

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

CHANGHOON HAHN^{1, 2, *} AND IQ COLLABORATORY

¹*Lawrence Berkeley National Laboratory, 1 Cyclotron Rd, Berkeley CA 94720, USA*

²*Berkeley Center for Cosmological Physics, University of California, Berkeley CA 94720, USA*

(Dated: DRAFT --- 950689c --- 2020-07-13 --- 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

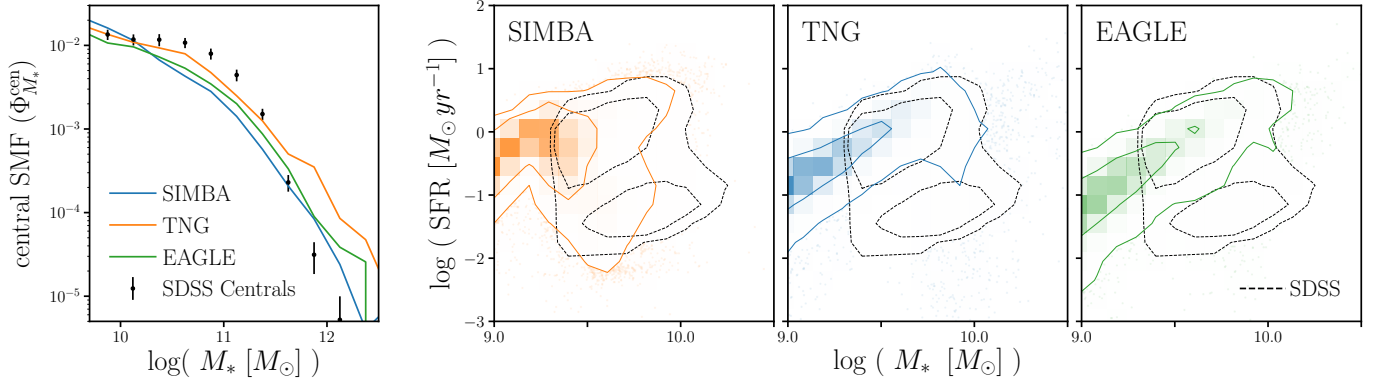


Figure 1. The stellar mass functions ($\Phi_{M_*}^{\text{cen}}$; left-most panel) and $M_* - \text{SFR}$ relation (right panels) of central galaxies from the SIMBA (orange), TNG (blue), and EAGLE (green) simulations. We include, for comparison, $\Phi_{M_*}^{\text{cen}}$ (black) and the $M_* - \text{SFR}$ 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 among the simulations highlight how 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 on top 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)¹. 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 simulation² (hereafter TNG) is a cosmological hydrodynamical simulation of comoving volume (110.7 Mpc)³ (Nelson et al. 2018; Pillepich et al. 2018; Springel et al. 2018). It improves on the original Illustris simulation³ (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 project⁴, 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)³ and is simulated with the ANARCHY code (Dalla Vecchia et al. in prep.; see also Appendix A of Schaye et al. 2015), a modified version of the GADGET-3

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

² <https://www.tng-project.org/>

³ <http://www.illustris-project.org>

⁴ <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 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 ≥ 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.

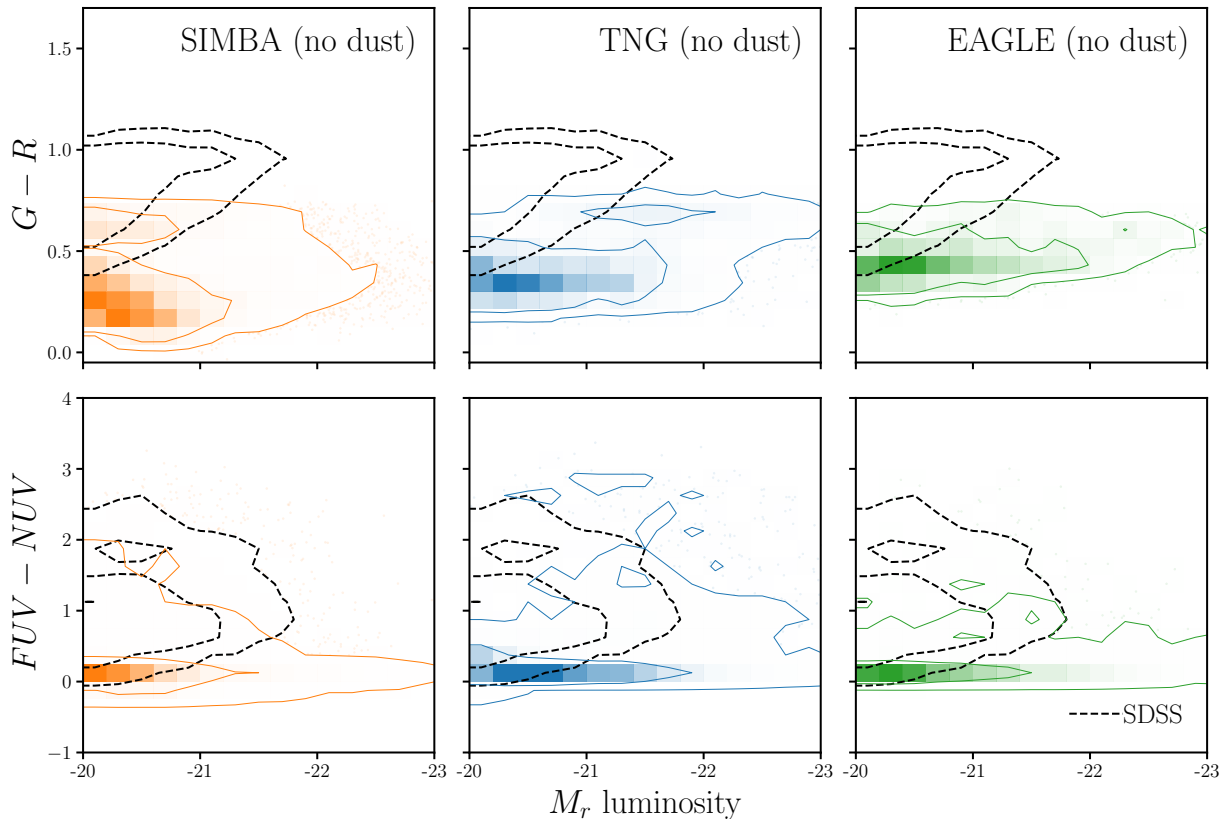


Figure 2. We present the optical and UV color-magnitude relations of central galaxies in the SIMBA (left), TNG (center), and EAGLE (right) simulations. The simulations above do *not* yet include the DEM or any prescription for dust attenuation. $(G-R) - M_r$ (top) and $(FUV-NUV) - M_r$ (bottom) are the 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 Atlas⁵.

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

⁵ <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, τ_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 τ_V . We use τ_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_{τ, M_*} quantifies the M_* dependence, $m_{\tau, \text{SFR}}$ quantifies the SFR dependence, and c_τ quantifies the overall amplitude. Since τ_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 τ_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. δ is the slope offset of the attenuation curve from k_{Cal} . Since δ correlates with galaxy properties (*e.g.* Leja et al. 2017; Salim et al. 2018), we parameterize δ with a similar M_* and SFR dependence as τ_V :

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

$D(\lambda)$ 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 λ_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 δ 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 δ from Kriek & Conroy (2013): $E_b = -1.9 \delta + 0.85$. We note that allowing the slope and amplitude of the E_b and δ 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_{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 τ_v and δ in Eqs. 4 and 6. Due to mass and temporal resolutions, some galaxies in the simulations have $\text{SFR} = 0$ — *i.e.* an unmeasurably low SFR (Hahn et al. 2019c). Eqs. 4 and 6 cannot be used to derive τ_v and δ for these galaxies. Since $\text{SFR} = 0$ galaxies do not account for a large fraction of our simulated galaxies, we directly sample their observables (G , R , NUV , and FUV) from the distribution of observables for SDSS quiescent galaxies. This way, the attenuation of $\text{SFR} = 0$ 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 $\text{SFR} = 0$ 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), τ_V , and δ are calculated for each galaxy using its i , M_* , and SFR. From τ_v and δ , 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.

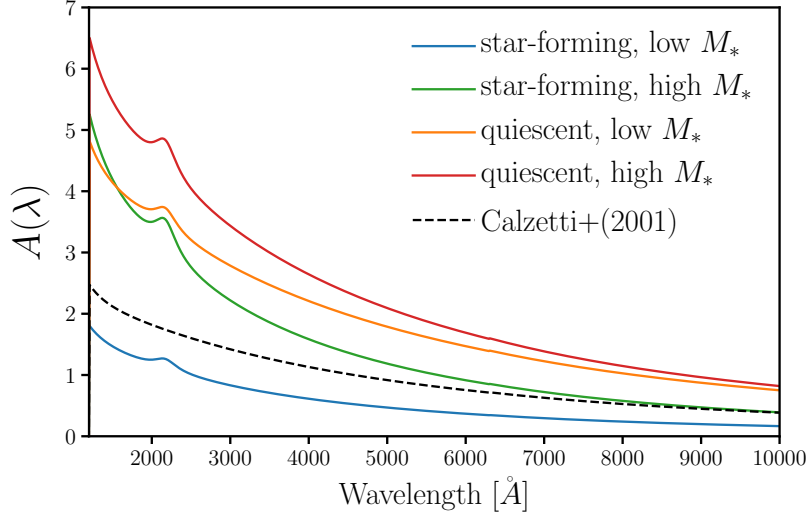


Figure 3. Attenuation curves, $A(\lambda)$, of our Dust Empirical Model (DEM) for galaxies with different SFR and M_* . We include $A(\lambda)$ for star-forming (SFR = $10^{0.5} M_\odot/\text{yr}$) low mass galaxy ($10^{10} M_\odot$; blue), high mass galaxy ($10^{11} M_\odot$; green) and quiescent (SFR = $10^{-2} M_\odot/\text{yr}$) 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 5, 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_{τ,M_*}	Slope of the optical depth, τ_V , $\log M_*$ dependence	flat $[-5., 5.]$
$m_{\tau,\text{SFR}}$	Slope of the optical depth, τ_V , $\log \text{SFR}$ dependence	flat $[-5., 5.]$
c_τ	amplitude of the optical depth, τ_V	flat $[0., 6.]$
m_{δ,M_*}	Slope of the attenuation curve slope offset, δ , $\log M_*$ dependence	flat $[-4., 4.]$
$m_{\delta,\text{SFR}}$	Slope of the attenuation curve slope offset, δ , $\log \text{SFR}$ dependence of	flat $[-4., 4.]$
c_δ	amplitude of the attenuation curve slope offset, δ	flat $[-4., 4.]$
f_{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_{τ,M_*} , $m_{\tau,\text{SFR}}$, c_τ 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_{δ,M_*} , $m_{\delta,\text{SFR}}$, c_δ are chosen to conservatively 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 note that uniform priors on the DEM parameters do not result in uniform priors on τ_V or δ (*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 τ_V or δ .

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:

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

θ_{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 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 ($i > 1$), the threshold, ϵ_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 ϵ_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, derived from ABC-PMC, of the DEM parameters for the SIMBA (orange), TNG (blue), and EAGLE (green) hydrodynamical simulations. The contours mark the 68% and 95% confidence intervals. The DEM parameters include m_{τ, M_*} , $m_{\tau, \text{SFR}}$, and c_τ , which parameterize the M_* dependence, SFR dependence, and amplitude of τ_V , as well as m_{δ, M_*} , $m_{\delta, \text{SFR}}$, and c_δ , which parameterize the M_* dependence, SFR dependence, and amplitude of δ (Table 1).

5. RESULTS

In Figure 5, we present the optical and UV color-magnitude relations predicted by the DEM with median of the 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. In the UV color-magnitude space, the simulations with DEM

CH: something about clear bimodality with the red sequence and blue cloud In contrast, *using DEMs we produce $(G-R)-M_r$ and $(FUV-NUV)-M_r$ relations consistent with SDSS for all of the simulations.*

We examine the observables more closely in Figure ??, where we present the $G-R$ (left) and $FUV-NUV$ (right) distributions in M_r bins for the DEM models of the SIMBA (orange), TNG (blue), and EAGLE (green) simulations. The top panels contain galaxies with $-20 > M_r > -21$ and the bottom panels contain galaxies with $-21 > M_r > -22.5$. In each panel we include the SDSS distributions for comparison (grey shaded). These distributions are number densities. The positions of the red sequence and blue cloud in the $G-R$ distributions for the DEM models are consistent with the SDSS distribution. They also produce $FUV-NUV$ distributions with peaks consistent with observations. We also find good agreement in the overall normalization (number density), especially for the $-20 > M_r > -21$ bin.

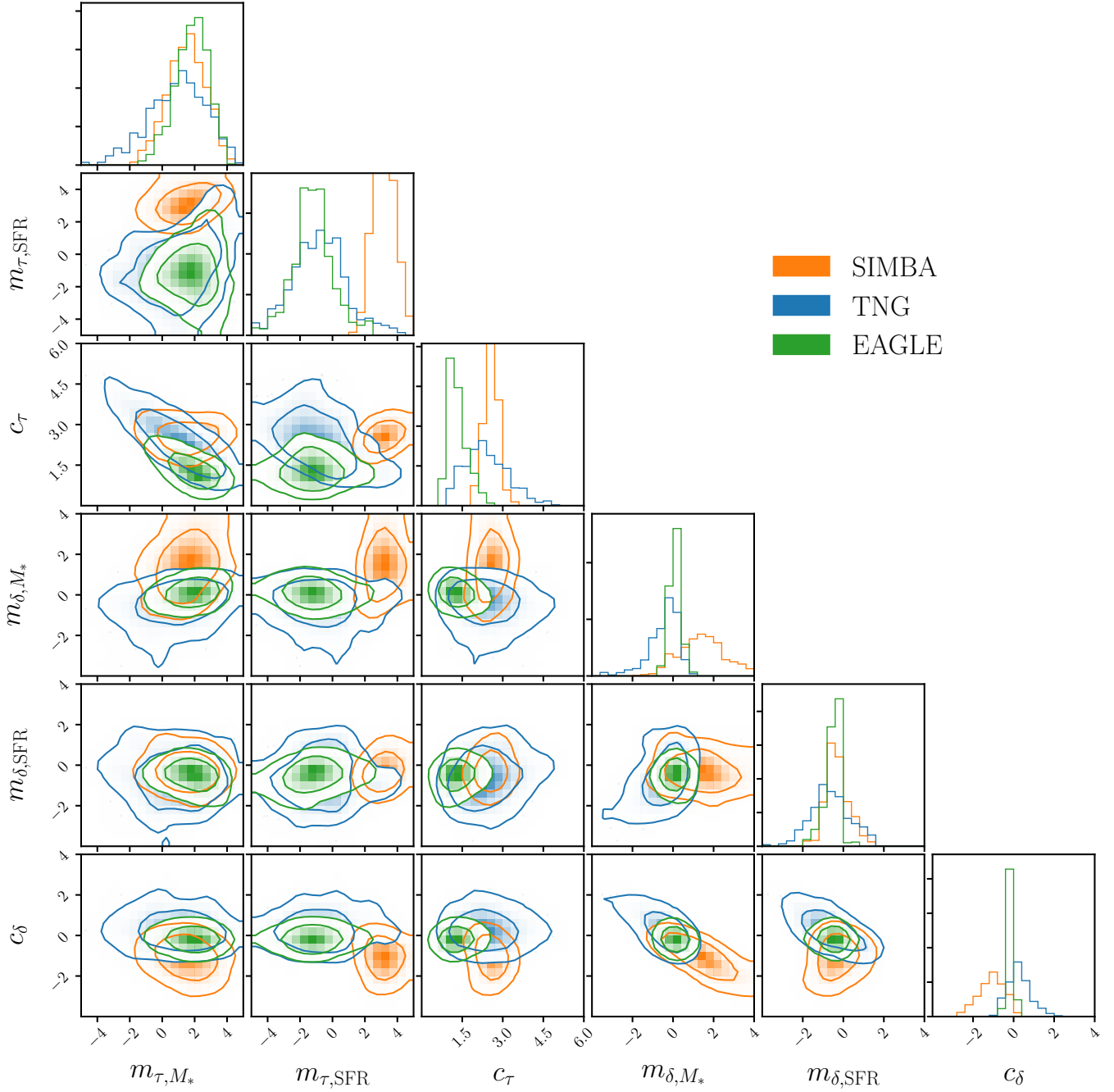


Figure 4. Posterior distributions of DEM parameters for the SIMBA (orange), TNG (blue), and EAGLE (green) hydro simulations. The contours mark the 68% and 95% confidence intervals. The posteriors are derived using Approximate Bayesian Computation with Population Monte Carlo (Section 4). **CH: REVISIT:**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

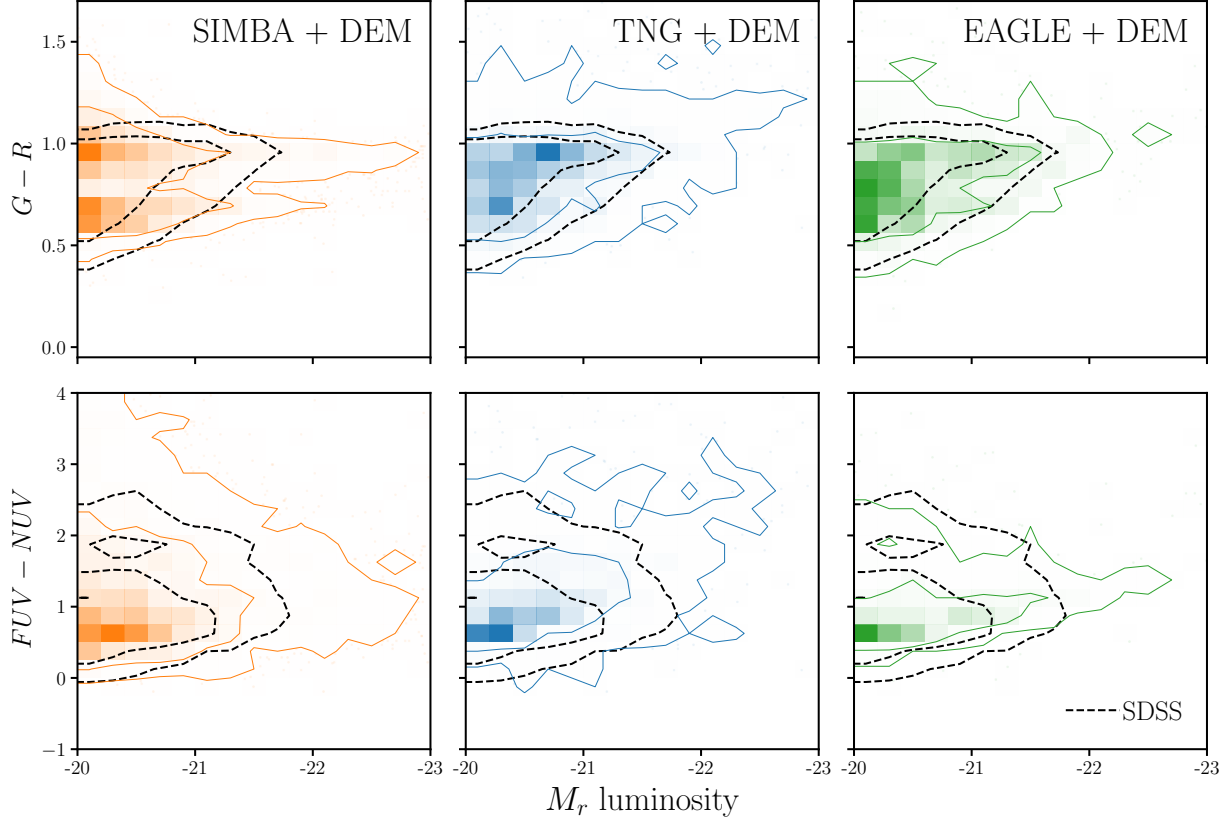


Figure 5. $(G - R) - M_r$ color-magnitude (top panels) and $(FUV - NUV) - M_r$ (bottom panels) relations predicted by the median DEM posteriors for the SIMBA (orange), TNG (blue), and EAGLE (green) hydro simulations. For comparison, we include the observables for SDSS in the left-most panel (Section 2.5). The median posterior DEMs 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.

There are also a few discrepancies between the DEM models and SDSS observations. First, the red sequence in the observed $G - R$ distribution has a sharp red color cut-off. In contrast, we do not find as sharp of a $G - R$ cut off in the DEM models, and find a small fraction of galaxies redder than observations. Furthermore, while the DEM models for TNG and EAGLE agree with SDSS, the DEM model for SIMBA overpredicts blue galaxies. **CH: what do we want to do with SIMBA?** We also find that the TNG and EAGLE DEM models slightly overpredict galaxies in the higher luminosity bin and produce a broader $G - R$ distribution. Nonetheless, Figure ?? demonstrates that overall the DEM models for SIMBA, TNG, and EAGLE produce observables are in good agreement with SDSS.

Previous works in the literature have also presented models that predict colors and luminosities for different simulations and dust models and compared them to observations. For instance, Trayford et al. (2015) calculate colors and luminosities for $z = 0.1$ galaxies using EAGLE with the GALAXEV population synthesis models and a two-component screen model for dust. Compared to GAMA obser-

vations, their model produces a bluer red sequence and overpredicts luminous blue galaxies (Trayford et al. 2015, Figure 5). Although a detailed comparison is difficult since they compare all galaxies, not just centrals, we note that the DEM models find good agreement in the positions of the red sequences. Furthermore, for TNG and EAGLE, the DEM models do not overpredict blue galaxies. Even for SIMBA, the DEM model overpredict blue galaxies by a smaller amount.

More recently, Trayford et al. (2017) calculated optical colors for the EAGLE simulation using SKIRT, a Monte Carlo radiative transfer code (Camps & Baes 2015), to model the dust. Trayford et al. (2017) compares all galaxies so again, we do not include a direct comparison. Compared to GAMA, while they find good agreement with observations at intermediate masses, $10^{10.5} < M_* < 10^{10.8} M_\odot$, they again find a bluer red-sequence at $10^{11.2} < M_* < 10^{11.5}$. While we do not present comparisons in M_* bins, which compare SED derived M_* to the predicted M_* from simulations, we find that the position of the red sequence in the DEM models are in good agreement with SDSS even at $M_* > 10^{11.2}$. Our models, however, predict an overall broader color distribution at the high mass end. Using the same Trayford et al. (2017) EAGLE and SKIRT framework, Baes et al. (2019) compared the predicted cosmic spectral energy distributions (CSED) to observations. While this comparison averages over the galaxy populations, they find the EAGLE-SKIRT CSED overestimates the observed CSED in the UV regime. Moreover, the $FUV - NUV$ color of their CSED is significantly higher than GAMA $FUV - NUV$. The DEM models on the other hand, predict $FUV - NUV$ in good agreement with observations.

Besides with EAGLE, Nelson et al. (2018) calculated optical colors for the TNG simulations with a dust model that includes attenuation due to dense gas birth clouds surrounding young stellar populations and also attenuation due to simulated distribution of neutral gas and metals. Although they compare the color distribution for all galaxies in bins of M_* , so we cannot directly compare to the DEM models, compared to SDSS they find bluer red sequence peak position and narrower blue cloud. We find neither of these discrepancies between the DEM models and SDSS. **CH: restatement of how the DEM models have the flexibility to reproduce the optical and UV color-magnitude relationship.**

The simulations with DEMs predict observables in agreement with observations despite the significant differences in the SMFs and M_* -SFR relations (Figure ??). In other words, the DEM has the flexibility to reproduce observations even when simulations predict galaxy populations with significantly different physical properties. We emphasize that the DEM is based on the standard prescriptions for dust attenuation and, thus, serve as a flexible parameterization within the bounds of our current understanding of dust in galaxies.

Figures 5 and ?? highlights two key points. First, any comparison of simulations must account for dust. Dust entirely changes the predictions of simulations in observables-space. Without dust, we did not find bimodality in the color-magnitude relation. Fortunately, the DEM provides a simple framework for including dust motivated by attenuation laws and correlations with the physical properties of galaxies. Second, the current limitations in our understanding of dust in galaxies significantly impedes our ability to understand galaxy formation from simulations. To robustly interpret any comparison of simulations, we would need to marginalize over dust (*e.g.* DEM parameters). Since DEMs can

produce consistent observables for a range of simulations, marginalizing over dust would leave little constraining power on the subgrid prescriptions (*i.e.* galaxy physics) of the simulations.

The DEMs demonstrate that simulations can closely match observations by varying dust attenuation. They therefore illustrate how dust is a major bottleneck for directly interpreting galaxy simulations for insights into galaxy formation. In addition, DEMs also provide some insight into dust. Given our parameterization (Section 3), it is especially easy to interpret correlation between dust attenuation and galaxy physical properties. For instance, the posteriors of DEM parameters in Figure 4 reveal a number of consistent trends across the simulations. In all three simulations, we find significant positive M_* dependence of τ_V : $m_{\tau, M_*} \sim 2$. Regardless of the underlying hydro simulations, we find that *galaxies with higher M_* have overall higher dust attenuation*.

This M_* dependence is consistent with previous works in the literature. The seminal work of Burgarella et al. (2005), for instance, found significant positive M_* dependence in FUV attenuation in NUV-selected and FIR-selected samples of Buat et al. (2005) and Iglesias-Páramo et al. (2006). Garn & Best (2010) also find positive M_* dependence in Balmer decrement-based $H\alpha$ attenuation in $\sim 90,000$ SDSS star-forming galaxies. Battisti et al. (2016) similarly find higher Balmer optical depth for higher M_* in $\sim 10,000$ star-forming galaxies from GALEX and SDSS. 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). citation is a bit SDSS heavy. Anything else in the literature?

TODO

In addition to the M_* dependence, the DEM posteriors also allow us to examine the correlation between dust attenuation and star formation. For TNG and EAGLE, we infer DEM posteriors with $m_{\tau, \text{SFR}} \sim -1$ (Figure 4). This means that TNG and EAGLE require higher attenuation for galaxies with lower SFR — *i.e.* quiescent galaxies have higher dust attenuation overall. While previous works that have examined the relationship between dust attenuation and SFR in observations (*e.g.* Garn & Best 2010; Reddy et al. 2015; Battisti et al. 2016, 2017), they focus solely on star-forming galaxies. While they find that star-forming galaxies with higher SFR have higher attenuation, much of this trend is driven by the star-forming sequence (more massive star-forming galaxies have higher SFR; Garn & Best 2010; Battisti et al. 2017). At fixed M_* , observations find no strong SFR dependence for the SF population.

For SIMBA, unlike for TNG and EAGLE, we infer $m_{\tau, \text{SFR}} \sim 3$: galaxies with higher SFR have higher dust attenuation (Figure 4). In fact, the attenuation is so high for star-forming galaxies that they populate the red sequence rather than the blue cloud in the color-magnitude relation. This extreme SFR dependence in the dust attenuation that results in a contradiction of established color-SFR relations, is due to the fact that SIMBA predicts a population of star-forming with exceptionally high SFR, that seemingly lie above the SFS (Figure 1). In a SIMBA DEM model with $m_{\tau, \text{SFR}} < 0$, these high SFR galaxies would be high luminosity blue galaxies, not found in observations. CH: If we impose a $m_{\tau, \text{SFR}} < 0$ prior for the SIMBA DEM model, we struggle to reproduce observables consistent with SDSS. The difference in $m_{\tau, \text{SFR}}$ among the hydro simulations demonstrates that, in addition to insights on dust in galaxies, our DEM approach can also highlight differences and limitations among simulations. Moreover, it further highlights that dust attenuation can be adjusted, within priors set

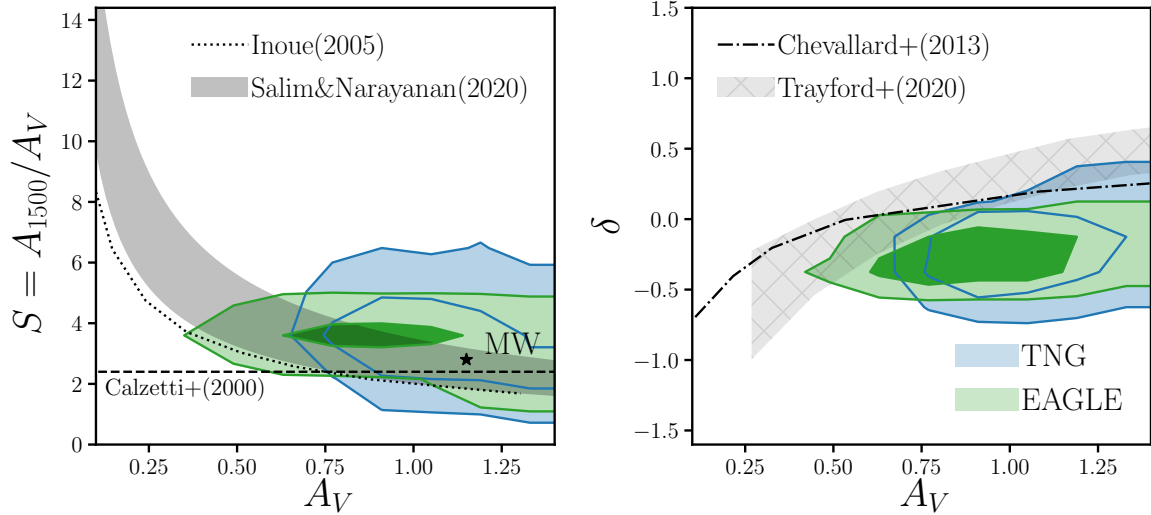


Figure 6. The attenuation-slope relation for the TNG (bue) and EAGLE (green) DEM models. We present the relation using two different measurements of slope, commonly used in the literature: $S = A(1500\text{\AA})/A_V$ (left panel) and the slope offset from the Calzetti (2001) curve, δ (right panel). The 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)

by observations, so that any simulation can match observations. Given this significant discrepancy for SIMBA, we focus the rest of the discussion on TNG and EAGLE. **what subgrid physics is causing SIMBA to overpredict low mass starbursts compared to TNG and EAGLE?** TODO

In addition to the dependence of A_V on galaxy properties, our results also shed light on the slope of dust attenuation as well as its correlation with galaxy properties. If we measure the slope of the attenuation curve of the DEM models using the quantity $S = A(3000\text{\AA})/A_V$, which is easier to constrain observationally, we find that the TNG and EAGLE DEM models both predict slopes within $2 < S < 5$ and centered around $S \sim 3.5$ (Figure 6). In comparison, previous observational works that measure dust attenuation in star-forming galaxies also find slopes within the range $2 < S < 5$ (Calzetti et al. 2000; Burgarella et al. 2005; Johnson et al. 2007; ?; Wild et al. 2011; Battisti et al. 2016, 2017). Leja et al. (2017) and Salim et al. (2018) find a broader range of slopes in their galaxy samples, $2 < S < 15$. However, their galaxies with steeper slopes, $S > 6$, have $A_V < 0.4$, which we are not in our sample (we discuss this in further detail below). Excluding those galaxies, we find good agreement. Also for reference, the Small Magellanic Cloud and the Milky Way have attenuation slopes $S_{SMC} = 4.8$ and $S_{MW} = 2.8$. **CH: The dust attenuation curve slopes predicted by the TNG and EAGLE DEM models are in excellent overall agreement with observations.**

In addition to the amplitude, we can also examine the correlation between attenuation curve slopes and galaxy properties. For TNG and EAGLE, we find little M_* and SFR dependence in δ :

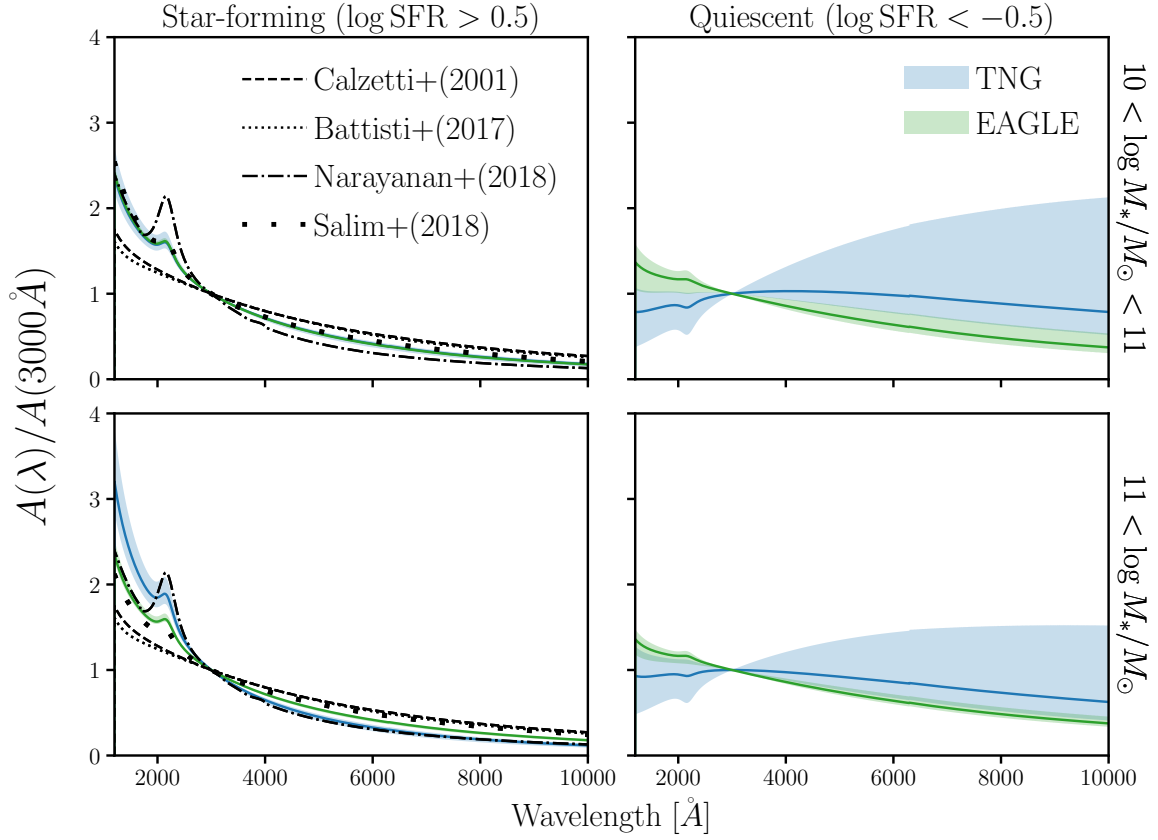


Figure 7. Attenuation curves of the TNG (blue) and EAGLE (green) DEM models for low (top) and high M_* (bottom), star-forming (left) and quiescent galaxies (right). The attenuation curves are normalized at 3000\AA : $A(\lambda)/A(3000\text{\AA})$. We mark the 1σ standard deviation of the attenuation 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.

both m_{δ, M_*} and $m_{\delta, \text{SFR}}$ are consistent with 0. To date most observational works do not examine the correlation between slope and galaxy properties and have been limited to star-forming galaxies. Leja et al. (2017) recently measured the slope of the attenuation curves of galaxies, some of which are non-star-forming. However, with only 129 galaxies the trends are difficult to meaningfully interpret. More recently, Salim et al. (2018) measured attenuation curves of 23,000 galaxies from the GALEX-SDSS-WISE Legacy Catalog 2 (GSWLC2; Salim & Boquien 2019). While they also mainly focus on star-forming galaxies, Salim et al. (2018) also find that quiescent galaxies have significantly steeper curves. This is in disagreement with our $m_{\delta, \text{SFR}} \sim 0$; however, we note that Salim et al. (2018) also find significantly steeper curves for the starburst population (galaxies above the SFS). CH: Leja

et al. (2017) finds the opposite — composites and AGN galaxies have shallower slopes, although with limited statistics (129 galaxies). Since the DEM model only includes a linear dependence on SFR, it's hard to tell. Furthermore, for star-forming galaxies, Salim et al. (2018) find a significant M_* dependence on δ , unlike our $m_{\delta, M_*} \sim 0$ constraint. They find, however, this dependence is entirely driven by the underlying fundamental trend between slope and A_V .

At low attenuation, dust scattering dominates absorption. In this regime, the attenuation curve steepens because red light tends to scatter isotropically while blue light tends to forward scatter, which causes more optical-to-IR light to escape the galaxy than UV light (Witt & Gordon 2000; ?). On the other hand, at high attenuation dust absorption is dominant and the attenuation curve is shallower (Chevallard et al. 2013). In Figure 6, we present the A_V and slope relationship for attenuation curves from the TNG (blue) and EAGLE (green) DEM models. On the left panel we quantify slope using $S = A(1500\text{\AA})/A_V$; on the right panel, we use δ . In both panels, the DEM models predict consistent A_V –slope relations where the slope is steeper at lower attenuation. Furthermore, although the comparison is limited by the A_V range of the DEM models, we find that the DEM A_V –slope relations are consistent with the $A_V - S$ relation of 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 (Witt & Gordon 2000; Inoue 2005; Chevallard et al. 2013). Our DEM models are in good agreement with Witt & Gordon (2000). However, Inoue (2005) and the radiative transfer models considered in Chevallard et al. (2013) predict shallower attenuation curves than our DEM models. Narayanan et al. (2018), not included in Figure 6, also predict shallower attenuation curves (Salim & Narayanan 2020). Overall, our DEM models better reproduce the A_V –slope relations of observations than these theoretical predictions. CH: also include Trayford et al. (2020)

The DEM does not explicitly model the complexities of dust-star geometry on a galaxy-by-galaxy basis. However, through the slab model and also the dependence of the attenuation on galaxy properties, it includes significant variation in the attenuation. In Figure 7, we present the normalized attenuation curves of the SIMBA (orange), 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\AA : $A(\lambda)/A(3000\text{\AA})$. We represent the variation in the attenuation curves (1σ standard deviation about the median) 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). We note that the Calzetti et al. (2000) and Battisti et al. (2017) attenuation curves are for star-forming galaxies with significantly lower M_* . The Battisti et al. (2017) curve, for instance, is derived from $M_* < 10^{9.9} M_\odot$ star-forming galaxies. Meanwhile, the Salim et al. (2018) curves we include are specifically for $10^{9.5} < M_* < 10^{10.5} M_\odot$ star-forming galaxies (top left), $10^{10.5} < M_* M_\odot$ star-forming galaxies (bottom left), and quiescent galaxies (top and bottom right panels).

Focusing first on the attenuation curves of star-forming galaxies (left panels), we find that the Calzetti et al. (2000) and ? curves are significantly shallower than the TNG and EAGLE DEM attenuation curves. However, at least part of the discrepancy is driven by differences in M_* . And,

although we do not probe down to low M_* , the TNG and EAGLE attenuation curves get shallower at lower M_* so we expect better agreement with Calzetti et al. (2000) and ? at the M_* range they probe. Next, compared to Salim et al. (2018), we find good agreement among the attenuation curves for star-forming galaxies. In addition, the variation we find in the DEM attenuation curves of star-forming galaxies is consistent with the variation Salim et al. (2018) presents in their Figure 9. The DEM attenuation curves are also in excellent agreement with the median attenuation curve of Narayanan et al. (2018) star-forming galaxies. While Narayanan et al. (2018) find larger attenuation curve variation than the DEM models, this is partly due to the fact that their cosmological zoom-in simulations span a broader M_* range (Abruzzo et al. 2018). Overall, the DEM models find significant variation in the attenuation curves of star-forming galaxies consistent with both observations and simulations.

In addition to constraining the attenuation curve of star-forming galaxies, the DEM models also shed light on dust attenuation in quiescent galaxies, which is particularly valuable due to the observational challenges of measuring dust attenuation in quiescent galaxies. For instance, MIR emission from active galactic nuclei (AGN) heating nearby dust complicates methods that rely on IR luminosity to measure dust attenuation Kirkpatrick et al. (2015). Even SED fitting methods require accounting for AGN MIR emission (??Salim et al. 2018) in addition to the overall challenge of breaking the degeneracy between SFH and metallicity to fit the continuum of quiescent galaxy SEDs (CH: cite?). With DEM models, which take a forward modeling approach with optical and UV data, we do not face these issues and can constrain the attenuation curves of quiescent galaxies. For both TNG and EAGLE, we find that quiescent galaxies have shallower attenuation curve with large variations in the slopes (Figure 7). We also find that they have higher $A_V \gtrsim 1.25$. CH: explanation for why DEM model does this In Leja et al. (2017), they similar find composite and AGN galaxies to have shallow attenuation curves with higher A_V ; however, the comparison is limited due to their smaller sample size (129 galaxies). In contrast, Salim et al. (2018) find that quiescent galaxies in GSWLC2 have significantly steeper curves; they, however, focus their analysis mainly on star-forming galaxies.

In our DEM model, we neglect the galaxies with $\text{SFR}=0$ from simulations by directly sampling the quiescent galaxy observables (Section A). Since galaxies with $\text{SFR}=0$ predicted by the simulations do not have recent star-formation and also have 0 gas mass@tjitske is this for all sims, we would expect them to also have no dust. However, a DEM model where we do not attenuate $\text{SFR}=0$ galaxies struggles to reproduce observations. This likely highlights the limitations of hydrodynamical simulations near the mass and temporal resolutions as well as the limitations of subgrid prescriptions for gas. CH: Uncertainties in the dust destruction timescale may also contribute to the explanation (e.g. Jones & Nuth 2011; Slavin et al. 2015). We emphasize that our prescription for $\text{SFR}=0$ galaxies ensures that they do not significantly impact the constraints on DEM parameters.

For the DEM model in this paper, we fix the amplitude of the UV dust bump to the amplitude of δ (Eq. ??). If, rather than fixing m_E and c_E in Eq. ??, we allow them to be free parameters we find little constraining power on these parameters. More importantly, we find consistent constraints for the other DEM parameters, which implies that fixing the UV bump parameters does not impact our results.

TODO

A major assumption in the DEM model is how we sample the A_V using the slab model. For each simulated galaxy, the DEM model uniformly samples the inclination and then assigns A_V with the slab model (Eq. 3). As we detail in Section 3, this slab model is consistent inclination dependence of attenuation found in both observations (*e.g.* Salim & Narayanan 2020) and simulations (Chevallard et al. 2013; Narayanan et al. 2018; Trayford et al. 2020, *e.g.*). Moreover, the slab model successfully reproduces the A_V distribution of SDSS galaxies (Figure 8). To ensure that our results do not hinge on the slab model, we conduct our analysis using a truncated normal distribution based DEM model to sample A_V (Appendix B). This model is more flexible than the slab model. For instance, we include physical parameter dependence for the scatter of A_V , and not just for τ_V . We refer readers to Appendix B for details. We find that adding our results are not impacted by the change in the A_V sampling model or the added flexibility. Therefore, our results are not driven by our choice of the slab model.

CH: How about our prior choice?

CH: what we learn about in the simulations

CH: what we learn about dust paragraph on restating how we can learn about dust through DEMs based on trends we see across all simulations. summarize main findings again.

CH: observables unaffected by dust/dem? Are there observables that hydro sims + DEMs cannot reproduce? What does that say about the hydro sims? What observables are unaffected by DEMs? We should chase those observables.

CH: don't over-interpret observations We clearly have to be careful with overinterpreting hydro sims because modifying dust allows us to reproduce whatever we want.

- Should we bother calibrating our empirical and semi-analytic models to hydrodynamic simulations when the hydro sims also require marginalizing over dust parameters? Does this mean that if our goal is to make realistic mocks, we can be relatively careless about

CH: What are some applications for DEMs? Realistic mock catalogs that reproduce observations in observable-space rather than physical parameter space. **CH:** talk about IQ paper where we use best-fit dust

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 τ_V , the attenuation depends solely on the orientation of the galaxy. This simplification, ignores any complexities in the star-to-dust geometry that impact the shape of the attenuation curve (Witt Gordon 1996, 2000, Seon Drain 2016).

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

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

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

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

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

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

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

REFERENCES

Abazajian K. N., et al., 2009, *The Astrophysical Journal Supplement Series*, 182, 543

Abruzzo M. W., Narayanan D., Davé R., Thompson R., 2018, arXiv e-prints, 1803, arXiv:1803.02374

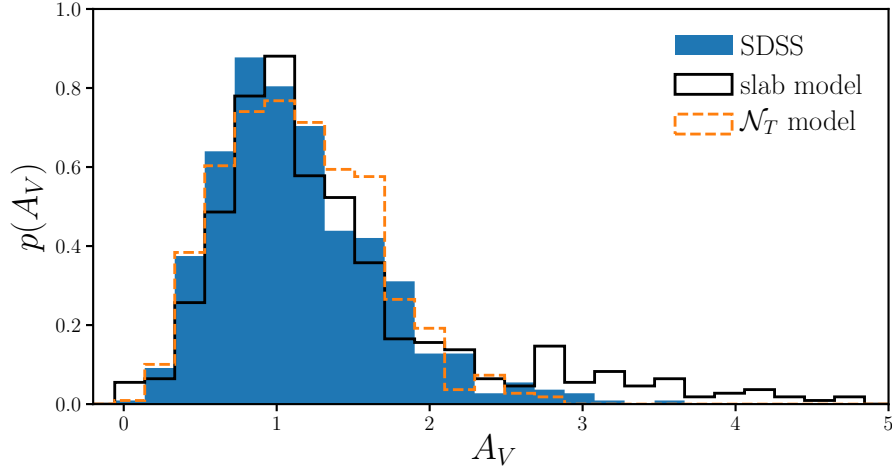


Figure 8. Comparison of A_V distribution of SDSS star-forming galaxies (blue) to predictions from the slab model (Eq. 3; black). **detail on how SDSS SF galaxies are classified.** The slab model assumes that there’s a slab of dust in front of a galaxy. We use $\tau_V = 2$ for the slab model above. Regardless of τ_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).

- 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
- Buat V., et al., 2005, *The Astrophysical Journal Letters*, 619, L51
- 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
- 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., Rafiefferantsoa 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., Rafiefferantsoa 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
- 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
- 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 \[astro-ph\]](#)
- Iglesias-Páramo J., et al., 2006, [The Astrophysical Journal Supplement Series](#), 164, 38
- 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
- Jones A. P., Nuth J. A., 2011, [Astronomy and Astrophysics](#), 530, A44
- 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
- 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., Boquien M., 2019, [The Astrophysical Journal](#), 872, 23
- Salim S., Narayanan D., 2020, [arXiv:2001.03181 \[astro-ph\]](#)
- 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
- Slavin J. D., Dwek E., Jones A. P., 2015, [The Astrophysical Journal](#), 803, 7
- 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
- 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