

## Empirical Dust Models

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 --- ffe4844 --- 2020-05-08 --- NOT READY FOR DISTRIBUTION)

### ABSTRACT

dust

*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](#); [?; ?](#); [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.

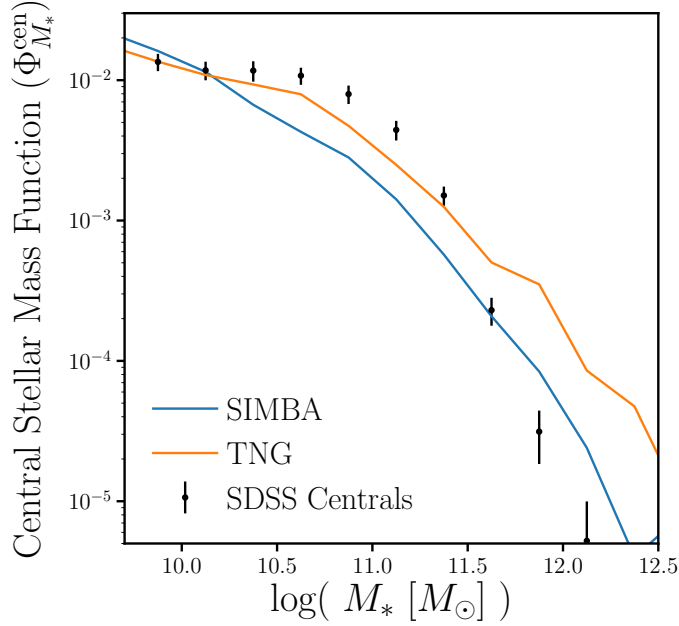
Moreover, because the radiative transfer method is expensive it's hard to compare many different simulations. Not only that, observables generated from simulations that take into radiative transfer dust models complicates simulation to simulation comparisons. Because you're simultaneously comparing the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

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

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

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



**Figure 1.** The stellar mass functions of central galaxies,  $\Phi_{M_*}^{\text{cen}}$ , of the SIMBA (orange) and TNG (blue) simulations compared to the SDSS  $\Phi_{M_*}^{\text{cen}}$ . The uncertainties for the SDSS  $\Phi_{M_*}^{\text{cen}}$  are derived using jackknife resampling and SDSS centrals are identified using a halo-based group finder (Section 2.1). For SIMBA and TNG galaxies, we calculate  $M_*$  as the total stellar mass within the host halo, excluding contributions from any subhalos; centrals are classified based on their individual definition (Section 2.2 and 2.3). *should we include total SMF...? The simulations and observations have loosely consistent  $\Phi_{M_*}^{\text{cen}}$ .*

observations to constrain the DEM. This framework allows us to learn about attenuation curves given a model for galaxy formation.

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.

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

## 2. DATA

### 2.1. SDSS DR7 Central Galaxies

Throughout the paper we compare the simulations and models described below to the observed SDSS central galaxy sample from the Tinker et al. (2011) group catalog. The group catalog, first, selects 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 NYU Value-Added Galaxy Catalog (VAGC; Blanton et al. 2005) of SDSS DR7 (Abazajian et al. 2009). The stellar masses are estimated using the `kcorrect` code (Blanton & Roweis 2007) assuming a Chabrier (2003) initial mass function.

Central galaxies are then 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. 2018).

## 2.2. *Illustris TNG*

describe what galaxy properties (SFH, ZH, etc) are available

TODO

## 2.3. *SIMBA*

describe what galaxy properties (SFH, ZH, etc) are available

TODO

In Figure 1, we compare the stellar mass function (SMF) of our SDSS central galaxy sample along with central galaxy SMFs of the SIMBA (orange) and TNG (blue) simulations. The uncertainties for the SDSS SMF are derived from jackknife resampling. Although we present the SMFs for reference, we do not use stellar masses throughout the paper since they are inconsistently defined among simulations and observations. Instead, we compare between the simulations and SDSS using luminosity,  $M_r$ , which we consistently forward model and measure in the simulations. In these comparisons, we restrict ourselves to galaxies brighter than  $M_r < -20$ , where our SDSS central galaxy sample is complete.

## 2.4. *Spectral Energy Distributions*

describe how the SED is generated using the SFH and ZHs

TODO

## 2.5. *Forward Modeling SDSS Photometry and Spectra*

# 3. DUST EMPIRICAL MODELING

## 3.1. *Fiducial DEM*

motivation for the DEM model

TODO

We begin by defining 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,

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

For the normalization of the attenuation curve,  $A_V$ , we use the slab model from Somerville & Primack (1999); Somerville et al. (2012). In the slab model the amplitude of attenuation depends on the inclination angle,  $i$ , and the optical depth,  $\tau_V$ :

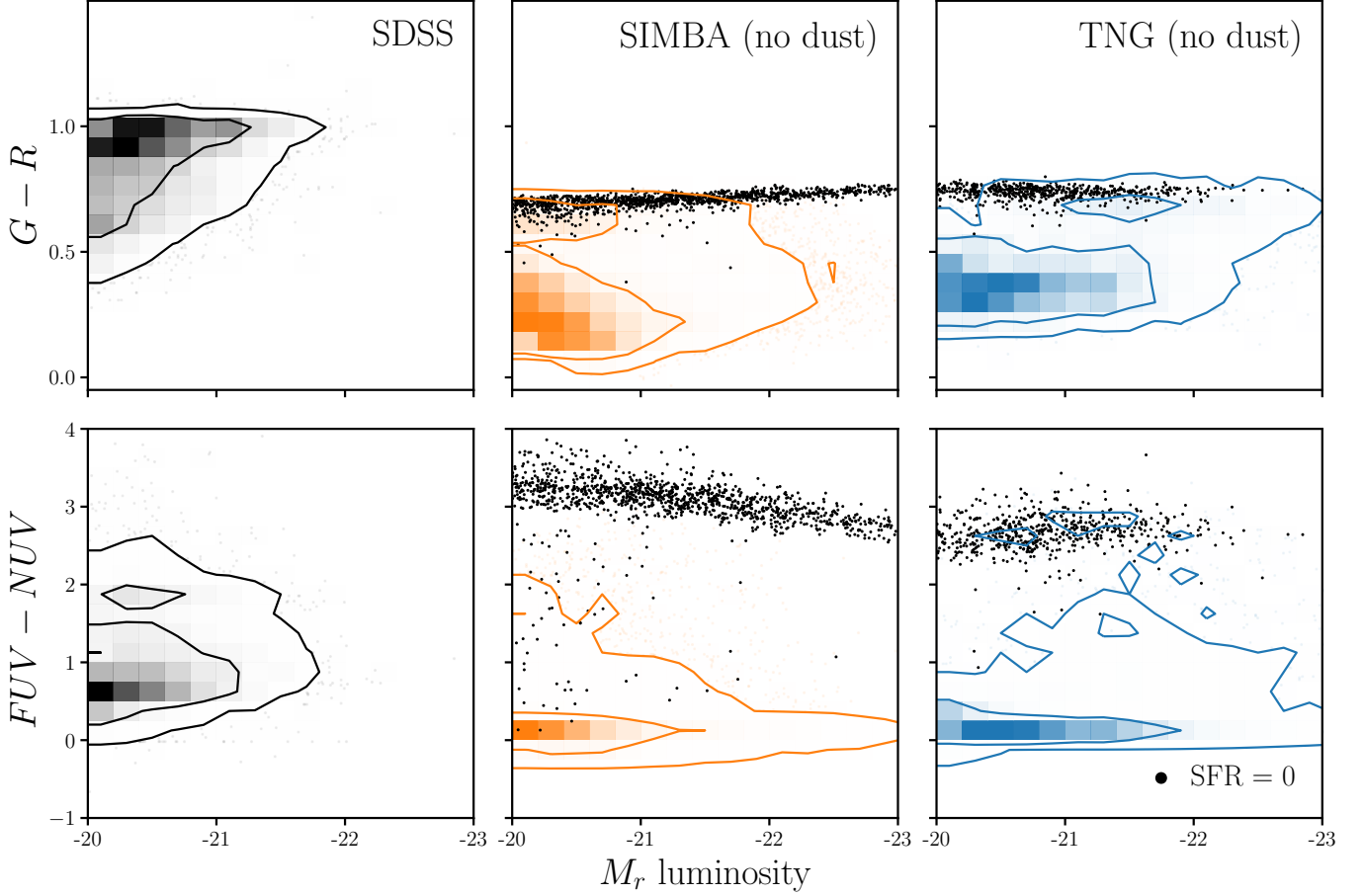
$$A_V = -2.5 \log \left[ \frac{1 - e^{-\tau_V \sec i}}{\tau_V \sec i} \right] \quad (3)$$

justification of why this is enough. We sample  $i$  uniformly.

TODO

Recently, Salim & Narayanan (2020) find significant dependence in  $A_V$  on both  $M_*$  and SFR. We include this dependence through  $\tau_V$ , which we flexibly parameterize as

$$\tau_V(M_*, \text{SFR}) = m_{\tau,1} \log \left( \frac{M_*}{10^{10} M_\odot} \right) + m_{\tau,2} \log \text{SFR} + c_\tau. \quad (4)$$



**Figure 2.** We present the distributions of the main observables used throughout the paper for SDSS (left), SIMBA (center), and TNG (right) centrals. We do *not* include any prescription for dust for the simulated galaxies. The top panels present the  $G - R$  versus  $M_r$  color magnitude relations while the bottom panels present the  $FUV - NUV$  versus  $M_r$  relations. The observables for the SIMBA and TNG simulated galaxies are derived using forward modeling and are therefore consistent with SDSS measurements (Section 2.5). The contours for SIMBA and TNG do not include galaxies with  $SFR = 0$ , which we mark separately in black. Despite similar SMFs, *the simulations without any dust prescription show stark differences with observations in the color-magnitude observable-space.*

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

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

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  $\lambda_V$  is the  $V$  band wavelength.  $\delta$ , the slope of the attenuation curve, also correlates with galaxy properties. So we parameterize  $\delta$  and

$$\delta(M_*, \text{SFR}) = m_{\delta,2} \log \left( \frac{M_*}{10^{10} M_\odot} \right) + m_{\delta,2} \log \text{SFR} + c_\delta \quad (6)$$

$D(\lambda)$  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. In our DEM, we assume fixed  $\lambda_0 = 2175 \text{\AA}$  and  $\Delta \lambda = 350 \text{\AA}$ .

Kriek & Conroy (2013) and Tress et al. (2018) found evidence that  $E_b$  correlates with the slope of the attenuation curve for star-forming galaxies  $z \sim 2$ . This dependence was confirmed with simulations in ?.  $E_b$ :

$$E_b = m_E \delta + c_E \quad (8)$$

We also split the attenuation on the star light and nebular emission

$$F_o(\lambda) = F_i^{\text{star}}(\lambda) 10^{-0.4A(\lambda)} + F_i^{\text{neb}}(\lambda) 10^{-0.4A_{\text{neb}}(\lambda)} \quad (9)$$

where we parameterize

$$A_{\text{neb}}(\lambda) = f_{\text{neb}} A(\lambda) \quad (10)$$

mention of how we treat  $\text{SFR} = 0$  galaxies

TODO

### 3.2. 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 a strong simplifying assumption that the dust in galaxies details. and that the attenuation depends on the orientation of the galaxy sentence or two about how we know the slab model doesn't work.

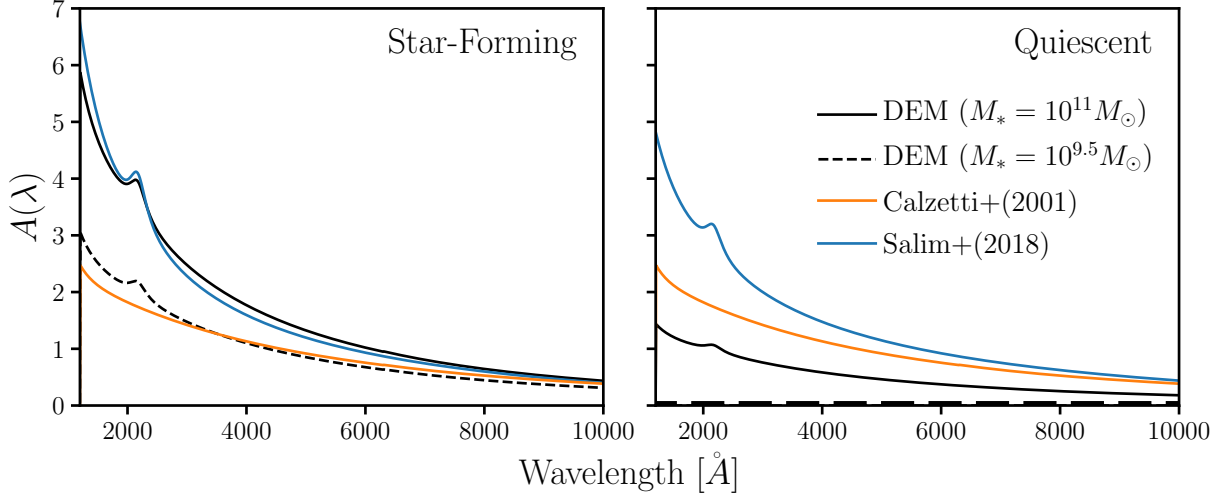
TODO  
TODO

Furthermore, the slab model predicts  $A_V$  distribution significantly different than  $A_V$  distributions measured from observations. In Figure ??, we compare the  $A_V$  distribution predicted by the slab model for  $\tau_V =$  to the  $A_V$  distribution of SDSS galaxies from ?. The comparison clearly reveals the discrepancy between the SDSS  $A_V$  distribution and the slab model prediction. The slab model predicts a sharp cutoff at lower  $A_V$  end of the distribution, which is not found in observables. The SDSS  $A_V$  measurements can be negative in ? because ....

TODO

To ensure that our results do not depend significantly on the slab model, we implement a more flexible alternative model for sampling  $A_V$  based on a truncated normal distribution:

$$A_V \sim \mathcal{N}_T(\mu_{A_V}, \sigma_{A_V}) = \frac{\mathcal{N}(\mu_{A_V}, \sigma_{A_V})}{1 - \Phi\left(-\frac{\mu_{A_V}}{\sigma_{A_V}}\right)}. \quad (11)$$



**Figure 3.** Comparison of the attenuation curve for our fiducial dust empirical model (DEM; black) to common attenuation curves in the literature (Calzetti 2001, orange; Salim et al. 2018, blue). We compare the curves for typical star-forming galaxies with  $\text{SFR} = 10^{0.5} M_{\odot}/\text{yr}$  in the left panel and for quiescent galaxies with  $\text{SFR} = 10^{-2} M_{\odot}/\text{yr}$  in the right. For our fiducial DEM, we use parameter values at **the center of the priors listed in Table 1** and include attenuation curves for  $M_{*} = 10^{9.5}$  (dashed) and  $10^{11} M_{\odot}$  (solid). Our fiducial DEM is flexibly parameterized to incorporate both  $M_{*}$  and SFR dependence in the attenuation curve (Section 3.1).

$\mathcal{N}$  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.

There’s evidence that  $\mu_{A_V}$  and  $\sigma_{A_V}$  depend on  $M_{*}$  and SFR, so we parameterize them in a similar fashion as  $\tau_V$  in Eq. ??:

$$\mu_{A_V} = m_{\mu,1}(\log M_{*} - 10.) + m_{\mu,2} \log \text{SFR} + c_{\mu} \quad (12)$$

$$\sigma_{A_V} = m_{\sigma,1}(\log M_{*} - 10.) + m_{\sigma,2} \log \text{SFR} + c_{\sigma} \quad (13)$$

$$(14)$$

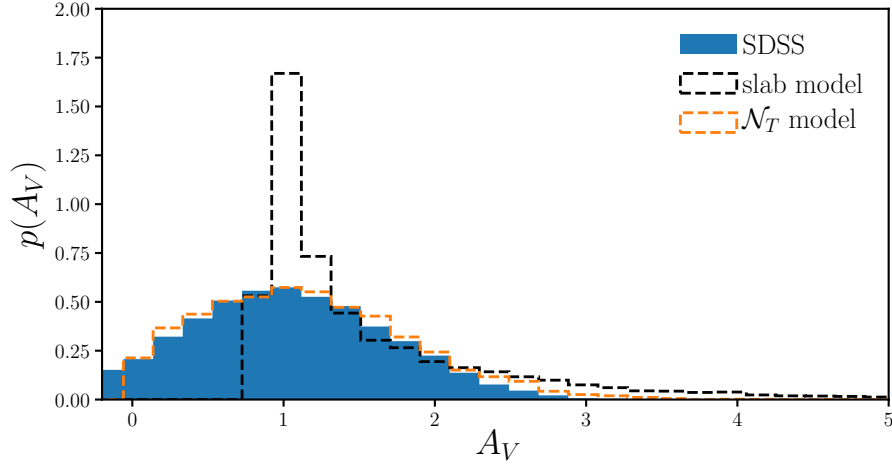
### 3.3. Likelihood-Free Inference

Approximate Bayesian Computation with Population Monte Carlo Hahn et al. (2017),

To compare the outputs of our DEMs to observations, we first measure the color-magnitude observables ( $G - R$ ,  $FUV - NUV$ , and  $M_r$ ) as described in Section 2.5 consistent with SDSS measurements. Afterwards, we compare the forward modeled observables to SDSS using a L2 norm distance metric:

$$\bar{\rho}(\theta) = \sum_{i=1}^n [X_i^{\text{SDSS}} - X_i^{\text{model}}(\theta)]^2. \quad (15)$$

$X^{\text{SDSS}}$  and  $X^{\text{model}}(\theta)$  are  $n$ -dimensional data vectors of the SDSS and model observables. In our case, we use a 3-dimensional histogram along  $G - R$ ,  $FUV - NUV$ ,



**Figure 4.** 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 3.2).

[Ishida et al. \(2015\)](#)

Paragraph on the priors we choose.

#### 4. RESULTS

- There isn’t a whole lot of flexibility for  $\text{SFR}=0$  galaxies predicted by simulations and they do not agree well with observations [A](#).

#### 5. 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

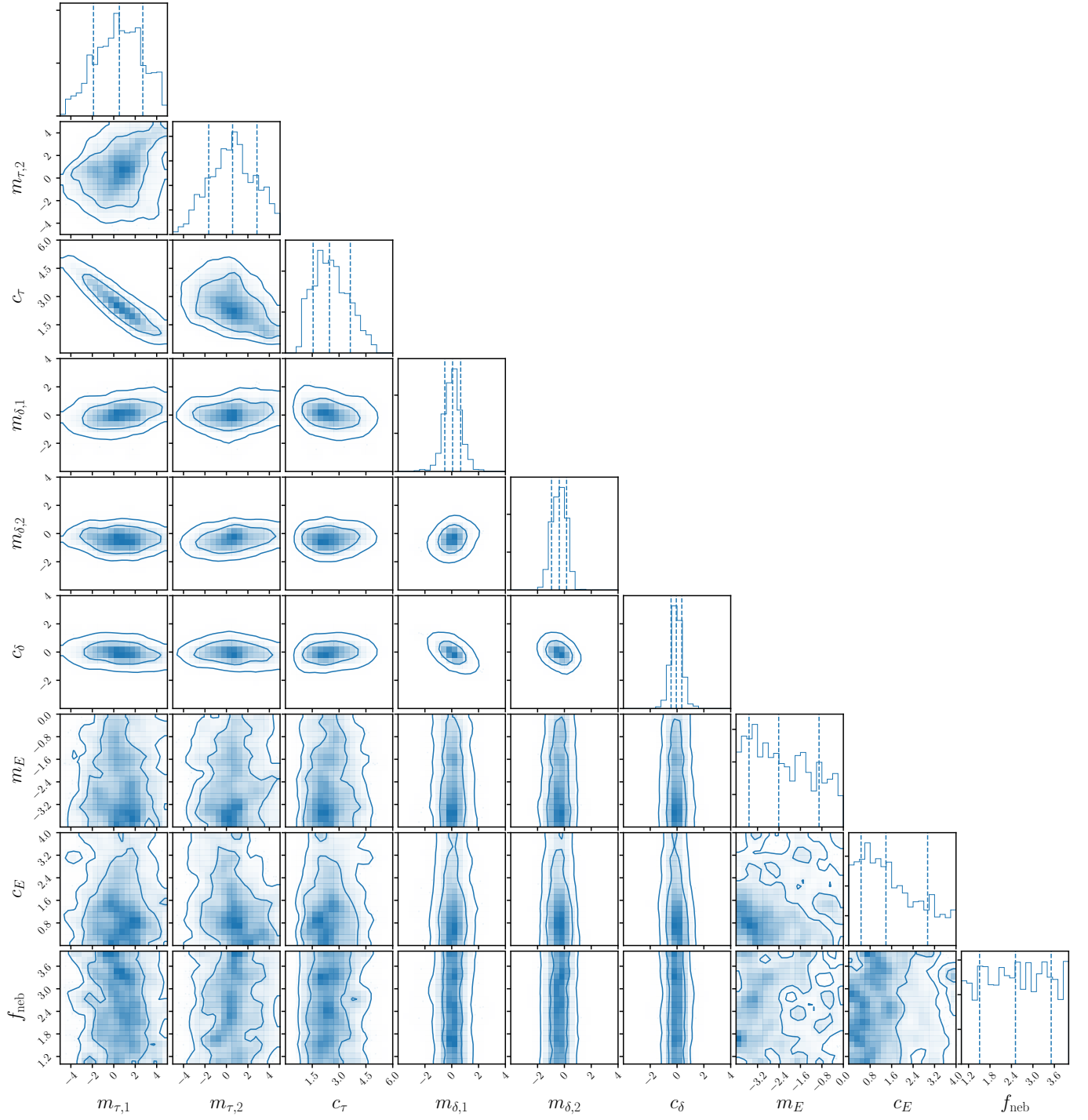
##### A. RESOLUTION EFFECTS

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

#### REFERENCES

Abazajian K. N., et al., 2009, [The Astrophysical Journal Supplement Series](#), 182, 543

Blanton M. R., Roweis S., 2007, [The Astronomical Journal](#), 133, 734

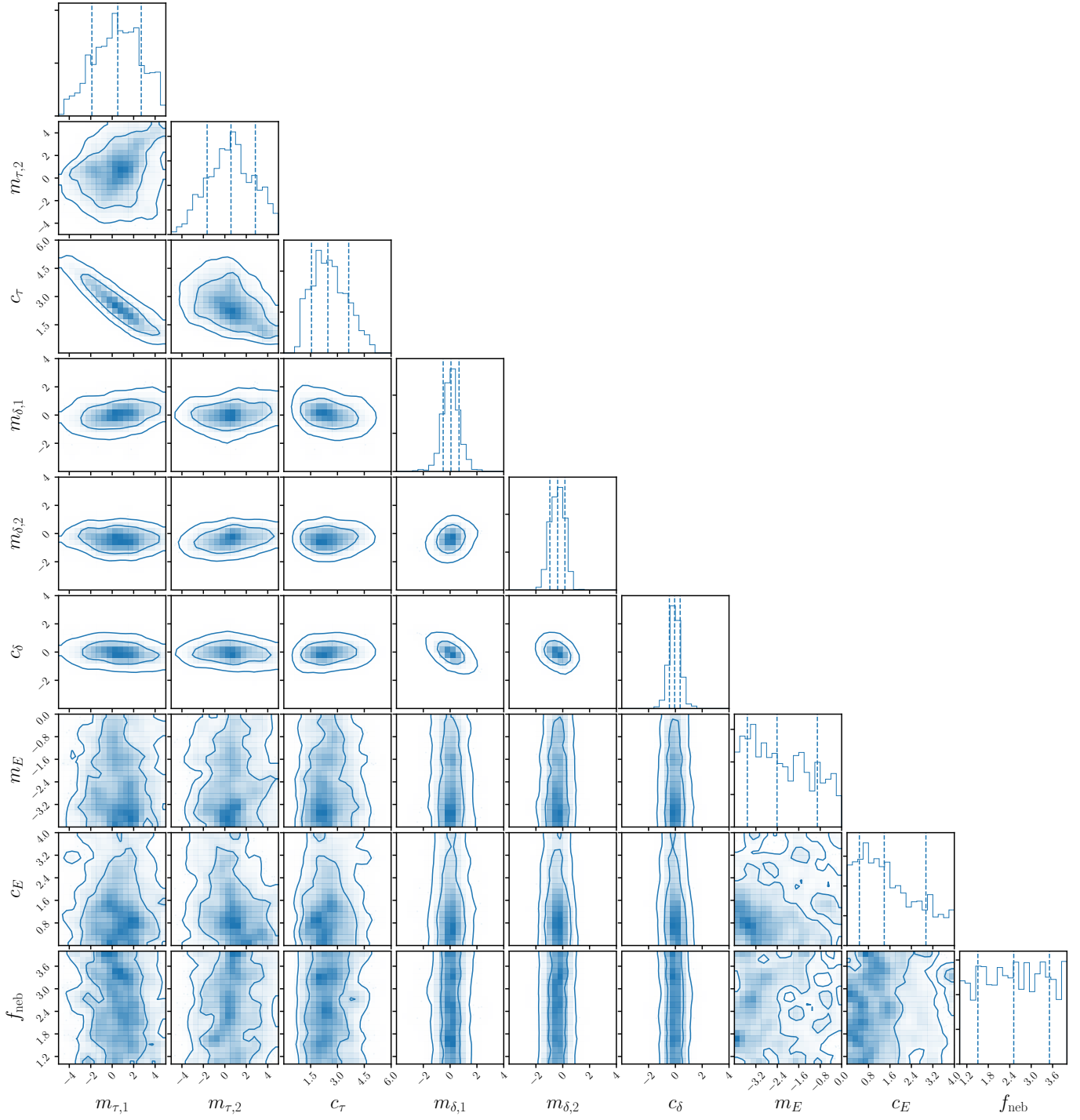


**Figure 5.** Posterior distribution of our fiducial DEM parameters derived using Approximate Bayesian Computation (ABC). (Section 3.3)

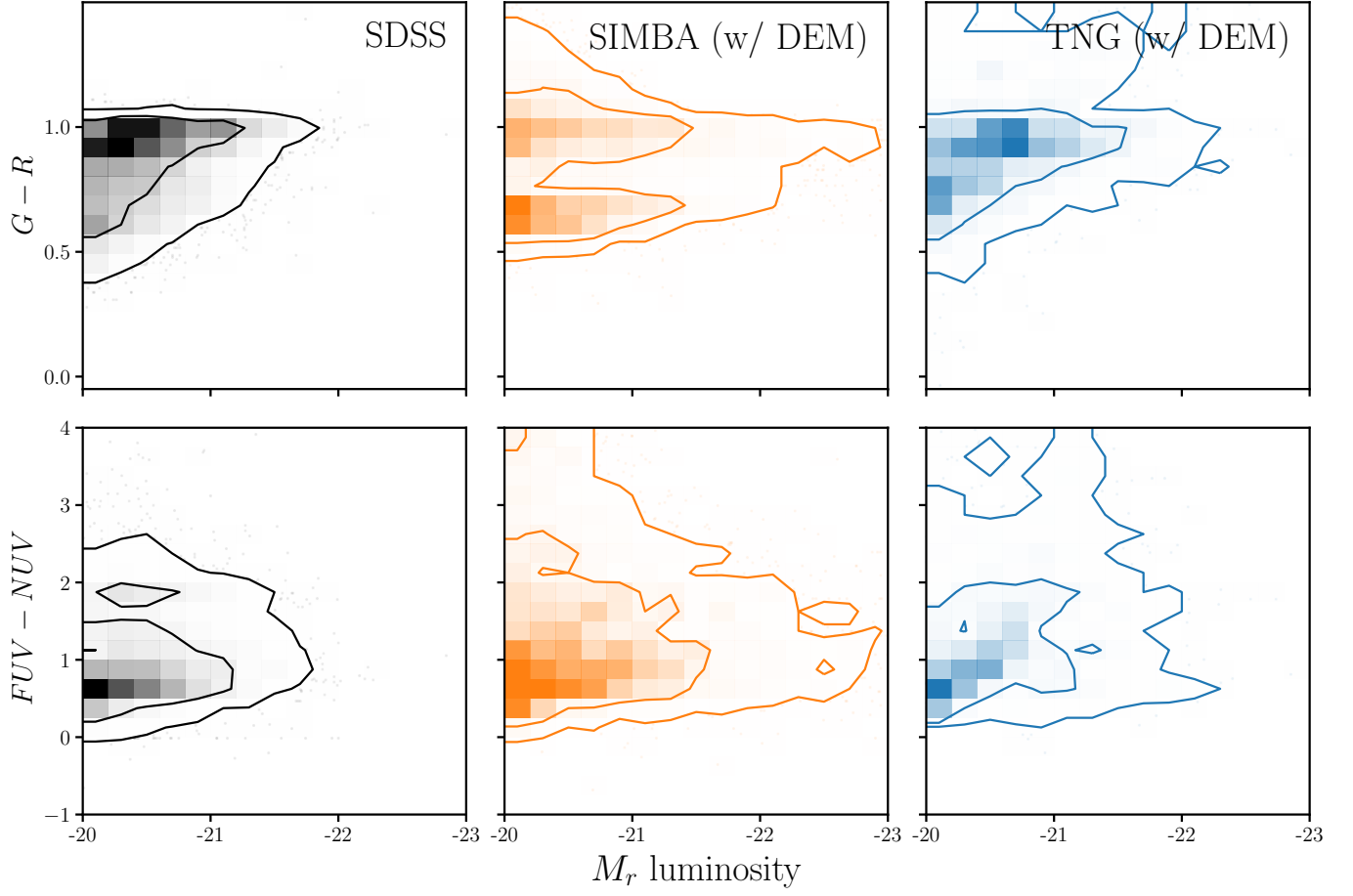


**Table 1.** Parameters of the Dust Empirical Models

Parameter	Definition	prior
DEM with slab model		
$m_{\tau,1}$	Slope of the $\log M_*$ dependence of optical depth, $\tau_V$	
$m_{\tau,2}$	Slope of the $\log \text{SFR}$ dependence of optical depth, $\tau_V$	
$c_{\tau}$	amplitude of the optical depth, $\tau_V$	
DEM with $\mathcal{N}_T$ model		
$m_{\mu,1}$	Slope of the $\log M_*$ dependence of optical depth, $\tau_V$	
$m_{\mu,2}$	Slope of the $\log \text{SFR}$ dependence of optical depth, $\tau_V$	
$c_{\mu}$	amplitude of the optical depth, $\tau_V$	
$m_{\sigma,1}$	Slope of the $\log M_*$ dependence of optical depth, $\tau_V$	
$m_{\sigma,2}$	Slope of the $\log \text{SFR}$ dependence of optical depth, $\tau_V$	
$c_{\sigma}$	amplitude of the optical depth, $\tau_V$	
$m_{\delta,1}$	Slope of the $\log M_*$ dependence of the attenuation curve slope, $\delta$	
$m_{\delta,2}$	Slope of the $\log \text{SFR}$ dependence of the attenuation curve slope, $\delta$	
$c_{\delta}$	amplitude of the attenuation curve slope, $\delta$	
$m_E$	slope of the $\delta$ dependence of UV dust bump strength, $E_b$	
$c_E$	amplitude of UV dust bump strength, $\delta$	
$f_{\text{neb}}$	fraction of nebular attenuation curve	



**Figure 6.** Posterior distribution of our fiducial DEM parameters derived using Approximate Bayesian Computation (ABC). (Section 3.3)



**Figure 7.** Comparison of the observables predicted by the simulations with the posterior DEM.

- Blanton M. R., et al., 2005, [The Astronomical Journal](#), 129, 2562
- Calzetti D., 2001, [New Astronomy Reviews](#), 45, 601
- Chabrier G., 2003, [Publications of the Astronomical Society of the Pacific](#), 115, 763
- Hahn C., Vakili M., Walsh K., Hearin A. P., Hogg D. W., Campbell D., 2017, [Monthly Notices of the Royal Astronomical Society](#), 469, 2791
- Ishida E. E. O., et al., 2015, [Astronomy and Computing](#), 13, 1
- Kriek M., Conroy C., 2013, [The Astrophysical Journal Letters](#), 775, L16
- Narayanan D., Conroy C., Davé R., Johnson B. D., Popping G., 2018, [The Astrophysical Journal](#), 869, 70
- Noll S., Burgarella D., Giovannoli E., Buat V., Marcillac D., Muñoz-Mateos J. C., 2009, [Astronomy and Astrophysics](#), 507, 1793
- Salim S., Narayanan D., 2020, arXiv:2001.03181 [astro-ph]
- Salim S., Boquien M., Lee J. C., 2018, [The Astrophysical Journal](#), 859, 11
- 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
- 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., 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