

## IQ Collaboratory III Empirical Dust Attenuation Models: Taking Hydrodynamical Simulations with a Grain of Dust

CHANGHOON HAHN,<sup>1,\*</sup> TJITSKE K. STARKENBURG,<sup>2</sup> ENA CHOI,<sup>3</sup> CLAIRE DICKEY,<sup>4</sup>  
KARTHEIK G. IYER,<sup>5</sup> RACHEL S. SOMERVILLE,<sup>6,7</sup> AND IQ COLLABORATORY

<sup>1</sup>*Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton NJ 08544, USA*

<sup>2</sup>*Center for Interdisciplinary Exploration and Research in Astrophysics (CIERA) and  
Department of Physics and Astronomy, 1800 Sherman Ave, Evanston IL 60201, USA*

<sup>3</sup>*Quantum Universe Center, Korea Institute for Advanced Study, Hoegiro 85, Seoul 02455, Korea*

<sup>4</sup>*Department of Astronomy, Yale University, New Haven CT 06520 USA*

<sup>5</sup>*Dunlap Institute for Astronomy and Astrophysics, University of Toronto, 50 St George St, Toronto, ON M5S 3H4,  
Canada*

<sup>6</sup>*Department of Physics and Astronomy, Rutgers University, 136 Frelinghuysen Road, Piscataway, NJ 08854, USA*

<sup>7</sup>*Center for Computational Astrophysics, Flatiron Institute, 162 5th Ave, New York, NY 10010, USA*

(Dated: DRAFT -- 921d497 -- 2021-02-15 -- NOT READY FOR DISTRIBUTION)

### ABSTRACT

We present the Empirical Dust Attenuation (EDA) framework for applying dust attenuation to galaxies from theoretical models. EDA provides a flexible prescription for assigning realistic dust attenuation to simulated galaxies related to their physical properties ( $M_*$ , SSFR) and using the Noll et al. (2009) attenuation curve parameterization.

Using the EDA, we forward model synthetic observations for a  $M_r < -20$  complete galaxy sample for three state-of-the-art large-scale cosmological hydrodynamical simulations: SIMBA, IllustrisTNG, and EAGLE. We then compare the optical and UV color-magnitude relations,  $(g-r) - M_r$  and  $(FUV-NUV) - M_r$ , for the forward modeled simulations to SDSS observations using likelihood-free inference. As expected, for all three of the simulations dust attenuation is necessary to match the observations. We find that we can accurately reproduce the color-magnitude relations of the observational sample for TNG and EAGLE. For SIMBA, we struggle to reproduce observations due to an overprediction of low-mass galaxies with high star-formation rates. Examining the attenuation curves predicted by EDA for TNG and EAGLE, we find good agreement with observed attenuation-slope relations and attenuation curves of star-forming galaxies. Additionally, EDA provides unprecedented predictions for the attenuation curves of quiescent galaxies, which are challenging to measure observationally. To match observations, simulated quiescent galaxies require significant UV and optical attenuation but with much shallower slopes than star-forming galaxies. Overall, we find that more massive galaxies in the simulations require higher dust attenuation, while galaxies with higher specific star formation rates have steeper attenuation curves. In addition to im-

proving our understanding of dust in galaxies, our results underscore the advantages of EDA and the forward modeling approach.

*Keywords:* keyword1 – keyword2 – keyword3

## 1. INTRODUCTION

Dust in the interstellar medium of a galaxy can dramatically impact its spectral energy distribution (SED). The combined effect of dust on a galaxy’s SED is typically described using an attenuation curve,  $A(\lambda)$ , which has now been broadly characterized by observations. In the UV, attenuation curves steeply rise due to absorption by small grains. At 2175Å, in the near-UV (NUV), there is an absorption bump referred to as the “UV dust bump”. At longer optical wavelengths, the curves take on a power-law shape. Finally, dust reemits the attenuated light in the optical and UV in the infrared (for an overview see Calzetti 2001; Draine 2003; Galliano et al. 2018). By impacting the SED, dust also affects the physical properties that are inferred from the SED, such as star formation rate (SFR), stellar mass ( $M_*$ ), or stellar ages (see reviews by Walcher et al. 2011; Conroy 2013). Assumptions on dust attenuation can dramatically vary these properties (Kriek & Conroy 2013; Reddy et al. 2015; Salim et al. 2016; Salim & Narayanan 2020). Since these properties are the building blocks to our understanding of galaxies and how they evolve, a better understanding of dust not only provides insights into dust, but also underpins all galaxy studies.

To better understand dust in galaxies, many observational works have examined trends between dust attenuation and galaxy properties. For example, UV and optical attenuation are found to correlate with  $M_*$ , SFR, and metallicity in star-forming galaxies (Garn & Best 2010; Battisti et al. 2016). The slope of the attenuation curves in star-forming galaxies also correlate with galaxy properties, such as  $M_*$ , specific SFR (SSFR), metallicity, and axial ratio (Wild et al. 2011; Battisti et al. 2017). Recently, Salim et al. (2018) argue that these correlations stem from the underlying “attenuation-slope relation”, a trend between the amplitude of attenuation and slope. Despite the progress, there is still no clear consensus on the connection between attenuation curves and galaxy properties. Also, studies so far have focused solely on star-forming galaxies and little is known about dust attenuation in quiescent galaxies. Furthermore, galaxy properties and dust attenuation measured from galaxy SEDs are subject to variations, inconsistencies, and biases of different methodologies, which can be significant even for the same observations (*e.g.* Speagle et al. 2014; Katsianis et al. 2020). SED fitting can also impose undesirable priors on derived galaxy properties (Carnall et al. 2018; Leja et al. 2019) and suffer from parameter degeneracies that are poorly understood.

Significant progress has also been made in theoretically modeling dust. Simulations can now model the radiative transfer of stellar light through a dusty ISM for a wide range of configurations: from simple slab-like dust geometries (*e.g.* Witt & Gordon 1996, 2000; Seon & Draine 2016) to 3D hydrodynamic simulations of entire galaxies (*e.g.* Jonsson 2006; Rocha et al. 2008; Hayward & Smith 2015; Natale et al. 2015; Hou et al. 2017). Radiative transfer models have even been applied to cosmological hydrodynamical simulations (*e.g.* Camps & Baes 2015; Narayanan et al. 2018; Trayford et al. 2020).

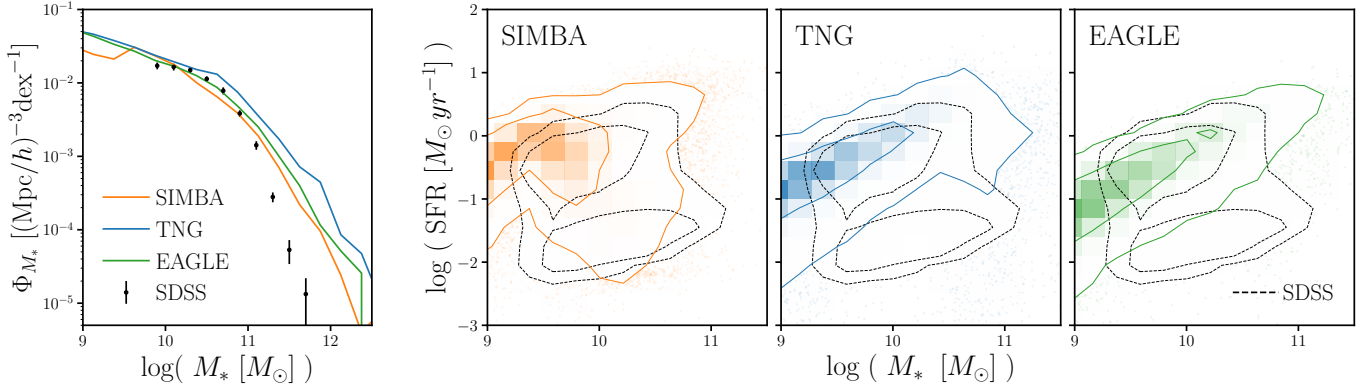
\* hahn.changhoon@gmail.com

Dust has also been examined in a cosmological context using semi-analytic models (SAMs; *e.g.* Granato et al. 2000; Fontanot et al. 2009; Wilkins et al. 2012; Gonzalez-Perez et al. 2013; Popping et al. 2017). Yet there are still major limitations in modeling dust. Dust models in cosmological simulations currently do not reproduce the redshift evolution of dust properties (Somerville et al. 2012; Yung et al. 2019; Vogelsberger et al. 2020). Also, radiative transfer models produce attenuation-slope relations that are significantly steeper than observations. Many models also require significant hand-tuning (*e.g.* propagating rays/photons into particular cells) and make assumptions on the underlying dust grain models (see Steinacker et al. 2013, for a review).

We take a different approach from the observational and theoretical works above — *we investigate dust attenuation using a forward modeling approach to compare simulations to observations*. Our “forward model” starts with three major large-scale hydrodynamical simulations: EAGLE, IllustrisTNG, and SIMBA. We use their outputs (*e.g.* star formation history) to build SEDs for each simulated galaxy. We then apply dust attenuation to all the SEDs using an Empirical Dust Attenuation framework, which we describe shortly. Finally, we apply a realistic noise model and survey selection function on the attenuated SEDs and construct synthetic photometric observations. Afterwards, we compare the synthetic observations from the forward model to actual observations. The comparisons are made in observational space, so they are not impacted by the inconsistencies of observational methods in measuring galaxy properties. Furthermore, since the forward models can directly include the selection functions and observational systematic effects, forward modeling makes it easier to account for these effects and to exploit the full observational data set.

An essential step in our forward model is applying the Empirical Dust Attenuation (EDA) framework, which provides a flexible and computationally inexpensive prescription for applying dust attenuation. The EDA first assigns attenuation curves to every simulated galaxy. In this work, we use the Noll et al. (2009) parameterization for the attenuation curves and determine the amplitude and slope of the curves from the simulated galaxy’s  $M_*$  and specific star formation rate (SSFR) as well as a randomly sampled inclination ( $i$ ). We use the same parameterization as observational studies so that the EDA attenuation curves can easily be compared to observational constraints. Also, the  $M_*$  and SSFR dependence is motivated by observations (*e.g.* Garn & Best 2010; Wild et al. 2011; Battisti et al. 2016; Leja et al. 2017; Salim et al. 2018; Salim & Narayanan 2020) and allows us to easily explore the correlation between galaxy properties and dust attenuation. After the assignment, we simply apply the attenuation curves to the SEDs of the simulated galaxies. The EDA, unlike radiative transfer models, does not produce realistic dust attenuation for individual galaxies. However, as we later demonstrate, it produces realistic distributions of dust attenuation for galaxy populations. Unlike radiative transfer models, however, the EDA does not produce realistic dust attenuation for individual galaxies. The EDA provides an empirical mapping framework for dust attenuation, analogous to the halo occupation or abundance matching frameworks in galaxy formation (for a review see Wechsler & Tinker 2018).

In principle, a radiative transfer model can be used instead of the EDA in the forward modeling approach; however, radiative transfer models are computationally expensive. Applying a range of radiative transfer dust models to multiple simulations for comparisons would require huge com-



**Figure 1.** The stellar mass functions,  $\Phi_{M_*}$  (left-most panel), and  $M_*$  – SFR relation (right panels) of galaxies in three cosmological hydrodynamic simulations: SIMBA (orange), TNG (blue), and EAGLE (green). For reference, we include  $\Phi_{M_*}$  (black) and the  $M_*$  – SFR relation (black dashed) of from SDSS observations. Uncertainties for the SDSS  $\Phi_{M_*}$  are derived using jackknife resampling. We describe the simulations and observations in Section 2. *There are significant differences between the  $M_*$  – SFR relations of SDSS and the simulations, which are driven by the inconsistencies between the simulations’ theoretical  $M_*$  and SFR predictions and the SDSS measurements. A forward modeling approach, where we construct synthetic observations for the simulations, reduces the inconsistencies and enables an apples-to-apples comparison between simulations and observation. Furthermore, differences in  $\Phi_{M_*}$  and the  $M_*$  – SFR relations among the hydrodynamic simulations highlight how they predict galaxy populations with significantly different physical properties.*

putational resources. Using them with Monte Carlo methods for parameter exploration would be prohibitive. On the other hand, with the EDA, we can apply a wide range of realistic dust attenuation to simulated galaxies in a matter of seconds. We can easily explore and sample the dust parameter space and infer the relationship between dust attenuation and galaxy properties. That is the focus of this paper. Beyond investigating dust, the EDA also provides a framework where we can treat dust as *nuisance* parameters and tractably marginalize over dust attenuation. In the subsequent paper of the IQ series, Starkenburg et al. (in preparation), we will use the EDA framework to compare star formation quenching in cosmological galaxy formation models after marginalizing over dust attenuation.

In Section 2, we describe the three large-scale cosmological hydrodynamical simulations (SIMBA, IllustrisTNG, and EAGLE) that we use in our forward model. We also describe the observed SDSS galaxy sample used for the comparison. Next, we present the specific EDA prescription used in this work in Section 3 and the likelihood-free inference method for comparing the simulations to observations in Section 4. Finally, in Section 5, we present the results of our comparison and discuss their implications on dust attenuation and its connection to galaxy properties.

## 2. DATA

In this paper, we present the Empirical Dust Attenuation (EDA) model and demonstrate how it can be used in a forward modeling approach to compare galaxy populations in simulations and observations. For our simulations, we use three large-scale cosmological hydrodynamical simulations: the IllustrisTNG (hereafter TNG), EAGLE, and SIMBA. For our observations, we use a galaxy

sample derived from SDSS. Below, we briefly describe the hydrodynamical simulations and the SDSS observations used throughout this work.

In Figure 1, we present the stellar mass functions,  $\Phi_{M_*}$  (left-most panel), and  $M_* - \text{SFR}$  relations (right panels) of galaxies in the SIMBA (orange), TNG (blue), and EAGLE (green) cosmological hydrodynamic simulations. For reference, we include  $\Phi_{M_*}$  and the  $M_* - \text{SFR}$  relation for SDSS observations. For the simulations,  $M_*$  is the total stellar mass within the subhalo and SFR is the instantaneous SFR in the dense and cold star-forming gas. *We do not impose any selection cuts on the simulations in Figure 1 and include both central and satellite galaxies.* For SDSS,  $M_*$  is estimated using `kcorrect` (Blanton & Roweis 2007) assuming a Chabrier (2003) initial mass function and SFR is from the current release of Brinchmann et al. (2004)<sup>1</sup>. The uncertainties for the SDSS SMF are derived from jackknife resampling. *We find striking differences between the  $M_* - \text{SFR}$  relation of SDSS and the simulations, which underscore the need for a forward modeling approach. While the  $M_*$  and SFR of simulations represent theoretical predictions, the values for SDSS are subject to, for example, SFR measurement limits, inconsistencies among SFR tracers, and aperture effects. By constructing synthetic observations for the simulations with forward models, the simulations can be more consistently compared to observations (see e.g. Dickey et al. 2020, ; Starkenburg et al. in prep.)* Figure 1 also illustrates that the hydrodynamical simulations predict different SMFs and  $M_* - \text{SFR}$  relations from each other. *This difference, which was also recently highlighted in Hahn et al. (2019c), demonstrates that the hydrodynamical simulations predict galaxy populations with significantly different physical properties from one another.*

## 2.1. IllustrisTNG100

The IllustrisTNG100 simulation<sup>2</sup> is a cosmological hydrodynamic simulation of comoving volume  $(110.7 \text{ Mpc})^3$ , with a particle mass resolution of  $7.6 \times 10^6 M_\odot$  for dark matter and  $1.4 \times 10^6 M_\odot$  for baryonic particles (Nelson et al. 2018; Pillepich et al. 2018; Springel et al. 2018). It improves on the original Illustris simulation<sup>3</sup> (Vogelsberger et al. 2014; Genel et al. 2014; public data release by Nelson et al. 2015), by including magneto-hydrodynamics and updated treatments for galactic winds, metal enrichment, and AGN feedback. Most notably, TNG uses a new implementation for feedback from SMBH, where feedback energy is injected in the form of a kinetic AGN-driven wind at low SMBH accretion rates (Weinberger et al. 2018). This new implementation has been shown to alleviate discrepancies found between the original Illustris and observations for  $> 10^{13-14} M_\odot$  massive halos.

## 2.2. EAGLE

The Virgo Consortium’s EAGLE project<sup>4</sup> (Schaye et al. 2015; Crain et al. 2015; McAlpine et al. 2016) is a publicly available suite of cosmological hydrodynamic simulations constructed using AN-ARCHY (Dalla Vecchia et al. in prep.; see also Appendix A of Schaye et al. 2015), a modified version of the GADGET-3 code (Springel 2005). We use the L0100Ref simulation, which has a comoving

<sup>1</sup> <http://www.mpa-garching.mpg.de/SDSS/DR7/>

<sup>2</sup> <https://www.tng-project.org/>

<sup>3</sup> <http://www.illustris-project.org>

<sup>4</sup> <http://www.eaglesim.org>

volume of  $(100 \text{ Mpc})^3$ , and a baryonic mass resolution of  $1.81 \times 10^6 M_\odot$ . EAGLE has subgrid models for star formation, stellar mass loss, metal enrichment and stellar feedback that stochastically inject thermal energy in the ISM as in (Dalla Vecchia & Schaye 2012). The feedback energy from AGN is also added to surrounding gas stochastically (Booth & Schaye 2009). Parameters of the stellar feedback and SMBH accretion are calibrated to broadly reproduce the  $z = 0$  stellar mass function and galaxy stellar size-stellar mass relation. Meanwhile, the AGN feedback efficiency is calibrated to match the SMBH-galaxy mass relation.

### 2.3. SIMBA

The SIMBA simulation suite (Davé et al. 2019), the successor to MUFASA (Davé et al. 2016, 2017a,b), is a cosmological hydrodynamical simulation constructed using GIZMO, a meshless finite mass hydrodynamics code (Hopkins 2015; Hopkins et al. 2017). Of the simulations, we use ‘m100n1024’, which has a box size of  $(100 h^{-1} \text{ Mpc})^3$  and baryonic mass resolution of  $1.82 \times 10^7 M_\odot$ . The simulation uses the same subgrid models as MUFASA for  $\text{H}_2$  based star formation, decoupled two-phase winds for star formation driven galactic winds, and feedback from Type I supernovae and AGB stars. SIMBA uses a two-mode SMBH accretion model, torque-limited accretion for cold gas (Anglés-Alcázar et al. 2017) and Bondi-based accretion for hot gas, and two-mode AGN feedback.

### 2.4. SDSS Galaxies

For our observations, we use a galaxy sample derived from SDSS observations. We impose a  $M_r < -20$  completeness limit on a volume-limited sample from Tinker et al. (2011). The Tinker et al. (2011) sample is derived from the SDSS DR7 (Abazajian et al. 2009) NYU Value-Added Galaxy Catalog (VAGC; Blanton et al. 2005) and has a  $M_* > 10^{9.4} h^{-2} M_\odot$  completeness limit. In total, our SDSS sample has 16,970 galaxies. In this work, we focus on observables that can be consistently defined and derived in both simulations and observations: the  $r$ -band absolute magnitude,  $M_r$ , the optical  $g-r$  color, and the  $FUV-NUV$  color. For the SDSS sample, we use  $FUV$ ,  $NUV$ ,  $r$  and  $g$  band absolute magnitudes from the NASA-Sloan Atlas<sup>5</sup>, which is re-reduction of SDSS DR8 (Aihara et al. 2011) that includes an improved background subtraction (Blanton et al. 2011) and near and far UV photometry from GALEX. These absolute magnitudes are derived using `kcorrect` (Blanton & Roweis 2007), assuming a Chabrier (2003) initial mass function.

### 2.5. Forward Modeling Observations

One of the main goals of this work is to conduct an “apples-to-apples” comparison between the simulations and observations. A crucial step in this comparison is to *forward model* the observables for the simulations. The simulations can then be directly compared to observations in observational-space. This means that the comparison does not rely on measured galaxy properties, which are impacted by variations, inconsistencies, and biases of different methods. The comparison can also include selection functions and systematic effects of the observations through the forward model. In this work, we use  $r$ -band luminosity ( $M_r$ ), optical color ( $g-r$ ), and UV color ( $FUV-NUV$ ) as our observables.

<sup>5</sup> <http://nsatlas.org/>



As the first step in our forward model, we construct SEDs for all of the simulated galaxies based on their star formation and metallicity histories (SFH and ZH) using the Flexible Stellar Population Synthesis model (FSPS; Conroy et al. 2009, 2010). For each simulated galaxy, we bin the total stellar mass formed by age ( $t$ ) and metallicity ( $Z$ ). We use a consistent  $t$ ,  $Z$  grid for all of the simulations to account for the variable time and mass resolutions. For each point in the  $t$ ,  $Z$  grid, we generate a spectrum assuming a simple stellar population (SSP) using FSPS and take the mass-weighted linear combination of them to produce the galaxy SED. We use a Chabrier (2003) initial mass function. For further details on how we construct the SEDs, we refer readers to Starkenberg et al. (in prep.).

Next, we apply dust attenuation to the SEDs. We use the EDA prescription to assign dust attenuation curves for each simulated galaxy based on its physical properties, a randomly sampled inclination, and EDA model parameters. We describe the EDA framework in detail later in Section 3. We measure the observables by convolving the attenuated SEDs with transmission curves of the GALEX  $FUV$ , GALEX  $NUV$ , SDSS  $g$ , and SDSS  $r$  broadband filters. We add realistic noise to  $M_r$ ,  $g-r$ , and  $FUV-NUV$  by sampling from the observed uncertainty distributions of the NASA Sloan-Atlas. Lastly, we apply the  $M_r < -20$  absolute magnitude completeness limit of our SDSS sample to the simulated galaxies.

In Figure 2, we present the forward modeled optical and UV color-magnitude relations,  $(g-r) - M_r$  (top) and  $(FUV-NUV) - M_r$  (bottom), for simulated galaxies in SIMBA (left), TNG (center) and EAGLE (right) *assuming no dust attenuation*. We mark the 68 and 95% contours and include, for reference, the optical and UV color-magnitude relations of our SDSS sample (black dashed). Comparison to SDSS observations clearly demonstrate that *without dust attenuation, the hydrodynamical simulations cannot reproduce the observed optical or UV color-magnitude relations*.

### 3. THE EMPIRICAL DUST ATTENUATION FRAMEWORK

In this section, we describe the Empirical Dust Attenuation (EDA) framework and present the EDA prescription we use in this work. The EDA is a flexible framework for applying dust attenuation curves to simulated galaxy populations. For each simulated galaxy, the EDA assigns a dust attenuation curve that is parameterized as a function of the galaxy’s properties ( $M_*$ , SSFR), the EDA parameters, and randomly sampled inclination. With the EDA, we can apply a wide variety of dust attenuation that include correlation between dust attenuation and physical galaxy properties.

We begin by defining the dust attenuation curve,  $A(\lambda)$ , as

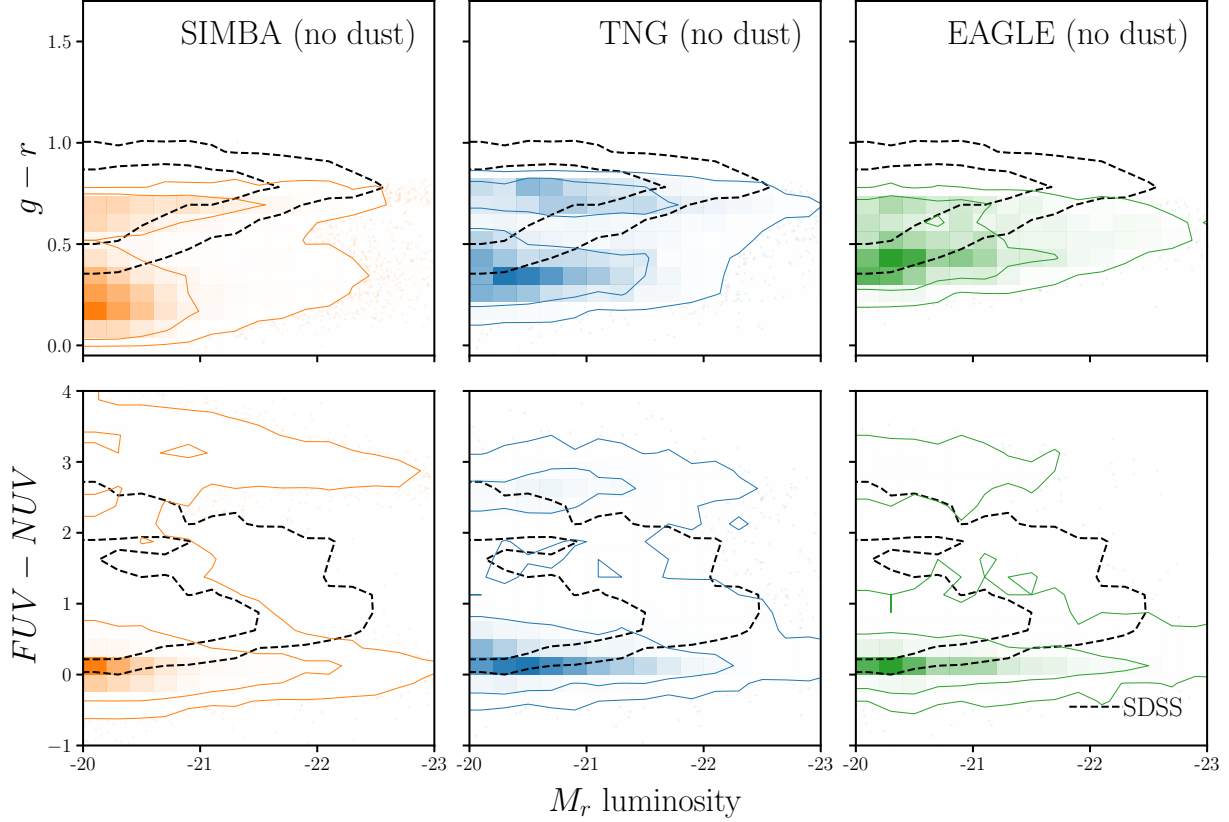
$$F_o(\lambda) = F_i(\lambda)10^{-0.4A(\lambda)} \quad (1)$$

where  $F_o$  is the observed flux and  $F_i$  is the intrinsic flux. We normalize the attenuation to the  $V$  band attenuation,

$$A(\lambda) = A_V \frac{k(\lambda)}{k_V} \quad (2)$$

so that  $A_V$  determines the amplitude of the attenuation, while  $k(\lambda)$  determines the wavelength dependence.

The EDA framework assigns  $A(\lambda)$  to every galaxy in the simulations using some flexible prescription. For the EDA prescription in this work, we assign  $A_V$  for each galaxy using the slab model,



**Figure 2.** We present the forward modeled optical and UV color-magnitude relations of SIMBA (left), TNG (center), and EAGLE (right) galaxies *assuming no dust attenuation*. We present  $(g-r) - M_r$  in the top panels and  $(FUV-NUV) - M_r$  in the bottom panels. The contours represent the 68 and 95% of the distribution. We derive observables  $M_r$ ,  $g-r$ , and  $FUV-NUV$  for the simulations using our forward model (Section 2.5). For comparison, we include the color-magnitude relations of our SDSS sample (black dashed; Section 2.4). *Without dust attenuation, the hydrodynamical simulations do not reproduce the SDSS optical or UV color-magnitude relations.*

where  $A_V$  is a function of galaxy inclination,  $i$ , and its optical depth,  $\tau_V$  (*e.g.* Somerville & Primack 1999; Somerville et al. 2012):

$$A_V = -2.5 \log \left[ \frac{1 - e^{-\tau_V \sec i}}{\tau_V \sec i} \right]. \quad (3)$$

Since observations have established that  $A_V$  correlates with galaxy properties (*e.g.* Garn & Best 2010; Battisti et al. 2016; Salim & Narayanan 2020), we include correlations between  $A_V$  and galaxy  $M_*$  and SSFR through  $\tau_V$  (see Appendix A for details). We parameterize  $\tau_V$  using a linear  $M_*$  and SSFR dependence:

$$\tau_V(M_*, \text{SFR}) = m_{\tau, M_*} \log \left( \frac{M_*}{10^{10} M_\odot} \right) + m_{\tau, \text{SSFR}} \log \left( \frac{\text{SSFR}}{10^{-10} \text{yr}^{-1}} \right) + c_\tau. \quad (4)$$

$m_{\tau, M_*}$ ,  $m_{\tau, \text{SSFR}}$ , and  $c_\tau$  represent the  $M_*$  dependence, the SSFR dependence, and amplitude of  $\tau_V$ . Since  $\tau_V$  is optical depth, we impose a  $\tau_V \geq 0$  limit. *For each galaxy, we uniformly sample  $\cos i$  from*



0 to 1. By sampling  $\cos i$ , our EDA prescription includes significant variance in  $A(\lambda)$  so galaxies with the same galaxy properties do not have the same dust attenuation.

We use the slab model primarily as a flexible prescription for  $A_V$  that depend on a randomly sampled  $i$ , with *loose* physical motivations. For star-forming galaxies, which typically have disc-like morphologies, the slab model produces  $A_V$  that is correlated with  $i$  in a way consistent with observations: edge-on galaxies have higher  $A_V$  than face-on galaxies (*e.g.* Conroy et al. 2010; Wild et al. 2011; Battisti et al. 2017; Salim & Narayanan 2020). Nevertheless, the slab model is a naive approximation; in reality,  $A_V$  depends on the detailed star-to-dust geometry. Furthermore, all galaxies in the simulations are assigned  $A_V$  from the slab model. For quiescent galaxies, which typically have elliptical morphologies, the slab model serves only as an *empirical* prescription for statistically sampling  $A_V$ . However, the purpose EDA is to assign an accurate distribution of dust attenuation curves for the galaxy population — *not* to accurately model dust attenuation for individual galaxies. In Appendix A, we demonstrate that the slab model based EDA can produce  $A_V$  values that match  $p(A_V)$  of observations, which include quiescent galaxies. Hence, the slab model is a sufficiently flexible prescription for sampling  $A_V$  for all galaxies.

For the wavelength dependence of the attenuation curve,  $k(\lambda)$ , we use the Noll et al. (2009) parameterization:

$$k(\lambda) = (k_{\text{Cal}}(\lambda) + D(\lambda)) \left( \frac{\lambda}{\lambda_V} \right)^\delta. \quad (5)$$

Here  $k_{\text{Cal}}(\lambda)$  is the Calzetti (2001) curve:

$$k_{\text{Cal}}(\lambda) = \begin{cases} 2.659(-1.857 + 1.040/\lambda) + R_V, & 6300A \leq \lambda \leq 22000A \\ 2.659(-2.156 + 1.509/\lambda - 0.198/\lambda^2 + 0.011/\lambda^3) + R_V & 1200A \leq \lambda \leq 6300A \end{cases}$$

where  $\lambda_V = 5500A$  is the  $V$  band wavelength and  $\delta$  is the slope offset of the attenuation curve from  $k_{\text{Cal}}$ . Since  $\delta$  correlates with galaxy properties (*e.g.* Wild et al. 2011; Battisti et al. 2016; Leja et al. 2017; Salim et al. 2018, ; see also Appendix A), we parameterize  $\delta$  with a similar  $M_*$  and SSFR dependence as  $\tau_V$ :

$$\delta(M_*, \text{SFR}) = m_{\delta, M_*} \log \left( \frac{M_*}{10^{10} M_\odot} \right) + m_{\delta, \text{SFR}} \log \left( \frac{\text{SSFR}}{10^{-10} \text{yr}^{-1}} \right) + c_\delta. \quad (6)$$

$D(\lambda)$  in Eq. 5 is the UV dust bump, which we parameter using the standard Lorentzian-like Drude profile:

$$D(\lambda) = \frac{E_b(\lambda - \Delta\lambda)^2}{(\lambda^2 - \lambda_0^2)^2 + (\lambda - \Delta\lambda)^2} \quad (7)$$

where  $\lambda_0$ ,  $\Delta\lambda$ , and  $E_b$  are the central wavelength, full width at half maximum, and strength of the bump, respectively. We include the UV dust bump since we use UV color as one of our observables.

We assume fixed  $\lambda_0 = 2175A$  and  $\Delta\lambda = 350A$ . Kriek & Conroy (2013) and Tress et al. (2018) find that  $E_b$  correlates with the  $\delta$  for star-forming galaxies  $z \sim 2$ . Narayanan et al. (2018) confirmed this dependence in simulations. However, we assume a fixed relation between  $E_b$  and  $\delta$  from Kriek & Conroy (2013):  $E_b = -1.9 \delta + 0.85$ . Allowing the slope and amplitude of the  $E_b$  and  $\delta$  relation to

**Table 1.** Free parameters of the Empirical Dust Attenuation Model

Parameter	Definition	prior
$m_{\tau, M_*}$	$M_*$ dependence of the optical depth, $\tau_V$	flat $[-5., 5.]$
$m_{\tau, \text{SSFR}}$	SSFR dependence of $\tau_V$	flat $[-5., 5.]$
$c_\tau$	amplitude of $\tau_V$	flat $[0., 6.]$
$m_{\delta, M_*}$	$M_*$ dependence of $\delta$ , the attenuation curve slope offset	flat $[-4., 4.]$
$m_{\delta, \text{SFR}}$	SSFR dependence of $\delta$	flat $[-4., 4.]$
$c_\delta$	amplitude of $\delta$	flat $[-4., 4.]$

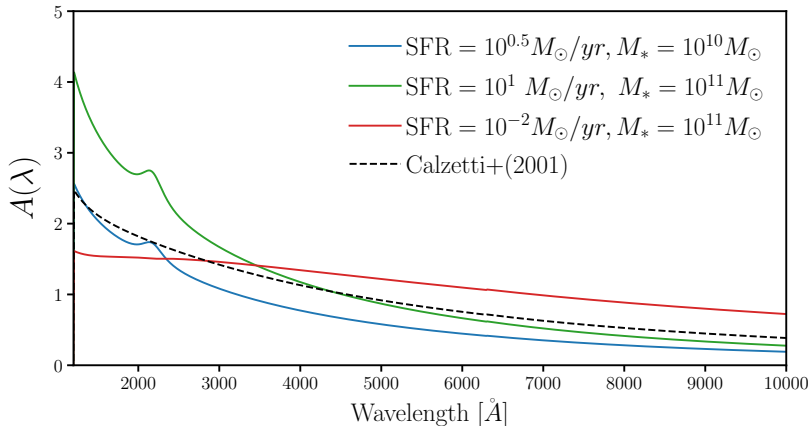
vary does *not* impact our results; however, we also do not derive any meaningful constraints on them. In Table 1, we list and describe all of the free parameters of the EDA.

SSFR of galaxies are used to calculate  $\tau_V$  and  $\delta$  in Eqs. 4 and 6. However, due to mass and temporal resolutions, some galaxies in the simulations have  $\text{SFR} = 0$  — *i.e.* an unmeasurably low SFR (Hahn et al. 2019c). They account for 17, 19, 9% of galaxies in SIMBA, TNG, and EAGLE, respectively. Since Eqs. 4 and 6 depend on  $\log \text{SSFR}$ , they cannot be used in the equations to derive  $\tau_V$  and  $\delta$  for these galaxies. To account for this issue, we assign  $\text{SFR}_{\min}$ , the minimum non-zero SFR in the simulations, to  $\text{SFR} = 0$  galaxies when calculating  $\tau_V$  and  $\delta$ . For SIMBA, TNG, and EAGLE,  $\text{SFR}_{\min} = 0.000816, 0.000268, \text{ and } 0.000707 M_\odot/\text{yr}$ . Although this assumes that  $\text{SFR} = 0$  galaxies have similar dust properties as the galaxies with  $\text{SFR} = \text{SFR}_{\min}$ , since the simulations have very low  $\text{SFR}_{\min}$  we expect galaxies with  $\text{SFR} = \text{SFR}_{\min}$  to have little recent star-formation and low gas mass, similar to  $\text{SFR} = 0$  galaxies.

In practice, to apply the EDA to a simulated galaxy population, we first assign a randomly sampled inclination,  $i$ , to each galaxy ( $\cos i$  uniformly sampled from 0 to 1).  $\tau_V$  and  $\delta$  are calculated for the galaxy based on its  $M_*$ , SSFR and the EDA parameters. We then combine  $A_V$  from  $i$  and  $\tau_V$  with  $k(\lambda)$  from  $\delta$  to determine  $A(\lambda)$  for each galaxy. Afterwards, we attenuate the galaxy SEDs using Eq. 1 and use the attenuated SEDs to calculate the observables:  $g, r, \text{NUV}$ , and  $FUV$  absolute magnitudes. In Figure 3, we present attenuation curves,  $A(\lambda)$ , generated by the EDA for galaxies with different SFR and  $M_*$  values. We include star-forming galaxies with  $\{M_*, \text{SFR}\} = \{10^{10} M_\odot, 10^{0.5} M_\odot/\text{yr}\}$  (blue),  $\{10^{11} M_\odot, 10^1 M_\odot/\text{yr}\}$  (green) and a quiescent galaxy with  $\{10^{11} M_\odot, 10^{-2} M_\odot/\text{yr}\}$  (red). We use an arbitrarily set of EDA parameters ( $m_{\tau, M_*}, m_{\tau, \text{SSFR}}, c_\tau, m_{\delta, M_*}, m_{\delta, \text{SFR}}, c_\delta$ ) within the prior range listed in Table 1. We set  $i = 0$  (edge-on) for all  $A(\lambda)$  in Figure 3 for simplicity. In practice the EDA uniformly samples  $\cos i$  from 0 to 1 for each galaxy. For comparison, we include the Calzetti (2001) attenuation curve. Even when we set  $i = 0$ , the EDA produces attenuation curves with a wide range of amplitude and slope to galaxies based on their physical properties.

#### 4. LIKELIHOOD-FREE INFERENCE: APPROXIMATE BAYESIAN COMPUTATION

With our forward model, which includes the EDA prescription for dust attenuation, we can now generate synthetic observations for simulated galaxies and make an “apples-to-apples” comparison to SDSS. Next, we want to use this comparison to infer the posterior probability distribution of



**Figure 3.** Attenuation curves,  $A(\lambda)$ , assigned by our Empirical Dust Attenuation (EDA) prescription to edge-on galaxies with different SFR and  $M_*$  values for an arbitrary set of EDA parameters. We include  $A(\lambda)$  for star-forming galaxies with  $\{M_*, \text{SFR}\} = \{10^{10} M_\odot, 10^{0.5} M_\odot/\text{yr}\}$  (blue),  $\{10^{11} M_\odot, 10^1 M_\odot/\text{yr}\}$  (green) and a quiescent galaxy with  $\{10^{11} M_\odot, 10^{-2} M_\odot/\text{yr}\}$  (red). We set  $i = 0$  for all the galaxies in the figure for simplicity but in practice the EDA uniformly samples  $\cos i$  from 0 to 1 for each galaxy. For comparison, we include the Calzetti (2001) attenuation curve. *The EDA provides a flexible prescription for assigning dust attenuation to galaxies based on their inclination, physical properties ( $M_*$  and SSFR), and the EDA parameters.*

the EDA parameters. Typically in astronomy, this inference is done assuming a Gaussian likelihood to compare the “summary statistic” (*e.g.* SMF) of the model to observations and some sampling method (*e.g.* Markov Chain Monte Carlo) to estimate the posterior distribution. The functional form of the likelihood, however, depends on the summary statistic and assuming an incorrect form of the likelihood can significantly bias the inferred posteriors (*e.g.* Hahn et al. 2019b). In this work, we use the optical and UV color-magnitude relations as our summary statistic. Since the statistic is three-dimensional histogram, its likelihood is *not* Gaussian but rather a Poisson distribution.

Rather than *incorrectly* assuming a Gaussian likelihood or attempting to estimate the true Poisson likelihood of the optical and UV color-magnitude relations, we use Approximate Bayesian Computation (hereafter ABC; Diggle & Gratton 1984; Tavaré et al. 1997; Pritchard et al. 1999; Beaumont et al. 2009; Del Moral et al. 2012) for our inference. ABC is a likelihood-free (or “simulation-based”) parameter inference framework that approximates the posterior probability distribution,  $p(\theta | \text{data})$ , without requiring evaluations of the likelihood. Instead, ABC only requires a forward model of the observed data, a prior that can be sampled, and a distance metric that quantifies the “closeness” to the observed data. Since ABC does not require evaluating the likelihood, it does not assume any functional form of the likelihood and so we avoid any biases from such assumptions. Furthermore, it also allows us to infer the posterior using summary statistics with likelihoods that are difficult or intractable to directly estimate (Hahn et al. 2017a).

In the simplest version of ABC, with a rejection sample framework (Pritchard et al. 1999), a proposal set of parameter values are drawn from the prior. The forward model is run with the proposal parameter values. The output of the forward model is then compared to the observed data

using a distance metric that quantifies the “closeness” of the forward model output to the observed data. If the distance is within some small threshold, we keep the proposed parameters; otherwise, we discard them. Proposals are drawn until enough pass the threshold to sample the posterior. A rejection sampling framework requires a large number of evaluations of the forward model, which can be computationally costly. Many variations of ABC with more efficient sampling strategies have now been applied to astronomy and cosmology (*e.g.* Cameron & Pettitt 2012; Weyant et al. 2013; Ishida et al. 2015; Lin et al. 2016; Alsing et al. 2018). Among these methods, we use ABC with Population Monte Carlo (PMC) importance sampling (Hahn et al. 2017a,b, 2019a).

ABC-PMC begins with an arbitrarily large threshold  $\epsilon_1$  and  $N$  proposals  $\bar{\theta}_1$  sampled from the prior distribution. Each proposal is assigned a weight  $w_1^i = 1/N$ . Then for subsequent iterations ( $n > 1$ ), the threshold,  $\epsilon_n$ , is set to the median distance of the previous iteration’s proposals. New proposals are drawn from the previous iteration’s proposals perturbed by a kernel and kept if their distance is below  $\epsilon_n$ . This is repeated until we assemble a new set of  $N$  proposals  $\bar{\theta}_n$ . The entire process is repeated for the next iteration until convergence is confirmed. We use the Python implementation of Akeret et al. (2015)<sup>6</sup>. For further details on the ABC-PMC implementation, we refer readers to Hahn et al. (2017b) and Hahn et al. (2019a).

In this work, we use ABC-PMC with uninformative uniform priors on each of the EDA parameters and choose ranges that encompass constraints in the literature. The prior ranges of  $m_{\tau,M_*}$ ,  $m_{\tau,\text{SSFR}}$ ,  $c_{\tau}$  are chosen to conservatively include the  $A_V$  range and  $M_*$  and SFR dependence of Narayanan et al. (2018) and Salim & Narayanan (2020). Meanwhile, the prior ranges of  $m_{\delta,M_*}$ ,  $m_{\delta,\text{SFR}}$ ,  $c_{\delta}$  are chosen to conservatively include the  $\delta$  range and  $M_*$  and SFR dependence of Leja et al. (2017) and Salim et al. (2018). We list the range of the priors in Table 1. For our forward model, we use the model described in Section 2.5: we construct SEDs for every simulated galaxies from the hydrodynamic simulations, apply dust attenuation with our EDA, calculate the observables ( $M_r$ ,  $g-r$ , and  $FUV-NUV$ ), add uncertainties to them, and apply a  $M_r < -20$  completeness limit. We use the optical and UV color-magnitude relation,  $(g-r) - M_r$  and  $(FUV-NUV) - M_r$  as our summary statistic to fully exploit the  $(M_r, g-r, FUV-NUV)$  observational-space. We measure the color-magnitude relations by calculating the number density in bins of  $(g-r, FUV-NUV, M_r)$  with widths  $(0.0625, 0.25, 0.5)$  mags. For our distance metric,  $\rho$ , we use the L2 norm between the summary statistics of the SDSS observation,  $X^{\text{SDSS}}$  and of our forward model,  $X^{\text{FM}}(\theta_{\text{EDA}})$ :

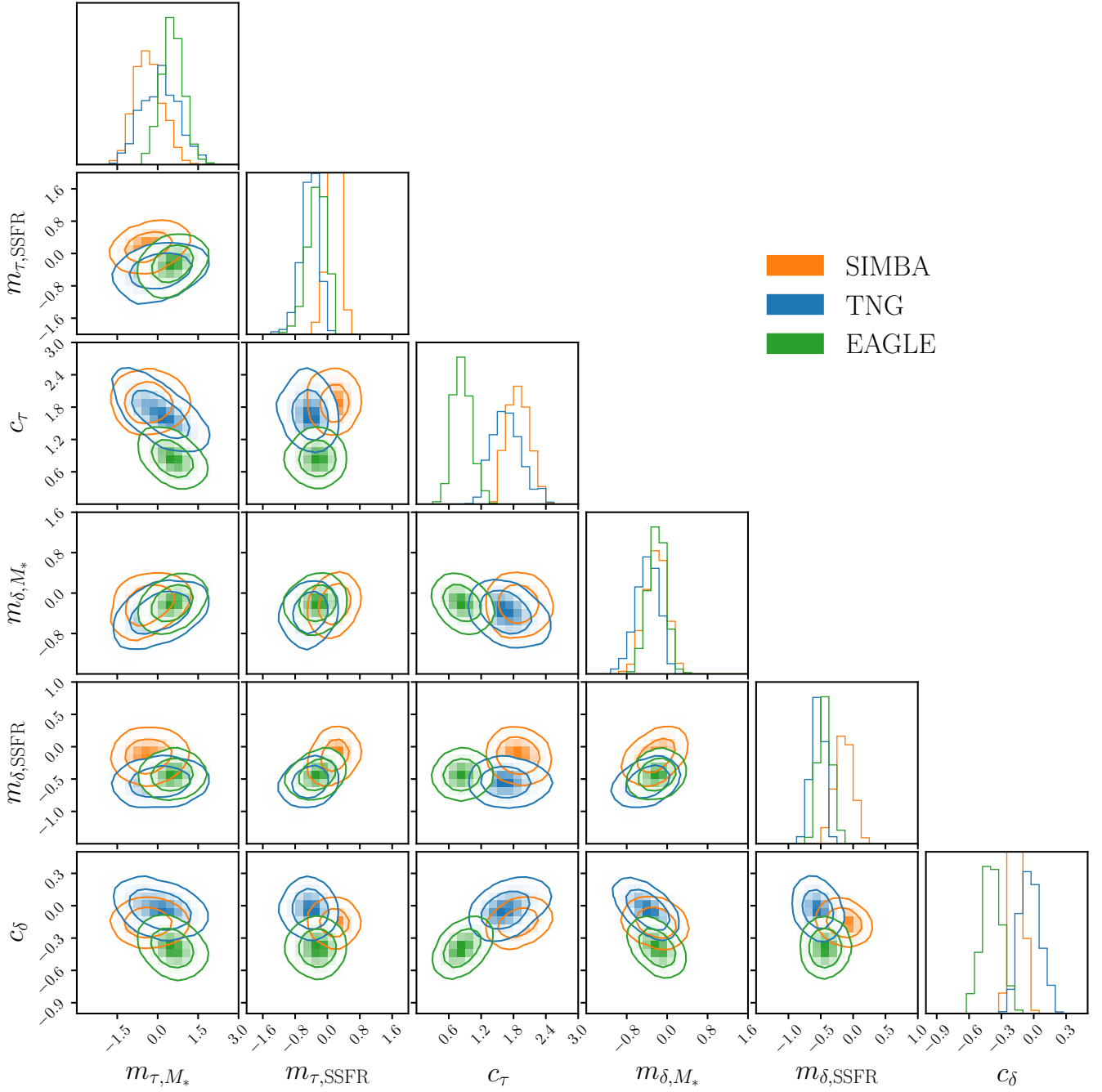
$$\rho(\theta_{\text{EDA}}) = [X^{\text{SDSS}} - X^{\text{FM}}(\theta_{\text{EDA}})]^2. \quad (8)$$

In Figure 4, we present the posterior distributions of the EDA parameters derived using ABC-PMC for the SIMBA (orange), TNG (blue), and EAGLE (green) hydrodynamical simulations. The contours mark the 68 and 95 percentiles of the distributions.

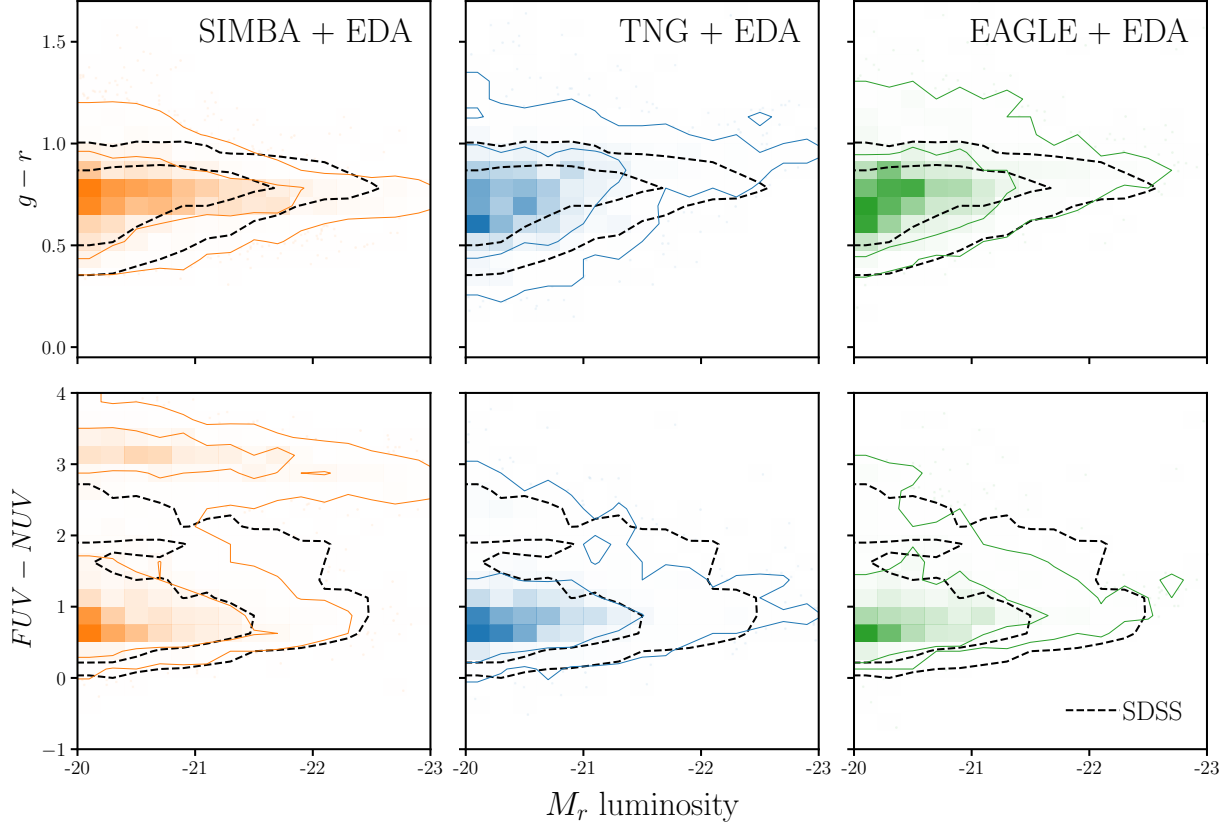
## 5. RESULTS

Without dust attenuation, all of the hydrodynamical simulations struggle to reproduce the  $(g-r) - M_r$  and  $(FUV-NUV) - M_r$  relations of SDSS (Figure 2). In the optical, the simulations predict

<sup>6</sup> <https://abcpmc.readthedocs.io/en/latest/index.html>



**Figure 4.** Posterior distributions of the EDA parameters for the SIMBA (orange), TNG (blue), and EAGLE (green) hydrodynamical simulations. The contours mark the 68 and 95 percentiles of the distributions. The posteriors are derived using the likelihood-free inference method: Approximate Bayesian Computation with Population Monte Carlo (Section 4). We focus on the EDA posteriors for TNG and EAGLE since the EDA struggles to reproduce SDSS observations with SIMBA, which predicts an overabundance of low mass galaxies with high SFR.



**Figure 5.** The optical,  $(g-r) - M_r$  (top), and UV,  $(FUV-NUV) - M_r$  (bottom), color-magnitude relations predicted by our EDA prescription for the SIMBA (orange), TNG (blue), and EAGLE (green) hydrodynamical simulations. For the EDA parameters of each simulation, we use the median of the posterior distributions inferred using ABC. For comparison, we include the color-magnitude relations of SDSS (black dashed). Comparing the color-magnitude relations above to those without dust attenuation in Figure 2, we see that dust *dramatically* impacts the color-magnitude relations. Dust attenuation must be accounted for when interpreting and comparing simulations. Furthermore, with our EDA prescription, all three simulations reproduce the color-magnitude relations of SDSS observations. *Since the different simulations can reproduce observations just by varying dust, dust significantly limits our ability to constrain the underlying physical processes of galaxy formation models.*

significantly bluer colors and find broader differences in color,  $\sim 0.5 \text{ mag}$ , between star-forming and quiescent galaxies. In the UV, they predict quiescent galaxies with  $\text{SSFR} < 10^{-12} \text{ yr}^{-1}$  that have  $FUV-NUV$  colors much redder than SDSS observations. Meanwhile, for the rest of the galaxies, the simulations predict significantly bluer UV colors than SDSS.

*With our EDA prescription, however, all three simulations reproduce the color-magnitude relations of SDSS observations.* In Figure 5, we present the optical and UV color-magnitude relations predicted by the EDA for the SIMBA (orange), TNG (blue), and EAGLE (green) simulations. For the EDA parameters, we use the median of the inferred posterior distributions (Figure 4). We include the color-magnitude relations of SDSS observations (black-dashed) comparison. The contours mark the 68 and 95 percentiles of the distributions.

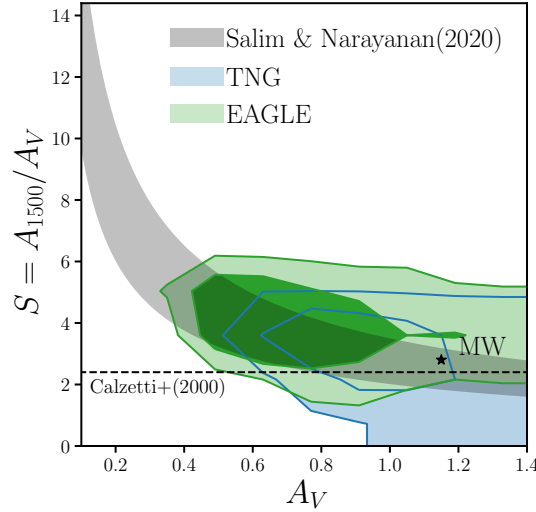


Dust dramatically impacts the observables of simulations. The EDA affects the the optical and UV color-magnitude relations in three major ways to produce good agreement with SDSS. First, the EDA significantly reddens simulated galaxies in both the optical and UV. Overall,  $g-r$  colors are  $\sim 0.25$  mag redder and  $FUV-NUV$  colors are  $\sim 0.5$  mag redder. Second, the EDA significantly attenuates ( $\sim 0.5$  mag) intrinsically blue galaxies (*i.e.* star-forming galaxies with  $\log \text{SSFR} > -10.5$ ). As a result, there are no luminous optically blue galaxies ( $M_r < -21$  and  $g-r < 0.5$ ) — consistent with observations. Lastly, in the UV, the EDA attenuates low SSFR quiescent galaxies that are intrinsically red in the UV ( $FUV-NUV \sim 3$  mag) by  $\sim 1$  mag. Thus, unlike in Figure 2, the UV color-magnitude relations from the EDA do not have a significant number of galaxies with high  $FUV-NUV$ , in agreement with SDSS.

For SIMBA, although the EDA prescription predicts optical and UV color-magnitude relations consistent with observations, there are still some discrepancies in the color-magnitude relations with SDSS. Furthermore, the overall agreement is significantly worse than TNG and EAGLE: the median ABC distance metric (Eq. 8) of the posterior  $\rho = 3.7 \times 10^{-3}, 3.2 \times 10^{-3}$  for TNG and EAGLE while  $\rho = 6.6 \times 10^{-3}$  for SIMBA. Most notably, SIMBA+EDA produces a population of galaxies with  $FUV-NUV > 2.5$  mag that lie well above the SDSS observations. The inferred EDA parameters for SIMBA also differ significantly from the parameters of TNG and EAGLE (Figure 4). For instance, we infer the opposite  $M_*$  and SSFR dependence for  $\tau_V$  than TNG or EAGLE. These discrepancies are primarily driven by the excess of high SSFR  $> 10-9.5 \text{ yr}^{-1}$  “starburst” galaxies with  $M_* < 10^{10} M_\odot$ , which lie significantly above the SFS and are *only* present in SIMBA (Figure 1). This starburst population, which has also been identified in Davé et al. (2019) (see their Figures 5 and 6) is caused by excess gas in low-mass galaxies at  $z = 0$ . If we exclude SSFR  $> 10^{-9.5} - 9.5 \text{ yr}^{-1}$  starburst galaxies and run the EDA for SIMBA using median values of the TNG and EAGLE posteriors for the EDA parameters, we find similar level of agreement with SDSS as TNG and EAGLE.

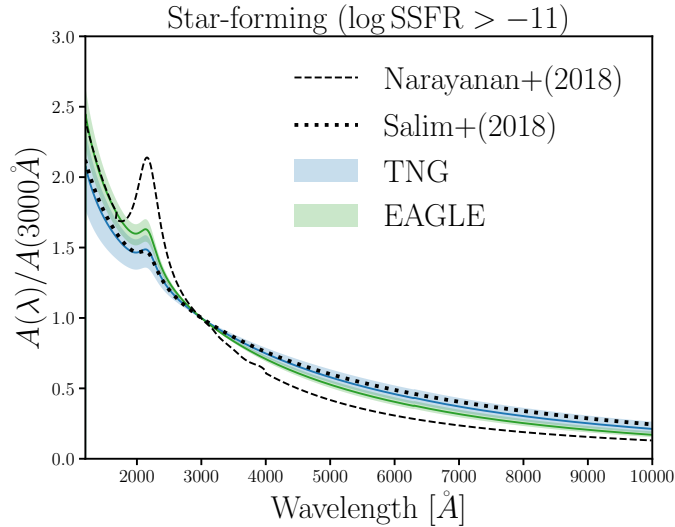
Without dust attenuation, the excess starburst galaxies in SIMBA are blue  $g-r \sim 0.1$  and have high luminosity,  $M_r < -22$ . Since such high luminosity blue galaxies do not exist in observations, they need to be both strongly attenuated and reddened. In our EDA prescription, dust attenuation and reddening is a linear function of  $\log \text{SSFR}$  so it cannot strongly attenuate and redden just the starburst population. Consequently, the quiescent galaxies are not sufficiently attenuated. Although we can increase the flexibility of our EDA prescription, the attenuation and reddening necessary for the starburst galaxies conflict with the attenuation-slope relation well-established in observations and simulations (Inoue 2005; Chevallard et al. 2013; Salim et al. 2018; Salim & Narayanan 2020; Trayford et al. 2020). The SIMBA starbursts require both high attenuation and steeper slopes but, according to the attenuation-slope relations, galaxies with higher attenuation have shallower slopes. Since the excess starburst population is a known discrepancy that dramatically impacts the inferred EDA parameters, we do not explore extending our EDA prescription and focus on the TNG and EAGLE simulations for the rest of the paper.

Previous works in the literature have also compared simulations with different dust prescriptions to observations in color-magnitude space. For EAGLE, Trayford et al. (2015) calculate colors and luminosities with the GALAXEV population synthesis models and a two-component screen model



**Figure 6.** The attenuation-slope relation of the attenuation curves predicted by our EDA prescription for the median posterior parameter values of TNG (blue) and EAGLE (green). For comparison, we include the observed attenuation-slope relation from GSWLC2 (Salim & Narayanan 2020). We also include the Milky Way (star) and mark the slope of the Calzetti (2001) curve (dashed). We use  $A_V$  and  $S = A(1500\text{\AA})/A_V$  as measurements of attenuation and slope, respectively. We derived the posteriors of the EDA parameters from comparing the UV and optical color-magnitude relation and do not fit any observed dust attenuation measurements. Yet, *we find excellent agreement between the attenuation-slope relation predicted by the EDA and observations.*

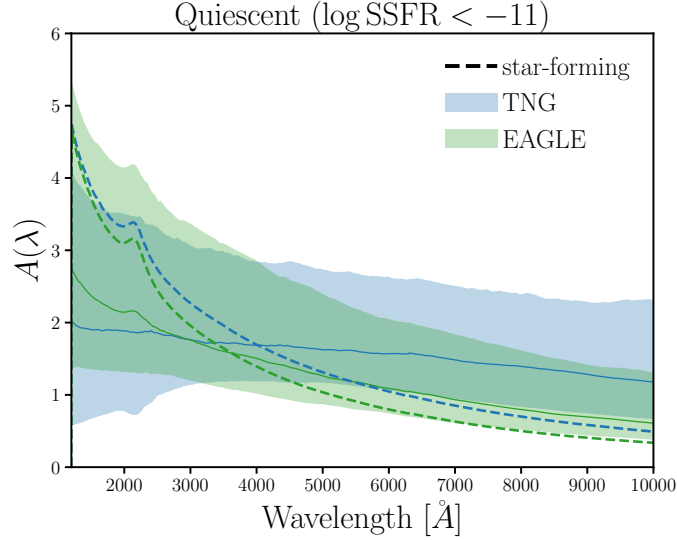
for dust. More recently, Trayford et al. (2017) calculated optical colors for EAGLE using SKIRT, a Monte Carlo radiative transfer code (Camps & Baes 2015), to model the dust. At stellar masses and luminosities comparable to our SDSS sample, both Trayford et al. (2015) and Trayford et al. (2017) produce red sequences bluer than in GAMA observations. Also, Trayford et al. (2015) predict an excess of luminous blue galaxies. Although a detailed comparison is difficult since both works compare to different observations, we note that with the EDA, EAGLE is able to successfully reproduce the position of the SDSS red sequence and does not predict a significant excess of luminous blue galaxies. Also using EAGLE and SKIRT, Baes et al. (2019) find that they overestimate the observed cosmic SED (CSED) in the UV regime and produce significantly higher  $FUV-NUV$  color than GAMA. The EDA for TNG and EAGLE predict  $FUV-NUV$  in good agreement with SDSS. For TNG, Nelson et al. (2018) calculate optical colors using a dust model that includes attenuation due to dense gas birth clouds surrounding young stellar populations and also due to simulated distribution of neutral gas and metals. They find bluer red sequence peaks and a narrower blue cloud compared to SDSS. We find neither of these discrepancies for the TNG+EDA. The EDA provides a simpler empirical framework for applying dust attenuation than the dust models in these works. Yet, with its flexibility, we are able to produce optical and UV color-magnitude relations that are in good agreement with observations. Furthermore, with its low computation costs we were able to fully explore our dust parameters.



**Figure 7.** The normalized attenuation curves of star-forming galaxies predicted by the EDA for median posterior parameter values of TNG (blue) and EAGLE (green). Galaxies are classified as star-forming using a  $\log \text{SSFR} > -11 \text{ yr}^{-1}$  cut. The attenuation curves are normalized at  $3000\text{\AA}$  and we mark the  $1\sigma$  standard deviation of the attenuation curves with the shaded region. For comparison, we include  $A(\lambda)/A(3000\text{\AA})$  measurements from the (Narayanan et al. 2018) radiative transfer simulation (dashed) and Salim et al. (2018) observations (dotted). *The EDA predict attenuation curves of star-forming galaxies that are in good agreement with the attenuation curves measured from the simulation and observations.*

### 5.1. Reproducing Dust Observations

With our EDA prescription, we are able to accurately reproduce the observed optical and UV color-magnitude relations with our simulations. In addition to reproducing observations, since the EDA assigns dust attenuation curves for each simulated galaxies, we can also compare the EDA dust attenuation curves to dust attenuation measured from observations. We begin with the well-established attenuation-slope relation: galaxies with higher dust attenuations have shallower attenuation curves. This relation is a consequence of dust scattering dominating absorption at low attenuation while dust absorption dominating at high attenuation (Gordon et al. 1994; Witt & Gordon 2000; Draine 2003; Chevallard et al. 2013). In Figure 6, we present the attenuation-slope relation of the dust attenuation curves predicted by the EDA for the median posteriors of TNG (blue) and EAGLE (green). For comparison, we include the observed attenuation-slope relations of GSWLC2 galaxies (grey shaded; Salim & Narayanan 2020), the Milky Way (star), and mark the slope of the Calzetti (2001) curve (dashed). For attenuation we use  $A_V$  and for slope we use the UV-optical slope,  $S = A(1500\text{\AA})/A_V$ , commonly found in the literature. The contours mark the 68 and 95 percentiles. We note that the different in the  $A_V$  ranges is due to the  $M_r$  completeness limit imposed by our forward model (Section 2.5). The GSWLC2 sample in Salim & Narayanan (2020) extends down to  $M_* \sim 10^{8.5} M_\odot$ ; however, the TNG and EAGLE samples do not extend below  $M_* \sim 10^{10} M_\odot$ . *The EDA predicts attenuation-slope relations for TNG and EAGLE that are in excellent agreement with observations.*



**Figure 8.** The attenuation curves of quiescent galaxies predicted by the EDA for median posterior parameter values of TNG (blue) and EAGLE (green). Galaxies are classified as quiescent using a  $\text{SSFR} < 10^{-11} \text{yr}^{-1}$  cut. We mark the  $1\sigma$  standard deviation of the attenuation curves with the shaded region and include the predicted attenuation curves of star-forming galaxies for comparison (dashed). For both TNG and EAGLE, the EDA predicts significant UV and optical dust attenuation in quiescent galaxies with an overall shallow attenuation curve.

In addition to the attenuation-slope relation, we can also directly compare the attenuation curves predicted by the EDA to measurements from observations for star-forming galaxies. In Figure 7, we present the normalized attenuation curves of star-forming galaxies predicted by the EDA for the median posterior parameter values of TNG (blue) and EAGLE (green). We define galaxies with  $\text{SSFR} > 10^{-11} \text{yr}^{-1}$  as star-forming. The attenuation curves are normalized at  $3000\text{\AA}$  and we present the variation in the attenuation curves in the shaded region,  $1\sigma$  standard deviation about the median. For comparison, we include  $A(\lambda)/A(3000\text{\AA})$  from the Narayanan et al. (2018) radiative transfer simulation (dashed) and observations (Salim et al. 2018, ; dotted). The attenuation curve from Salim et al. (2018) correspond to star-forming galaxies with  $M_* > 10^{10.5} M_\odot$ , a similar  $M_*$  range as our forward modeled TNG and EAGLE samples. Since we do not vary the UV bump in our EDA prescription, we ignore any discrepancies in the amplitudes of the bump. Overall, we find good agreement between the EDA attenuation curves for star-forming galaxies and the attenuation curves from observations and simulations.

In this section, we show that EDA predicts dust attenuation curves that are in good agreement with the well-established attenuation-slope relation and the attenuation curves of star-forming galaxies from observations and simulations. The EDA predictions are derived solely from matching the SDSS UV and optical color-magnitude relations — without fitting any observed dust attenuation measurements. The overall agreement demonstrates the robustness of the EDA framework and the specific parameterization of our EDA prescription.

## 5.2. The Attenuation Curves of Quiescent Galaxies

Beyond reproducing previous work, the EDA also makes predictions on the dust attenuation of quiescent galaxies, which is still poorly constrained by observations. There are many challenges to measuring the attenuation curves of quiescent galaxies directly from observations. Methods that rely on IR luminosities can be contaminated by MIR emission from AGN heating nearby dust (Kirkpatrick et al. 2015). SED fitting methods must also account for AGN MIR emission (Salim et al. 2016; Leja et al. 2018; Salim et al. 2018) and struggle to tightly constrain dust attenuation for quiescent galaxies due to the degeneracies with star formation history and metallicity.

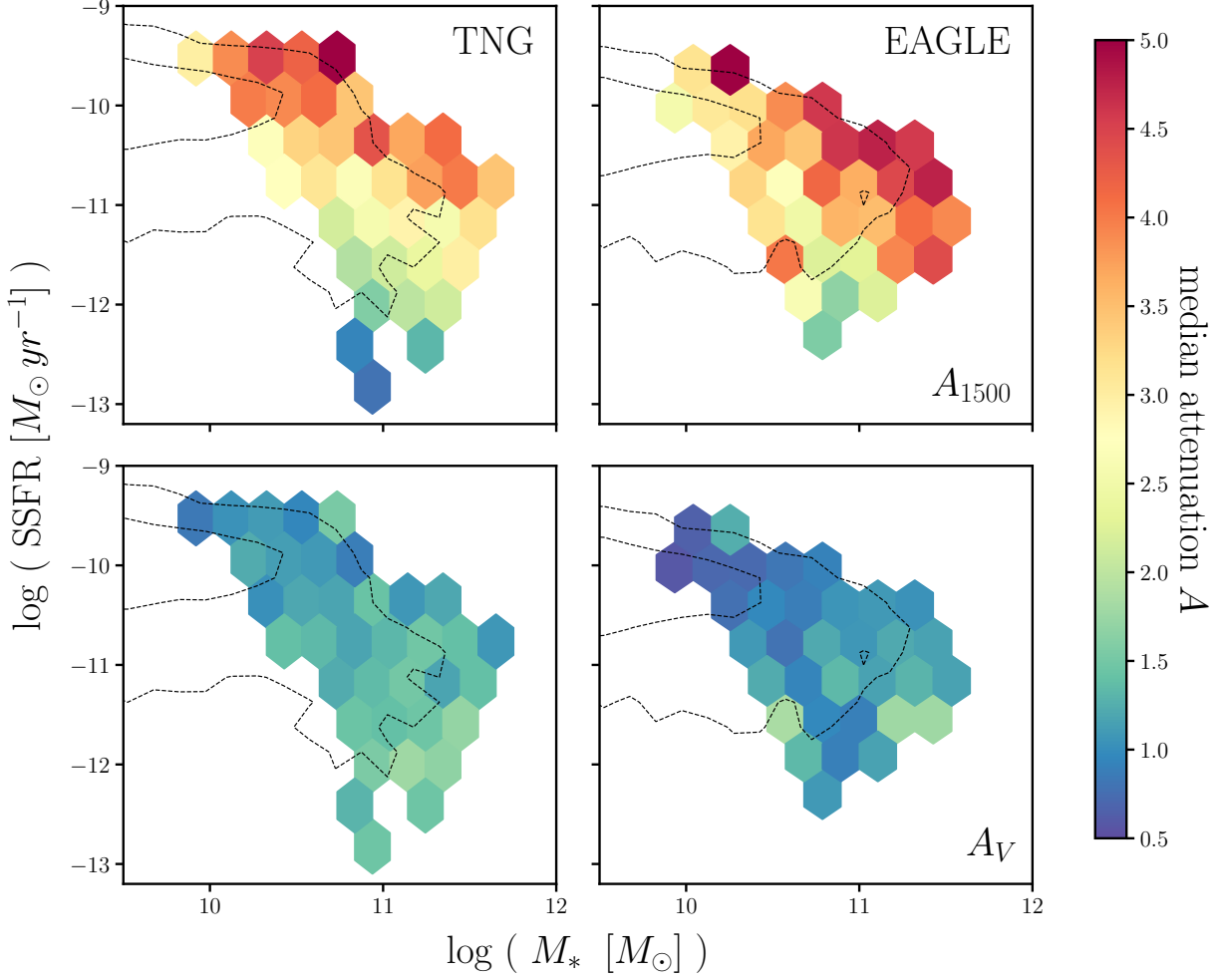
With a forward modeling approach, we circumvent these challenges. Instead, we derive the attenuation curves necessary for the simulated quiescent population to reproduce the observed optical and UV photometry. In Figure 8, we present the attenuation curves of quiescent galaxies predicted by the EDA for the median posterior parameter values of TNG (blue) and EAGLE (green). Quiescent galaxies are classified using a  $\text{SSFR} < 10^{-11} \text{yr}^{-1}$  cut. Unlike Figure 7, the attenuation curves are *not* normalized at  $3000\text{\AA}$ . For comparison, we include  $A(\lambda)$  of star-forming galaxies predicted by the EDA (dotted). Both TNG+EDA and EAGLE+EDA predict significant dust attenuation in quiescent galaxies:  $A(\lambda) \gtrsim 1$  throughout the UV and optical. Their attenuation curves are also significantly shallower than the attenuation curves of star-forming galaxies.

The quiescent galaxy dust attenuation we predict is not a consequence of the parameterization of our EDA prescription. Without dust both TNG and EAGLE predict quiescent galaxies that are too luminous and bluer in the optical than SDSS observations (Figure 2). Therefore, to match the SDSS color-magnitude relations, significant attenuation is necessary to lower their luminosities but with a shallow slope that reddens their  $g-r$  colors while maintaining their  $FUV-NUV$  colors. This is also the reason why TNG has a shallower slope than EAGLE: without dust, TNG has a quiescent population that is redder in both UV and optical.

Dust attenuation in quiescent galaxies, which is typically neglected, has significant implications. For instance, it strengthens the evidence for the UV upturn phenomenon, the unexpected detections of UV flux in quiescent galaxies (*e.g.* Code 1969; O’Connell 1999; Le Cras et al. 2016; Ali et al. 2018; Dantas et al. 2021). Better constraints on the attenuation in quiescent galaxies may help discern among the different hypotheses: residual star formation activity (*e.g.* Kaviraj et al. 2007), post-main-sequence stellar evolutionary phases (*e.g.* Yi et al. 1997), or binary systems (*e.g.* Han et al. 2007). Since the attenuation curves of quiescent galaxies are difficult to measure from observations, the EDA predictions highlight the advantages of a forward modeling approach and its complementarity with standard approaches. We, however, note that since we only vary the EDA parameters in our forward model, we do not explore whether the discrepancy between the quiescent populations in simulations and observations is due to shortcomings of the galaxy formation models. In principle, and in future work, we can vary both the EDA parameters and the parameters of the galaxy formation models and infer them simultaneously with a forward modeling approach.

### 5.3. The Galaxy – Dust Connection

With the EDA framework, we can also shed light on the connection between the physical properties of the simulated galaxies and dust attenuations. In our EDA prescription, we included a flexible  $M_*$  and SSFR dependence in both the amplitude and slope of the attenuation curve (Eqs 4 and 6). Hence,



**Figure 9.**  $M_*$  and SSFR dependence of dust attenuation at 1500Å ( $A_{1500}$ ; top) and at 5500Å ( $A_V$  bottom) predicted by the EDA for TNG (left) and EAGLE (right). The colormap in each hexbin represents the median attenuation for all simulated galaxies in the bin (right color bar). We only include bins with more than 10 galaxies that satisfy our  $M_r < -20$  completeness limit. Overall, TNG and EAGLE galaxies with higher  $M_*$  have higher dust attenuation — consistent with the literature. Furthermore, since previous works have primarily focused on star-forming galaxies, the EDA provides new insight into the SSFR dependence of dust attenuation: simulated galaxies with higher SSFR have steeper attenuation curves.

we can reveal the  $M_*$  and SSFR dependence of dust attenuation through both the EDA parameter constraints (Figure 4) and the predicted attenuation curves.

Focusing first on the amplitude of dust attenuation, we find that TNG has little  $M_*$  dependence in  $\tau_V$ :  $m_{\tau, M_*} = 0.14^{+0.64}_{-0.58}$ . EAGLE has a more significant positive  $M_*$  dependence:  $m_{\tau, M_*} = 0.53^{+0.36}_{-0.36}$ . Though neither TNG nor EAGLE has a strong dependence, V-band dust attenuation is higher for more massive galaxies. Meanwhile, we find significant SSFR dependences in both TNG ( $m_{\tau, SSFR} = -0.42^{+0.2}_{-0.18}$ ) and EAGLE ( $m_{\tau, SSFR} = -0.24^{+0.22}_{-0.19}$ ): galaxies with higher SSFR have lower V-band dust attenuation. For the slope of the dust attenuation, we find some  $M_*$  dependence in both TNG ( $m_{\delta, M_*} = -0.36^{+0.23}_{-0.19}$ ) and EAGLE ( $m_{\delta, M_*} = -0.2^{+0.16}_{-0.16}$ ). More massive galaxies have slightly steeper



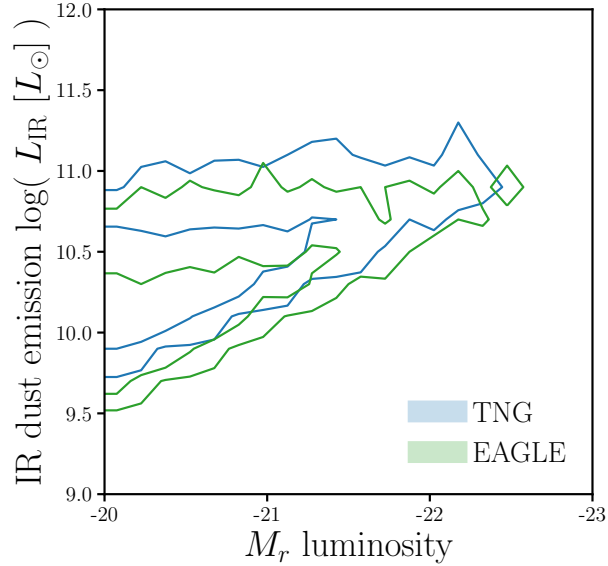
attenuation curves. We also find significant SSFR dependence in both TNG ( $m_{\delta, \text{SSFR}} = -0.55^{+0.08}_{-0.08}$ ) and EAGLE ( $m_{\delta, \text{SSFR}} = -0.43^{+0.08}_{-0.08}$ ). Galaxies with higher SSFR have steeper attenuation curves.

We take a closer look at the  $M_*$  and SSFR dependence of the attenuation curve in Figure 9. We present dust attenuation at  $1500\text{\AA}$  ( $A_{1500}$ ; top) and  $5500\text{\AA}$  ( $A_V$ ; bottom) as a function of  $\log M_*$  and  $\log \text{SSFR}$  predicted by the EDA for TNG (left) and EAGLE (right). For each hexbin, the colormap represents the median attenuation for all simulated galaxies in the bin. We only include bins with more than 10 galaxies that satisfy our  $M_r < -20$  completeness limit. In each of the panels, we find that TNG and EAGLE galaxies with higher  $M_*$  have higher dust attenuation — consistent with the literature. Burgarella et al. (2005), for instance, found significant positive  $M_*$  dependence in  $FUV$  attenuation in NUV-selected and FIR-selected samples. Garn & Best (2010) and Battisti et al. (2016) also find higher attenuation in more massive SDSS star-forming galaxies. Most recently, Salim et al. (2018) find higher  $V$  and  $FUV$  attenuation for more massive star-forming galaxies in GSWLC2.

For the SSFR dependence, we find that galaxies with higher SSFR have higher  $A_{1500}$  (top) but lower  $A_V$  (bottom). The opposite SSFR dependence in  $A_{1500}$  versus  $A_V$  is a result of the significant SSFR dependence on the slope of the attenuation curve. Galaxies with higher SSFR have significantly steeper slopes so even though they have lower  $A_V$ , they have higher  $A_{1500}$ . At the lowest SSFR end, quiescent galaxies in TNG and EAGLE have nearly flat attenuation curves. Since observations have only focused on star-forming galaxies due to the difficulty in measuring dust attenuation in quiescent galaxies, the EDA predictions provide new insight into the SSFR dependence of dust attenuation. In summary, from the EDA predictions, we find that *TNG and EAGLE galaxies with higher  $M_*$  require overall higher dust attenuation and galaxies with higher SSFR require steeper attenuation curves.*

#### 5.4. Discussion

We make a number of assumptions and choices in our EDA prescription. First, we use the slab model (Eq. 3) to assign  $A_V$  as a function of the randomly sampled  $i$  and  $\tau_V$ . This choice is loosely motivated by the fact that the slab model reproduces the correlation between attenuations and inclination found in star-forming galaxies from observations (Conroy 2010; Wild et al. 2011; Battisti et al. 2017; Salim & Narayanan 2020) as well as simulations (*e.g.* Chevallard et al. 2013; Narayanan et al. 2018; Trayford et al. 2020). More importantly, the slab model can reproduce the  $A_V$  distribution of SDSS star-forming galaxies as well as the GSWLC2 sample, which includes quiescent galaxies (Appendix A). We therefore conclude that the slab model provides a sufficiently flexible empirical prescription for sampling  $A_V$  in our EDA. In our EDA, we also use a parameterization of  $\tau_V$  and  $\delta$  that depend linearly on  $\log M_*$  and  $\log \text{SSFR}$ . While the  $M_*$  and SSFR dependence of  $A_V$  is well-motivated and is found in, for instance, the Salim et al. (2018) GSWLC2 catalog (Appendix A), the linear dependence was chosen primarily for its simplicity. The EDA framework can be easily extended to more flexible parameterizations. In fact, a more flexible parameterization would likely reduce some of the discrepancies with the SDSS color-magnitude relations. The EDA produces broader distributions of optical colors than SDSS. Few galaxies in SDSS have  $g-r > 1$ , while some galaxies in the EDA broadly extend beyond this cut-off. In the UV, the EDA struggles to accurately reproduce the redder portions ( $FUV-NUV > 1.5$ ) of the UV color-magnitude relation. The main challenges for a more flexible parameterization would be model selection and finding a well-motivated parameteri-



**Figure 10.** IR dust emission luminosity predicted by the EDA with median parameter values of the TNG (blue) and EAGLE (green) posteriors as a function of  $M_r$ . The dust emission is estimated assuming the da Cunha et al. (2008) energy balance. Despite reproducing the same SDSS UV and optical color-magnitude relations, the EDA predicts significantly different IR dust emission for TNG and EAGLE. Therefore, including IR observations will significantly improve the constraints on EDA parameters and allow us to better differentiate galaxy formation models.

zation. Notwithstanding, for SDSS observations, our EDA prescription using parameter values from the TNG and EAGLE posteriors find good agreement.

We demonstrate that accounting for dust attenuation is essential when comparing simulations to observations. None of the simulations reproduce the UV and optical color-magnitude relation without dust attenuation (Figure 2). Furthermore, the fact that we can use the EDA to reproduce SDSS observations for different hydrodynamical simulations highlights how our current lack of understanding of dust limits our ability to closely compare galaxy formation models. Our EDA prescription is built on what we currently know about dust attenuation in galaxies: *e.g.* the Noll et al. 2009 parameterization, the UV bump, the slab model, etc. Yet with the EDA, two simulations that predict galaxy populations with significantly different physical properties (Figure 1) can reproduce the same SDSS observations. This suggests that dust is highly degenerate with the differences between simulations. Put another way — if we were to marginalize over dust in our comparison to observations, we would not be able to differentiate between the different galaxy physics prescriptions in the simulations. Hence, current limitations in our understanding of dust is a major bottleneck for investigating galaxy formation using simulations. In the next paper of the series, Starkenburg et al. (in preparation), we will examine whether we can compare the prescriptions for star formation quenching in different galaxy formation models once we include the EDA framework.

Fortunately, there are many avenues for improving our understanding of dust with a forward modeling approach. In this work, we used a restrictive  $M_r < -20$  complete SDSS galaxy sample. Figure 9 illustrates that our completeness cut limits the number of quiescent galaxies below  $M_* <$

$10^{11}M_{\odot}$ . Instead of imposing a completeness limit, we can include the actual SDSS selection function in the forward model (*e.g.* Dickey et al. 2020). This would allow us to compare the simulations with EDA to the entire SDSS sample, a substantially larger sample with a wider range of galaxies.

Upcoming surveys, such as the Bright Galaxy Survey (BGS) of the Dark Energy Spectroscopic Instrument (DESI; DESI~Collaboration et al. 2016; Ruiz-Macias et al. 2020) and galaxy evolution survey of the Prime Focus Spectrograph (PFS; Takada et al. 2014; Tamura et al. 2016), will also vastly expand galaxy observations. BGS, for instance, will measure  $10\times$  the number of galaxy spectra as SDSS out to  $z \sim 0.4$  and with its  $r \sim 20$  magnitude limit will probe a significant number of low redshift dwarf galaxies. Such an observational sample will allow us to place tighter constraints on the EDA parameters, which may enable comparisons of the underlying galaxy formation models, and shed light on dust in a broader range of galaxies.

In this work, we also only used observables derived from UV and optical photometry, which means that we have only examined one side of the impact that dust has on galaxy spectra. While dust attenuates light in the optical and UV, it emits light in IR. In fact, even though the TNG and EAGLE simulations reproduce the same UV and optical color-magnitude relations with the EDA, they predict significantly different dust emission in the IR. In Figure 10, we present IR dust emission luminosity,  $L_{\text{IR}}$ , predicted by the EDA with median parameter values of the TNG (blue) and EAGLE (green) posteriors as a function of the  $r$ -band absolute magnitude,  $M_r$ . The dust emissions are estimated using the standard energy balance assumption — *i.e.* all starlight attenuated by dust is reemitted in the IR (da Cunha et al. 2008).

Despite reproducing the same SDSS UV and optical color-magnitude relations, the EDA predicts significantly different IR dust emission for TNG and EAGLE. For TNG, the EDA predicts an overall  $\sim 0.3$  dex ( $2\times$ ) higher dust emissions than for EAGLE. Higher dust emission for TNG is consistent with the higher  $c_{\tau}$  we infer for TNG (Figure 4). It is also consistent with the fact that TNG predicts bluer galaxies and more luminous quiescent galaxies with red  $FUV-NUV$  color than EAGLE (Figure 2). Since IR dust emission measures the total dust attenuation, IR observations would specifically constrain the EDA and therefore break degeneracies between dust and the galaxy physics in simulations. Upcoming will provide crucial observation on this front. BGS, for instance, will have IR photometry from NEOWISE (Meisner et al. 2018). Moreover, *James Webb Space Telescope (JWST)* and *Nancy Grace Roman Space Telescope* will also provide valuable IR observations.

## 6. SUMMARY

In this work, we present the EDA, a framework for statistically applying dust attenuation to simulated galaxy populations. It uses a parameterization of the attenuation curves motivated from observations (Noll et al. 2009) and assigns attenuation curves to simulated galaxies based on their physical properties ( $M_*$  and SSFR). We apply the EDA to three state-of-the-art hydrodynamical simulations (SIMBA, TNG, and EAGLE) and forward model the optical and UV color-magnitude relations. We then compare the forward modeled simulations to a  $M_r < -20$  complete SDSS galaxy sample using likelihood-free inference. Based on this comparison, we find the following results:

- Dust attenuation is essential for simulations to reproduce observations. Without dust attenuation, all of the hydrodynamical simulations struggle to reproduce the observed UV and optical color-magnitude relation.
- With the EDA, the TNG and EAGLE simulations are able to produce UV and optical color-magnitude relations in good agreement with SDSS observations. SIMBA, however, overpredicts a substantial population of low mass high SFR galaxies. In order for SIMBA to reproduce SDSS, these galaxies require both high attenuation and reddening, which goes against the observed attenuation-slope relation.
- The attenuation curves predicted by the EDA for TNG and EAGLE are in excellent agreement with the observed attenuation-slope relation. They also closely reproduce the observed attenuation curves of star-forming galaxies. The success of the EDA in reproducing these observations, which were not included in the comparison, demonstrates the robustness of the EDA framework.
- Lastly, the EDA sheds light on dust attenuation in quiescent galaxies, which remains poorly understood due to observations challenges. **Simulated quiescent galaxies require significant UV and optical attenuation with shallow attenuation curves.** For all galaxies, we find that more massive galaxies have higher overall dust attenuation while galaxies with higher SSFR have steeper attenuation curves.

Our results clearly demonstrate that the EDA and a forward modeling approach provides key insights into dust attenuation. For those uninterested in dust, the EDA also provides a computationally feasible framework for marginalizing over dust when comparing simulations to observations. In the case of SIMBA, we found that the EDA dust attenuation is insufficient to accurately reproduce observations due to its excess starburst population. For TNG and EAGLE, however, dust attenuation is highly degenerate with differences in their galaxy physics prescriptions. Even though TNG and EAGLE predict galaxy populations with significantly different physical properties, there is enough uncertainty in our understanding of dust that by adjusting attenuation alone both TNG and EAGLE can reproduce the same SDSS observations. This also suggests that any comparisons across simulations must marginalize over dust attenuation or run the risk of overinterpretation. Therefore, our current understanding of dust, or lack of, limit our ability to distinguish between the various hydrodynamical models and is a major bottleneck for investigating galaxy formation using simulations.

The forward modeling approach we present offers many avenues for improving on our understanding of dust. In this paper, we used a restrictive  $M_r$  complete SDSS galaxy sample. Comparison to a larger observed galaxy sample will place tighter constraints on EDA parameters and enable better differentiation between the simulations. One way to expand the observed galaxy sample would be to remove the  $M_r$  completeness limit by including the selection function to our forward model. Upcoming surveys, such as the DESI Bright Galaxy Survey and the PFS Galaxy Evolution Survey, will also soon provide much larger observational galaxy samples. Furthermore, IR observations, which measure dust emission and trace dust attenuation, also have the potential to tightly constrain the

EDA parameters and therefore break degeneracies between dust and the galaxy physics in simulations. In the next paper of the series, we will use the forward modeling approach with the EDA to investigate star formation quenching in galaxy formation models. In other future works, we will apply the EDA and a forward modeling approach to more statistically powerful samples and include IR observables in order to tightly constrain and reveal new insights into dust attenuation.

## ACKNOWLEDGEMENTS

It's a pleasure to thank Marla Geha, Shy Genel, Daniel Kelson, Mariska Kriek, Ariyeh H. Maller, Desika Narayanan, Katherine Suess, Aaron Yung ... TBD. This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of High Energy Physics, under contract No. DE-AC02-05CH11231. This project used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

We thank the Illustris collaboration and the Virgo Consortium for making their simulation data publicly available, and the SIMBA collaboration for sharing their data with us. The EAGLE and SIMBA simulations were performed using the DiRAC-2 facility at Durham, managed by the ICC, and the PRACE facility Curie based in France at TGCC, CEA, Bruyères-le-Châtel.

This research was supported in part through the computational resources and staff contributions provided by the Quest high performance computing facility at Northwestern University, which is jointly supported by the Office of the Provost, the Office for Research, and Northwestern University Information Technology.

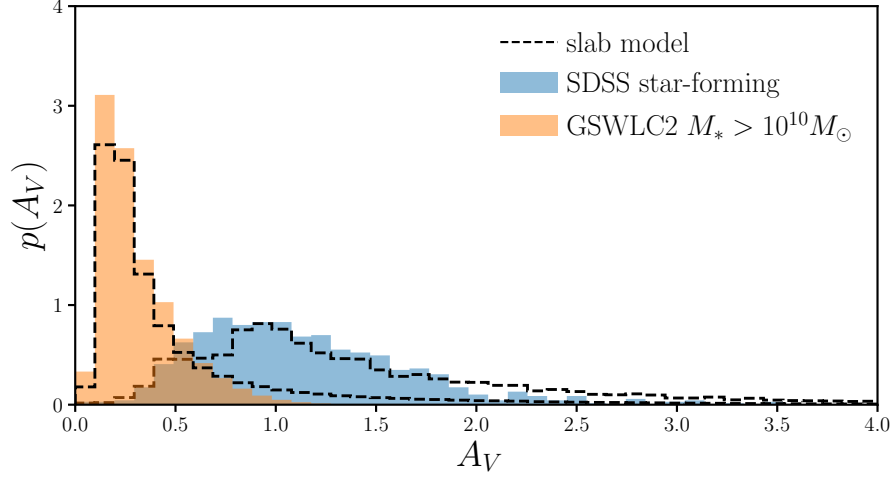
The data used in this work were, in part, hosted on facilities supported by the Scientific Computing Core at the Flatiron Institute, a division of the Simons Foundation, and the analysis was largely done using those facilities. The IQ (Isolated & Quiescent) Collaboratory thanks the Flatiron Institute for hosting the collaboratory and its meetings. The Flatiron Institute is supported by the Simons Foundation. Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P.

Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS web site is [www.sdss.org](http://www.sdss.org). The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions.

## APPENDIX

### A. THE SLAB MODEL BASED EDA

In our EDA prescription, we use the slab model to determine  $A_V$ , the amplitude of attenuation, as a function of a randomly sampled inclination,  $i$ , and  $\tau_V$  (see Eq. 3 in Section 3). The slab model is based on the assumption that dust in galaxies have slab-like geometry and illuminated by the stellar radiation source (Somerville & Primack 1999). For a given  $\tau_V$ , the attenuation depends solely on the orientation of the galaxy. While this simplification reproduces the correlation between  $A_V$  and  $i$  found in observed star-forming galaxies (*e.g.* Conroy et al. 2010; Wild et al. 2011; Battisti



**Figure 11.** The  $A_V$  distributions,  $p(A_V)$ , generated from the slab model (Eq. 3; black dash) compared to  $p(A_V)$  of star-forming galaxies in SDSS (blue) and of  $M_* > 10^{10} M_\odot$  galaxies in the Salim et al. (2018) GSWLC2 sample (orange). The  $A_V$  values for both observations are derived using SED fitting (Brinchmann et al. 2004; Salim et al. 2018). Meanwhile, for the slab model, we generate  $A_V$  values for each galaxy in the SDSS and GSWLC2 samples using Eq. 3 with its measured  $M_*$  and SSFR. The significant difference between the  $p(A_V)$  of SDSS and GSWLC2 is due to inconsistencies in the  $A_V$  measurements of the two catalogs. It illustrates the challenges in observationally measuring  $A_V$  and, again, highlights the advantages of our forward modeling approach. Despite the significant differences between the two, the slab model is able to generate  $p(A_V)$  in good agreement with  $p(A_V)$  from both observations with parameter values chosen within the Table ?? prior range. Therefore, the slab model provides a sufficiently flexible prescription for our EDA.

et al. 2017; Salim & Narayanan 2020), it ignores the detailed star-to-dust geometry that impacts the attenuation curve. It also does not provide a well-motivated prescription for quiescent galaxies, which typically have elliptical morphologies. Despite its limitations, the slab model provides a robust empirical prescription that allows us to produce realistic distributions of  $A_V$ .

In Figure 11, we compare the  $A_V$  distributions,  $p(A_V)$ , of star-forming galaxies in SDSS (blue) and galaxies in the Salim et al. (2018) GSWLC2 sample (orange) to the  $p(A_V)$  generated from the slab model (black dashed). The  $A_V$  values of the SDSS are derived using SED fitting from the Brinchmann et al. (2004) MPA-JHU catalog. The  $A_V$  values of the GSWLC2 galaxies are also derived from SED fitting UV and optical photometry from GALEX and SDSS observations as well as mid-IR photometry from WISE. We include all GSWLC2 galaxies, including quiescent galaxies, above  $M_* > 10^{10} M_\odot$ . We generate two  $p(A_V)$  with the slab model for the SDSS and GSWLC2 samples separately. For each SDSS/GSWLC2 galaxy, we determine  $A_V$  by uniformly sampling  $\cos i$  from 0 to 1 and derive  $\tau_V$  (Eq. 4) with the galaxy’s measured  $M_*$  and SSFR. We pick  $m_{\tau, M_*}$ ,  $m_{\tau, \text{SSFR}}$ ,  $c_\tau$  values within the prior range (Table 1) by hand to roughly reproduce the SDSS and GSWLC2  $p(A_V)$  distributions.

Galaxies in SDSS and GSWLC2 have substantially different  $p(A_V)$ . This is due to inconsistencies in the  $A_V$  measurements of MPA-JHU and GSWLC2 — even for the same galaxy, the  $A_V$  measure-

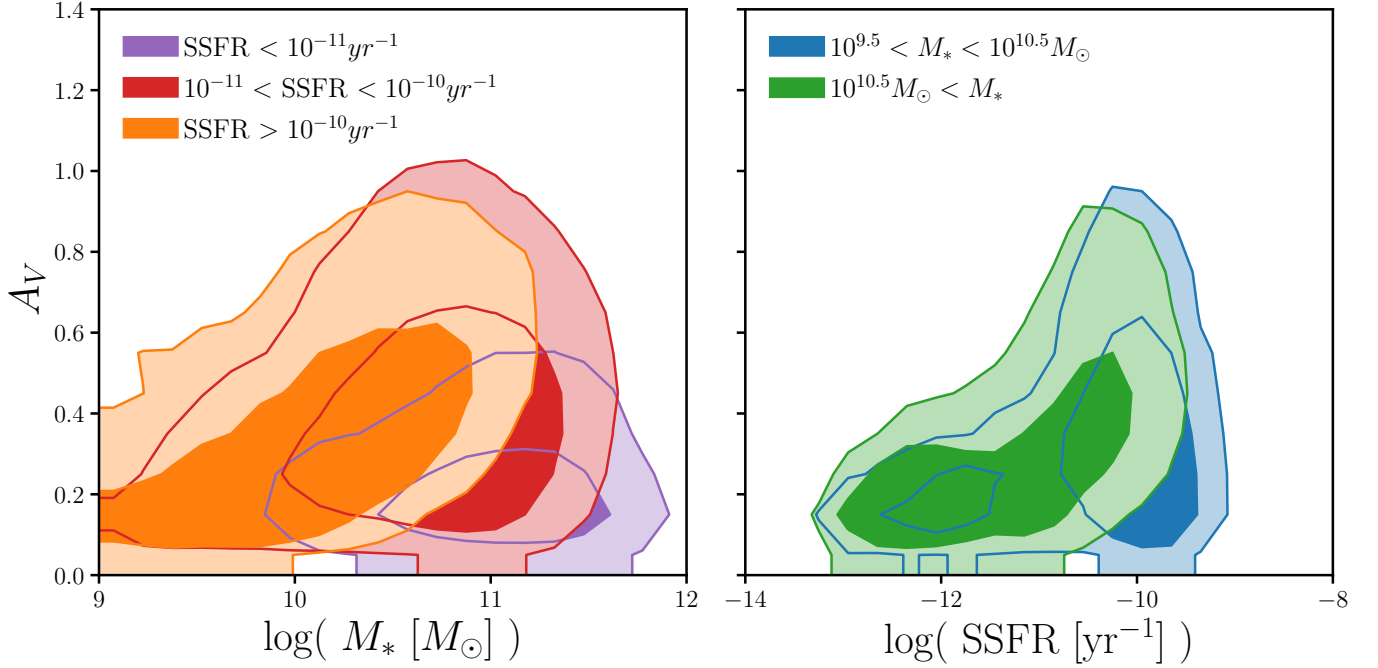


ments from the two catalogs differ significantly. This difference in  $p(A_V)$  illustrates the challenges in directly measuring dust attenuation and yet again highlights the advantages of our forward modeling approach. Despite the dramatic differences between the two, the slab model can produce  $p(A_V)$  in good agreement with both observed distributions. We therefore conclude that the slab model provides a sufficiently flexible prescription to sample a realistic distribution of  $A_V$ .

In addition to the slab model, in the EDA, we also use a linear dependence on  $M_*$  and SSFR in the  $V$  band optical depth,  $\tau_V$  (see Eq. 4). This parameterization is motivated by observations that find significant correlation between  $A_V$  and  $M_*$  and SSFR (*e.g.* Garn & Best 2010; Battisti et al. 2016; Salim & Narayanan 2020). We take a closer look at this correlation using the GSWLC2 sample in Figure 12. We present the dependence of  $A_V$  on  $M_*$  (left panel) and SSFR (right panel). In the left panel, we divide the GSWLC2 galaxies by SSFR:  $\text{SSFR} < 10^{-11} \text{yr}^{-1}$  (purple),  $10^{-11} < \text{SSFR} < 10^{-10} \text{yr}^{-1}$  (red), and  $10^{-10} < \text{SSFR}$  (orange). For each of the SSFR bins, we find significant  $\log M_*$  dependence in  $A_V$ : galaxies with higher SSFR have a stronger  $M_*$  dependence. In the right, we divide the galaxies by  $M_*$ :  $10^{9.5} < M_* < 10^{10.5} M_\odot$  (blue) and  $10^{10.5} M_\odot < M_*$  (green). Although both  $M_*$  bins have some SSFR dependence, the dependence is stronger for galaxies with  $M_* > 10^{10.5} M_\odot$ . This stellar mass limit roughly corresponds to galaxies that are included in our forward model (see Figure 9). The  $M_*$  and SSFR dependence we find in  $A_V$  from the GSWLC2 sample is consistent with previous observations and further motivates our EDA prescription.

## REFERENCES

- Abazajian K. N., et al., 2009, *The Astrophysical Journal Supplement Series*, 182, 543
- Aihara H., et al., 2011, *The Astrophysical Journal Supplement Series*, 193, 29
- Akeret J., Refregier A., Amara A., Seehars S., Hasner C., 2015, *Journal of Cosmology and Astroparticle Physics*, 2015, 043
- Ali S. S., Bremer M. N., Phillipps S., De Propris R., 2018, *Monthly Notices of the Royal Astronomical Society*, 476, 1010
- Alsing J., Wandelt B., Feeney S., 2018, arXiv:1801.01497 [astro-ph]
- Anglés-Alcázar D., Davé R., Faucher-Giguère C.-A., Özel F., Hopkins P. F., 2017, *Monthly Notices of the Royal Astronomical Society*, 464, 2840
- Baes M., Trčka A., Camps P., Nersesian A., Trayford J., Theuns T., Dobbels W., 2019, arXiv:1901.08878 [astro-ph]
- Battisti A. J., Calzetti D., Chary R.-R., 2016, *The Astrophysical Journal*, 818, 13
- Battisti A. J., Calzetti D., Chary R.-R., 2017, *The Astrophysical Journal*, 840, 109
- Beaumont M. A., Cornuet J.-M., Marin J.-M., Robert C. P., 2009, *Biometrika*, 96, 983
- Blanton M. R., Roweis S., 2007, *The Astronomical Journal*, 133, 734
- Blanton M. R., et al., 2005, *The Astronomical Journal*, 129, 2562
- Blanton M. R., Kazin E., Muna D., Weaver B. A., Price-Whelan A., 2011, *The Astronomical Journal*, 142, 31
- Booth C. M., Schaye J., 2009, *Monthly Notices of the Royal Astronomical Society*, 398, 53
- Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, *Monthly Notices of the Royal Astronomical Society*, 351, 1151
- Burgarella D., Buat V., Iglesias-Páramo J., 2005, *Monthly Notices of the Royal Astronomical Society*, 360, 1413
- Calzetti D., 2001, *New Astronomy Reviews*, 45, 601
- Cameron E., Pettitt A. N., 2012, *Monthly Notices of the Royal Astronomical Society*, 425, 44
- Camps P., Baes M., 2015, *Astronomy and Computing*, 9, 20



**Figure 12.** Dependence of  $A_V$  on  $M_*$  (left) and SSFR (right) for the Salim et al. (2018) GSWLC2 sample. In the left panel, we divide the GSWLC2 sample into bins of SSFR:  $\text{SSFR} < 10^{-11} \text{yr}^{-1}$  (purple),  $10^{-11} < \text{SSFR} < 10^{-10} \text{yr}^{-1}$  (red), and  $10^{-10} < \text{SSFR}$  (orange). In each of the SSFR bins, we find significant  $M_*$  dependence. In the right panel, we divide the sample into bins of  $M_*$ :  $10^{9.5} < M_* < 10^{10.5} M_\odot$  (blue) and  $10^{10.5} M_\odot < M_*$  (green). In the  $M_* > 10^{10.5} M_\odot$  bin, which roughly corresponds to our SDSS sample, we find significant SSFR dependence. The  $M_*$  and SSFR dependence in  $A_V$  we find in GSWLC2 is consistent with previous works and provides further motivation for our EDA prescription.

- Carnall A. C., Leja J., Johnson B. D., McLure R. J., Dunlop J. S., Conroy C., 2018, arXiv:1811.03635 [astro-ph]
- Chabrier G., 2003, *Publications of the Astronomical Society of the Pacific*, 115, 763
- Chevallard J., Charlot S., Wandelt B., Wild V., 2013, *Monthly Notices of the Royal Astronomical Society*, 432, 2061
- Code A. D., 1969, *Publications of the Astronomical Society of the Pacific*, 81, 475
- Conroy C., 2010, *Monthly Notices of the Royal Astronomical Society*, 404, 247
- Conroy C., 2013, *Annual Review of Astronomy and Astrophysics*, 51, 393
- Conroy C., Gunn J. E., White M., 2009, *The Astrophysical Journal*, 699, 486
- Conroy C., White M., Gunn J. E., 2010, *The Astrophysical Journal*, 708, 58
- Crain R. A., et al., 2015, *Monthly Notices of the Royal Astronomical Society*, 450, 1937
- DESI~Collaboration et al., 2016, arXiv:1611.00036 [astro-ph]
- Dalla Vecchia C., Schaye J., 2012, *Monthly Notices of the Royal Astronomical Society*, 426, 140
- Dantas M. L. L., Coelho P. R. T., Sánchez-Blázquez P., 2021, *Monthly Notices of the Royal Astronomical Society*, 500, 1870
- Davé R., Thompson R., Hopkins P. F., 2016, *Monthly Notices of the Royal Astronomical Society*, 462, 3265
- Davé R., Rafieferantsoa M. H., Thompson R. J., 2017a, arXiv:1704.01135 [astro-ph]
- Davé R., Rafieferantsoa M. H., Thompson R. J., Hopkins P. F., 2017b, *Monthly Notices of the Royal Astronomical Society*, 467, 115
- Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, *Monthly Notices of the Royal Astronomical Society*, 486, 2827
- Del Moral P., Doucet A., Jasra A., 2012, *Statistics and Computing*, 22, 1009

- Dickey C. M., et al., 2020, arXiv e-prints, 2010, arXiv:2010.01132
- Diggle P. J., Gratton R. J., 1984, *Journal of the Royal Statistical Society. Series B (Methodological)*, 46, 193
- Draine B. T., 2003, *The Astrophysical Journal*, 598, 1017
- Fontanot F., Somerville R. S., Silva L., Monaco P., Skibba R., 2009, *Monthly Notices of the Royal Astronomical Society*, 392, 553
- Galliano F., Galametz M., Jones A. P., 2018, *Annual Review of Astronomy and Astrophysics*, 56, 673
- Garn T., Best P. N., 2010, *Monthly Notices of the Royal Astronomical Society*, 409, 421
- Genel S., et al., 2014, *Monthly Notices of the Royal Astronomical Society*, 445, 175
- Gonzalez-Perez V., Lacey C. G., Baugh C. M., Frenk C. S., Wilkins S. M., 2013, *Monthly Notices of the Royal Astronomical Society*, 429, 1609
- Gordon K. D., Witt A. N., Carruthers G. R., Christensen S. A., Dohne B. C., 1994, *The Astrophysical Journal*, 432, 641
- Granato G. L., Lacey C. G., Silva L., Bressan A., Baugh C. M., Cole S., Frenk C. S., 2000, *The Astrophysical Journal*, 542, 710
- Hahn C., Vakili M., Walsh K., Hearin A. P., Hogg D. W., Campbell D., 2017a, *Monthly Notices of the Royal Astronomical Society*, 469, 2791
- Hahn C., Tinker J. L., Wetzel A. R., 2017b, *The Astrophysical Journal*, 841, 6
- Hahn C., Tinker J. L., Wetzel A., 2019a, arXiv:1910.01644 [astro-ph]
- Hahn C., Beutler F., Sinha M., Berlind A., Ho S., Hogg D. W., 2019b, *Monthly Notices of the Royal Astronomical Society*, 485, 2956
- Hahn C., et al., 2019c, *The Astrophysical Journal*, 872, 160
- Han Z., Podsiadlowski P., Lynas-Gray A. E., 2007, *Monthly Notices of the Royal Astronomical Society*, 380, 1098
- Hayward C. C., Smith D. J. B., 2015, *Monthly Notices of the Royal Astronomical Society*, 446, 1512
- Hopkins P. F., 2015, *Monthly Notices of the Royal Astronomical Society*, 450, 53
- Hopkins P. F., et al., 2017, arXiv:1707.07010 [astro-ph]
- Hou K.-C., Hirashita H., Nagamine K., Aoyama S., Shimizu I., 2017, *Monthly Notices of the Royal Astronomical Society*, 469, 870
- Inoue A. K., 2005, *Monthly Notices of the Royal Astronomical Society*, 359, 171
- Ishida E. E. O., et al., 2015, *Astronomy and Computing*, 13, 1
- Jonsson P., 2006, *Monthly Notices of the Royal Astronomical Society*, 372, 2
- Katsianis A., et al., 2020, *Monthly Notices of the Royal Astronomical Society*, 492, 5592
- Kaviraj S., et al., 2007, *The Astrophysical Journal Supplement Series*, 173, 619
- Kirkpatrick A., Pope A., Sajina A., Roebuck E., Yan L., Armus L., Díaz-Santos T., Stierwalt S., 2015, *The Astrophysical Journal*, 814, 9
- Kriek M., Conroy C., 2013, *The Astrophysical Journal Letters*, 775, L16
- Le Cras C., Maraston C., Thomas D., York D. G., 2016, *Monthly Notices of the Royal Astronomical Society*, 461, 766
- Leja J., Johnson B. D., Conroy C., van Dokkum P. G., Byler N., 2017, *The Astrophysical Journal*, 837, 170
- Leja J., Johnson B. D., Conroy C., van Dokkum P., 2018, *The Astrophysical Journal*, 854, 62
- Leja J., Carnall A. C., Johnson B. D., Conroy C., Speagle J. S., 2019, *ApJ*, 876, 3
- Lin C.-A., Kilbinger M., Pires S., 2016, *Astronomy and Astrophysics*, 593, A88
- McAlpine S., et al., 2016, *Astronomy and Computing*, 15, 72
- Meisner A. M., Lang D., Schlegel D. J., 2018, *Research Notes of the American Astronomical Society*, 2, 1
- Narayanan D., Conroy C., Davé R., Johnson B. D., Popping G., 2018, *The Astrophysical Journal*, 869, 70
- Natale G., Popescu C. C., Tuffs R. J., Debattista V. P., Fischera J., Grootes M. W., 2015, *Monthly Notices of the Royal Astronomical Society*, 449, 243
- Nelson D., et al., 2015, *Astronomy and Computing*, 13, 12
- Nelson D., et al., 2018, *Monthly Notices of the Royal Astronomical Society*, 475, 624
- Noll S., Burgarella D., Giovannoli E., Buat V., Marcillac D., Muñoz-Mateos J. C., 2009, *Astronomy and Astrophysics*, 507, 1793

- O'Connell R. W., 1999, [Annual Review of Astronomy and Astrophysics](#), 37, 603
- Pillepich A., et al., 2018, [Monthly Notices of the Royal Astronomical Society](#), 473, 4077
- Popping G., Somerville R. S., Galametz M., 2017, [Monthly Notices of the Royal Astronomical Society](#), 471, 3152
- Pritchard J. K., Seielstad M. T., Perez-Lezaun A., Feldman M. W., 1999, [Molecular Biology and Evolution](#), 16, 1791
- Reddy N. A., et al., 2015, [The Astrophysical Journal](#), 806, 259
- Rocha M., Jonsson P., Primack J. R., Cox T. J., 2008, [Monthly Notices of the Royal Astronomical Society](#), 383, 1281
- Ruiz-Macias O., et al., 2020, arXiv:2007.14950 [astro-ph]
- Salim S., Narayanan D., 2020, arXiv:2001.03181 [astro-ph]
- Salim S., et al., 2016, [The Astrophysical Journal Supplement Series](#), 227, 2
- Salim S., Boquien M., Lee J. C., 2018, [The Astrophysical Journal](#), 859, 11
- Schaye J., et al., 2015, [Monthly Notices of the Royal Astronomical Society](#), 446, 521
- Seon K.-I., Draine B. T., 2016, [The Astrophysical Journal](#), 833, 201
- Somerville R. S., Primack J. R., 1999, [Monthly Notices of the Royal Astronomical Society](#), 310, 1087
- Somerville R. S., Gilmore R. C., Primack J. R., Domínguez A., 2012, [Monthly Notices of the Royal Astronomical Society](#), 423, 1992
- Speagle J. S., Steinhardt C. L., Capak P. L., Silverman J. D., 2014, [The Astrophysical Journal Supplement Series](#), 214, 15
- Springel V., 2005, [Monthly Notices of the Royal Astronomical Society](#), 364, 1105
- Springel V., et al., 2018, [Monthly Notices of the Royal Astronomical Society](#), 475, 676
- Steinacker J., Baes M., Gordon K. D., 2013, [Annual Review of Astronomy and Astrophysics](#), 51, 63
- Takada M., et al., 2014, [Publications of the Astronomical Society of Japan](#), 66, R1
- Tamura N., et al., 2016, in [Ground-Based and Airborne Instrumentation for Astronomy VI](#). eprint: arXiv:1608.01075, p. 99081M, doi:10.1117/12.2232103
- Tavare S., Balding D. J., Griffiths R. C., Donnelly P., 1997, [Genetics](#), 145, 505
- Tinker J., Wetzel A., Conroy C., 2011, preprint, 1107, arXiv:1107.5046
- Trayford J. W., et al., 2015, [Monthly Notices of the Royal Astronomical Society](#), 452, 2879
- Trayford J. W., et al., 2017, [Monthly Notices of the Royal Astronomical Society](#), 470, 771
- Trayford J. W., Lagos C. d. P., Robotham A. S. G., Obreschkow D., 2020, [Monthly Notices of the Royal Astronomical Society](#), 491, 3937
- Tress M., et al., 2018, [Monthly Notices of the Royal Astronomical Society](#), 475, 2363
- Vogelsberger M., et al., 2014, [Monthly Notices of the Royal Astronomical Society](#), 444, 1518
- Vogelsberger M., et al., 2020, [Monthly Notices of the Royal Astronomical Society](#), 492, 5167
- Walcher J., Groves B., Budavári T., Dale D., 2011, [Astrophysics and Space Science](#), 331, 1
- Wechsler R. H., Tinker J. L., 2018, preprint, 1804, arXiv:1804.03097
- Weinberger R., et al., 2018, [Monthly Notices of the Royal Astronomical Society](#), 479, 4056
- Weyant A., Schafer C., Wood-Vasey W. M., 2013, [The Astrophysical Journal](#), 764, 116
- Wild V., Charlot S., Brinchmann J., Heckman T., Vince O., Pacifici C., Chevallard J., 2011, [Monthly Notices of the Royal Astronomical Society](#), 417, 1760
- Wilkins S. M., Gonzalez-Perez V., Lacey C. G., Baugh C. M., 2012, [Monthly Notices of the Royal Astronomical Society](#), 424, 1522
- Witt A. N., Gordon K. D., 1996, [The Astrophysical Journal](#), 463, 681
- Witt A. N., Gordon K. D., 2000, [The Astrophysical Journal](#), 528, 799
- Yi S., Demarque P., Kim Y.-C., 1997, [The Astrophysical Journal](#), 482, 677
- Yung L. Y. A., Somerville R. S., Finkelstein S. L., Popping G., Davé R., 2019, [Monthly Notices of the Royal Astronomical Society](#), 483, 2983
- da Cunha E., Charlot S., Elbaz D., 2008, [Monthly Notices of the Royal Astronomical Society](#), 388, 1595