The Astrophysical Journal, 926:122 (20pp), 2022 February 20 https://doi.org/10.3847/1538-4357/ac4253

IQ Collaboratory. III. The Empirical Dust Attenuation Framework—Taking Hydrodynamical Simulations with a Grain of Dust

ChangHoon Hahn[1], Tjitske K. Starkenburg[2], Daniel Anglés-Alcázar[3,4], Ena Choi[5], Romeel Davé[6,7,8], Claire Dickey[9], Kartheik G. Iyer[10], Ariyeh H. Maller[11], Rachel S. Somerville[4,12], Jeremy L. Tinker[13], and L. Y. Aaron Yung[14]
[1] Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton NJ 08544, USA; changhoon.hahn@princeton.edu
[2] Center for Interdisciplinary Exploration and Research in Astrophysics (CIERA) and Department of Physics and Astronomy, 1800 Sherman Ave, Evanston IL 60201, USA
[3] Department of Physics, University of Connecticut, 196 Auditorium Road, U-3046, Storrs, CT 06269-3046, USA
[4] Center for Computational Astrophysics, Flatiron Institute, 162 5th Ave, New York, NY 10010, USA
[5] Quantum Universe Center, Korea Institute for Advanced Study, Hoegiro 85, Seoul 02455, Republic of Korea
[6] Institute for Astronomy, Royal Observatory, Univ. of Edinburgh, Edinburgh EH9 3HJ, UK
[7] University of the Western Cape, Bellville, Cape Town 7535, South Africa
[8] South African Astronomical Observatories, Observatory, Cape Town 7925, South Africa
[9] Department of Astronomy, Yale University, New Haven CT 06520 USA
[10] Dunlap Institute for Astronomy and Astrophysics, University of Toronto, 50 St George St, Toronto, ON M5S 3H4, Canada
[11] Department of Physics, New York City College of Technology, City University of New York, 300 Jay St., Brooklyn, NY 11201, USA
[12] Department of Physics and Astronomy, Rutgers University, 136 Frelinghuysen Road, Piscataway, NJ 08854, USA
[13] Center for Cosmology and Particle Physics, Department of Physics, New York University, 4 Washington Place, New York, NY 10003, USA
[14] Astrophysics Science Division, NASA Goddard Space Flight Center, 8800 Greenbelt Rd, Greenbelt, MD 20771, USA
Received 2021 June 17; revised 2021 November 7; accepted 2021 November 20; published 2022 February 17

Abstract

We present the empirical dust attenuation (EDA) framework—a flexible prescription for assigning realistic dust attenuation to simulated galaxies based on their physical properties. We use the EDA to forward model synthetic observations for three state-of-the-art large-scale cosmological hydrodynamical simulations: SIMBA, IllustrisTNG, and EAGLE. We then compare the optical and UV color–magnitude relations, $(g-r)-M_r$ and (far-UV −near-UV) $-M_r$, of the simulations to a $M_r<-20$ and UV complete Sloan Digital Sky Survey galaxy sample using likelihood-free inference. Without dust, none of the simulations match observations, as expected. With the EDA, however, we can reproduce the observed color–magnitude with all three simulations. Furthermore, the attenuation curves predicted by our dust prescription are in good agreement with the observed attenuation–slope relations and attenuation curves of star-forming galaxies. However, the EDA does not predict star-forming galaxies with low A_V since simulated star-forming galaxies are intrinsically much brighter than observations. Additionally, the EDA provides, for the first time, predictions on the attenuation curves of quiescent galaxies, which are challenging to measure observationally. Simulated quiescent galaxies require shallower attenuation curves with lower amplitude than star-forming galaxies. The EDA, combined with forward modeling, provides an effective approach for shedding light on dust in galaxies and probing hydrodynamical simulations. This work also illustrates a major limitation in comparing galaxy formation models: by adjusting dust attenuation, simulations that predict significantly different galaxy populations can reproduce the same UV and optical observations.

Unified Astronomy Thesaurus concepts: Galaxy formation (595); Galaxy physics (612); Astronomical simulations (1857); Hydrodynamical simulations (767); Astrophysical dust processes (99); Galaxy evolution (594)

1. Introduction

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

To better understand dust in galaxies, many observational works have examined trends between dust attenuation and galaxy properties. For example, UV and optical attenuation are found to correlate with galaxy properties such as M_*, SFR, and metallicity in star-forming galaxies (Garn & Best 2010; Battisti et al. 2016; for a recent review see Salim & Narayanan 2020). The slopes of the attenuation curves in star-forming galaxies also correlate with galaxy properties, such as M_*, specific SFR (SSFR), metallicity, and axial ratio (Wild et al. 2011; Battisti et al. 2017). Despite the progress, there is still no clear consensus on the connection between dust attenuation and galaxy properties. Furthermore, studies so far have focused mainly on star-forming galaxies and

little is known about dust attenuation in quiescent galaxies. A major limitation of observational approaches is that dust attenuation and galaxy properties measured from galaxy SEDs are model dependent and subject to variations, inconsistencies, and biases of different methodologies. Different methods can measure vastly different derived values, even for the same observations (e.g., Speagle et al. 2014; Katsianis et al. 2020, see also the Appendix). SED modeling can also impose undesirable priors on derived galaxy properties (Carnall et al. 2019; Leja et al. 2019) and suffer from parameter degeneracies that are poorly understood.

Significant progress has also been made in theoretically modeling dust. Simulations can now model the radiative transfer of stellar light through a dusty interstellar medium for a wide range of configurations: from simple slab-like dust geometries (e.g., Witt & Gordon 1996, 2000; Seon & Draine 2016) to 3D hydrodynamical simulations of entire galaxies (e.g., Jonsson 2006; Rocha et al. 2008; Hayward & Smith 2015; Natale et al. 2015; Hou et al. 2017). Radiative transfer models have even been applied to cosmological hydrodynamical simulations (e.g., Camps & Baes 2015; Narayanan et al. 2018; Cochrane et al. 2019; Rodriguez-Gomez et al. 2019; Trayford et al. 2020). Dust has also been examined in a cosmological context using semianalytic models (e.g., Granato et al. 2000; Fontanot et al. 2009; Wilkins et al. 2012; Gonzalez-Perez et al. 2013; Popping et al. 2017). Yet there are still major limitations in modeling dust. Dust models in cosmological simulations currently do not reproduce the redshift evolution of dust properties (Somerville et al. 2012; Yung et al. 2019; Vogelsberger et al. 2020, but see Li et al. 2019), and radiative transfer models produce attenuation-slope relations that are significantly steeper than observations. Many models also require significant hand-tuning (e.g., propagating rays/photons into particular cells) and make assumptions on the underlying dust grain models (see Steinacker et al. 2013, for a review). Lastly, radiative transfer models are computationally expensive. Applying a range of radiative transfer dust models to multiple simulations for comparisons would require huge computational resources. Using them with Monte Carlo sampling methods for parameter exploration or to marginalize over the impact of dust would be prohibitive.

We take a different approach from the observational and theoretical works above: we investigate dust attenuation using a forward modeling approach to compare simulations to observations. Our "forward model" starts with three major large-scale hydrodynamical simulations: EAGLE (Schaye et al. 2015), IllustrisTNG (Nelson et al. 2019), and SIMBA (Davé et al. 2019). We use their outputs (e.g., SFH) to build SEDs for each simulated galaxy. We then apply dust attenuation to the SEDs using the empirical dust attenuation (EDA) framework, which we describe shortly. We construct photometry from the attenuated SEDs and afterward apply a realistic noise model and sample selection function to construct synthetic observations. Finally, we compare the synthetic observations to actual observations and constrain the free parameters of our EDA prescription. The EDA parameter constraints then provide insight into dust attenuation and its connection to galaxies.

The EDA framework provides a flexible and computationally inexpensive prescription for statistically assigning attenuation curves to simulated galaxies based on their physical properties. In this work, we assign attenuation curves based on M_*, SSFR, and the EDA parameters, as well as randomly sampled galaxy inclinations that introduce stochasticity. The EDA parameters set the strength of the M_* and SSFR dependence for the amplitude and slope of the attenuation curves. Unlike radiative transfer models, the EDA does not seek to produce realistic dust attenuation for individual galaxies. However, as we later demonstrate, it produces realistic distributions of dust attenuation for galaxy populations. The EDA provides an empirical framework for dust attenuation, analogous to the halo occupation framework in galaxy formation (for a review, see Wechsler & Tinker 2018).

With a forward modeling approach, we compare simulations to observations entirely in observational space, so we are not impacted by the inconsistencies of observational methods for measuring galaxy properties. Forward models can also directly include the selection functions and observational systematic effects so they can account for these effects to exploit the full observational data set. Furthermore, since we use the EDA in our forward model, we can apply a wide range of realistic dust attenuation curves to simulated galaxies in a matter of seconds, unlike radiative transfer models. This means we can easily explore and sample the dust parameter space and infer the relationship between dust attenuation and galaxy properties. That is the focus of this paper. Beyond investigating dust, the EDA also provides a framework through which we can treat dust as *nuisance* parameters and tractably marginalize over dust attenuation. In the subsequent paper in the IQ series, T. K. Starkenburg et al. (in preparation), we will use the EDA framework to compare star formation quenching in cosmological galaxy formation models after marginalizing over dust attenuation.

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

2. Data

In this paper, we present the EDA model and demonstrate how it can be used in a forward modeling approach to compare galaxy populations in simulations and observations. For our simulations, we use three large-scale cosmological hydrodynamical simulations: the IllustrisTNG (hereafter TNG), EAGLE, and SIMBA. For our observations, we use a galaxy sample derived from SDSS. Below, we briefly describe the simulations and the SDSS observations used throughout this work.

In Figure 1, we present the stellar mass functions (SMFs), Φ_{M_*} (left-most panel), and M_*–SFR relations (right panels) of galaxies in SIMBA (orange), TNG (blue), and EAGLE (green). For reference, we include Φ_{M_*} and the M_*–SFR relation for SDSS observations. For the simulations, M_* is the total stellar mass within the subhalo and SFR is the instantaneous SFR in the dense and cold star-forming gas. We do not impose any selection cuts on the simulations in Figure 1. For SDSS, we use `kcorrect` M_* assuming a Chabrier (2003) initial mass function (Blanton & Roweis 2007) and SFR from the current release of Brinchmann et al. (2004).[15] The uncertainties for the SDSS SMF are derived from jackknife resampling. Although we find striking differences between the M_*–SFR relation of SDSS and the

[15] http://www.mpa-garching.mpg.de/SDSS/DR7/

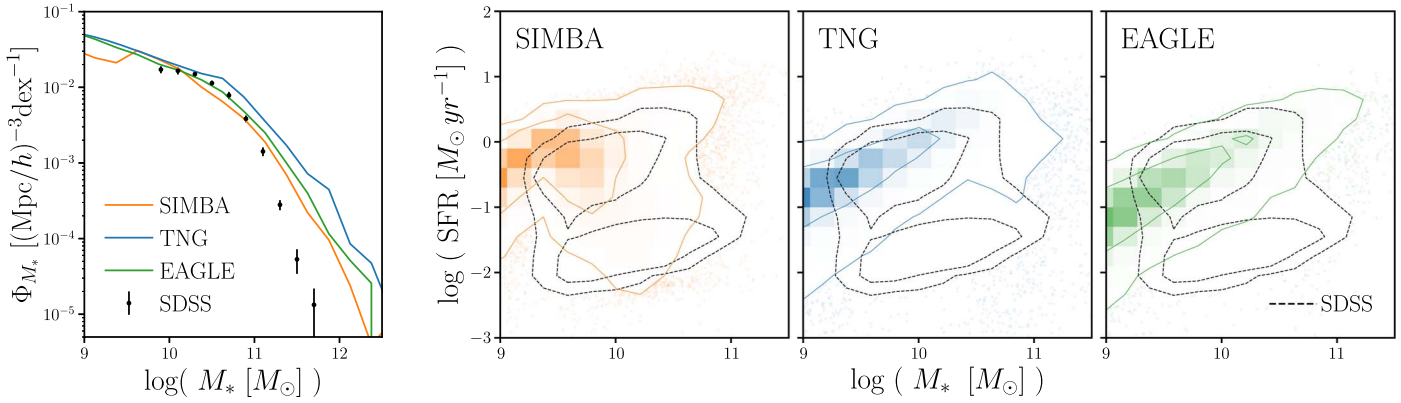

Figure 1. The SMFs, Φ_{M_*} (left-most panel), and M_*–SFR relation (right panels) of galaxies in three cosmological hydrodynamic simulations: SIMBA (orange), TNG (blue), and EAGLE (green). For reference, we include Φ_{M_*} and the M_*–SFR relation of SDSS observations (black). Uncertainties for the SDSS Φ_{M_*} are derived using jackknife resampling. We describe the simulations and observations in Section 2. Although, we see significant differences between the M_*–SFR relations of SDSS and the simulations, the SDSS measurements cannot be directly compared to the simulations' theoretical M_* and SFR predictions. A forward modeling approach, where we construct synthetic observations for the simulations, enables an apples-to-apples comparison between simulations and observation. Furthermore, differences in Φ_{M_*} and the M_*–SFR relations among the hydrodynamic simulations highlight how they predict galaxy populations with significantly different physical properties.

simulations, the M_* and SFR of simulations are theoretical predictions, while the SDSS values are measurements, which are subject to, for instance, SFR measurement limits, inconsistencies among SFR tracers, and aperture effects. This inconsistency prevents a meaningful direct comparison. However, with a forward modeling approach of constructing synthetic observations, the simulations can be directly compared to observations (see, e.g., Dickey et al. 2021, T. K. Starkenburg et al. 2021, in preparation). Figure 1 also illustrates that the hydrodynamical simulations predict significantly different SMFs and M_*–SFR relations from each other. This difference, which was also recently highlighted in Hahn et al. (2019c), demonstrates that the hydrodynamical simulations predict galaxy populations with significantly different physical properties from one another.

2.1. IllustrisTNG100

The IllustrisTNG100 simulation[16] is a cosmological hydrodynamic simulation of comoving volume $(110.7\,\mathrm{Mpc})^3$, with a particle mass resolution of $7.6 \times 10^6\,M_\odot$ for dark matter and $1.4 \times 10^6\,M_\odot$ for baryonic particles (Nelson et al. 2018; Pillepich et al. 2018; Springel et al. 2018; Marinacci et al. 2018; Naiman et al. 2018). It improves on the original Illustris simulation[17] (Vogelsberger et al. 2014; Genel et al. 2014; public data release by Nelson et al. 2015) by including magnetohydrodynamics and updated treatments for galactic winds, metal enrichment, and active galactic nucleus (AGN) feedback. Most notably, TNG uses a new implementation for feedback from supermassive black holes (SMBHs), where feedback energy is injected in the form of a kinetic AGN-driven wind at low SMBH accretion rates (Weinberger et al. 2018). This new implementation has been shown to alleviate discrepancies found between the original Illustris and observations for $> 10^{13-14}\,M_\odot$ massive halos.

2.2. EAGLE

The Virgo Consortium's EAGLE project[18] (Schaye et al. 2015; Crain et al. 2015; McAlpine et al. 2016) is a publicly available suite of cosmological hydrodynamic simulations constructed using ANARCHY (C. Dalla Vecchia et al., in preparation; 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 $(100\,\mathrm{Mpc})^3$ and a baryonic mass resolution of $1.81 \times 10^6\,M_\odot$. EAGLE has subgrid models for star formation, stellar mass loss, metal enrichment, and stellar feedback that stochastically inject thermal energy in the interstellar medium 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 $z = 0$ SMF 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, 2017b), is a cosmological hydrodynamical simulation constructed using the GIZMO meshless finite-mass hydrodynamics code (Hopkins 2015; Hopkins et al. 2018). Of the suite, we use "m100n1024", which has a box size of $(100\,h^{-1}\,\mathrm{Mpc})^3$ and baryonic mass resolution of $1.82 \times 10^7\,M_\odot$. The simulation uses some of the same subgrid models as MUFASA for H_2-based star formation, decoupled two-phase winds for star formation–driven galactic winds with velocity and mass loading calibrated to reproduce the FIRE simulations (Muratov et al. 2015; Anglés-Alcázar et al. 2017b), and feedback from Type I supernovae and AGB stars. SIMBA uses a two-mode SMBH accretion model, torque-limited accretion for cold gas (Anglés-Alcázar et al. 2017a) and Bondi-based accretion for hot gas, and AGN feedback in jet, radiative, and X-ray modes.

2.4. SDSS Galaxies

For our observations, we begin with the volume-limited Tinker et al. (2011) sample derived from the SDSS DR7 (Abazajian et al. 2009) NYU Value-Added Galaxy Catalog (Blanton et al. 2005), which has a $M_* > 10^{9.7}\,M_\odot$ completeness limit. However, rather than M_*, we focus on observables that can be consistently defined and derived in both simulations and observations: the r-band absolute magnitude,

[16] https://www.tng-project.org/
[17] http://www.illustris-project.org
[18] http://www.eaglesim.org

M_r, the optical $g-r$ color, and the far-UV − NUV color. We use far-UV (FUV), NUV, r and g band absolute magnitudes from the NASA-Sloan Atlas[19] (NSA), which is a re-reduction of SDSS DR8 (Aihara et al. 2011), which includes an improved background subtraction (Blanton et al. 2011) and NUV and FUV photometry from GALEX. These absolute magnitudes are derived using `kcorrect` (Blanton & Roweis 2007), assuming a Chabrier (2003) initial mass function.

We impose a $M_r < -20$ completeness limit on the Tinker et al. (2011) sample as well as completeness limits in the FUV and NUV bands. The `kcorrect` UV absolute magnitudes are poorly constrained for galaxies with low UV fluxes. We compare the reconstructed FUV and NUV fluxes from `kcorrect` to the measured fluxes and determine the flux limits above which the fluxes are in good agreement. The flux limits correspond to completeness limits of $M_{\rm FUV} < -13.5$ and $M_{\rm NUV} < -14.0$. In Figure 2, we present the M_*–SSFR relation of our observational sample (blue). We include the original Tinker et al. (2011) SDSS sample (black dashed line) for comparison. In total, our SDSS sample has 4451 star-forming and quiescent galaxies with $M_* \gtrsim 10^{10}\ M_\odot$.

2.5. Forward Modeling Observations

One of the main goals of this work is to conduct an apples-to-apples comparison between the simulations and observations. A crucial step in this comparison is to *forward model* the observables from the simulations. The simulations can then be directly compared to observations in observational space, instead of relying on measured galaxy properties, which are impacted by variations, inconsistencies, and biases of different methods (Dickey et al. 2021). The comparison can also include selection functions and observational systematic effects through the forward model. In this work, we use r-band luminosity (M_r), optical color ($g-r$), and UV color (FUV − NUV) as our observables.

First, we construct SEDs for all of the simulated galaxies based on their SFH and metallicity history using the flexible stellar population synthesis model (`FSPS`; Conroy et al. 2009, 2010) with the MILES spectral library (Sánchez-Blázquez et al. 2006), MIST isochrones (Paxton et al. 2011, 2013, 2015; Choi et al. 2016; Dotter 2016), and Chabrier (2003) initial mass function. For each simulated galaxy, we bin the total stellar mass formed by age (t) and metallicity (Z). We use the same t, Z grid for all of the simulations to account for the variable time and mass resolutions. We assume each (t, Z) bin is a single stellar population and generate a spectrum assuming the use of `FSPS` and take the mass-weighted linear combination of them to produce the galaxy SED. For further details on how we construct the SEDs, we refer readers to T. K. Starkenberg et al. (in preparation).

Next, we apply dust attenuation to the SEDs using the EDA prescription, which assigns dust attenuation curves to each simulated galaxy based on its physical properties and EDA model parameters (Section 3). We then convolve the attenuated SEDs with the transmission curves of the GALEX FUV, GALEX NUV, SDSS g, and SDSS r broadband filter to construct the observables. We add realistic noise to M_r, $g-r$, and FUV − NUV by sampling from the observed uncertainty distributions of the NSA. Lastly, we apply the same $M_r < -20$, $M_{\rm FUV} < -13.5$, and $M_{\rm NUV} < -14$ absolute magnitude completeness limits of our SDSS sample to the simulated galaxies.

[19] http://nsatlas.org/

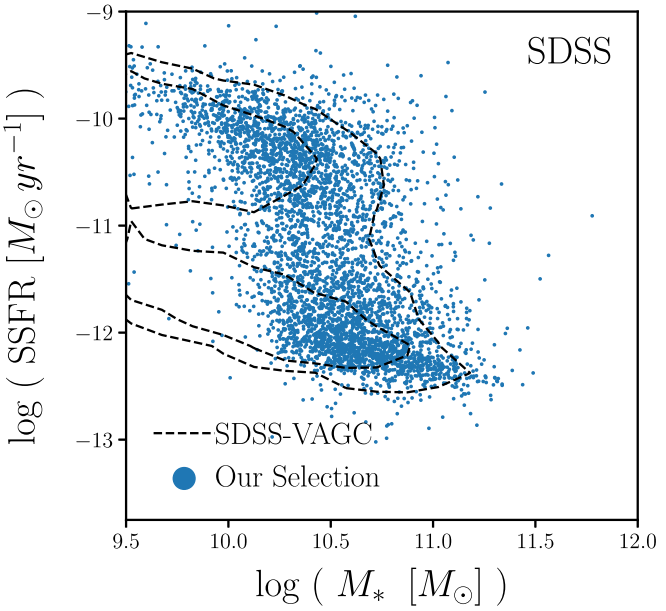

Figure 2. We derive our observational sample (blue) from the Tinker et al. (2011) SDSS sample (black dashed line) by imposing $M_r < -20$, $M_{\rm FUV} < -13.5$, and $M_{\rm NUV} < -14.0$ completeness limits. We describe the galaxy sample and completeness limits in Section 2.4. M_* is estimated using `kcorrect`, and SSFR is from Brinchmann et al. (2004). Our SDSS sample has 4451 galaxies that include both star-forming and quiescent galaxies with $M_* \gtrsim 10^{10}\ M_\odot$.

In Figure 3, we present the forward-modeled optical and UV color–magnitude relations, $(g-r)-M_r$ (top) and (FUV − NUV) − M_r (bottom) for simulated galaxies in SIMBA (left), TNG (center), and EAGLE (right) *assuming no dust attenuation*. We mark the 68% and 95% contours and include, for reference, the optical and UV color–magnitude relations of our SDSS sample (black dashed line). The comparison to SDSS observations clearly demonstrates that *without dust attenuation, the hydrodynamical simulations do not reproduce the observed optical or UV color–magnitude relations.*

3. The EDA Framework

In this section, we describe the EDA framework and present the EDA prescription used in this work to apply dust attenuation to our simulated galaxies. We begin by defining the dust attenuation curve, $A(\lambda)$, as

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

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

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

so that A_V determines the amplitude of the attenuation, while $k(\lambda)$ determines the wavelength dependence. The EDA assigns an A_V and $k(\lambda)$ for each simulated galaxy. For A_V, we use the slab model (e.g., Somerville & Primack 1999; Somerville et al. 2012), where A_V is a function of galaxy inclination, i, and galaxy properties:

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

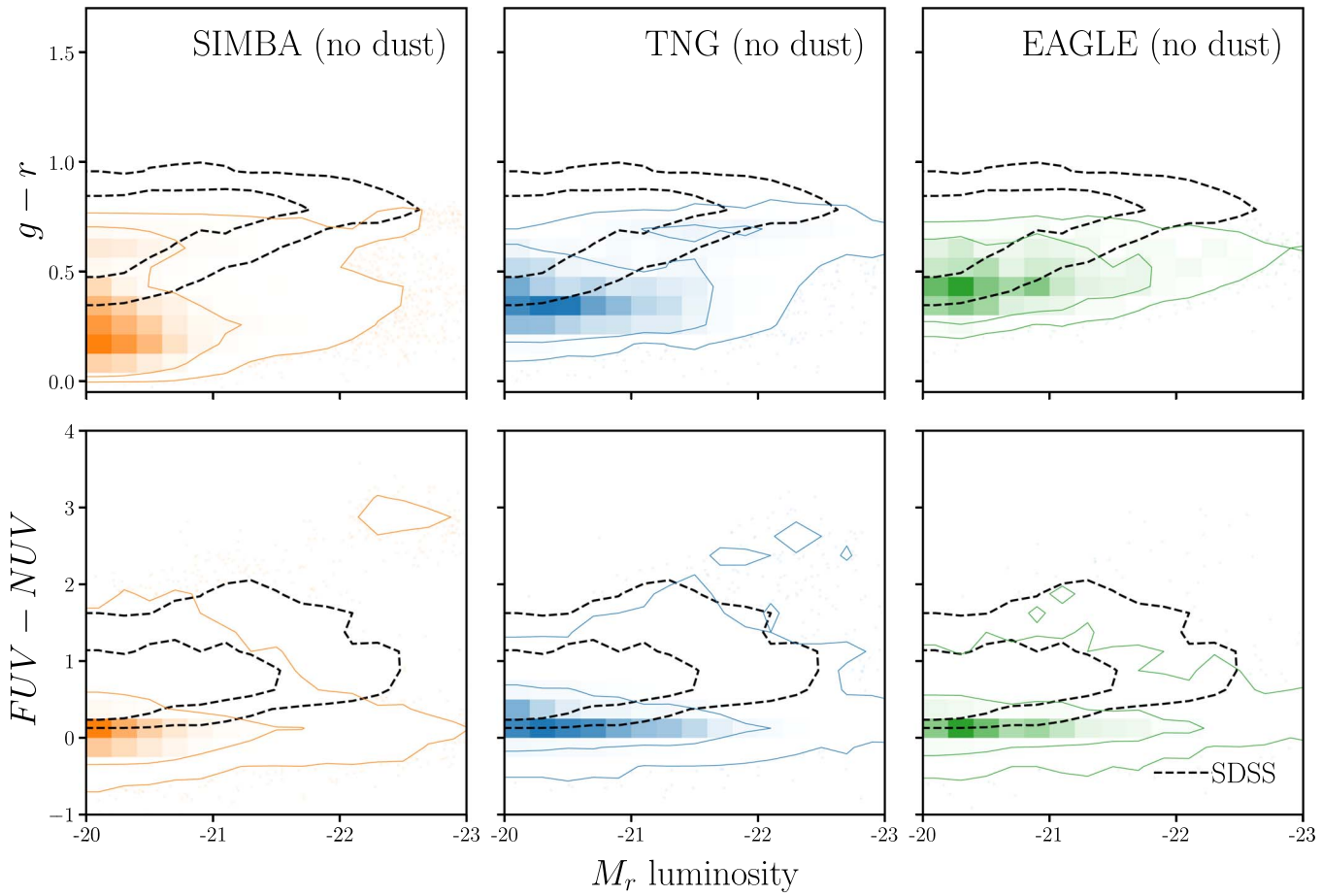

Figure 3. We present the forward-modeled optical and UV color–magnitude relations of SIMBA (left), TNG (center), and EAGLE (right) galaxies *assuming no dust attenuation*. We present $(g-r)-M_r$ in the top panels and (FUV − NUV) − M_r in the bottom panels. The contours represent the 68% and 95% of the distribution. We derive observables M_r, $g-r$, and FUV − NUV for the simulations using our forward model (Section 2.5). For comparison, we include the color–magnitude relations of our SDSS sample (black dashed line; Section 2.4). Without dust attenuation, the hydrodynamical simulations do not reproduce the SDSS optical or UV color–magnitude relations.

τ_V is the V-band optical depth that depends linearly on $\log M_*$ and logSSFR:

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

and m_{τ,M_*}, $m_{\tau,\mathrm{SSFR}}$, and c_τ represent the M_* dependence, the SSFR dependence, and amplitude of τ_V, respectively. Since τ_V is optical depth, we impose a $\tau_V \geqslant 0$ limit. For each galaxy, we uniformly sample $\cos i$ from 0 to 1 to introduce stochasticity. This produces significant variance in A_V, so galaxies with the same properties do not have identical dust attenuation.

Our τ_V parameterization is based on correlations between dust attenuation and galaxy properties that have been established by observations (e.g., Garn & Best 2010; Battisti et al. 2016; Salim & Narayanan 2020). Previous works have parameterized dust attenuation based on other galaxy properties such as gas density, gas metallicity, or star-gas geometry, motivated by the fact that dust attenuation on small scales depends on local stellar and gas properties (e.g., Somerville & Primack 1999; Somerville et al. 2012; Steinacker et al. 2013; Camps & Baes 2015; Narayanan et al. 2018; Trayford et al. 2020; Vogelsberger et al. 2020). Galaxies in the SIMBA, TNG, and EAGLE, however, have substantially different gas masses and metallicities (Davé et al. 2020, A. H. Maller et al. 2021, in preparation). If we were to parameterize τ_V using these properties, their differences would dominate any comparison of dust attenuation. In the Appendix, we confirm the correlation between A_V and the properties M_* and SSFR in the Salim et al. (2018) GSWLC2 sample (Figure 14).

In our EDA, we use the slab model because it provides a simple prescription for generating a distribution of A_V that depends on randomly sampled i, with loose physical motivations. For star-forming galaxies, which typically have disk-like morphologies, the slab model produces A_V, which is correlated with i in a way consistent with observations: edge-on galaxies have higher A_V than face-on galaxies (e.g., Conroy et al. 2010; Wild et al. 2011; Battisti et al. 2017; Salim & Narayanan 2020). Nevertheless, the slab model is a simplification. In reality, A_V depends on the detailed star-to-dust geometry. Furthermore, we assign A_V to all galaxies, not just star-forming. For quiescent galaxies, which typically have elliptical morphologies, the slab model serves only as an *empirical* prescription for statistically sampling A_V. The EDA seeks to assign an accurate distribution of dust attenuation curves for an ensemble of galaxies—*not* to accurately model dust attenuation for individual galaxies. In this regard, we demonstrate in the Appendix that the slab model can match the observed distribution of A_V, even for samples that include quiescent galaxies.

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

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

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

$$k_{\rm Cal}(\lambda) = \begin{cases} 2.659(-1.857 + 1.040/\lambda) + R_V, & 6300A \leqslant \lambda \leqslant 22000A \\ 2.659(-2.156 + 1.509/\lambda & \\ -0.198/\lambda^2 + 0.011/\lambda^3) + R_V & 1200A \leqslant \lambda \leqslant 6300A \end{cases}, \quad (6)$$

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

$$\delta(M_*, \text{SFR}) = m_{\delta,M_*} \log\left(\frac{M_*}{10^{10} M_\odot}\right) + m_{\delta,\text{SFR}} \log\left(\frac{\text{SSFR}}{10^{-10}\ \text{yr}^{-1}}\right) + c_\delta. \quad (7)$$

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

$$D(\lambda) = \frac{E_b(\lambda\ \Delta\lambda)^2}{(\lambda^2 - \lambda_0^2)^2 + (\lambda\ \Delta\lambda)^2}, \quad (8)$$

where $\lambda_0 = 2175$ Å, $\Delta\lambda = 350$ Å, and E_b are the central wavelength, FWHM, and strength of the bump, respectively. We include the UV dust bump since we use UV color as one of our observables. Kriek & Conroy (2013) and Tress et al. (2018) find that E_b correlates with δ for star-forming galaxies at $z \sim 2$. Narayanan et al. (2018) confirmed this dependence in simulations. Hence, we assume a fixed relation between E_B and δ: $E_b = -1.9\,\delta + 0.85$ (Kriek & Conroy 2013). Allowing the slope and amplitude of the E_B and δ relation to vary does *not* impact our results; however, we also do not derive any meaningful constraints on them. In Table 1, we list and describe all of the free parameters of our EDA prescription.

The SSFRs of galaxies are used to calculate τ_V and δ in Equations (4) and (7). However, due to mass and temporal resolution limits, some galaxies in the simulations have SFR = 0, that is, an unmeasurably low SFR (Hahn et al. 2019c). They account for 17%, 19%, and 9% of galaxies in SIMBA, TNG, and EAGLE, respectively. Since Equations (4) and (7) depend on logSSFR, they cannot be used in the equations to derive τ_V and δ for these galaxies. To account for this issue, we assign $\text{SFR}_{\rm min}$, the minimum nonzero instantaneous SFR in each simulation, to SFR = 0 galaxies when calculating τ_V and δ. For SIMBA, TNG, and EAGLE, $\text{SFR}_{\rm min} = 0.000816$, 0.000268, and 0.000707 $M_\odot\ \text{yr}^{-1}$, respectively. Although this assumes that SFR = 0 galaxies have similar dust properties as the galaxies with SFR = $\text{SFR}_{\rm min}$, since the simulations have very low $\text{SFR}_{\rm min}$, we expect galaxies with SFR = $\text{SFR}_{\rm min}$ to have little recent star formation and low gas mass, similar to SFR = 0 galaxies.

In summary, to apply the EDA to a simulated galaxy population, we first assign a randomly sampled i to each galaxy ($\cos i$ uniformly sampled from 0 to 1); τ_V and δ are calculated for the galaxy based on its M_*, SSFR, and the EDA parameters.

Table 1
Free Parameters of the EDA Model

Parameter	Definition	Prior
m_{τ,M_*}	M_* dependence of the optical depth, τ_V	flat [−5., 5.]
$m_{\tau,\text{SSFR}}$	SSFR dependence of τ_V	flat [−5., 5.]
c_τ	amplitude of τ_V	flat [0., 6.]
m_{δ,M_*}	M_* dependence of δ, the attenuation curve slope offset	flat [−4., 4.]
$m_{\delta,\text{SFR}}$	SSFR dependence of δ	flat [−4., 4.]
c_δ	amplitude of δ	flat [−4., 4.]

We then calculate A_V and $k(\lambda)$ to determine $A(\lambda)$ for each galaxy. Afterward, we attenuate the galaxy SEDs using Equation (1) and use the attenuated SEDs to calculate the observables: g, r, NUV, and FUV absolute magnitudes. In Figure 4, we present attenuation curves, $A(\lambda)$, generated by the EDA for galaxies with different SFR and M_* values. We present star-forming galaxies with $\{M_*, \text{SFR}\} = \{10^{10}\ M_\odot,\ 10^{0.5}\ M_\odot\ \text{yr}^{-1}\}$ (blue) and $\{10^{11}\ M_\odot,\ 10^{1}\ M_\odot\ \text{yr}^{-1}\}$ (green) and a quiescent galaxy with $\{10^{11}\ M_\odot,\ 10^{-2}\ M_\odot\ \text{yr}^{-1}\}$ (red). We use an arbitrary set of EDA parameters (m_{τ,M_*}, $m_{\tau,\text{SSFR}}$, c_τ, m_{δ,M_*}, $m_{\delta,\text{SFR}}$, c_δ) within the prior range listed in Table 1. We set $i = 0$ (edge-on) for all $A(\lambda)$ in Figure 4 for simplicity. For comparison, we include the Calzetti (2001) attenuation curve. The EDA produces attenuation curves with a wide range of amplitudes and slopes for galaxies based on their physical properties.

4. Likelihood-free Inference: Approximate Bayesian Computation

With our forward model, which includes the EDA prescription for dust attenuation, we can now generate synthetic observations for simulated galaxies and make an apples-to-apples comparison to SDSS. Next, we want to use this comparison to infer the posterior probability distribution of the EDA parameters. Typically in astronomy, this inference is done assuming a Gaussian likelihood to compare the "summary statistic" (e.g., SMF) of the model to observations and some sampling method (e.g., Markov Chain Monte Carlo) to estimate the posterior distribution. The functional form of the likelihood, however, depends on the summary statistic, and assuming an incorrect form of the likelihood can significantly bias the inferred posteriors (e.g., Hahn et al. 2019b). In this work, we use the optical and UV color–magnitude relations as our summary statistic. Since this statistic is a three-dimensional histogram, the likelihood is *not* Gaussian. Furthermore, since the bins are not independent, the true likelihood is difficult to analytically write down.

Rather than *incorrectly* assuming a Gaussian likelihood or attempting to estimate the true likelihood of the optical and UV color–magnitude relations, we use approximate Bayesian computation (hereafter ABC; Diggle & Gratton 1984; Tavare et al. 1997; Pritchard et al. 1999; Beaumont et al. 2009; Del Moral et al. 2012) for our inference. ABC is a likelihood-free (or "simulation-based") parameter inference framework that approximates the posterior probability distribution, $p(\theta \mid \text{data})$, 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

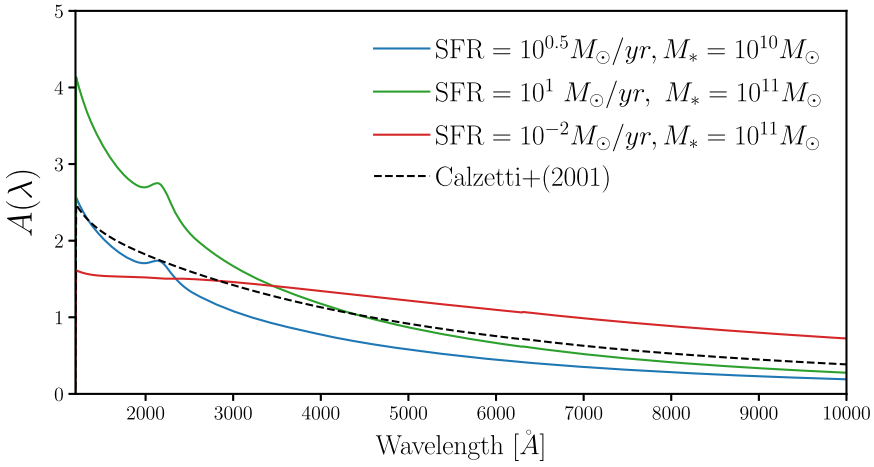

Figure 4. Attenuation curves, $A(\lambda)$, assigned by our EDA prescription to edge-on galaxies with different SFR and M_* values for an arbitrary set of EDA parameters. We include $A(\lambda)$ for star-forming galaxies with $\{M_*, \mathrm{SFR}\} = \{10^{10}\ M_\odot, 10^{0.5}\ M_\odot\ \mathrm{yr}^{-1}\}$ (blue) and $\{10^{11}\ M_\odot, 10^{1}\ M_\odot\ \mathrm{yr}^{-1}\}$ (green) and a quiescent galaxy with $\{10^{11}\ M_\odot, 10^{-2}\ M_\odot\ \mathrm{yr}^{-1}\}$ (red). We set $i = 0$ for all the galaxies in the figure for simplicity, but in practice the EDA uniformly samples $\cos i$ from 0 to 1 for each galaxy. For comparison, we include the Calzetti (2001) attenuation curve. The EDA provides a flexible prescription for assigning dust attenuation to galaxies based on their physical properties (M_* and SSFR) and the EDA parameters.

form of the likelihood, so we avoid any biases from such assumptions. It also expands the summary statistics we can use to infer the posteriors and, therefore, provides a general inference framework for a forward modeling approach.

In the simplest version of ABC, with rejection sampling (Pritchard et al. 1999), a proposal set of parameter values is drawn from the prior. The forward model is run with the proposal parameter values. The output of the forward model is then compared to the observed data using a distance metric. If the distance is within some small threshold, we keep the proposed parameters; otherwise, we discard them. Proposals are drawn until enough pass the threshold to sample the posterior. A rejection sampling framework requires a large number of evaluations of the forward model, which can be computationally costly. Many variations of ABC with more efficient sampling strategies have now been applied to astronomy and cosmology (e.g.,Cameron & Pettitt 2012; Weyant et al. 2013; Ishida et al. 2015; Lin et al. 2016; Alsing et al. 2018). Among these methods, we use ABC with population Monte Carlo (PMC) importance sampling (Hahn et al. 2017a, 2017b, 2019a).

ABC-PMC begins with an arbitrarily large threshold ϵ_1 and N proposals $\bar{\theta}_1$ sampled from the prior distribution. Each proposal is assigned a weight $w_1^i = 1/N$. Then for subsequent iterations ($n > 1$), the threshold, ϵ_n, is set to the median distance of the previous iteration's proposals. New proposals are drawn from the previous iteration's proposals perturbed by a kernel and kept if their distance is below ϵ_n. This is repeated until we assemble a new set of N proposals $\bar{\theta}_n$. The entire process is repeated for the next iteration until convergence is confirmed. We use the Python implementation of Akeret et al. (2015).[20] For further details on the ABC-PMC implementation, we refer readers to Hahn et al. (2017b) and Hahn et al. (2019a).

In this work, we use ABC-PMC with uninformative uniform priors on each of the EDA parameters and choose ranges that encompass constraints in the literature. The prior ranges of m_{τ,M_*}, $m_{\tau,\mathrm{SSFR}}$, c_τ 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_{δ,M_*}, $m_{\delta,\mathrm{SFR}}$, c_δ include the δ 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 use the forward model described in Section 2.5, where we construct SEDs for every simulated galaxy from SIMBA, TNG, and EAGLE, apply dust attenuation with our EDA, calculate the observables (M_r, $g-r$, and FUV − NUV), add realistic noise, and apply the $M_r < -20$, $M_{\mathrm{FUV}} < -13.5$, and $M_{\mathrm{NUV}} < -14$ completeness limits. We use the optical and UV color–magnitude relation, $(g-r) - M_r$ and (FUV − NUV) − M_r, as our summary statistic to fully exploit the (M_r, $g-r$, FUV − NUV) observational space. We measure the color–magnitude relations by calculating the number density in bins of ($g-r$, FUV − NUV, M_r) with widths (0.0625, 0.25, 0.5) mags. For our distance metric, ρ, we use the L2 norm between the number density of the SDSS observation, n^{SDSS}, and of our forward model, $n^{\mathrm{FM}}(\theta_{\mathrm{EDA}})$:

$$\rho(\theta_{\mathrm{EDA}}) = \sum_{i \in \mathrm{bins}} [n_i^{\mathrm{SDSS}} - n_i^{\mathrm{FM}}(\theta_{\mathrm{EDA}})]^2. \quad (9)$$

In ABC, the distance metric is only used to make the approximation: $p(\theta \mid \mathrm{data}) \approx p(\theta \mid \rho < \epsilon)$. Therefore, while an uninformative distance metric (i.e., the condition $\rho < \epsilon$ does not imply our forward model reproduces observations precisely) can result in less precise posteriors, it will be unbiased. Nevertheless, we confirm that a different distance metric (e.g., L1 norm) does *not* impact our results. In Figure 5, we present the posterior distributions of the EDA parameters derived using ABC-PMC for the SIMBA (orange), TNG (blue), and EAGLE (green). The contours mark the 68th and 95th percentiles of the distributions.

5. Results

Without dust attenuation, all of the hydrodynamical simulations struggle to reproduce the $(g-r) - M_r$ and (FUV − NUV) − M_r relations of SDSS (Figure 3). Both in the optical and UV, the simulations predict galaxies significantly bluer than SDSS galaxies. The simulations also predict optically blue luminous galaxies with $M_r < -21.5$ that are not found in the observations; this is particularly noticeable for SIMBA and TNG. Simulated galaxies in SIMBA also have a significantly broader distribution

[20] https://abcpmc.readthedocs.io/en/latest/index.html

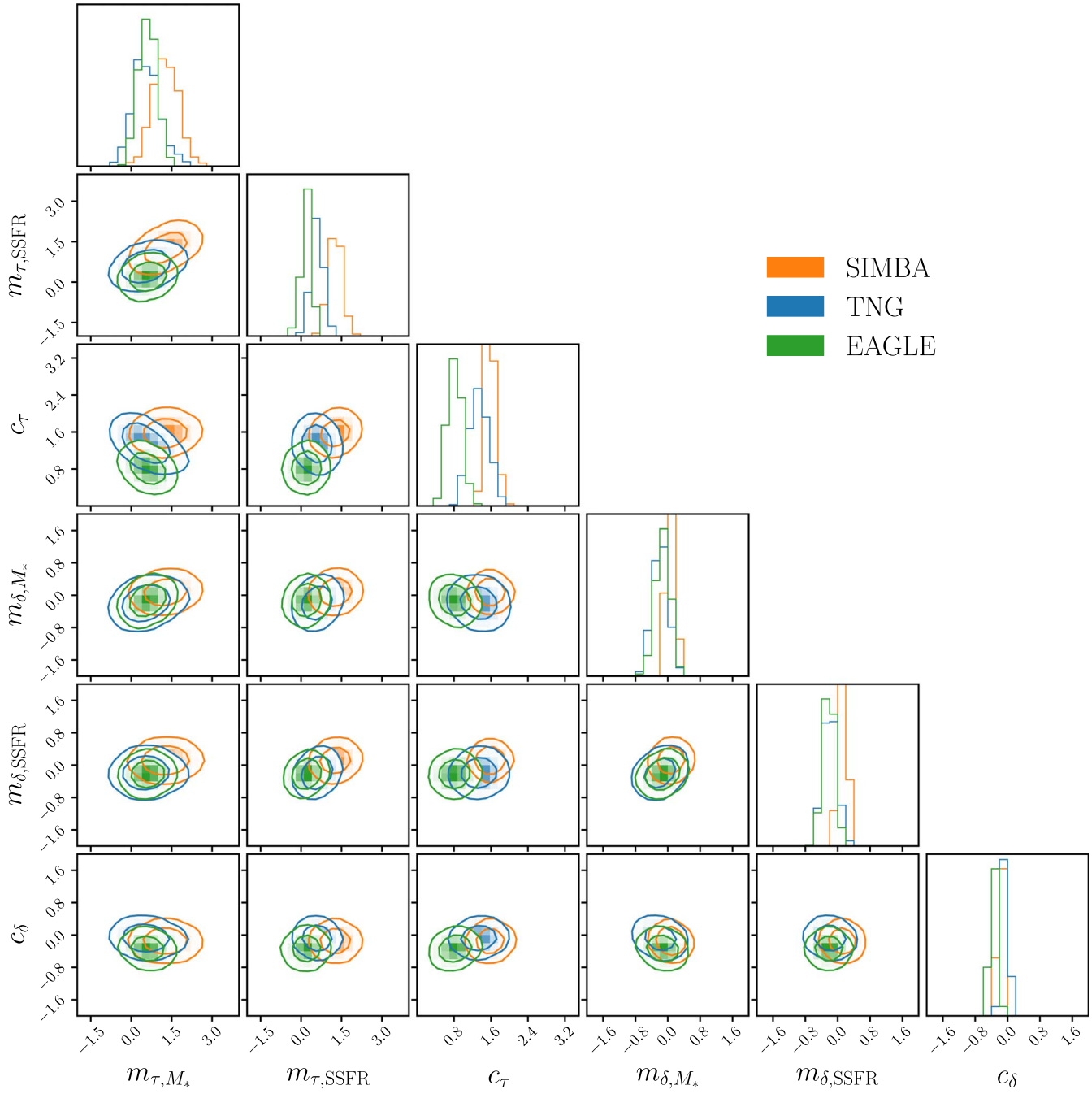

Figure 5. Posterior distributions of the EDA parameters for the SIMBA (orange), TNG (blue), and EAGLE (green) hydrodynamical simulations derived from comparing the simulations to SDSS with a forward modeling approach. The EDA parameters determine the M_* dependence, SSFR dependence, and amplitude of τ_V and δ (Table 1). The contours mark the 68th and 95th percentiles of the distributions. The posteriors are derived from likelihood-free inference using ABC-PMC (Section 4).

of $g-r$ colors than SDSS galaxies. Meanwhile, all of the simulations predict a broader distribution of FUV − NUV color than SDSS. In fact, SIMBA and TNG predict a significant number of luminous galaxies, $M_r < -22$, with FUV − NUV > 2 colors, beyond SDSS observations.

With our EDA prescription, all three simulations produce color–magnitude relations that are highly consistent with SDSS observations. In Figure 6, 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 values of the inferred posterior distributions (Figure 5). We include the color–magnitude relations of SDSS observations (black dashed line) for comparison. The contours mark the 68th and 95th percentiles of the distributions.

Dust dramatically impacts the observables of simulations. The EDA affects the optical and UV color–magnitude relations in three major ways to produce good agreement with SDSS. First, the EDA significantly reddens the simulated galaxies in the optical: $g-r$ colors are $\gtrsim 0.25$ mag redder than the optical color–magnitude relation in Figure 3 and match the $g-r$ distribution of SDSS. Second, the EDA reddens nonquiescent (logSSFR > -11) galaxies in the UV by $\gtrsim 0.5$ mag. While quiescent galaxies have

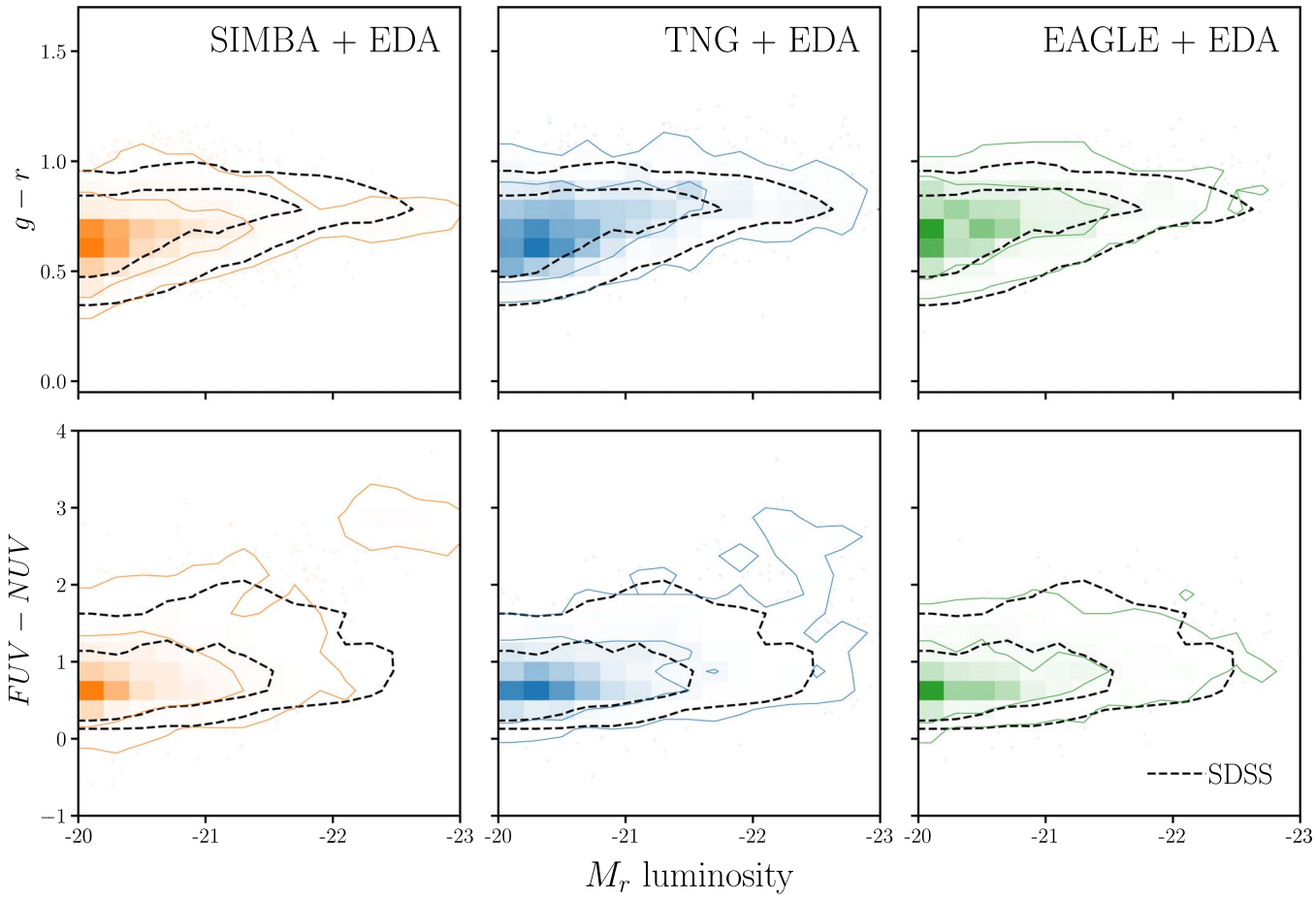

Figure 6. The optical, $(g - r) - M_r$ (top), and UV, (FUV − NUV) − M_r (bottom) color–magnitude relations predicted by our EDA prescription for the SIMBA (orange), TNG (blue), and EAGLE (green) hydrodynamical simulations. For the EDA parameters of each simulation, we use the median of the posterior distributions inferred using ABC. For comparison, we include the color–magnitude relations of SDSS (black dashed line). Comparing the color–magnitude relations above to those without dust attenuation in Figure 3, we see that dust *dramatically* impacts the color–magnitude relations. Dust attenuation must be accounted for when interpreting and comparing simulations. Furthermore, with our EDA prescription, all three simulations reproduce the color–magnitude relations of SDSS observations. Since the different simulations can reproduce observations just by varying dust, dust significantly limits our ability to constrain the underlying physical processes of galaxy formation models.

intrinsically red UV colors that generally agree with SDSS (FUV − NUV > 0.5), the rest of the galaxies are intrinsically bluer in the UV than observations. Lastly, the EDA attenuates nonquiescent galaxies so there are no longer luminous galaxies that are blue in the optical or UV—consistent with observations.

Despite the substantial improvement in the color–magnitude relation agreement with the EDA, there is still one significant discrepancy: the presence of luminous $M_r > -22$ galaxies that are UV red, not found in observations (FUV − NUV > 2). This galaxy population consists of quiescent galaxies with SFR $\lesssim 10^{-2}\ M_\odot\ \mathrm{yr}^{-1}$ and is especially pronounced in the UV color–magnitude of SIMBA but also found in TNG. They are also present in the UV color–magnitude predictions without dust attenuation (Figure 3). Since they are the most luminous galaxies in our sample, dust attenuation cannot remove them from our sample. Furthermore, dust reddening would only increase the discrepancies with observations since they have intrinsic FUV − NUV > 2. In other words, the excess luminous UV-red galaxies predicted by SIMBA and TNG are irreconcilable with dust attenuation.

In order to understand the origin of the luminous UV-red galaxies in SIMBA (left) and TNG (right), we examine their star formation histories in Figure 7. The top and center panels mark the luminous UV-red galaxies on the optical and UV color–magnitude relations predicted by the EDA (red). The bottom panels present the median specific SFH (SSFH), SSFR as a function of lookback time, t_{lookback}, with the shaded regions representing the 68th percentile. For comparison, we include the SSFHs of other quiescent galaxies with matching luminosities, SSFR $< 10^{-11}\ \mathrm{yr}^{-1}$ and $M_r < -22$ (orange). The SSFHs reveal that, unlike other quiescent galaxies, the luminous UV-red galaxies of SIMBA and TNG have almost no star formation within the last $t_{\mathrm{lookback}} \lesssim 2$ Gyr. With no recent star formation contributing to the SED in FUV wavelengths, these galaxies have red FUV − NUV color. These luminous UV-red galaxies in SIMBA and TNG suggest that star formation quenching is too efficient in the most massive quiescent galaxies.

The SSFHs in Figure 7 also reveal that luminous UV-red galaxies in SIMBA have a substantially different SSFH than other quiescent galaxies. In addition to the lack of recent star formation, the luminous UV-red galaxies also have significantly shorter star-forming timescales. They peak their star formation earlier than other quiescent galaxies, at $t_{\mathrm{lookback}} \sim 11$ Gyr, and have a more rapid decline in star formation. In contrast, the luminous UV-red galaxies in TNG have overall similar SSFHs

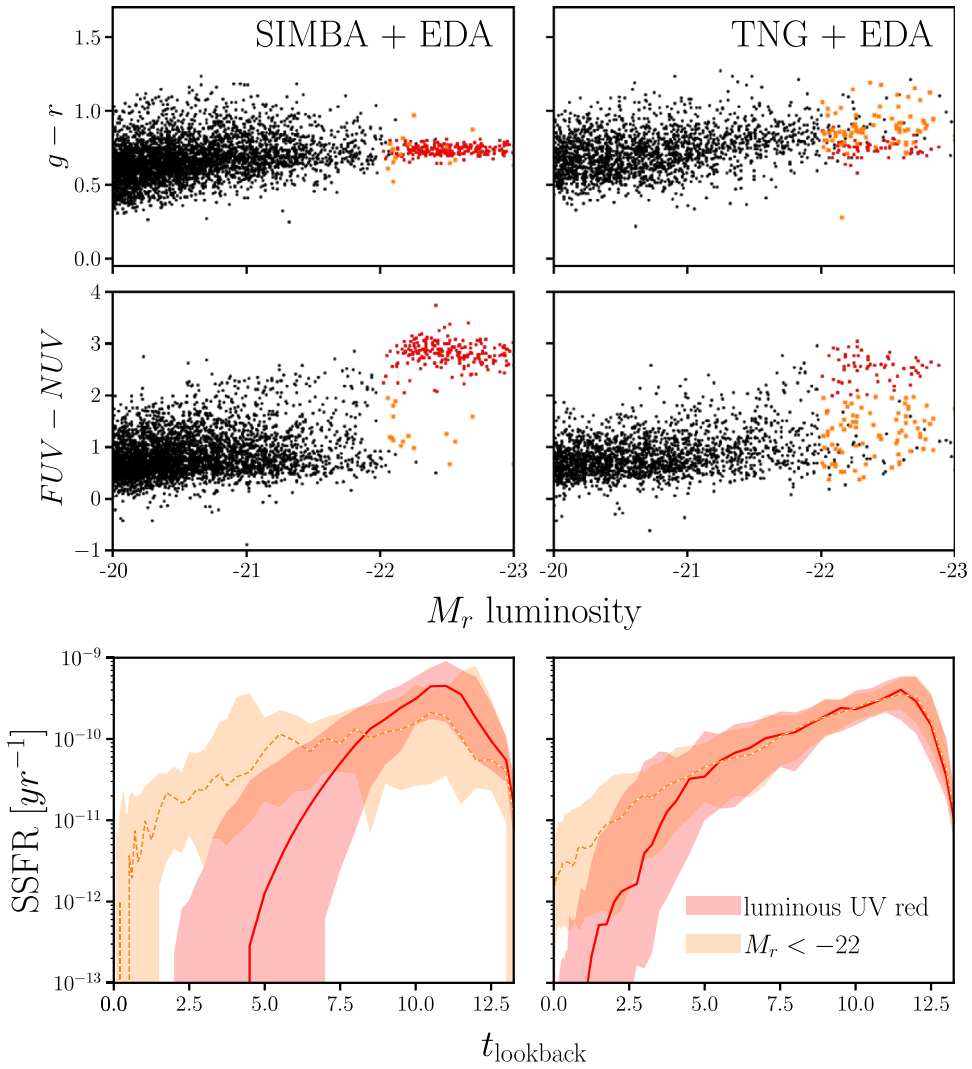

Figure 7. The SFHs of luminous UV-red galaxies (red) in SIMBA (left) and TNG (right) that cause the discrepancy between the color–magnitude relations predicted by the EDA and SDSS observations. We include the SFHs of quiescent galaxies with matching luminosities, $M_r < -22$, for comparison (orange). The top and center panels mark the luminous UV-red and the other quiescent $M_r < -22$ galaxies in the EDA predicted optical and UV color–magnitude relations, respectively. In the bottom panels, we present the median SSFH of these galaxies with the shaded regions representing the 68th percentiles of the SSFH. In both TNG and SIMBA, the luminous UV-red galaxies have negligible star formation within the last 2 Gyr, unlike the other quiescent galaxies. This suggests that SIMBA and TNG may be predicting luminous UV-red galaxies not found in observations because their prescription for star formation quenching is too efficient in the most massive galaxies.

to other quiescent galaxies. This difference in SFH suggests that a distinct star formation quenching mechanism is responsible for the luminous red galaxies in SIMBA. In another paper of the IQ series (J. Choi et al. in preparation), we examine this SFH difference in further detail and present its impact on the quiescent fraction evolution over $0 < z < 3$.

In principle, the simulations could remove the luminous UV-red galaxies from the sample if star formation quenching in these galaxies was even more efficient. With absolutely no recent star formation, their UV SED would fall below our $M_{\rm FUV}$ or $M_{\rm NUV}$ completeness limits and would not be included in the comparison. However, a large portion of the luminous quiescent galaxies are UV red, so the simulations, especially SIMBA, would struggle to reproduce the luminous end of the SDSS red sequence and the luminosity function of the observations. Furthermore, reducing the total number of massive quiescent galaxies that pass the selection function would conflict with our observations, which include a significant number of massive quiescent galaxies (Figure 2).

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) calculated 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 in the UV regime and produce significantly higher FUV − NUV color than GAMA.

The EDA for EAGLE 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 with SDSS. We find neither of these discrepancies for the TNG+EDA. The EDA provides a simpler empirical framework for applying dust attenuation than the dust models in these works. Yet, with its flexibility and low computation cost, we are able to fully explore our dust parameters and produce optical and UV color–magnitude relations that are in good agreement with observations.

5.1. Comparison to Dust Observations

In addition to reproducing observations, the EDA assigns dust attenuation curves to each simulated galaxy so we can compare the EDA attenuation curves to dust attenuation measured from observations. We begin with the well-established attenuation-slope relation: star-forming galaxies with higher dust attenuation have shallower attenuation curves. This relation is a consequence of dust scattering dominating absorption at low attenuation, while dust absorption dominates at high attenuation (Gordon et al. 1994; Witt & Gordon 2000; Draine 2003; Chevallard et al. 2013). In Figure 8, we present the attenuation-slope relation of star-forming galaxies with $\mathrm{SSFR} > 10^{-11}\,\mathrm{yr}^{-1}$ using the dust attenuation curves predicted by the EDA for the median posteriors of SIMBA (left), TNG (center) and EAGLE (right). For comparison, we include the observed attenuation-slope relations of GSWLC2 galaxies (gray shaded; Salim & Narayanan 2020). For attenuation, we use A_V; for slope, we use the UV-optical slope $S = A(1500\,\mathring{A})/A_V$. The contours mark the 68th and 95th percentiles.

Most noticeably, we find that the EDA does not predict $A_V < 0.3$ for star-forming galaxies. This is a consequence of SIMBA, TNG, and EAGLE predicting star-forming galaxies that are intrinsically more luminous than observations. All of the simulations have star-forming galaxies with intrinsic $M_r < -21$ and $g - r < 0.5$ (Figure 3). This is further corroborated by the $\mathrm{SFR} - M^*$ relations in Figure 1, where the simulations all have star-forming galaxies with $M_* > 10^{11}\,M_\odot$, not found in SDSS. To reproduce the SDSS optical color–magnitude relation, these galaxies would need to be significantly reddened and attenuated so any dust prescription would need to assign high A_V to star-forming galaxies. Nevertheless, for $A_V > 0.3$, we find good agreement between the attenuation-slope relation predicted by the EDA and observations. We refrain from further scrutinizing on A_V values since they can vary *significantly* between different observational measurements—even for the same galaxy. SDSS star-forming galaxies, for instance, have significantly higher $A_V > 0.3$ according to the Brinchmann et al. (2004) measurements (Appendix).

In addition to the attenuation-slope relation, we can also directly compare the attenuation curves predicted by the EDA to measurements from observations for star-forming galaxies. In Figure 9, we present the normalized attenuation curves of star-forming galaxies predicted by the EDA for the median posterior parameter values of SIMBA (orange), TNG (blue), and EAGLE (green). We again define galaxies with $\mathrm{SSFR} > 10^{-11}$ yr^{-1} as star forming. The attenuation curves are normalized at 3000 Å, and we present the variation in the attenuation curves in the shaded region, 68th percentile. For comparison, we include $A(\lambda)/A(3000\,\mathring{A})$ from the Narayanan et al. (2018) radiative transfer simulation (dashed line) and observations (Salim et al. 2018, dotted line). The attenuation curve from Salim et al. (2018) corresponds to star-forming galaxies with $M_* > 10^{10.5}\,M_\odot$, a similar M_* range as our forward-modeled samples. Since we do not vary the UV bump in our EDA prescription, we ignore any discrepancies in the amplitudes of the bump. Overall, we find good agreement between the EDA attenuation curves for star-forming galaxies and the attenuation curves from observations and simulations.

5.2. The Attenuation Curves of Quiescent Galaxies

In addition to star-forming galaxies, the EDA also predicts dust attenuation of quiescent galaxies. This is particularly valuable since dust attenuation in quiescent galaxies is still poorly constrained by observations due to challenges in directly measuring it from observations. For instance, methods that rely on IR luminosities can be contaminated by mid-IR emission from AGN heating nearby dust (Kirkpatrick et al. 2015). SED fitting methods must also account for AGN MIR emission (Salim et al. 2016; Leja et al. 2018; Salim et al. 2018). They also struggle to tightly constrain dust attenuation for quiescent galaxies due to the degeneracies with SFH and metallicity. With a forward modeling approach, we circumvent these challenges. Instead, we derive the attenuation curves necessary for the simulated quiescent population to reproduce the observed optical and UV photometry.

In Figure 10, we present the attenuation curves of quiescent galaxies predicted by the EDA for the median posterior parameter values of SIMBA (left), TNG (center), and EAGLE (right). Quiescent galaxies are classified using an $\mathrm{SSFR} < 10^{-11}$ yr^{-1} cut. Unlike Figure 9, we do not normalize the attenuation curves at 3000 Å. For comparison, we include $A(\lambda)$ of star-forming galaxies predicted by the EDA for the corresponding simulation in each panel (dotted line). In all three simulations, quiescent galaxies have attenuation curves with lower amplitudes and shallower slopes than star-forming galaxies. We note that for SIMBA, the attenuation curve is flat even though the inferred $m_{\delta,\mathrm{SSFR}} > 0$ (Table 2) because of the low amplitude of the attenuation curve for quiescent galaxies.

We predict $A(\lambda)$ with lower amplitude because quiescent galaxies in SIMBA, TNG, and EAGLE are intrinsically only slightly more luminous than observations. In the top panels of Figure 3, we see that the most luminous galaxies with the highest $g - r$ color are < 0.5 mag brighter than the tip of the red sequence in the SDSS color–magnitude relation. For SIMBA+EDA, where we predict $A(\lambda) \sim 0$, the most luminous and optically red galaxies have comparable M_r as the tip of the SDSS red sequence. In contrast, the most luminous blue, star-forming, galaxies are > 1 mag brighter than the luminous end of the SDSS blue cloud. Despite having lower attenuation than star-forming galaxies, in TNG and EAGLE we predict significant dust attenuation in quiescent galaxies, $A_V > 0.5$. Although this could be because TNG and EAGLE produce quiescent galaxies that are intrinsically too luminous, the presence of dust attenuation in quiescent galaxies, which is typically neglected, has significant implications. For instance, it strengthens the evidence for the UV upturn phenomenon, the unexpected detections of UV flux in quiescent galaxies (e.g., Code 1969; O'Connell 1999; Le Cras et al. 2016; Ali et al. 2018; Dantas et al. 2021). Constraints on the attenuation in

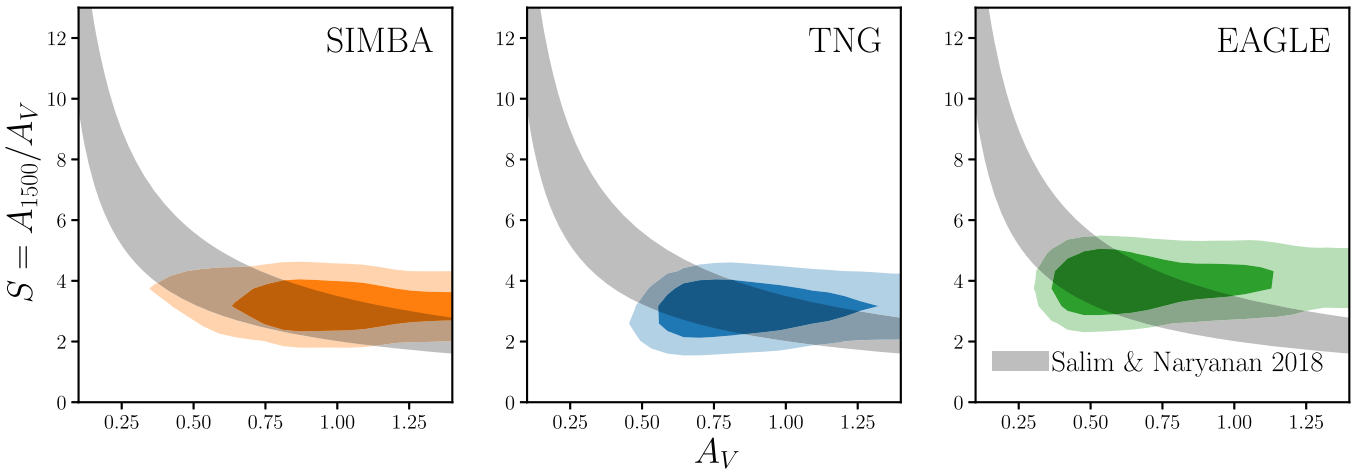

Figure 8. The attenuation-slope relation of star-forming galaxies (SSFR $> 10^{-11}$ yr^{-1}), using the attenuation curves predicted by our EDA prescription for the median posterior parameter values of SIMBA (left), TNG (center), and EAGLE (right). For comparison, we include the observed attenuation-slope relation from GSWLC2 (Salim & Narayanan 2020). We use A_V and $S = A(1500\ \text{Å})/A_V$ as measurements of attenuation and slope, respectively. *The EDA does not predict* $A_V < 0.3$ *because star-forming galaxies in the simulations are intrinsically too luminous and require significant attenuation to match observations.* Beyond $A_V > 0.3$, however, there is good agreement between the attenuation-slope relation predicted by the EDA and observations.

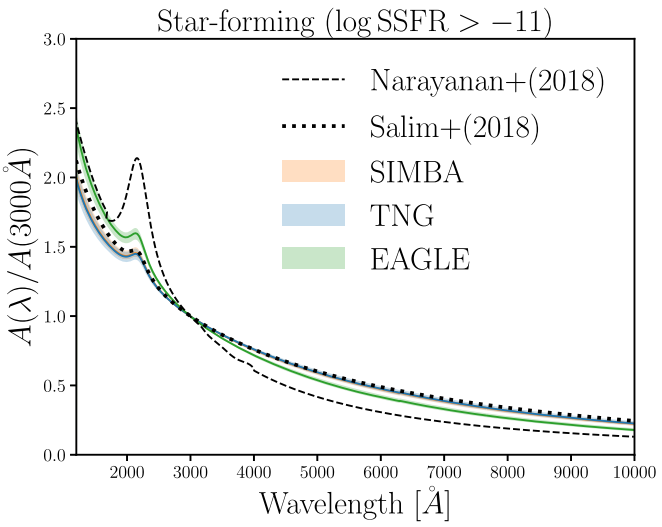

Figure 9. The normalized attenuation curves of star-forming galaxies predicted by the EDA for median posterior parameter values of SIMBA (orange), TNG (blue), and EAGLE (green). We classify galaxies with SSFR $> 10^{-11}$ yr^{-1} as star forming. The attenuation curves are normalized at 3000 Å and we mark the 68th percentile of the attenuation curves with the shaded region. For comparison, we include $A(\lambda)/A(3000\ \text{Å})$ measurements from the Narayanan et al. (2018) radiative transfer simulation (dashed line) and Salim et al. (2018) observations (dotted line). The EDA predicted attenuation curves of star-forming galaxies are in good agreement with the attenuation curves measured from the simulation and observations in the literature.

quiescent galaxies may help discern among the different hypotheses: residual star formation activity (e.g., Kaviraj et al. 2007), post-main-sequence stellar evolutionary phases (e.g., Yi et al. 1997), or binary systems (e.g., Han et al. 2007). Since the attenuation curves of quiescent galaxies are difficult to measure from observations, the EDA predictions highlight the advantages of forward modeling and its complementarity with standard approaches.

In Figure 10, we also find that quiescent galaxies have shallower attenuation curves than star-forming galaxies. This is because SIMBA, TNG, and EAGLE all predict intrinsically UV-red galaxies that do not require significant reddening. This is especially true for SIMBA and TNG, which predict a significant number of galaxies with intrinsic FUV − NUV > 1 (Figures 3 and 7). These galaxies are quiescent (SSFR $< 10^{-11}$ yr^{-1}) and have high FUV − NUV due to a lack of recent star formation contributing to the SED. When we examine their SFHs, we find that, although they have more star formation than the luminous UV-red galaxies discussed earlier, they have little star formation in the last 1 Gyr. This implies that SIMBA, whose quiescent galaxies have the shallowest attenuation curve, has the most efficient star formation quenching among the simulations.

The mass resolution of the simulations can impact the SFHs of quiescent galaxies and, thus, their observables. The SFHs of simulated galaxies cannot include any star formation below the resolution limit, which can affect the SEDs we compute from them. For recent star formation, this can have a significant impact on the SED, especially in the FUV and NUV (Leja et al. 2017). SIMBA, TNG, and EAGLE have mass resolutions of $M_{\rm lim} = 1.82 \times 10^7 M_\odot$, $1.4 \times 10^6\ M_\odot$, and $1.81 \times 10^6\ M_\odot$, respectively. Even if we were to include a < 100 Myr old stellar population with total mass of $M_{\rm lim}$ in their SFHs, the impact on FUV − NUV is small: $\lesssim 0.1$ mag for SIMBA and $\lesssim 0.01$ mag for TNG and EAGLE. Hence, mass resolution does not significantly impact the dust attenuation we predict for quiescent galaxies.

Despite the advantages of our forward modeling approach in deriving dust attenuation curves for quiescent galaxies, we caution readers that we only vary the EDA parameters in this work. The EDA predictions we present assume that the simulations accurately model the star formation and metallicity histories of quiescent galaxies. Shortcomings in the galaxy formation models, and not the dust attenuation, may be responsible for some of the differences between the simulations and observations. For instance, we already find that quenching is too efficient in certain SIMBA and TNG quiescent galaxies, which produce luminous UV-red galaxies not found in SDSS (Section 5). TNG and EAGLE may also be producing quiescent galaxies that are overall intrinsically too luminous, which then requires significant dust attenuation to match observations. Our results may also be impacted by our choice and limitations in stellar population synthesis since different stellar libraries produce different UV

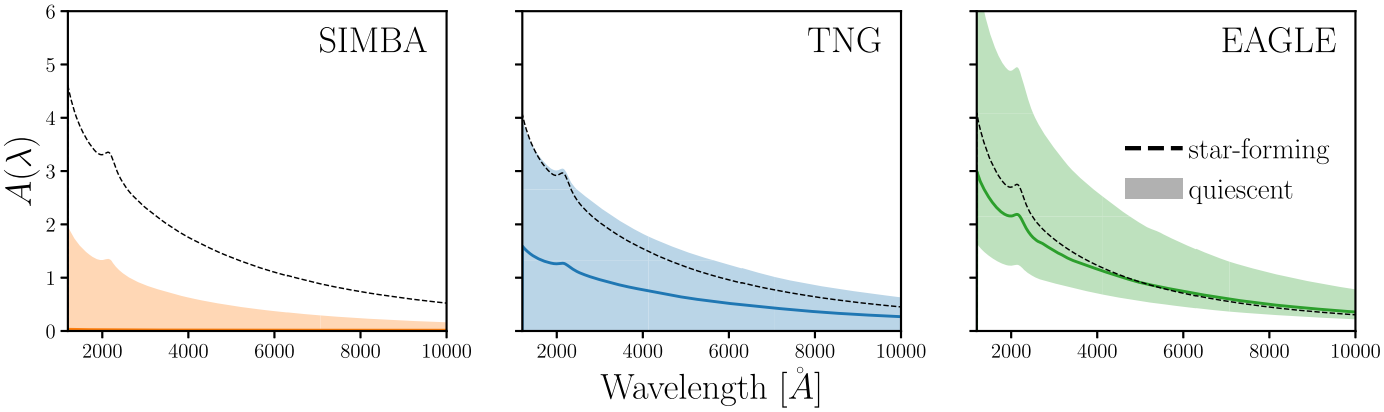

Figure 10. The attenuation curves of quiescent galaxies predicted by the EDA for median posterior parameter values of SIMBA (left), TNG (center), and EAGLE (right). Galaxies with SSFR $< 10^{-11}$ yr^{-1} are classified as quiescent. We mark the 68th percentile of the attenuation curves with the shaded region and include the predicted attenuation curves of star-forming galaxies for comparison (dashed line). In all three simulations, *the EDA predicts attenuation curves that have lower amplitudes and shallower slopes than star-forming galaxies.*

Table 2
Inferred EDA Model Parameters

	m_{τ,M_*}	$m_{\tau,\mathrm{SSFR}}$	c_τ	m_{δ,M_*}	$m_{\delta,\mathrm{SSFR}}$	c_δ
SIMBA	$1.27^{+0.46}_{-0.46}$	$1.28^{+0.24}_{-0.23}$	$1.58^{+0.12}_{-0.12}$	$0.07^{+0.12}_{-0.11}$	$0.13^{+0.10}_{-0.10}$	$-0.18^{+0.04}_{-0.04}$
TNG	$0.57^{+0.44}_{-0.53}$	$0.62^{+0.21}_{-0.20}$	$1.34^{+0.19}_{-0.21}$	$-0.18^{+0.20}_{-0.19}$	$-0.19^{+0.15}_{-0.16}$	$-0.07^{+0.08}_{-0.08}$
EAGLE	$0.59^{+0.33}_{-0.33}$	$0.18^{+0.20}_{-0.17}$	$0.81^{+0.14}_{-0.15}$	$-0.13^{+0.17}_{-0.18}$	$-0.22^{+0.14}_{-0.14}$	$-0.34^{+0.08}_{-0.08}$

fluxes. In principle, we can vary both the EDA parameters and the parameters of the galaxy formation models and infer them simultaneously with a forward modeling approach using different stellar population synthesis models. We will explore this in future work.

5.3. The Galaxy–Dust Connection

In this section, we examine the connection between dust attenuation and galaxy properties in more detail. First, we examine the galaxy-dust connection using the M_*- and SSFR-dependent parameterization of our EDA prescription (Equations (4) and (7)). In Table 2, we list the median values and the 68% confidence interval of the inferred EDA parameter posteriors for the three simulations. We find significant SSFR dependence, $m_{\tau,\mathrm{SSFR}} > 0$, consistent with what we found in the previous section: quiescent galaxies have lower dust attenuation than star-forming galaxies. We also find significant M_* dependence in τ_V: $m_{\tau,M_*} > 0$. V-band dust attenuation is higher for more massive galaxies. There is, however, little M_* dependence in the slope of the dust attenuation.

We take a closer look at the M_* and SSFR dependence of the attenuation curve in Figure 11, where we present dust attenuation at 1500 Å (A_{1500}; top) and 5500 Å (A_V; bottom) as a function of $\log M_*$ and logSSFR predicted by the EDA for SIMBA (left), TNG (center), and EAGLE (right). For each hexbin, the color map represents the median attenuation for all simulated galaxies in the bin. We only include bins with more than 10 galaxies. We include, for reference, the $M_* - \mathrm{SSFR}$ relation of all galaxies in the simulations in black dashed contours, which mark the 68th and 95th percentiles. We do not include a direct comparison with observations because there are large variations between different measurements (Appendix, see also Figure 15).

In each panel, we find that SIMBA, TNG, and EAGLE galaxies with higher M_* have higher dust attenuation at constant SSFR—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 found higher attenuation in more massive SDSS star-forming galaxies. Most recently, Salim et al. (2018) found higher V and FUV attenuation for more massive star-forming galaxies in GSWLC2. For the SSFR dependence, we find that galaxies with higher SSFR have higher A_{1500} (top) and A_V (bottom) as well as steeper slopes. The SSFR dependence is not as prominent in EAGLE (see also Table 2), which has a narrower SSFR distribution than SIMBA and TNG with no starburst galaxies or quiescent galaxies with SSFR $< 10^{-12}$ yr^{-1}. EAGLE has fewer intrinsically luminous star-forming galaxies or UV-red galaxies (Figure 3) and a narrower intrinsic $g - r$ and FUV $-$ NUV color distributions. To reproduce observations, it requires an overall attenuation and reddening without a significant SSFR dependence. Nevertheless, in all simulations, star-forming galaxies have slopes that are consistent with observations (Section 5.1), while quiescent galaxies with the lowest SSFR have nearly flat attenuation curves. In summary, we find that SIMBA, TNG, and EAGLE galaxies with higher M_* require overall higher dust attenuation and galaxies with higher SSFR require steeper attenuation curves.

5.4. Discussion

We make a number of assumptions and choices in our EDA prescription. First, we use the slab model (Equation (3)) to assign A_V as a function of $\tau_V(M_*, \mathrm{SSFR})$ and randomly sampled i. This choice is motivated by the fact that the slab model reproduces the correlation between attenuations and inclination found in star-forming galaxies from observations (Conroy 2010; Wild et al. 2011; Battisti et al. 2017; Salim & Narayanan 2020) and

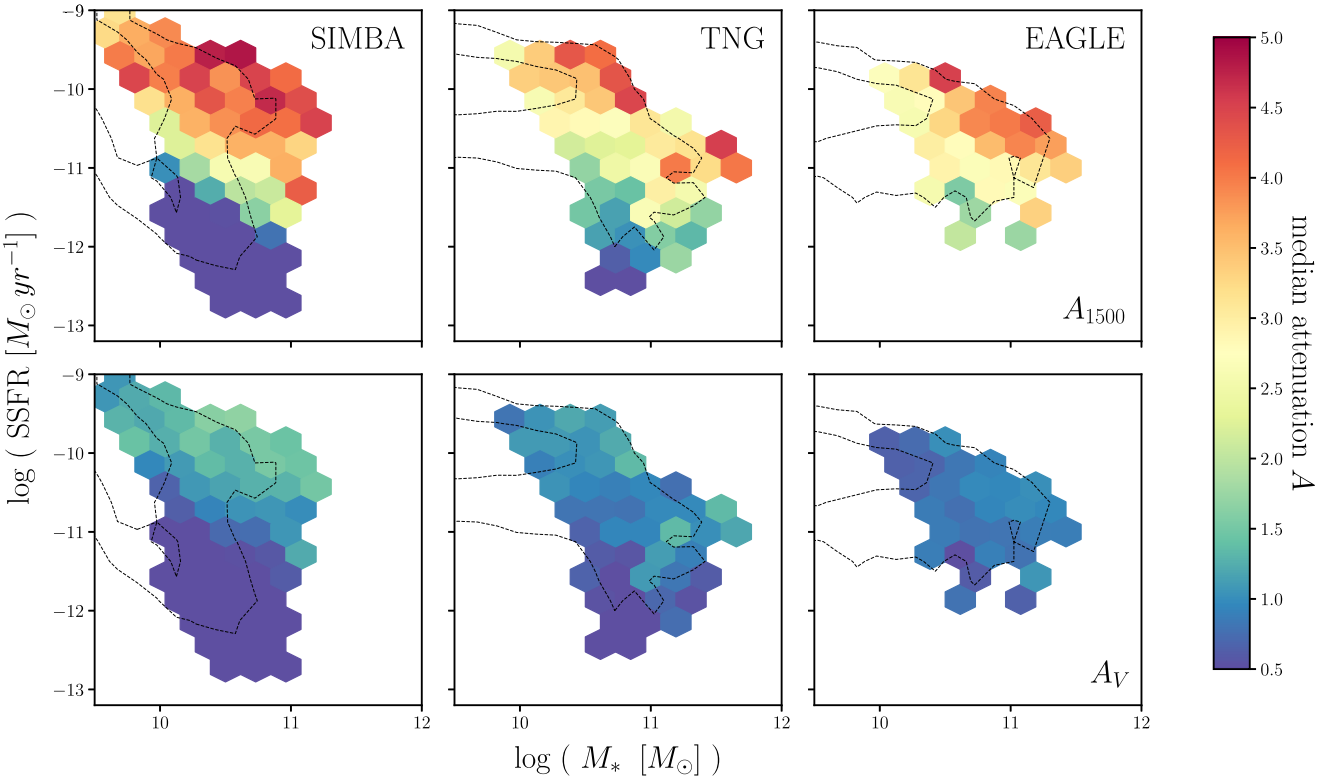

Figure 11. M_* and SSFR dependence of dust attenuation at 1500 Å (A_{1500}; top) and at 5500 Å (A_V bottom) predicted by the EDA for SIMBA (left), TNG (center), and EAGLE (right). The color map in each hexbin represents the median attenuation for all simulated galaxies in the bin (right color bar). We only include bins with more than 10 galaxies. For reference, we include in each panel the M_* – SSFR relation of all galaxies from the simulations (black dashed line). Overall, simulated galaxies with higher M_* have higher dust attenuation at constant SSFR—consistent with the literature. Furthermore, since previous works have primarily focused on star-forming galaxies, the EDA provides new insight into the SSFR dependence of dust attenuation: simulated galaxies with higher SSFR have steeper attenuation curves.

simulations (e.g., Chevallard et al. 2013; Narayanan et al. 2018; Trayford et al. 2020). More importantly, we use the slab model because it is flexible enough to reproduce a wide range of A_V distributions. In the Appendix, we demonstrate that we can reproduce $p(A_V)$ of SDSS star-forming galaxies as well as the GSWLC2 sample, which includes quiescent galaxies. We also use a parameterization of τ_V and δ that depends linearly on $\log M_*$ and logSSFR. While the M_* and SSFR dependence of A_V is well motivated and is found in, for instance, the Salim et al. (2018) GSWLC2 catalog (Appendix), the linear dependence was chosen primarily for its simplicity. The EDA framework can be easily extended to more flexible parameterizations. Although we already find good agreement with SDSS observations, a more flexible parameterization can produce even better agreement with the SDSS color–magnitude relations. The main challenges for a more flexible parameterization would be model selection and finding a better motivated parameterization.

We demonstrate in this work that accounting for dust attenuation is essential when comparing simulations to observations. After all, none of the simulations reproduce the UV and optical color–magnitude relation without dust (Figure 3). Moreover, this work also demonstrates that our current lack of understanding of dust limits our ability to closely compare galaxy formation models. Our EDA prescription is built on what we currently know about dust attenuation in galaxies: for example, the Noll et al. (2009) parameterization, the UV bump, the M_*, SSFR dependence, and so on. Yet, with the EDA, simulations that predict galaxy populations with significantly different physical properties (Figure 1) can reproduce the same SDSS observations. For instance, SIMBA has significantly fewer massive galaxies above $M_* > 10^{11}\ M_\odot$ than TNG or EAGLE (see SMFs in Figure 1). It also has $M_* < 10^{10}\ M_\odot$ starburst galaxies with SSFR $> 10^{-9.5}\,\mathrm{yr}^{-1}$ (see also Davé et al. 2019 and Figures 5 and 6) that are not found in TNG or EAGLE (Figure 11). Meanwhile, star-forming galaxies in TNG have significantly higher SFRs than in EAGLE. Despite all these differences, by adjusting dust attenuation, the simulations consistently reproduce the observations.

All 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 are a major bottleneck for investigating galaxy formation with simulations. In the next paper of the series, T. K. Starkenburg et al. (in preparation), we will examine whether we can compare the prescriptions for star formation quenching in different galaxy formation models once we include the EDA framework.

Fortunately, there are many avenues for improving our understanding of dust with a forward modeling approach. In this work, we used a restrictive SDSS galaxy sample with a $M_r < -20$, $M_{\mathrm{FUV}} < 13.5$, and $M_{\mathrm{NUV}} < -14$ completeness limit. This selection excludes star-forming galaxies below $M_* \lesssim 10^{10}\ M_\odot$ and quiescent galaxies below $M_* \lesssim 10^{10.5}\ M_\odot$ (Figure 11). Instead of imposing this completeness limit, we can include the actual SDSS selection function in the forward model (e.g., Dickey et al. 2021). This would allow us to compare the simulations with EDA to the entire SDSS sample, a substantially larger sample with a wider range of galaxies. Upcoming surveys, such as the Bright Galaxy Survey (BGS) of

the Dark Energy Spectroscopic Instrument (DESI; DESI Collaboration et al. 2016; Ruiz-Macias et al. 2021) and the galaxy evolution survey of the Prime Focus Spectrograph (PFS; Takada et al. 2014; Tamura et al. 2016), will vastly expand galaxy observations. BGS, for instance, will measure $10\times$ the number of galaxy spectra as SDSS out to $z\sim0.4$ and with its $r\sim20$ magnitude limit will probe a broader range of galaxies. Such observations will allow us to place tighter constraints on the EDA parameters, which may enable comparisons of the underlying galaxy formation models and shed light on dust in a broader range of galaxies.

In this work, we only use observables derived from UV and optical photometry, which means that we have only examined one side of the impact that dust has on galaxy spectra. While dust attenuates light in the optical and UV, it emits light in IR. In fact, even though the simulations reproduce the same SDSS UV and optical color–magnitude relations with the EDA, they predict significantly different dust emission in the IR. In Figure 12, we present IR dust emission luminosity, $L_{\rm IR}$, predicted by the EDA with median parameter values of the SIMBA (orange), 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—that is, all starlight attenuated by dust is reemitted in the IR (da Cunha et al. 2008). Most noticeably, SIMBA and TNG have bimodal distributions of dust emission, while EAGLE only has luminous IR dust emissions. This is because EAGLE requires significant dust attenuation in all galaxies, while SIMBA and TNG require quiescent galaxies to have substantially lower dust attenuation than star-forming galaxies (Figure 10). The luminous mode of the $L_{\rm IR}$ distributions, however, is in good agreement for all simulations, since they all predict comparable dust attenuation in star-forming galaxies (Figures 8 and 10). When we compare the IR dust emission of SIMBA+EDA and TNG+EDA further, we find that TNG +EDA produces more luminous galaxies with high IR dust emission ($M_r<-22$ and $L_{\rm IR}>10^9\ L_\odot$) since it has more intrinsically luminous star-forming galaxies (Figure 3). On the other hand, SIMBA+EDA has more luminous galaxies with fainter IR dust emissions, which correspond to the anomalous luminous UV-red galaxies highlighted in Figure 7.

While dust attenuation can be adjusted to reproduce UV and optical observations, since IR dust emission measures the total attenuation, IR observations can place a limit on the total impact of dust and thereby break the degeneracies between dust and the galaxy physics in simulations. While some upcoming surveys, such as BGS, will have existing near-IR photometry from NEOWISE (Meisner et al. 2018), future observations will dramatically expand the information we have in IR. Nancy Grace Roman Space Telescope and James Webb Space Telescope, for instance, will provide valuable near and mid-IR observations. Meanwhile, IR observations at even longer wavelengths will come from Atacama Large Millimeter/submillimeter Array or future facilities such as the Next-Generation Very Large Array and Origins Space Telescope.

6. Summary

In this work, we present the EDA, a framework for applying dust attenuation to simulated galaxy populations. It uses a parameterization of the attenuation curves motivated from observations (Noll et al. 2009) and assigns attenuation curves to simulated galaxies based on their physical properties (M_* and SSFR). We apply the EDA to three state-of-the-art hydrodynamical simulations (SIMBA, TNG, and EAGLE) and forward model the optical and UV color–magnitude relations. We then compare the forward-modeled simulations to a $M_r<-20$, $M_{\rm FUV}<-13.5$, $M_{\rm NUV}<-14.0$ complete SDSS galaxy sample using likelihood-free inference. Based on this comparison, we find the following results:

1. Dust attenuation is essential for our hydrodynamical simulations to reproduce observations. Without dust, SIMBA, TNG, and EAGLE all struggle to reproduce the observed UV and optical color–magnitude relation. They predict galaxies that are intrinsically much bluer and more luminous in the optical and with broader distributions of FUV − NUV colors than SDSS.
2. With the EDA, each of the simulations are able to produce UV and optical color–magnitude relations in good agreement with SDSS observations. However, we find that star formation quenching is too efficient for some of the most massive quiescent galaxies in SIMBA and TNG. When forward modeled, these galaxies appear as luminous UV-red galaxies outside the observed UV color–magnitude relation that cannot be reconciled by dust.
3. The attenuation curves of star-forming galaxies predicted by the EDA for the simulations are in good agreement with the observed attenuation-slope relation. They also closely reproduce the observed attenuation curves of star-forming galaxies. The simulations, however, predict star-forming galaxies that are intrinsically more luminous than observations. Hence, the EDA does not predict star-forming galaxies with $A_V<0.3$, as found in the literature.

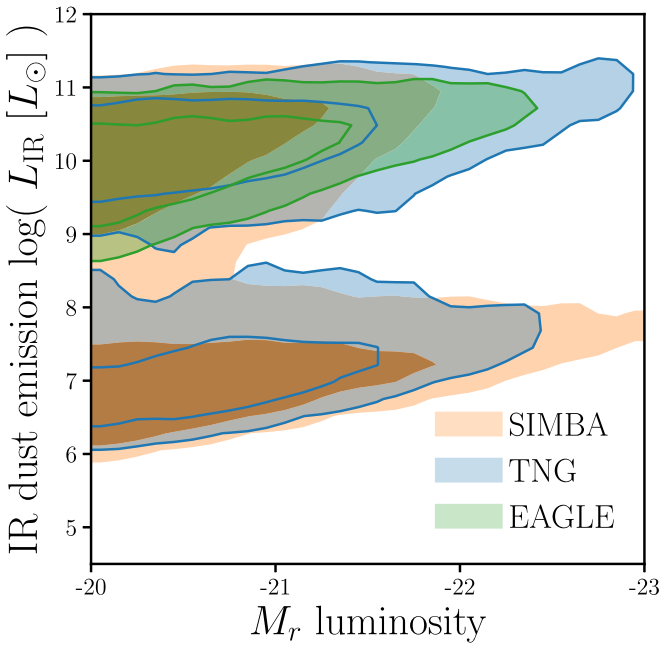

Figure 12. IR dust emission luminosity predicted by the EDA with median parameter values of the SIMBA (orange), TNG (blue), and EAGLE (green) posteriors as a function of M_r. The dust emission is estimated assuming the da Cunha et al. (2008) energy balance. Despite reproducing the same SDSS UV and optical color–magnitude relations, because the simulations require different amounts of dust attenuation to do this, *the EDA predicts significantly different IR dust emissions*. Therefore, including IR observations will significantly improve the constraints on EDA parameters and allow us to better differentiate galaxy formation models.

4. Lastly, we predict the attenuation curves of quiescent galaxies, which observations struggle to directly measure. We find that quiescent galaxies have $A(\lambda)$ curves with lower amplitudes and shallower slopes than star-forming galaxies. Over the entire population, we find that more massive galaxies have higher overall dust attenuation, while galaxies with higher SSFR have steeper attenuation curves.

Our results clearly demonstrate that the EDA and a forward modeling approach provide key insights into dust attenuation. For those uninterested in dust, the EDA also provides a computationally feasible framework for marginalizing over dust when comparing simulations to observations. However, we find that dust attenuation is *highly* degenerate with differences in their galaxy physics prescriptions. Even though the simulations predict galaxy populations with significantly different physical properties, there is enough uncertainty in our understanding of dust that by adjusting attenuation they can all reproduce the same SDSS observations. This also suggests that any comparisons across simulations must marginalize over dust attenuation or run the risk of overinterpretation. Therefore, our current understanding of dust, or lack of, limits our ability to distinguish between the various hydrodynamical models and is a major bottleneck for investigating galaxy formation using simulations.

The forward modeling approach we present offers many avenues for improving on our understanding of dust. In this paper, we used a restrictive observational galaxy sample from SDSS. Comparison to a larger observed 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 completeness limits and include the SDSS survey selection in our forward model. Upcoming surveys, such as the DESI BGS and the PFS Galaxy Evolution Survey, will also soon provide much larger observational galaxy samples. Furthermore, IR observations, which measure dust emission and trace the total dust attenuation, have the potential to tightly constrain the EDA parameters and therefore break degeneracies between dust and the galaxy physics in simulations. In the next paper of the series, we will use the forward modeling approach with the EDA to investigate star formation quenching in galaxy formation models. In other future works, we will apply the EDA and a forward modeling approach to more statistically powerful samples and include IR observables in order to tightly constrain and reveal new insights into dust attenuation.

It is a pleasure to thank Michael Blanton, Nicholas T. Faucher, Marla Geha, Shy Genel, Jenny E. Green, Daniel Kelson, Mariska Kriek, Peter Melchior, Desika Narayanan, Samir Salim, and Katherine Suess for valuable discussions and comments. 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. C. H. is supported by the AI Accelerator program of the Schmidt Futures Foundation.

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

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

The data used in this work were, in part, hosted on facilities supported by the Scientific Computing Core at the Flatiron Institute, a division of the Simons Foundation, and the analysis was largely done using those facilities. The IQ (Isolated & Quiescent) Collaboratory thanks the Flatiron Institute for hosting the collaboratory and its meetings. The Flatiron Institute is supported by the Simons Foundation. Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS website is www.sdss.org. The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions.

Appendix
The Slab Model–Based EDA

In our EDA prescription, we use the slab model to determine A_V, the amplitude of attenuation, as a function of a randomly sampled inclination, i, and τ_V (see Equation (3) in Section 3). The slab model is based on the assumption that dust in galaxies has slab-like geometry and is illuminated by the stellar radiation source (Somerville & Primack 1999). For a given τ_V, the attenuation depends solely on the orientation of the galaxy. While this simplification reproduces the correlation between A_V and i found in observed star-forming galaxies (e.g., Conroy et al. 2010; Wild et al. 2011; Battisti et al. 2017; Salim & Narayanan 2020), it ignores the detailed star-to-dust geometry that impacts the attenuation curve. It also does not provide a physically motivated prescription for quiescent galaxies, which typically have elliptical morphologies. Despite its limitations, the slab model provides a robust empirical prescription that allows us to produce realistic distributions of A_V.

In Figure 13, we compare the A_V distributions, $p(A_V)$, of star-forming galaxies in SDSS (blue) and galaxies in the Salim et al. (2018) GSWLC2 sample (orange) to $p(A_V)$ generated from the slab model (black dashed line). The A_V values of the SDSS are derived using SED fitting from the Brinchmann et al. (2004) MPA-JHU catalog. The GSWLC2 A_V values are also derived from SED fitting UV and optical photometry from GALEX and SDSS observations as well as mid-IR photometry from WISE. The GSWLC2 $p(A_V)$ includes all galaxies, including quiescent ones, above $M_* > 10^{10}\ M_\odot$. We generate two $p(A_V)$ with the slab model for the SDSS and GSWLC2 samples separately. For each SDSS/GSWLC2 galaxy, we determine A_V by uniformly sampling $\cos i$ from 0 to 1 and derive τ_V (Equation (4)) with the galaxy's measured M_* and SSFR. We pick m_{τ,M_*}, $m_{\tau,\mathrm{SSFR}}$, c_τ values within the prior range (Table 1) by hand to roughly reproduce the SDSS and GSWLC2 $p(A_V)$ distributions.

Galaxies in SDSS and GSWLC2 have substantially different $p(A_V)$. While the galaxy populations only partially overlap, this difference is *mostly* due to inconsistencies in the A_V

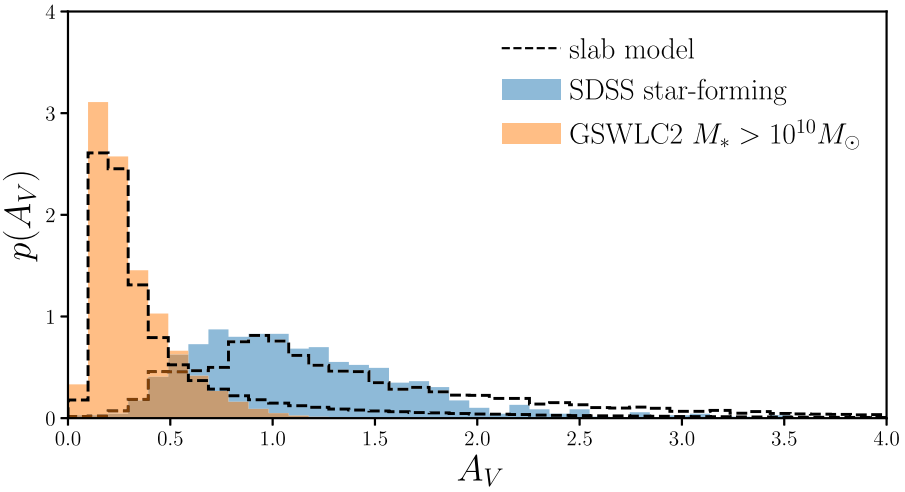

Figure 13. The A_V distributions, $p(A_V)$, generated from the slab model (Equation (3); black dashed line) compared to $p(A_V)$ of star-forming galaxies our SDSS sample (blue; Section 2.4) and of $M_* > 10^{10}\,M_\odot$ star-forming and quiescent galaxies in the GSWLC2 sample (orange). The A_V values for both observations are derived using SED fitting but with different bands and methodologies (Brinchmann et al. 2004; Salim et al. 2018). For the slab model, we generate A_V values for each galaxy in the SDSS and GSWLC2 samples using Equation (3) with its measured M_* and SSFR and randomly sampled i. Despite the significant differences between the $p(A_V)$ of SDSS and GSWLC2, the slab model is able to generate $p(A_V)$ in good agreement with both observations using parameter values within the Table 1 prior range. Therefore, the slab model provides a sufficiently flexible prescription for our EDA.

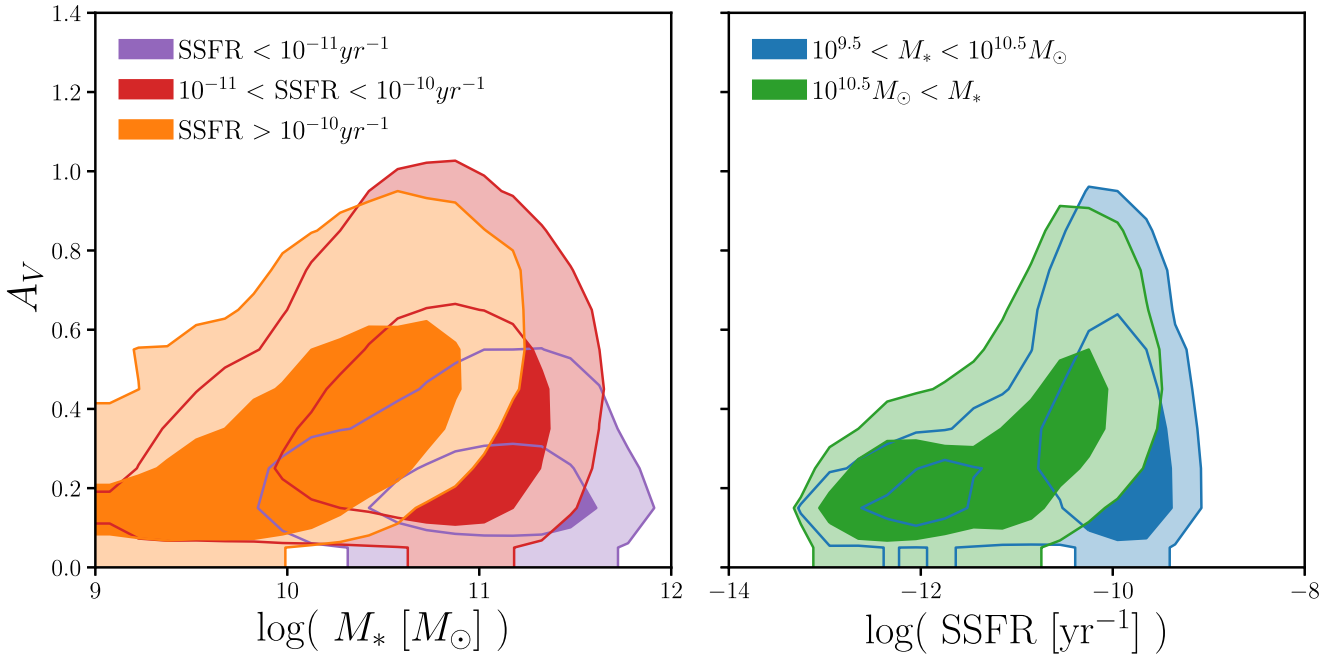

Figure 14. Dependence of A_V on M_* (left) and SSFR (right) for the Salim et al. (2018) GSWLC2 sample. In the left panel, we divide the GSWLC2 sample into bins of SSFR: SSFR $< 10^{-11}$ yr^{-1} (purple), $10^{-11} <$ SSFR $< 10^{-10}$ yr^{-1} (red), and $10^{-10} <$ SSFR (orange). In each of the SSFR bins, we find significant M_* dependence. In the right panel, we divide the sample into bins of M_*: $10^{9.5} < M_* < 10^{10.5}$ $M_\odot$ (blue) and $10^{10.5} M_\odot < M_*$ (green). In the $M_* > 10^{10.5}$ $M_\odot$ bin, which roughly corresponds to our SDSS sample, we find significant SSFR dependence. The M_* and SSFR dependence in A_V we find in GSWLC2 is consistent with previous works and provides further motivation for our EDA prescription.

measurements of MPA-JHU and GSWLC2—even for the same galaxy. This difference in $p(A_V)$ illustrates the challenges in directly measuring dust attenuation. Despite the dramatic differences between the two, the slab model can produce $p(A_V)$ in good agreement with both observed distributions. We therefore conclude that the slab model provides a sufficiently flexible prescription to sample a realistic distribution of A_V.

In addition to the slab model, in the EDA, we also use a linear dependence on M_* and SSFR in the V-band optical depth, τ_V (see Equation (4)). This parameterization is motivated by observations that find significant correlation between A_V and M_* and SSFR (e.g., Garn & Best 2010; Battisti et al. 2016; Salim & Narayanan 2020). We take a closer look at this correlation using the GWSLC2 sample in Figure 14. We present the dependence of A_V on M_* (left panel) and SSFR (right panel). In the left panel, we divide the GSWLC2 galaxies by SSFR: SSFR $< 10^{-11}\,\mathrm{yr}^{-1}$ (purple), $10^{-11} <$ SSFR $< 10^{-10}\,\mathrm{yr}^{-1}$ (red), and $10^{-10} <$ SSFR (orange). For $10^{-10} <$ SSFR bins, we find significant M_* dependence in A_V: more massive galaxies have higher A_V. In the right panel, we divide

the galaxies by M_*: $10^{9.5} < M_* < 10^{10.5}$ $M_\odot$ (blue) and $10^{10.5}$ $M_\odot < M_*$ (green). In both M_* bins, galaxies with higher SSFR have higher A_V. The dependence is stronger for galaxies with $M_* > 10^{10.5}$ $M_\odot$, which roughly corresponds to the M_* limit of our forward model (see Figure 11). Overall, the M_* and SSFR dependence we find in A_V from the GSWLC2 sample is consistent with previous observations and further motivates our EDA prescription.

In Figure 15, we present the M_* and SSFR dependence of dust attenuation in SDSS galaxies, which contains 2361 galaxies that pass our selection cut and are also in the GSWLC2 and MPA-JHU samples. In the top panel, we present A_{1500} from GSWLC2 as a function of M_* and SSFR. In the bottom panels, we present A_V from GSWLC2 (left), MPA-JHU (center), and NSA (right). The NSA A_V measurements are derived assuming an intrinsic Balmer decrement of 2.85, $R_V = 3.1$ and O'Donnell (1994) extinction. The color map in each hexbin represents the median attenuation for galaxies in the bin, the same as in Figure 11. Bins with less than five galaxies are omitted. For each observational sample (column), we use M_* and SSFR from the respective samples for consistency. We find the same M_* and SSFR dependence of A_V as Figure 14 even after our selection cut (bottom left). The bottom panels highlight that even for the same galaxies, A_V from GSWLC2, MPA-JHU, and NSA have significant differently amplitudes. M_* and SSFR are also significantly different across the samples. Since observations have large discrepancies among dust attenuation measurements and a detailed comparison is beyond the scope of this work, we refrain from comparing our EDA predicted dust attenuation (Section 5) to observations.

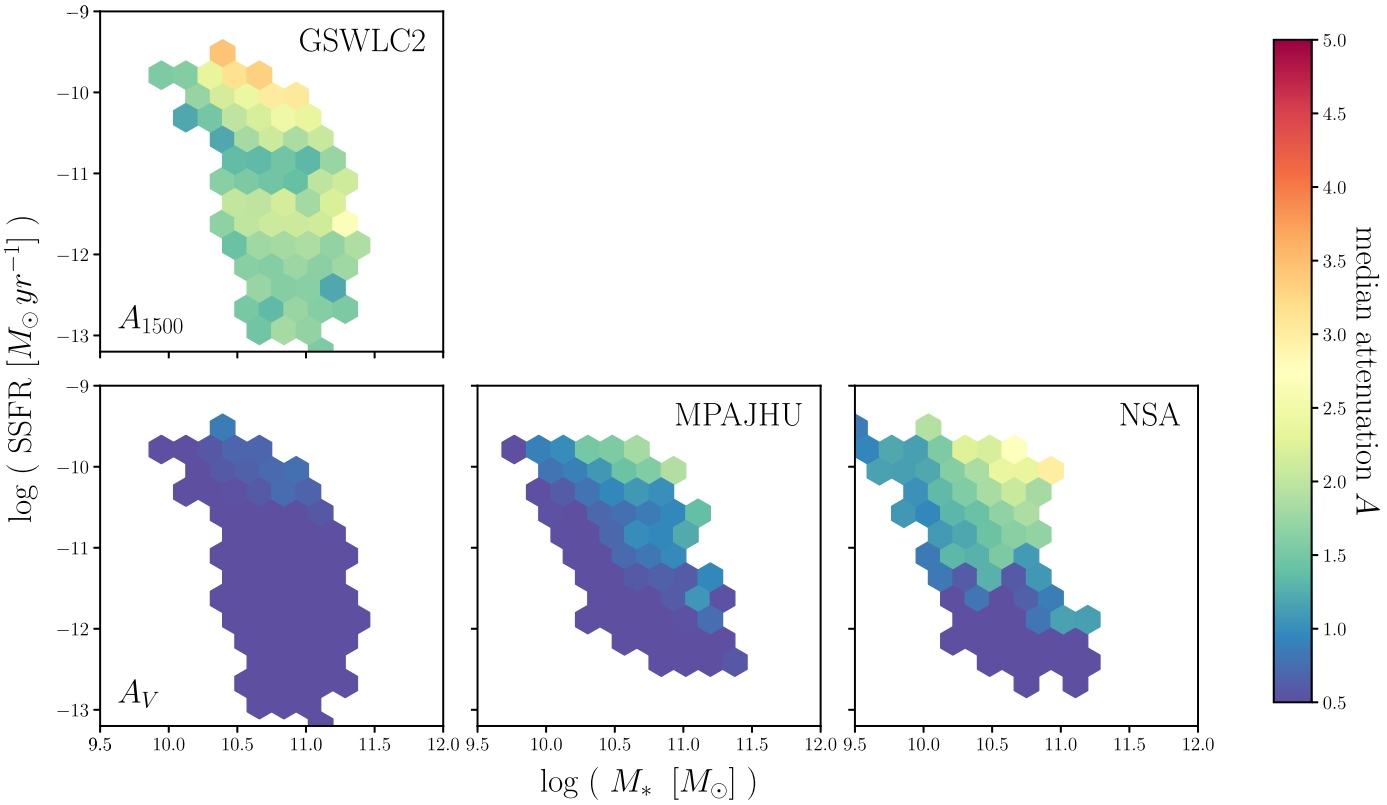

Figure 15. M_* and SSFR dependence of dust attenuation at 1500 Å (A_{1500}; top) and at 5500 Å (A_V; bottom) of SDSS galaxies. The sample includes 2361 galaxies that pass our selection cut (Section 2.4) and are also part of the GSWLC2 and MPA-JHU samples. In the left panels, we use A_{1500} and A_V from GSWLC2. In the center panel, we use A_V from MPA-JHU. In the right panel, we use A_V from the NSA. The M_* and SSFR in each panel are from the respective samples. As in Figure 11, the color map in each hexbin represents the median attenuation for all galaxies in the bin (color bar). We only include bins with more than five galaxies. The bottom panels illustrate that A_V measurements from GSWLC2, MPA-JHU, and NSA differ *significantly* even for the same galaxies. We, therefore, do not directly compare our EDA predictions to observations.

ORCID iDs

ChangHoon Hahn https://orcid.org/0000-0003-1197-0902
Tjitske K. Starkenburg https://orcid.org/0000-0003-2539-8206
Daniel Anglés-Alcázar https://orcid.org/0000-0001-5769-4945
Ena Choi https://orcid.org/0000-0002-8131-6378
Romeel Davé https://orcid.org/0000-0003-2842-9434
Claire Dickey https://orcid.org/0000-0002-1081-3991
Kartheik G. Iyer https://orcid.org/0000-0001-9298-3523
Ariyeh H. Maller https://orcid.org/0000-0003-2060-8331
Jeremy L. Tinker https://orcid.org/0000-0003-3578-6149
L. Y. Aaron Yung https://orcid.org/0000-0003-3466-035X

References

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, ApJS, 182, 543
Aihara, H., Allende Prieto, C., An, D., et al. 2011, ApJS, 193, 29
Akeret, J., Refregier, A., Amara, A., Seehars, S., & Hasner, C. 2015, JCAP, 2015, 043
Ali, S. S., Bremer, M. N., Phillipps, S., & De Propris, R. 2018, MNRAS, 476, 1010
Alsing, J., Wandelt, B., & Feeney, S. 2018, MNRAS, 477, 2874
Anglés-Alcázar, D., Davé, R., Faucher-Giguère, C.-A., Özel, F., & Hopkins, P. F. 2017a, MNRAS, 464, 2840
Anglés-Alcázar, D., Faucher-Giguère, C.-A., Kereš, D., et al. 2017b, MNRAS, 470, 4698
Baes, M., Trčka, A., Camps, P., et al. 2019, MNRAS, 484, 4069
Battisti, A. J., Calzetti, D., & Chary, R.-R. 2016, ApJ, 818, 13
Battisti, A. J., Calzetti, D., & Chary, R.-R. 2017, ApJ, 840, 109
Beaumont, M. A., Cornuet, J.-M., Marin, J.-M., & Robert, C. P. 2009, Biometrika, 96, 983
Blanton, M. R., Kazin, E., Muna, D., Weaver, B. A., & Price-Whelan, A. 2011, AJ, 142, 31
Blanton, M. R., & Roweis, S. 2007, AJ, 133, 734
Blanton, M. R., Schlegel, D. J., Strauss, M. A., et al. 2005, AJ, 129, 2562
Booth, C. M., & Schaye, J. 2009, MNRAS, 398, 53
Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, MNRAS, 351, 1151
Burgarella, D., Buat, V., & Iglesias-Páramo, J. 2005, MNRAS, 360, 1413
Calzetti, D. 2001, NewAR, 45, 601
Cameron, E., & Pettitt, A. N. 2012, MNRAS, 425, 44
Camps, P., & Baes, M. 2015, A&C, 9, 20
Carnall, A. C., Leja, J., Johnson, B. D., et al. 2019, ApJ, 873, 44
Chabrier, G. 2003, PASP, 115, 763
Chevallard, J., Charlot, S., Wandelt, B., & Wild, V. 2013, MNRAS, 432, 2061
Choi, J., Dotter, A., Conroy, C., et al. 2016, ApJ, 823, 102
Cochrane, R. K., Hayward, C. C., Anglés-Alcázar, D., et al. 2019, MNRAS, 488, 1779
Code, A. D. 1969, PASP, 81, 475
Conroy, C. 2010, MNRAS, 404, 247
Conroy, C. 2013, ARA&A, 51, 393
Conroy, C., Gunn, J. E., & White, M. 2009, ApJ, 699, 486
Conroy, C., White, M., & Gunn, J. E. 2010, ApJ, 708, 58
Crain, R. A., Schaye, J., Bower, R. G., et al. 2015, MNRAS, 450, 1937
da Cunha, E., Charlot, S., & Elbaz, D. 2008, MNRAS, 388, 1595
Dalla Vecchia, C., & Schaye, J. 2012, MNRAS, 426, 140
Dantas, M. L. L., Coelho, P. R. T., & Sánchez-Blázquez, P. 2021, MNRAS, 500, 1870
Davé, R., Anglés-Alcázar, D., Narayanan, D., et al. 2019, MNRAS, 486, 2827
Davé, R., Crain, R. A., Stevens, A. R. H., et al. 2020, MNRAS, 497, 146
Davé, R., Rafieferantsoa, M. H., & Thompson, R. J. 2017a, MNRAS, 471, 1671
Davé, R., Rafieferantsoa, M. H., Thompson, R. J., & Hopkins, P. F. 2017b, MNRAS, 467, 115
Davé, R., Thompson, R., & Hopkins, P. F. 2016, MNRAS, 462, 3265
Del Moral, P., Doucet, A., & Jasra, A. 2012, Stat. and Comput., 22, 1009
DESI Collaboration, Aghamousa, A., guillar, J., et al. 2016, arXiv:1611.00036 [astro-ph]
Dickey, C. M., Starkenburg, T. K., Geha, M., et al. 2021, ApJ, 915, 53
Diggle, P. J., & Gratton, R. J. 1984, J. R. Stat. Soc. Series B Stat. Methodol., 46, 193
Dotter, A. 2016, ApJS, 222, 8
Draine, B. T. 2003, ApJ, 598, 1017
Fontanot, F., Somerville, R. S., Silva, L., Monaco, P., & Skibba, R. 2009, MNRAS, 392, 553
Galliano, F., Galametz, M., & Jones, A. P. 2018, ARA&A, 56, 673
Garn, T., & Best, P. N. 2010, MNRAS, 409, 421
Genel, S., Vogelsberger, M., Springel, V., et al. 2014, MNRAS, 445, 175
Gonzalez-Perez, V., Lacey, C. G., Baugh, C. M., Frenk, C. S., & Wilkins, S. M. 2013, MNRAS, 429, 1609
Gordon, K. D., Witt, A. N., Carruthers, G. R., Christensen, S. A., & Dohne, B. C. 1994, ApJ, 432, 641
Granato, G. L., Lacey, C. G., Silva, L., et al. 2000, ApJ, 542, 710
Hahn, C., Beutler, F., Sinha, M., et al. 2019b, MNRAS, 485, 2956
Hahn, C., Starkenburg, T. K., Choi, E., et al. 2019c, ApJ, 872, 160
Hahn, C., Tinker, J. L., & Wetzel, A. 2019a, arXiv:1910.01644 [astro-ph]
Hahn, C., Tinker, J. L., & Wetzel, A. R. 2017b, ApJ, 841, 6
Hahn, C., Vakili, M., Walsh, K., et al. 2017a, MNRAS, 469, 2791
Han, Z., Podsiadlowski, P., & Lynas-Gray, A. E. 2007, MNRAS, 380, 1098
Hayward, C. C., & Smith, D. J. B. 2015, MNRAS, 446, 1512
Hopkins, P. F. 2015, MNRAS, 450, 53
Hopkins, P. F., Weltzel, A., Kereš, D., et al. 2018, MNRAS, 477, 1578
Hou, K.-C., Hirashita, H., Nagamine, K., Aoyama, S., & Shimizu, I. 2017, MNRAS, 469, 870
Ishida, E. E. O., Vitenti, S. D. P., Penna-Lima, M., et al. 2015, A&C, 13, 1
Jonsson, P. 2006, MNRAS, 372, 2
Katsianis, A., Gonzalez, V., Barrientos, D., et al. 2020, MNRAS, 492, 5592
Kaviraj, S., Schawinski, K., Devriendt, J. E. G., et al. 2007, ApJS, 173, 619
Kirkpatrick, A., Pope, A., Sajina, A., et al. 2015, ApJ, 814, 9
Kriek, M., & Conroy, C. 2013, ApJL, 775, L16
Le Cras, C., Maraston, C., Thomas, D., & York, D. G. 2016, MNRAS, 461, 766
Leja, J., Carnall, A. C., Johnson, B. D., Conroy, C., & Speagle, J. S. 2019, ApJ, 876, 3
Leja, J., Johnson, B. D., Conroy, C., & van Dokkum, P. 2018, ApJ, 854, 62
Leja, J., Johnson, B. D., Conroy, C., van Dokkum, P. G., & Byler, N. 2017, ApJ, 837, 170
Li, Q., Narayanan, D., & Davé, R. 2019, MNRAS, 490, 1425
Lin, C.-A., Kilbinger, M., & Pires, S. 2016, A&A, 593, A88
Marinacci, F., Vogelsberger, M., Pakmor, R., et al. 2018, MNRAS, 480, 5113
McAlpine, S., Helly, J. C., Schaller, M., et al. 2016, A&C, 15, 72
Meisner, A. M., Lang, D., & Schlegel, D. J. 2018, RNAAS, 2, 1
Muratov, A. L., Kereš, D., Faucher-Giguère, C.-A., et al. 2015, MNRAS, 454, 2691
Naiman, J. P., Pillepich, A., Springel, V., et al. 2018, MNRAS, 477, 1206
Narayanan, D., Conroy, C., Davé, R., Johnson, B. D., & Popping, G. 2018, ApJ, 869, 70
Natale, G., Popescu, C. C., Tuffs, R. J., et al. 2015, MNRAS, 449, 243
Nelson, D., Pillepich, A., Genel, S., et al. 2015, A&C, 13, 12
Nelson, D., Pillepich, A., Springel, V., et al. 2018, MNRAS, 475, 624
Nelson, D., Springel, V., Pillepich, A., et al. 2019, ComAC, 6, 2
Noll, S., Burgarella, D., Giovannoli, E., et al. 2009, A&A, 507, 1793
O'Connell, R. W. 1999, ARA&A, 37, 603
O'Donnell, J. E. 1994, ApJ, 422, 158
Paxton, B., Bildsten, L., Dotter, A., et al. 2011, ApJS, 192, 3
Paxton, B., Cantiello, M., Arras, P., et al. 2013, ApJS, 208, 4
Paxton, B., Marchant, P., Schwab, J., et al. 2015, ApJS, 220, 15
Pillepich, A., Springel, V., Nelson, D., et al. 2018, MNRAS, 473, 4077
Popping, G., Somerville, R. S., & Galametz, M. 2017, MNRAS, 471, 3152
Pritchard, J. K., Seielstad, M. T., Perez-Lezaun, A., & Feldman, M. W. 1999, Mol. Biol. Evol., 16, 1791
Reddy, N. A., Kriek, M., Shapley, A. E., et al. 2015, ApJ, 806, 259
Rocha, M., Jonsson, P., Primack, J. R., & Cox, T. J. 2008, MNRAS, 383, 1281
Rodriguez-Gomez, V., Snyder, G. F., Lotz, J. M., et al. 2019, MNRAS, 483, 4140
Ruiz-Macias, O., Zarrouk, P., Cole, S., et al. 2021, MNRAS, 502, 4328
Salim, S., Boquien, M., & Lee, J. C. 2018, ApJ, 859, 11
Salim, S., & Narayanan, D. 2020, ARA&A, 58, 529
Salim, S., Lee, J. C., Janowiecki, S., et al. 2016, ApJS, 227, 2
Sánchez-Blázquez, P., Peletier, R. F., Jiménez-Vicente, J., et al. 2006, MNRAS, 371, 703
Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, MNRAS, 446, 521
Seon, K.-I., & Draine, B. T. 2016, ApJ, 833, 201
Somerville, R. S., Gilmore, R. C., Primack, J. R., & Domínguez, A. 2012, MNRAS, 423, 1992
Somerville, R. S., & Primack, J. R. 1999, MNRAS, 310, 1087
Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, ApJS, 214, 15
Springel, V. 2005, MNRAS, 364, 1105

Springel, V., Pakmor, R., Pillepich, A., et al. 2018, MNRAS, 475, 676
Steinacker, J., Baes, M., & Gordon, K. D. 2013, ARA&A, 51, 63
Takada, M., Ellis, R. S., Chiba, M., et al. 2014, PASJ, 66, R1
Tamura, N., Takato, N., Shimono, A., et al. 2016, Proc. SPIE, 9908, 99081M
Tavare, S., Balding, D. J., Griffiths, R. C., & Donnelly, P. 1997, Genetics, 145, 505
Tinker, J., Wetzel, A., & Conroy, C. 2011, arXiv:1107.5046
Trayford, J. W., Camps, P., Theuns, T., et al. 2017, MNRAS, 470, 771
Trayford, J. W., Lagos, C. d. P., Robotham, A. S. G., & Obreschkow, D. 2020, MNRAS, 491, 3937
Trayford, J. W., Theuns, T., Bower, R. G., et al. 2015, MNRAS, 452, 2879
Tress, M., Mármol-Queraltó, E., Ferreras, I., et al. 2018, MNRAS, 475, 2363
Vogelsberger, M., Genel, S., Springel, V., et al. 2014, MNRAS, 444, 1518
Vogelsberger, M., Nelson, D., Pillepich, A., et al. 2020, MNRAS, 492, 5167
Walcher, J., Groves, B., Budavári, T., & Dale, D. 2011, Ap&SS, 331, 1
Wechsler, R. H., & Tinker, J. L. 2018, ARA&A, 56, 435
Weinberger, R., et al. 2018, MNRAS, 479, 4056
Weyant, A., Schafer, C., & Wood-Vasey, W. M. 2013, ApJ, 764, 116
Wild, V., Charlot, S., Brinchmann, J., et al. 2011, MNRAS, 417, 1760
Wilkins, S. M., Gonzalez-Perez, V., Lacey, C. G., & Baugh, C. M. 2012, MNRAS, 424, 1522
Witt, A. N., & Gordon, K. D. 1996, ApJ, 463, 681
Witt, A. N., & Gordon, K. D. 2000, ApJ, 528, 799
Yi, S., Demarque, P., & Kim, Y.-C. 1997, ApJ, 482, 677
Yung, L. Y. A., Somerville, R. S., Finkelstein, S. L., Popping, G., & Davé, R. 2019, MNRAS, 483, 2983