



How Well Can We Measure the Stellar Mass of a Galaxy: The Impact of the Assumed Star Formation History Model in SED Fitting

Sidney Lower¹ , Desika Narayanan^{1,2,3} , Joel Leja^{4,8} , Benjamin D. Johnson⁴ , Charlie Conroy⁴ , and Romeel Davé^{5,6,7}

¹ Department of Astronomy, University of Florida, 211 Bryant Space Science Center, Gainesville, FL 32611, USA

² University of Florida Informatics Institute, 432 Newell Drive, CISE Bldg. E251 Gainesville, FL 32611, USA

³ Cosmic Dawn Centre at the Niels Bohr Institute, University of Copenhagen and DTU-Space, Technical University of Denmark, Denmark

⁴ Center for Astrophysics | Harvard & Smithsonian, 60 Garden St., Cambridge, MA 02138, USA

⁵ Institute for Astronomy, Royal Observatory, University of Edinburgh, Edinburgh, EH9 3HJ, UK

⁶ University of the Western Cape, Bellville, Cape Town 7535, South Africa

⁷ South African Astronomical Observatory, Observatory, Cape Town 7925, South Africa

Received 2020 June 5; revised 2020 September 23; accepted 2020 October 7; published 2020 November 19

Abstract

The primary method for inferring the stellar mass (M_*) of a galaxy is through spectral energy distribution (SED) modeling. However, the technique rests on assumptions such as the galaxy star formation history (SFH) and dust attenuation law that can severely impact the accuracy of derived physical properties from SED modeling. Here we examine the effect that the assumed SFH has on the stellar properties inferred from SED fitting by ground-truthing them against mock observations of high-resolution cosmological hydrodynamic galaxy formation simulations. Classically, SFHs are modeled with simplified parameterized functional forms, but these forms are unlikely to capture the true diversity of galaxy SFHs and may impose systematic biases with underreported uncertainties on results. We demonstrate that flexible nonparametric SFHs outperform traditional parametric forms in capturing variations in galaxy SFHs and, as a result, lead to significantly improved stellar masses in SED fitting. We find a decrease in the average bias of 0.4 dex with a delayed- τ model to a bias under 0.1 dex for the nonparametric model, though this is heavily dependent on the choice of prior for the nonparametric model. Similarly, using nonparametric SFHs in SED fitting results in increased accuracy in recovered galaxy star formation rates and stellar ages.

Unified Astronomy Thesaurus concepts: [Astronomy data modeling \(1859\)](#); [Hydrodynamical simulations \(767\)](#); [Stellar masses \(1614\)](#); [Galaxy properties \(615\)](#); [Spectral energy distribution \(2129\)](#); [Radiative transfer simulations \(1967\)](#)

1. Introduction

The ability to accurately infer the physical properties of galaxies is critical for our understanding of galaxy formation and evolution. Modeling the ultraviolet (UV) to infrared (IR) spectral energy distributions (SEDs) of galaxies is one of the main methodologies used to derive the physical properties of galaxies, such as the stellar mass (M_*), star formation rate (SFR), and stellar age. These techniques, pioneered by Tinsley (1968), Spinrad & Taylor (1971), and Faber (1972), have seen an explosion of interest and activity as space-based missions such as Galaxy Evolution Explorer (GALEX) and the Hubble Space Telescope (HST) have opened up UV and optical wavelengths for galaxies near and far, respectively. Similarly, advances in IR and submillimeter detector technology have opened up IR windows that provide constraints for SED models that consider energy balance between UV/optical photons and thermal IR emission from dust.

The abundance of panchromatic data available has spurred the development of many SED modeling codes (e.g., CIGALE, Boquien et al. 2019; FAST, Kriek et al. 2009, 2018; MAGPHYS, da Cunha et al. 2008) that were developed to estimate physical properties from observed broadband data. These codes rely on models describing the stellar populations, nebular emission, and dust content in the galaxy, along with an optimization method to fit the SED and return the resulting physical parameters.

The basic components in an SED model include information about stellar populations—the stellar initial mass function (IMF), stellar isochrones and spectral templates, and star formation history (SFH)—along with emission from nebular regions and active galactic nuclei (AGNs), dust emission, and attenuation from dust. The robustness of an SED model and our ability to accurately recover physical properties of a galaxy depend on our confidence in each model component to accurately capture the complexity of the many physical processes that occur in a galaxy (see, e.g., Conroy 2013, for an in-depth review).

Despite the widespread use of SED modeling by the observational galaxy community, it remains difficult to establish the efficacy of the technique owing to the many weakly constrained components and relative lack of ground truth (see Mobasher et al. 2015; see also reviews by Conroy 2013; Walcher et al. 2011). Indeed, some efforts have emerged in recent years to provide such a ground truth in the context of extensive comparisons between SED modeling codes (e.g., Hunt et al. 2019) or comparisons between input mock SEDs drawn from known physical properties and the output properties from SED modeling. Examples of the latter context range from testing on a library of SEDs from an empirical mock catalog (Mobasher et al. 2015; Leja et al. 2019a) to SED modeling of idealized galaxies (Hayward & Smith 2015) and galaxies from a cosmological simulation (Iyer & Gawiser 2017; Katsianis et al. 2020).

One of the more influential yet poorly constrained components of SED fitting is the assumed form of the SFH

⁸ NSF Astronomy and Astrophysics Postdoctoral Fellow.

(e.g., Ocvirk et al. 2006; Iyer & Gawiser 2017; Carnall et al. 2019). The most common models for SFHs are parameterized by a simple functional form, and the parameters varied in the SED fit describe that functional form. Hereafter we refer to these as “parametric” SFHs. Examples include the τ and delayed- τ models, which model the SFH as exponentially declining with some characteristic e -folding time. Although these models describe the SFHs of galaxies in a closed box (i.e., isolated with no inflow of pristine gas) where gas forms stars with constant star formation efficiency, the restricted nature of the functional forms does not match the diversity of true galaxy SFHs (Gallagher et al. 1984; Sandage 1986; Lee et al. 2009; Oemler et al. 2013; Simha et al. 2014; Diemer et al. 2017). Carnall et al. (2019) have shown that the cosmic SFR densities (CSFRDs) inferred from delayed- τ SFH fits to galaxies from the GAMA survey (Driver et al. 2011) are incompatible with the CSFRDs predicted by the Universe-Machine (Behroozi et al. 2019). Leja et al. (2019b) found that the backward-evolved stellar mass functions (SMFs) inferred from SED fits to galaxies from the 3D-HST survey using a delayed- τ SFH model are in tension with observed SMFs at $z = 3$. The assumed SFH and associated priors can also strongly bias the inferred physical properties of galaxies (Simha et al. 2014; Acquaviva et al. 2015; Salmon et al. 2015; Ciesla et al. 2017; Iyer & Gawiser 2017; Carnall et al. 2019; Curtis-Lake et al. 2020). For instance, Michałowski et al. (2012) found that the assumed SFH model had the largest impact out of all other SED model components on the stellar masses inferred for observations of submillimeter galaxies (SMGs). Similarly, Dudzevičiūtė et al. (2020) found that the average difference between the stellar masses predicted by the MAGPHYS SED model and the true stellar masses for galaxies from the EAGLE cosmological simulation was close to 0.5 dex. This bias, attributed to the assumed SFH module, is consistent with the earlier results of Michałowski et al. (2012).

Parametric SFHs with more flexibility have also been explored in the literature (e.g., Papovich et al. 2011; Simha et al. 2014; Ciesla et al. 2017; Carnall et al. 2018). One example is the lognormal parameterization, which has been shown to reproduce the evolution of the cosmic SFR history (Gladders et al. 2013) and provide a reasonable match to the Illustris galaxy SFHs (Diemer et al. 2017). However, the lognormal parameterization still suffers from stellar age biases similar to the simpler parametric forms (Ciesla et al. 2017; Iyer & Gawiser 2017; Carnall et al. 2019; Leja et al. 2019a). This happens because the SFH is constrained to either have recent star formation or have a population of old stars (Leja et al. 2019a)—a consequence of the inflexible mathematical form. Moreover, Diemer et al. (2017) fit a lognormal SFH directly to the SFHs of the Illustris galaxies instead of via SED modeling, where all galaxy properties must be inferred simultaneously; thus, biases such as outshining (Papovich et al. 2001) did not affect the fit.

An alternative to the parametric models described above are nonparametric forms. Nonparametric SFH models, defined as models that do not explicitly assume a functional form, have been shown to have the flexibility to reproduce the variation in SFHs seen in observations and galaxy formation simulations. Examples include models sampled from a diverse basis of analytical SFH models (Iyer & Gawiser 2017), flexible piecewise linear functions in time (Reichardt et al. 2001;

Heavens et al. 2004; Tojeiro et al. 2007; Kelson 2014; Leja et al. 2017, 2019a; Morishita et al. 2019), models drawn from simulations or semianalytical models (SAMs; Brammer et al. 2008; Pacifici et al. 2012, 2016; Zhang et al. 2017), and models utilizing machine-learning methods (Iyer et al. 2019; Lovell et al. 2019). Most early implementations of nonparametric SED fitting methods relied on spectroscopic data, but models like those presented in Iyer & Gawiser (2017) and Leja et al. (2017) have been shown to produce reasonable results with broadband data only.

While ground-truth tests of parametric SFHs have generally found poor agreement with both real and simulated SFHs, the opposite is typically true for nonparametric models. Focusing on results from broadband SEDs, Iyer et al. (2019) validated the reconstruction of SFHs using Gaussian processes with a sample of galaxies from the Santa Cruz SAMs and the MUFASA hydrodynamical cosmological simulation and found good agreement between the mock galaxy SFHs and the SFHs reconstructed from Gaussian processes. Leja et al. (2019b) demonstrated that galaxies selected from the 3D-HST photometry catalog (Skelton et al. 2014) were inferred to be systematically more massive and older when modeled with the nonparametric SFHs in PROSPECTOR compared to previously published results using parametric SFHs. However, extensive tests and ground-truthing of the results from SED fitting have lacked in ground-truth sample technique (as in the case of, e.g., Leja et al. 2019a, where the mock galaxy SFHs used to generate synthetic SEDs were relatively simple), sample size, or a combination of both (as in the case of, e.g., Hayward & Smith 2015, where just two scenarios of simplified idealized simulated galaxies were used).

Thus, in this paper we advance our understanding of the efficacy of SED modeling by fitting nonparametric SFH models to mock SEDs generated from 3D radiative transfer on a large galaxy sample from a hydrodynamical cosmological simulation. First, we focus on the impact of the assumed SFH on the stellar masses and other galaxy properties inferred from SED fitting by isolating the SED fits from dust. Specifically, we fix the dust-to-stellar geometries in our radiative transfer calculations and match that geometry in our SED fits to effectively obtain the baseline uncertainties on the inferred galaxy properties that can be attributed to just the assumed form of the SFH model. This is distinguished from previous observationally based studies, where all components of the SED model are necessarily tested at once. In other words, we ignore the uncertainty from other model components to first understand the biases and uncertainties imposed by just the SFH model. While the main goal of this paper is to isolate the impact of galaxy SFHs on SED fitting techniques, specifically to ground-truth the efficacy of various SFH models, we then briefly generalize these results by rerunning our dust radiative transfer using the intrinsic dust-to-stellar geometries and dust properties from the simulations.

The paper is organized as follows: In Section 2 we describe our numerical methods (including cosmological simulations, radiative transfer, and SED fitting). In Section 3 we describe the results of the SED fitting through comparisons to the true values from the simulations. In Section 4 we discuss the results, the possible origins of fitting failures, and the inclusion of realistic dust in the mock SEDs. In Section 5 we conclude and propose a pathway toward improving our dust models to better accommodate realistic galaxies in SED fitting.

2. Numerical Methods

2.1. Overview

For this analysis, we fit the SEDs of galaxies from a cosmological hydrodynamical simulation to determine the robustness of stellar masses estimated from SED modeling. The simulated galaxy SEDs are generated with post-processing radiative transfer that propagates the intrinsic stellar SEDs through dust in the interstellar medium (ISM).

We fit these mock SEDs using the Bayesian inference code PROSPECTOR. We “observe” the mock SEDs in the same broadband filters for all galaxies and use the same models for stellar metallicity and dust attenuation when fitting, so that the only difference between the results shown below originates from differences in the assumed SFH model in the SED fit. In the remainder of this section, we describe these methods in more detail.

2.2. The SIMBA Galaxy Formation Model

We first need a population of model galaxies to generate the SEDs from. For this, we use the SIMBA cosmological simulation, described in full in Davé et al. (2019). Briefly, SIMBA is the descendant of the simulation suite MUFASA and relies on GIZMO’s meshless finite mass hydrodynamics. New subresolution prescriptions for stellar and AGN feedback, as well as black hole growth, have enabled SIMBA to accurately reproduce observables like the galaxy SMF and the star-forming main sequence. SIMBA additionally includes an on-the-fly self-consistent model for the formation, growth, and destruction of dust that reproduces both the $z = 0$ dust mass function and the scaling between the dust-to-gas ratio and metallicity (Li et al. 2019).

We employ a box with $25/h$ Mpc side length with 512³ particles, resulting in a baryon mass resolution of $1.4 \times 10^6 M_{\odot}$. To identify galaxies, we have employed a modified version of CAESAR⁹ (Thompson 2014). We focus on the $z = 0$ snapshot, in which there are ~ 1600 galaxies identified with a minimum of 32 bound star particles with a 6D friends-of-friends galaxy finder. The choice of 32 star particles is motivated by the mass resolution of this simulation. These galaxies lie within a mass range of 4.4×10^7 – $1.4 \times 10^{12} M_{\odot}$.

2.3. Dust Radiative Transfer

We use the 3D radiative transfer code POWDERDAY¹⁰ (Narayanan et al. 2020) to construct the synthetic SEDs by first generating with FSPS (Conroy et al. 2009; Conroy & Gunn 2010) the dust-free SEDs for the star particles within each cell using the stellar ages and metallicities as returned from the cosmological simulations. For these, we assume a Kroupa (2002) stellar IMF and the MIST stellar isochrones (Paxton et al. 2011; Choi et al. 2016; Dotter 2016).

Traditionally, POWDERDAY then propagates the emission from these stars through the diffuse dusty ISM using HYPERION as the dust radiative transfer solver (Robitaille 2011; Robitaille et al. 2012). However, this then imposes the uncertainty of the diverse attenuation laws that vary from galaxy to galaxy on our SED fits (Narayanan et al. 2018a, 2018b; Salim & Narayanan 2020). We therefore abandon the diffuse dust in our POWDERDAY radiative

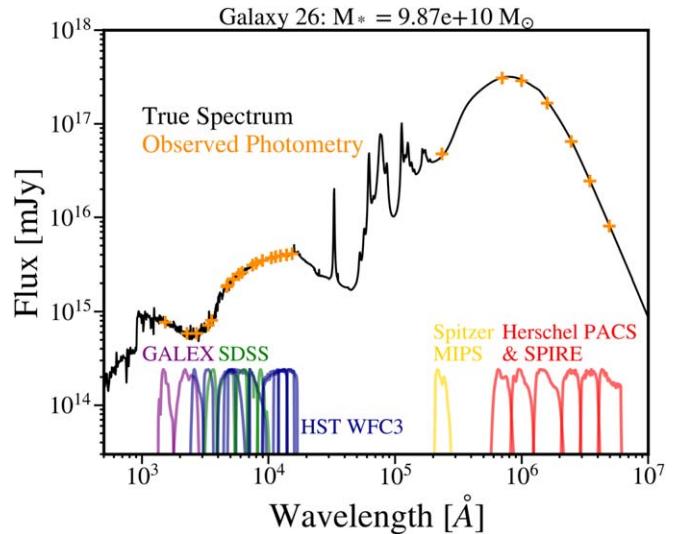


Figure 1. Example POWDERDAY mock SED with “observed” photometric bands highlighted, spanning from GALEX far-UV to Herschel SPIRE, totaling 25 bands, resulting in almost complete coverage across all wavelength regimes. All galaxies are observed with the same filters, and photometric errors are fixed at 3%. Photometry is “observed” for each filter and convolved over the filter bandwidth. We ignore near-IR photometry, spanning from ~ 2 to $20 \mu\text{m}$, to avoid dependence on the ultrasmall grain size fractions chosen for the POWDERDAY calculations.

transfer simulations and instead employ a dust screen surrounding all stars. This is akin to how PROSPECTOR treats dust obscuration and therefore allows us to isolate the impact of the galaxy SFH on our SED fits. In the dust screen setup for POWDERDAY, we assume a uniform dust screen around all stars with an optical depth of $\tau_{\text{uniform}} = 0.7$. Younger stars (< 10 Myr old) have an additional assumed source of extinction from their birth clouds that have an optical depth of $\tau_{\text{BC}} = 0.7$. This fiducial dust screen model ensures an apples-to-apples comparison between the creation of the SEDs and the technique used to fit them. We then generalize this comparison in Section 4.4, where we rerun our dust radiative transfer with POWDERDAY using the intrinsic dust-to-stellar geometries and dust properties from the simulations.

The result of the POWDERDAY radiative transfer is the UV–far-IR spectrum for each galaxy. We extract model photometry from these dust spectra, selecting 25 bands from the GALEX far-UV filter at 1542 Å through the Herschel SPIRE band at 500 μm as shown in Figure 1 and Table 1. The SED coverage is comparable to galaxies in, e.g., the CANDELS GOOD-S field (Guo et al. 2013). And while SED coverage necessarily impacts the accuracy of galaxy properties inferred from SED fitting, we leave an in-depth investigation into these factors to a future work. Further, we fixed uncertainties to 3% of the flux value, since the aim of this study is to analyze not the effect of photometric uncertainties but rather the systematics that arise from the use of various SFH models.

2.4. SED Fitting

In order to model the mock SEDs generated by POWDERDAY, we use the Bayesian inference code PROSPECTOR¹¹ (Leja et al. 2017, 2019a). PROSPECTOR derives galaxy physical properties using stellar population synthesis evolved with dust

⁹ <https://github.com/dnarayanan/caesar>

¹⁰ <https://github.com/dnarayanan/powderday>

¹¹ <https://github.com/bd-j/prospector>

Table 1
Table of the 25 Filters Used to Extract Photometry from the Synthetic POWDERDAY Spectra

Instrument	Filter	Effective Wavelength (Å)
GALEX	Far-UV	1549
	Near-UV	2304
HST/WFC3	F275W	2720
	F336W	3359
	F475W	4732
	F555W	5234
	F606W	5780
	F814W	7977
	F105W	10431
	F110W	11203
	F125W	12364
	F140W	13735
SDSS	F160W	15279
	<i>u</i>	3594
	<i>g</i>	4640
	<i>r</i>	6122
	<i>i</i>	7439
Spitzer/MIPS	<i>z</i>	8897
	24 μ m	232096
Herschel/PACS	Blue	689247
	Green	979036
	Red	1539451
	PSW	2428393
Herschel/SPIRE	PMW	3408992
	PLW	4822635

Note. Broadband fluxes are assigned a 3% fractional uncertainty.

within the framework of FSPS. POWDERDAY and PROSPECTOR rely on the same spectral libraries, IMF, and stellar isochrones within FSPS. And though the assumptions made when modeling stellar spectra, especially concerning the impact of late stellar evolutionary stages on a galaxy’s SED, are important and the physical parameters derived from SED can be greatly influenced by these assumptions (e.g., Santini et al. 2015; Mobasher et al. 2015), this exploration is ultimately outside the scope of this paper, as we are focusing on the targeted question of understanding the impact of the assumed SFH on derived stellar masses. Additionally, the use of the fiducial dust screen model enables the subsequent SED fitting procedures and results to be isolated from assumptions about the dust attenuation model. In other words, the models described below will vary owing to assumptions about the SFH only, effectively isolating our results to the differences between these models and providing a baseline uncertainty estimate arising from just the SFH model.

PROSPECTOR uses a Bayesian inference framework via DYNESTY nested sampling (Speagle 2020) to fit the observed SEDs and provide posterior distribution functions (PDFs) for physical parameters such as stellar mass, stellar metallicity, and SFR. The power of DYNESTY lies in its ability to efficiently sample multimodel distributions and have well-defined stopping criteria based on evaluations of Bayesian evidence ensuring model convergence. Below we describe the SED model components and their prior distributions, which are summarized in Table 2. Because DYNESTY is based on Bayesian inference, all results in our analysis are sampled from the resulting PDF from the nested sampling iterations (as opposed to so-called “best-fit” parameters quoted by a χ^2

minimization algorithm), with uncertainties quoted as the 16th through 84th percentiles of the posterior distributions.

2.4.1. Star Formation History

PROSPECTOR includes several models for a galaxy’s SFH, including the commonly used delayed- τ model, along with flexible nonparametric models. The nonparametric models are constrained by priors that either can enforce continuity (i.e., the SFH is smooth rather than bursty) or allow for episodes of star formation bursts. We show the prior probability distribution for one such model in Figure 2. As a means for comparison, we considered three simple parametric SFHs, an example of which is also shown as a prior probability distribution in the left panel of Figure 2. All SFH models used in this analysis are described in depth in Carnall et al. (2019) and Leja et al. (2019a), while we provide a more top-level view here. Three commonly used parametric models are considered (delayed- τ , delayed- τ with a burst component, and a constant SFR across time), as well as a nonparametric SFH model.

Delayed- τ : the delayed- τ model is an exponentially declining SFH, parameterized by the e -folding time that is informed by a log-uniform prior. The free parameters for this model also include the stellar mass formed by the galaxy and the maximum age of the stellar population.

Delayed- τ + burst: this is the same as above but including a random burst of star formation. During a burst, some fraction of mass is formed in an instantaneous burst of star formation.

Constant: the constant SFH model is set to a uniform value for all times. These are often used for modeling star-forming galaxies at high redshift.

Nonparametric: the nonparametric SFHs as implemented in PROSPECTOR fit for the fractional stellar mass formed in a particular time bin, independent of the total mass formed (i.e., the shape of the SFH is separate from the normalization). For this model, the marginal probability distribution on the specific SFR (sSFR) in each bin follows a Dirichlet prior centered on a constant SFR(t). The time bins can be adjusted in both number and size by the user but remain fixed during the fit. An additional parameter, called the concentration parameter, is required to specify the prior and controls the concentration of stellar mass formation across time bins. Lower concentration values (~ 0.2) result in more bursty SFHs, while higher values (~ 1.0) result in smoother SFHs. We chose a value of 0.7 to allow for short-timescale variations in the SFH, though we briefly explore the impact of this prior choice in Appendix A. We also choose 10 time bins spaced logarithmically in time motivated by Ocvirk et al. (2006), who found that the ability to distinguish separate stellar populations is proportional to their difference in age in logarithmic time. The last two bins do not follow this prescription and instead span the previous 100 Myr and 100–300 Myr, allowing for a minimally young population of stars. The choices for this model are motivated largely by the impact of each choice on the priors for sSFR and stellar age: priors that are too narrow in inferred property space can result in biased estimates, while priors that are too wide can hinder model convergence. Though this model depends on the choice of time bins, as shown in Appendix A of Leja et al. (2019a) and our Appendix A, the stellar masses inferred from this model are robust against perturbations in the number and spacing of time bins, while the assumed prior on the fractional

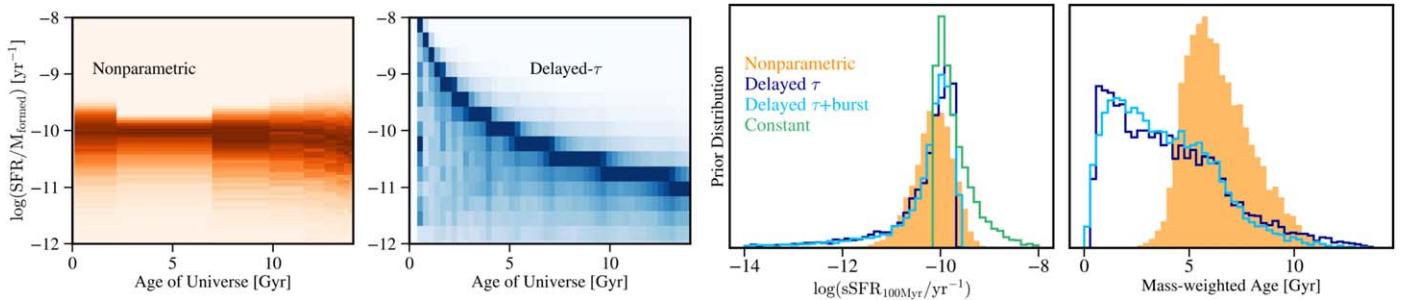


Figure 2. Left: prior probability distributions for the nonparametric SFH model and the delayed- τ model. Middle: effective prior distribution on the sSFR averaged over the past 100 Myr. Right: effective prior distribution on the mass-weighted stellar age. The constant SFH mass-weighted age prior is not computed, as this is simply $0.5 \times t_H$.

Table 2
Table of SED Models and Associated Parameters and Prior Distributions for PROSPECTOR SED Fitting

	Model	Parameters	Prior Distribution
SFH	Delayed τ	Age τ	Uniform(0.01, 13.8) Gyr LogUniform(0.001, 10) Gyr $^{-1}$
	Delayed τ +burst	Age, τ Burst time as a fraction of age Burst M_* fraction	As above Uniform(0.5, 1.0) * Age Uniform(0.0, 5.0) * M_*
	Constant	Age	As above
	Nonparametric	M_* fraction per time bin ($N = 10$) Concentration parameter α	Dirichlet Fixed at 0.7
M_* -Z	Gallazzi et al. (2005)	Log(M_* formed) Stellar metallicity Z_*	Uniform(7, 13) M_\odot ClippedNormal(-1.9, 0.19)
Dust emission	Draine & Li (2007)	Minimum radiation field U_{\min} Warm dust fraction γ PAH mass fraction q_{PAH}	Uniform(0.1, 30) Uniform(0.0, 1.0) Fixed at 5.86%
Dust absorption	Fixed dust screen	Uniform dust screen opacity τ_{uniform} Birth cloud dust opacity τ_{BC} Power-law index	Fixed at 0.7 Fixed at 0.7 Fixed at -0.7 (uniform) & -1.0 (birth cloud)

Note. All POWDERDAY SEDs are fit four separate times with a different SFH model. Three commonly used parametric SFH models are considered (delayed- τ , delayed- τ with a burst component, and a constant SFR across time) in addition to a nonparametric SFH model. The dust and stellar mass–stellar metallicity models are kept the same between runs, though the parameters are allowed to vary.

masses has a much larger impact. We will refer to this model as “nonparametric” throughout our following analysis.

2.4.2. Metallicity and Dust

Dust and metallicity also significantly impact the SED of a galaxy. Below we describe the models for each component. The same models and parameter prior distributions are used for each of the SFH model fits.

Metallicity: Following Leja et al. (2019b), we use a prior on the stellar metallicities as a modified version of the stellar mass–stellar metallicity relationship from $z = 0$ Sloan Digital Sky Survey (SDSS) data (Gallazzi et al. 2005). In practice, FSPS, and subsequently PROSPECTOR, will effectively assume a uniform metallicity across the galaxy for all time, constrained by the stellar mass–stellar metallicity relation. This is in contrast to the metallicity and chemical enrichment history present in the SIMBA simulation and PROSPECTOR stellar SEDs, where each star particle in a galaxy has an individual metallicity that contributes to the galaxy’s overall enrichment history. This complexity has yet to be implemented in most SED fitting codes, in part due to the degeneracies arising between metallicity, dust, and stellar age.

Dust emission: Constrained by energy balance, where the thermal emission from dust in the far-IR is assumed to be equal to the stellar light absorbed by dust (da Cunha et al. 2008), dust emission is modeled following Draine & Li (2007), which describes dust emission using three parameters: U_{\min} , which is the the minimum radiation field strength in units of the MW value, q_{PAH} , which is the mass fraction of dust in polycyclic aromatic hydrocarbon (PAH) form, and γ , which is the fraction of dust in high radiation fields. U_{\min} and γ are allowed to vary in the fits, but the PAH mass fraction is fixed to the true value (5.86%) because the synthetic PAH spectra are not sampled.

Dust attenuation: The dust attenuation model is fixed to match the POWDERDAY dust screen model. Thus, the SEDs have a uniform dust component with an optical depth of 0.7 affecting all stars and a birth cloud dust component affecting young stars (less than 10 Myr old) with an optical depth of 0.7.

3. Results

We fit SEDs for all SFH models described above. The following sections detail the results of the output of each SED fit from the fiducial run (i.e., with dust attenuation fixed to the true model). In general, we find that the commonly used

parametric SFHs struggle to reproduce the physical properties of the SIMBA galaxies, including the stellar mass, mass-weighted age, and the recent rate of star formation. For instance, the three parametric models systematically underestimate the stellar masses of the SIMBA galaxies by 0.4 dex on average. The properties mentioned above are directly dependent on the assumed SFH model: the stellar mass is the integral of the SFH across time (modulo the fraction of stars that now exist as stellar remnants and mass loss from post-main-sequence stars) and will depend on the amplitude of the SFH across time. The mass-weighted age of the galaxy will depend on the shape of the SFH, and the SFR depends only on the SFH over the past 100 Myr. We examine the efficacy of each SFH model to recover the above properties in the following sections.

3.1. Stellar Mass Recovery

In Figure 3, we examine the impact of the SFH model on the derived stellar mass (M_*) of our galaxies. In the top panel, we show a comparison between the stellar masses inferred from the various SFH models described above. The contours show the delayed- τ , delayed- τ +burst, and constant parametric SFHs, while the orange points show the distribution of galaxies when using the flexible nonparametric SFH model. We compare the derived M_* on the y -axis to the true M_* on the x -axis. The derived quantities are the median of the stellar mass PDF for each galaxy. The stated biases for each model are the average offset between the inferred stellar mass and the true value. The scatter is the 1σ standard deviation of this distribution.

The nonparametric models recover the true stellar masses with significantly better accuracy afforded by the flexibility of the SFH model. To quantify this, we use two plots to show the improved accuracy and uncertainty estimates afforded by the nonparametric model. On the bottom left, we show the cumulative inferred stellar mass offset for each SFH model. The stellar masses inferred from the nonparametric model are all within 0.4 dex of the true stellar mass, compared to the stellar masses inferred from the τ models, which suffer larger offsets. On the bottom right, we show the fraction of true stellar masses contained within the 1σ region of each galaxy stellar mass posterior, i.e., the true stellar mass falls between the estimated 16th and 84th mass percentiles. In other words, the uncertainty quantified by the stellar mass posterior width includes the true stellar mass. While the stellar mass PDFs inferred from the nonparametric SFH capture the true stellar mass for 50% of galaxies, the PDFs from the τ models include the true stellar mass value for only 20% of galaxies. The other 80% of galaxies have a stellar mass that is offset from the true value, with error bars that do not capture the true value. Thus, the reported model uncertainties do not reflect the systematic biases imposed on the inferred galaxy properties.

These three plots demonstrate the significantly improved accuracy afforded by the nonparametric SFHs as compared to the traditional parametric forms. Though M_* is traditionally considered the most robust property derived from SED fitting, the results here paint a different picture: parametric SFHs fail at recovering the true stellar mass for a majority of galaxies, across all stellar mass ranges, even when we fix the dust attenuation model to be the same in the SED generation and SED fits. Furthermore, the uncertainties associated with the inferred physical properties from parametric SFHs tend to be underreported because the model priors and subsequently

postiors are highly informative, meaning that stellar masses inferred from parametric SFH models will be systematically underestimated with underreported uncertainties (i.e., the uncertainties do not increase with increasing bias). The average offset in stellar mass for the nonparametric model is -0.02 dex with an average uncertainty of 0.11 dex, whereas for the delayed- τ model these values change to 0.38 and 0.19 dex, respectively.

3.2. Star Formation History Recovery

A natural question is how well a given model recovers the true SFH of a galaxy. In Figure 4, we show a randomly chosen galaxy's SFH and compare the recovered SFH for both a parametric and nonparametric SFH model. For this particular galaxy, the nonparametric model reasonably matches the true SFH with the median fit (solid orange line), and with the 1σ region (orange shaded region) also capturing the stochastic behavior in the true SFH over short timescales. The median delayed- τ model, however, fails to match the amplitude of the true SFH over much of cosmic time, which will result in an inferred stellar mass that is overestimated. The 1σ region for the delayed- τ also covers the true SFH but shows the large dispersion in SFH solutions throughout the fit.

That the nonparametric SFHs are more accurate at recovering the stellar mass of a galaxy is mainly attributed to the fact that they are significantly more flexible and thus better at describing the various SFHs seen in the SIMBA galaxy formation simulation. With only a small number of parameters describing the width and location of the curve, the three parametric SFHs (delayed- τ , delayed- τ with a burst component, and constant) struggle to match the true SFH for most galaxies. This will affect not only the stellar masses inferred from each model but also the stellar ages and SFRs. The two delayed- τ models struggle to match the true SFHs that rise over time, as only very large values of τ allow for a slower decline at late times.

For massive galaxies ($M_* > 10^{11} M_\odot$) at $z = 0$, the exponential decline of the parametric SFHs can match the true galaxy SFHs at late times, as these massive galaxies are typically quenched or quiescent, but only at the expense of missing the large, extended early periods of star formation and thus missing out on the bulk of formed stellar mass. The lower-mass galaxies tend to be bluer, star-forming galaxies with SFHs ill-suited for the exponential decline at late times, so that the true SFHs are not well recovered and the stellar mass estimates will be worse, a point confirmed by Figure 3.

Building on Figure 4, in Figure 5 we compare the offsets between the inferred SFHs and the true SFHs for all galaxies again for the nonparametric model and the delayed- τ model. The solid lines refer to the median offset for the entire galaxy distribution for each model, while the shaded regions include the 16th through 84th percentiles. The offsets were calculated between the median model SFH and the true SFH in 100 Myr intervals over the entire history. The median SFH offsets for both models are centered around zero for most of cosmic time, but the delayed- τ model has a much larger dispersion. Iyer & Gawiser (2017) found similar results to the Dense Basis nonparametric SFH method for stellar mass, SFR, and stellar age when fitting mock broadband SEDs from simulations and a sample of galaxies from CANDELS and comparing these results using SpeedyMC (Acquaviva et al. 2011, 2015).

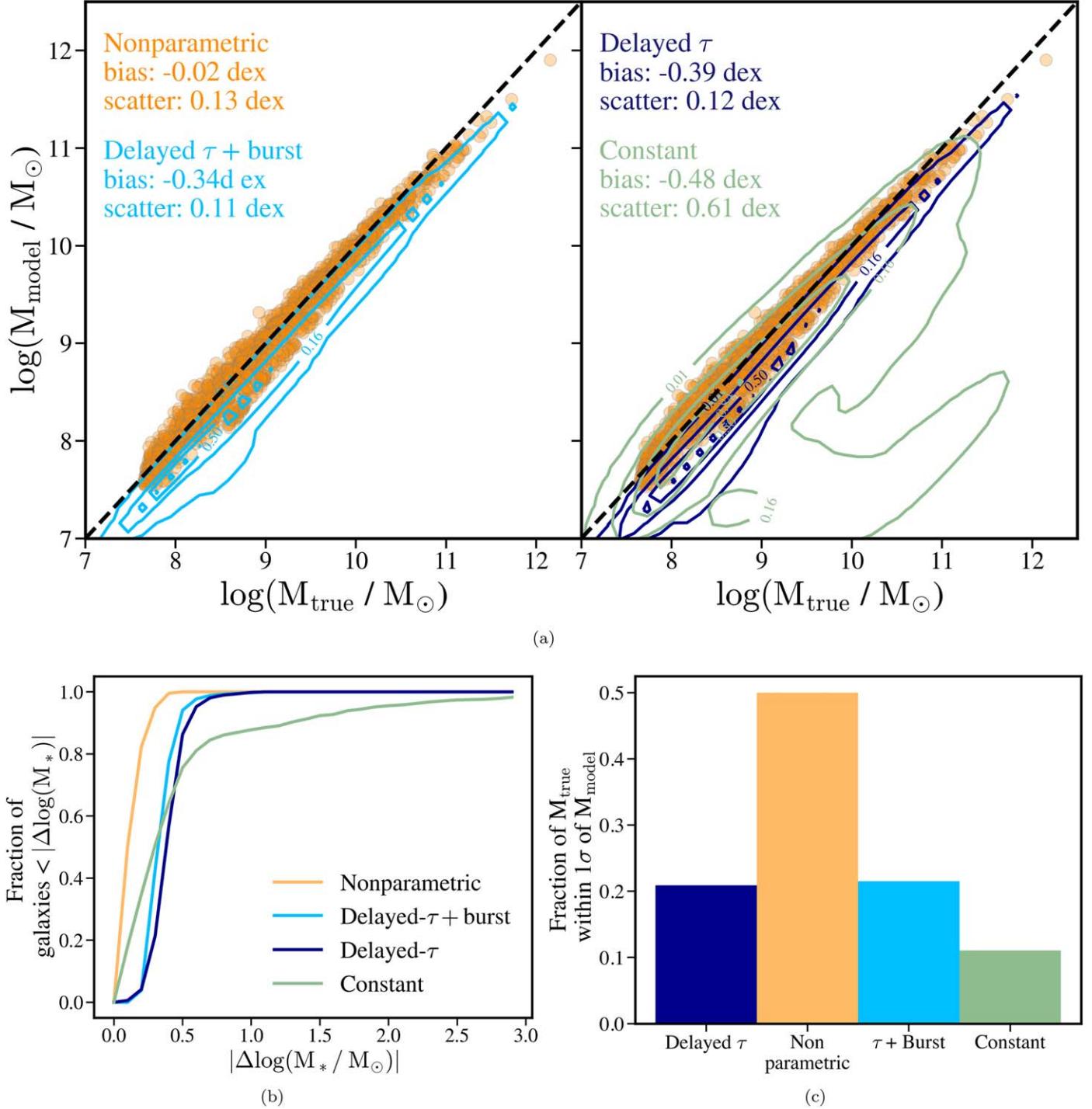


Figure 3. Top: comparison of inferred stellar M_{*} to true stellar M_{*} of SIMBA-simulated galaxies for all SFH models. The bias is the average offset between the inferred stellar mass and the true. The scatter is the 1σ standard deviation of this distribution. The masses inferred from the nonparametric SFH are shown in orange. Light-blue contours show the delayed- τ + burst SFH (a parametric model). The right panel is the same as the left, but green contours are for the constant SFH model and dark-blue contours show the delayed- τ model. Contour levels for the three parametric models highlight the 16th, 50th, and 84th percentiles. The black dashed line is the 1:1 relation (i.e., the ideal case where the inferred mass perfectly matches the true mass). Bottom left: cumulative fraction of galaxies with inferred M_{*} offsets. The stellar mass offsets are calculated as the absolute value of the difference between the $\log(\text{inferred } M_{*})$ and $\log(\text{true } M_{*})$. Bottom right: fraction of true galaxy M_{*} that are within 1σ of the median inferred M_{*} for each SFH model, where 1σ includes the 16th through 84th percentiles of the stellar mass posteriors. The stellar mass PDFs inferred from the nonparametric model capture the true stellar mass for more than 50% of galaxies, compared to just 20% for the τ models.

3.3. Ages and Star Formation Rates

The mass-weighted stellar age of a galaxy depends on the shape of the SFH, and the accuracy of these inferred properties therefore depends on the model SFH accurately matching the true galaxy SFH. We show the offsets from the true values for

the mass-weighted stellar ages and SFRs, along with stellar mass, in Figure 6. The offsets are defined as the difference between the median inferred value for each property and the corresponding true value for each galaxy. The M_{*} offsets are shown in the left panel, alongside the offsets in SFR over the

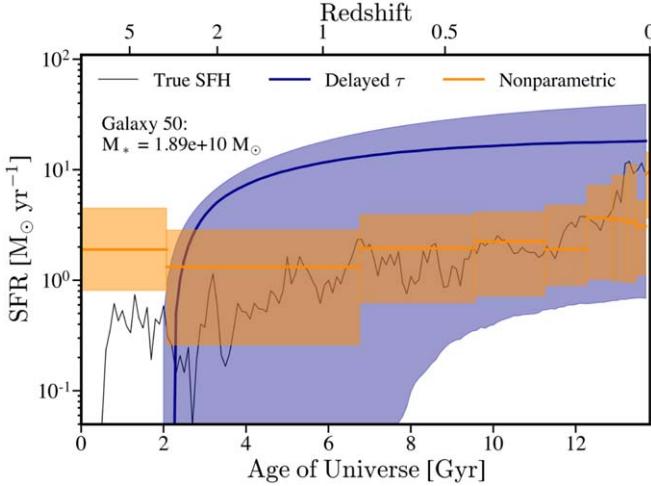


Figure 4. SFH for an example galaxy. The true galaxy SFH is shown in black. Two of the model SFHs are shown, with the nonparametric shown in orange and the parametric (delayed- τ) shown in blue. The 50th percentile value is shown as the solid line, while the shaded regions include the 16th through 84th percentiles.

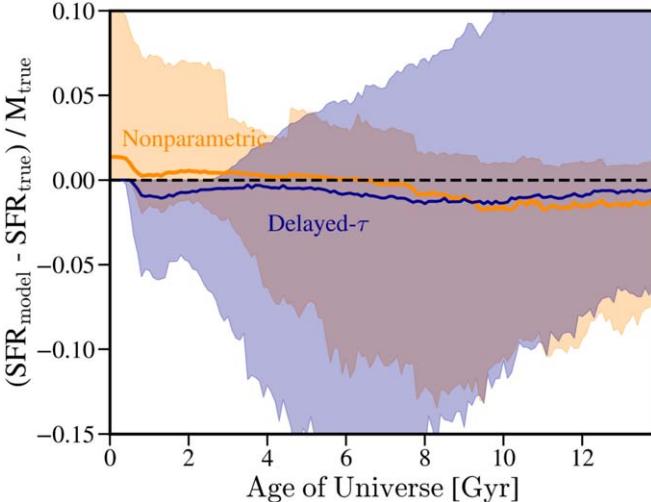


Figure 5. Comparison of the SFH residuals normalized by stellar mass of all galaxies for the nonparametric model and the delayed- τ model across time. The black dashed line represents the ideal scenario of a perfect match. Offsets were calculated between the median model SFH and true SFH every 100 Myr across the entire history. The solid lines refer to the median offset for each model, while the shaded regions include the 16th through 84th percentiles. The nonparametric model outperforms the delayed- τ model on average for all times.

past 100 Myr in the middle panel and the mass-weighted stellar age in the right panel, derived from the average of the inferred SFHs over the past 100 Myr and the full SFH, respectively.

The nonparametric model shows significantly better fits for each galaxy property when compared to the parametric models. The stellar mass is the most robustly inferred property, followed by the SFR and mass-weighted stellar age. The parametric models struggle to match all three properties simultaneously. For example, the constant SFH tends to capture the late-time SFR of the SIMBA galaxies, but the mass-weighted stellar ages cannot be accurately derived, as the age inferred from a constant SFH is just $0.5 \times t_H$. All three parametric models systematically underestimate the stellar mass and mass-weighted age of all galaxies. This bias is a

consequence of the use of strong priors and the behavior of these priors when fitting photometry that tends to prefer younger stellar populations, as shown in Figure 2 and Carnall et al. (2019) and Leja et al. (2019a). A drawback of the delayed- τ -like models is the trade-off between correctly inferring stellar age and SFR for a declining SFH. Unless the prior space is manipulated to allow for a rising SFR by, e.g., allowing for very large τ values (as in the case of, e.g., Acquaviva et al. 2011; Ciesla et al. 2017; Aufort et al. 2020), the SFRs will be biased low. On the other hand, the SFRs may be correctly inferred if the peak of the SFH is placed relatively close to the time of observation, but this will occur at the expense of inferring the correct stellar age.

4. Discussion

On average, the nonparametric SFH model outperforms the parametric SFHs on all metrics, including recovering the M_* , mass-weighted stellar ages, and late-time SFR of the SIMBA galaxies. The flexible nonparametric model used here is able to more accurately infer the physical properties and growth histories of galaxies from the SIMBA cosmological simulation. The flexibility of the nonparametric model, compared to the relatively inflexible parametric models, is twofold: (1) the time resolution allows the SFH model to describe the shorter-timescale fluctuations in the galaxy SFH, and (2) the prior on the fractional SFR in each time bin results in effectively unbiased priors on sSFR and stellar age. Though the priors on these properties are not flat, they are unbiased in the sense that the center of the prior distributions does not prefer high or low values for each property; the median of the mass-weighted stellar age distribution is centered on the Madau & Dickinson (2014) estimate of the median stellar age derived from the CSFRD, while the median sSFR is centered at $\log(1/t_{\text{univ}}) \sim -10.1 \text{ yr}^{-1}$, which corresponds to a constant SFR. The latter point is in contrast with the τ models, where younger stellar populations are preferred owing to the priors imposed on the shape of the SFH.

Moreover, the danger in applying a τ model to a sample of galaxies lies in the false constraints on galaxy properties imposed by the SFH priors. This results in biased inferred galaxy properties with underestimated uncertainties (i.e., the uncertainties do not increase with bias). In this section, we discuss the ways in which each SFH model impacts the inferred galaxy properties and why some models perform better than others. We also discuss the importance of carefully chosen priors when fitting SED photometry, for both nonparametric and parametric SFH models, and briefly discuss how including diffuse dust in this analysis impacts the inferred stellar masses.

4.1. Parametric Models

All three parametric SFH models considered (constant SFR across time, delayed- τ exponentially declining SFH, and delayed- τ with an additional burst component) have been shown to systematically underestimate the stellar masses of the SIMBA galaxies. This is true for galaxies of all masses and sSFRs. We show the stellar mass offsets as a function of galaxy stellar mass in the top right panel of Figure 7 for the two delayed- τ models (the constant model is neglected for the sake of clarity). The solid lines refer to the running median of the stellar mass offset distribution for each model. The τ models systematically underestimate the stellar masses by approximately 0.4 dex for all galaxies. Compared to previous studies

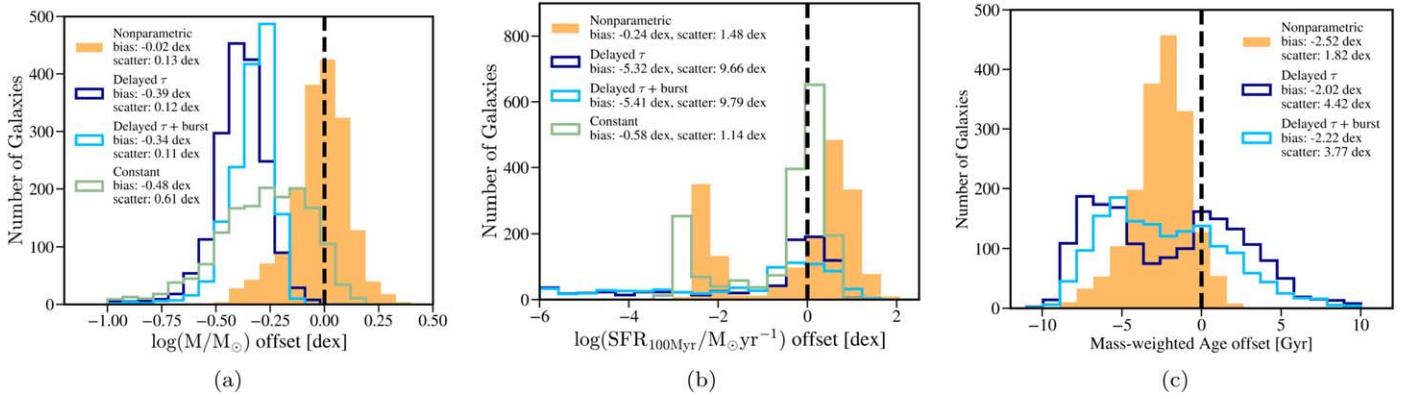


Figure 6. Offsets from the true values of inferred galaxy properties. Properties inferred from the nonparametric model lie much closer to offsets of 0 than the parametric models. The stated bias for each model is the average offset between the inferred galaxy property and the true value. The scatter is the 1 σ standard deviation of this distribution. Left: inferred stellar mass offsets for each SFH; 13% of galaxies fit with a constant SFR have M_* offsets > 1 dex. Middle: same as the left panel, but for SFR over the past 100 Myr. Right: same as the left and middle panels, but for mass-weighted age. Note that the mass-weighted age for the constant SFH is $0.5 \times t_H$, so we neglect plotting this inferred distribution.

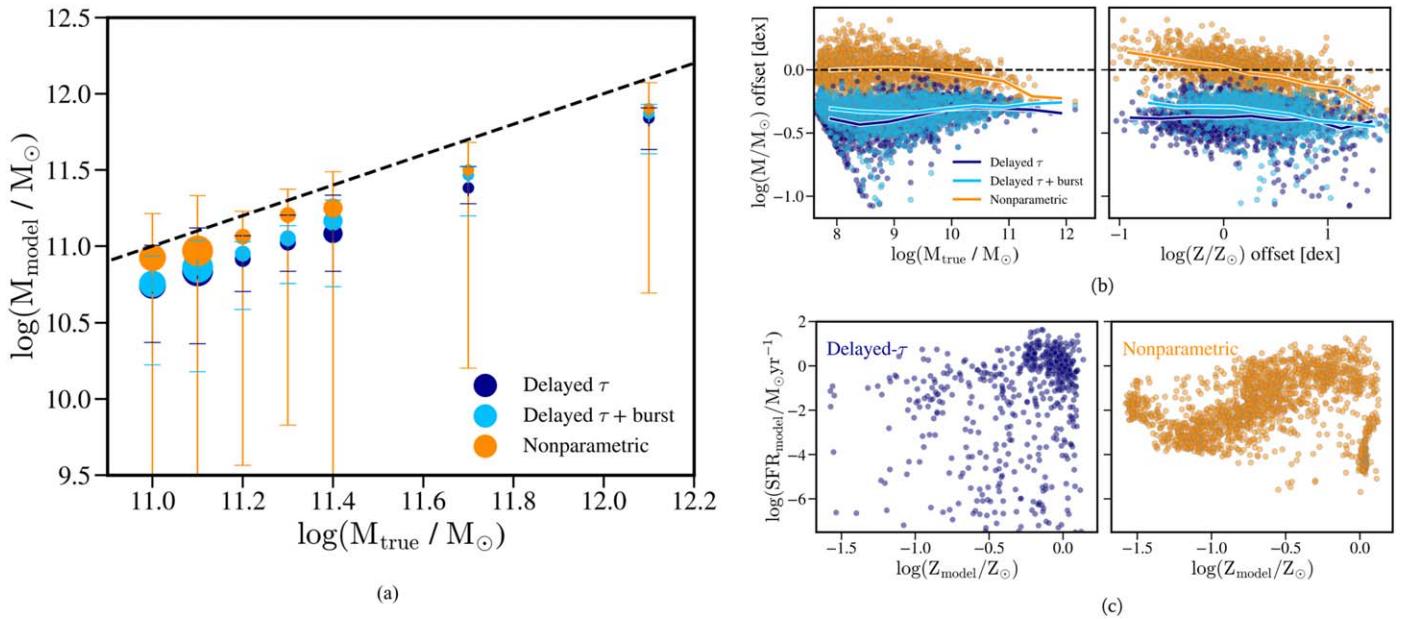


Figure 7. Left: binned 1:1 stellar mass comparison including uncertainties for high-mass galaxies. The sizes of the points reflect the number of galaxies in each bin. The average uncertainties for each mass bin cover the biases for the nonparametric model, while the uncertainties are smaller than the biases for the parametric models for the highest-mass bins. Top right: stellar mass offset as a function of true stellar mass (left) and metallicity offset (right). Bottom right: inferred SFRs as a function of inferred metallicities for the delayed- τ and nonparametric SFH models.

on the effect of assuming a parametric SFH model (e.g., Pforr et al. 2013; Carnall et al. 2019), our inferred stellar mass offsets for the τ models are similar. Pforr et al. (2012) found average stellar mass offsets of 0.6 dex for mock galaxies at $z \sim 0.5$ when reddening effects were considered. Ruling out unrealistic dust and age solutions lowered these median offsets to 0.2 dex. Though we could, in principle, apply a correction for the offset in the τ models, we choose not to do so because such a correction would depend on the model assumptions made here, both in the SED fitting procedure and in the SIMBA model, and because the scatter of the tau stellar mass distributions is sufficiently large. Our aim of this analysis is not to provide correction estimates for SED models or to advocate for any one set of SED models or parameters but to provide ground-truth tests for these assumptions. Additionally, for real observations, due to the complex degeneracies in SED fitting, an inaccurately inferred SFH typically means that the inferred dust, metallicity,

nebulosity, properties, etc., are also wrong. An overall correction applied to mass is not only dependent on the model assumptions in this analysis but also ultimately unsatisfactory because stellar mass is not the only output that depends on an accurate SFH and would otherwise negatively impact results inferred for real observations.

The inability of the parametric models to accurately recover the stellar mass of the SIMBA galaxies is a consequence of the restrictive nature of these models. We find that unless the inferred e -folding time is sufficiently large to allow prolonged star formation, especially at early times, the stellar masses recovered from these models will be underestimated. A by-eye analysis of the library of SFHs measured from the SIMBA simulation shows that the τ and delayed- τ SFH models are a poor match for a majority of the star-forming galaxy population at redshift $z = 0$ unless the e -folding time is sufficiently large ($\tau \sim 5$ Gyr – 1) to model the slow decline of quenching or to

model a rising SFH. This is a consequence of the specific galaxy formation models implemented in SIMBA but is not unique to SIMBA or galaxy formation simulations in general. While a delayed- τ SFH model may match the SFR over time of an isolated closed-box galaxy (i.e., no gas inflows or outflows such that star formation is an exponential function of the in situ gas mass and gas depletion time), it is not flexible enough to describe the shorter-timescale fluctuations seen in both simulated and observed SFHs. Flexibility in the form of additional bursts of SFH or the use of a lognormal model with a large volume of parameter prior space to sample from, as presented in Diemer et al. (2017), can result in more accurate stellar masses but does not remedy the wide dispersion in late-time SFR and stellar age as discussed below. And though the photometry inferred from each model is a reasonable fit to the observed SEDs, the inferred stellar masses differ from the true values by -0.38 dex on average. Again focusing on the τ models, the inferred value of τ and the placement of the peak of the SFH will impact the inferred stellar age and recent SFR. These models can also underestimate the recent SFR of the galaxy by several orders of magnitude, evidenced by Figure 6(b), and severely underestimate the age of the galaxy in order to match the observed SFR, evidenced by Figure 6(c). These biases are driven by the strong priors such that the parametric model is forced between including either recent star formation or an older population of stars, but not both unless the model is distorted and forced away from an exponential decline. The strongly peaked priors are also responsible for the uncertainties that do not increase in step with biases. We show in the left panel of Figure 7 a binned 1:1 stellar mass plot, zoomed in to show galaxies in the highest-mass bins. The size of the points represents the number of galaxies in each bin. The uncertainties for a majority of the bins do not compensate for the large biases seen with either τ model.

4.2. Nonparametric Model

Again in the top right panel of Figure 7, we show the stellar mass offsets for the nonparametric SFH model in orange. This model, in contrast with the parametric models, achieves much more accurately inferred stellar masses for galaxies of all masses. The average stellar mass offset improves from 0.38 dex with the delayed- τ model to under 0.1 dex for the nonparametric model. We note the increasing trend of the magnitude of offsets for high-mass galaxies ($M_* > 10^{11} M_\odot$). However, only 20 galaxies populate this region of stellar mass space, so it is difficult to draw conclusions about the performance of this model on high-mass galaxies. We show a zoomed-in 1:1 stellar mass plot for these galaxies in the left panel of Figure 7. Though the stellar mass offsets grow for increasing stellar masses, the uncertainties increase in step for the nonparametric model.

For the handful of high-mass galaxies in this particular SIMBA snapshot, star formation has effectively stopped anywhere between 1 and 4 Gyr ago. To accurately infer the stellar mass of a galaxy, the model must match the true SFH at early times when the older stars that dominate the stellar mass are formed. However, early bursts of star formation, even for prolonged periods of time, do not leave obvious artifacts on a galaxy's SED, so accurately matching the early SFH is difficult for any SED model. Anecdotally, many of the massive galaxies from this SIMBA snapshot experienced prolonged periods of enhanced star formation peaking at $\sim 100 M_\odot \text{ yr}^{-1}$ around

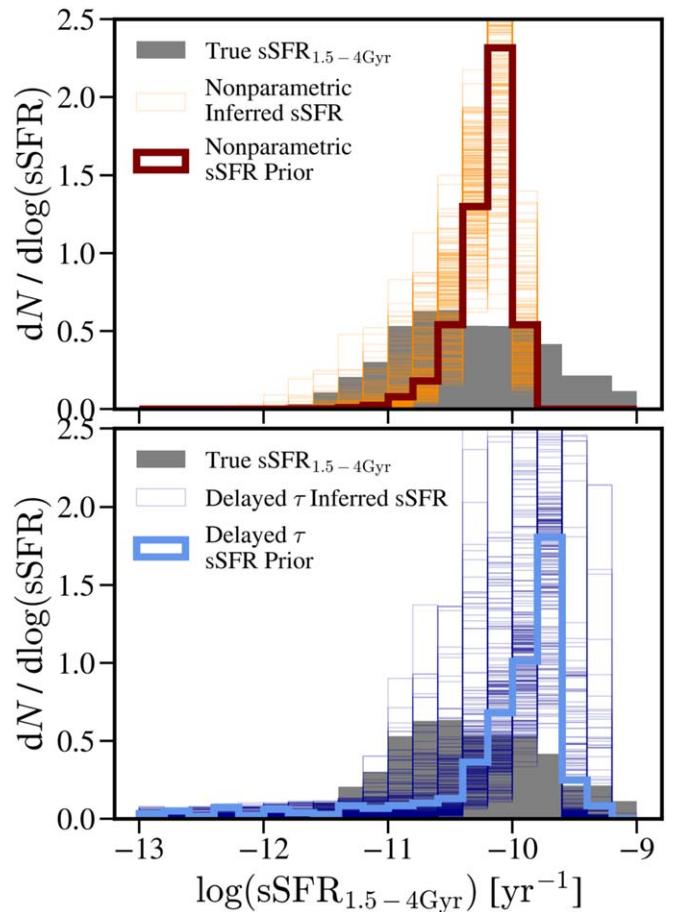


Figure 8. Distributions of the average sSFRs during the first 1.5 – 4 Gyr. The true distribution for the SIMBA galaxies is shown by the gray histogram. Top: nonparametric inferred sSFR posteriors for 300 randomly selected galaxies. The prior distribution on early sSFR is shown with the bold maroon line. Bottom: same as the top panel, but for the delayed- τ model. The prior distribution is shown with the bold light-blue line, with individual posteriors drawn in dark blue.

10–12 Gyr ago that are not recovered well by any star formation model considered here. This is demonstrated in Figure 8, where we show the distributions of the average sSFR over the first 2–4 Gyr for the SIMBA galaxies compared to the inferred early sSFRs from two of the SFH models. As explored in Iyer et al. (2019), inferred SFRs from more than a couple gigayears ago are dominated by the prior set on the SFH rather than the fit to photometry, as little evidence exists in the observed SEDs of these early star-forming episodes. This problem is not unique to PROSPECTOR or the nonparametric SFH model used here but is found in all models, as SFHs are only minimally informed by broadband photometry owing to a lack of SED features left by old stellar populations. The distribution of early sSFRs inferred from the nonparametric SFH model is narrowly peaked at $\log(\text{sSFR}) \sim -10.5$, so that galaxies with smaller or larger sSFRs at early times will have inferred stellar masses that deviate from the true value, as long-lived stars with solar masses or lower will dominate the stellar mass content of a galaxy.

For galaxies that are actively star-forming, the problem of “outshining” will augment the above difficulties, as the massive stars that are formed recently will outshine the older stellar populations that dominate a galaxy's stellar mass such that the

inferred stellar mass will be heavily dependent on the priors informing the stellar age distribution (e.g., Papovich et al. 2001; Pforr et al. 2012). In other words, the stellar masses inferred for these systems are dominated by the constraints from the prior (shown in Figures 2 and 8) in addition to the constraints from photometry. In both cases, for quiescent and actively star-forming galaxies, the underlying cause of the differences in model and true stellar masses is the inability to capture early, prolonged star formation activity with the model SFH.

Besides the assumed SFH, assumptions about a galaxy's stellar metallicity and the mapping from metallicity to the SED will also impact the inferred stellar mass. In this analysis, we used the Gallazzi et al. (2005) stellar mass–stellar metallicity relation. However, as noted in Section 2.4, this relatively simplistic model does not entirely capture the growth history of the SIMBA galaxies, as the simulated star particles have metallicities that are neither uniform nor static through time. Though we isolated our results from the inclusion of a diffuse dust component, in truth our assumptions in this metallicity model will also impact the stellar masses inferred by PROSPECTOR. In the top right panel of Figure 7, we show the stellar mass offsets as a function of the stellar metallicity offsets. The solid lines refer to the running median of each distribution. For the sake of comparison, we take the true galaxy metallicity to be the mass-weighted stellar metallicity for each galaxy. This way, we have an aggregate metallicity to compare to the inferred stellar metallicities. We see a strong inverse correlation between the offset distribution for galaxies fit with the nonparametric model but not for either parametric model. This trend is also seen in fits where the stellar metallicity is not tied to the stellar mass. One explanation for this is the inability of the uniform and static metallicity model to recover the distribution in true stellar metallicities. A model describing the chemical enrichment history of galaxies could be implemented but would be minimally constrained by broadband photometry and most likely highly degenerate with the SFH and dust model. In the absence of dust, stars with lower inferred metallicity will have excess UV compared to stars with higher inferred metallicity. This change in UV light in the SED can impact both the inferred SFR and, to a lesser degree, the inferred stellar mass. In Figure 7(c) we see that a trend also exists between the inferred SFR and metallicity from the nonparametric model, where generally galaxies with larger inferred SFRs will have large metallicities. These recent SFR bins are covariant with metallicity, and allowing freedom in these, which is not allowed by the tau models, introduces a covariance with stellar mass that could explain the trend we see in Figure 7(b).

4.3. Model Priors, Uncertainties, and Degeneracies

An underlying theme of this analysis is the flexibility afforded by the nonparametric SFH model. This flexibility is in the form of both model variability and minimally informative (i.e., only moderately peaked) priors. Here we briefly comment on the dependence of our results on the choice of hyperparameters for the nonparametric model (concentration of mass formation, number of time bins). We also comment on the change in performance in the presence of additional uncertainties from the stellar population synthesis models and observational noise.

4.3.1. Tuning of the Nonparametric Model

The nonparametric model hyperparameters and other priors described in Section 2.4.1 were chosen for maximum flexibility. While the choice of prior for the degree of mass concentration (i.e., large variations over short timescales versus a smoother SFH) matters for the nonparametric SFH (Appendix A), we find that the results presented here are not heavily dependent on the choice of time bins. The priors on either end of the spectrum in terms of favoring smooth SFHs result in stellar mass offsets that are smaller than the parametric models but are much larger compared to the fiducial nonparametric SFH model used in this analysis. The choice of prior is equivalent to constraining the shape of the SFH; thus, the strongest prior on the inferred stellar mass is set by the choice of an SFH form that a priori assumes a certain shape, as is the case for the parametric models. Moreover, the priors on the τ model parameters (i.e., τ and the time at which the SFH peaks) generally cannot be tuned to a given data set because (i) that could exclude large regions of parameter space for the inferred properties and (ii) the parameters couple differently in different contexts. For instance, if we want to do well in recovering SFHs for galaxies with rising SFRs at late times, we could impose a prior on τ that would favor large e -folding times. But this would necessarily bias the results against galaxies that do not have rising SFHs. Furthermore, the influence of the priors of the parametric model parameters like τ and t_{age} on the subsequent galaxy physical properties is not straightforward; setting an uninformative prior on τ will not result in a flat prior on SFR, sSFR, or stellar age. The effective priors on galaxy properties are primarily driven by the choice of a declining exponential form in the first place. Comparing the delayed- τ model to the simpler τ model, which does not allow for rising SFHs even with large τ , Wuys et al. (2011) found little difference between ages and SMFs inferred from either model, especially for galaxies subjected to outshining. In other words, modifying the τ model to allow for rising SFRs did not result in significantly better results for stellar mass or age. Because the nonparametric model does not favor any SFH shape beforehand, the model can tackle the wide diversity of SFHs as seen in SIMBA and elsewhere.

4.3.2. Uncertainties from Stellar Evolution and Observational Noise

A major source of uncertainty in SED modeling originates from stellar evolution; the choice of stellar isochrones, spectral libraries, and IMF contributes to large uncertainties in stellar mass (e.g., Conroy 2013; Pforr et al. 2013; Santini et al. 2015). Our results have been independent of these uncertainties since we mirror the stellar population synthesis (SPS) models between the synthetic POWDERDAY SEDs and the SED modeling with PROSPECTOR. One such aspect of SPS modeling is the assumed stellar IMF. Remodeling the POWDERDAY SEDs with a mismatched IMF (Chabrier 2003 versus Kroupa 2002), we find that, even in the presence of SPS uncertainty, the distribution of stellar mass offsets between the nonparametric model and the delayed- τ model remains clear (Appendix B). However, the assumed stellar IMF is just one component of an SPS model. The assumed stellar spectral library and whether to use an empirically based or theoretical library, along with the choice of stellar isochrones, have been shown to impart serious uncertainties on inferred stellar masses (Conroy et al. 2009). These uncertainties stem from difficulty

in sampling rare stars (e.g., massive, low metallicity) and stars in relatively short-lived phases (e.g., thermally pulsating AGB stars), as well as difficulties in modeling stellar atmospheres to produce theoretical spectra. The issue of how to model thermally pulsating AGB stars is of particular interest, as it has been shown to largely impact inferred stellar masses since these stars dominate the near-IR luminosity of intermediate-age galaxies (Maraston et al. 2006; Conroy et al. 2009).

Further observational uncertainties on inferred stellar masses originate from photometric noise, especially for data sets containing spectral line and continuum information. The additional information provided by spectra helps to discern different stellar populations and provides more information about the metal content in galaxy, potentially helping to avoid degeneracies between these parameters. However, spectra tend to suffer from calibration issues, and the information provided by spectra is dependent on the signal-to-noise ratio of said spectra. Though we do not consider spectra from nebular emission in this analysis, we conduct a brief test of the performance of the SFH models in the presence of noise by perturbing the input broadband photometry points within the 3% uncertainties. We show in Appendix B, Figure 11 that fits to these noisy SEDs produce results similar to the above analysis; thus, the results presented here, isolated from additional sources of uncertainty, hold. However, in future work we will conduct a more in-depth investigation into how these results change as a function of varying signal-to-noise ratio and SED coverage; for a thorough investigation into the efficacy of these models for varying signal-to-noise ratios of mock empirical spectra, see Leja et al. (2019a).

4.3.3. Uncertainties from Model Degeneracies

How an SED model handles degeneracies between model parameters and their mapping to the observed SED is a large difference between the parametric and nonparametric SFH models studied here. For the parametric SFHs, placing a strong prior on, e.g., the shape of SFH will affect the properties derived from the model and the correlations between properties by ruling out certain parameter combinations from the beginning. Though this can break degeneracies between, e.g., dust and stellar age, it is done so through strongly peaked priors with little physical basis and results in strongly biased predictions. For instance, we see in Figures 2 and 6(c) that the parametric models all favor younger stellar populations than the true ages even in the absence of realistic dust, an aspect of the τ model priors noted by Carnall et al. (2019). Because many degenerate, though physically plausible, solutions are ruled out a priori by the model choice (e.g., two stellar populations produced by two separate star bursts), the inferred uncertainties on galaxy properties like stellar mass are underestimated, as shown in Figures 3(c) and 7(a). The nonparametric SFH, on the other hand, is subject to the same degeneracies but does not impose such strong biases on inferred properties as a result of the carefully applied priors. In the Bayesian framework, an overly complex model will provide unconstrained results; thus, if the chosen number of time bins is too great relative to the input information or the model solutions are highly degenerate, the galaxy properties inferred with a flexible nonparametric SFH model will return increasingly larger posteriors. We see evidence of this in the middle panel of Figure 7, where average uncertainties for some of the high stellar mass bins extend more than 2 dex.

Focusing on the stellar mass inferred from the PROSPECTOR SED fits, the nonuniform sensitivity to variations in the SFH of a galaxy makes it difficult to untangle the contribution of old stars, young stars, metallicity, and dust to the integrated SED of a galaxy and can result in inaccurately inferred galaxy properties. It is worth restating that the results from the specific implementation of nonparametric SFHs presented here outmatch the accuracy of the three commonly used parametric forms considered, as shown in Figures 3 and 6. The use of more simple parametric SFH forms to determine the physical properties of galaxies will result in systematically biased stellar mass estimates with SFRs and stellar ages that are not well constrained, even in a best-case scenario of 25 broadband photometry points and a constrained dust model.

4.4. The Impact of Diffuse Dust

The results discussed so far do not consider the impact of diffuse galactic dust on the inferred stellar mass from SED modeling. Instead, we forced the POWDERDAY dust radiative transfer simulations to employ a dust screen model around stars in order to force an apples-to-apples comparison with the SED fitting techniques that also model dust via a screen model. This allowed us to isolate the differences between the input POWDERDAY SEDs and the output PROSPECTOR SEDs and focus on how the assumed SFH model affects the inferred stellar mass. However, by neglecting the true dust distribution and dust-to-star geometry, we underestimate the uncertainty with which we can infer the stellar mass of a galaxy. In this section, we now briefly explore how including diffuse dust impacts the results presented so far (noting that a full exploration of the impact of dust attenuation will be presented in future work).

4.4.1. Diffuse Dust Radiative Transfer

For the exploration of diffuse dust on our results, we use POWDERDAY to perform radiative transfer on the SIMBA galaxies, as described in Section 2.3. In this situation, however, we abandon our initial assumption of a uniform dust optical depth for all stars. Instead, the FSPS stellar SEDs are propagated through the dusty ISM. The diffuse dust content is derived from the on-the-fly self-consistent model in SIMBA (Li et al. 2019), and this dust is assumed to have extinction properties following the carbonaceous and silicate mix of Draine & Li (2007), which follows the Weingartner & Draine (2001) size distribution and the Draine (2003) renormalization relative to hydrogen. We assume $R_V \equiv A_V/E(B-V) = 3.15$. We do not assume further extinction from subresolution birth clouds. PAHs are included following the Robitaille et al. (2012) model, in which PAHs are assumed to occupy a constant fraction of the dust mass (here, modeled as grains with size $a < 20 \text{ \AA}$) and occupying 5.86% of the dust mass. The dust emissivities follow the Draine & Li (2007) model, though they are parameterized in terms of the mean intensity absorbed by grains, rather than the average interstellar radiation field as in the original Draine & Li model.

4.4.2. Impact of Diffuse Dust on the Inferred Stellar Mass

To include the impact of diffuse dust in this analysis, we again fit the SIMBA broadband SEDs with PROSPECTOR, this time allowing a flexible dust attenuation curve following the parameterization presented in Kriek & Conroy (2013). In

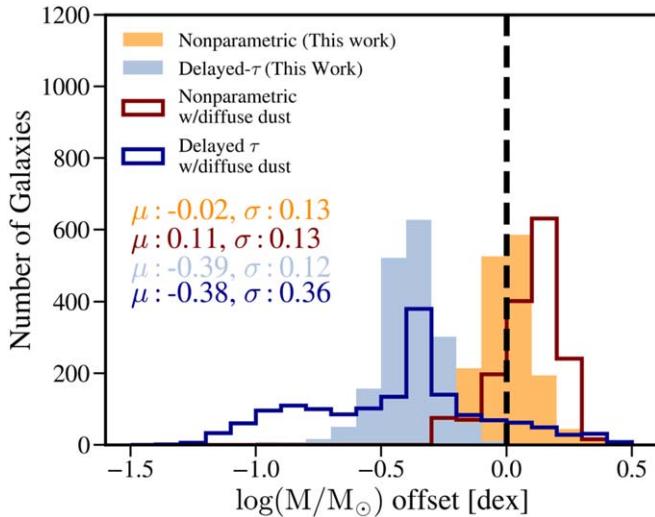


Figure 9. Distribution of stellar mass offsets for the dust screen models (filled histograms) and the models including diffuse dust and a variable dust attenuation curve (open histograms). The median and 1σ width of each distribution are shown.

Figure 9, we compare the stellar mass offsets from two of the SFH models with and without diffuse dust for a subset of SIMBA galaxies. The stellar masses inferred from the nonparametric SFH model are only marginally affected by the addition of diffuse dust. The difference in the median of the distribution of stellar mass offsets between the two data sets is statistically insignificant. The delayed- τ model, on the other hand, is significantly affected by the addition of diffuse dust: the magnitude of stellar mass offsets extends to just over 1 dex. And while, like the nonparametric model, the median of the stellar mass offset distribution remains largely unchanged for the delayed- τ model, the dispersion of the distribution has increased, such that a larger fraction of galaxies have larger offsets compared to the SED fits without diffuse dust.

The addition of diffuse dust and variable attenuation curve not only increases the number of model parameters PROSPECTOR must fit simultaneously but also introduces degeneracies between model parameters that are difficult to untangle, even with the constraints on, e.g., stellar metallicity in place. The addition of diffuse dust amplifies the drawbacks of using a parametric SFH model shown previously. In future work, we will explore these drawbacks in detail by considering a flexible dust attenuation curve to capture the diversity of attenuation curves present in the universe and to follow up this analysis by determining the impact of observational sources of uncertainty on inferred galaxy properties.

5. Conclusions

We have used simulated galaxies from the SIMBA cosmological simulation to understand our current ability to infer the stellar mass of a galaxy with SED modeling. In particular, we assessed the impact of the SFH model in SED fitting. We considered four SFH models, three commonly used parametric models and one nonparametric model, included in the PROSPECTOR modeling framework. We have demonstrated that the biases in stellar mass estimates decrease significantly with the use of nonparametric SFHs, falling below 0.09 dex for a majority of modeled SIMBA galaxies. The conclusions from our analysis are as follows:

1. Stellar masses derived from the PROSPECTOR nonparametric SFH models are much more accurate on average than those derived from parametric models, for galaxies of all stellar masses, ages, and morphologies. Figure 3(a) shows the median derived M_* for galaxies assuming different SFHs models. We find that the offset between the inferred M_* and the true M_* decreases from 0.4 dex on average when modeled with a delayed- τ SFH to -0.02 dex when modeled with a nonparametric SFH. Outliers exist for galaxies at the high-mass end, caused by failures to recover early periods of intense star formation. An important note is that while our results improve when using this particular nonparametric model, the differences between nonparametric and parametric begin to smear out once diffuse dust, noise, and other model uncertainties are considered. That said, we present here an estimate for the baseline M_* uncertainties achievable with current broadband SED models.
2. Parametric SFHs suffer from biases that are larger than their associated uncertainties, as explored in this analysis and in, e.g., Simha et al. (2014), Salmon et al. (2015), Carnall et al. (2019), and Curtis-Lake et al. (2020). The danger in applying a delayed- τ SFH to a sample of galaxies, aside from the relatively poor match to true SFHs on average, lies in the false constraints imposed on galaxy properties by the SFH priors. As shown in Figure 3(c), we find that the nonparametric SFHs tend to capture the true mass value within the 1σ M_* posterior width for a much higher fraction of galaxies compared to the three parametric models considered.
3. Nonparametric SFHs in PROSPECTOR are also able to match the true SIMBA SFHs across time significantly better than the parametric models, as shown in Figure 5. This increase in accuracy is owed to the well-constrained (i.e., through the choice of prior) flexibility permitted by the nonparametric model, so that the SFR at any one epoch is not determined by the SFR at another time. Mass recovery can be further improved by using more discerning data such as spectra or narrowband photometry, which could provide better constraints for early star formation activity.

Though the nonparametric SFH model used here outperformed the other parametric models on all metrics, it is important to note that these models must be used carefully. As described in Leja et al. (2019a), the priors that are chosen to constrain a nonparametric SFH are the primary drivers of the size of the inferred M_* posterior. For the nonparametric models in PROSPECTOR, this means that the stellar mass posteriors, while wider than those of the parametric SFH models, are able to capture the true stellar mass of the simulated galaxies at much higher fractions. The PROSPECTOR nonparametric priors, including the Dirichlet prior chosen for this analysis, perform much better than the parametric SFHs across the board, but the degree to which this performance improves is dependent on the choice of prior, as shown in Appendix A.

The difficulty in SED modeling lies in the fact that the SFH is only moderately constrained by broadband photometry, so priors must be carefully implemented to allow a diverse range of SFHs to be modeled while simultaneously fitting for dust and other model parameters. On this point, the PROSPECTOR nonparametric models are significantly better than the parametric SFHs used here and some more simplistic implementations of

nonparametric SFHs at producing meaningful error bars thanks to the carefully chosen priors, and SED modeling can only benefit from other thoughtful implementations of model priors. Finally, cosmological simulations can play an important role in future work to constrain priors not only for SFHs but also for dust attenuation laws. We can also develop nonparametric models for dust attenuation in a similar way, but the increase in computational resources and model degeneracies warrant caution. As such, we hope to explore further improvements to SED modeling and deriving physical properties from broadband photometry.

D.N. acknowledges support from the US NSF via grants AST-1715206 and AST-1909153 and from the Space Telescope Science Institute via grant HST-AR-15043.001-A. J.L. is supported by an NSF Astronomy and Astrophysics Postdoctoral Fellowship under award AST-1701487. C.C. acknowledges support from the Packard Foundation.

The SIMBA simulation used in this work was run on the ARCHER U.K. National Supercomputing Service, <http://www.archer.ac.uk>.

Software: SIMBA (Davé et al. 2019), CAESAR (Thompson 2014), POWDERDAY (Narayanan et al. 2015), HYPERION (Robitaille 2011), FSPS (Conroy et al. 2009), PYTHON-FSPS (Foreman-Mackey et al. 2014), PROSPECTOR (Johnson & Leja 2017), NUMPY (van der Walt et al. 2011), MATPLOTLIB (Caswell et al. 2018), ASTROPY (Virtanen et al. 2020).

Appendix A Nonparametric SFH Hyperparameters

The nonparametric SFH used in this analysis (described in full in Leja et al. 2019a, along with all other models available in PROSPECTOR) is described by N bins of constant SFR, where $N = 10$ and the fraction of mass formed in each bin of star formation is constrained by a transformation of the Dirichlet

prior (Betancourt & Girolami 2013) parameterized by α , which sets the concentration of mass formation. Modulating the value of α modulates the preference for all stellar mass to be formed in one bin ($\alpha < 1$) versus a smoother distribution of stellar mass ($\alpha > 1$). In practice, the fractional mass formed in each bin is fit, as opposed to the actual value of mass formed in each bin, so as to avoid sampling from a large volume of prior space and to separate fitting the shape of the SFH (done by the fractional mass) from the surviving stellar mass (the normalization of the SFH). Many tests were run to determine the dependence of the inferred stellar masses on these hyperparameters. Figure 10 shows the inferred stellar mass offsets for different choices of time bins and priors. We also include summary statistics in Table 3. The choice of priors includes changing α in the Dirichlet prior along with a different prior distribution, called the Continuity prior, which favors small changes between adjacent time bins of star formation (similar to Dirichlet priors with large α values, though star formation between bins is explicitly tied to each other). The dominant factor is the choice of prior constraining the fractional mass in each time bin, highlighting the importance of choosing a prior suited to a particular data set.

The continuity prior was tuned in Leja et al. (2019a) to galaxy SFHs from the Illustris hydrodynamical simulation. As recently presented in Iyer et al. (2020), galaxy SFHs in Illustris have less power on short timescales compared to galaxies from SIMBA. This means that SIMBA galaxies tend to have more fluctuations on short timescales, so implementing an SFH prior favoring small fluctuations would tend to give worse stellar mass estimates. Furthermore, because the “correct” choice of prior for any sample of galaxies, both simulated and observed, will change depending on the class of galaxy (star-forming versus quiescent, local versus high redshift), we do not suggest that the particular priors we imposed on the SIMBA data set will be correct for all galaxies.

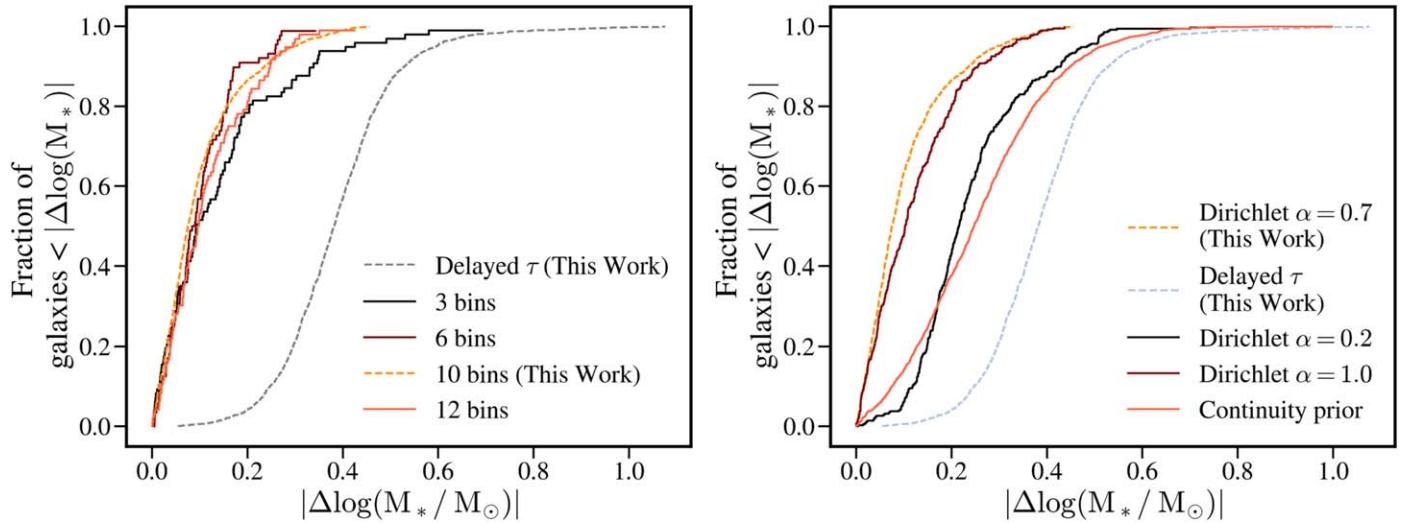


Figure 10. Cumulative fraction of galaxies with inferred stellar mass offsets for varying numbers of time bins (left) and prior on fractional mass per time bin used in the nonparametric SFH model (right). The data presented in this analysis are shown by the dashed lines.

Table 3
Summary Statistics for SED Models with Differing Parameter Choices

SFH Model	Data Set/Model	Median M* Offset	
Nonparametric	Concentration parameter	0.2 0.7^a 1.0 Continuity prior	0.22 ± 0.12 −0.02 ± 0.13 0.08 ± 0.13 0.24 ± 0.15
	Time resolution	3 bins 6 bins 10 bins 12 bins	−0.06 ± 0.21 −0.03 ± 0.12 −0.02 ± 0.13 −0.03 ± 0.14
	Initial mass function	Kroupa 2002 Chabrier 2003	−0.02 ± 0.13 −0.04 ± 0.21
	Noisy SED	3% uncertainties Perturbed w/3% unc.	−0.02 ± 0.13 0.15 ± 0.72
	Initial mass function	Kroupa 2002 Chabrier 2003	−0.39 ± 0.12 −0.48 ± 0.89
	Noisy SED	3% uncertainties Perturbed w/3% unc.	−0.39 ± 0.12 −0.47 ± 1.16

Notes. Median stellar mass offsets are shown for each SED model distribution, with the associated 1σ width, and the fiducial model choices are highlighted in bold. The choice of prior for the nonparametric model (Dirichlet+concentration parameter and continuity prior) has the largest effect on the stellar masses inferred from the nonparametric model. Using a perturbed input SED or a mismatched IMF slightly increases the median stellar mass offsets for each respective SFH model, though effectively each distribution is smeared out by the increased uncertainty in the inputs/model. *Concentration parameter*: The concentration parameter controls for the distribution of mass formation across bins. Low values ($\alpha < 1$) prefer to put all of the weight in one bin, while higher values more evenly distribute the weight across all bins. The continuity prior is an alternative nonparametric model prior available in PROSPECTOR and favors smoother SFHs, explicitly weighting against sharp changes in mass formation between adjacent time bins. *Time resolution*: The nonparametric model depends on a choice of time bins, both in number and in location. Based on Ocvirk et al. (2006), we set our time bins to be evenly spaced in logarithmic time and focus on the impact the time resolution (number of bins) has on the inferred stellar masses. *IMF and Noise*: For both the nonparametric and delayed- τ SFH model, we tested the impact of the assumed IMF and the input SED noise level.

^a Fiducial model choices are highlighted in bold.

Appendix B Variations in the SPS Model and Synthetic SEDs

To quantify the impact of the assumed SPS model, we changed the model IMF in PROSPECTOR to no longer match the IMF used in the POWDERDAY radiative transfer. Though the IMF represents just one uncertainty in an SPS model, the impact on the inferred stellar mass offsets is noticeable for both parametric and nonparametric SFH models, shown in the left panel of Figure 11, with statistics given in Table 3. Even so, the average galaxy has a more accurately inferred stellar mass from the nonparametric model than the delayed- τ model. An IMF change, to first order, is a multiplicative offset in stellar mass

with no change to the observed photometry. Switching from a Kroupa (2002) IMF to a Chabrier (2003) one will decrease the mass by 0.03 dex; fitting a Salpeter (1955) IMF to photometry generated from a Chabrier (2003) IMF would consistently overestimate masses by 0.2 dex.

Similarly, we simulated the effect of photometric noise on the inferred stellar mass. This was done by modulating the input photometry within the 3% uncertainties. Again, we see noticeable decreases in accuracy from both nonparametric and parametric SFH models when noise is added, though the nonparametric SFHs continue to outperform the traditional parametric models, with offsets extending to 1.5 dex.

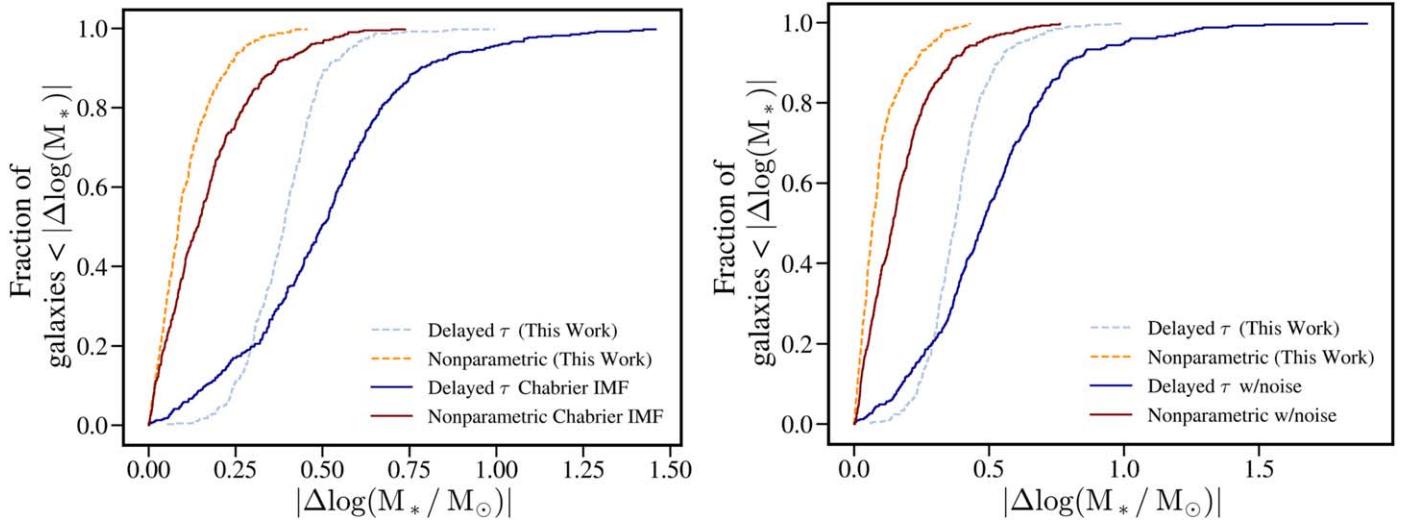


Figure 11. Distribution of stellar mass offsets for differences in the input and model IMF (left) and signal-to-noise of the input broadband SED (right). Summary statistics are given for each data set and model.

ORCID iDs

Sidney Lower <https://orcid.org/0000-0003-4422-8595>
 Desika Narayanan <https://orcid.org/0000-0002-7064-4309>
 Joel Leja <https://orcid.org/0000-0001-6755-1315>
 Benjamin D. Johnson <https://orcid.org/0000-0002-9280-7594>
 Charlie Conroy <https://orcid.org/0000-0002-1590-8551>
 Romeel Davé <https://orcid.org/0000-0003-2842-9434>

References

Acquaviva, V., Gawiser, E., & Guaita, L. 2011, *ApJ*, **737**, 47
 Acquaviva, V., Raichoor, A., & Gawiser, E. 2015, *ApJ*, **804**, 8
 Aufort, G., Ciesla, L., Pudlo, P., & Buta, V. 2020, *A&A*, **635**, A136
 Behroozi, P., Wechsler, R. H., Hearin, A. P., & Conroy, C. 2019, *MNRAS*, **488**, 3143
 Betancourt, M. J., & Girolami, M. 2013, arXiv:1312.0906
 Boquien, M., Burgarella, D., Roehlly, Y., et al. 2019, *A&A*, **622**, A103
 Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, *ApJ*, **686**, 1503
 Carnall, A. C., Leja, J., Johnson, B. D., et al. 2019, *ApJ*, **873**, 44
 Carnall, A. C., McLure, R. J., Dunlop, J. S., & Davé, R. 2018, *MNRAS*, **480**, 4379
 Caswell, T. A., Droettboom, M., Hunter, J., et al. 2018, Matplotlib/Matplotlib V3.0.1, Zenodo, doi:10.5281/zenodo.1482098
 Chabrier, G. 2003, *PASP*, **115**, 763
 Choi, J., Dotter, A., Conroy, C., et al. 2016, *ApJ*, **823**, 102
 Ciesla, L., Elbaz, D., & Fensch, J. 2017, *A&A*, **608**, A41
 Conroy, C. 2013, *ARA&A*, **51**, 393
 Conroy, C., & Gunn, J. E. 2010, *ApJ*, **712**, 833
 Conroy, C., Gunn, J. E., & White, M. 2009, *ApJ*, **699**, 486
 Curtis-Lake, E., Chevallard, J., & Charlot, S. 2020, arXiv:2001.08560
 da Cunha, E., Charlot, S., & Elbaz, D. 2008, *MNRAS*, **388**, 1595
 Davé, R., Anglés-Alcázar, D., Narayanan, D., et al. 2019, *MNRAS*, **486**, 2827
 Diemer, B., Sparre, M., Abramson, L. E., & Torrey, P. 2017, *ApJ*, **839**, 26
 Dotter, A. 2016, *ApJS*, **222**, 8
 Draine, B. T. 2003, *ARA&A*, **41**, 241
 Draine, B. T., & Li, A. 2007, *ApJ*, **657**, 810
 Driver, S. P., Hill, D. T., Kelvin, L. S., et al. 2011, *MNRAS*, **413**, 971
 Dudzevičiūtė, U., Smail, I., Swinbank, A. M., et al. 2020, *MNRAS*, **494**, 3828
 Faber, S. M. 1972, *A&A*, **20**, 361
 Foreman-Mackey, D., Sick, J., & Johnson, B. 2014, python-fsps: Python bindings to FSps, v0.1.1, Zenodo, doi:10.5281/zenodo.12157
 Gallagher, J. S. I., Hunter, D. A., & Tutukov, A. V. 1984, *ApJ*, **284**, 544
 Gallazzi, A., Charlot, S., Brinchmann, J., White, S. D. M., & Tremonti, C. A. 2005, *MNRAS*, **362**, 41
 Gladders, M. D., Oemler, A., Dressler, A., et al. 2013, *ApJ*, **770**, 64
 Guo, Y., Ferguson, H. C., Giavalisco, M., et al. 2013, *ApJS*, **207**, 24
 Hayward, C. C., & Smith, D. J. B. 2015, *MNRAS*, **446**, 1512
 Heavens, A., Panter, B., Jimenez, R., & Dunlop, J. 2004, *Natur*, **428**, 625
 Hunt, L. K., De Looze, I., Boquien, M., et al. 2019, *A&A*, **621**, A51
 Iyer, K., & Gawiser, E. 2017, *ApJ*, **838**, 127
 Iyer, K. G., Gawiser, E., Faber, S. M., et al. 2019, *ApJ*, **879**, 116
 Iyer, K. G., Tacchella, S., Genel, S., et al. 2020, *MNRAS*, **498**, 430
 Johnson, B., & Leja, J. 2017, BdJ/Prospector: Initial Release, vv0.1, Zenodo, doi:10.5281/zenodo.1116491
 Katsianis, A., Gonzalez, V., Barrientos, D., et al. 2020, *MNRAS*, **492**, 5592
 Kelson, D. D. 2014, arXiv:1406.5191
 Kriek, M., & Conroy, C. 2013, *ApJL*, **775**, L16
 Kriek, M., van Dokkum, P. G., Labbé, I., et al. 2009, *ApJ*, **700**, 221
 Kriek, M., van Dokkum, P. G., Labbé, I., et al. 2018, FAST: Fitting and Assessment of Synthetic Templates, Astrophysics Source Code Library, ascl:1803.008
 Kroupa, P. 2002, *Sci*, **295**, 82
 Lee, S.-K., Idzi, R., Ferguson, H. C., et al. 2009, *ApJS*, **184**, 100
 Leja, J., Carnall, A. C., Johnson, B. D., Conroy, C., & Speagle, J. S. 2019a, *ApJ*, **876**, 3
 Leja, J., Johnson, B. D., Conroy, C., et al. 2019b, *ApJ*, **877**, 140
 Leja, J., Johnson, B. D., Conroy, C., van Dokkum, P. G., & Byler, N. 2017, *ApJ*, **837**, 170
 Li, Q., Narayanan, D., & Davé, R. 2019, *MNRAS*, **490**, 1425
 Lovell, C. C., Acquaviva, V., Thomas, P. A., et al. 2019, *MNRAS*, **490**, 5503
 Madau, P., & Dickinson, M. 2014, *ARA&A*, **52**, 415
 Maraston, C., Daddi, E., Renzini, A., et al. 2006, *ApJ*, **652**, 85
 Michałowski, M. J., Dunlop, J. S., Cirasuolo, M., et al. 2012, *A&A*, **541**, A85
 Mobasher, B., Dahlen, T., Ferguson, H. C., et al. 2015, *ApJ*, **808**, 101
 Morishita, T., Abramson, L. E., Treu, T., et al. 2019, *ApJ*, **877**, 141
 Narayanan, D., Conroy, C., Davé, R., Johnson, B. D., & Popping, G. 2018a, *ApJ*, **869**, 70
 Narayanan, D., Davé, R., Johnson, B. D., et al. 2018b, *MNRAS*, **474**, 1718
 Narayanan, D., Turk, M., Feldmann, R., et al. 2015, *Natur*, **525**, 496
 Narayanan, D., Turk, M. J., Robitaille, T., et al. 2020, arXiv:2006.10757
 Ocvirk, P., Pichon, C., Lançon, A., & Thiébaut, E. 2006, *MNRAS*, **365**, 74
 Oemler, Augustus, J., Dressler, A., Gladders, M. G., et al. 2013, *ApJ*, **770**, 63
 Pacifici, C., Charlot, S., Blaizot, J., & Brinchmann, J. 2012, *MNRAS*, **421**, 2002
 Pacifici, C., Kassin, S. A., Weiner, B. J., et al. 2016, *ApJ*, **832**, 79
 Papovich, C., Dickinson, M., & Ferguson, H. C. 2001, *ApJ*, **559**, 620
 Papovich, C., Finkelstein, S. L., Ferguson, H. C., Lotz, J. M., & Giavalisco, M. 2011, *MNRAS*, **412**, 1123
 Paxton, B., Bildsten, L., Dotter, A., et al. 2011, *ApJS*, **192**, 3
 Pforr, J., Maraston, C., & Tonini, C. 2012, *MNRAS*, **422**, 3285
 Pforr, J., Maraston, C., & Tonini, C. 2013, *MNRAS*, **435**, 1389
 Reichardt, C., Jimenez, R., & Heavens, A. F. 2001, *MNRAS*, **327**, 849
 Robitaille, T. P. 2011, *A&A*, **536**, A79
 Robitaille, T. P., Churchwell, E., Benjamin, R. A., et al. 2012, *A&A*, **545**, A39
 Salim, S., & Narayanan, D. 2020, *ARA&A*, **58**, 529
 Salmon, B., Papovich, C., Finkelstein, S. L., et al. 2015, *ApJ*, **799**, 183

Salpeter, E. E. 1955, [ApJ](#), **121**, 161

Sandage, A. 1986, [A&A](#), **161**, 89

Santini, P., Ferguson, H. C., Fontana, A., et al. 2015, [ApJ](#), **801**, 97

Simha, V., Weinberg, D. H., Conroy, C., et al. 2014, arXiv:1404.0402

Skelton, R. E., Whitaker, K. E., Momcheva, I. G., et al. 2014, [ApJS](#), **214**, 24

Speagle, J. S. 2020, [MNRAS](#), **493**, 3132

Spinrad, H., & Taylor, B. J. 1971, [ApJS](#), **22**, 445

Thompson, R. 2014, pyGadgetReader: GADGET snapshot reader for python, Astrophysics Source Code Library, ascl:1411.001

Tinsley, B. M. 1968, [ApJ](#), **151**, 547

Tojeiro, R., Heavens, A. F., Jimenez, R., & Panter, B. 2007, [MNRAS](#), **381**, 1252

van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, [CSE](#), **13**, 22

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, [Nature Methods](#), **17**, 261

Walcher, C. J., Groves, B., Budavari, T., & Dale, D. 2011, [Ap&SS](#), **331**, 1

Weingartner, J. C., & Draine, B. T. 2001, [ApJ](#), **548**, 296

Wuyts, S., FörsterSchreiber, N. M., Lutz, D., et al. 2011, [ApJ](#), **738**, 106

Zhang, H.-X., Puzia, T. H., & Weisz, D. R. 2017, [ApJS](#), **233**, 13