

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

CHANGHOON HAHN<sup>1, 2, \*</sup> AND IQ COLLABORATORY

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

<sup>2</sup>*Berkeley Center for Cosmological Physics, University of California, Berkeley CA 94720, USA*

(Dated: DRAFT --- 2b67db9 --- 2020-07-16 --- NOT READY FOR DISTRIBUTION)

## ABSTRACT

We construct a dust empirical model (DEM) framework for applying dust attenuation to simulated galaxies based on handful of sensible assumptions that come from our understanding of dust attenuation. The DEM framework is essentially based on state-of-the-art attenuation curves with a flexible parameterization that allows us to statistically sample them. Applying DEMs to three different hydrodynamic simulations, SIMBA, Illustris TNG, and EAGLE, we are able to produce UV and optical photometry consistent with SDSS observations ( $(G - R) - M_r$  color-magnitude and  $(FUV - NUV) - M_r$  relations). This suggests that there's enough freedom in our current understanding (or lack) of dust for all simulations to reproduce observations. Meanwhile, the DEM provides some insights into dust as well as the subgrid physics that goes into the hydro simulations.

*Keywords:* keyword1 – keyword2 – keyword3

## 1. INTRODUCTION

dust is important because....

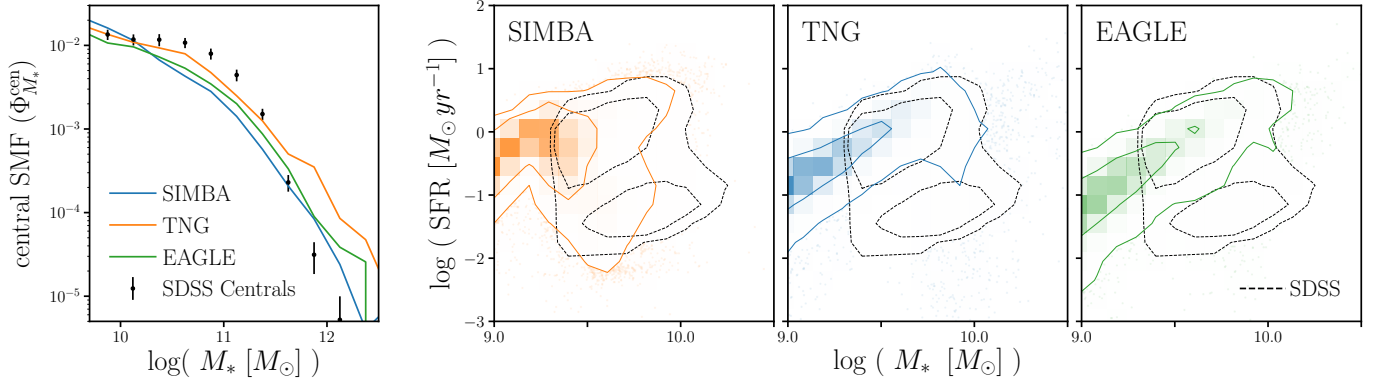
assumptions on the attenuation curve can dramatically impact the physical properties inferred from SED fitting (*e.g.* [Kriek & Conroy 2013](#); [?](#); [Reddy et al. 2015](#); [Salim & Narayanan 2020](#)).

motivation for an empirical dust attenuation model

attenuation vs extinction. While extinction curves have been derived from observations and theoretically, it's not easy to map this to attenuation curves. Attenuation curves are a product of complicated empirical processes since it accounts for light that gets scattered and star light that is not obscured

This makes modeling them in a complete physically motivated method expensive. People have done it [Narayanan et al. \(2018\)](#); [Trayford et al. \(2020\)](#). some detail about the radiative transfer method and such. But besides being expensive they have to make a number of assumptions anyway. *e.g.* [Narayanan et al. \(2018\)](#) assumes a fixed extinction curve.

\* [hahn.changhoon@gmail.com](mailto:hahn.changhoon@gmail.com)



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

Moreover, because the radiative transfer method is expensive it's hard to compare many different simulations. Not only that, observables generated from simulations that take into radiative transfer dust models complicates simulation to simulation comparisons. Because you're simultaneously comparing the galaxy formation prescription and all the dust prescription.

emphasize somewhere here that using a prescription trained one sim doesn't translate well to other sims.

Instead, we present a framework using flexible dust empirical models that paints attenuation curves onto galaxies. describe at a high level how we are parameterizing DEMs

emphasize comparisons in observable space (Nelson et al. (2018) talks about importance of color)

talk about the advantages: extremely flexible so it can encompass the wide variety of attenuation curves found in radiative transfer, easy to correlate the attenuation curve with galaxy properties.

Also DEMs make it possible to statistically apply attenuation curves for large galaxy population. Putting this ontop of simulations, we can use them to generate observables and compare them to observations to constrain the DEM. This framework allows us to learn about attenuation curves given a model for galaxy formation.

Trayford et al. (2015) uses an empirical dust model but doesn't implement an attenuation curve but rather multiplicative factors for the broadband photometry.

The other way around also works. If you don't care about dust at all, DEM provides a framework to easily marginalize over dust attenuation and treat dust as a nuisance parameter.

In this paper, we do above for multiple simulations.

Starkenburger et al. in prep will use this framework to marginalize over dust and compare galaxy populations predicted by multiple simulations .

**CH:** why do we only do centrals?

## 2. DATA

In this paper we apply dust empirical models (DEM) to galaxies in the Illustris TNG, EAGLE, and SIMBA cosmological hydrodynamical simulations. We then measure the  $r$ -band luminosity ( $M_r$ ), optical color ( $G-R$ ), and UV color ( $FUV-NUV$ ) of the DEM output from forward modeled spectral energy distributions (SED). These forward modeled observables, unlike physical properties such as  $M_*$  and SFR, are *consistently* defined and derived in both simulations and observations. Afterwards, we compare these predicted DEM observables to the central galaxies in SDSS DR7 observations. Below, we briefly describe the hydrodynamical simulations and the SDSS observations used throughout this work.

In Figure 1, we present the stellar mass functions ( $\Phi_{M_*}^{\text{cen}}$ ; left-most panel) and  $M_* - \text{SFR}$  relations (right panels) for central galaxies in the SIMBA (orange), TNG (blue), and EAGLE (green) simulations. We include, for reference,  $\Phi_{M_*}^{\text{cen}}$  and the  $M_* - \text{SFR}$  relation for our SDSS central galaxy sample. The uncertainties for the SDSS SMF are derived from jackknife resampling. For the simulations,  $M_*$  is the total stellar mass within the host halo, excluding any stellar mass in subhalos, and SFR is the instantaneous SFR derived from the rate of star formation in dense and cold gas. For SDSS,  $M_*$  is estimated using `kcorrect` (Blanton & Roweis 2007) assuming a Chabrier (2003) initial mass function and SFR are from the current release of Brinchmann et al. (2004)<sup>1</sup>. We describe the simulations and observations further in sections below. Figure 1 illustrates that the hydrodynamical simulations have significantly different SMFs and  $M_* - \text{SFR}$  relations. This difference, which was also recently highlighted in Hahn et al. (2019c), demonstrates that *the hydrodynamical simulations all predict central galaxy populations with significantly different physical properties*.

### 2.1. Illustris TNG

The Illustris TNG simulation<sup>2</sup> (hereafter TNG) is a cosmological hydrodynamical simulation of comoving volume (110.7 Mpc)<sup>3</sup> (Nelson et al. 2018; Pillepich et al. 2018; Springel et al. 2018). It improves on the original Illustris simulation<sup>3</sup> (Vogelsberger et al. 2014; Genel et al. 2014; public data release by Nelson et al. 2015), by including magneto-hydrodynamics and updated treatments for galactic winds, metal enrichment, and AGN feedback. Most notably, TNG uses a new implementation for feedback from SMBH (Weinberger et al. 2018), where feedback energy is injected in the form of a kinetic AGN-driven wind at low SMBH accretion rates. This new implementation has been shown to alleviate discrepancies found between the original Illustris and observations for  $> 10^{13-14} M_\odot$  massive halos. **details on the following properties that we use in the paper: SFH, ZH**

TODO

### 2.2. EAGLE

We use L0100Ref of the Virgo Consortium’s EAGLE project<sup>4</sup>, a publicly available suite of cosmological hydrodynamic simulations (Schaye et al. 2015; Crain et al. 2015; McAlpine et al. 2016). The simulation has a comoving volume of (100 Mpc)<sup>3</sup> and is simulated with the ANARCHY code (Dalla Vecchia et al. in prep.; see also Appendix A of Schaye et al. 2015), a modified version of the GADGET-3

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

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

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

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

code (Springel 2005). It has subgrid models for star formation, stellar mass loss, metal enrichment and stellar feedback that stochastically inject thermal energy in the ISM as in (Dalla Vecchia & Schaye 2012); the feedback energy from AGN is also added to surrounding gas stochastically (Booth & Schaye 2009). Parameters of the stellar feedback and SMBH accretion are calibrated to broadly reproduce the  $z = 0$  stellar mass function and galaxy stellar size-stellar mass relation. Meanwhile, the AGN feedback efficiency is calibrated to match the SMBH-galaxy mass relation. details on the following properties that we use in the paper: SFH, ZH TODO

### 2.3. SIMBA

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

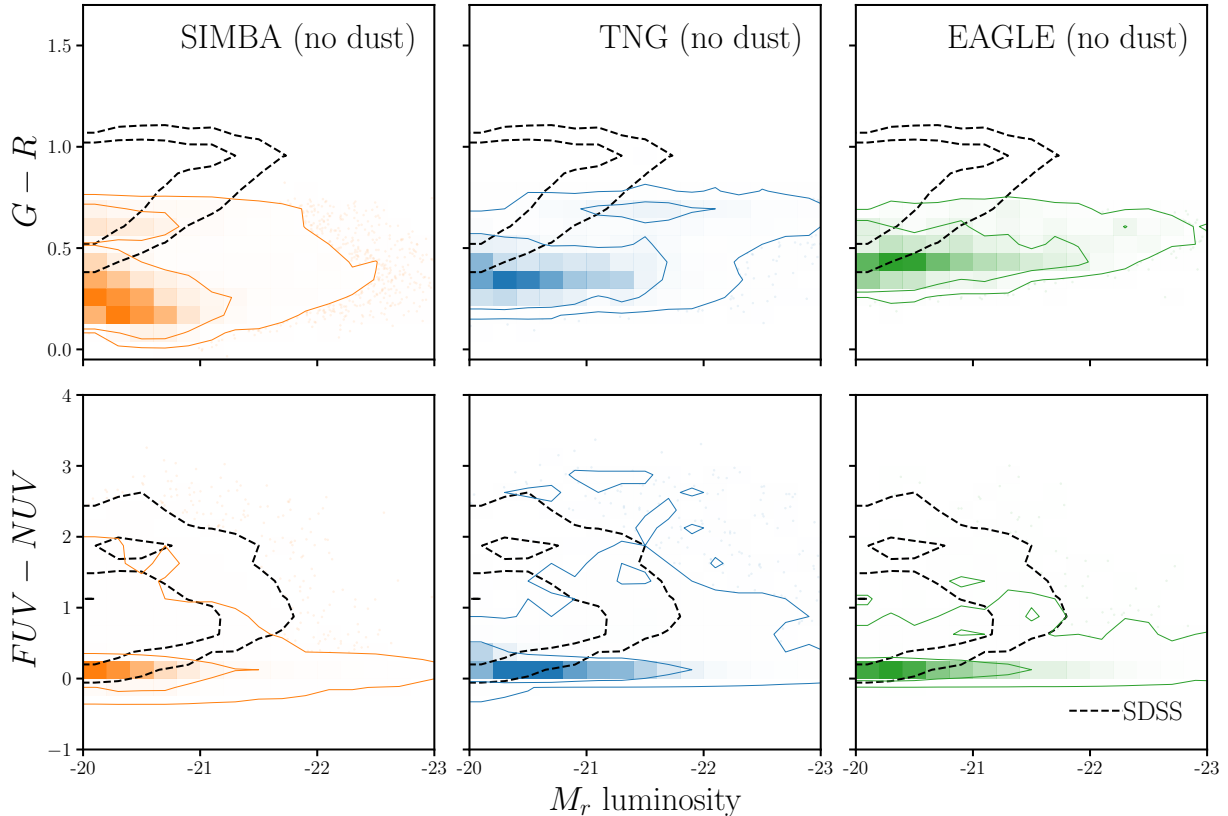
### 2.4. Forward Modeled Spectral Energy Distributions

describe how the SED is generated using the SFH and ZHs TODO

In Figure 2 we present the optical and UV color-magnitude relations,  $(G-R) - M_r$  (top) and  $(FUV-NUV) - M_r$  (bottom), for central galaxies of the SIMBA (left), TNG (center) and EAGLE (right) simulations. The  $G-R$  and  $FUV-NUV$  colors are derived from the forward modeled SED and absolute magnitudes. The observables for the simulations do not include any prescription for dust attenuation. Comparison to the observables of SDSS centrals (black dashed) clearly demonstrate that *without any dust attenuation, the hydrodynamical simulations cannot reproduce the observed optical or UV color-magnitude relations.*

### 2.5. SDSS DR7 Central Galaxies

The goal of the DEM we present in this paper is to provide a flexible model that can reproduce observations using the hydrodynamical simulations above. Throughout this work, we use the Tinker et al. (2011) SDSS central galaxy sample as our observation. This sample is constructed by first selecting a volume-limited sample of galaxies at  $z \approx 0.04$  with  $M_r < -18$  and complete above  $M_* > 10^{9.4} h^{-2} M_\odot$  from the SDSS DR7 (Abazajian et al. 2009) NYU Value-Added Galaxy Catalog (VAGC; Blanton et al. 2005). Then, central galaxies are identified using a halo-based group finder that uses the abundance matching ansatz to iteratively assign halo masses to groups. Every group contains one central galaxy, which by definition is the most massive, and a group can contain  $\geq 0$  satellites. As with any group finder, galaxies are misassigned due to projection effects and redshift space distortions; however, the central galaxy sample has a purity of  $\sim 90\%$  and completeness of  $\sim 95\%$  (Tinker et al.



**Figure 2.** We present the optical and UV color-magnitude relations of central galaxies in the SIMBA (left), TNG (center), and EAGLE (right) simulations. The simulations above do *not* yet include the DEM or any prescription for dust attenuation.  $(G-R) - M_r$  (top) and  $(FUV-NUV) - M_r$  (bottom) are the main observables used throughout the paper. They are derived from forward modeled SEDs and, thus, are consistently defined and measured as SDSS observations (Section 2.4). For comparison, we include the distributions of SDSS centrals (black dashed). *Without dust, the hydrodynamical simulations cannot reproduce the observed optical or UV color-magnitude.*

2018). Finally, we impose a  $r$ -band absolute magnitude cut of  $M_r < -20$  where our SDSS central galaxy sample is complete. In this work, we focus on observables that can be consistently defined and derived in both simulations and observables:  $M_r$ ,  $G-R$ , and  $FUV-NUV$ . For our SDSS sample, we use  $r$  and  $g$  band absolute magnitudes from VAGC and GALEX  $FUV$  and  $NUV$  from the NASA-Sloan Atlas<sup>5</sup>.

### 3. DUST EMPIRICAL MODELING

In this section we present the dust empirical model (DEM), a flexible model for applying attenuation curves to galaxy populations that allows us to incorporate intrinsic variations in dust attenuation as well as correlation to physical galaxy properties. Later, we demonstrate that we can accurately

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

reproduce observations with the DEM and use it to test galaxy formation models and shed light on dust in galaxies.

We define the dust attenuation curve  $A(\lambda)$  as

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

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

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

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

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

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

For all of our galaxies, we uniformly sample  $i$ . Then, we include the correlation between  $A_V$  and galaxy properties ( $M_*$  and SFR), found in both observations and simulations (*e.g.* Narayanan et al. 2018; Salim & Narayanan 2020), in  $\tau_V$ . We use  $\tau_V$  with a simple and flexible linear  $M_*$  and SFR dependence:

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

$m_{\tau, M_*}$  quantifies the  $M_*$  dependence,  $m_{\tau, \text{SFR}}$  quantifies the SFR dependence, and  $c_\tau$  quantifies the overall amplitude. Since  $\tau_V$  is optical depth, we impose a  $\tau_V \geq 0$  limit. We note that the slab model is a naive approximation. In reality,  $A_V$  for a galaxy will depend on complexities of its star-to-dust geometry, variations in the extinction curves, and other properties beyond just inclination and  $\tau_V$ . The purpose of the DEM, however, is not to accurately model dust attenuation for individual galaxies, but rather to accurately model the distribution of dust attenuation of galaxy populations. In this sense, the slab model qualitatively reproduces the correlation between  $A_V$  and  $i$  found in the literature: edge-on galaxies have higher  $A_V$  than face-on galaxies (*e.g.* Salim & Narayanan 2020). More importantly, the distribution of  $A_V$ ,  $p(A_V)$ , produced using the slab model with uniformly sampled inclinations closely matches the  $p(A_V)$  of our SDSS sample (Figure 8). Also, we demonstrate that replacing the slab model with a more flexible prescription for sampling  $A_V$  does not significantly impact our analysis (Appendix B). We therefore conclude that the slab model is a sufficiently flexible empirical prescription for sampling  $A_V$ .

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

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



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

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

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

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

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

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

where  $\lambda_0$ ,  $\Delta\lambda$ , and  $E_b$  are the central wavelength, FWHM, and strength of the bump, respectively. We assume fixed  $\lambda_0 = 2175\text{\AA}$  and  $\Delta\lambda = 350\text{\AA}$ . Kriek & Conroy (2013) and Tress et al. (2018) find that  $E_b$  correlates with the  $\delta$  for star-forming galaxies  $z \sim 2$ . Narayanan et al. (2018) confirmed this dependence in simulations. In our DEM, we assume a fixed relation between  $E_b$  and  $\delta$  from Kriek & Conroy (2013):  $E_b = -1.9 \delta + 0.85$ . Allowing the slope and amplitude of the  $E_b$  and  $\delta$  relation to vary, does *not* impact our results.

Next, to attenuate the galaxy SEDs, we apply  $A(\lambda)$  we separately to the star light and nebular emission:

$$F_o(\lambda) = F_i^{\text{star}}(\lambda)10^{-0.4A(\lambda)} + F_i^{\text{neb}}(\lambda)10^{-0.4A_{\text{neb}}(\lambda)}. \quad (8)$$

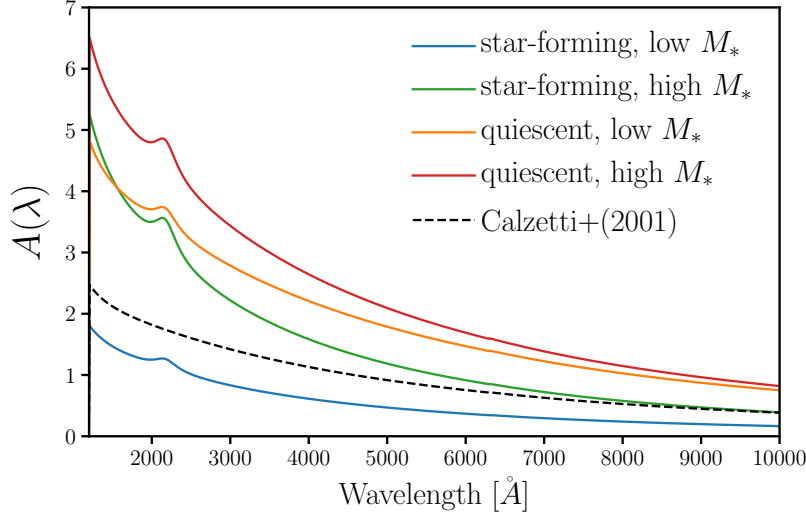
We parameterize

$$A_{\text{neb}}(\lambda) = f_{\text{neb}}A(\lambda) \quad (9)$$

and allow  $f_{\text{neb}}$  to vary freely. In Table 1, we list and describe all of the free parameters in the DEM.

SFR of galaxies are used to calculate  $\tau_v$  and  $\delta$  in Eqs. 4 and 6. Due to mass and temporal resolutions, some galaxies in the simulations have SFR = 0 — *i.e.* an unmeasurably low SFR (Hahn et al. 2019c). Eqs. 4 and 6 cannot be used to derive  $\tau_v$  and  $\delta$  for these galaxies. Since SFR = 0 galaxies do not account for a large fraction of our simulated galaxies, we directly sample their observables ( $G$ ,  $R$ ,  $NUV$ , and  $FUV$ ) from the distribution of observables for SDSS quiescent galaxies. This way, the attenuation of SFR = 0 galaxies does not impact the rest of the DEM parameters. In Appendix A, we discuss the resolution effects in more detail and demonstrate that our results are *not* impacted by other prescriptions for attenuating SFR = 0 galaxies.

In practice, for a simulated galaxy population, we first uniformly sample inclinations,  $i$ , and assign them to each galaxy. Then for a given set of DEM parameter values (Table 1),  $\tau_V$ , and  $\delta$  are calculated for each galaxy using its  $i$ ,  $M_*$ , and SFR. From  $\tau_v$  and  $\delta$ , we get  $A_V$  and  $k(\lambda)$ , which together gives  $A(\lambda)$  for all the galaxies. Afterwards, we attenuate the galaxy SEDs using Eq. 8 and use the attenuated SEDs to calculate our observables: the  $G$ ,  $R$ ,  $NUV$ , and  $FUV$  absolute magnitudes.



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

In Figure 3, we present the DEM  $A(\lambda)$  for galaxies with different SFR and  $M_*$ : star-forming (SFR =  $10^{0.5} M_\odot/\text{yr}$ ) with low mass ( $10^{10} M_\odot$ ; blue), with high mass ( $10^{11} M_\odot$ ; green) and quiescent (SFR =  $10^{-2} M_\odot/\text{yr}$ ) with low mass ( $10^{10} M_\odot$ ; orange), with high mass ( $10^{11} M_\odot$ ; red). All galaxies are edge-on (*i.e.*  $i = 0$ ) and we use  $\{m_{\tau, M_*}, m_{\tau, \text{SFR}}, c_\tau, m_{\delta, M_*}, m_{\delta, \text{SFR}}, c_\delta\} = \{2., -2., 2., -0.1, -0.1, -0.2\}$ , which were arbitrarily chosen within the prior range listed in Table 1. For comparison, we include the Calzetti (2001) attenuation curve. The DEM we describe in this section provides a flexible model for assigning dust attenuation to galaxies based on their physical properties.

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

The DEM provides a flexible model to apply dust attenuation to central galaxies from the hydrodynamic simulations (Section 2) and derive observables which can be directly compared to SDSS observations. For the comparison and the inference of DEM parameters, we use Approximate Bayesian Computation (hereafter ABC; Diggle & Gratton 1984; Tavaré et al. 1997; Pritchard et al. 1999; Beaumont et al. 2009; Del Moral et al. 2012). ABC is a simulation-based (or “likelihood-free”) parameter inference framework that approximates the posterior probability distribution,  $p(\theta | \text{data})$ , without requiring evaluations of the likelihood. Instead, ABC only requires a forward model of the observed data, a prior that can be sampled, and a distance metric that quantifies the “closeness” to the observed data. Since ABC does not require evaluating the likelihood, it does not assume any functional form of the likelihood, which can significantly bias the inferred posterior Hahn et al. (2019b). It also



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

Parameter	Definition	prior
$m_{\tau,M_*}$	Slope of the optical depth, $\tau_V$ , $\log M_*$ dependence	flat $[-5., 5.]$
$m_{\tau,\text{SFR}}$	Slope of the optical depth, $\tau_V$ , $\log \text{SFR}$ dependence	flat $[-5., 5.]$
$c_\tau$	amplitude of the optical depth, $\tau_V$	flat $[0., 6.]$
$m_{\delta,M_*}$	Slope of the attenuation curve slope offset, $\delta$ , $\log M_*$ dependence	flat $[-4., 4.]$
$m_{\delta,\text{SFR}}$	Slope of the attenuation curve slope offset, $\delta$ , $\log \text{SFR}$ dependence of	flat $[-4., 4.]$
$c_\delta$	amplitude of the attenuation curve slope offset, $\delta$	flat $[-4., 4.]$
$f_{\text{neb}}$	nebular attenuation fraction	flat $[1., 4.]$

enables us to estimate the posterior using observables with difficult or intractable likelihoods [Hahn et al. \(2017a\)](#).

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

The forward model in our scenario is the hydrodynamic simulation combined with the DEM. Given a set of values for the DEM parameters, our forward model produces  $G$ ,  $R$ ,  $NUV$ , and  $FUV$  absolute magnitudes, which can be directly compared to SDSS observations. We use uninformative uniform priors on each of the DEM parameters and choose ranges to encompass constraints in the literature. The prior ranges of  $m_{\tau,M_*}$ ,  $m_{\tau,\text{SFR}}$ ,  $c_\tau$  are chosen to conservatively include the  $A_V$  range and  $M_*$  and SFR dependence of [Narayanan et al. \(2018\)](#) and [Salim & Narayanan \(2020\)](#). Meanwhile, the prior ranges of  $m_{\delta,M_*}$ ,  $m_{\delta,\text{SFR}}$ ,  $c_\delta$  are chosen to conservatively include the  $\delta$  range and  $M_*$  and SFR dependence of [Leja et al. \(2017\)](#) and [Salim et al. \(2018\)](#). We list the range of the priors in Table 1. We note that uniform priors on the DEM parameters do not result in uniform priors on  $\tau_V$  or  $\delta$  (*e.g.* [Handley & Millea 2019](#)). However, we are interested in marginalizing over dust attenuation and understanding the dependence of dust attenuation on galaxy properties, so we use uninformative priors on the DEM parameters and not on the derived  $\tau_V$  or  $\delta$ .

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

$$\bar{\rho}(\theta_{\text{DEM}}) = [X^{\text{SDSS}} - X^{\text{FM}}(\theta_{\text{DEM}})]^2. \quad (10)$$

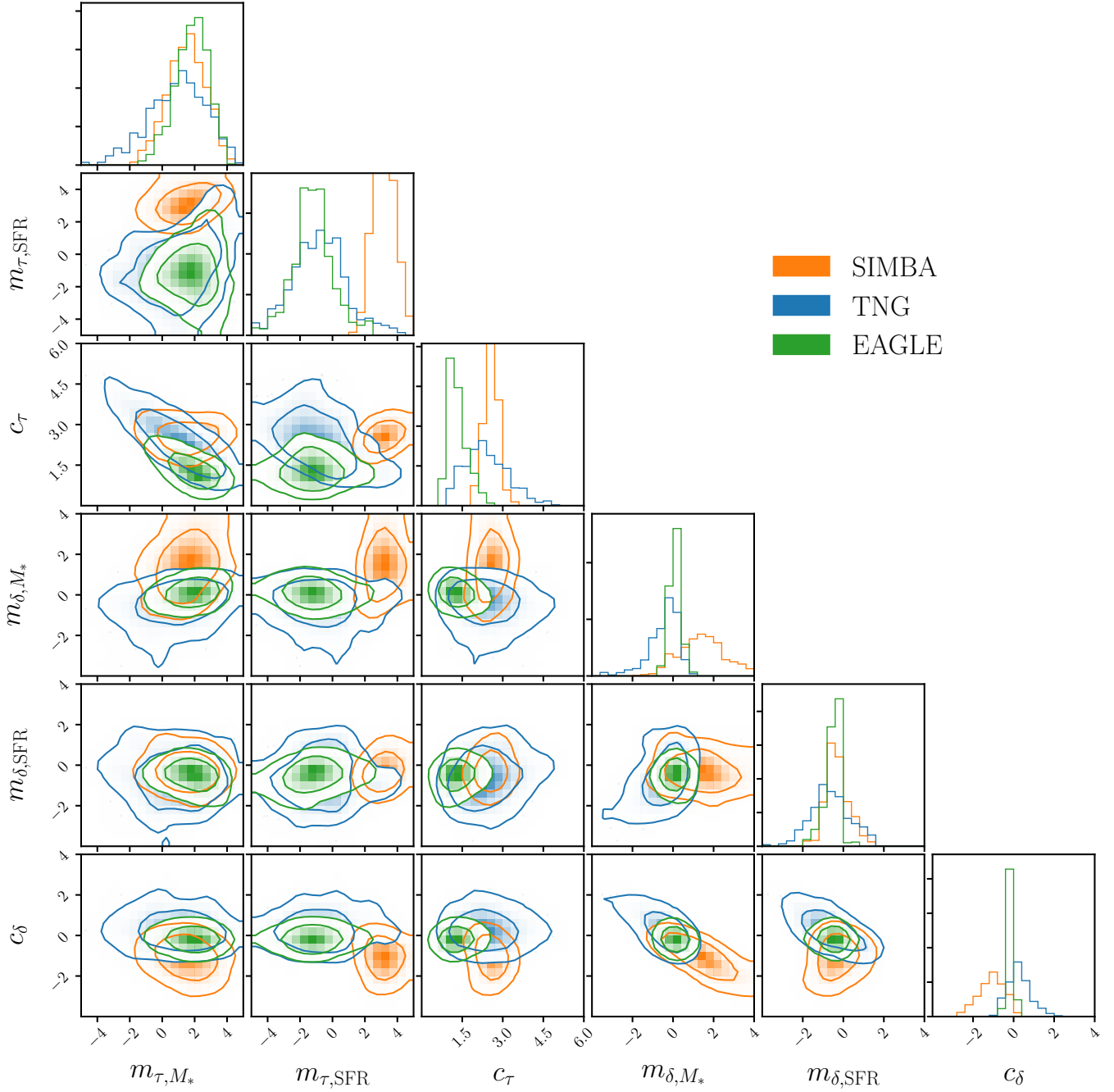
$\theta_{\text{DEM}}$  are the DEM parameters. The summary statistics are based on the optical and UV color-magnitude,  $(G-R)-R$  and  $(FUV-NUV)-R$ , relations of central galaxies brighter than  $M_r < -20$ , where our SDSS central galaxy sample is complete (Figure 2). More specifically, for  $X$ ,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

ABC-PMC begins by first with an arbitrarily large threshold  $\epsilon_1$  and  $N$  proposals  $\bar{\theta}_1$  sampled from the prior distribution. Each proposal is assigned a weight  $w_1^i = 1/N$ . Then for subsequent iterations ( $i > 1$ ), the threshold,  $\epsilon_i$ , is set to the median distance of the previous iteration’s proposals. New proposals are drawn from the previous iteration’s proposals perturbed by a kernel and kept if their distance is below  $\epsilon_i$ . This is repeated until we assemble a new set of  $N$  proposals  $\bar{\theta}_i$ . The entire process is repeated for the next iteration until convergence is confirmed. For further details on the ABC-PMC implementation, we refer readers to Hahn et al. (2017b) and Hahn et al. (2019a). In Figure 4, we present the posterior distributions of the DEM parameters derived from ABC-PMC for the SIMBA (orange), TNG (blue), and EAGLE (green) hydrodynamical simulations. The contours mark the 68% and 95% confidence intervals.

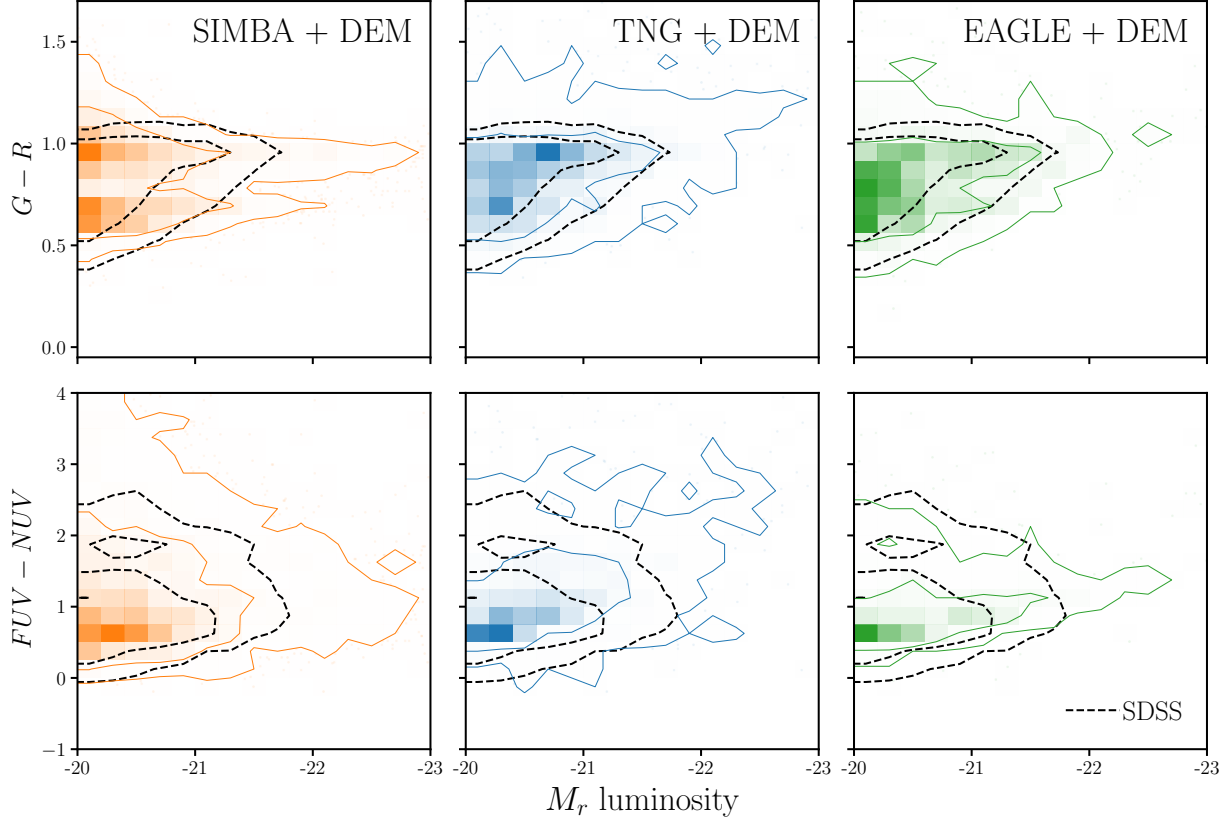
## 5. RESULTS

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

Previous works in the literature have also compared colors and luminosities predicted by simulations to observations. For EAGLE, Trayford et al. (2015) calculate colors and luminosities with the GALAXEV population synthesis models and a two-component screen model for dust. More recently, Trayford et al. (2017) calculated optical colors for EAGLE using SKIRT, a Monte Carlo radiative transfer code (Camps & Baes 2015), to model the dust. Both Trayford et al. (2015) and Trayford et al. (2017) produce bluer red sequences compared to GAMA observations, at  $10^{11.2} < M_* < 10^{11.5}$  for Trayford et al. (2017). Although a detailed comparison is difficult since both works examine all galaxies, not only centrals, the DEM accurately reproduces the position of the SDSS red sequence, even at high  $M_*$ . Trayford et al. (2015) also predict significant more luminous blue galaxies than observations or the DEM. Using the same Trayford et al. (2017) SKIRT framework, Baes et al. (2019) find that they overestimate the observed cosmic spectral energy distributions (CSED) in the UV regime. Moreover, the  $FUV-NUV$  color of their CSED is significantly higher than GAMA  $FUV-NUV$ . The DEM, on the other hand, predict  $FUV-NUV$  in good agreement with SDSS. For TNG, Nelson et al. (2018) calculate optical colors using a dust model that includes attenuation due to dense gas birth clouds surrounding young stellar populations and also due to simulated distribution of neutral



**Figure 4.** Posterior distributions of DEM parameters for the SIMBA (orange), TNG (blue), and EAGLE (green) hydro simulations. The contours mark the 68% and 95% confidence intervals. The posteriors are derived using Approximate Bayesian Computation with Population Monte Carlo (Section 4). **CH: REVISIT:** In all simulations, dust attenuation increases for higher  $M_*$  galaxies ( $m_{\tau, M_*} \sim 2$ ). The simulations also have consistent optical depth amplitudes ( $c_{\tau}$ ). However, the SFR dependence of  $\tau_V$  is different among the simulations. For TNG and EAGLE, star-forming galaxies have lower  $\tau_V$ ; for SIMBA quiescent galaxies have lower  $\tau_V$ . Meanwhile, for the slope offset of the attenuation curve,  $\delta$ , we find little  $M_*$  and SFR dependence in the simulations and that the amplitude ( $c_{\tau}$ ) is consistent with 0.



**Figure 5.** The optical and UV color-magnitude relations predicted by the DEM with the median ABC posteriors for the SIMBA (orange), TNG (blue), and EAGLE (green) hydrodynamical simulations. For comparison, we include the  $(G-R) - M_r$  (top panels) and  $(FUV-NUV) - M_r$  (bottom panels) relations for SDSS (black dashed). With the DEM, the simulations produce dramatically different observables than when we do not include any dust prescription (Figure 2). Hence, dust must be account for when interpreting and comparing simulations. Moreover, with the DEMs, all three simulations produce observables consistent with SDSS. Since different simulations can produce reproduce observations by varying dust, dust significantly limits our ability to constrain the physical processes that go into galaxy simulations.

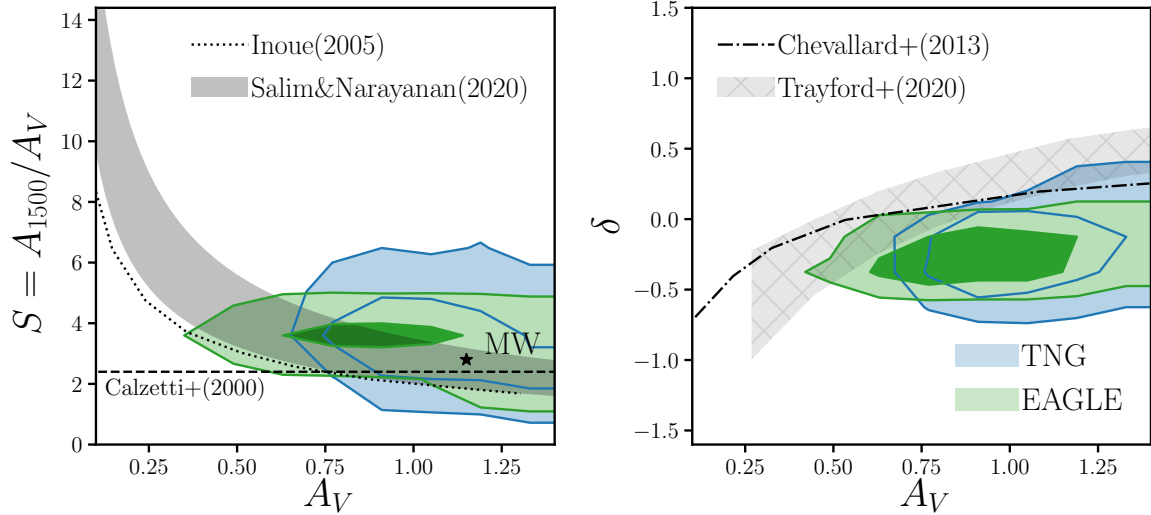
gas and metals. They find bluer red sequence peaks and a narrower blue cloud compared to SDSS. Although they compare the color distribution for all galaxies in  $M_*$  bins, we find neither of these discrepancies with the DEM. *With the DEM, we produce optical and UV color-magnitude relations that are in good agreement with observations, better than previous works, for the SIMBA, TNG, and EAGLE hydrodynamical simulations.*

Figures 5 clearly illustrates that any comparison of simulations must include dust attenuation. Dust dramatically changes the predicted observables of simulations. Without dust (Figure 2), we did not find a clearly bimodality in the optical color-magnitude relation and the simulations predicted UV colors outside of the range of observations. But with a simple framework for dust motivated by attenuation laws and correlation with galaxy properties, such as the DEM, simulations can reproduce observations. Our results also highlight another key point. Robustly interpreting subgrid physics in simulations requires marginalizing over dust. Even for three simulations that produce significantly

different SMFs and  $M_* - \text{SFR}$  relations (Figure 1), the DEM is able to produce observables that agree with observations. In fact for SIMBA, the DEM reproduces the observations by assigning higher attenuation to star-forming galaxies so that they populate the red sequence while quiescent galaxies populate the blue cloud. This is due to the large number of low mass star-forming SIMBA galaxies that lie well above the SFS (Figure 1), which would otherwise all be luminous blue galaxies not found in observations. Our current understanding of dust, which is encapsulated in the DEM, has enough flexibility to reproduce observations for simulations that predict galaxy populations with different physical properties, even if it means contradicting the established relationship between color and SFR. Then marginalizing over dust would leave little constraining power on the subgrid galaxy physics of the simulations. Therefore, *current limitations in our understanding of dust in galaxies significantly impedes our ability to investigate galaxy formation from simulations.*

In addition to reproducing observations, DEM also provides insight into dust in galaxies. Given the parameterization of the DEM, it is especially easy to interpret correlation between dust attenuation and galaxy physical properties. In all three simulations, we find significant positive  $M_*$  dependence of  $\tau_V$ ,  $m_{\tau, M_*} \sim 2$  (Figure 4), consistent with previous works in the literature. Burgarella et al. (2005), for instance, found significant positive  $M_*$  dependence in  $FUV$  attenuation in NUV-selected and FIR-selected samples. Garn & Best (2010) and Battisti et al. (2016) also find positive  $M_*$  dependence in SDSS star-forming galaxies. Most recently, Salim et al. (2018) find higher  $V$  and  $FUV$  attenuation for more massive star-forming galaxies in the GALEX-SDSS-WISE Legacy Catalog 2 (GSWLC2). In addition to the  $M_*$  dependence, the DEM posteriors also reveal the correlation between dust attenuation and star formation. Ignoring SIMBA, which flips the color versus SFR relation, we find  $m_{\tau, \text{SFR}} \sim -1$  — galaxies with lower SFR have higher attenuation. Observations that examine the relationship between dust attenuation and SFR (*e.g.* Garn & Best 2010; Reddy et al. 2015; Battisti et al. 2016, 2017; Salim et al. 2018) have thus far focused primarily on star-forming galaxies. With the DEM, we confirm that *galaxies with higher  $M_*$  have overall higher dust attenuation and find that galaxies with lower SFR have overall higher dust attenuation.*

Our results also shed light on the slope of dust attenuation. In Figure 6, we present the attenuation-slope relation for TNG (blue) and EAGLE (green) with the DEM. The left and right panels present two different measurements of the slope  $S = A(3000\text{\AA})/A_V$ , which is easier to constrain in observations, and  $\delta$ , the slope offset from the Calzetti (2001) curve that we use in the DEM. We include the attenuation and slope for the Milky Way (star) for reference. TNG and EAGLE both predict slopes within  $2 < S < 5$  and centered around  $S \sim 3.5$ . In comparison, for the same  $A_V > 0.4$  range as the DEM, observations find slopes within the range  $2 < S < 5$  (Calzetti et al. 2000; Burgarella et al. 2005; Johnson et al. 2007; Conroy 2010; Wild et al. 2011; Battisti et al. 2016, 2017; Leja et al. 2017; Salim et al. 2018) in good agreement. We also find that the DEM predicts steeper attenuation curves at lower attenuation. This is consistent with the established attenuation-slope relation. At low attenuation, dust scattering dominates absorption so the attenuation curve steepens because red light scatters isotropically while blue light scatters forward (Gordon et al. 1994; Witt & Gordon 2000; Draine 2003). At high attenuation dust absorption is dominant and the attenuation curve is shallower (Chevallard et al. 2013). For the  $A_V$  range probed by the DEM, the  $A_V$ -slope

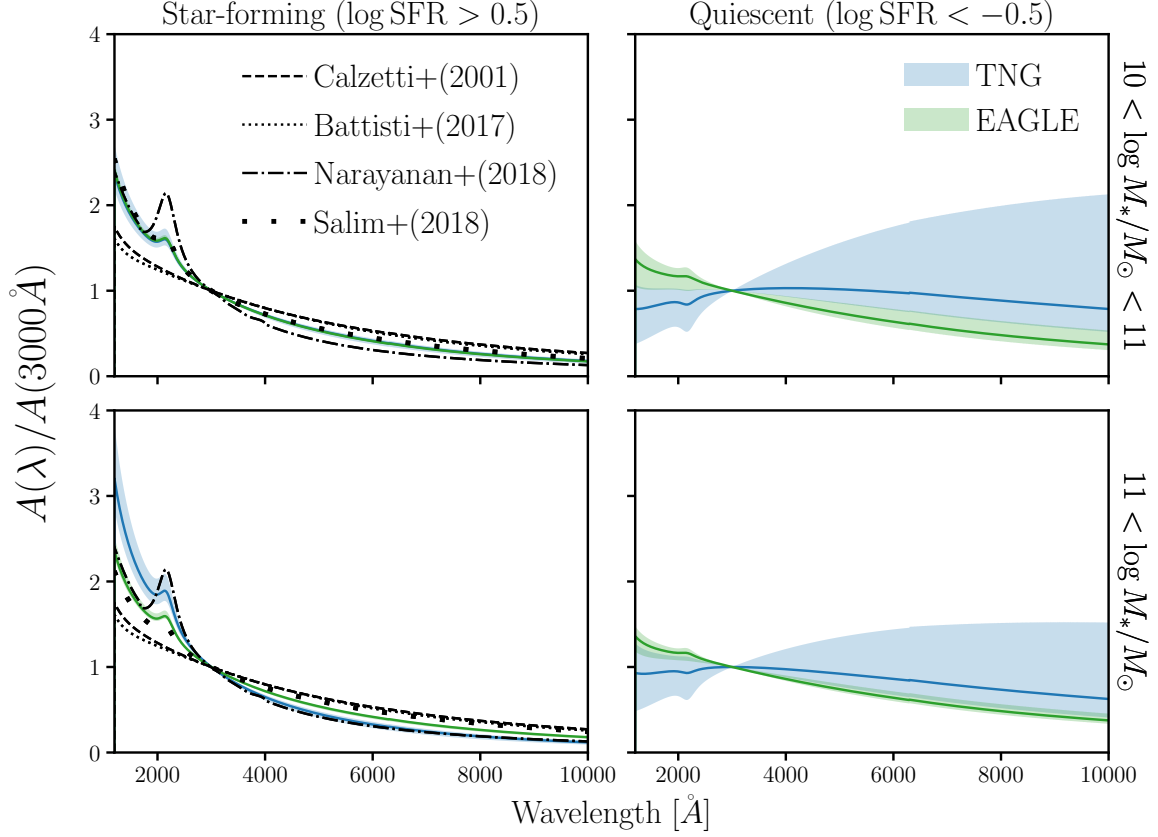


**Figure 6.** The attenuation-slope relation for TNG (blue) and EAGLE (green) simulations with the DEM. We present the relation using two different measurements of slope, commonly used in the literature:  $S = A(1500\text{\AA})/A_V$  (left panel) and the slope offset from the Calzetti (2001) curve,  $\delta$  (right panel). The DEM models predict an attenuation-slope relation, where the slope is steeper at lower attenuation, consistent with both observations and simulation. The DEM models only include massive galaxies, hence, they do not include many galaxies with low attenuation. At  $A_V > 0.5$ , however, the DEM models are in good agreement with observations Salim & Narayanan (2020). In fact, the DEM models match the observed attenuation-slope relation better than radiative transfer simulations, which predict attenuation curves that are too shallow (Inoue 2005; Chevallard et al. 2013; Trayford et al. 2020)

relation is in good agreement with GSWLC2 galaxies (black shaded Salim & Narayanan 2020). They are also consistent with Leja et al. (2017). We also compare our results to theoretical predictions from radiative transfer models, Inoue (2005) (dotted), the radiative transfer models considered in Chevallard et al. (2013) (dot dashed), and Trayford et al. (2020) (light shaded), which all predict shallower attenuation curves than observations. This is also the case for the Narayanan et al. (2018) attenuation curves (not included). Overall, *the dust attenuation curve slopes from the DEM are in excellent agreement with observations and better reproduces the observed attenuation-slope relation than radiative transfer models.*

In Figure 7, we present the normalized attenuation curves of the TNG (blue) and EAGLE (green) DEM for low (top) and high  $M_*$  (bottom), star-forming (left) and quiescent galaxies (right). The attenuation curves are normalized at  $3000\text{\AA}$  in order to emphasize their slopes. Also, although the DEM does not model the complexities of dust-star geometry, it includes significant variations in the dust attenuation through the slab model and the dependence on galaxy properties, which we present with the shaded region ( $1\sigma$  standard deviation about the median). For comparison, we include  $A(\lambda)/A(3000\text{\AA})$  from observations (Calzetti et al. 2000; Battisti et al. 2017; Salim et al. 2018) as well as from simulations (Narayanan et al. 2018). The Calzetti et al. (2000) and Battisti et al. (2017) attenuation curves are derived from  $M_* < 10^{9.9} M_\odot$  star-forming galaxies. Meanwhile, the Salim et al.





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

(2018) curves are for star-forming galaxies with  $10^{9.5} < M_* < 10^{10.5} M_\odot$  star-forming galaxies (top left) and  $10^{10.5} < M_* M_\odot$  star-forming galaxies (bottom left).

The attenuation curves for star-forming centrals in TNG and EAGLE are in good agreement with the Salim et al. (2018) curves as well as the median curve of Narayanan et al. (2018) star-forming galaxies. They are, however, steeper than the Calzetti et al. (2000) and Battisti et al. (2017) curves. Figure 7 also reveals a possible  $M_*$  dependence in the slopes, especially for TNG. Lower  $M_*$  star-forming galaxies have slightly shallower attenuation curves. If this  $M_*$  dependence of  $\delta$  in the DEM extends below our  $M_*$  range, we expect a better agreement the Calzetti et al. (2000) and Battisti et al. (2017) curves as well. Besides their slopes, we also find significant variations in  $A(\lambda)$ . This

is generally consistent with Salim et al. (2018) and Narayanan et al. (2018). They present larger variations; however, this is due to their broader  $M_*$  range. Our results imply that even among star-forming galaxies with similar  $M_*$ , there is significant variation in their attenuation curves.

The DEM also sheds light on dust attenuation in quiescent galaxies, which is particularly valuable since there are many challenges to measuring attenuation curves for quiescent galaxies in observations. For instance, methods that rely on IR luminosities can be contaminated by MIR emission from AGN heating nearby dust Kirkpatrick et al. (2015). Even SED fitting methods require accounting for AGN MIR emission (Salim et al. 2016; Leja et al. 2018; Salim et al. 2018) in addition to the challenge in breaking the degeneracy between SFH and metallicity to fit the continuum (CH: cite?). With the DEM, which forward models optical and UV photometry, we do not face these issues. For both TNG and EAGLE, quiescent galaxies have shallower attenuation curve with large variations in the slopes (Figure 7). They also have higher  $A_V \gtrsim 1.25$ . CH: explanation for why DEM model does this

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

In the DEM, we make a few other assumptions and choices. For simulated galaxies with  $\text{SFR} = 0$ , we directly sample their observables from the distributions of SDSS quiescent galaxies. These  $\text{SFR} = 0$  galaxies do not have recent star-formation and also have 0 gas mass CHH: @tjitske is this for all sims so we would expect them to also have no dust. However, without attenuating these galaxies, the simulations struggle to reproduce observations. Our prescription for  $\text{SFR} = 0$  galaxies ensures that  $\text{SFR} = 0$  galaxies do not impact our results, without delving into the issue further. In Appendix A, we describe our prescription in detail and discuss the limitations of the simulations near the mass and temporal resolutions or their gas prescriptions. Another assumption in the DEM is how we assign  $A_V$  using the slab model (Eq. 3). Although the slab model is consistent with the correlation between attenuation and inclination in observations (*e.g.* Conroy 2010; Salim & Narayanan 2020) and simulations (Chevallard et al. 2013; Narayanan et al. 2018; Trayford et al. 2020, *e.g.*) and reproduces the SDSS  $A_V$  distribution (Figure 8), we test the robustness of our results by replacing it with a more flexible truncated normal distribution in Appendix B. Replacing the slab model, does not significantly impact our results. We therefore conclude that our results are robust to the assumptions and choices we make in the DEM.

In this work, we present the DEM, an empirical framework for including dust attenuation in simulated galaxy populations. We apply the DEM to the SIMBA, TNG, and EAGLE hydrodynamical simulations and forward model optical and UV color-magnitude relations. For all three simulations,

TODO

we reproduce SDSS observations with the DEM. However, based on the inferred DEM parameter posteriors, we find that SIMBA requires an extreme dust attenuation that reverses the established relationship between color and SFR. SIMBA overpredicts a large starburst population at  $< 10^{10} M_{\odot}$  and otherwise struggles to reproduce observations. **any specific subgrid physics in SIMBA that's responsible.**

TODO

Focusing on the DEM for TNG and EAGLE, we find significant  $M_*$  and SFR dependences in the amplitude of dust attenuation,  $A_V$ . More massive galaxies have higher dust attenuation; galaxies with lower SFR have higher dust attenuation. Also, we find that the DEM attenuation curves closely reproduce the observed attenuation–slope relation, better than radiative transfer models. Furthermore, the DEM produces attenuation curves that are in good agreement with the literature for star-forming galaxies and is able to constrain the attenuation curves for quiescent galaxies, which have few constraints from observations. Based on these attenuation curves, quiescent galaxies have shallower attenuation curves than star-forming galaxies. Lastly, we find significant variation in the attenuation curves even for galaxies with comparable  $M_*$  and SFR.

By reproducing SDSS observations with the DEM for different hydrodynamical simulations, we demonstrate that accounting for dust attenuation is essential to reproduce observations. However, based on our current understanding of dust there is enough flexibility to reproduce observations even for simulations that predict galaxy populations with significantly different physical properties. Since adjusting dust alone can reproduce observations, dust is highly degenerate with the variations in subgrid physics across simulations. In other words, if we were to marginalize over dust we would not be able to differentiate between the various hydrodynamical models using observations. So detailed comparisons across simulations and to observations likely overinterpret the differences and similarities found in simulations. Therefore, the current limitations in our understanding of dust is a major bottleneck for investigating galaxy formation using simulations.

The DEM provides a simple empirical model for applying dust attenuation to galaxies. With only a few parameters, it allows  $M_*$  and SFR dependences in dust attenuation curves, reproduces the attenuation–slope relation, and produces attenuation curves with significant variation. Combined with a simulation-based inference method such as ABC-PMC, it provides a direct framework for inferring dust attenuation from observations. Upcoming surveys such as the Bright Galaxy Survey of the Dark Energy Spectroscopic Instrument (DESI; [?, CH: Ruizet et al. 2020](#)) ([desicollaboration2016](#)), the Galaxy Evolution Survey of the Prime Focus Spectrograph (PFS; [Takada et al. 2014](#); [Tamura et al. 2016](#)), and the Wide-Area VISTA Extragalactic Survey (WAVES; [Driver et al. 2016, 2019](#)) will provide more statistically powerful observations to tightly constrain dust attenuation. For those uninterested in dust, the DEM provides a straightforward framework for marginalizing over dust.

**IQ3 description?**

TODO

## 6. SUMMARY

### ACKNOWLEDGEMENTS

It's a pleasure to thank ... This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of High Energy Physics, under contract No. DE-AC02-05CH11231.

This project used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

## APPENDIX

### A. RESOLUTION EFFECTS

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

### B. BEYOND THE SLAB DEM

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

Besides its simplifications, the slab model predicts  $A_V$  distribution with significant differences than the  $A_V$  distributions measured from observations. In Figure 8, we compare the  $A_V$  distribution predicted by the slab model (black) to the  $A_V$  distribution of star-forming galaxies in our SDSS sample (blue). The  $A_V$  values are derived using SED fitting from the Brinchmann et al. (2004) MPA-JHU catalog and how are the SF galaxies classified. The slab model  $A_V$  values are derived using Eq. 3 and 4 with  $M_*$ s and SFRs from the same SDSS sample and the inclinations,  $i$ , are uniformly sampled over the range  $[0, \pi/2]$ . With  $\{m_{\tau,1}, m_{\tau,2}, c_{\tau}\}$  chosen to reproduce the observed  $A_V$  distribution, the slab model can reproduce the overall shape. However, it predicts an extended high  $A_V$  tail not found in observations. TODO

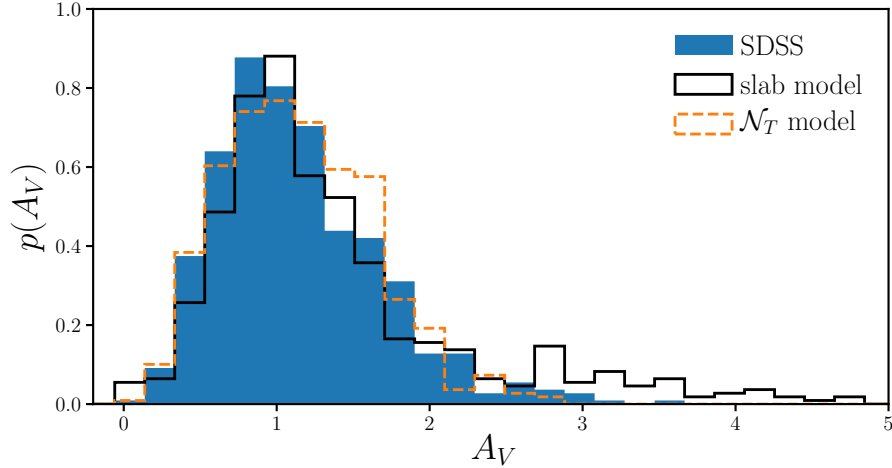
Given these shortcomings of the slab model, we want to ensure that our results do not hinge on the slab model. Modeling the star-to-dust geometries with increased complexities, however, would involve expensive hydrodynamic simulations and dust radiative transfer calculations (*e.g.* Narayanan et al.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

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

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

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

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



**Figure 8.** Comparison of  $A_V$  distribution of SDSS star-forming galaxies (blue) to predictions from the slab model (Eq. 3; black). [detail on how SDSS SF galaxies are classified.](#) The slab model assumes that there’s a slab of dust in front of a galaxy. We use  $\tau_V = 2$  for the slab model above. Regardless of  $\tau_V$ , however, the slab model predicts a significantly more asymmetric and peaked  $A_V$  distribution than observations. Given this disagreement, *we include in our analysis a DEM with an empirical prescription for  $A_V$  based on a truncated normal distribution, which better reproduce the observed  $A_V$  distribution* (Section B).

The  $A_V$  distribution from our truncated normal (orange dashed) closely reproduces the observed SDSS  $A_V$  distribution (Figure 5).  $N_T$  is able to reproduce the overall skewness but unlike the slab model, it does not have a long high  $A_V$  tail. With more free parameters and a functional form that closely resembles the observed  $A_V$  distribution, the truncated normal model provides a flexible alternative to the slab model and we include it in our analysis.

## REFERENCES

- Abazajian K. N., et al., 2009, [The Astrophysical Journal Supplement Series](#), 182, 543
- Alsing J., Wandelt B., Feeney S., 2018, [arXiv:1801.01497 \[astro-ph\]](#)
- Anglés-Alcázar D., Davé R., Faucher-Giguère C.-A., Özel F., Hopkins P. F., 2017, [Monthly Notices of the Royal Astronomical Society](#), 464, 2840
- Baes M., Trčka A., Camps P., Nersesian A., Trayford J., Theuns T., Dobbels W., 2019, [arXiv:1901.08878 \[astro-ph\]](#)
- Battisti A. J., Calzetti D., Chary R.-R., 2016, [The Astrophysical Journal](#), 818, 13
- Battisti A. J., Calzetti D., Chary R.-R., 2017, [The Astrophysical Journal](#), 840, 109
- Beaumont M. A., Cornuet J.-M., Marin J.-M., Robert C. P., 2009, [Biometrika](#), 96, 983
- Blanton M. R., Roweis S., 2007, [The Astronomical Journal](#), 133, 734
- Blanton M. R., et al., 2005, [The Astronomical Journal](#), 129, 2562
- Booth C. M., Schaye J., 2009, [Monthly Notices of the Royal Astronomical Society](#), 398, 53
- Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, [Monthly Notices of the Royal Astronomical Society](#), 351, 1151
- Burgarella D., Buat V., Iglesias-Páramo J., 2005, [Monthly Notices of the Royal Astronomical Society](#), 360, 1413
- Calzetti D., 2001, [New Astronomy Reviews](#), 45, 601



- Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, *The Astrophysical Journal*, 533, 682
- Cameron E., Pettitt A. N., 2012, *Monthly Notices of the Royal Astronomical Society*, 425, 44
- Camps P., Baes M., 2015, *Astronomy and Computing*, 9, 20
- Chabrier G., 2003, *Publications of the Astronomical Society of the Pacific*, 115, 763
- Chevallard J., Charlot S., Wandelt B., Wild V., 2013, *Monthly Notices of the Royal Astronomical Society*, 432, 2061
- Conroy C., 2010, *Monthly Notices of the Royal Astronomical Society*, 404, 247
- Crain R. A., et al., 2015, *Monthly Notices of the Royal Astronomical Society*, 450, 1937
- Dalla Vecchia C., Schaye J., 2012, *Monthly Notices of the Royal Astronomical Society*, 426, 140
- Davé R., Thompson R., Hopkins P. F., 2016, *Monthly Notices of the Royal Astronomical Society*, 462, 3265
- Davé R., Rafieferantsoa M. H., Thompson R. J., 2017a, arXiv:1704.01135 [astro-ph]
- Davé R., Rafieferantsoa M. H., Thompson R. J., Hopkins P. F., 2017b, *Monthly Notices of the Royal Astronomical Society*, 467, 115
- Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, *Monthly Notices of the Royal Astronomical Society*, 486, 2827
- Del Moral P., Doucet A., Jasra A., 2012, *Statistics and Computing*, 22, 1009
- Diggle P. J., Gratton R. J., 1984, *Journal of the Royal Statistical Society. Series B (Methodological)*, 46, 193
- Draine B. T., 2003, *The Astrophysical Journal*, 598, 1017
- Driver S. P., Davies L. J., Meyer M., Power C., Robotham A. S. G., Baldry I. K., Liske J., Norberg P., 2016, *The Universe of Digital Sky Surveys*, 42, 205
- Driver S. P., et al., 2019, *The Messenger*, 175, 46
- Garn T., Best P. N., 2010, *Monthly Notices of the Royal Astronomical Society*, 409, 421
- Genel S., et al., 2014, *Monthly Notices of the Royal Astronomical Society*, 445, 175
- Gordon K. D., Witt A. N., Carruthers G. R., Christensen S. A., Dohne B. C., 1994, *The Astrophysical Journal*, 432, 641
- Hahn C., Vakili M., Walsh K., Hearin A. P., Hogg D. W., Campbell D., 2017a, *Monthly Notices of the Royal Astronomical Society*, 469, 2791
- Hahn C., Tinker J. L., Wetzel A. R., 2017b, *The Astrophysical Journal*, 841, 6
- Hahn C., Tinker J. L., Wetzel A., 2019a, arXiv:1910.01644 [astro-ph]
- Hahn C., Beutler F., Sinha M., Berlind A., Ho S., Hogg D. W., 2019b, *Monthly Notices of the Royal Astronomical Society*, 485, 2956
- Hahn C., et al., 2019c, *The Astrophysical Journal*, 872, 160
- Handley W., Millea M., 2019, *Entropy*, 21, 272
- Hopkins P. F., 2015, *Monthly Notices of the Royal Astronomical Society*, 450, 53
- Hopkins P. F., et al., 2017, arXiv:1707.07010 [astro-ph]
- Inoue A. K., 2005, *Monthly Notices of the Royal Astronomical Society*, 359, 171
- Ishida E. E. O., et al., 2015, *Astronomy and Computing*, 13, 1
- Johnson B. D., et al., 2007, *The Astrophysical Journal Supplement Series*, 173, 392
- Kirkpatrick A., Pope A., Sajina A., Roebuck E., Yan L., Armus L., Díaz-Santos T., Stierwalt S., 2015, *The Astrophysical Journal*, 814, 9
- Kriek M., Conroy C., 2013, *The Astrophysical Journal Letters*, 775, L16
- Leja J., Johnson B. D., Conroy C., van Dokkum P. G., Byler N., 2017, *The Astrophysical Journal*, 837, 170
- Leja J., Johnson B. D., Conroy C., van Dokkum P., 2018, *The Astrophysical Journal*, 854, 62
- Lin C.-A., Kilbinger M., Pires S., 2016, *Astronomy and Astrophysics*, 593, A88
- McAlpine S., et al., 2016, *Astronomy and Computing*, 15, 72
- Narayanan D., Conroy C., Davé R., Johnson B. D., Popping G., 2018, *The Astrophysical Journal*, 869, 70
- Nelson D., et al., 2015, *Astronomy and Computing*, 13, 12
- Nelson D., et al., 2018, *Monthly Notices of the Royal Astronomical Society*, 475, 624
- Noll S., Burgarella D., Giovannoli E., Buat V., Marcillac D., Muñoz-Mateos J. C., 2009, *Astronomy and Astrophysics*, 507, 1793
- Pillepich A., et al., 2018, *Monthly Notices of the Royal Astronomical Society*, 473, 4077



- Pritchard J. K., Seielstad M. T., Perez-Lezaun A., Feldman M. W., 1999, [Molecular Biology and Evolution](#), 16, 1791
- Reddy N. A., et al., 2015, [The Astrophysical Journal](#), 806, 259
- Salim S., Narayanan D., 2020, arXiv:2001.03181 [astro-ph]
- Salim S., et al., 2016, [The Astrophysical Journal Supplement Series](#), 227, 2
- Salim S., Boquien M., Lee J. C., 2018, [The Astrophysical Journal](#), 859, 11
- Schaye J., et al., 2015, [Monthly Notices of the Royal Astronomical Society](#), 446, 521
- Somerville R. S., Primack J. R., 1999, [Monthly Notices of the Royal Astronomical Society](#), 310, 1087
- Somerville R. S., Gilmore R. C., Primack J. R., Domínguez A., 2012, [Monthly Notices of the Royal Astronomical Society](#), 423, 1992
- Springel V., 2005, [Monthly Notices of the Royal Astronomical Society](#), 364, 1105
- Springel V., et al., 2018, [Monthly Notices of the Royal Astronomical Society](#), 475, 676
- Takada M., et al., 2014, [Publications of the Astronomical Society of Japan](#), 66, R1
- Tamura N., et al., 2016, in *Ground-Based and Airborne Instrumentation for Astronomy VI*. eprint: arXiv:1608.01075, p. 99081M, doi:10.1117/12.2232103
- Tavare S., Balding D. J., Griffiths R. C., Donnelly P., 1997, *Genetics*, 145, 505
- Tinker J., Wetzel A., Conroy C., 2011, preprint, 1107, arXiv:1107.5046
- Tinker J. L., Hahn C., Mao Y.-Y., Wetzel A. R., Conroy C., 2018, [Monthly Notices of the Royal Astronomical Society](#), 477, 935
- Trayford J. W., et al., 2015, [Monthly Notices of the Royal Astronomical Society](#), 452, 2879
- Trayford J. W., et al., 2017, [Monthly Notices of the Royal Astronomical Society](#), 470, 771
- Trayford J. W., Lagos C. d. P., Robotham A. S. G., Obreschkow D., 2020, [Monthly Notices of the Royal Astronomical Society](#), 491, 3937
- Tress M., et al., 2018, [Monthly Notices of the Royal Astronomical Society](#), 475, 2363
- Vogelsberger M., et al., 2014, [Monthly Notices of the Royal Astronomical Society](#), 444, 1518
- Weinberger R., et al., 2018, [Monthly Notices of the Royal Astronomical Society](#), 479, 4056
- Weyant A., Schafer C., Wood-Vasey W. M., 2013, [The Astrophysical Journal](#), 764, 116
- Wild V., Charlot S., Brinchmann J., Heckman T., Vince O., Pacifici C., Chevallard J., 2011, [Monthly Notices of the Royal Astronomical Society](#), 417, 1760
- Witt A. N., Gordon K. D., 2000, [The Astrophysical Journal](#), 528, 799