., et al., 2014, *A&A*, **563**, A31
- Rodriguez-Gomez V., et al., 2018, preprint, (arXiv:1809.08239)
- Rogers A. B., McLure R. J., Dunlop J. S., 2013, *MNRAS*, **429**, 2456
- Rosdahl J., et al., 2018, *MNRAS*, **479**, 994
- Saftly W., Baes M., Camps P., 2014, *A&A*, **561**, A77
- Salmon B., et al., 2016, *ApJ*, **827**, 20
- Sánchez-Blázquez P., et al., 2006a, *MNRAS*, **371**, 703
- Sánchez-Blázquez P., et al., 2006b, *MNRAS*, **371**, 703
- Schechter P., 1976, *ApJ*, **203**, 297
- Schenker M. A., et al., 2013, *ApJ*, **768**, 196
- Schmidt K. B., et al., 2014, *ApJ*, **786**, 57
- Scoville N., Faisst A., Capak P., Kakazu Y., Li G., Steinhardt C., 2015, *ApJ*, **800**, 108
- Seaton M. J., 1979, *MNRAS*, **187**, 73P
- Shimizu I., Inoue A. K., Okamoto T., Yoshida N., 2014, *MNRAS*, **440**, 731
- Siana B., et al., 2009, *ApJ*, **698**, 1273
- Sijacki D., Vogelsberger M., Genel S., Springel V., Torrey P., Snyder G. F., Nelson D., Hernquist L., 2015, *MNRAS*, **452**, 575
- Silva L., Granato G. L., Bressan A., Danese L., 1998, *ApJ*, **509**, 103
- Smit R., Bouwens R. J., Franx M., Illingworth G. D., Labbé I., Oesch P. A., van Dokkum P. G., 2012, *ApJ*, **756**, 14
- Sobral D., Smail I., Best P. N., Geach J. E., Matsuda Y., Stott J. P., Cirasuolo M., Kurk J., 2013, *MNRAS*, **428**, 1128
- Sobral D., et al., 2015, *MNRAS*, **451**, 2303
- Somerville R. S., Gilmore R. C., Primack J. R., Domínguez A., 2012, *MNRAS*, **423**, 1992
- Song M., et al., 2016, *ApJ*, **825**, 5

- Springel V., 2010, *MNRAS*, **401**, 791
- Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, *MNRAS*, **328**, 726
- Springel V., et al., 2018, *MNRAS*, **475**, 676
- Stanway E. R., Bunker A. J., McMahon R. G., 2003, *MNRAS*, **342**, 439
- Stark D. P., Ellis R. S., Bunker A., Bundy K., Targett T., Benson A., Lacy M., 2009, *ApJ*, **697**, 1493
- Stark D. P., Schenker M. A., Ellis R., Robertson B., McLure R., Dunlop J., 2013, *ApJ*, **763**, 129
- Steidel C. C., Hamilton D., 1993, *AJ*, **105**, 2017
- Steidel C. C., Giavalisco M., Pettini M., Dickinson M., Adelberger K. L., 1996, *ApJ*, **462**, L17
- Tacchella S., Trenti M., Carollo C. M., 2013, *ApJ*, **768**, L37
- Tacchella S., Bose S., Conroy C., Eisenstein D. J., Johnson B. D., 2018, preprint, ([arXiv:1806.03299](https://arxiv.org/abs/1806.03299))
- Tilvi V., et al., 2013, *ApJ*, **768**, 56
- Torrey P., Vogelsberger M., Genel S., Sijacki D., Springel V., Hernquist L., 2014, *MNRAS*, **438**, 1985
- Trayford J. W., et al., 2015, *MNRAS*, **452**, 2879
- Trayford J. W., et al., 2017, *MNRAS*, **470**, 771
- Vijayan A. P., Clay S. J., Thomas P. A., Yates R. M., Wilkins S. M., Henriques B. M., 2019, arXiv e-prints,
- Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, *MNRAS*, **436**, 3031
- Vogelsberger M., et al., 2014a, *MNRAS*, **444**, 1518
- Vogelsberger M., et al., 2014b, *Nature*, **509**, 177
- Vogelsberger M., McKinnon R., O’Neil S., Marinacci F., Torrey P., Kannan R., 2018a, arXiv e-prints,
- Vogelsberger M., et al., 2018b, *MNRAS*, **474**, 2073
- Weinberger R., et al., 2017, *MNRAS*, **465**, 3291
- Whitaker K. E., Pope A., Cybulski R., Casey C. M., Popping G., Yun M. S., 2017, *ApJ*, **850**, 208
- White S. D. M., Rees M. J., 1978, *MNRAS*, **183**, 341
- Wilkins S. M., Bunker A. J., Ellis R. S., Stark D., Stanway E. R., Chiu K., Lorenzoni S., Jarvis M. J., 2010, *MNRAS*, **403**, 938
- Wilkins S. M., Bunker A. J., Stanway E., Lorenzoni S., Caruana J., 2011, *MNRAS*, **417**, 717
- Wilkins S. M., Bunker A., Coulton W., Croft R., di Matteo T., Khandai N., Feng Y., 2013, *MNRAS*, **430**, 2885
- Wilkins S. M., Feng Y., Di-Matteo T., Croft R., Stanway E. R., Bunker A., Waters D., Lovell C., 2016, *MNRAS*, **460**, 3170
- Wilkins S. M., Feng Y., Di Matteo T., Croft R., Lovell C. C., Waters D., 2017, *MNRAS*, **469**, 2517
- Wiseman P., Schady P., Bolmer J., Krühler T., Yates R. M., Greiner J., Fynbo J. P. U., 2017, *A&A*, **599**, A24
- Xu H., Wise J. H., Norman M. L., Ahn K., O’Shea B. W., 2016, *ApJ*, **833**, 84
- Yajima H., Nagamine K., Thompson R., Choi J.-H., 2014, *MNRAS*, **439**, 3073
- Yung L. Y. A., Somerville R. S., Finkelstein S. L., Popping G., Davé R., 2018, preprint, ([arXiv:1803.09761](https://arxiv.org/abs/1803.09761))
- Zafar T., Watson D., 2013, *A&A*, **560**, A26
- van der Burg R. F. J., Hildebrandt H., Erben T., 2010, *A&A*, **523**, A74

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