

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

CHANGHOON HAHN,<sup>1,\*</sup> TJITSKE K. STARKENBURG,<sup>2</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*

(Dated: DRAFT -- 9314c56 -- 2020-11-24 -- NOT READY FOR DISTRIBUTION)

## ABSTRACT

We present the Empirical Dust Attenuation (EDA) model, an empirical framework for applying dust attenuation to simulated galaxy populations. The EDA uses a parameterization of the attenuation curve based on observations and a statistical prescription for assigning the curves that includes correlations with galaxy properties. We apply the EDA to three state-of-the-art hydrodynamical simulations — SIMBA, Illustris TNG, and EAGLE — and forward model their optical and UV color-magnitude relations,  $(G-R) - M_r$  and  $(FUV-NUV) - M_r$ . Then we compare them to SDSS observations using simulation-based inference. For TNG and EAGLE, with the EDA we accurately reproduce the SDSS color magnitude relations. Meanwhile, for SIMBA, we struggle to reproduce observations due to an overprediction of starburst galaxies. The bestfit EDA for TNG and EAGLE also predict attenuation-slope relations and attenuation curves of star-forming galaxies in good agreement with observations. Beyond reproducing observations, the EDA also sheds light on dust attenuation in quiescent galaxies, which is challenging to measure in observations. Quiescent galaxies have significant UV and optical with shallow attenuation curves. The EDA, more broadly, reveals the connections between galaxy properties and dust attenuation: more massive galaxies have higher overall dust attenuation while galaxies with higher SSFR have steeper attenuation curves. In addition to improving our understanding of dust in galaxies, our results underscore the advantages of the EDA and a forward modeling approach.

*Keywords:* keyword1 – keyword2 – keyword3

## 1. INTRODUCTION

Dust in the interstellar medium of a galaxy can dramatically impact the observed light from the galaxy over the full range of the electromagnetic spectrum. Dust emits in the infrared (IR) and modifies the galaxy’s stellar radiation through absorption and scatter from the near-infrared (NIR)

\* hahn.changhoon@gmail.com

to ultraviolet (UV). It can, therefore, dramatically impact the physical properties of galaxies we infer from optical and UV light, such as star formation rate (SFR), stellar mass ( $M_*$ ), or stellar ages (see reviews by [Walcher et al. 2011](#); [Conroy 2013](#)). Since these properties are the building blocks for our understanding of galaxies and how they evolve, a better understanding of dust not only provides insights into the physical processes related to dust, but it is also crucial to all galaxy studies.

The combined effect of dust on the spectral energy distribution (SED) of a galaxy is typically described using an attenuation curve,  $A(\lambda)$ . Observations have now established the major features in  $A(\lambda)$ . In the UV, the curve steeply rises due to absorptions by small grains. At 2175, in the near-UV (NUV), there is an absorption bump referred to as the UV dust bump. Then at longer wavelengths, the curves take a power-law shape. For an overview, we refer to reviews by [Calzetti \(2001\)](#); [Draine \(2003\)](#); [Galliano et al. \(2018\)](#). Attenuation curves, however, are not universal. Observations find a wide range of attenuation curves in the local Universe ([Wild et al. 2011](#); [Battisti et al. 2017](#); [Salim et al. 2018](#); [Salim & Narayanan 2020](#)) as well as at high redshifts (*e.g.* [Reddy et al. 2015](#); [Salmon et al. 2016](#)).

To understand the origin of this variation in attenuation curves, a number of observational works have examined the correlation between dust attenuation and galaxy properties. Using 23,000 star-forming galaxies in the Sloan Digital Sky Survey (SDSS), [Wild et al. \(2011\)](#) found that the slope of the attenuation curves varies strongly with galaxy axial ratio and weakly with specific SFR. Similarly, [Battisti et al. \(2017\)](#), from 5,500 star-forming galaxies, found tentative trends with stellar age, specific star formation rate, stellar mass, and metallicity. Meanwhile, [Salim et al. \(2018\)](#), from 230,000 galaxies in the GALEX-SDSS-WISE Legacy Catalog 2 (GSWLC2), find a significant  $M_*$  dependence on the slope. They argue that this is caused by the underlying “attenuation-slope relation”, a trend between the amplitude of the  $V$ -band attenuation ( $A_V$ ) and slope, where galaxies with higher  $A_V$  have shallower slopes. At low attenuation, dust scattering dominates absorption so the attenuation curve steepens because red light scatters isotropically while blue light scatters forward ([Gordon et al. 1994](#); [Witt & Gordon 2000](#); [Draine 2003](#)). At high attenuation dust absorption is dominant and the attenuation curve is shallower ([Chevallard et al. 2013](#)). Nevertheless, there is still no clear consensus on the connection between attenuation curves and galaxy properties. Also, studies so far have focused only on star-forming galaxies and little is known about dust attenuation in quiescent galaxies.

Alongside observations, theoretical efforts that model radiative transfer of stellar light through a dusty ISM also provide insights into dust attenuation. Radiative transfer models span a wide range of geometric configurations of stars and dust. For instance, models focused on isolating the physical effects of dust have considered simple slab or shell-like dust geometries illuminated by stellar radiation (*e.g.* [Witt & Gordon 1996, 2000](#); [Seon & Draine 2016](#)). Other models, focused on modeling dust attenuation in galaxies as a whole, have applied 3D dust radiative transfer in hydrodynamic simulations of idealized galaxies (*e.g.* [Jonsson 2006](#); [Rocha et al. 2008](#); [Hayward & Smith 2015](#); [Natale et al. 2015](#); [Hou et al. 2017](#)). Dust attenuation has also been examined in a cosmological context using semi-analytic models (SAMs) that do not track baryonic growth directly but make simple physically motivated assumptions about the resulting galaxy properties from dark matter growth (*e.g.* [Granato et al. 2000](#); [Fontanot et al. 2009](#); [Wilkins et al. 2012](#); [Gonzalez-Perez et al. 2013](#)). Lastly, radiative

transfer models have recently been applied to cosmological hydrodynamical simulations (*e.g.* Camps & Baes 2015; Narayanan et al. 2018; Trayford et al. 2020). As in observations, simulations find significant variation in attenuation curves.

Despite their progress, there are still major challenges for radiative transfer models. For instance, they produce attenuation-slope relations that are significantly steeper than observations. Also, many models require significant hand-tuning (*e.g.* propagating rays/photons into particular cells). Assumptions in the underlying dust grain models also produce systematic uncertainties that are difficult to quantify (see Steinacker et al. 2013, for a review). Lastly, radiative transfer models are computationally expensive. So applying multiple radiative transfer dust models to multiple simulations for comparisons is prohibitive. Using Markov Chain Monte Carlo sampling with them for parameter exploration or to marginalize over the impact of dust would be *intractable*.

In this paper, we take a different approach and present the Empirical Dust Attenuation (EDA) model, a flexible framework for statistically applying dust attenuation to galaxy populations. The EDA model assigns, to every galaxy, a different dust attenuation curve. The attenuation curves are parameterized with a function-form used in observational studies (Noll et al. 2009) and the parameters of the curve for each galaxy (*e.g.* optical depth, slope) is sampled from distributions set by the EDA model parameters and the galaxy’s properties. By sampling the attenuation curve parameters, it produces realistic variations among the attenuation curves. We emphasize that the EDA does not seek to produce realistic dust attenuation for individual galaxies, like radiative transfer models. Instead, it aims to produce realistic dust attenuations for the entire galaxy sample and enable galaxy population-to-population comparisons.

There are a number of advantages to our EDA approach. The EDA uses the same functional-form for the attenuation curves as observational studies, so predictions of the model can easily be compared to observational constraints. We also formulate the EDA parameters so that they are easily interpretable and directly reveal correlations between dust attenuation and galaxy properties. Lastly, the EDA model is computationally inexpensive.

The EDA model, with its flexibility and speed, provides a key crucial in a *forward modeling approach* to comparing simulations to observations (*e.g.* Nelson et al. 2018; Baes et al. 2019; Trčka et al. 2020; ?).. In the standard approach, the galaxy properties (*e.g.* SFR,  $M_*$ ) predicted by simulations are compared to those derived from observations. Instead, in the forward modeling approach, observables (*e.g.* magnitude, color) are *forward modeled* for each galaxy in the simulations. Then the simulations are compared to observations directly in observational space. With a forward modeling approach, comparisons are not limited by variations, inconsistencies, and biases of different observational methods for measuring galaxy properties (*e.g.* different tracers of SFR or  $M_*$ ). Selection functions and systematic effects can also be included in the forward model.

Of course to produce realistic observables of simulated galaxies, the forward model must include some prescription for applying dust attenuation. In principle, a radiative transfer model can be used for this purpose; however, its computational cost would severely limit any exploration of the “dust parameters”. With the EDA, however, we can apply a wide range of realistic dust attenuation to simulated galaxies in a matter of seconds. We can, therefore, easily explore and sample the EDA pa-



**Figure 1.** The stellar mass functions,  $\Phi(M_*)$  (left-most panel), and  $M_* - \text{SFR}$  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  $M_* - \text{SFR}$  relation (black dashed) of our observational sample from SDSS. Uncertainties for the SDSS  $\Phi(M_*)$  are derived using jackknife resampling. We describe the simulations and observations above in Section 2. *Differences in  $\Phi(M_*)$  and the  $M_* - \text{SFR}$  relations among the hydrodynamic simulations highlight how they predict galaxy populations with significantly different the physical properties.*

parameter space to infer the relationship between dust attenuation and galaxy properties. *For readers uninterested in dust*, the EDA provides a way to treat dust parameters as *nuisance* parameters and tractably marginalize over them. In other words, the EDA can be used to effectively marginalize over dust attenuation.

In this work, we present a simple EDA model that uses the Noll et al. (2009) attenuation curve parameterization and includes correlations between dust attenuation and galaxy  $M_*$  and SFR (Section 3). We apply the EDA to three state-of-the-art cosmological large-scale hydrodynamical simulations, SIMBA, IllustrisTNG, and EAGLE (Section 2), and compare them to observed SDSS galaxies (Section 2.4). In Section 3, we describe our EDA model in detail. Finally, in Section 5, we present the results of our comparison and discuss their implications for our understanding of dust attenuation as well as of its connection to galaxy properties.

## 2. DATA

In this paper, we present the Empirical Dust Attenuation (EDA) model — a flexible prescription for applying dust attenuation to simulated galaxy populations. We also demonstrate how the EDA can be used in a forward modeling approach to compare galaxy populations in simulations and observations. For our simulations, we use three cosmological hydrodynamical simulations: the Illustris TNG (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. We include, for reference,  $\Phi(M_*)$  and the  $M_* - \text{SFR}$  relation for our

observational galaxy sample from SDSS. For the simulations,  $M_*$  is the total stellar mass within the host halo, excluding any stellar mass in subhalos; SFR is the instantaneous SFR derived from dense and cold gas. 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. Figure 1 illustrates that the hydrodynamical simulations predict significantly different SMFs and  $M_* - \text{SFR}$  relations. 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*.

### 2.1. Illustris TNG

The Illustris TNG simulation<sup>2</sup> is a cosmological hydrodynamic simulation of comoving volume  $(110.7 \text{ Mpc})^3$  (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 (Weinberger et al. 2018), where feedback energy is injected in the form of a kinetic AGN-driven wind at low SMBH accretion rates. 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 volume of  $(100 \text{ Mpc})^3$ . 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

<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>



two-phase winds for star formation driven galactic winds, and feedback from Type I supernovae and AGB stars. Meanwhile, it uses updated models for AGN feedback and on-the-fly dust model. 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 start with the  $M_* > 10^{9.4} h^{-2} M_\odot$  complete Tinker et al. (2011) volume-limited sample and then impose a  $M_r < -20$  completeness cut. The original 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). In this work, we focus on observables that can be consistently defined and derived in both simulations and observables:  $M_r$ ,  $G-R$ , and  $FUV-NUV$ . For our 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

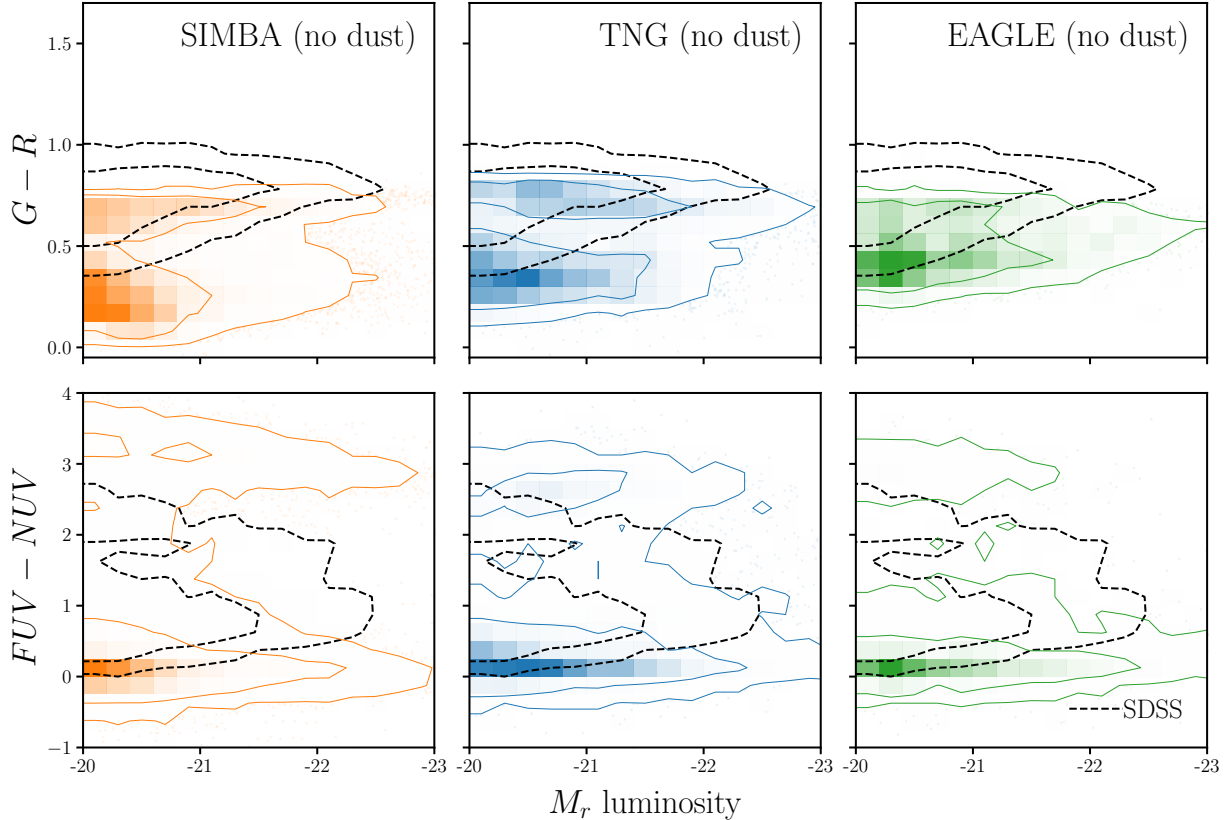
One of the main goals of this work is to conduct an “apples-to-apples” comparison between the simulations above and observations using a forward modeling approach. A crucial step is to forward model the observables for our simulations — in this work, we use  $r$ -band luminosity ( $M_r$ ), optical color ( $G-R$ ), and UV color ( $FUV-NUV$ ). As the first step, 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 the spectrum of a simple stellar population (SSP) using FSPS and take the linear combination of them, weighted by the total stellar mass formed in the bin, 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.).

In the next step of the forward model, we use the EDA model assign dust attenuation curves for each of the simulated galaxies based on their physical properties and apply dust attenuation to their SEDs. We describe the EDA in detail in Section 3.

From the dust attenuated SEDs, we measure the observables by convolving them with transmission curves of the GALEX  $FUV$ , GALEX  $NUV$ , SDSS  $g$ , and SDSS  $r$  broadband filters. Lastly, we add realistic noise to  $M_r$ ,  $G-R$ , and  $FUV-NUV$ . We sample the noise from normal distributions, where the standard deviations are drawn from  $\chi^2$  fits to the the observed uncertainty distributions from NASA-Sloan Atlas.

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



**Figure 2.** We present the forward modeled optical and UV color-magnitude relations of simulated galaxies in SIMBA (left), TNG (center), and EAGLE (right). The simulations above do *not* yet include the EDA or any prescription for dust attenuation.  $(G-R) - M_r$  (top) and  $(FUV-NUV) - M_r$  (bottom) are the main observables used throughout the paper. They are derived using SEDs constructed from the star-formation and metallicity of the simulated galaxies. These observables are consistently defined and measured as our SDSS sample (Section 2.5), which we include for comparison (black dashed). *Without dust, the hydrodynamical simulations cannot reproduce the observed optical or UV color-magnitude.*

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 of SIMBA (left), TNG (center) and EAGLE (right). The observables for the simulations do not yet include any prescription for dust attenuation. Comparison to SDSS observations (black dashed) clearly demonstrate that *without dust attenuation, the hydrodynamical simulations cannot reproduce the observed optical or UV color-magnitude relations.*

### 3. THE EMPIRICAL DUST ATTENUATION MODEL

In this section we present the Empirical Dust Attenuation (EDA) model, a flexible prescription for applying dust attenuation curves to galaxy populations. With the EDA, we can apply a wide variety of dust attenuation and include correlation between dust attenuation and physical galaxy properties. Later, we demonstrate that we can accurately reproduce SDSS observations with the EDA and use it to test galaxy formation models and shed light on dust in galaxies.

We define 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 at the  $V$  band,

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

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

To determine  $A(\lambda)$  for each galaxy, we first assign  $A_V$  using the slab model from [Somerville & Primack \(1999\)](#); [Somerville et al. \(2012\)](#). In the slab model,  $A_V$  is calculated based on the inclination of the galaxy,  $i$ , and its optical depth,  $\tau_V$ :

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

For each galaxy, we uniformly sample  $i$  from 0 to  $\pi/2$ . Then in  $\tau_V$  we include the correlation between  $A_V$  and galaxy properties ( $M_*$  and SFR), found in both observations and simulations (*e.g.* [Narayanan et al. 2018](#); [Salim & Narayanan 2020](#)). We parameterize  $\tau_V$  with a simple and flexible linear  $M_*$  and SFR dependence:

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

$m_{\tau, M_*}$  quantifies the  $M_*$  dependence,  $m_{\tau, \text{SSFR}}$  quantifies the SSFR dependence, and  $c_\tau$  quantifies the overall amplitude. Since  $\tau_V$  is optical depth, we impose a  $\tau_V \geq 0$  limit. The slab model is a naive approximation. In reality,  $A_V$  for a galaxy depends on complexities of its star-to-dust geometry, variations in the extinction curves, and other properties beyond just inclination and  $\tau_V$ . The purpose of the EDA, however, is not to accurately model dust attenuation for individual galaxies, but rather to accurately model the distribution of dust attenuation for galaxy populations. In this regard, the slab model qualitatively reproduces the correlation between  $A_V$  and  $i$  found in the literature: 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](#)). More importantly, the  $A_V$  distribution,  $p(A_V)$ , produced using the slab model with uniformly sampled inclinations closely matches the  $p(A_V)$  of our SDSS sample (Figure 10). Replacing the slab model with a more flexible prescription for sampling  $A_V$  does not significant impact our analysis (Appendix B). Therefore, we conclude that the slab model is a sufficient, flexible empirical prescription for sampling  $A_V$ .

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:

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



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

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

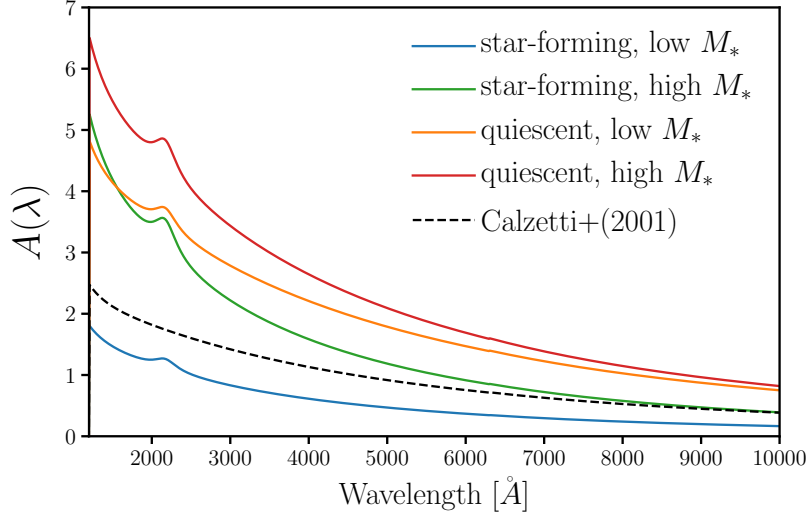
$D(\lambda)$  in Eq. 5 is the UV dust bump, which we parameter using the standard Lorentzian-like Drude profile:

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

where  $\lambda_0$ ,  $\Delta\lambda$ , and  $E_b$  are the central wavelength, FWHM, and strength of the bump, respectively. We assume fixed  $\lambda_0 = 2175$  and  $\Delta\lambda = 350$ . 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. In our EDA model, 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 vary does *not* impact our results. In Table 1, we list and describe all of the free parameters in 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  $\text{SFR}_{\text{min}}$ , the minimum non-zero SFR in the simulations, to  $\text{SFR} = 0$  galaxies when calculating  $\tau_V$  and  $\delta$ . For SIMBA, TNG, and EAGLE,  $\text{SFR}_{\text{min}} = 0.000816, 0.000268$ , and  $0.000707 M_\odot/\text{yr}$ . Although this assumes that  $\text{SFR} = 0$  galaxies have similar dust properties as the galaxies with  $\text{SFR} = \text{SFR}_{\text{min}}$ , since the simulations have very low  $\text{SFR}_{\text{min}}$  we expect galaxies with  $\text{SFR} = \text{SFR}_{\text{min}}$  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 begin by assigning a uniformly sample inclination,  $i$ , to each galaxy. Then based on the EDA parameters,  $\tau_V$ , and  $\delta$  are calculated for the galaxy using its  $i$ ,  $M_*$ , and SFR. From  $\tau_V$  and  $\delta$ , we determine  $A_V$  and  $k(\lambda)$ , which together gives  $A(\lambda)$ . 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  $M_*$ : star-forming ( $\text{SFR} = 10^{0.5} M_\odot/\text{yr}$ ) with low mass ( $10^{10} M_\odot$ ; blue), with high mass ( $10^{11} M_\odot$ ; green) and quiescent ( $\text{SFR} = 10^{-2} M_\odot/\text{yr}$ ) with low mass ( $10^{10} M_\odot$ ; orange), with high mass ( $10^{11} M_\odot$ ; red). All galaxies are edge-on (*i.e.*  $i = 0$ ) and we use EDA parameters:  $\{m_{\tau, M_*}, m_{\tau, \text{SSFR}}, c_\tau, m_{\delta, M_*}, m_{\delta, \text{SFR}}, c_\delta\} = \{2., -2., 2., -0.1, -0.1, -0.2\}$ , arbitrarily chosen within the prior range listed in Table 1. For comparison, we include the Calzetti (2001) attenuation curve. The EDA model provides a flexible prescription for applying dust attenuation with a wide range of amplitude and slope to galaxies based on their physical properties.



**Figure 3.** Attenuation curves,  $A(\lambda)$ , generated by the Empirical Dust Attenuation (EDA) model for galaxies with different SFR and  $M_*$ . We include  $A(\lambda)$  for star-forming (SFR =  $10^{0.5} M_\odot/\text{yr}$ ) low mass galaxy ( $10^{10} M_\odot$ ; blue), high mass galaxy ( $10^{11} M_\odot$ ; green) and quiescent (SFR =  $10^{-2} M_\odot/\text{yr}$ ) low mass galaxy ( $10^{10} M_\odot$ ; orange), high mass galaxy ( $10^{11} M_\odot$ ; red). All galaxies are edge-on (*i.e.*  $i = 0$ ) and we use EDA parameter values near the center of our prior range (Table 1):  $\{m_{\tau, M_*}, m_{\tau, \text{SSFR}}, c_\tau, m_{\delta, M_*}, m_{\delta, \text{SFR}}, c_\delta\} = \{2., -2., 2., -0.1, -0.1, -0.2\}$ . For comparison, we include the Calzetti (2001) attenuation curve. In the EDA, the amplitude and slope of  $A(\lambda)$  depend on  $M_*$ , and SSFR (Eqs. 4 and 6). *The EDA provides a flexible model for applying dust attenuation to galaxies based on their physical properties.*

**Table 1.** Parameters of the Dust Empirical Model

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

#### 4. SIMULATION-BASED INFERENCE: APPROXIMATE BAYESIAN COMPUTATION

With a forward modeling approach and the EDA model, we can generate observables for simulated galaxies and make an “apples-to-apples” comparison to SDSS observations (Section 2). For the comparison, which we can use to infer EDA parameters, 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). ABC is a simulation-based (or “likelihood-free”) 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, which can significantly bias the inferred posterior (Hahn et al. 2019b). It also enables us to estimate the posterior using observables with difficult or intractable likelihoods (Hahn et al. 2017a).

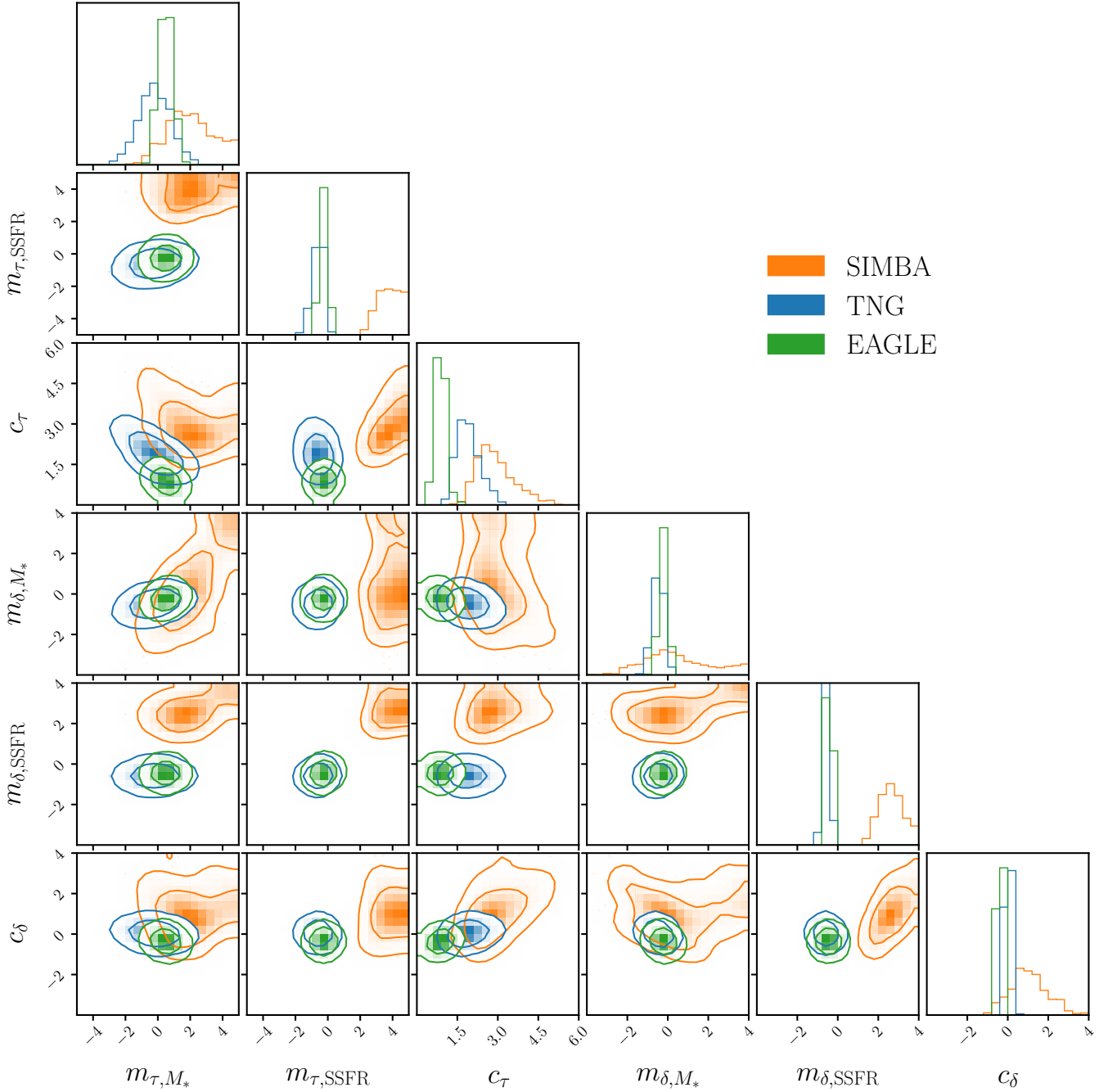
In the simplest version of ABC, with a rejection sample framework (Pritchard et al. 1999), a proposal set of parameter values are drawn from the prior. The forward model is run with the proposal parameter values. The output of the forward model is then compared to the observed data using the distance metric. If the distance is within some small distance threshold, we keep the proposed parameters; otherwise, we discard them. Proposals are drawn until enough of them 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 in conjunction with Population Monte Carlo (PMC) importance sampling (Hahn et al. 2017a,b, 2019a).

The forward model in our scenario starts with the galaxies from our hydrodynamic simulations. Then we construct SEDs using their SFH and ZH, apply dust attenuation with the EDA model, and calculate the observables from the attenuated SED. For each hydro simulation and a set of EDA parameters, our forward model produces  $G$ ,  $R$ ,  $NUV$ , and  $FUV$  absolute magnitudes. We use uninformative uniform priors on each of the EDA parameters and choose ranges to encompass constraints in the literature. The prior ranges of  $m_{\tau, M_*}$ ,  $m_{\tau, \text{SFR}}$ ,  $c_\tau$  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. We note that uniform priors on the EDA parameters do not necessarily produce uniform priors on  $\tau_V$  or  $\delta$  (*e.g.* Handley & Millea 2019). However, we are interested in marginalizing over dust attenuation and understanding the dependence of dust attenuation on galaxy properties, so we use uninformative priors on the EDA parameters and not on  $\tau_V$  or  $\delta$ .

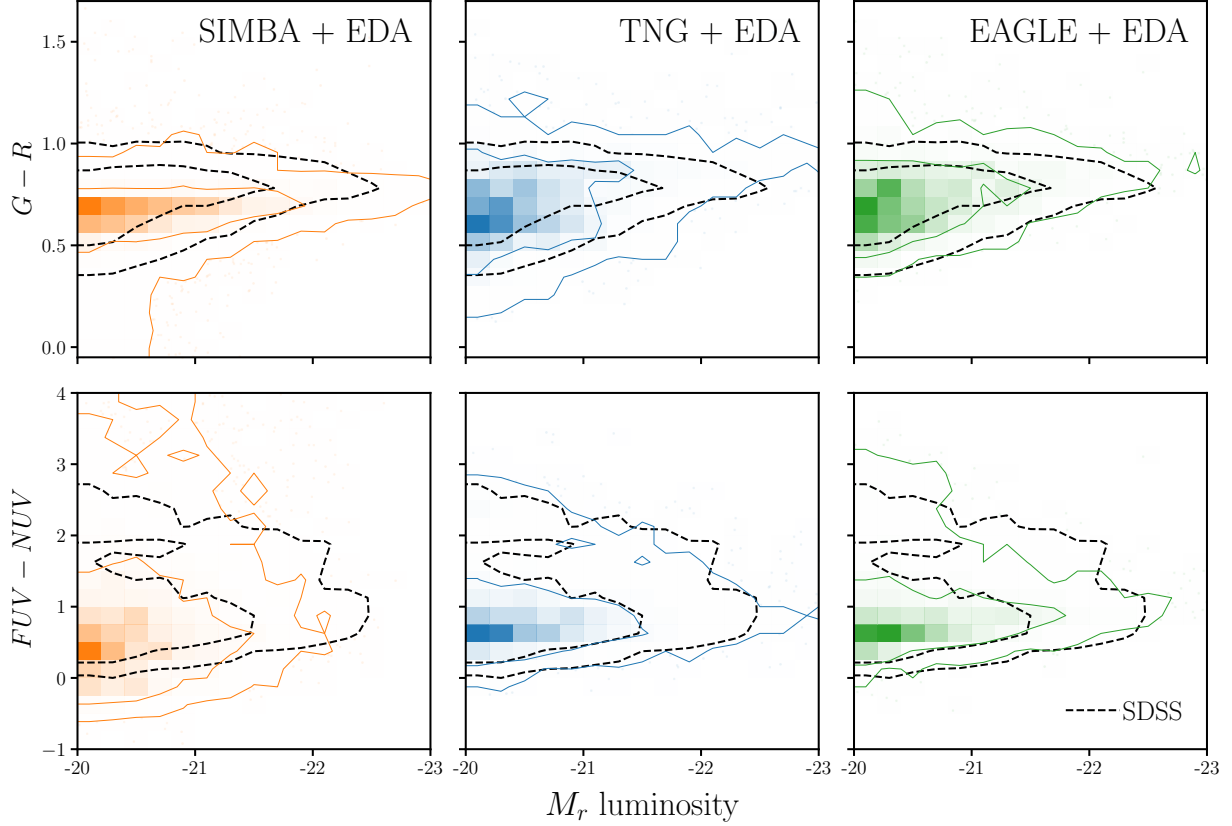
ABC also requires a distance metric that quantifies the “closeness” of the forward model output to the observed data. For our distance metric, we use the L2 norm between the summary statistics,  $X$ , of the SDSS observation and our forward model:

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

$\theta_{\text{EDA}}$  are the EDA parameters. The summary statistics are based on the optical and UV color-magnitude relations,  $(G-R) - R$  and  $(FUV-NUV) - R$ , of galaxies brighter than  $M_r < -20$ , the completeness limit of our SDSS sample (Figure 2). More specifically, for  $X$ , we calculate the number density in 3D bins of  $G-R$ ,  $FUV-NUV$ , and  $M_r$  with widths 0.0625, 0.25, and 0.5 mags. We choose this summary statistic to fully exploit the observable-space predicted by the forward model. Later in Section 5 we discuss other observables that could be used in the analysis.



**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% confidence intervals. The posteriors are derived using the simulation-based inference method: Approximate Bayesian Computation with Population Monte Carlo (Section 4). We focus on the EDA posteriors for TNG and EAGLE since SIMBA requires dust attenuation to reverse the relationship between color and SFR in order to account for it overpredicting low mass starburst galaxies. Based on the posteriors, we find that *galaxies with higher  $M_*$  have overall higher dust attenuation and galaxies with lower SFR have overall higher dust attenuation.*



**Figure 5.** The optical and UV color-magnitude relations predicted by the DEM with the median ABC posteriors for the SIMBA (orange), TNG (blue), and EAGLE (green) hydrodynamical simulations. For comparison, we include the  $(G-R)-M_r$  (top panels) and  $(FUV-NUV)-M_r$  (bottom panels) relations for SDSS (black dashed). With the DEM, the simulations produce dramatically different observables than without any dust prescription (Figure 2). Hence, dust must account for when interpreting and comparing simulations. Moreover, with the DEMs, all three simulations produce color-magnitude relations consistent with SDSS. Since different simulations can produce reproduce observations by varying dust, dust significantly limits our ability to constrain the physical processes that go into galaxy simulations.

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 ( $i > 1$ ), the threshold,  $\epsilon_i$ , 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  $\epsilon_i$ . This is repeated until we assemble a new set of  $N$  proposals  $\bar{\theta}_i$ . The entire process is repeated for the next iteration until convergence is confirmed. For further details on the ABC-PMC implementation, we refer readers to [Hahn et al. \(2017b\)](#) and [Hahn et al. \(2019a\)](#). In Figure 4, we present the posterior distributions of the EDA parameters derived from ABC-PMC for the SIMBA (orange), TNG (blue), and EAGLE (green) hydrodynamical simulations. The contours mark the 68% and 95% confidence intervals.

## 5. RESULTS

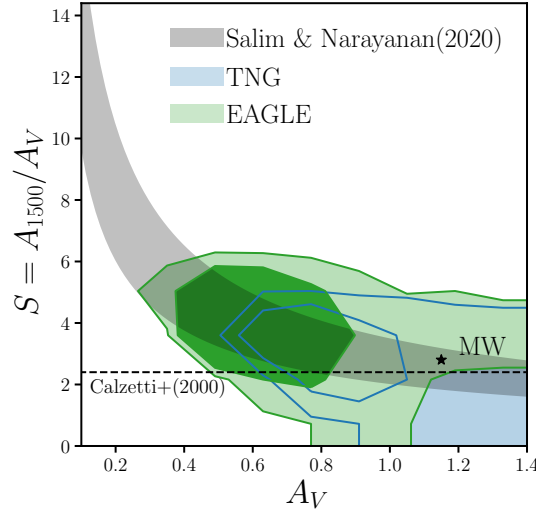
Without dust attenuation, all of the hydrodynamical simulations struggle to reproduce the  $(G-R)-M_r$  and  $(FUV-NUV)-M_r$  relations of SDSS (Figure 2). In optical, the simulations predict significantly bluer colors and find broader differences in color,  $\sim 0.5\ mag$ , between star-forming and quiescent galaxies. In the UV, they predict a subpopulation of galaxies that have red  $FUV-NUV$  colors,  $FUV-NUV \sim 3\ mag$ , beyond the SDSS observations. These are quiescent galaxies with particularly low SSFR,  $\log SSFR < -12\ yr^{-1}$ . For the rest of the galaxies, the simulations predict significantly bluer UV colors than SDSS.

On the other hand, *with the EDA model all of the simulations are able to reproduce the color-magnitude relations of SDSS*. 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 posterior distributions inferred using ABC (Figure 4). The contours mark the 68% and 95% confidence intervals 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 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  $\sim 0.5\ mag$  redder. Second, the EDA significantly attenuates, by  $\sim 0.5\ mag$ , intrinsically blue galaxies (*i.e.* star-forming galaxies with  $\log SSFR > -10.5$ ). 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,  $FUV-NUV \sim 3$ , by  $\sim 1\ mag$ . Thus, unlike in Figure 2, the UV color-magnitude relations from the EDA does not produce galaxies with high  $FUV-NUV$  and is in good agreement with SDSS.

For SIMBA, the EDA substantially improves the agreement with observations; however, there are still some significant discrepancies in the color-magnitude relations with SDSS. SIMBA produces a  $G-R$  and  $FUV-NUV$  color distributions that are respectively narrower and broader than SDSS. Also, the inferred EDA parameters for SIMBA differ significantly from the parameters of TNG and EAGLE (Figure 4). These discrepancies are primarily driven by the large population of high  $\log SSFR > -9.5$  “starburst” galaxies that lie above the SFS, which are *only* present in SIMBA (Figure 1). Without dust attenuation, these starburst galaxies are blue  $G-R \sim 0.1$  and have high luminosity,  $M_r < -22$ . Such high luminosity blue galaxies are not present in observations, so these SIMBA galaxies need to be both strongly attenuated and reddened. This would, however, contradict the attenuation-slope relation established in both observations and simulations where galaxies with higher attenuation have shallower slopes (Inoue 2005; Chevallard et al. 2013; Salim et al. 2018; Salim & Narayanan 2020; Trayford et al. 2020). To further examine the impact of the starburst galaxies in SIMBA, we try running the EDA for SIMBA excluding galaxies with  $\log SSFR > -9.5$  and with median parameter values of the TNG and EAGLE posteriors. Without starburst galaxies, SIMBA is able to produce color-magnitude relations in good agreement with observations. *Hence, we conclude that the excess starburst population in SIMBA is in tension with observations.* For the rest of the paper, we focus on the TNG and EAGLE simulations.





**Figure 6.** The attenuation-slope relation of the attenuation curves predicted by the EDA model for median posterior parameter values of TNG (blue) and EAGLE (green) simulations. We use  $A_V$  and  $S = A(3000)/A_V$  for attenuation and slope, respectively. 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 derived the posteriors of the EDA model from comparing the UV and optical color-magnitude relation and do not fit any observed dust attenuation measurements. Yet, using a forward modeling approach, we find excellent agreement between the attenuation-slope relation predicted by the EDA and observations.

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 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 predicts  $FUV-NUV$  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 with the EDA. The EDA is a simpler prescription 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.

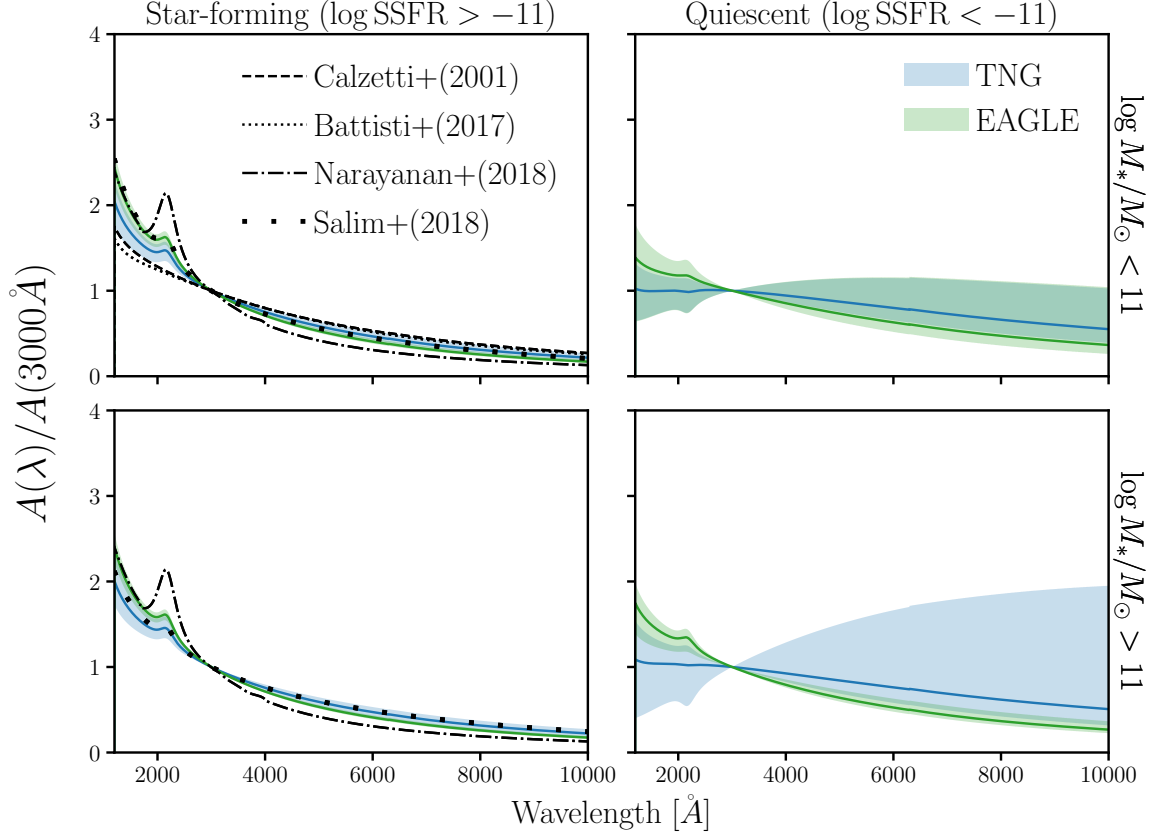
### 5.1. Reproducing Dust Observations

We demonstrated above that we can use the EDA model in a forward model to accurately reproduce the observed color-magnitude relations. In the forward model, the EDA assigns dust attenuation curves to each simulated galaxy. This means that we can compare the dust attenuation assigned to the simulated galaxies (*i.e.* predicted by the EDA) to the dust attenuation measured from observed galaxies. In Figure 6, we present the attenuation-slope relation of the EDA dust attenuation curves for the median posterior parameter values of TNG (blue) and EAGLE (green) simulations. For attenuation, we use  $A_V$  and for slope we use the UV-optical slope,  $S = A(3000)/A_V$ , commonly found in the literature. To compare with observations, we include the attenuation-slope relations of GSWLC2 galaxies (grey; Salim & Narayanan 2020), the Milky Way (star), and mark the slope of the Calzetti (2001) curve (dashed).

*The attenuation-slope relation predicted by the EDA is in excellent agreement with observations.* In both the EDA and observations, galaxies with higher attenuation have shallower slopes. The EDA attenuation curves do not span the entire range of the observations due to the stellar mass ( $M_* \gtrsim 10^{10} M_\odot$ ) and luminosity limit ( $M_r < -20$ ) of our sample. Meanwhile, the Salim & Narayanan (2020) GSWLC2 sample extends down to  $M_* \sim 10^{8.5} M_\odot$  where they find star-forming galaxies have lower attenuations. The attenuation-slope relation is a consequence of the fact that dust scattering dominates absorption at low attenuation while dust absorption dominates at high attenuation (Gordon et al. 1994; Witt & Gordon 2000; Draine 2003; Chevallard et al. 2013). The constraints on the EDA parameters were derived solely from comparing the the UV and optical color-magnitude relations. This underscores the advantages of our forward modeling approach — without fitting to any dust measurements, we are able to reproduce the observed properties of dust using the EDA.

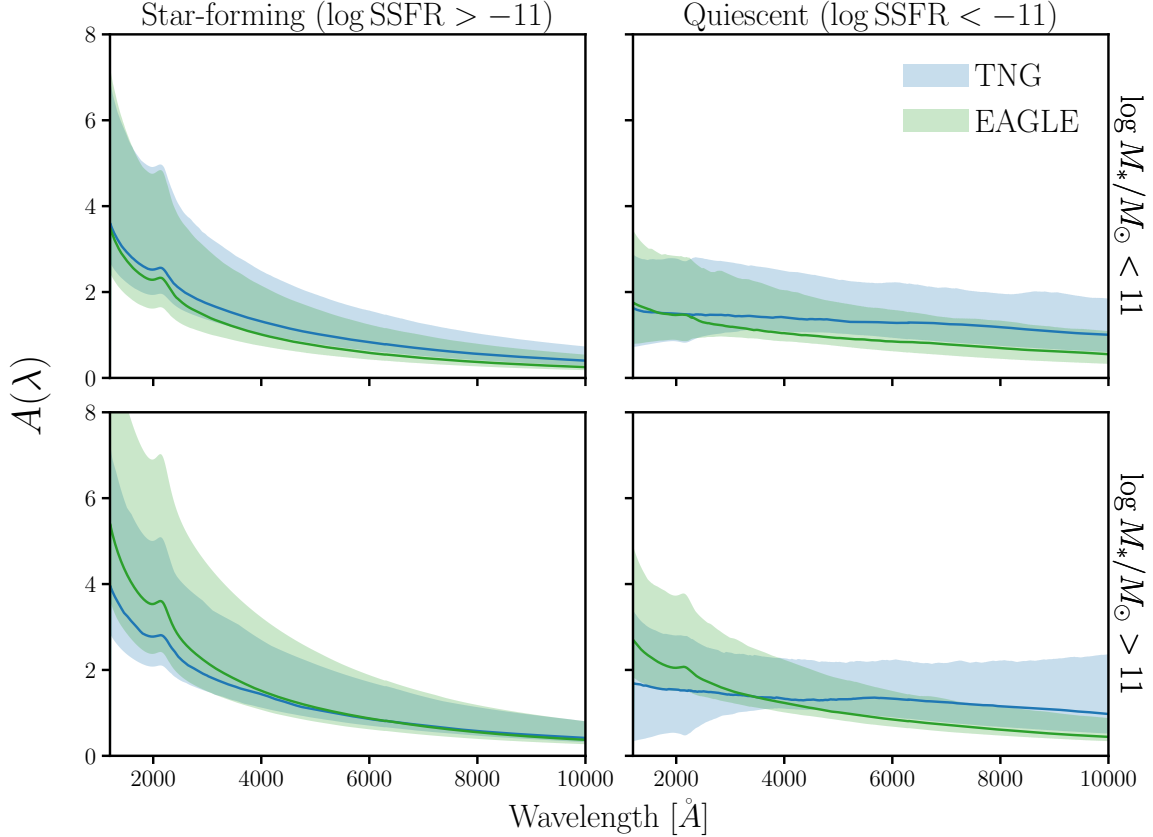
We can more closely compare the attenuation curves predicted by the EDA to observations. In the left panels of Figure 7, we present the attenuation curves of star-forming galaxies predicted by the EDA model for the median posterior parameter values of TNG (blue) and EAGLE (green). In the top panel, we present galaxies with  $M_* < 10^{11} M_\odot$ ; in the bottom we present galaxies with  $M_* > 10^{11} M_\odot$ . We define galaxies with  $\log \text{SSFR} > -11$  as star-forming. The attenuation curves are normalized at 3000 and we present the variation in the attenuation curves in the shaded region,  $1\sigma$  standard deviation about the median. For comparison, we include  $A(\lambda)/A(3000)$  from observations: Calzetti et al. (2000), Battisti et al. (2017), and Salim et al. (2018). We also include  $A(\lambda)/A(3000)$  measured from the Narayanan et al. (2018) simulation. The left panels of Figure 8 is the same as in Figure 7, except with attenuation curves that are not normalized. The normalized attenuation curves in Figure 7 emphasize the slope while the curves in Figure 8 emphasize the amplitude.

Based on the EDA, star-forming galaxies have higher dust attenuation at lower wavelengths. This is driven by both TNG and EAGLE predicting star-forming galaxies that are bluer than observations in the optical and UV wavelengths (Figure 2). As a result, the EDA assigns attenuation curves that significantly redden star-forming galaxies. This is also why TNG, which predicts a bluer star-forming population, has slightly higher attenuation. In Figure 8, we also find that more massive star-forming galaxies have higher attenuation. This is because the simulations overpredict luminous blue star-forming galaxies, which must be attenuated to reproduce observations.



**Figure 7.** Attenuation curves predicted by the EDA for median posterior parameter values of the TNG (blue) and EAGLE (green) simulations. In the left and right panels, we present star-forming and quiescent galaxies classified using a  $\log \text{SSFR} = -11$  cut. In the top and bottom panels, we present galaxies with  $M_* < 10^{11} M_\odot$  and  $M_* > 10^{11} M_\odot$ , respectively. The attenuation curves are normalized at 3000 and we mark the  $1\sigma$  standard deviation of the attenuation curves with the shaded region. For comparison, we include  $A(\lambda)/A(3000)$  measurements from observations (Calzetti et al. 2000; Battisti et al. 2017; Salim et al. 2018) and simulations (Narayanan et al. 2018). For star-forming galaxies, the Calzetti et al. (2000) and Battisti et al. (2017) attenuation curves are shallower than the EDA attenuation curves, but they probe less massive galaxies than our sample. For Salim et al. (2018), which probe a similar  $M_*$  range, we find good agreement. We also find good agreement with median attenuation curve of star-forming galaxies in Narayanan et al. (2018). With EDA, we can also constrain the attenuation curves of quiescent galaxies, which are challenging to directly measure in observations. Quiescent galaxies have significantly shallower attenuation curves than star-forming galaxies.

The EDA attenuation curves for both TNG and EAGLE are in good agreement with the observed Salim et al. (2018) attenuation curves. They are also consistent with the median curve of Narayanan et al. (2018). We ignore the discrepancies between the EDA and Narayanan et al. (2018) attenuation curves in the amplitude of UV dust bump, since we keep the UV bump fixed in the EDA. The EDA attenuation curves are noticeably steeper than the Calzetti et al. (2000) and Battisti et al. (2017) curves. These attenuation curves, however, are derived from  $M_* < 10^{9.9} M_\odot$  star-forming galaxies — well below our  $M_*$  range. Overall, the EDA predicts attenuation curves that are in good agreement



**Figure 8.** Same as Figure 7 except the attenuation curves are not normalized at 3000. Based on the EDA, quiescent galaxies have significant UV and optical dust attenuation. Furthermore, they have significantly lower attenuation in the UV compared to star-forming galaxies and have shallow attenuation curves overall.

with both observations and radiative transfer simulations for star-forming galaxies. Again, the fact that we reproduce the detailed dust attenuation curves of star-forming galaxies in observations and simulations with the EDA, without fitting for them, highlights the advantages of a forward modeling approach.

### 5.2. The Attenuation Curves of Quiescent Galaxies

We have demonstrated so far that the EDA is able reproduce the observed UV and optical color-magnitude relations and also predict dust attenuation curves of star-forming galaxies consistent with observations and radiative transfer simulations. In addition, the EDA also predicts dust attenuation curves of quiescent galaxies. This is particularly valuable since there are many challenges to measuring attenuation curves for 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. \(2015\)](#). Even SED fitting methods require accounting for AGN MIR emission ([Salim et al. 2016](#); [Leja et al. 2018](#); [Salim et al. 2018](#)). SED fitting methods also struggle to tightly constrain dust attenuation for quiescent galaxies since they are limited by the degeneracies with star formation history and metallicity.

With a forward modeling approach, we circumvent these challenges. We derive the attenuation curves necessary for quiescent galaxy population in simulations to reproduce the observed optical and UV photometry. In right panels of Figure 7, we present the attenuation curves of quiescent galaxies predicted by the EDA model for the median posterior parameter values of TNG (blue) and EAGLE (green). In the top and bottom panels, we present galaxies with  $M_* < 10^{11} M_\odot$  and  $M_* > 10^{11} M_\odot$ , respectively. We define galaxies with  $\log \text{SSFR} < -11$  as quiescent. The right panels of Figure 8 are the same as in Figure 7, except the attenuation curves in Figure 8 are not normalized.

For both TNG and EAGLE posteriors, the EDA predicts significant dust attenuation in quiescent galaxies. Compared to star-forming galaxies, however, they have lower attenuation in the UV and much shallower attenuation curves. The amplitude of the EDA quiescent galaxy attenuation curves is driven by the fact that both TNG and EAGLE — without dust — predict quiescent galaxies that are too luminous compared to observations. Hence, significant attenuation is necessary to lower their luminosity. Meanwhile, the shallow slope is driven by the simulations predicting quiescent galaxies that are bluer in the optical but redder in the UV than observations. The EDA optically reddens the quiescent galaxies but maintains a shallow enough slope to reproduce the UV color-magnitude relation. This is also why TNG has a shallower slope than EAGLE: TNG has an optically redder quiescent population and more quiescent galaxies with high  $FUV-NUV$  color.

Given the challenges in observationally measuring attenuation curves of quiescent galaxies, the predictions of the bestfit EDA models for TNG and EAGLE highlight the advantages of a forward modeling approach and provide valuable insights into dust attenuation in quiescent galaxies. *Quiescent galaxies have significant UV and optical attenuation with shallow attenuation curves.*

### 5.3. The Galaxy – Dust Connection

With the EDA model, simulations can reproduce the observed color-magnitude relations. The bestfit EDA models predict attenuation curves that are in good agreement with observations — the observed attenuation–slope relation and the observed attenuation curves of star-forming galaxies. They also predict the attenuation curves of quiescent galaxies, which are poorly constrained by observations. Furthermore, since the EDA models predict dust attenuation for all galaxies in the sample, they allow us to shed light on the connection between galaxy properties and dust attenuation. More specifically, the EDA parameters constraints (Figure 4) and the predicted attenuation curves (Figures 7 and 8) reveal the stellar mass and SSFR dependence of dust attenuation.

Focusing first on the  $M_*$  dependence, we find that TNG has little  $M_*$  dependence in  $\tau_V$  ( $m_{\tau, M_*} = -0.15^{+0.83}_{-0.91}$ ) while EAGLE has a significant  $M_*$  dependence ( $m_{\tau, M_*} = 0.55^{+0.40}_{-0.37}$ ). Meanwhile, we find significant  $M_*$  dependence in  $\delta$  for both TNG ( $m_{\delta, M_*} = -0.44^{+0.27}_{-0.26}$ ) and EAGLE ( $m_{\delta, M_*} = -0.20^{+0.17}_{-0.17}$ ) — *i.e.* more massive galaxies have steeper slopes (Figure 4). The  $M_*$  dependence in  $\tau_V$  and  $\delta$  are further illustrated in the attenuation curves in Figure 8. For TNG, there is little difference in the  $V$ -band attenuation between the top and bottom panels. However, more massive galaxies have steeper slopes with similar  $A_V$  so they have significantly higher UV attenuation. For EAGLE, more massive galaxies have both higher  $V$ -band and UV attenuation. Overall, the EDA predicts that more massive galaxies have higher attenuation, which is 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.

Next, we examine the SSFR dependence of dust attenuation. For both TNG and EAGLE, we find a significantly negative SSFR dependence in  $\tau_V$  ( $m_{\tau, \text{SSFR}} = -0.56^{+0.36}_{-0.30}$  and  $-0.24^{+0.21}_{-0.20}$ , respectively) as well as  $\delta$  ( $m_{\delta, \text{SSFR}} = -0.57^{+0.14}_{-0.112}$  and  $-0.43^{+0.10}_{-0.09}$ , respectively). In other words, galaxies with higher SSFR have lower  $V$ -band attenuation and steeper attenuation curves. From the attenuation curves in Figure 8, we similarly find that star-forming galaxies have attenuation curves with slightly lower  $A_V$  but substantially higher UV attenuation. Quiescent galaxies, on the other hand, have significantly shallower attenuation curves. Although observations have examined the SSFR dependence of dust attenuation, they cannot be compared to the predicts of our EDA models since they focus on star-forming galaxies, due to the difficulty in observationally constraining the attenuation curve in quiescent galaxies (*e.g.* Garn & Best 2010; Reddy et al. 2015; Battisti et al. 2016, 2017; Salim et al. 2018). In summary, from the bestfit EDA models of TNG and EAGLE, we find that *galaxies with higher  $M_*$  have overall higher dust attenuation and galaxies with higher SSFR have steeper attenuation curves.*

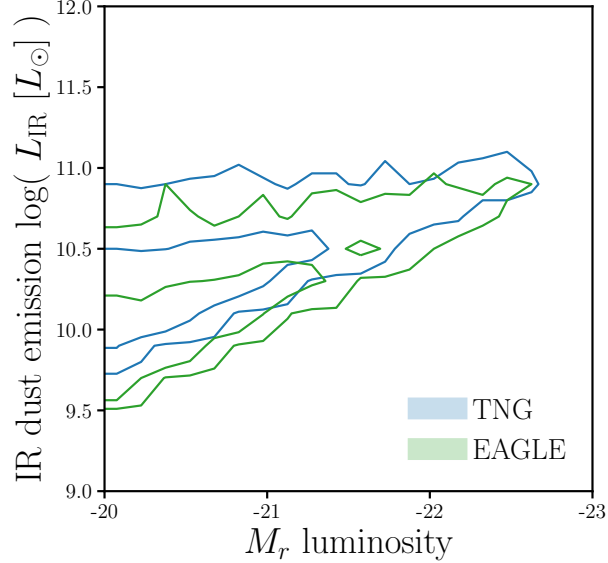
#### 5.4. Discussion

In the EDA we make a number of assumptions and choices. For instance, EDA assigns  $A_V$  using the slab model (Eq. 3). We use the slab model because it reproduces the correlation between attenuation and inclination in observations (Conroy 2010; Wild et al. 2011; Battisti et al. 2017; Salim & Narayanan 2020) and simulations (*e.g.* Chevallard et al. 2013; Narayanan et al. 2018; Trayford et al. 2020). It also reproduces the SDSS  $A_V$  distribution (Figure 10). If we replace the slab model with a more flexible model for sampling  $A_V$  using truncated normal distributions, we find that our results are not significantly impacted (see Appendix B for details). Therefore, we conclude that our results do not significantly depend on our choice of the slab model.

Besides the slab model, we use a simple parameterization of  $\tau_V$  and  $\delta$  in the EDA. Both  $\tau_V$  and  $\delta$  depend linearly on  $\log M_*$  and  $\log \text{SSFR}$ . This linear parameterization was choice for solely for its simplicity; the EDA can easily be 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 parameterization. Nevertheless — for SDSS observations — the EDA using parameter values from the TNG and EAGLE posteriors find good agreement.

The fact that we can use the EDA to reproduce SDSS observations for different hydrodynamical simulations highlight two key points. First, it demonstrates 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). Second, it also highlights limitation in our current understanding of dust. The EDA is built on our current understanding





**Figure 9.** IR dust emission 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  $\tau$  energy balance. Despite reproducing the same SDSS UV and optical color-magnitude relations, *the bestfit EDA models for TNG and EAGLE predict significantly different IR dust emission*. Therefore, including IR observations will significantly improve the constraints on EDA parameters and allow us to better differentiate galaxy formation models.

of dust attenuation of 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.

Fortunately, there are many avenues for improving our understanding of dust using a forward modeling approach. In this work, we used a restrictive  $M_r < -20$  complete SDSS galaxy sample. Instead of imposing a completeness limit, we can include the actual SDSS selection function in the forward model (*e.g.* [?](#) ). This would allow us to compare the simulations with EDA to a substantially larger galaxy sample. Upcoming surveys, such as the Bright Galaxy Survey (BGS) of the Dark Energy Spectroscopic Instrument (DESI; [DESI~Collaboration et al. 2016](#); [?](#) ) 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  $z \sim 0.4$ . Comparing to a larger more statistically powerful observation sample will place tighter constraints on EDA parameters and enable an actual comparison of the simulations.

Also, in this work, we only used observables derived from UV and optical photometry. This only examines one side of the impact of dust 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 9, we present IR dust emission,  $L_{\text{IR}}$ , 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 (?).

Despite reproducing the same SDSS UV and optical color-magnitude relations, the bestfit EDA models for TNG and EAGLE predict significantly different IR dust emission. The bestfit EDA model for TNG 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. A number of upcoming observations (*e.g.* @tjitske observations you mentioned) will probe the IR. BGS galaxies will also have IR photometry from NEOWISE (?).

TODO

## 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 a flexible method for sampling the attenuation curve parameters that includes correlations with galaxy 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 simulation-based 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 starburst galaxy population. In order to reproduce observations, these starburst galaxies would require both high attenuation and reddening, which goes against the observed attenuation-slope relation.
- The attenuation curves predicted by the bestfit 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, highlights the advantages of a forward modeling approach.

- Lastly, the EDA sheds light on dust attenuation in quiescent galaxies, which remains poorly understood due to observations challenges. From the EDA, we find that quiescent galaxies have significant UV and optical attenuation with shallow attenuation curves. Moreover, since the EDA predict dust attenuation for all galaxies, it reveals the connection between galaxy properties and dust attenuation. 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 can provide key insights into dust attenuation. For those uninterested in dust, the EDA also provides a framework for marginalizing over dust when comparing simulations to observations. In the case of SIMBA, we found that dust attenuation is insufficient to accurately reproduce observations due to its starburst population. For TNG and EAGLE, however, dust attenuation is highly degenerate with the differences they have 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 — differences between simulations may not produce any change in the observables once dust attenuation is taken into account. Therefore, our current understanding of dust, or lack of understanding, 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 relatively 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 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 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 Daniel Kelson, Mariska Kriek, Desika Narayanan, Katherine Suess. 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.

## APPENDIX

## A. RESOLUTION EFFECTS

Figure demonstrating imprint SFR=0 leave on the observable space and how we deal with them so we can ignore them...

## B. BEYOND THE SLAB DEM

A major assumption of our fiducial DEM is that we sample the amplitude of attenuation from the slab model. The slab model makes the simplifying assumption that dust in galaxies are in a slab-like geometry and illuminated by the stellar radiation source (Somerville & Primack 1999). Then, for a given  $\tau_V$ , the attenuation depends solely on the orientation of the galaxy. This simplification, ignores any complexities in the star-to-dust geometry that impact the shape of the attenuation curve (Witt Gordon 1996, 2000, Seon Drain 2016). TODO

Besides its simplifications, the slab model predicts  $A_V$  distribution with significant differences than the  $A_V$  distributions measured from observations. In Figure 10, we compare the  $A_V$  distribution predicted by the slab model (black) to the  $A_V$  distribution of star-forming galaxies in our SDSS sample (blue). The  $A_V$  values are derived using SED fitting from the Brinchmann et al. (2004) MPA-JHU catalog and how are the SF galaxies classified. The slab model  $A_V$  values are derived using Eq. 3 and 4 with  $M_*$ s and SFRs from the same SDSS sample and the inclinations,  $i$ , are uniformly sampled over the range  $[0, \pi/2]$ . With  $\{m_{\tau,1}, m_{\tau,2}, c_{\tau}\}$  chosen to reproduce the observed  $A_V$  distribution, the slab model can reproduce the overall shape. However, it predicts an extended high  $A_V$  tail not found in observations. TODO

Given these shortcomings of the slab model, we want to ensure that our results do not hinge on the slab model. Modeling the star-to-dust geometries with increased complexities, however, would involve expensive hydrodynamic simulations and dust radiative transfer calculations (e.g. Narayanan et al. 2018)jonsson2006, rocha2008, natale2015,hayward smith2015,hou2017,trayford2020. We instead take an empirical approach and implement a flexible model for sampling  $A_V$  based on a truncated normal distribution: TODO

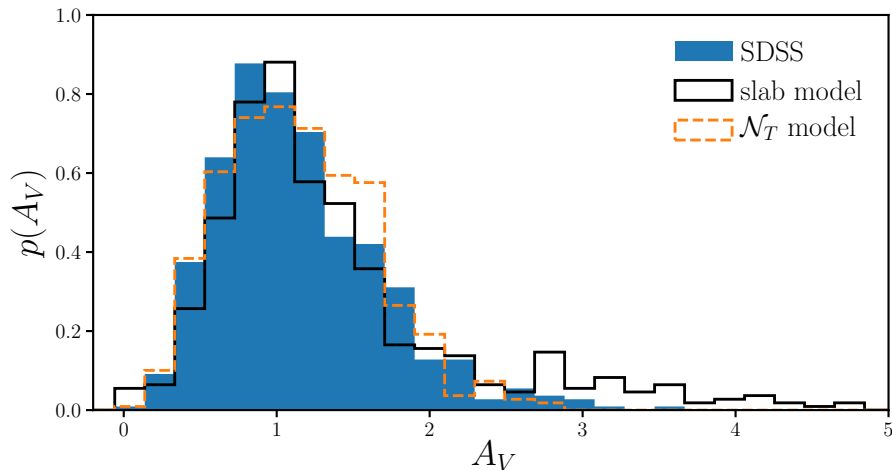
$$A_V \sim \mathcal{N}_T(\mu_{A_V}, \sigma_{A_V}) = \frac{\mathcal{N}(\mu_{A_V}, \sigma_{A_V})}{1 - \Phi\left(-\frac{\mu_{A_V}}{\sigma_{A_V}}\right)}. \quad (\text{B1})$$

Here,  $\mathcal{N}$  is the standard normal distribution and  $\Phi(x) = \frac{1}{2} (1 + \text{erf}(x/\sqrt{2}))$  is the cumulative distribution function of  $\mathcal{N}$ .  $\mu_{A_V}$  and  $\sigma_{A_V}$  are the mean and variance of the truncated normal distribution. Similar to Eq. 4, we allow  $\mu_{A_V}$  and  $\sigma_{A_V}$  to depend on the physical properties of galaxies:

$$\mu_{A_V} = m_{\mu,1}(\log M_* - 10.) + m_{\mu,2} \log \text{SFR} + c_{\mu} \quad (\text{B2})$$

$$\sigma_{A_V} = m_{\sigma,1}(\log M_* - 10.) + m_{\sigma,2} \log \text{SFR} + c_{\sigma}. \quad (\text{B3})$$

The  $A_V$  distribution from our truncated normal (orange dashed) closely reproduces the observed SDSS  $A_V$  distribution (Figure 5).  $N_T$  is able to reproduce the overall skewness but unlike the slab



**Figure 10.** Comparison of  $A_V$  distribution of SDSS star-forming galaxies (blue) to predictions from the slab model (Eq. 3; black). **detail on how SDSS SF galaxies are classified.** The slab model assumes that there’s a slab of dust in front of a galaxy. We use  $\tau_V = 2$  for the slab model above. Regardless of  $\tau_V$ , however, the slab model predicts a significantly more asymmetric and peaked  $A_V$  distribution than observations. Given this disagreement, *we include in our analysis a DEM with an empirical prescription for  $A_V$  based on a truncated normal distribution, which better reproduce the observed  $A_V$  distribution (Section B).*

model, it does not have a long high  $A_V$  tail. With more free parameters and a functional form that closely resembles the observed  $A_V$  distribution, the truncated normal model provides a flexible alternative to the slab model and we include it in our analysis.

## 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
- 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

- Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, *The Astrophysical Journal*, 533, 682
- 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
- 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
- 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
- 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
- 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
- Handley W., Millea M., 2019, *Entropy*, 21, 272
- 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
- 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
- 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
- Lin C.-A., Kilbinger M., Pires S., 2016, [Astronomy and Astrophysics](#), 593, A88
- McAlpine S., et al., 2016, [Astronomy and Computing](#), 15, 72
- 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
- Pillepich A., et al., 2018, [Monthly Notices of the Royal Astronomical Society](#), 473, 4077
- 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
- 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
- Salmon B., et al., 2016, [The Astrophysical Journal](#), 827, 20
- 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
- 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
- Trčka A., et al., 2020, arXiv:2003.12576 [astro-ph]
- 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
- Walcher J., Groves B., Budavári T., Dale D., 2011, [Astrophysics and Space Science](#), 331, 1
- 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