

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

CHANGHOON HAHN,<sup>1,2,3,\*</sup> TJITSKE K. STARKENBURG,<sup>4</sup> AND IQ COLLABORATORY

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

<sup>2</sup>*Lawrence Berkeley National Laboratory, 1 Cyclotron Rd, Berkeley CA 94720, USA*

<sup>3</sup>*Berkeley Center for Cosmological Physics, University of California, Berkeley CA 94720, USA*

<sup>4</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 --- 92da70c --- 2020-10-09 --- NOT READY FOR DISTRIBUTION)

## ABSTRACT

We present the Empirical Dust Attenuation (EDA) models, an empirical framework for applying dust attenuation to simulated galaxy populations. The EDA uses a parameterization of the attenuation curve based on observational constraints and a statistical prescription for sampling 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. Then we forward model the spectral energy distributions and optical and UV color-magnitude relations,  $(G-R) - M_r$  and  $(FUV-NUV) - M_r$ , for the simulated galaxies and compare them to the central galaxies in SDSS using Approximate Bayesian Computation. For all three simulations, we are able to reproduce SDSS observations with the EDA. We find that SIMBA overpredicts low mass starburst galaxies and requires extreme EDA that reverses the established relation between galaxy color and SFR. Since we can reproduce observations for disparate galaxy populations by varying dust attenuation, we demonstrate that dust is highly degenerate with sub-grid physics in galaxy formation models. In addition to the observed color-magnitude relations, the EDA also reproduces the observed attenuation-slope relation better than radiative transfer models. The predicted attenuation curves for star-forming galaxies are in good agreement with the literature. Furthermore, using the EDA we find correlations between dust attenuation and galaxy properties: more massive galaxies have higher dust attenuation; galaxies with lower SFR have higher dust attenuation **TKS: this seems weird, we need to look into this and rephrase I think**. Finally, we find that quiescent galaxies have shallower attenuation curves and larger variation than star-forming galaxies.

*Keywords:* keyword1 – keyword2 – keyword3

\* hahn.changhoon@gmail.com

## 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. In the infrared (IR), dust produces emissions and from near-infrared (NIR) to ultraviolet (UV), dust modifies the galaxy’s stellar radiation through absorption and scattering. Dust 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, dust inevitably plays a crucial role for galaxy evolution. Beyond providing insights into the physical processes related to dust, better understanding dust is consequential for 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  $2175\text{\AA}$ , in the near-UV (NUV), there is absorption bump referred to as the UV dust bump. Then at longer wavelengths, the curves on 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. Large statistical samples of local galaxies find a wide range of attenuation curves ([Wild et al. 2011](#); [Battisti et al. 2017](#); [Salim et al. 2018](#); [Salim & Narayanan 2020](#)). Observations also reveal a diversity of attenuation curve at high redshifts (*e.g.* [Reddy et al. 2015](#); [Salmon et al. 2016](#)). For a given amount of optical attenuation, the far-UV (FUV) attenuation can be  $\sim 2 - 7\times$  greater.

To understand the origin of this variation in attenuation curves, previous works have examined the correlation between them and galaxy properties. Using 23,000 star-forming galaxies in the Sloan Digital Sky Survey (SDSS), [Wild et al. \(2011\)](#) find 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, find only 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 dependence is caused by underlying trends between the amplitude of the  $V$ -band attenuation ( $A_V$ ) and slope. Based on this “attenuation-slope relation”, galaxies with higher  $A_V$  have shallower slopes. Nevertheless, there is still no clear consensus on the connection between attenuation curves and galaxy properties. Furthermore, most 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

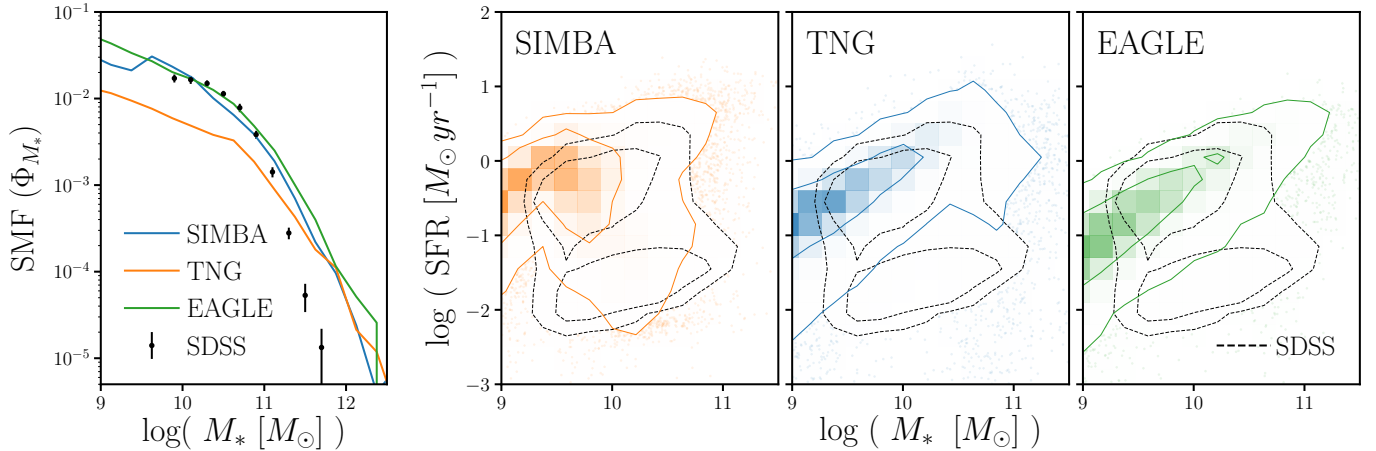
semi-analytic models (SAMs) that do not track baryonic growth directly but make simple physically motivated assumptions about the resulting galaxy properties from dark matter growth (*e.g.* Granato et al. 2000; Fontanot et al. 2009; Wilkins et al. 2012; Gonzalez-Perez et al. 2013). Lastly, radiative transfer models have recently been applied to cosmological hydrodynamical simulations (Camps & Baes 2015; Narayanan et al. 2018; Trayford et al. 2020, *e.g.*). As in observations, these simulations find significant variation in attenuation curves. They also reproduce the shape of the attenuation-slope relation.

Despite their progress, there are still major challenges for radiative transfer models. For instance, computational constraints can result in under-resolved dust and radiation field grids. 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). Furthermore, radiative transfer models produce attenuation-slope relations that are significantly steeper than observations. Finally, radiative transfer models are computationally expensive. So applying multiple radiative transfer dust models to multiple simulations for comparisons is prohibitive. Marginalizing over the impact of dust, for comparisons between simulations and observations aimed at examining subgrid galaxy physics, would be intractable.

In this paper, we take a different approach than these radiative transfer models. We present the Dust Empirical Model (DEM), which provides a framework for statistically assigning dust attenuation to galaxy populations. It uses a parameterization of the attenuation curves motivated from observational studies (*e.g.* Noll et al. 2009) and a flexible method for sampling the attenuation curve parameters (*e.g.* optical depth, slope) that include correlations with galaxy properties. There are a number of advantages to the DEM. Given the DEM parameterization, constraints on the DEM can be directly compared to observed attenuation curves and correlations to galaxy properties can be easily interpreted. The DEM can also produce the wide range of attenuation curves found in both observations and radiative transfer models. Lastly, the empirical approach of the DEM makes it computationally inexpensive.

The DEM can be included on top of simulations and used to forward model observables and statistically compared to observations (*e.g.* Nelson et al. 2018; Baes et al. 2019; Trčka et al. 2020; ?). From such comparisons, we can exploit the statistical power of large galaxy surveys (*e.g.* SDSS) to constrain the DEM. These constraints provide insights into attenuation curves and their connection to galaxy properties. Besides providing insights into dust, the DEM can also be used to tractably marginalize over the impact of dust by treating the DEM parameters as nuisance parameters. The posterior probability distributions of the DEM parameters, derived from the comparison to observations, can be used to effectively marginalize over the DEM and disentangle the effect of dust from the subgrid galaxy formation prescriptions of the simulations.

In this work, we present a simple DEM that uses the Noll et al. (2009) attenuation curve parameterization and includes correlations with galaxy  $M_*$  and SFR. We apply the DEM to three state-of-the-art cosmological large-scale hydrodynamical simulations, SIMBA, IllustrisTNG, and EAGLE, which we describe in Section 2 and compare them to observed SDSS galaxies, which we describe in Section 2.5. In Section 3, we describe our DEM in detail. Finally, in Section 5, we present the results



**Figure 1.** The stellar mass functions ( $\Phi_{M_*}$ ; left-most panel) and  $M_* - \text{SFR}$  relation (right panels) of simulated galaxies from SIMBA (orange), TNG (blue), and EAGLE (green). For reference, we include  $\Phi_{M_*}$  (black) and the  $M_* - \text{SFR}$  relation (black dashed) of our SDSS galaxy sample. Uncertainties for the SDSS  $\Phi_{M_*}$  are derived using jackknife resampling. In Section 2, we describe the simulations and observations above. Differences in  $\Phi_{M_*}$  and the  $M_* - \text{SFR}$  relations among the simulations highlight that the *hydrodynamical simulations predict galaxy populations with significantly different the physical properties*.

of our comparison and discuss insights the DEM provides for our understanding of dust attenuation as well as of its interplay with galaxy evolution in cosmological simulations.

## 2. DATA

In this paper we apply empirical dust attenuation (EDA) prescriptions to galaxies in the Illustris TNG, EAGLE, and SIMBA cosmological hydrodynamical simulations. For each galaxy in the EDAoutput, we forward model the spectral energy distributions (SED) and measure the  $r$ -band luminosity ( $M_r$ ), optical color ( $G - R$ ), and UV color ( $FUV - NUV$ ) from the SEDs. These forward modeled observables, unlike physical properties such as  $M_*$  and SFR, are *consistently* defined and derived in both simulations and observations. Afterwards, we compare the predicted EDAobservables to galaxies in the SDSS observations. 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) for simulated galaxies of SIMBA (orange), TNG (blue), and EAGLE (green). We include, for reference,  $\Phi_{M_*}$  and the  $M_* - \text{SFR}$  relation for our SDSS galaxy sample. The uncertainties for the SDSS SMF are derived from jackknife resampling. 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 are from the current release of Brinchmann et al. (2004)<sup>1</sup>. Figure 1 illustrates that the hydrodynamical simulations predict significantly different SMFs and  $M_* - \text{SFR}$  relations. This difference, which was also recently highlighted

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

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> (hereafter TNG) is a cosmological hydrodynamical 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. *details on the following properties that we use in the paper: SFH, ZH*

TODO

### 2.2. *EAGLE*

We use L0100Ref of the Virgo Consortium’s EAGLE project<sup>4</sup>, a publicly available suite of cosmological hydrodynamic simulations ([Schaye et al. 2015](#); [Crain et al. 2015](#); [McAlpine et al. 2016](#)). The simulation has a comoving volume of  $(100 \text{ Mpc})^3$  and is simulated with the ANARCHY code ([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](#)). It 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. *details on the following properties that we use in the paper: SFH, ZH*

TODO

### 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. 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. *details on the following properties that we use in the paper: SFH, ZH*

TODO

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

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

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



#### 2.4. Forward Modeled Spectral Energy Distributions

One of the main goals of this work is to conduct an “apples-to-apples” comparison between observations and the simulations above. A crucial step in this is to forward model the observables for our simulations. In this work, we use the following observables:  $r$ -band luminosity ( $M_r$ ), optical color ( $G - R$ ), and UV color ( $FUV - NUV$ ). We measure these observables from the SED we construct for all the simulated galaxies based on their star formation and metallicity histories 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 the SEDs are constructed, we refer readers to Starkenberg et al. (in prep.). Next, we measure the observables by convolving the SEDs by 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 observed uncertainty distributions from NASA-Sloan Atlas.

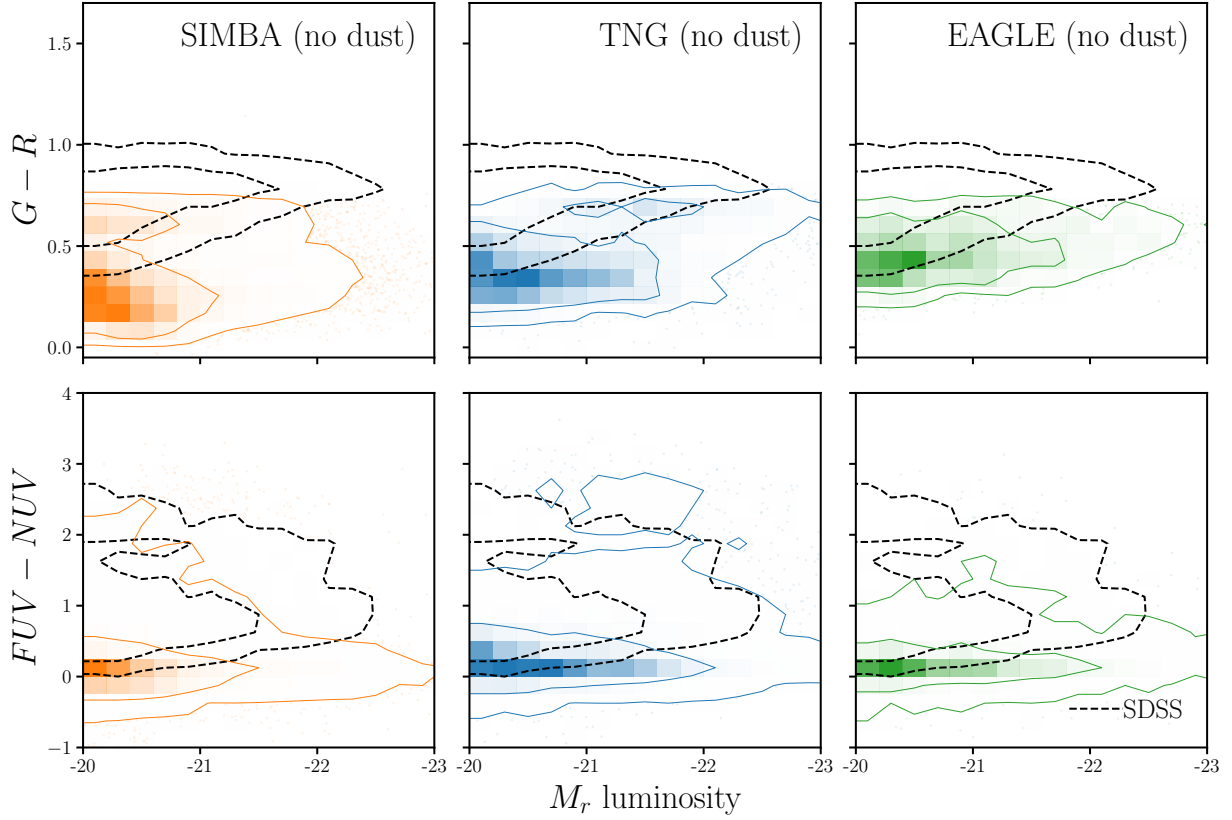
#### 2.5. SDSS Galaxies

The EDAModel provides a flexible prescription for assigning dust attenuation to simulated galaxies so that the hydrodynamical simulations, described above, can reproduce observations. Throughout this work, for our observations, we use a galaxy sample derived from SDSS. We impose a  $M_r < -20$  completeness cut on the Tinker et al. (2011) volume-limited sample that is complete above  $M_* > 10^{9.4} h^{-2} M_\odot$ . 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.

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. EMPIRICAL DUST ATTENUATION MODELING

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



**Figure 2.** We present the 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 from forward modeled SEDs and, thus, are consistently defined and measured as SDSS observations (Section 2.4). For comparison, we include the distributions of our SDSS sample (black dashed). *Without dust, the hydrodynamical simulations cannot reproduce the observed optical or UV color-magnitude.*

In this section we present the Empirical Dust Attenuation (EDA) model, a flexible prescription for applying attenuation curves to galaxy populations that allows us to incorporate intrinsic variations in dust attenuation as well as correlation to 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 from 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 all of our galaxies, we uniformly sample  $i$ . We then 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](#)), in  $\tau_V$ . We use  $\tau_V$  with a simple and flexible linear  $M_*$  and SFR dependence:

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

$m_{\tau, M_*}$  quantifies the  $M_*$  dependence,  $m_{\tau, \text{SFR}}$  quantifies the SFR dependence, and  $c_\tau$  quantifies the overall amplitude. Since  $\tau_V$  is optical depth, we impose a  $\tau_V \geq 0$  limit. We note that the slab model is a naive approximation. In reality,  $A_V$  for a galaxy will depend 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 sense, 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 8). Also, replacing the slab model with a more flexible prescription for sampling  $A_V$  does not significantly impact our analysis (Appendix B). We therefore conclude that the slab model is a sufficiently 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:

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

where  $\lambda_V = 5500\text{\AA}$  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 SFR dependence as  $\tau_V$ :

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



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

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

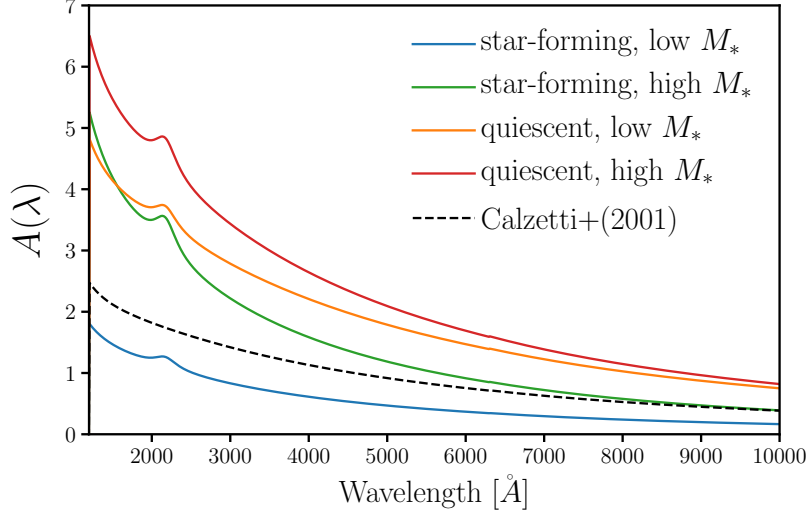
where  $\lambda_0$ ,  $\Delta\lambda$ , and  $E_b$  are the central wavelength, FWHM, and strength of the bump, respectively. We assume fixed  $\lambda_0 = 2175\text{\AA}$  and  $\Delta\lambda = 350\text{\AA}$ . Kriek & Conroy (2013) and Tress et al. (2018) find that  $E_b$  correlates with the  $\delta$  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.

SFR of galaxies are used to calculate  $\tau_V$  and  $\delta$  in Eqs. 4 and 6. 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). Eqs. 4 and 6 cannot be used to derive  $\tau_V$  and  $\delta$  for these galaxies. Since  $\text{SFR} = 0$  galaxies do not account for a large fraction of our simulated galaxies, we directly sample their observables ( $G, R, NUV$ , and  $FUV$ ) from the distribution of observables for SDSS quiescent galaxies. This way, we ensure that the attenuation of  $\text{SFR} = 0$  galaxies does not impact the rest of the EDA parameters. In Appendix A, we discuss the resolution effects in more detail and demonstrate that our results are *not* impacted by other prescriptions for attenuating  $\text{SFR} = 0$  galaxies.

To apply the EDA to a simulated galaxy population in practice, we begin by uniformly sample inclinations,  $i$ , and assign them to each galaxy. Then for a given set of EDA parameter values,  $\tau_V$ , and  $\delta$  are calculated for each galaxy using its  $i$ ,  $M_*$ , and SFR. From  $\tau_V$  and  $\delta$ , we determine  $A_V$  and  $k(\lambda)$ , which together gives  $A(\lambda)$  for each of the galaxies. Afterwards, we attenuate the galaxy SEDs using Eq. ?? and use the attenuated SEDs to calculate the observables:  $G, R, NUV$ , and  $FUV$  absolute magnitudes. In Figure 3, we present the EDA  $A(\lambda)$  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  $\{m_{\tau, M_*}, m_{\tau, \text{SFR}}, c_\tau, m_{\delta, M_*}, m_{\delta, \text{SFR}}, c_\delta\} = \{2., -2., 2., -0.1, -0.1, -0.2\}$ , which were arbitrarily chosen within the prior range listed in Table 1. For comparison, we include the Calzetti (2001) attenuation curve. The EDA we describe in this section provides a flexible model for assigning a wide range of dust attenuation to galaxies based on their physical properties.

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

The DEM provides a flexible model to apply dust attenuation to central galaxies from the hydrodynamic simulations (Section 2) and derive observables which can be directly compared to SDSS observations. For the comparison and the inference of DEM 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



**Figure 3.** Attenuation curves,  $A(\lambda)$ , of our 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  $\{m_{\tau, M_*}, m_{\tau, \text{SFR}}, c_\tau, m_{\delta, M_*}, m_{\delta, \text{SFR}}, c_\delta\} = \{2., -2., 2., -0.1, -0.1, -0.2\}$  near the center of our prior range (Table 1). For comparison, we include the Calzetti (2001) attenuation curve. In the EDA, the amplitude, slope, and UV dust bump of  $A(\lambda)$  depend on  $M_*$ , and SFR (Eqs. 4 and 6). *The EDA provides a flexible model for assigning 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 $[-5., 5.]$
$m_{\tau, \text{SFR}}$	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 $[-4., 4.]$

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. Then the output of the forward model is then compared to the observed data using the distance metric and a distance threshold. 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 is the hydrodynamic simulation combined with the DEM. Given a set DEM parameters, our forward model produces  $G$ ,  $R$ ,  $NUV$ , and  $FUV$  absolute magnitudes, which can be directly compared to SDSS observations. We use uninformative uniform priors on each of the DEM 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 DEM parameters do not result in 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 DEM parameters and not on the derived  $\tau_V$  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 of the SDSS observation and our forward model:

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

$\theta_{\text{DEM}}$  are the DEM parameters. The summary statistics are based on the optical and UV color-magnitude relations,  $(G-R)-R$  and  $(FUV-NUV)-R$ , of central galaxies brighter than  $M_r < -20$ , where our SDSS central galaxy sample is complete (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 included in the analysis.

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 DEM 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.

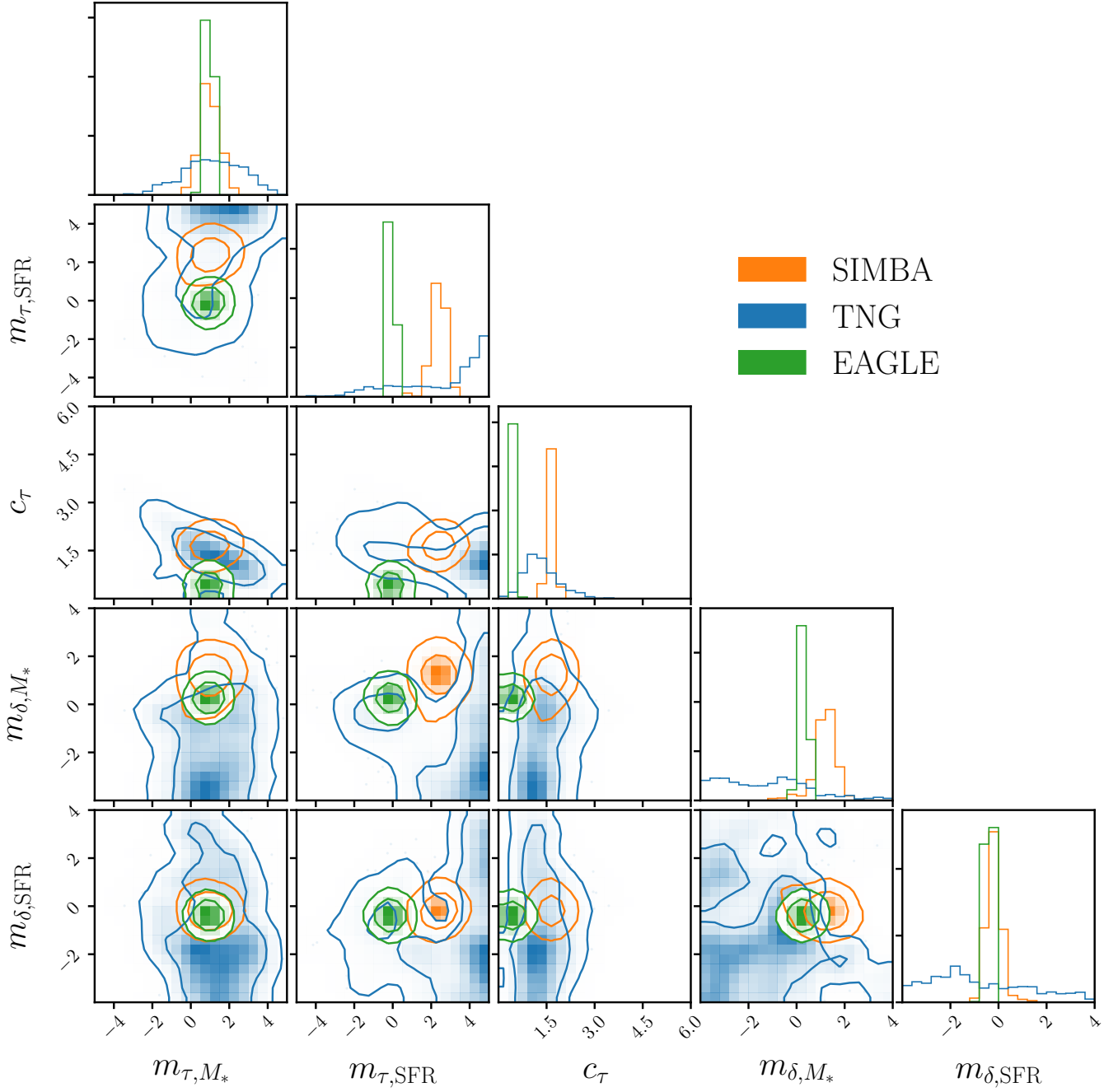
## 5. RESULTS

### 5.1. Reproducing SDSS Color-Magnitude Relations

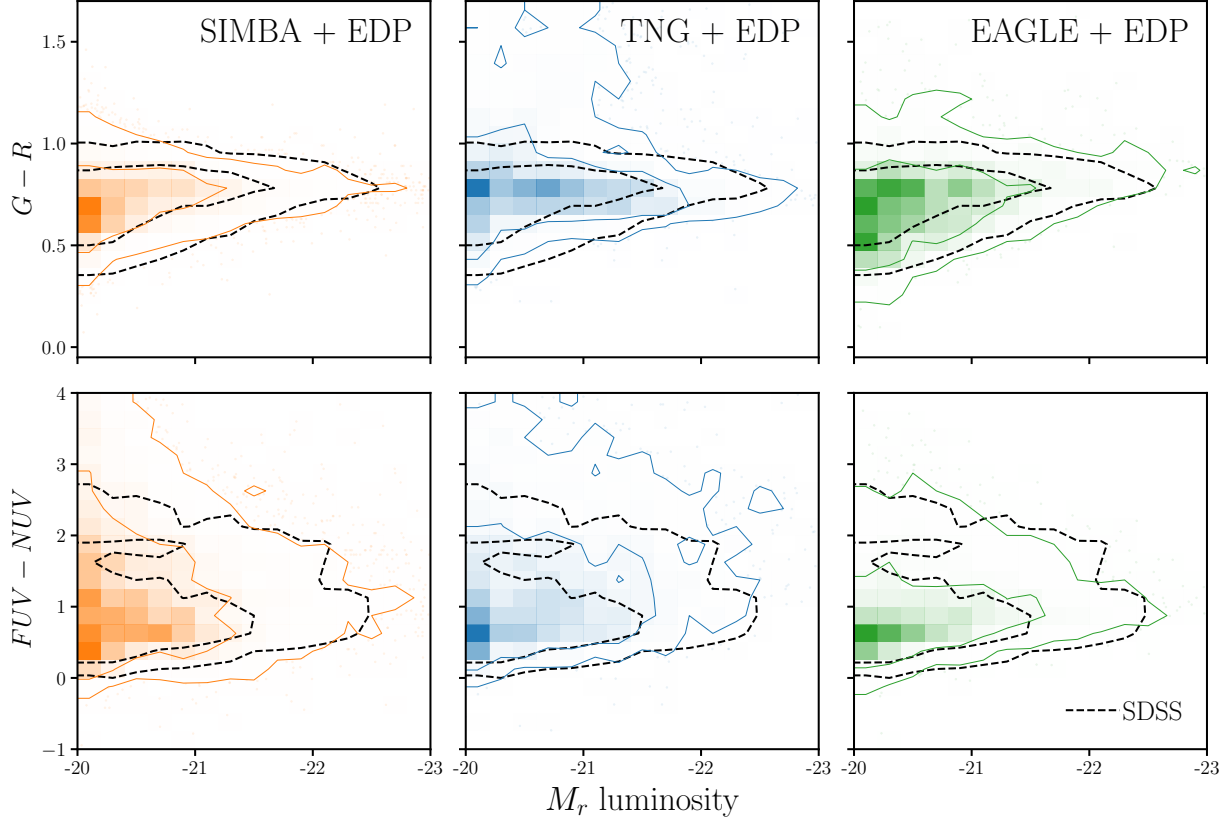
In Figure 5, we present the optical and UV color-magnitude relations predicted by the DEM with the median ABC posteriors for the SIMBA (orange), TNG (blue), and EAGLE (green) simulations. We include the SDSS observables for comparison (black dashed). Without any dust attenuation, we previously found that simulations predict dramatically different  $(G-R)-M_r$  and  $(FUV-NUV)-M_r$  relations than SDSS (Figure 2). In contrast, with the DEM, the optical color-magnitude relations have well-defined red sequences and blue clouds that are consistent with SDSS. The DEM also produces galaxies with  $FUV-NUV$  distributions that are consistent with SDSS. We also find good agreement in the galaxy number density at  $M_r < -20$ : **CH: numbers**

Previous works in the literature have also compared colors and luminosities predicted by simulations to observations. 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. Both Trayford et al. (2015) and Trayford et al. (2017) produce bluer red sequences compared to GAMA observations, at  $10^{11.2} < M_* < 10^{11.5}$  for Trayford et al. (2017). Although a detailed comparison is difficult since both works examine all galaxies, not only centrals, the DEM accurately reproduces the position of the SDSS red sequence, even at high  $M_*$ . Trayford et al. (2015) also predict significant more luminous blue galaxies than observations or the DEM. Also using EAGLE and SKIRT framework, 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 DEM, on the other hand, 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 with the DEM. *With the DEM, we produce optical and UV color-magnitude relations that are in good agreement with observations, better than previous works, for the SIMBA, TNG, and EAGLE hydrodynamical simulations.*

Figures 5 clearly illustrates that any comparison of simulations must include dust attenuation. Dust dramatically changes the predicted observables of simulations. Without dust (Figure 2), we did not find a clearly bimodality in the optical color-magnitude relation and the simulations predicted UV colors outside of the range of observations. But with attenuation curves assigned from the DEM, simulations successfully reproduce observations. Our results also highlight another key point. Even for three simulations that produce significantly different SMFs and  $M_* - \text{SFR}$  relations (Figure 1), the DEM is able to produce observables that agree with observations. In fact for SIMBA, the DEM reproduces the observations by assigning higher attenuation to star-forming galaxies so that they



**Figure 4.** Posterior distributions of DEM 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 Approximate Bayesian Computation with Population Monte Carlo (Section 4). We focus on the DEM 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 be 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.

populate the red sequence while quiescent galaxies populate the blue cloud. This is due to the large number of low mass star-forming SIMBA galaxies that lie well above the SFS (Figure 1), which would otherwise all be luminous blue galaxies not found in observations. Our current understanding of dust, which is encapsulated in the DEM, has enough flexibility to reproduce observations for simulations that predict galaxy populations with different physical properties, even if it contradicts the established relationship between color and SFR. This means that marginalizing over dust would leave little constraining power on the subgrid galaxy physics of the simulations. Therefore, *current limitations in our understanding of dust in galaxies significant impedes our ability to investigate galaxy formation from simulations.*

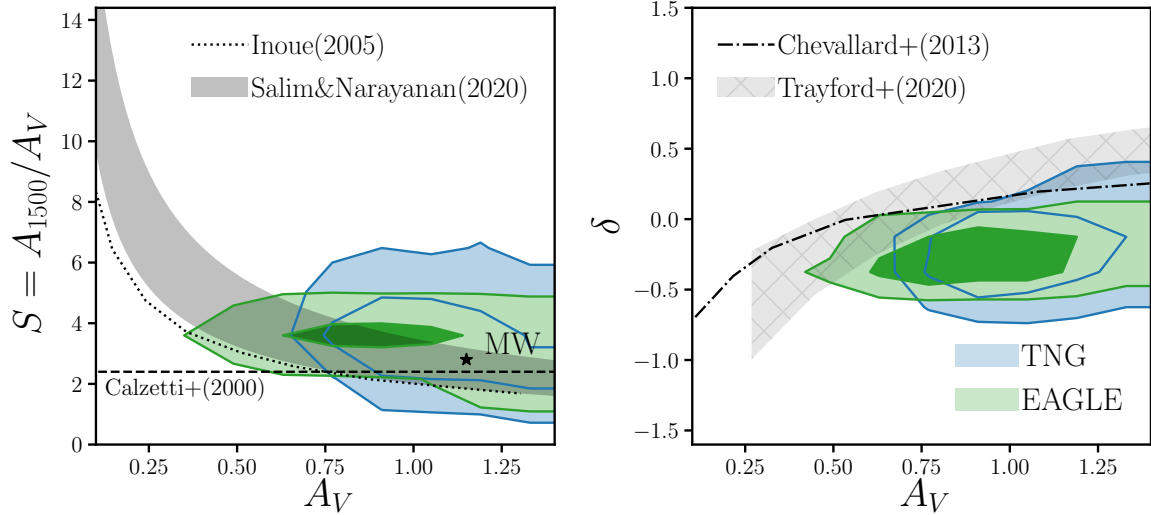
## 5.2. Reproducing the Observed Attenuation-Slope Relation



In addition to reproducing the observed color-magnitude relations, we also find good agreement between the attenuation curves of the DEM and observations. In Figure 6, we present the attenuation-slope relation for TNG (blue) and EAGLE (green) with the DEM. The left and right panels present two different measurements of the UV-optical slope  $S = A(3000\text{\AA})/A_V$ , which is easier to constrain in observations, and  $\delta$ , the slope offset from the Calzetti (2001) curve that use in the DEM. We include the attenuation and slope for the Milky Way (star) for reference. TNG and EAGLE both predict slopes within  $2 < S < 5$  and centered around  $S \sim 3.5$ . In comparison, for the same  $A_V > 0.4$  range as the DEM, observations find slopes within the range  $2 < S < 5$  (Calzetti et al. 2000; Burgarella et al. 2005; Johnson et al. 2007; Conroy 2010; Wild et al. 2011; Battisti et al. 2016, 2017; Leja et al. 2017; Salim et al. 2018) in good agreement. We also find that the DEM predicts steeper attenuation curves at lower attenuation. This is consistent with the established attenuation-slope relation. 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). For the  $A_V$  range probed by the DEM, the  $A_V$ -slope relation is in good agreement with GSWLC2 galaxies (black shaded Salim & Narayanan 2020). They are also consistent with Leja et al. (2017). We also compare our results to theoretical predictions from radiative transfer models, Inoue (2005) (dotted), the radiative transfer models considered in Chevallard et al. (2013) (dot dashed), and Trayford et al. (2020) (light shaded), which all predict shallower attenuation curves than observations. This is also the case for the Narayanan et al. (2018) attenuation curves (not included). *The attenuation curve slopes from the DEM for are in excellent agreement with observations and better reproduces the observed attenuation-slope relation than radiative transfer models.*

### 5.3. The Galaxy – Dust Attenuation Connection

Next, in Figure 7, we present the normalized attenuation curves of the TNG (blue) and EAGLE (green) DEM for low (top) and high  $M_*$  (bottom), star-forming (left) and quiescent galaxies (right). The attenuation curves are normalized at  $3000\text{\AA}$  in order to emphasize their slopes. The DEM includes significant variations in the dust attenuation through the slab model and the dependence on galaxy properties. We present this variation in the shaded region ( $1\sigma$  standard deviation about the median). For comparison, we include  $A(\lambda)/A(3000\text{\AA})$  from observations (Calzetti et al. 2000; Battisti et al. 2017; Salim et al. 2018) as well as from simulations (Narayanan et al. 2018). The DEM attenuation curves for star-forming centrals in TNG and EAGLE are in good agreement with the Salim et al. (2018) attenuation curves for  $10^{9.5} < M_* < 10^{10.5} M_\odot$  (top left) and  $10^{10.5} < M_* M_\odot$  star-forming galaxies (bottom left). They are also consistent with the median curve of Narayanan et al. (2018) star-forming galaxies. On the other hand, they are steeper than the Calzetti et al. (2000) and Battisti et al. (2017) curves. However, we note that the Calzetti et al. (2000) and Battisti et al. (2017) curves are derived from  $M_* < 10^{9.9} M_\odot$  star-forming galaxies (below the our  $M_*$  range). We find that lower  $M_*$  star-forming galaxies have shallower attenuation curves, especially for TNG. If this  $M_*$  dependence continues below our  $M_*$  range, we expect a better agreement the Calzetti et al. (2000) and Battisti et al. (2017) curves as well. In Salim et al. (2018) and Narayanan et al. (2018),



**Figure 6.** The attenuation-slope relation of the DEM for TNG (blue) and EAGLE (green). We present the relation using two different measurements of slope, commonly used in the literature:  $S = A(1500\text{\AA})/A_V$  (left panel) and the slope offset from the Calzetti (2001) curve,  $\delta$  (right panel). The DEM models predict an attenuation-slope relation, where the slope is steeper at lower attenuation, consistent with both observations and radiative transfer models. We only include massive galaxies with  $M_r < -20$  so galaxies in our sample have  $A_V \gtrsim 0.4$ . In this  $A_V$  range, the DEM for TNG and EAGLE are in good agreement with observations Salim & Narayanan (2020). In fact, the DEMs match the observed attenuation-slope relation better than radiative transfer model that predict attenuation curves that are too shallow (Inoue 2005; Chevallard et al. 2013; Trayford et al. 2020).

they find large variations in the slopes of the attenuation curve. This is, however, over their entire  $M_*$  range. For a narrow  $M_*$  range of star-forming galaxies, we find significantly less variation. *Overall, the DEM attenuation curves for star-forming galaxies are in good agreement with attenuation curves in the literature from both observations and simulations.*

So far, we have demonstrated that with the DEM we successfully reproduce the observed color-magnitude relation as well as the attenuation-slope relation. We also find good agreement between the attenuation curves of the DEM and in the literature for star-forming galaxies. Besides reproducing observations and trends in the literature, the constraints on the DEM provide insights into dust attenuation in galaxies. In fact, the parameterization of the DEM makes it especially easy to interpret correlation between dust attenuation and galaxy physical properties. In Figure 4, for all three simulations we find we find significant positive  $M_*$  dependence of  $\tau_V$  ( $m_{\tau, M_*} \sim 2$ ) consistent with previous works in 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 positive  $M_*$  dependence in SDSS star-forming galaxies. Most recently, Salim et al. (2018) find higher  $V$  and  $FUV$  attenuation for more massive star-forming galaxies in GSWLC2. In addition to the  $M_*$  dependence, the DEM posteriors also reveal the correlation between dust attenuation and star formation. Ignoring SIMBA, which flips the color versus SFR relation, we find  $m_{\tau, \text{SFR}} \sim -1$  — galaxies with lower SFR have higher attenuation. Observations that examine the

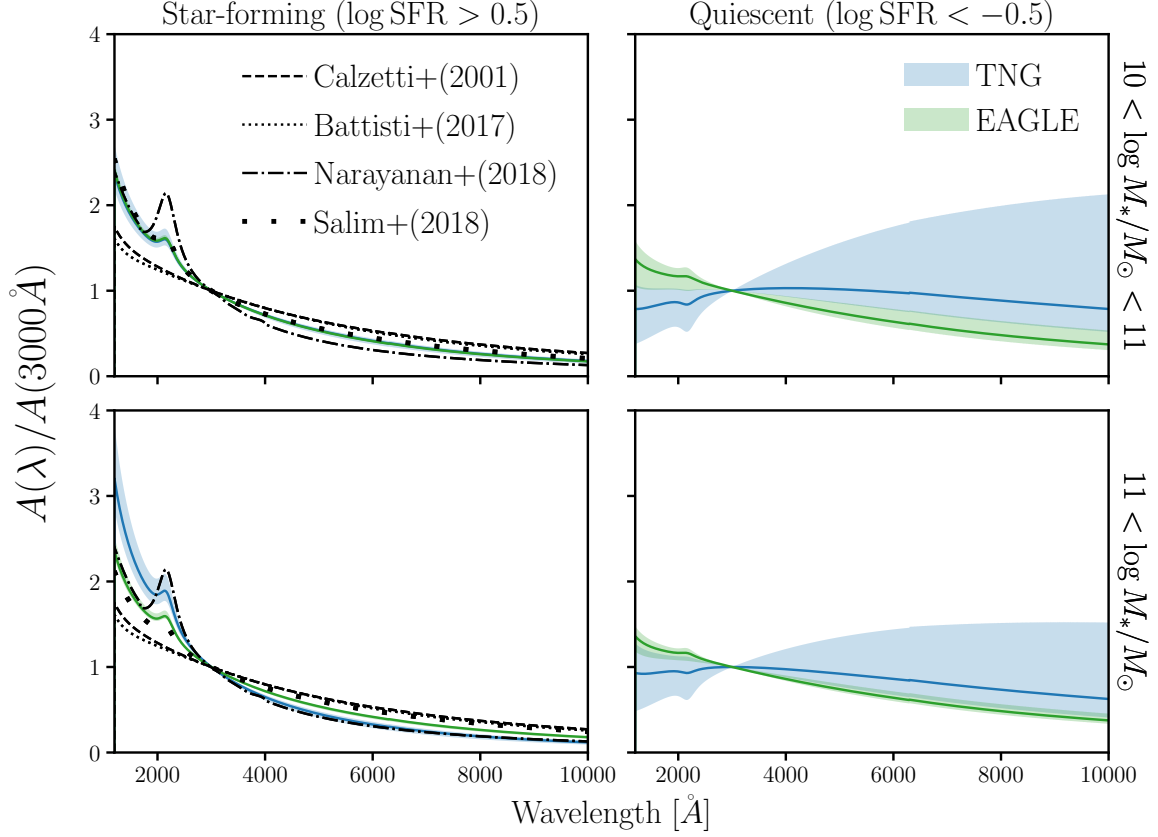
relationship between dust attenuation and SFR (*e.g.* Garn & Best 2010; Reddy et al. 2015; Battisti et al. 2016, 2017; Salim et al. 2018) have thus far focused primarily on star-forming galaxies. With the DEM, we confirm that *galaxies with higher  $M_*$  have overall higher dust attenuation and find that galaxies with lower SFR have overall higher dust attenuation.*

In addition to the correlation with galaxy properties, the DEM also reveals dust attenuation in quiescent galaxies (left panels of Figure 7). This is particularly valuable since there are many challenges to measuring attenuation curves for quiescent galaxies directly from observations. For instance, 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 due to the degeneracies with star formation history and metallicity (CH: cite?). By forward modeling the optical and UV photometry with the DEM, we do not face these issues. For both TNG and EAGLE, we find that quiescent galaxies have significantly shallower attenuation curve than star-forming galaxies. They also have significantly larger variations than star-forming galaxies.

The DEM produces optical and UV color-magnitude relations overall consistent with SDSS. There are, however, still a few discrepancies between the DEM observables and SDSS. For instance, the DEM produces broader distributions overall than observations. Galaxies in SDSS sharply cut-off above the red sequence, while some galaxies in the DEM broadly extend beyond the cut-off. The DEM also produces galaxies more luminous galaxies than SDSS. Nevertheless, the DEM better reproduces observations than other works. Furthermore, we chose linear parameterization for  $\tau_V$  and  $\delta$  in the DEM for simplicity. However, the empirical framework of the DEM can easily be extended to more flexible parameterizations that better reproduce observations — the only challenge would be to find a well-motivated parameterization from observations.

In the DEM, we make a few other assumptions and choices. For simulated galaxies with  $\text{SFR} = 0$ , we directly sample their observables from the distributions of SDSS quiescent galaxies. These  $\text{SFR} = 0$  galaxies do not have recent star-formation and also have 0 gas mass CHH: @tjitske is this for all sims so we would expect them to also have no dust. However, without attenuating these galaxies, the simulations struggle to reproduce observations. Our prescription for  $\text{SFR} = 0$  galaxies ensures that  $\text{SFR} = 0$  galaxies do not impact our results, without delving into the issue further. In Appendix A, we describe our prescription in detail and discuss the limitations of the simulations near the mass and temporal resolutions or their gas prescriptions. Another assumption in the DEM is how we assign  $A_V$  using the slab model (Eq. 3). The slab model is consistent with 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). Furthermore, it is able to reproduce the SDSS  $A_V$  distribution (Figure 8). When we further test the robustness of our results by replacing the slab model with a more flexible truncated normal distribution in Appendix B. We find that our results are not significant impacted. We therefore conclude that our results are robust to the assumptions and choices we make in the DEM.

TODO



**Figure 7.** Attenuation curves of the DEM for TNG (blue) and EAGLE (green) for low (top) and high  $M_*$  (bottom), star-forming (left) and quiescent galaxies (right). The attenuation curves are normalized at  $3000\text{\AA}$ :  $A(\lambda)/A(3000\text{\AA})$ . We mark the  $1\sigma$  standard deviation of the attenuation curves with the shaded region. For comparison, we include measurements of  $A(\lambda)/A(3000\text{\AA})$  from observations (Calzetti et al. 2000; Battisti et al. 2017; Salim et al. 2018) as well as from 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 DEM attenuation curves. However, this is primarily driven by the differences in  $M_*$  ranges. 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 DEM, we can also constrain the attenuation curves of quiescent galaxies, which are challenging to directly constrain from observations. Quiescent galaxies have significantly shallower attenuation curves and larger variations than star-forming galaxies.

In this work, we present the DEM, a framework for statistically assigning attenuation curves to galaxy populations. We apply the DEM to the SIMBA, TNG, and EAGLE hydrodynamical simulations and forward model optical and UV color-magnitude relations. With the DEM, we reproduce the color-magnitude relations of the SDSS observations for all three simulations. Also, the attenuation curves of the DEM closely reproduce the observed attenuation–slope relation, better than radiative transfer models. Furthermore, the DEM produces attenuation curves that are in good agreement with the literature for star-forming galaxies.

Focusing on the DEM for TNG and EAGLE, we find significant  $M_*$  and SFR dependences in the amplitude of dust attenuation,  $A_V$ . More massive galaxies have higher dust attenuation; galaxies with lower SFR have higher dust attenuation. The DEM is also able to constrain the attenuation curves for quiescent galaxies, which have few constraints from observations. We find that quiescent galaxies have shallower attenuation curves with larger variation than star-forming galaxies.

By reproducing SDSS observations with the DEM for different hydrodynamical simulations, we demonstrate that accounting for dust attenuation is essential to reproduce observations. However, based on our current understanding of dust there is enough flexibility to reproduce observations even for simulations that predict galaxy populations with significantly different physical properties. In fact, for SIMBA the inferred dust attenuation reverses the established relationship between color and SFR in order to account for SIMBA overpredicting a large starburst population at  $< 10^{10} M_\odot$ . Since adjusting dust alone can reproduce observations, dust is highly degenerate with the variations in subgrid physics across simulations. In other words, if we were to marginalize over dust we would not be able to differentiate between the various hydrodynamical models using observations. So detailed comparisons across simulations and to observations likely overinterpret the differences and similarities found in simulations. Therefore, the current limitations in our understanding of dust is a major bottleneck for investigating galaxy formation using simulations.

## 6. SUMMARY

In this work, we present the DEM, an empirical framework for including dust attenuation in 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 SFR). We apply the DEM to three state-of-the-art hydrodynamical simulations (SIMBA, TNG, and EAGLE) and forward model the optical and UV color-magnitude relations. Afterwards, we compare these forward modeled simulations to observed central galaxies in SDSS using simulation-based inference. Based on this comparison, we find the following results:

- Dust attenuation is essential for simulations to reproduce observations. With the DEM, we are able to reproduce SDSS observations for all three hydrodynamical simulations. However, SIMBA requires an extreme dust attenuation that reverses the established relationship between color and SFR in order to account for SIMBA overpredicting starburst galaxies at  $< 10^{10} M_\odot$ .
- The DEM attenuation curves are in good agreement with major observational constraints. They closely reproduce the observed attenuation–slope relation, better than radiative transfer models. We also find significant variation in the attenuation curves as in observations. For star-forming galaxies, the DEM attenuation curves are in good agreement with the literature.
- Focusing on the DEM for TNG and EAGLE, we find significant  $M_*$  and SFR dependences in  $A_V$ . More massive galaxies have higher dust attenuation; galaxies with lower SFR have higher dust attenuation.

- Lastly, the DEM is able to constrain the attenuation curves of quiescent galaxies, which are poorly constrained by observations. We find that quiescent galaxies have shallower attenuation curves with higher  $A_V$  and larger variation than star-forming galaxies.

Our results clearly demonstrate that the DEM can be used to provide insight into dust attenuation. For those uninterested in dust, the DEM also provides a straightforward framework for marginalizing over dust. Although accounting for dust is necessary to reproduce observations, our limited understanding of dust attenuation allows simulations that predict galaxy populations with significantly different physical properties to reproduce the same observable. Since varying attenuation alone can entirely reproduce observations, dust is highly degenerate with differences in subgrid physics across simulations. After marginalizing over dust, observations do not have the constraining power to differentiate between the various hydrodynamical models. Hence, detailed comparisons across simulations and to observations likely overinterpret the differences and similarities found in simulations. Therefore, we demonstrate that the current limitations in our understanding of dust is a major bottleneck for investigating galaxy formation using simulations.

Even with the limited statistical power of our  $M_r$  complete SDSS central galaxy sample, we derived tight constraints on dust attenuation. Upcoming surveys such as the Bright Galaxy Survey of the Dark Energy Spectroscopic Instrument (DESI; [DESI Collaboration et al. 2016](#), [CH: Ruizet et al. 2020](#)), the Galaxy Evolution Survey of the Prime Focus Spectrograph (PFS; [Takada et al. 2014](#); [Tamura et al. 2016](#)), and the Wide-Area VISTA Extragalactic Survey (WAVES; [Driver et al. 2016, 2019](#)) will provide much more statistically powerful observations and at higher redshifts. With these observations, the DEM will be able to more 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).

Besides its simplifications, the slab model predicts  $A_V$  distribution with significant differences than the  $A_V$  distributions measured from observations. In Figure 8, 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.

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:

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

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:

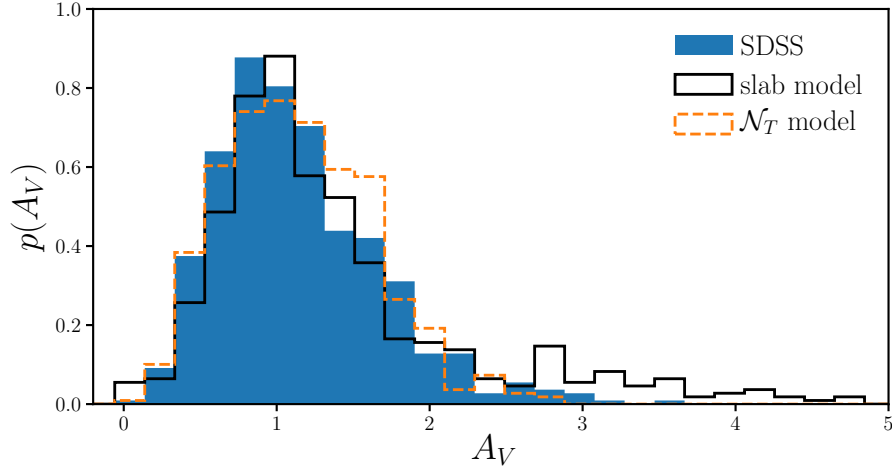
$$\mu_{A_V} = m_{\mu,1}(\log M_* - 10.) + m_{\mu,2} \log \text{SFR} + c_{\mu} \quad (\text{B2})$$

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

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 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

- |   |   |
|---|---|
| <p>Abazajian K. N., et al., 2009, <i>The Astrophysical Journal Supplement Series</i>, 182, 543</p> <p>Aihara H., et al., 2011, <i>The Astrophysical Journal Supplement Series</i>, 193, 29</p> <p>Alsing J., Wandelt B., Feeney S., 2018, arXiv:1801.01497 [astro-ph]</p> | <p>Anglés-Alcázar D., Davé R., Faucher-Giguère C.-A., Özel F., Hopkins P. F., 2017, <i>Monthly Notices of the Royal Astronomical Society</i>, 464, 2840</p> <p>Baes M., Trčka A., Camps P., Nersesian A., Trayford J., Theuns T., Dobbels W., 2019, arXiv:1901.08878 [astro-ph]</p> |
|---|---|



**Figure 8.** 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).

- 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
- Driver S. P., Davies L. J., Meyer M., Power C., Robotham A. S. G., Baldry I. K., Liske J., Norberg P., 2016, *The Universe of Digital Sky Surveys*, 42, 205
- Driver S. P., et al., 2019, *The Messenger*, 175, 46
- 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
- Johnson B. D., et al., 2007, *The Astrophysical Journal Supplement Series*, 173, 392
- 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