



# Nonparametric Star Formation History Reconstruction with Gaussian Processes. I. Counting Major Episodes of Star Formation

Kartheik G. Iyer<sup>1</sup> , Eric Gawiser<sup>1,2</sup> , Sandra M. Faber<sup>3</sup>, Henry C. Ferguson<sup>4</sup> , Jeyhan Kartaltepe<sup>5</sup> , Anton M. Koekemoer<sup>4</sup> ,  
Camilla Pacifici<sup>4</sup> , and Rachel S. Somerville<sup>1,2</sup>

<sup>1</sup> Department of Physics and Astronomy, Rutgers, The State University of New Jersey, 136 Frelinghuysen Road, Piscataway, NJ 08854-8019 USA  
[kgil@physics.rutgers.edu](mailto:kgil@physics.rutgers.edu)

<sup>2</sup> Center for Computational Astrophysics, Flatiron Institute, 162 5th Ave, New York, NY 10010, USA

<sup>3</sup> University of California Observatories/Lick Observatory, University of California, Santa Cruz, CA 95064, USA

<sup>4</sup> Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA

<sup>5</sup> School of Physics and Astronomy, Rochester Institute of Technology, 84 Lomb Memorial Drive, Rochester, NY 14623, USA

Received 2019 January 9; revised 2019 May 2; accepted 2019 May 7; published 2019 July 12

## Abstract

The star formation histories (SFHs) of galaxies contain imprints of the physical processes responsible for regulating star formation during galaxy growth and quenching. We improve the Dense Basis SFH reconstruction method of Iyer & Gawiser, introducing a nonparametric description of the SFH based on the lookback times at which a galaxy assembles certain quantiles of its stellar mass. The method uses Gaussian processes to create smooth SFHs independent of any functional form, with a flexible number of parameters that is adjusted to extract the maximum amount of information from the SEDs being fit. Applying the method to reconstruct the SFHs of 48,791 galaxies with  $H < 25$  at  $0.5 < z < 3.0$  across the five Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey fields, we study the evolution of galaxies over time. We quantify the fraction of galaxies that show multiple major episodes of star formation, finding that the median time between two peaks of star formation is  $\sim 0.42_{-0.10}^{+0.15} t_{\text{univ}}$  Gyr, where  $t_{\text{univ}}$  is the age of the universe at a given redshift and remains roughly constant with stellar mass. Correlating SFHs with morphology allows us to compare the timescales on which the SFHs decline for different morphological classifications, ranging from  $0.60_{+1.54}^{-0.54}$  Gyr for galaxies with spiral arms to  $2.50_{+2.25}^{-1.50}$  Gyr for spheroids at  $0.5 < z < 1.0$  with  $10^{10} < M_* < 10^{10.5} M_{\odot}$ . The Gaussian process-based SFH description provides a general approach to reconstruct smooth, flexible, nonparametric SFH posteriors for galaxies that can be incorporated into Bayesian SED fitting codes to minimize the bias in estimating physical parameters due to SFH parameterization.

*Key words:* galaxies: evolution – galaxies: fundamental parameters – galaxies: star formation – galaxies: statistics – techniques: photometric

## 1. Introduction

Galaxies are massive, turbulent systems, shaped by physical processes that regulate star formation across many orders of magnitude in spatial and temporal scales; e.g., see White & Rees (1978), Searle et al. (1973), Hopkins et al. (2014), and Genel et al. (2019), as well as the reviews by Somerville & Davé (2015) and Naab & Ostriker (2017). Despite this apparent chaos, observations of ensembles of galaxies across cosmic time reveal several correlations, such as the Tully–Fisher relation (Tully & Fisher 1977), the SFR– $M_*$  correlation (Daddi et al. 2007; Elbaz et al. 2007; Noeske et al. 2007), the black hole mass–velocity dispersion correlation (Gebhardt et al. 2000), and the mass–metallicity correlation (Tremonti et al. 2004). Median trends constructed using these scaling relations indicate an equilibrium mode of galaxy growth through baryon cycling, punctuated by mergers and followed by eventual quiescence (Davé 2008; Peng et al. 2010; Tacchella et al. 2016; Behroozi et al. 2018). However, these trends can be equivalently recovered through stochastic evolution (Kelson 2014) or simple parametric models with minimal physics (Abramson et al. 2016). Given information about the present state of a galaxy, a lot can be said about its evolutionary history, because the gas inflow rate depends to first order on the halo formation rate, which is driven by gravity and evolves over approximately one Hubble time. It is important to

determine the extent to which a galaxy’s evolution is driven by evolving physical processes, such as baryon cycling and star formation suppression. These processes depend on the physical conditions of galaxies, such as their size, morphology, and stellar mass, as opposed to the stochastic processes governing halo and galaxy mergers and the creation and destruction of molecular clouds that regulate in situ star formation, which remains broadly invariant across many orders of magnitude, as inferred from the SFR– $M_*$  correlation extending from galaxy-wide scales (Whitaker et al. 2014; Kurczynski et al. 2016) down to kpc scales in resolved observations (Hsieh et al. 2017).

A key observable that correlates the present state of a galaxy with its evolutionary history is its star formation history (SFH)—a record of when a galaxy formed its stars. The SFHs of galaxies bear imprints from all the physical processes that shape galaxy growth by regulating star formation. This includes inflows and outflows of gas, mergers between galaxies, and feedback due to supernovae and Active Galactic Nuclei (AGNs), which leave imprints on the SFH on timescales ranging from  $< 1$  Myr to  $> 10$  Gyr (Somerville et al. 2008, 2015; Inutsuka et al. 2015; Sparre et al. 2015; Torrey et al. 2018; Behroozi et al. 2018; Matthee & Schaye 2019; Weinberger et al. 2018). Summary statistics of the SFH allow us to calculate traditionally estimated quantities like the stellar masses, star formation rates at the epoch of observation, and mass- and light-weighted ages of individual galaxies (Bell et al. 2007). Additional information

about the shape of galaxy SFHs allows us to infer whether the galaxy is actively forming stars, or if it formed most of its stars in the distant past (Kauffmann et al. 2003; Brinchmann et al. 2004). Nonparametric estimates of the median SFH for a sample of galaxies let us better understand the width and strength of major episodes of star formation (Heavens et al. 2000; Tojeiro et al. 2007; Pacifici et al. 2016; Iyer & Gawiser 2017), the origin and evolution of scaling relations (Iyer et al. 2018; Matthee & Schaye 2019; Torrey et al. 2018), mass functions (C. Pacifici 2019, in preparation), and the cosmic star formation rate density (Leja et al. 2019b).

With the advent of large surveys and the impending arrival of the next generation of surveys with *JWST*, developing more sophisticated nonparametric techniques of recovering galaxy SFHs will allow us to estimate galaxy properties out to higher redshifts. More importantly, higher-precision multiwavelength data allow us to go beyond traditional parametric forms and estimate finer features, such as SFHs with multiple strong episodes of star formation (Iyer & Gawiser 2017; Morishita et al. 2018), that can be caused by violent events like mergers or smoother events like stripping followed by inflow of pristine gas (Kelson 2014; Boogaard et al. 2018; Tacchella et al. 2018; Torrey et al. 2018). For example, Morishita et al. (2018) find that even old, quiescent galaxies sometimes show evidence for multiple episodes of star formation. Correlating these features of the SFH with other observables, such as the metallicity of the galaxy, evidence of recent mergers, and morphological and environmental conditions, will allow us to test different models that can help explain the diversity seen in SFHs at a particular epoch.

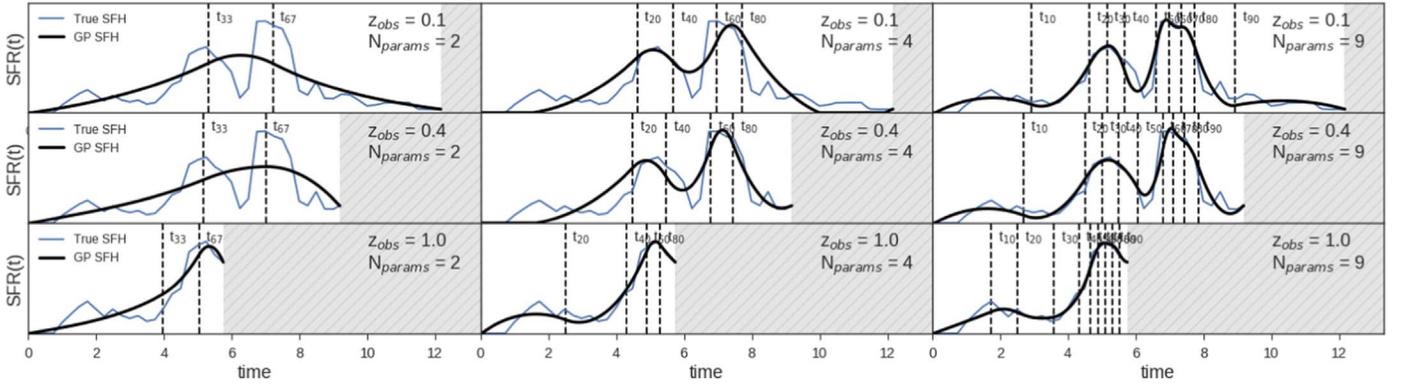
In the observational domain, the integrated light from distant galaxies contains a host of information about the physical processes that shape them during their formative phases (Tinsley 1968; Bruzual & Charlot 2003; Conroy et al. 2009; Conroy & Gunn 2010). Because stellar populations of different ages have distinct spectral characteristics, careful analysis of multiwavelength spectral energy distributions (SEDs) allows us, in principle, to disentangle these different populations (Heavens et al. 2000; Reichardt et al. 2001; Tojeiro et al. 2007; Dye 2008; Acquaviva et al. 2011; Pacifici et al. 2012, 2016; Smith & Hayward 2015; Domínguez Sánchez et al. 2016; Ciesla et al. 2017; Iyer & Gawiser 2017; Lee et al. 2018; Carnall et al. 2018; Leja et al. 2019a). SED fitting allows us to estimate the SFHs for a much larger population of galaxies, but requires much more sophisticated analysis to avoid biases. Assumptions of simple parametric forms for SFHs lead to biases, as shown in Iyer & Gawiser (2017), Ciesla et al. (2017), Lee et al. (2018), and Carnall et al. (2019). On the other hand, more complicated parametric forms, as well as methods that estimate the SFR in time bins, require us to estimate a much larger number of parameters. Therefore, we are in a situation where we would like to estimate as much information about the SFH as possible, encoded in the smallest possible number of estimated variables, while attempting to be as nonparametric as possible in order to avoid biases. As Ivezić et al. (2014) states, “nonparametric” does not imply an absence of parameters, but rather an independence from the model structure such as a functional form for the SFH. Ideally, the nature and number of parameters that specify the shape of a galaxy’s SFH are flexible and determined by the data. A nonparametric method in the context of SFHs should avoid assuming a simple parametric form (e.g., exponentially declining or lognormal SFHs), be

flexible enough to describe any function in SFH space, and infer the number of parameters from the quality of the data being analyzed.

In Iyer & Gawiser (2017), we introduced the Dense Basis method, which uses a basis of SFHs comprised of four different functional families and all of their combinations, determining the optimal number of SFH components using a statistical test. While this approach produces a basis of SFHs that is effectively dense in SED space and was shown to minimize the bias and scatter due to SFH parameterization, it still retains a minor dependence on the functional families under consideration. Additionally, it cannot be flexibly incorporated into an MCMC or nested sampling framework, and it becomes computationally expensive as we go to large numbers of SFH components—making it inefficient at extracting all the SFH information present in high-quality spectrophotometric data.

In this work, we introduce an improved version of the Dense Basis method that uses nonparametric SFHs constructed using Gaussian Process Regression (GPR) (also called kriging or Wigner–Kolmogorov prediction; see Chiles & Delfiner (2009), Wackernagel (2003), and Krivoruchko (2011)), using the lookback times at which a galaxy assembled certain quantiles of its overall stellar mass. Gaussian processes (Rasmussen & Williams 2006) extend the Gaussian probability distribution to the space of functions, allowing us to describe posteriors in SFH space without the need for parametric forms or bins in time. Historically, GPR developed as a method to interpolate between noisy data points. In our case, we have two sets of noisy data: the noisy SEDs themselves, and the noisy estimates of physical properties (such as the SFH) estimated using them. While Leistedt & Hogg (2017) use the former, in this paper we exclusively use the GPR to perform interpolation in SFH space in order to create smooth curves from sparse constraints on the SFH determined through SED fitting. Compared to traditional interpolation methods that require us to specify a smoothing scale or bandwidth, Gaussian processes allow us to specify a covariance function, also called the kernel, which defines how SFRs separated by a time  $\Delta t$  are related. In combination with constraints on the SFR at certain times from the observable SEDs, the Gaussian process routine is used to compute the likelihood of the SFR at every point in lookback time. While retaining all the advantages of the previous method, this has the additional advantages of being completely independent of any choice of functional form, scaling linearly with the number of SFH parameters, and providing a modular framework capable of being incorporated into any existing SED fitting routine. We establish the robustness of the method using semi-analytic models and hydrodynamical simulations for which the true SFHs are known, and apply it to galaxies at  $0.5 < z < 3.0$  across the five Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS) fields in order to study the evolution of galaxies around cosmic noon using their SFHs.

This paper is structured as follows. Section 2 describes the methodology, including the formalism used for constructing smooth, nonparametric SFHs and incorporating them into an SED-fitting framework. Section 3 contains a suite of validation tests for the methodology developed and applied in this work. In Section 4, we introduce the CANDELS data set that we use for the current analysis. In Section 5, we describe the SFHs reconstructed from the CANDELS sample, including the fraction of galaxies with multiple strong episodes of star formation, the evolution of this fraction with time, and implications for the



**Figure 1.** Applying Gaussian process regression to create smooth approximations of a single SFH from the Santa Cruz semi-analytic model (Somerville et al. 2008; Porter et al. 2014) at different redshifts of interest using the parameterization described in Section 2.1. The nine panels show how versatile the method is at capturing details of the SFH using a varying number of parameters, which can be tuned based on the quality of the observations, such as rest-frame wavelength coverage or S/N. In SFH space, this demonstrates the versatility of the method at describing arbitrary SFH shapes, with the accuracy of the approximation increasing with the number of parameters. An advantage of the Gaussian process formalism is that it makes it possible to describe an SFH with multiple episodes using as few as three parameters.

timescales of quenching followed by rejuvenation as well as for the morphological transformation of galaxies as they approach quiescence. Section 6 considers caveats and future directions for applying the Dense Basis method, and Section 7 concludes. Throughout this paper, magnitudes are in the AB system; we use a standard  $\Lambda$ CDM cosmology, with  $\Omega_m = 0.3$ ,  $\Omega_\Lambda = 0.7$ , and  $H_0 = 70 \text{ km Mpc}^{-1} \text{ s}^{-1}$ .

## 2. Methodology

### 2.1. Star Formation Histories

The main thrust of the Dense Basis method is to encode the maximum amount of information about the SFHs of galaxies using a minimal number of parameters. In this respect, the formalism used to describe SFHs in this work can be used as a module in existing sophisticated inference methods developed in public SED fitting codes like Bagpipes (Carnall et al. 2018), Beagle (Chevallard & Charlot 2016), CIGALE (Noll et al. 2009), and Prospector (Leja et al. 2017) to flexibly describe SFHs. As shown in our validation tests (Section 3), this description minimizes the bias in estimating SFHs at all lookback times, compared to existing parametric and nonparametric methods (Lee et al. 2010; Ciesla et al. 2017; Iyer & Gawiser 2017; Carnall et al. 2019; Leja et al. 2019a).

This subsection describes the methodology for creating SFHs using this formalism for a given number of parameters, and the following subsections handle the construction of the SED fitting machinery needed to infer the optimal number of SFH parameters corresponding to the amount of information available in individual galaxy SEDs.

We define an SFH by the tuple:  $(M_*, \text{SFR}, \{t_X\})$ , where  $M_*$  is the stellar mass, SFR is the star formation rate at the epoch of observation, and the set  $\{t_X\}$  contains the  $N$  “shape” parameters that describe the SFH. The shape parameters,  $\{t_X\}$ , are  $N$  lookback times at which the galaxy formed equally spaced quantiles of its total mass (Pacifi et al. 2016; Behroozi et al. 2018). For the first few values of  $N$ , we can write

$$\begin{aligned} N = 1 \quad P &= \{t_{50}\} \\ N = 2 \quad P &= \{t_{33}, t_{67}\} \\ N = 3 \quad P &= \{t_{25}, t_{50}, t_{75}\} \\ N = 4 \quad P &= \{t_{20}, t_{40}, t_{60}, t_{80}\}. \\ &\dots \end{aligned}$$

This can be seen graphically in Figure 1, where we show the  $\{t_X\}$  parameters with  $N = 2, 4, 9$  (vertical dashed black lines) for a mock SFH (blue line), along with our construction of the SFH using these parameters (black solid lines). As expected, the shape of the mock SFH is better captured as  $N$  increases, with multiple episodes captured using four parameters. In practice, we find that it is possible to recover multiple episodes of star formation with as few as three  $\{t_X\}$  parameters. Together with the stellar mass and SFR, this tuple describes a set of integral constraints that describe the shape and overall normalization of the SFH. For a galaxy at redshift  $z$ , when the universe was  $t_z$  Gyr old, the constraints are:

$$\begin{aligned} M_*(t_z) &= \int_{t=0}^{t_z} \text{SFH}(t) f_{\text{ret}}(t - t_z, Z) dt \\ &\times \left\{ \frac{iM_{*,\text{tot}}(t_z)}{N + 1} = \int_{t=0}^{t_{X,i}} \text{SFH}(t) dt \right\} \forall i \in N \end{aligned} \quad (1)$$

$$M_* = \frac{\int_{t=0}^{t_z} \text{SFH}(t) f_{\text{ret}}(t - t_z, Z) dt}{\int_{t=0}^{t_z} \text{SFH}(t) dt} M_{*,\text{tot}}. \quad (2)$$

The second line is a set of  $N$  equations, one for each parameter in the set  $\{t_X\}$  that requires that the galaxy form “ $x$ ” fraction of its total mass by time  $t_X$ .

This description of a galaxy’s SFH already offers several advantages over methods found in the literature, a few of which are summarized below.

1. Not being restricted to a particular functional form minimizes bias due to SFH parameterization (Iyer et al. 2018).
2. Describing an SFH using  $\{t_X\}$  reduces the discrepancy in S/N per parameter, in comparison to methods that determine the SFR in bins of lookback time: for example,  $t_{20}$  might be less well-constrained compared to  $t_{80}$ , but the overall signal depends on the shape of the SFH. For example, the parameterization will not try to constrain the SFR in the first year after the Big Bang unless enough stars were formed that early to provide a discernible signal in the SED. This can be compared to methods that adaptively choose time bins, e.g., VESPA (Tojeiro et al. 2007).

3. This provides a novel framework for compressing the amount of information present in an SFH to a small set of numbers (given a way to reconstruct an SFH from a tuple), and hence for comparing SFHs across different simulations and observations on the same footing.
4. The distribution of different  $\{t_X\}$  among galaxies at a given epoch within a simulation can be extremely useful in defining and checking the physical assumptions of the SFH priors during SED fitting.

Reconstructing an SFH from the tuple  $(M_*, \text{SFR}, \{t_X\})$  can be done in multiple ways, but we seek to minimize the information lost in doing so, while remaining computationally inexpensive. As with all compression methods, we approach the true SFH as the number of parameters  $N$  in the set  $\{t_X\} \rightarrow \infty$ , but we would like to minimize the loss even with a relatively small number of parameters.

Reconstructing an SFH requires quantifying the integral constraints as points on a fractional mass  $(M_{*,\text{tot}}(t))$ —cosmic time ( $t$ ) plane and drawing a piecewise smooth curve passing through these points. Rescaling the mass axis and differentiating this cumulative curve would then yield the SFH as a star formation rate at each lookback time. The simplest approximation would be to connect each point such that the resulting SFH is piecewise linear, but while this provides an acceptable solution, it is not very physical—in the sense that taking the derivative yields an SFR with jump discontinuities. While generalizing to polynomials for the interpolation causes problems with the derivative going negative in parts of the SFH, methods such as tensioned cubic splines and piecewise-cubic Hermite polynomial interpolation (De Boor et al. 1978; Butt & Brodlye 1993) provide more sophisticated solutions to this problem.

In this work, we use GPR (Rasmussen & Williams 2006; Leistedt & Hogg 2017) implemented through the `george` python package (Ambikasaran et al. 2015; Foreman-Mackey 2015) to create a smooth SFH along with uncertainties following a physically motivated covariance function (kernel) for a given SFH tuple. The Gaussian process framework uses a set of constraints, given by the set of Equation (1), along with a covariance function (or kernel) to estimate the probability of  $\text{SFH}(t)$  at a given time  $t$ . We use a Matern32 kernel (Seeger 2004) in the present application, given by:

$$k(r^2) = (1 + \sqrt{3r^2}) \exp(-\sqrt{3r^2}) \quad (3)$$

where the covariance function contains a scale length hyperparameter ( $r^2 \propto (\Delta t)^2$ ) that sets how much the  $\text{SFR}(t)$  can vary from the  $\text{SFR}(t+\Delta t)$  separated by a time interval  $\Delta t$ . The hyperparameter in the kernels essentially encodes the amount of stochasticity in the SFHs, and is described further in Section 3.5. In practice, this can be thought of as setting the tension in a string that passes through all the constraints in cumulative mass–cosmic time space, and therefore affects the shape and amount of ringing that can happen between two quantiles. When using a spline to reconstruct the SFH, this ringing can sometimes cause the SFR to be negative. Our choice of physically motivated kernel minimizes this behavior and limits it to extreme SFH tuples (e.g.,  $t_{25} = 0.1$  Myr,  $t_{50} = 10$  Gyr,  $t_{75} = 10.1$  Gyr, i.e., the SFH formed the first and third quarters of its mass in bursts of 0.1 Myr, separated by a long period of relative quiescence), which account for  $<3\%$

across our basis. In these cases, we set the relevant portion (i.e.,  $\text{SFR}(t) < 0$ ) to 0 and check that the overall error in stellar mass due to this is  $<5\%$ . By doing so, we ensure that  $\text{SFR} \geq 0$  at all times across our entire basis. Setting the SFR to 0 in cases like this is physically motivated—stretches with  $\text{SFR} = 0$  might be found in nature as well, as seen in SFH reconstructions of local dwarfs (Weisz et al. 2011) or in high-resolution simulations (Wright et al. 2019), for galaxies with very bursty SFHs. Examples of approximating SFHs using this formalism are shown in Figure 1. The advantage of this method is that both parameters  $(M_*, \text{SFR}, \{t_X\})$  and uncertainties on these parameters can be passed as arguments while constructing the SFH posterior. This is tested using SFHs from the Santa Cruz semi-analytic models (SAM) (Somerville et al. 2008; Porter et al. 2014) and the MUFASA Davé et al. (2016) simulation to minimize loss during the reconstruction of SFHs; in both cases, the error in approximating the true SFH decreases as we increase the number of parameters. The Santa Cruz semi-analytic model uses dark matter halo merger trees from the Bolshoi-Planck simulation (Klypin et al. 2011) in combination with analytic recipes for the radiative cooling of gas, collapse of cold gas to form rotationally supported disks, star formation, feedback, and chemical enrichment. The MUFASA hydrodynamical simulation uses the GIZMO (Hopkins 2015) meshless code, including prescriptions for  $\text{H}_2$ -based star formation, chemical evolution, and kinetic outflows following scalings from the FIRE (Hopkins et al. 2014) simulation, as well as an evolving halo-mass-based quenching mechanism. Both simulations reproduce a wide range of present-day observables, including  $z \sim 0$  mass functions and scaling relations, with further details in Somerville et al. (2008), Porter et al. (2014), and Davé et al. (2016).

## 2.2. SED Fitting

While we are primarily interested in the SFH, determining it from the galaxy’s SED requires us to account for several other factors, such as the chemical enrichment, stellar initial mass function (IMF), dust attenuation model, and absorption by the IGM. We then formulate the SFH estimation as an inference problem given by:

$$P(\text{SFH}, A_V, z \dots | F_{\nu,j}^{\text{obs}}) = \frac{P(F_{\nu,j}^{\text{obs}} | \text{SFH}, A_V, Z, z \dots) P(\text{SFH}, A_V, Z, z \dots)}{P(F_{\nu,j}^{\text{obs}})} \quad (4)$$

The term  $P(F_{\nu,j}^{\text{obs}} | \text{SFH}, A_V, Z, z)$  is the likelihood, given by  $\mathcal{L} \propto \exp(-\chi^2/2)$ , where

$$\chi^2 = \sum_{j=1}^{N_{\text{filters}}} \left( \frac{F_{\nu,j}^{\text{obs}} - F_{\nu,j}^{\text{model}}(\text{SFH}, A_V, Z, z)}{\sigma_j} \right)^2 \quad (5)$$

The term  $P(\text{SFH}, A_V, Z, z \dots)$  denotes the prior distribution for the model. If we assume uncorrelated priors for all the parameters, this can be written as  $P(\text{SFH})P(A_V)P(Z)P(z) \dots$ . Here,  $F_{\nu,j}^{\text{obs}}$  is the observed photometry being fit, in the  $j$ th photometric filter. SFH denotes the SFH tuple  $(M_*, \text{SFR}, \{t_X\})$ ,  $A_V$  is the dust model,  $Z$  the stellar metallicity, and  $z$  is the redshift. In addition to this, we need to consider the stellar

population synthesis models, stellar IMF, absorption by the intergalactic medium, and a self-consistent implementation of nebular emission lines using CLOUDY through the Flexible Stellar Population Synthesis (FSPS) code. We adopt the Calzetti attenuation law for the dust attenuation (Calzetti 2001), with one free parameter (because the Calzetti law couples the birth-cloud attenuation to that from older stars), one for stellar metallicity, and a Chabrier IMF (Chabrier 2003) with no free parameters. We define our SFH parameterization in Section 2.1 with  $N+2$  parameters, where  $N$  is the number of SFH percentiles, given by the set  $\{t_x\}$ . We use FSPS (Conroy et al. 2009; Conroy & Gunn 2010) to generate spectra corresponding to the Basel stellar tracks and the Padova isochrones. With  $N$  SFH quantiles, the model then has  $N+5$  ( $M_*$ , SFR,  $A_V$ ,  $Z$ ,  $z$ ) free parameters that need to be determined from the data. We construct the method in such a way that  $N$  itself is a variable that is tuned to extract the maximum amount of information present in a galaxy’s SED. It is important to keep in mind that these modeling choices can impose an implicit prior on our SED fits. The effects of testing our model assumptions and priors are further explored in Section 3, where we find that our models and priors are suitably robust, considering the S/N and wavelength coverage of our data set.

We implement the posterior computation numerically, using a brute-force Bayesian approach similar to Pacifici et al. (2012, 2016) and Da Cunha et al. (2008), with a large pregrid of model SEDs constructed through random draws from the prior distributions corresponding to each free parameter in Equation (5). To ensure that the pregrid samples the priors finely enough and is effectively dense in SED space, we perform fits to a sample of 1000 galaxies while varying the size of our pregrid. Using this, we estimate the optimal size of the pregrid as the point where the improvement in median  $\chi^2$  for the sample as a function of pregrid size is negligible, leading to a pregrid with  $\sim 900,000$  SEDs.

To construct the pregrid, we draw random values from our prior distributions for stellar metallicity, dust attenuation, and SFH parameters ( $M_*$ , SFR,  $\{t_x\}$ ). For metallicity, we adopt a flat prior on  $\log Z/Z_\odot$  (Pacifici et al. 2016; Carnall et al. 2018; Leja et al. 2019a), an exponential prior on dust attenuation (Iyer & Gawiser 2017), and a Dirichlet prior for the lookback times  $\{t_x\}$  that specify the shape of the SFH. The Dirichlet prior is a generalization of the Beta distribution to  $N$  variables, such that a random draw yields  $N$  random numbers  $x_i$  that satisfy  $\sum_i x_i = 1$ . More details about this prior can be found in Leja et al. (2017, 2019a). In practice, for a galaxy SFH at redshift  $z$ , we generate the set  $\{t_x\}$  by performing a random draw multiplied by the age of the universe at that redshift, giving the set of lookback times at which a galaxy formed various quantiles of its stellar mass. The Dirichlet prior has a single tunable parameter  $\alpha$  that specifies how correlated the values are. In our case, values of this parameter  $\alpha < 1$  result in values that can be arbitrarily close, leading to extremely spiky SFHs because galaxies have to assemble a significant fraction of their mass in a very short period of time, while  $\alpha > 1$  leads to smoother SFHs with more evenly spaced values that nevertheless have considerable diversity. In practice, we use a value of  $\alpha = 5$ , which leads to a distribution of parameters that is similar to what we find in SAM and MUFASA. This can be seen in Figure 2. For each galaxy in our pregrid, we turn the

corresponding ( $M_*$ , SFR,  $\{t_x\}$ ) tuple into an SFH via our Gaussian process routine, assuming fiducial 1% uncertainties. Using this SFH, in combination with the values for metallicity, dust, and redshift, we generate a spectrum through the FSPS stellar population synthesis routine. Multiplying this spectrum with the appropriate filter transmission curves gives the SED corresponding to each tuple in the pregrid.

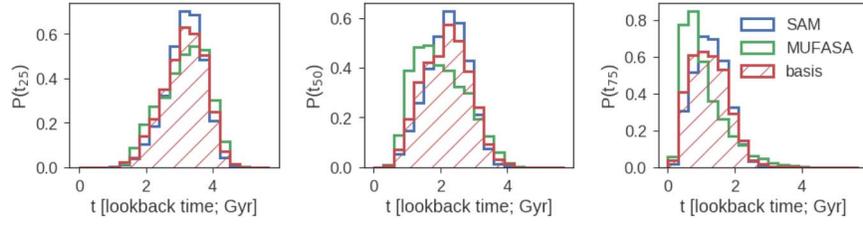
SFH uncertainties are computed after the fit is performed, using the posterior for each parameter describing the tuple ( $M_*$ , SFR,  $\{t_x\}$ ). For each galaxy, 100 self-consistent random draws are performed from the posterior with the covariances between parameters taken into account, and corresponding SFH realizations are constructed using the Gaussian process routine. For a given realization, if the set of parameters already exists in the pregrid, the corresponding precomputed SFH is used to decrease the computational cost. Figure 3 shows 20 draws from the SFH posterior for that galaxy using this approach, depicted as thin black lines in the SFH inset panel. The 68% confidence intervals are then constructed by taking the 16th to 84th percentiles of the SFR distribution at each point in look-back time.

### 2.3. Nonparametric SFH Reconstruction

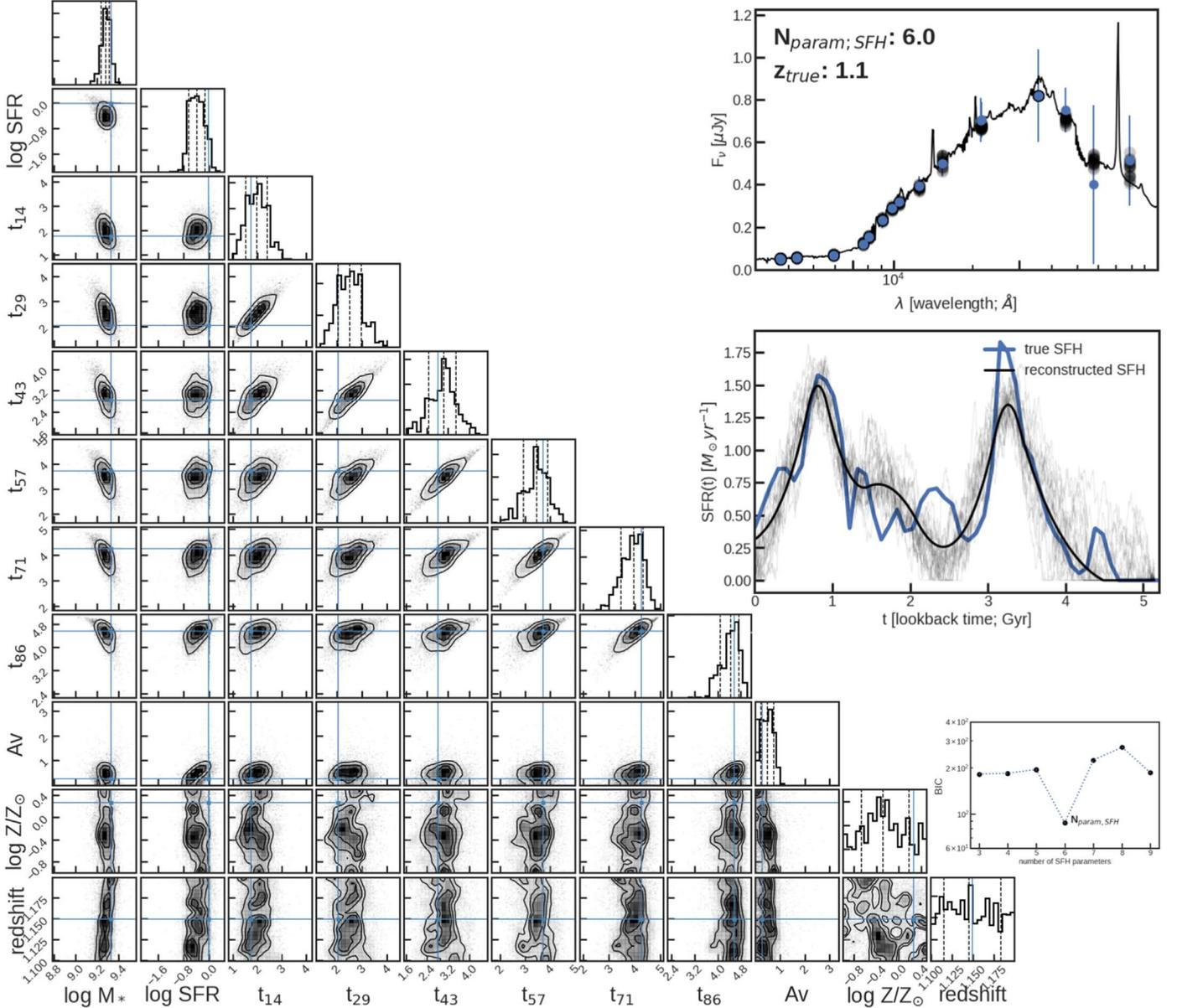
The Gaussian process–based SFH (GP-SFH) formalism allows us to gain independence from having to make a choice of functional form for the shape of the SFH, without having to bin SFHs in lookback time. This results in smooth, effectively nonparametric SFHs that minimize the bias and scatter in SFH reconstruction, as seen in Section 3. However, because we would like the method to be truly nonparametric, we require that the number of  $\{t_x\}$  parameters be optimized for the amount of information present in a given noisy SED. To implement this in practice, we generate pregrids for SFHs with the set  $\{t_x\}$  ranging from three to nine parameters, because it is difficult to specify the shape of complex SFHs with less than three parameters and it is impractical to recover more than nine from broadband SEDs. We then fit each observed SED with all seven pregrids, and obtain the most appropriate number of parameters using an appropriate model selection criterion. Ideally, the Bayesian evidence would be used for this model selection step (Liddle 2007). However, in practice, the numerical computation of the evidence is expensive due to the need for a nested sampler, and cannot be completed for the number of galaxies typically found in large photometric surveys. Having tested the properties of the likelihood surfaces using this method, however, we find that while they may be multimodal, they generally do not contain pathological features that necessitate this numerically expensive procedure. In light of this, we perform our model selection using an approximation of the evidence, given by the Bayesian Information Criterion (BIC) (Schwarz et al. 1978; Liddle 2007), defined as

$$\text{BIC} = k \ln(n) - 2 \ln \mathcal{L}_{\max} \quad (6)$$

where  $n$  is the number of photometric data points,  $k$  is the number of parameters in the SFH tuple, and  $\mathcal{L}$  is the maximized value of the likelihood function given by Equation (5). The latter term in this equation is a measure of how well the model describes the data, and the former term is a penalty for an increased number of parameters. The results from the two model selection criteria are equivalent when all the parameters



**Figure 2.** The prior distributions of SFHs in  $t_{25}$ ,  $t_{50}$ ,  $t_{75}$  space for SFHs of galaxies at  $z \sim 1$  from the Santa Cruz semi-analytic model and the MUFASA hydrodynamical simulation, in addition to the basis assembled using the Dirichlet prior we adopt.



**Figure 3.** An illustrative example of the SED fitting method, applied to mock noisy photometry for a galaxy from the Somerville et al. (2008) and Porter et al. (2014) semi-analytic model with more than one major episode of star formation, with the photometry simulated using the GOODS-S filter set. The top right panel shows the simulated photometry (blue error bars) and the spectrum (solid black line) corresponding to the median parameter values. The corresponding reconstructed SFH (solid black line) is shown in the middle right panel below it, with the true SFH (solid blue line) from the SAM shown for comparison. The thin black lines in both panels show draws from the posterior distribution for comparison. This particular SED was determined to contain enough information to estimate six correlated SFH parameters using the BIC model selection criterion, as shown in the inset panel (bottom right). The corner plot (left) shows the posteriors for each parameter using our brute-force Bayesian approach, with the blue lines representing the true values used to generate the mock noisy photometry. Although redshift is formally a free parameter, we fit galaxies at the  $(z_{\text{best}} \pm 0.05)$  from D. Kodra et al. (2019, in preparation), finding that the redshift posterior is effectively flat within this small dynamic range, as expected. Dashed black lines for each histogram show the median and 16–84th percentile range.

are independent (Kass & Wasserman 1995; Szydlowski et al. 2015). However, if parameters are strongly covariant, the volume will reflect this (because the intrinsic dimensionality is lower than the number of dimensions describing the likelihood) while the BIC does not. This can lead to the optimal number of parameters estimated with the BIC  $\leq$  the optimal number of parameters estimated using the evidence. Our validation tests show that using the BIC does not cause any significant issues in fitting an ensemble of SEDs to recover their SFHs. By finding the minima in the BIC, we find the number of SFH parameters that can be robustly extracted from the SED being fit. This leads to a truly nonparametric description of the SFH, based on the amount of information about the different stellar populations encoded in the galaxy’s SED. An example of this for a single galaxy is shown in the bottom right plot of Figure 2. Iyer & Gawiser (2017) used a similar nonparametric estimate of the number of SFH components using a F-test, but it was computationally expensive to implement as the number of SFH components grew. Tojeiro et al. (2007) and Dye (2008) also used nonparametric methods to estimate the optimal number of time bins during SFH reconstruction, although this is prone to bin edge effects, as described in Leja et al. (2019a). Our Gaussian process SFH reconstruction method can be summarized as follows:

1. We introduce a formalism based on Bayesian statistics to describe the SFH, which we will express not as a functional form but rather as N-tuples, plus  $M_*$  and SFR at the time of observation. To generate a curve  $\text{SFR}(t)$  from the tuples, the GPR needs a covariance function or kernel, which determines how SFR distributions at  $t$  and  $t + \Delta t$  are related, i.e., how smooth the SFHs are.
2. The Bayesian likelihood is computed for each of a sufficiently large number of model SFHs realized using GPR, having known values of the n-tuples,  $M_*$ , and SFR-today (the pregrid) drawn from our chosen prior. By estimating the likelihood of each basis N-tuple, we can compute posteriors for the N-tuples (and  $M_*$  and SFR-today) along with their credible intervals and covariances.
3. The GPR is now used to interpolate between these (noisy) estimates of the n-tuples to obtain draws from the posterior distribution. Each of these draws is a model SFH. The distribution of these draws at each point in time gives the uncertainty in the SFH at that time.
4. This procedure is repeated as we vary the number of parameters in the N-tuples, using a Bayesian model selection criterion to evaluate the optimal number of parameters.

#### 2.4. Quantifying the Number of Major Episodes of Star Formation in an SFH

Unlike simple SFH parameterizations commonly used in SED fitting, the Gaussian process-based SFHs can have multiple maxima, even with four or fewer  $t_x$  parameters, as seen in Figure 1. Hence, it is possible to analyze them using a peak-finding algorithm to quantify the number of major episodes of star formation in a galaxy’s past. For this particular analysis, we use the best-fit SFH for each galaxy, because the median SFH for each galaxy is biased toward smooth SFHs. This is an effect due to our choice of kernel, which prefers

smooth solutions when the  $\{t_x\}$  parameters are uncertain, as seen in the left column of Figure 1. With tighter constraints on  $\{t_x\}$  from higher-S/N data, this problem is alleviated. As a result, while the number of major episodes ( $N_{\text{ep}}$ ) estimated using this method for individual galaxies is susceptible to noise, the overall distribution is seen to be recovered without any significant bias, as shown in Section 3.

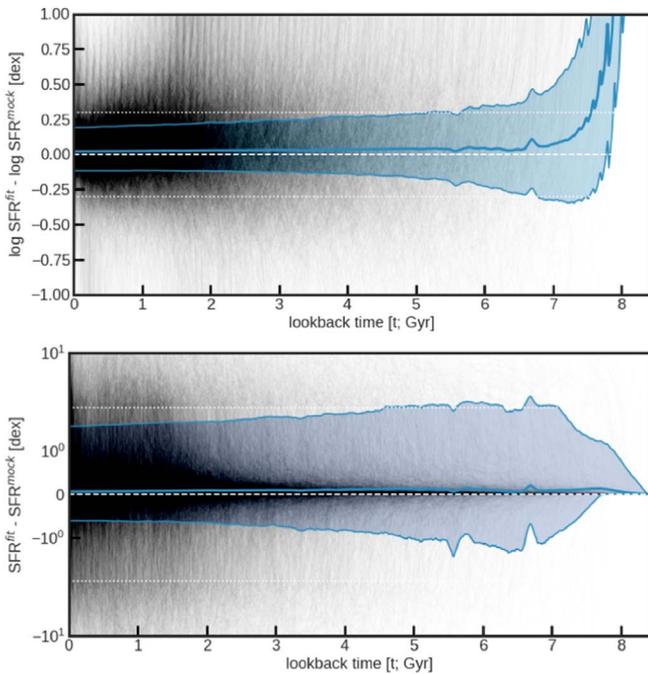
For each SFH, we quantify the number of episodes as follows: we first find the number of peaks in an SFH as the set of points that satisfy  $d\text{SFH}/dt = 0$  and  $d^2\text{SFH}/dt^2 < 0$ . To separate multiple peaks within an overall episode of star formation from different episodes, we impose a peak-prominence criterion by requiring that

$$(\log \text{SFR}_{\text{peak}} - \log \text{SFR}_{\text{min,local}}) > 1.5 + \frac{1.5}{4} \log \frac{M_*}{10^8 M_\odot} \quad (7)$$

where  $\text{SFR}_{\text{min,local}}$  is the minima between two peaks in the SFH. This condition is shown (in Section 3) to minimize the type-1 (overestimating  $N_{\text{ep}}$ ) and type-2 (underestimating  $N_{\text{ep}}$ ) errors in our validation sample. It arises because the sensitivity to star formation drops approximately logarithmically with time (Ocvirk et al. 2006). Because more massive galaxies tend to have older stellar populations, we found that a mass-independent peak-prominence criterion caused a mass-dependent bias in our estimates of the number of episodes. We correct for this effect by requiring a more (less) stringent dip in the SFH for a massive (low-mass) galaxy, and while this does not improve the result for every galaxy, it accurately recovers the distribution, which is the quantity that we are interested in.

### 3. Validation Tests

At each state of the method development, we perform validation using an ensemble of galaxies from two cosmological simulations: the Santa Cruz semi-analytic models (Somerville et al. 2008; Porter et al. 2014) and the MUFASA hydrodynamical simulation (Davé et al. 2016). Using two different simulations allows us to alleviate the potential biases induced by testing our method against a single simulation. For the tests described in this section, we created a mock catalog with 10,000 galaxies, sampling randomly from both simulations at  $0.5 < z < 3.0$ , matching our analysis sample. For each galaxy in our mock catalog, we draw a random redshift  $z_{\text{mock}}$  between 0.5 and 3. We then create synthetic spectra using FSPS with the corresponding SFH and mass-weighted metallicity, with a Calzetti dust attenuation  $A_V$  sampled from an exponential distribution as in Iyer & Gawiser (2017). We then multiply these spectra by the appropriate filter transmission curves corresponding to one of the five CANDELS fields, and perturb the photometry in each band by adding realistic noise derived from the median photometric uncertainties for the CANDELS catalog in the redshift range  $[z_{\text{mock}} - 0.1, z_{\text{mock}} + 0.1]$ . Figure 2 shows an example following this procedure, using a galaxy with an SFH that is not well-approximated by a simple parametric form currently used in the literature, but is recovered well with the GP-SFH approach. Using our mock catalog, we then perform a series of validation experiments. In Section 3.1, we consider the robustness of the reconstructed SFHs. In Section 3.2, we quantify the bias and scatter in estimating stellar masses, star formation rates, dust attenuation, and stellar metallicities. In Section 3.3, we consider the robustness of the uncertainties on



**Figure 4.** Comparing the ensemble of true SFHs from cosmological simulations to SFHs reconstructed from SEDs using the Gaussian Process Dense Basis method. Thin black lines show the difference in log SFR for individual galaxies, with the pointwise median in time shown as a solid blue line and the shaded blue region denoting the 16th to 84th percentiles. The left panel shows the difference in terms of  $\Delta \log \text{SFR}$ , while the right panel shows the difference in terms of  $\Delta \text{SFR}$  at each point in lookback time.

the reconstructed SFHs. In Section 3.4, we test the recovery of the number of major episodes of star formation for our mock sample, find a mass-dependent bias, and correct for it. In Section 3.5, we test our choice of Gaussian process kernel. In Section 3.6, we test the information loss and how well the method compares to existing nonparametric SFH descriptions.

### 3.1. SFH Robustness

To quantify possible biases in our SFH reconstruction, we plot the difference between the true and recovered SFH for our sample of validation galaxies in Figure 4. The two panels show the linear and logarithmic differences between the true and the recovered SFHs, which respectively represent the difference and ratio of the true versus predicted  $\text{SFR}(t)$ . This difference is smaller than 0.3 dex for most of the lookback time. While the reconstructed SFHs cannot recover every short episode of star formation, as seen in the errors for individual galaxies (thin black lines), the overall SFH for the ensemble of galaxies is unbiased out to nearly  $\sim 8$  Gyr. In the  $\Delta \log \text{SFR}$  error plot, the bias blows up as we approach the Big Bang. This is because the SFHs in the mocks can abruptly fall to 0 close to the Big Bang while the Gaussian process SFHs smoothly decline to 0 at  $t = 0$ , leading to a one-sided error. However, the arithmetic  $\Delta \text{SFR}$  plot shows that this is, in fact, a very small difference, exaggerated by the fact that  $\log \text{SFR}_{\text{true}} \rightarrow -\infty$  as  $\text{SFR}_{\text{true}} \rightarrow 0$  close to the Big Bang.

### 3.2. Parameter Robustness

In addition to the SFHs, Figure 5 shows the results of estimating traditional SED fit parameters—stellar mass, star formation rate averaged over the last 100 Myr, dust attenuation,

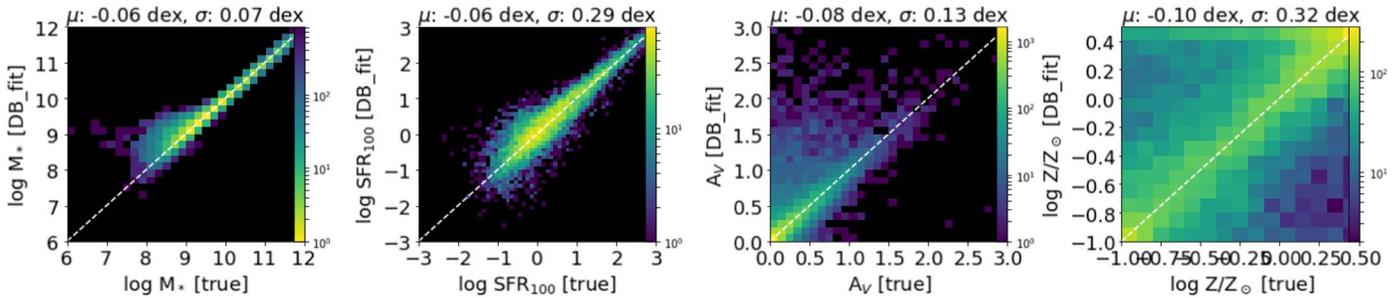
and stellar metallicity for the validation catalog. The stellar masses are constrained better than the traditional scatter around the mean of  $\sim 0.14$  dex (Mobasher et al. 2015). Star formation rates have a larger scatter (0.29 dex) due to degeneracies with dust attenuation (0.13 dex) and metallicity (0.32 dex). At low stellar masses, we are prone to overestimating the overall mass due to our choices of SFH prior dominating the fit for low-S/N SEDs. Restricting the fits to galaxies with  $H < 25$  is found to largely exclude these low-S/N objects, leading to more robust estimates of these parameters.

### 3.3. SFH Uncertainties

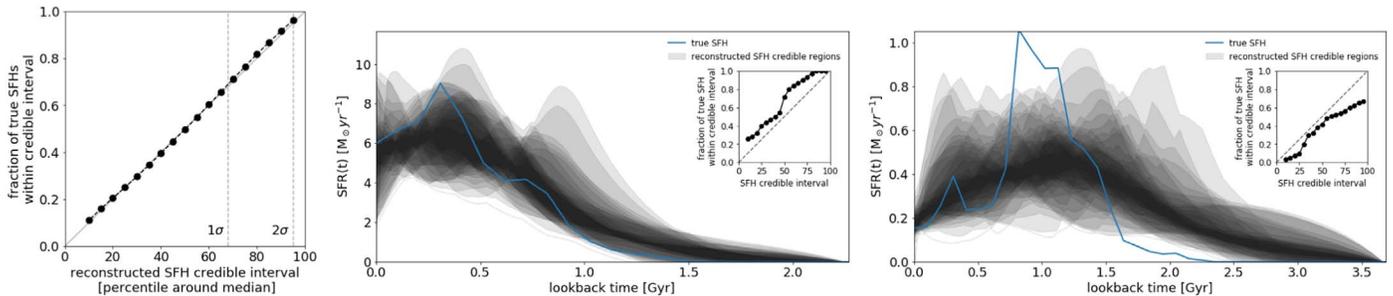
As described in Section 2, SFH uncertainties are computed at each point in lookback time for various percentiles, ranging from the 10th (55–65th percentile bounds) to the 95th (2.5–97.5th) percentile credible intervals of SFHs constructed from 100 draws from the SFH posterior ( $M_*$ , SFR,  $\{t_i\}$ ) using the Gaussian process routine. To check whether these represent true 68% uncertainties, we consider the uncertainties estimated for each of the 10,000 galaxies in the mock catalog that we fit and compute the fraction of the true SFH that lies within the uncertainties for each galaxy. This is similar to the concept of a “coverage probability” (Levasseur et al. 2017) used to evaluate the accuracy and precision of uncertainty estimates. The distribution of this fraction is given in the left panel of Figure 6, which shows the fraction of the truth that lies, on average, within a given credible interval. For an ideal method, we expect this to lie along the 1:1 line; our method is extremely close to this across the entire evaluated range. This indicates that our uncertainties are robust, expanding on the tests in Iyer & Gawiser (2017), which solely checked the 68% uncertainties. Figure 6 shows two examples: one where the uncertainties are representative of the truth, and another where both the reconstruction and the uncertainties miss a relatively short ( $< 0.5$  Gyr) episode of star formation that follows a relatively quiescent period, due to the smoother reconstructed SFH producing a comparable SED given the photometric noise.

### 3.4. Number of Episode Estimation

For each galaxy in the mock sample, we compute the number of major episodes of star formation ( $N_{\text{ep}}$ ) using the peak-finding algorithm described in Section 2. For the true SFHs, we compute the number of episodes requiring a dip of  $\log \text{SFR}_{\text{peak}} - \log \text{SFR}_{\text{min,local}} \geq 0.3$  dex, to model the scenario of a galaxy on the star-forming sequence dropping below the sequence before rejoining it. The distribution of galaxies with multiple strong episodes of star formation for the mock sample is then given by the thick dashed black lines in Figure 7. We then estimate the number of episodes for the reconstructed SFHs with the same criterion, and find a flatter trend than observed in the mocks. This is a combination of two effects. (1) The S/N contributed to the overall SED by a stellar population of a certain age decreases roughly logarithmically with lookback time (Ocvirk et al. 2006), leading to poorer constraints on intermediate and older stellar populations. (2) There is a mass-dependent correlation toward older stellar populations as galaxies grow more massive. As a result of this, galaxies that formed most of their mass prior to an extended period of quiescence can be erroneously classified as having multiple episodes, because when the SFR at the time of observation is low, it is easier to have a fluctuation of  $\geq 0.3$  dex



**Figure 5.** Results of estimating stellar masses, star formation rates, dust attenuation, and metallicities for the ensemble of mock galaxies described in Section 3. Each parameter is shown as a log-scale heatmap, with the adjacent color bar detailing the number of galaxies in a given bin. Bins with no galaxies are shown in black. The dashed white line shows the 1:1 relation, and each plot title contains the overall bias and scatter around the mean for the sample.



**Figure 6.** Validation performed to check the robustness of SFH uncertainties estimated from using the posterior from SED fitting. For our sample of 10,000 mock galaxies, we compute the fraction of the true SFHs that lie within the 10%–95% credible intervals, shown in the left panel. For robust uncertainties, we expect a 1:1 correspondence between the fraction of truth that lies within a credible interval and the interval itself. The middle and right panels show two examples of this computation for two galaxies in our sample: one where the majority of the true SFH lies within the uncertainties (middle), and another where a sharp peak is not recovered by the SFH reconstruction, due to photometric noise (right).

with a small variation in SFR. We find that introducing a stellar mass-dependent threshold to determine the number of peaks given by Equation (7) accounts for these mass-dependent effects on the distribution, allowing us to better reproduce the true trends in  $N_{\text{ep}}$  with stellar mass and redshift, as seen in Figure 7. The thin lines also show the number of overestimates and underestimates, such that  $N_{\text{ep,true}} = N_{\text{ep,rec}} - \text{underestimates} + \text{overestimates}$ . Because we are using the best-fit SFHs for this computation, we find that our results are quite sensitive to noise for individual galaxies. However, the distributions are estimated robustly, across a range of stellar masses and redshifts. Additionally, the differing behavior of the observational sample from our mock catalog is indicative of the fact that our mass-dependent threshold does not impose an artificial trend on the distribution. Reconciling the differences between the two distributions would be an interesting topic for further study.

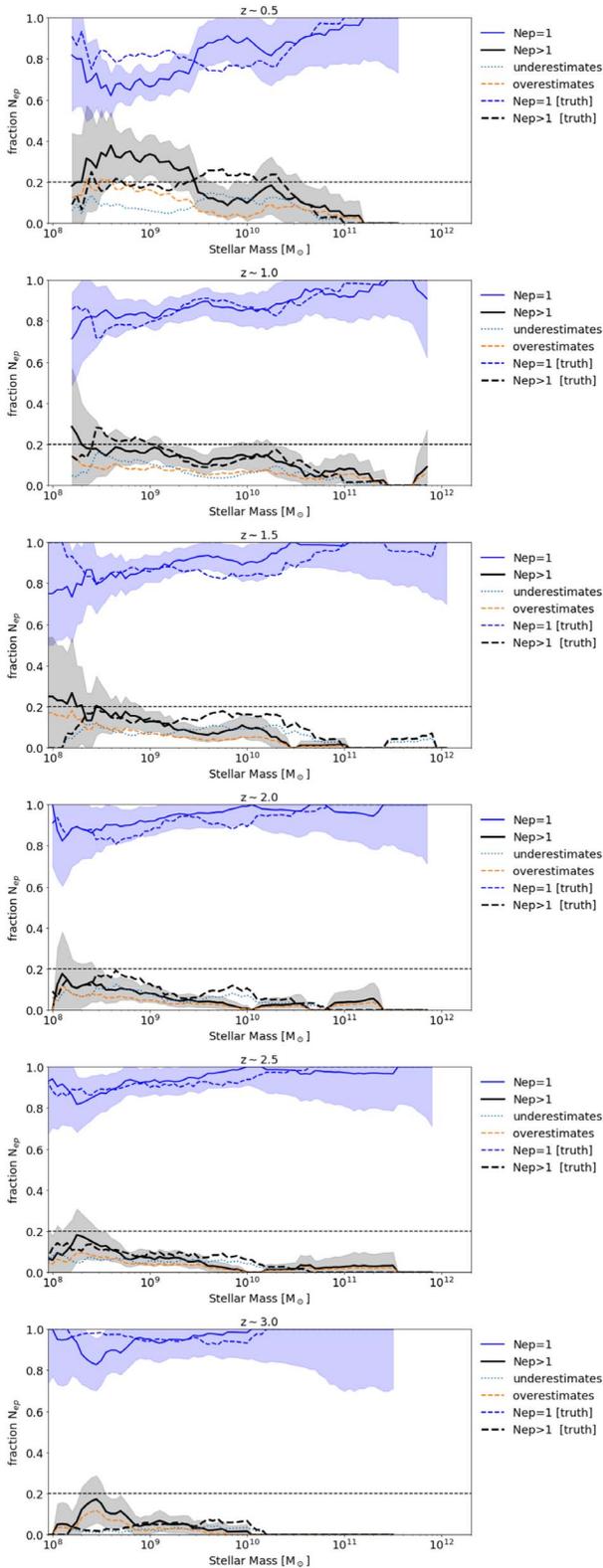
### 3.5. Effect of Varying GP Kernel and Hyperparameters

To generate a curve in SFR versus lookback time space, we use GPR to perform interpolation in cumulative mass ( $\int \text{SFR}(t) dt$ ) as a function of time, conditioned on constraints given by the N-tuples  $(\text{SFR}, M_*, \{t_X\})$ , as described in Section 2. The choice of kernel can make a significant difference to this interpolation, because it determines the extent to which the cumulative mass (and hence the SFR) at two times ( $t$  and  $t'$ ) are correlated (Rasmussen & Williams 2006; Foreman-Mackey 2015). In the current implementation, we choose to focus on kernels that are one-dimensional and stationary, which means that the kernel only depends on the separation between two times and not the times themselves.

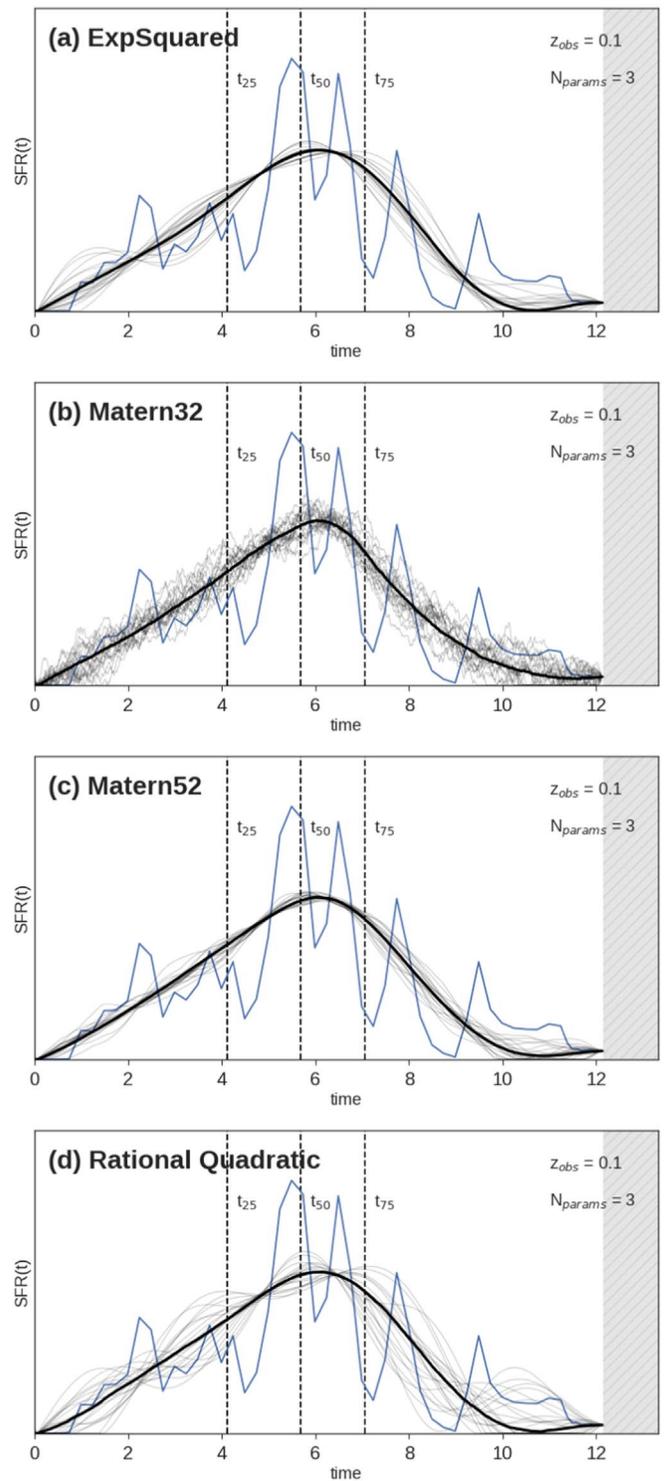
This choice of kernel is physically motivated in the sense that SFRs over short timescales are thought to be strongly correlated, with this correlation dropping off as the time separation increases, similar to the models explored by Kelson (2014) and Caplar & Tacchella (2019). To determine the most appropriate kernel from those widely used in the literature, we consider four stationary kernels implemented in the `george` Python package: the squared-exponential or Radial Basis kernel, the Matern32 and Matern52 kernels from the Matern class of kernels (Seeger 2004), and the Rational Quadratic kernel, which can be thought of as adding together many Radial Basis kernels with differing length scales. We test all of these different kernel functions using the SFHs from our mock catalog. An example of the reconstruction for a single SFH using different kernel choices is shown in Figure 8. In general, we find that the choice of kernel only makes a percent-level difference in computing the median SFH, although the uncertainties can be significantly different. We choose the Matern32 kernel because it gives the best reconstructions and most accurate uncertainties, as explored in Section 3.3. The Matern32 kernel as implemented in the `george` python package is given by the equation

$$k(r^2) = (1 + \sqrt{3r^2}) \exp(-\sqrt{3r^2}), \quad (8)$$

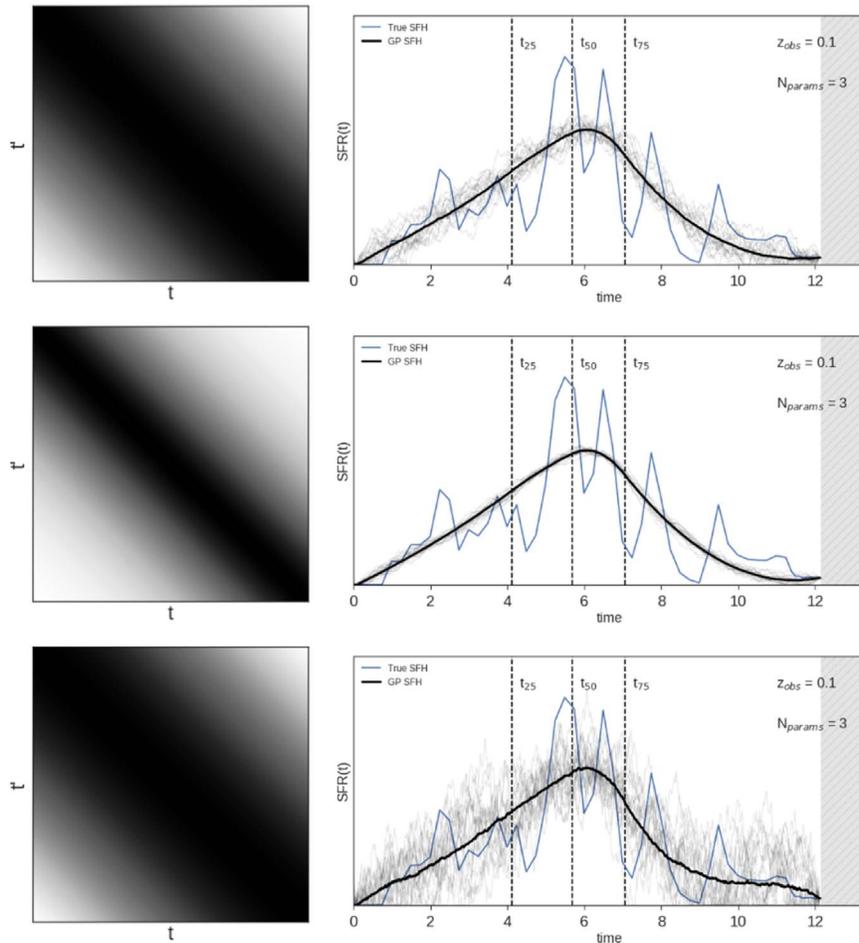
where  $r^2$  is related to the separation between two times. Because our kernel is one-dimensional and stationary (independent of choice of origin), this simplifies to  $r^2 \propto (t - t')^2$ , where  $r^2$  is related to the separation by a scale length. This scale length is determined directly within the `george` package by minimizing the log-likelihood of the data under the



**Figure 7.** Validation tests recovering the number of major episodes of star formation in a galaxy’s past by fitting the ensemble of mock noisy SEDs described in Section 3. The plots show the results of our analysis for the same redshift bins and mass range used for the main CANDELS sample. The solid lines show a sliding median within  $\pm 0.25$  dex in stellar mass for each quantity, and the shaded regions show the uncertainty for the estimates assuming Poisson noise.



**Figure 8.** The effect of varying the Gaussian process kernel, showing the SFH posterior constructed using the (a) Exponential-Squared (also called RBF), (b) Matern32 (used in this work), (c) Matern52, and (d) Rational Quadratic kernels. For each choice of kernel, we take the same input SFH (blue), compute the  $(M_*, \text{SFR}, \{t_X\})$  tuple, and then reconstruct an SFH using GPR with different kernel choices. For most well-motivated radial kernels, the choice of kernel only contributes to percent-level changes in the reconstructed SFH, which is a much smaller effect than the  $\sim 0.2\text{--}0.3$  dex uncertainties due to modeling systematics found in our validation tests (Figure 5).

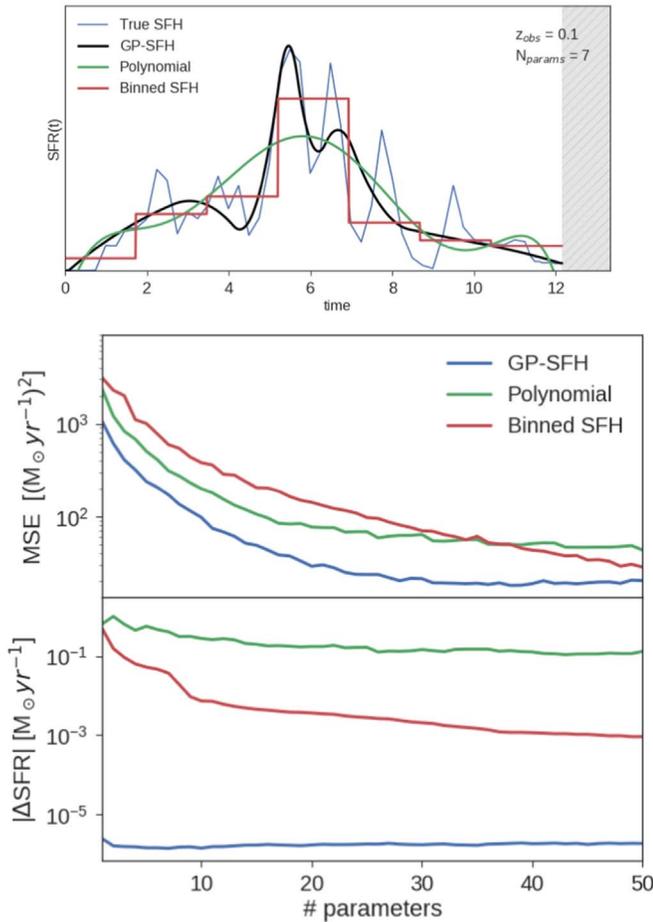


**Figure 9.** Varying the kernel hyperparameter controlling the scale length (amount of covariance between the cumulative SFH ( $\int \text{SFH}(t) dt$ ) at two times ( $t$  and  $t'$ ) does not significantly influence the reconstructed SFH obtained by sampling from the predictive posterior distribution conditioned on the tuple  $(M_*, \text{SFR}, t_{25}, t_{50}, t_{75})$ . The left panels show the covariance matrix between the cumulative SFH ( $\int \text{SFH}(t) dt$ ) at different points in time.

Gaussian process model. While the scale length varies with the number of SFH percentiles ( $N$ ), because there is more information to condition the likelihood with for greater  $N$ , it is fixed across the pregrid for a given choice of  $N$ . In Figure 9, we test the effect of increasing and reducing the scale length used in the Gaussian process kernel. This causes the covariance matrix to become more (less) coupled to SFRs separated by a given time interval. While this does not have a major effect on the reconstructed SFH, it can cause a decrease (increase) in the uncertainties estimated using random draws from the SFH posterior. Taken in conjunction with our validation of the SFH credible intervals in Section 3.3, we find that our choice of kernel and hyperparameters robustly reconstructs SFHs with accurate uncertainties. In addition to the basic kernels described in this section, it is possible to design more physically motivated kernels by computing the covariance function of SFHs from cosmological simulations. However, because this is subject to a wide variety of systematics depending on resolution, simulation size, and differing physical prescriptions, such an analysis requires a detailed comparison of different simulations. Although this out of the scope of the current paper, it is being actively investigated for a follow-up paper.

### 3.6. Minimizing Loss of Information

We test the effectiveness of the Gaussian process–based SFH reconstruction method in compressing data about the SFH and encoding the information in a small number of parameters. Because the SFH percentiles store information about the shape of the SFHs, the accuracy of describing an arbitrary SFH should decrease as we increase the number of parameters ( $N$ ). We see this to be the case in Figure 10, where the error in describing an ensemble of randomly drawn SFHs from our mock catalog decreases with increasing  $N$ . This analysis is done purely in SFH space as a test of the loss of information in going from the full SFH to the GP-SFH tuple. In addition to this, we compare the method with existing nonparametric techniques of approximating an arbitrary function using: 1) polynomial coefficients, and 2) binning the SFH linearly in time such that the SFR in any given bin is a constant. The latter approach and its variants—such as binning in logarithmic lookback time or using adaptive bins—are seen across the literature in works that explore nonparametric SFH reconstruction (Heavens et al. 2000; Tojeiro et al. 2007; Da Cunha et al. 2008; Dye 2008; Leja et al. 2017). While these are generally effective, they have the disadvantages of introducing unphysical discontinuities at bin edges. Repeating the test with these two methods, we find that their error approaches the GP-SFH error for large  $N$ , with the GP-SFH performing better for any given  $N$ .



**Figure 10.** Comparison of the Gaussian process–based method to other commonly used data compression methods, including a polynomial fit to SFHs and approximating a SFH using linear bins in time with constant SFR in each bin (Heavens et al. 2000; Tojeiro et al. 2007; Dye 2008). An example of this for a single SFH and  $N = 7$  is shown in the top panel. We use a sample of 1000 randomly selected SFHs from our mock catalogs for this test, where we approximate the known “true” SFH for each galaxy with three different methods, and quantify the error in the approximation using the mean-squared error (MSE) metric (middle panel) as well as the error in the SFR at the time of observation (lower panel) vs. the number of parameters. While the error reduces with increasing # parameters for all three methods, the GP-SFH method consistently performs better in both metrics, yielding better approximations of the true SFH for a fixed number of parameters.

Tuning the number of parameters during SED fitting therefore optimally estimates the amount of SFH information from each SED being fit.

#### 4. Data

In the current analysis, we use a sample of galaxies from the *HST*/F160w selected catalogs for the five CANDELS (Grogin et al. 2011; Koekemoer et al. 2011) fields covering a total area of  $\sim 800$  arcmin<sup>2</sup>: GOODS-S (Guo et al. 2013), GOODS-N (Barro et al. 2019), COSMOS (Nayyeri et al. 2017), EGS (Stefanon et al. 2017), and UDS (Galametz et al. 2013).

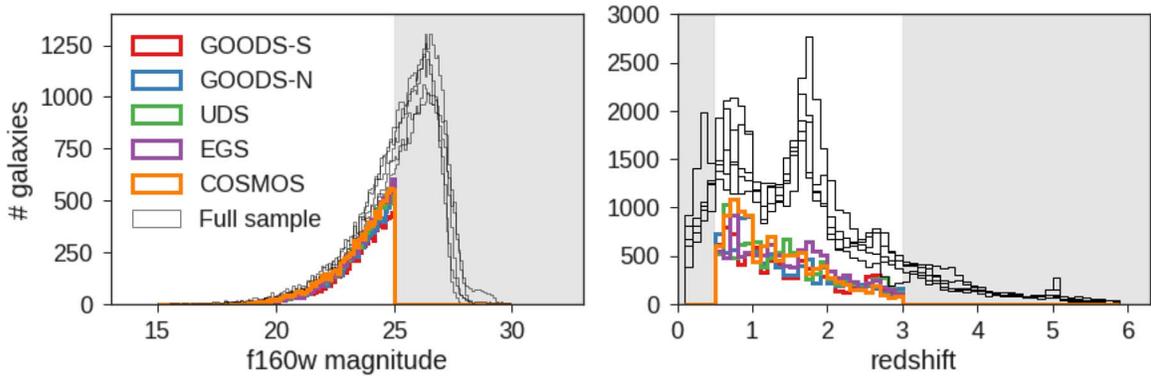
The GOODS-South (Guo et al. 2013) field contains 34,930 objects and covers an area of  $\sim 170$  arcmin<sup>2</sup>, with a  $5\sigma$  limiting depth of 27.4, 28.2, and 29.7 AB magnitudes in the three overlapping survey regions (CANDELS wide, deep, and HUDF regions). The GOODS-North field (Barro et al. 2019) contains 35,445 objects over a similar area, with a  $5\sigma$  limiting depth of 27.5 AB mag (Pacifci et al. 2016). The UKIRT Infrared Deep

**Table 1**

Collected Measurements Comprising the UV-to-NIR SEDs of Galaxies across the Five CANDELS Fields

Field	Filter Set
GOODS-S (Guo et al. 2013)	Blanco/CTIO U, VLT/VIMOS U, <i>HST</i> /ACS f435w, f606w, f775w, f814w, f850lp, <i>HST</i> /WFC3 f098m, f105w, f125w, f160w, VLT/ISAAC Ks, VLT/Hawk-I Ks, <i>Spitzer</i> /Irac 3.6, 4.5, 5.8, 8.0 $\mu$ m
GOODS-N (Barro et al. 2019)	KPNO U, LBC U, <i>HST</i> /ACS f435w, f606w, f775w, f814w, f850lp, <i>HST</i> /WFC3 f105w, f125w, f140w, f160w, f275w, MOIRCS K, CFHT Ks, <i>Spitzer</i> /Irac 3.6, 4.5, 5.8, 8.0 $\mu$ m
UDS (Galametz et al. 2013)	CFHT/MegaCam u, Subaru/Suprime-Cam B, V, Rc, i', z', <i>HST</i> /ACS f606w, f814w, <i>HST</i> /WFC3 f125w, f160w, VLT/Hawk-I Y, Ks, WFCAM/UKIRT J, H, K, <i>Spitzer</i> /Irac 3.6, 4.5, 5.8, 8.0 $\mu$ m
EGS (Stefanon et al. 2017)	CFHT/MegaCam u*, g', r', i', z', <i>HST</i> /ACS f606w, f814w, <i>HST</i> /WFC3 f125w, f140w, f160w, Mayall/NEWFIRM J1, J2, J3, H1, H2, K, CFHT/WIRCAM J, H, Ks, <i>Spitzer</i> /Irac 3.6, 4.5, 5.8, 8.0 $\mu$ m
COSMOS (Nayyeri et al. 2017)	CFHT/MegaCam u*, g*, r*, i*, z*, Subaru/Suprime-Cam B, g+, V, r+, i+, z+, <i>HST</i> /ACS f606w, f814w, <i>HST</i> /WFC3 f125w, f160w, Subaru/Suprime-cam IA484, IA527, IA624, IA679, IA738, IA767, IB427, IB464, IB505, IB574, IB709, IB827, NB711, NB816, VLT/VISTA Y, J, H, Ks, Mayall/NEWFIRM J1, J2, J3, H1, H2, K, <i>Spitzer</i> /Irac 3.6, 4.5, 5.8, 8.0 $\mu$ m

Sky Survey (UKIDSS) Ultra-Deep Survey (UDS) catalog (Galametz et al. 2013) contains 35,932 sources over an area of 201.7 arcmin<sup>2</sup> with a  $5\sigma$  limiting depth of 27.45 AB magnitudes. The Extended Growth Strip (EGS) catalog (Stefanon et al. 2017) contains 41,457 objects over an area of  $\approx 206$  arcmin<sup>2</sup> reaching a depth of 26.62 AB mag. The COSMOS field (Nayyeri et al. 2017) contains 38,671 objects covering an area of  $\approx 216$  arcmin<sup>2</sup> with a limiting depth of 27.6 AB mag. The catalogs select objects via SExtractor in dual-image mode using F160w as the detection band. The dual-image mode (Galametz et al. 2013) is optimized to detect both faint, small galaxies in “hot” mode without over-deblending large, resolved galaxies detected in “cold” mode. The *HST* (ACS and WFC3) bands were matched using a point-spread function to measure photometry, and TFIT (Laidler et al. 2007) was used to measure the photometry of ground-based and IRAC bands using the *HST* WFC3 imaging as a template. The SEDs in the five fields include a wide range of UV-to-NIR measurements, with the flux measured in 17, 18, 19, 23, and 43 photometric bands in GOODS-S, GOODS-N, UDS, EGS, and COSMOS, respectively. The photometric filters used for each field are detailed in Table 1.



**Figure 11.** The distribution of galaxies across the five CANDELS fields in *HST*/WFC3 F160w magnitude (left) and redshift (right). Thin black lines show the full distribution for each field and colored lines show the sample used in the current analysis after the selection described in Table 2. While most redshifts are photometric, the sample contains  $\approx 7000$  galaxies with spectroscopic redshifts. Gray regions show parts of the sample that we exclude in the current analysis.

**Table 2**  
The Number of Galaxies Used in Our Analysis, and the Effect of Each Step in the Selection Process

Field	All Objects	Good Flags <sup>a</sup>	$0.5 < z < 3$	$H < 25^b$	$\chi^2_{\text{red}} < 10$	Final Sample	$z_{\text{spec}}^c$
GOODS-S (Guo et al. 2013)	34,930	31,273	22,713	8,520	8,299	8,299	1,758
GOODS-N (Barro et al. 2019)	35,445	34,693	26,838	9,551	9,206	9,206	2,399
UDS (Galametz et al. 2013)	35,932	26,917	21,263	10,234	10,176	10,176	538
EGS (Stefanon et al. 2017)	41,457	31,714	24,444	10,554	10,261	10,261	1,671
COSMOS (Nayyeri et al. 2017)	38,671	30,070	22,092	10,883	10,849	10,849	705

**Note.**

<sup>a</sup> Galaxies with flags indicating no contamination by nearby objects, halos, or star spikes, as well as objects with a low stellarity classification given by the SExtractor CLASS STAR output.

<sup>b</sup> We select galaxies brighter than 25 mag in the *HST*/WFC3 F160W band, to ensure sufficient S/N in the rest-optical regime of the SED necessary for accurate SFH reconstruction.

<sup>c</sup> This column gives the number of galaxies with confirmed spectroscopic redshifts in each field for our final analysis sample.

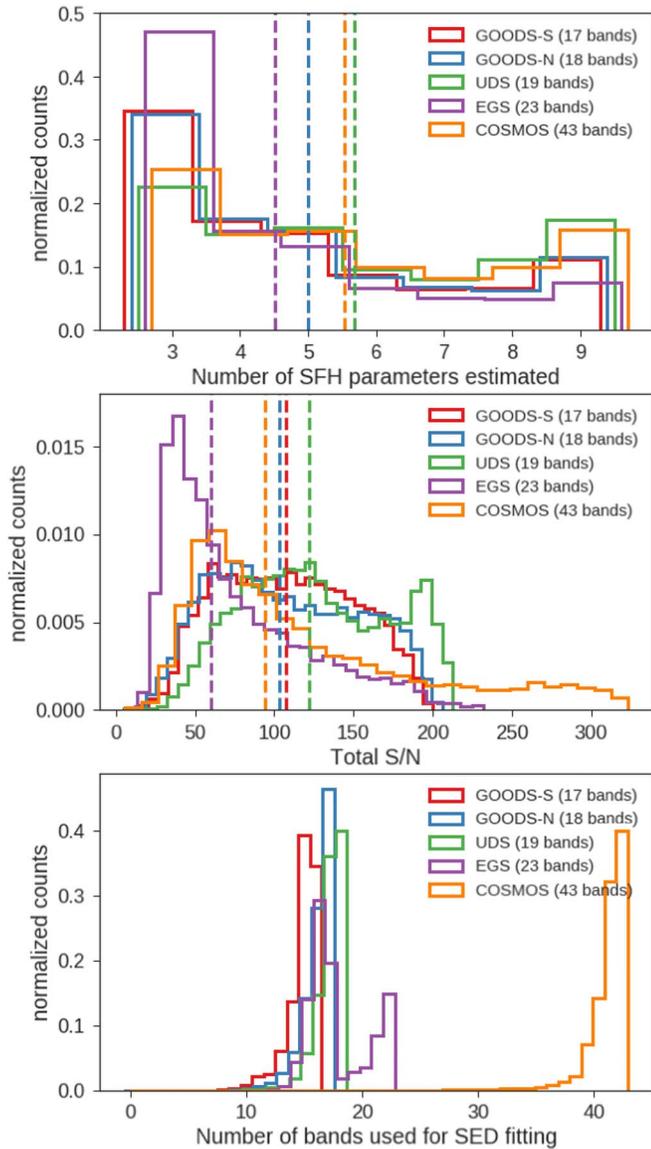
In this paper, we focus on galaxies at  $0.5 < z < 3.0$  because we do not have multiple reliable rest-UV measurements from the *HST* bands at  $z < 0.45$ , leading to larger uncertainties in SFR, and we cannot accurately constrain the  $1.6 \mu\text{m}$  bump at redshifts  $z > 3.06$ , which is important for robust estimation of stellar masses. In addition to estimating robust stellar masses and star formation rates, it is necessary to ensure robust S/N in the rest-optical portion of the SED, which is most sensitive to variations in the SFH. As a proxy for the total S/N, we restrict our sample to galaxies with  $H < 25$ , where H is the *HST*/WFC3 F160w band. In addition to restricting our sample to the those with good S/N, this cut also helps alleviate inhomogeneities in the depths of the different fields, especially GOODS-N and GOODS-S, which have a wedding-cake structure. After implementing these selection cuts, we are left with a total of 48,791 galaxies. The effect of each selection effect and the total number of galaxies used for the analysis is given in Table 2 and Figure 11. To perform our fits, we use an updated CANDELS photometric redshift catalog by D. Kodra et al. (2019, in preparation) containing an increased number of spectroscopic redshift measurements as well as photometric redshifts with Bayesian combined uncertainties estimated by comparing the redshift probability distributions of four different SED fitting methods. We perform our fits using their  $z_{\text{best}}$  binned to the resolution of our pregrid, with  $\delta z = 0.01$ .

## 5. Results

### 5.1. The SFHs of Galaxies at $0.5 < z < 3$

We now apply the Dense Basis SED fitting method as described in Section 2 to the sample of CANDELS galaxies at  $0.5 < z < 3.0$ , to reconstruct the SFH of each galaxy with uncertainties. Figure 12 shows the distribution of the number of SFH parameters that individual galaxies are best fit with in each of the five fields, with a slight trend toward increasing amounts of SFH information recovered with more photometric bands.

In Figure 13, we show the distributions of reconstructed SFHs in the five CANDELS fields and how they evolve with mass and redshift. The median and interquartile range are computed at each point in time as in Pacifici et al. (2016), who performed a similar calculation for quiescent galaxies using a basis of SFHs derived from a semi-analytic model (De Lucia & Blaizot 2007). In the current analysis, we have used a redshift bin width of  $\delta z = \pm 0.1$ , and a mass bin width of  $\delta M_* = \pm 0.25$  dex, such that the  $M_* \sim 10^{10} M_\odot$  bin includes all galaxies with  $10^{9.75} < M_* < 10^{10.25} M_\odot$ . Dashed black lines in each panel show the mean SFR ( $\equiv M_*/t_{\text{univ}}$ ), assuming constant SFR for that redshift and mass bin, and the  $f_{\text{samp}}$  at the top left corner of each panel shows the fraction of all galaxies in our sample at that redshift that fall in a particular mass bin. Because there can be a few galaxies that fall outside the plotted mass range, this may not sum to unity.



**Figure 12.** Top: distributions of the number of SFH parameters estimated while fitting the SEDs of galaxies at  $0.5 < z < 3.0$  in the five CANDELS fields, with the vertical dashed lines showing the mean value of the distribution. Middle: the distribution of total  $S/N = \sqrt{\sum_j (F_{i,j}/\sigma_{v,j})^2}$  for the five fields, with the vertical dashed lines showing the medians of each distribution. Bottom: the distribution of the number of photometric bands used in fitting the SEDs in each field. The bimodality in the EGS observations is due to partial coverage of the field with the six NEWFIRM bands (J1, J2, J3, H1, H2, K).

We see that the median SFH across the five CANDELS fields are remarkably similar, as expected. There are discrepancies in a few of the bins, most notably at  $2 < z < 3$  and  $M_* \sim 10^{10} M_\odot$  for UDS, which could be the result of correlated photometric noise or smearing effects in the photo- $z$  that was used for the calculation. In the highest redshift bin, the portion of the SED with wavelengths greater than rest-frame  $1.6 \mu\text{m}$  is not well-sampled, leading to poorer constraints on the SFHs in the last row.

In general, SFHs tend to rise at high redshifts and low stellar masses, similar to those from cosmological simulations, with a turnover and subsequent decline as we go to higher masses and lower redshifts. In agreement with Pacifici et al. (2012) and Behroozi et al. (2018), massive galaxies tend to peak earlier in

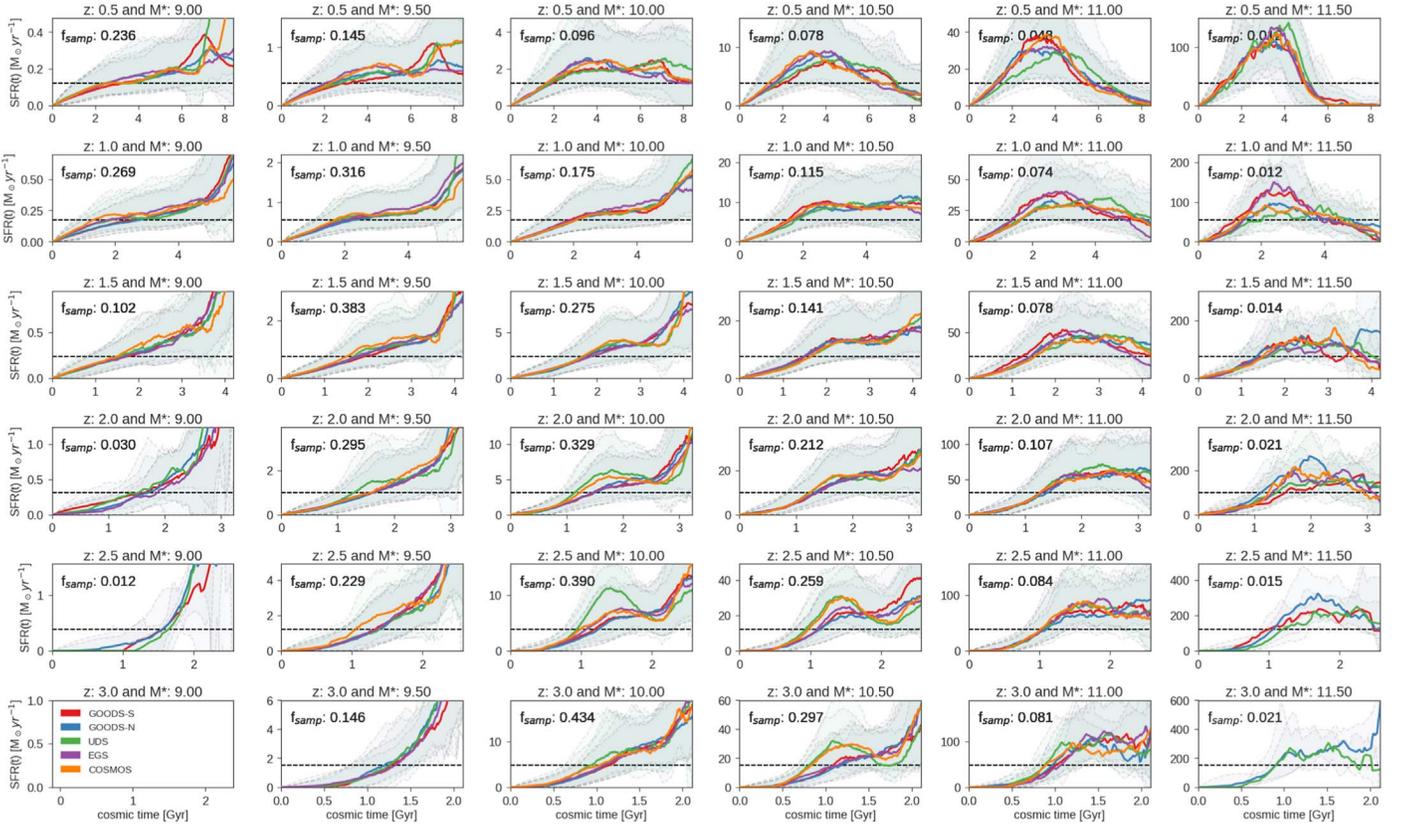
their SFHs, and galaxies on average tend to move toward quiescence at lower masses as the universe grows older, with the threshold changing from nearly  $10^{11.5} M_\odot$  at  $z \sim 3$  to  $10^{10} M_\odot$  at  $z \sim 0.5$ , without the need for any implicit assumptions about the SFHs, such as the well-motivated assumption of Behroozi et al. (2018) that earlier-forming halos get lower SFRs. The additional advantage of the Dense Basis method is that, in addition to the average SFHs, the individual SFHs of galaxies reconstructed from the observations allow us to explore the additional factors that drive the diversity of SFHs at a given mass and epoch. This allows us to extend our analysis beyond the physics encoded in the stellar mass–halo mass relation, which gives a constraint upon the first-order behavior of SFHs.

### 5.2. The Number of Major Episodes of Star Formation Experienced by Galaxies

Most studies of galaxy SFHs focus on the overall rise and fall of an ensemble of SFHs (Leitner 2012; Pacifici et al. 2016; Ciesla et al. 2017; Leja et al. 2017), which has led to a well-constrained understanding of the overall behavior seen in Figure 13. However, with smooth, nonparametric SFHs, it is now possible to ask questions about the *second-order* statistics of an ensemble of SFHs, analyzing the departures from this overall behavior in the form of periods of relative quiescence between episodes of star formation.

In Figure 14, we show the fraction of galaxies with different numbers of major episodes of star formation in a galaxy’s past at redshifts  $0.5 < z < 3.0$ . Because the number of episodes is a discrete quantity, Poisson noise dominates the formal uncertainties in an individual galaxy’s number of episodes while calculating functions of the sample, as discussed in Section 3. The different fields (colored lines), are in good agreement with each other and the median of the full sample (solid black line).

We find that, at low redshifts, the fraction of galaxies with multiple major episodes of star formation decreases as we go up in stellar mass above  $10^{10.5} M_\odot$ , in agreement with Iyer & Gawiser (2017). In addition to this, we find a slight decrease in the overall fraction of galaxies with multiple episodes with increasing redshift at any given mass, with the notable exception of  $M_* \approx 10^{10.5} M_\odot$ , which does not show a noticeable evolution with redshift. Although we have accounted for S/N variations, the decrease at lower masses could be at least partially due to having insufficient S/N to resolve multiple episodes of star formation as we go to lower masses and higher redshifts. A few explanations are possible for the behavior at high masses: AGN feedback quenching galaxies (Weinberger et al. 2018) could lead to SFHs that form most of their stars at by  $z \sim 3$ , which could look like a single early episode of star formation without sufficient S/N in the SEDs to resolve the older populations at  $z \sim 1$ . This is made more probable by the fact that, while most galaxies with multiple episodes are found to lie on the SFR– $M_*$  correlation, the greatest number of galaxies with low SFRs at the time of observation and multiple episodes occurs at masses close to  $10^{10.5} M_\odot$ . Another reason could be the central limit theorem (Kelson 2014): massive galaxies that grow primarily through mergers (Brinchmann et al. 2004; Bundy et al. 2005; Pérez-González et al. 2008) at early times could be composed of multiple progenitors. As the number of progenitors grows with mass, by the central limit theorem, their SFHs should look smoother than those for less-massive galaxies. Low-mass galaxies, by comparison, should

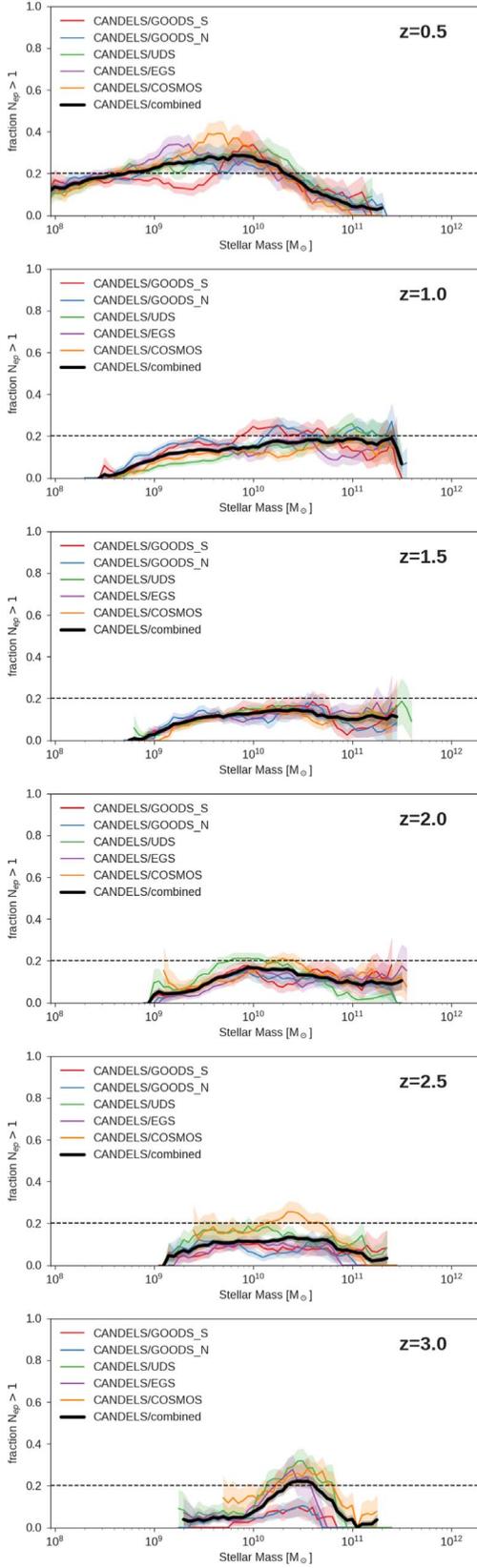


**Figure 13.** The median star formation histories (SFHs) of galaxies in the five CANDELS fields in bins of stellar mass (horizontal) and redshift (vertical), showing how galaxies evolve with cosmic time as they grow in stellar mass. Solid colored lines show the median SFH for each field separately, showing a remarkable similarity across the different fields in the majority of the bins. The shaded regions show the 16th–84th percentiles in the SFHs, highlighting the diversity in SFHs as a function of stellar mass and epoch. The dashed black line shows the mean SFR ( $\equiv M_*/t_{\text{univ}}$ ) assuming constant SFR for that redshift and mass bin, and the  $f_{\text{samp}}$  at the top left corner of each panel shows the fraction of all galaxies in our sample at that redshift that fall in a particular mass bin. In good agreement with cosmological simulations, as well as semi-analytical and empirical models, galaxy SFHs tend to rise with time at low masses and high redshifts, then start to turn over at high masses, with the turnover mass decreasing as we go to lower redshifts. However, in addition to the average SFH behavior, we also have access to the individual SFH for each galaxy, which now opens up the possibility of repeating this analysis to trace the evolution of SFHs with quantities like  $t_{50}$ , metallicity, morphology, central density, size, environment, and other probes of galaxy evolution.

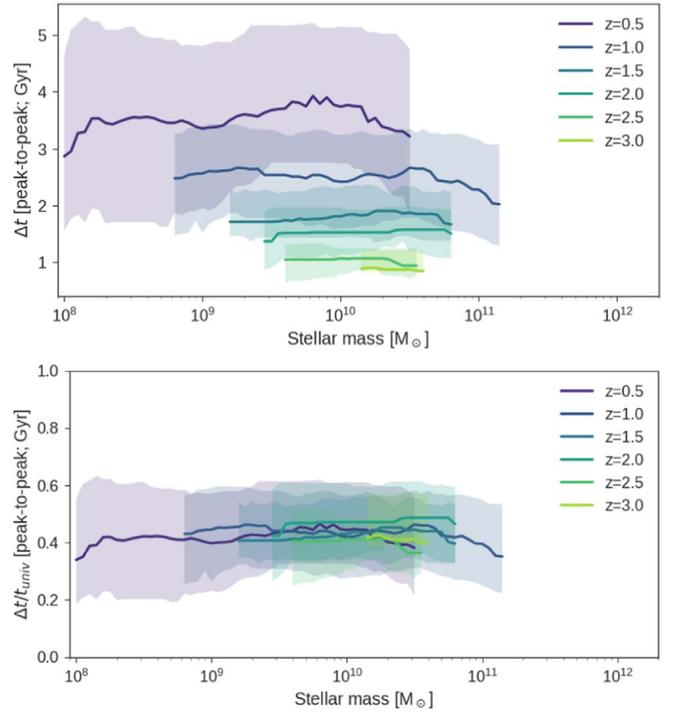
have more stochastic SFHs because they are growing most of their mass in situ, which would be in close agreement with the findings of Guo et al. (2016), Matthee & Schaye (2019), Emami et al. (2018), Shivaie et al. (2015), and Broussard et al. (2019).

If a galaxy is found to have multiple strong episodes of star formation in its lifetime, an interesting question would be whether the galaxy was actively forming stars and the star formation was temporarily suppressed by a quenching attempt (short time interval between peaks), as opposed to a galaxy that was on its way to quiescence but restarted star formation due to an inflow of pristine gas or merger. This is especially interesting within the context of rejuvenation of galaxy SFRs (Fang et al. 2012), because it sets timescales for how long a galaxy spends off the star-forming sequence when it makes such an excursion. To quantify this, we measure the time interval between multiple peaks for the subsample of galaxies with  $N_{\text{ep}} > 1$ . We plot this as a function of redshift and mass in Figure 15, finding that although the separation does not vary strongly with mass, it does show a strong trend with redshift. However, upon normalizing by the age of the universe at different redshifts, this trend is significantly decreased, leaving us with a roughly constant timescale across which galaxy rejuvenation occurs, given by  $t_{\Delta\text{peak}} \sim 0.42^{+0.15}_{-0.10} t_{\text{univ}}$  Gyr, where  $t_{\text{univ}}$  is the age of the universe at the redshift of

observation. This is similar to the result by Abramson et al. (2016), which found the transit time through the green valley (i.e., half the time between two peaks for a rejuvenating SFH) to be  $\sim 0.2 t_{\text{univ}}$  Gyr, roughly independent of redshift (see also Pandya et al. (2017), which studies the transition timescales in massive CANDELS galaxies using a statistical analysis of their number densities). This is also related to the results of Pacifici et al. (2016), who that found that the width of the SFH for quiescent galaxies is roughly constant across stellar mass and redshift when the age of the universe is factored out, as well as those of Muzzin et al. (2014), who found that the post-starburst spectra of galaxies at  $z \sim 1$  are fit well with a quenching timescale of  $0.4^{+0.4}_{-0.3}$  Gyr. Fang et al. (2012) identify a subsample of galaxies at  $z \sim 0.1$  that could linger in the green valley for  $\mathcal{O}(\text{Gyr})$ . The astute reader may wonder how significant it is to find that two episodes are typically separated by roughly half the age of the universe at the time of observation, as this also corresponds to the median period of a generic sine wave possessing two peaks without that interval. Given the present data quality, it is difficult to test this further by investigating the separation between the earliest two star formation episodes in galaxies whose SFHs show 3–5 major episodes of star formation, but such a test should be done with higher-S/N spectrophotometry. At present, we can compare the fit for separation between episodes against the predictions of



**Figure 14.** The fraction of galaxies that show multiple major episodes of star formation in their SFHs as a function of stellar mass at various redshifts for galaxies in the five CANDELS fields. The solid line shows a running median within a bin of  $\pm 0.25$  dex in stellar mass, and the shaded regions show the uncertainty for the estimates assuming Poisson noise.

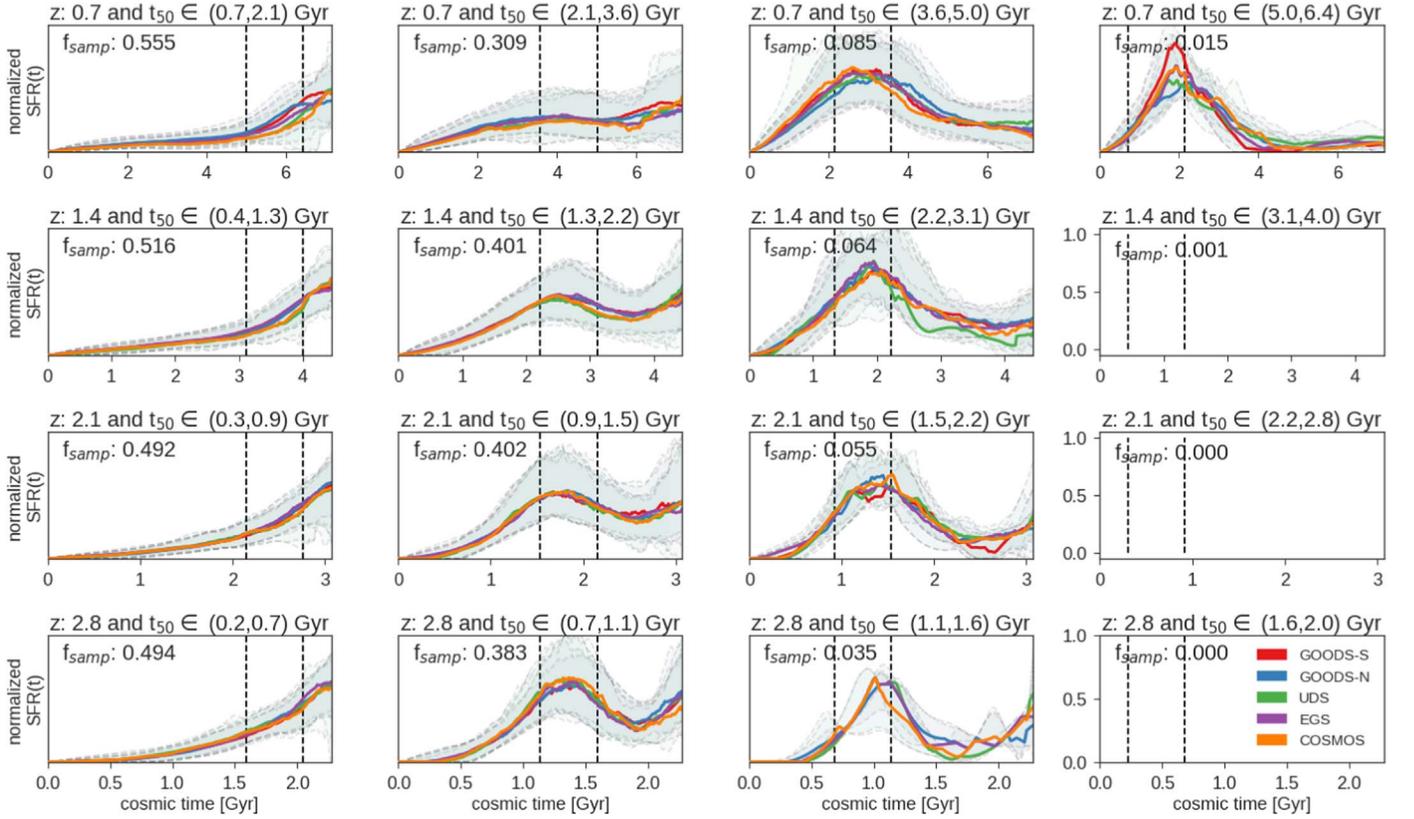


**Figure 15.** The separation between multiple peaks of star formation, as a function of mass and redshift for the subsample of galaxies that have  $N_{\text{ep}} > 1$ . The redshift bins are the same as Figure 14, and the solid line and shaded region show the median and 16–84th percentiles, respectively. The top panel shows the distribution across stellar mass at different redshifts. The bottom shows the same, but divided by the age of the universe at that epoch.

galaxy formation models, finding similar trends albeit with a slightly smaller value of  $\approx 0.3 \pm 0.15 t_{\text{univ}}$  Gyr. The consistency of our result with a variety of similar results across a range of redshifts summarized above is an additional reassuring check. In Behroozi et al. (2018), galaxy rejuvenation is a generic feature of a population, with the timescale depending on the mode by which a halo is accreting mass, whether through mergers, accretion from another halo, or infall. In this scenario, rejuvenation occurs more often when the quenched population evolves more slowly than the halo dynamical time, during which it can switch between modes of accretion, increasing or decreasing the SFR as a result. This results in estimates of the fraction of galaxies with past rejuvenation as a function of mass, decreasing as redshift increases and showing a trend in stellar mass that is consistent with our results. Our result seems to indicate that the rejuvenation timescales remain relatively constant over cosmic time. It is important to note that the scatter in this quantity is quite large, and the evolution in the quenched fraction happens most rapidly at redshifts  $0 < z < 0.5$  (Muzzin et al. 2013; Behroozi et al. 2018; Donnari et al. 2019; Hahn et al. 2019), so extrapolation to that regime needs to be tested with further data.

### 5.3. The Different Demographics of Galaxies

In keeping with the finding of Pacifici et al. (2016) that the widths of the SFHs of passive galaxies are roughly constant upon factoring out the age of the universe, and our similar finding for the time interval between two peaks of star formation, we consider galaxy SFHs binned in  $t_{50}$ . We bin



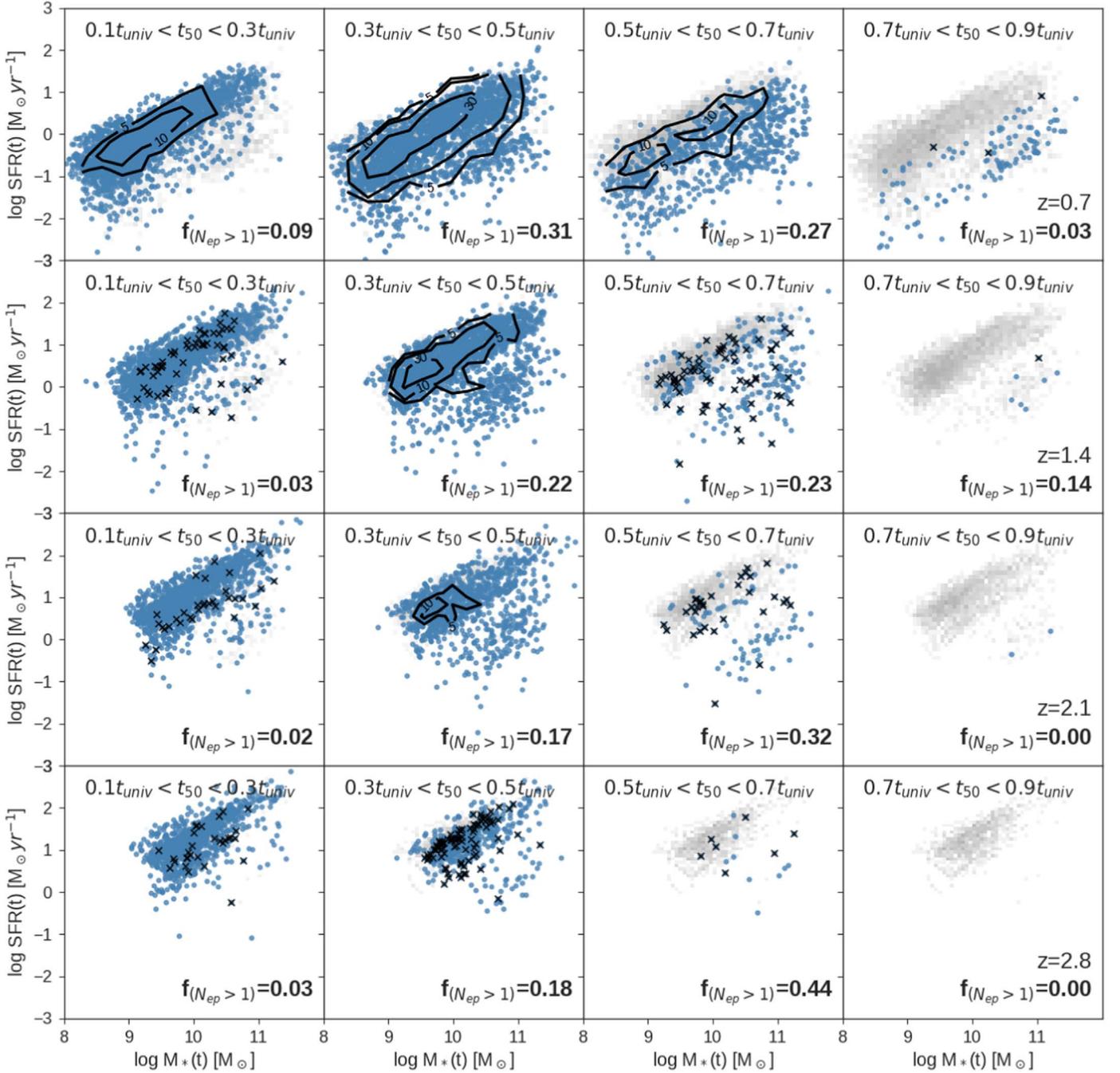
**Figure 16.** SFHs split into linearly increasing bins of  $t_{50}$ , the lookback time at which a galaxy assembled 50% of its stellar mass, from  $0.1t_{\text{univ}} < t_{50} < 0.3t_{\text{univ}}$  (formed recently),  $0.3t_{\text{univ}} < t_{50} < 0.5t_{\text{univ}}$ ,  $0.5t_{\text{univ}} < t_{50} < 0.7t_{\text{univ}}$ , and  $0.7t_{\text{univ}} < t_{50} < 0.9t_{\text{univ}}$  (formed earliest), in four bins of redshift. The plotting scheme and colors are the same as Figure 13. In each bin, the SFHs are normalized to the same mass because we are most interested in the diversity of SFH shapes for the entire demographic. Vertical dashed black lines show the  $t_{50}$  bounds for each panel. We see that the SFHs in a bin broadly tend to describe one of four demographics of galaxies: starbursting galaxies at high redshifts and late bloomers at  $z \sim 0.7$  (Dressler et al. 2016) can be found in the first column from the left, star-forming galaxies contribute to the median SFH in columns 1–3, post-starburst or green-valley galaxies (Fang et al. 2012) in columns 2–3, and quiescent galaxies in columns 3–4. The  $f_{\text{samp}}$  at the top left of each panel shows the fraction of galaxies at each redshift that fall into each demographic. In conjunction with the  $UVJ$  diagram and position on the  $\text{SFR}-M_*$  plane, the SFHs of galaxies allow for additional diagnostics regarding its evolutionary phase.

galaxies in four bins,  $0.1t_{\text{univ}} < t_{50} < 0.3t_{\text{univ}}$ ,  $0.3t_{\text{univ}} < t_{50} < 0.5t_{\text{univ}}$ ,  $0.5t_{\text{univ}} < t_{50} < 0.7t_{\text{univ}}$ , and  $0.7t_{\text{univ}} < t_{50} < 0.9t_{\text{univ}}$ , at different redshifts. We show the results in Figure 16. As in Figure 13, each panel lists the fraction of the sample in a particular bin. However, in this case, the fractions are no longer tracing the mass function of galaxies. Instead, the four bins in time serve as proxies for galaxy SFHs at different stages of their lifetimes. This enables us to identify different populations of galaxies, including starbursting galaxies, late bloomers (Dressler et al. 2016), star-forming galaxies, post-starburst or green-valley galaxies (Fang et al. 2012), and quiescent galaxies, using different redshift-dependent  $t_{50}$  cuts, either independently or in combination with other factors like  $t_{25}$ ,  $t_{75}$ , size, and morphology. To further interpret these SFHs, Figure 17 displays the positions of the galaxies within each panel in Figure 16 on the  $\text{SFR}-M_*$  plane. We find that the left (right) columns in both figures select star-forming (quiescent) galaxies that lie on (off) the star-forming sequence. At intermediate values of  $t_{50}$ , the populations are a combination of star-forming and quiescent galaxies, with the selection gradually shifting from star-forming to quiescent galaxies. The intermediate  $t_{50}$  panels also show an excess of galaxies with multiple strong episodes of star formation ( $N_{\text{ep}} > 1$ ). While this is an intuitive result, because galaxies that have assembled most of their mass recently or those that have long since shut off their star formation are not very likely to contain multiple episodes of star formation, it has

important implications for analyses that assume galaxies evolve along smooth SFH trajectories (Leitner 2012) or are described by simple parametric forms (Dressler & Abramson 2014; Ciesla et al. 2017; Lee et al. 2018).

#### 5.4. Correlation with Morphology

The morphologies of galaxies are seen to strongly correlate with their stellar masses and redshifts (Conselice 2014), as well as sSFR (Whitaker et al. 2015). While this is a combined effect of the different processes that regulate star formation within galaxies, including mergers, gas accretion through inflows, and stellar and AGN feedback (Boselli & Gavazzi 2006; Hopkins et al. 2008, 2014; Anglés-Alcázar et al. 2017; Weinberger et al. 2018), it is difficult to observationally disentangle the relative strengths of these effects. However, the different timescales that these processes act on enable us to discriminate between the relative effects of these processes, if we can observationally constrain the timescales on which morphological transformation occurs across a population of galaxies—as was done for groups in, e.g., Kovač et al. (2010). The reconstructed SFHs of galaxies provide a direct probe of these timescales by correlating the morphologies of galaxies with their SFHs as compared to indirect measurements of timescales through the frequencies of different morphologies, which are subject to a variety of systematics and selection effects.



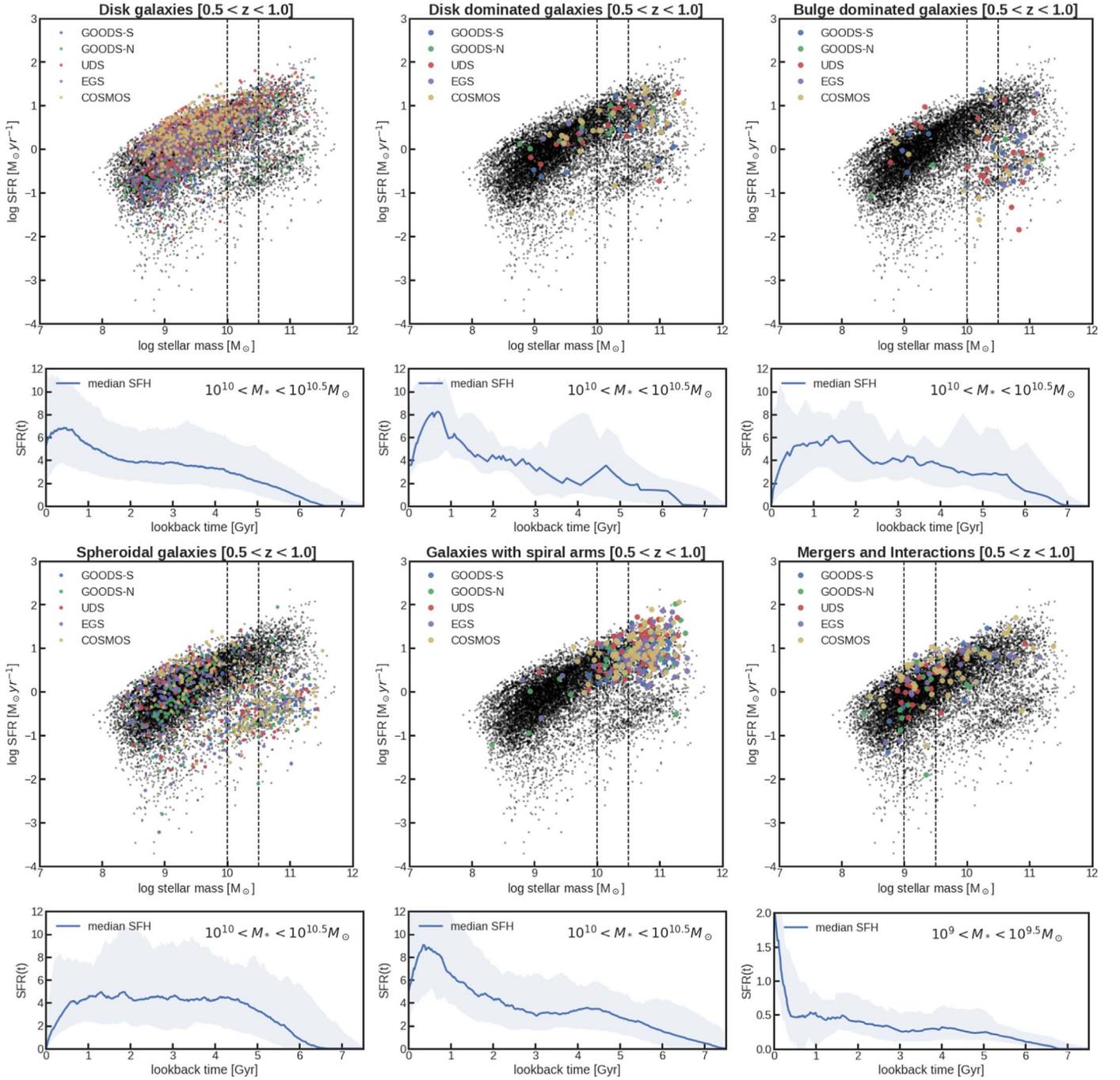
**Figure 17.** Positions on the SFR– $M_*$  plane for the galaxies shown in each panel of Figure 16. The underlying gray heatmap shows the full sample at a given redshift, and blue points show all galaxies satisfying  $0.1t_{univ} < t_{50} < 0.3t_{univ}$  (formed recently),  $0.3t_{univ} < t_{50} < 0.5t_{univ}$ ,  $0.5t_{univ} < t_{50} < 0.7t_{univ}$ , and  $0.7t_{univ} < t_{50} < 0.9t_{univ}$  (formed earliest) within a redshift bin. The black crosses are a subset of the blue points that are identified as having more than one major episode of star formation during their lifetimes ( $N_{ep} > 1$ ), with this fraction of galaxies given in the bottom right corner. Black contours are used where there are more than 100 galaxies with  $N_{ep} > 1$  in a given panel.

We use the CANDELS wide morphology catalogs by Kartaltepe et al. (2015) to study the SFHs of galaxies with different morphological features at  $0.5 < z < 1.0$ . We limit our redshift range to avoid the effects of small number statistics of classifications as we go to higher redshifts. The morphology catalogs contain visual classifications for over 50,000 objects spanning  $0 < z < 4$  with  $f160w < 24.5$ , which gives a large overlap with our sample. The catalogs contain flags for main morphology class (disk, spheroid, peculiar/irregular, point

source/compact, and unclassifiable), a class for mergers and other interactions, and structure flags for bars, tidal features, spiral arms, and more. For each class and flag, the catalog reports the fraction of classifiers who were confident about the existence of that feature.

We use this to analyze the SFHs of six classes of galaxies, described as follows.

1. Disk: ( $f_{disk} > 0.9$ ) AND ( $f_{sph}, f_{irreg} < 0.1$ ). This includes the set of all galaxies classified as disky.



**Figure 18.** The star formation histories of galaxies at  $0.5 < z < 1.0$  with different morphological features identified in Kartaltepe et al. (2015). For each class, the top panel shows the position of the galaxies under consideration (colored points) in the SFR– $M_*$  plane, with the full population shown as black dots. The bottom panel shows the median SFH (blue solid line) and diversity (16th to 84th percentile; shown as a shaded blue region) of all the galaxies with  $M_* \in [10^{10}, 10^{10.5}] M_\odot$  that satisfy a particular morphological criterion, except for the last bin, where we have chosen a lower mass bin due to insufficient statistics. The second and third panels show galaxies that have both a disk and bulge component, which are then broken down into disk-dominated versus bulge-dominated galaxies.

2. Disk-dominated galaxies: ( $f_{\text{disk dom}} > 0.9$ ). Disks with a central bulge where the disk dominates the structure.
3. Bulge-dominated galaxies: ( $f_{\text{bulge dom}} > 0.9$ ). Disks with a central bulge where the bulge dominates the structure.
4. Spheroidal: ( $f_{\text{sphk}} > 0.9$ ) AND ( $f_{\text{disk}}, f_{\text{irreg}} < 0.1$ ). This includes the set of all galaxies classified as spheroidal.
5. Galaxies with spiral arms: ( $f_{\text{arms}} > 0.9$ ).
6. Mergers and interactions: galaxies that either appear to have undergone a merger, as evidenced by tidal features, structures such as loops, or highly irregular outer

isophotes ( $f_{\text{merger}} > 0.9$ ), or are interacting with a companion galaxy within the segmentation map from SExtractor ( $f_{\text{intl}} > 0.9$ ).

The results of this analysis are shown in Figure 18. The Figure contains six sets of panels, one for each subsample of galaxies. The top panel shows where the galaxies lie on the SFR– $M_*$  correlation, and the bottom panel shows the median SFH for the subsample of these galaxies at  $M_* \in [10^{10}, 10^{10.5}] M_\odot$ . This is useful to test feedback-driven models of quenching

that posit a correlation between bulge–total ratios and SFH shape (Zolotov et al. 2015; Belfiore et al. 2016; Tacchella et al. 2016; Abramson et al. 2018). While the SFHs of our pure disk population at  $M_* \sim 10^{10.25} M_\odot$  seem to actively form stars throughout their lifetime, with the maximal peak in their SFHs’ maximum SFR close to the time of observation, the galaxies containing a disk and bulge component seem to show a downward trend in their median SFHs, with disk-dominated galaxies peaking earlier on average than pure disks without a bulge, followed by a decline in SFR. This trend continues to bulge-dominated galaxies and spheroids, showing an evolution in timescales that can be tested with simulations implementing different models for quiescence. For each population, we consider the time since the SFH of each galaxy peaked ( $t_{\text{fall}}$ ) and use this distribution to quantify the timescale on which galaxy SFHs began their decline. For each morphologically distinct population, we find the median and 68% of the  $t_{\text{fall}}$  distribution for galaxies at  $0.5 < z < 1.0$  and  $10^{10} < M_* < 10^{10.5} M_\odot$ , given by the following.

1. Mergers:  $t_{\text{fall}}: 0.00_{+0.39}^{-0.00}$  Gyr.
2. Galaxies with spiral arms:  $t_{\text{fall}}: 0.60_{+1.54}^{-0.54}$  Gyr.
3. Disks:  $t_{\text{fall}} = 0.81_{+2.56}^{-0.80}$  Gyr.
4. Disk-dominated galaxies:  $t_{\text{fall}}: 0.70_{+2.73}^{-0.38}$  Gyr.
5. Bulge-dominated galaxies:  $t_{\text{fall}}: 2.15_{+3.07}^{-1.55}$  Gyr.
6. Spheroids:  $t_{\text{fall}}: 2.50_{+2.25}^{-1.60}$  Gyr.

In the absence of a bulge component, we also see that galaxies with spiral arms inhabit the high-stellar-mass, high-SFR portion of the SFR– $M_*$  plane, continuing to actively form stars until they lose rotational support or start forming bulges. We also see that mergers and interactions show a noticeable increase in recent SFR, over timescales within the last  $\sim 0.5$  Gyr of their SFHs. The timescales for these morphological transformations can be further constrained by determining the resolved SFHs of individual galaxies using IFU surveys like SDSS-IV MaNGA and CALIFA (Delgado et al. 2014; Belfiore et al. 2016).

## 6. Discussion

### 6.1. Improvements from Better Data Sets, Models, and Priors

Because nonparametric methods make no explicit assumption about the form of SFHs, they are only as good as the data being used for SED fitting. In this regard, there are three main avenues for data-driven improvement: better wavelength resolution, better wavelength coverage, and better S/N. Spectroscopy contains more information about stellar populations of different ages and metallicity, as compared to broadband photometry, but often suffers from wavelength-dependent flux calibration issues that need to be accounted for prior to fitting. Panchromatic SEDs allow us to test models of dust attenuation and re-emission to better constrain dust effects while estimating the SFHs of galaxies.

The SFH reconstructions we obtain are also subject to several modeling uncertainties: Stellar Population Synthesis models can introduce systematics into the mapping between physical parameters and observed photometry (Conroy & Gunn 2010; Han & Han 2019). Differences in dust models can introduce systematics into the measurement of recent SFR, which would then propagate into differences in the SFH. C. Pacifici et al. (2019, in preparation) compares the results from 14 different SED fitting codes applied to the same sample

of CANDELS/GOODS-S galaxies at  $z \sim 1$ . This allows us to examine the effects of intercode variability and model assumptions for dust, IMF, SFH, and SPS models, including the effects of binary populations. Additionally, Han & Han (2019) tested multiple SPS models, SFH assumptions and dust models using a comprehensive Bayesian formalism that allowed them to estimate the Bayesian evidence in comparing different models.

Finally, the choices of prior assumed during SED fitting are extremely influential in the estimates of physical parameters and their covariances. While we have tried to be agnostic about the priors in this work, it is important to note that an informative prior could be especially useful while fitting noisy, low-S/N data with limited wavelength coverage. Predictive checks could be put in place to ensure that the priors do not introduce significant biases into the estimates, or cause artificially tight correlations due to regression to the mean. These informative priors could be developed by studying the distributions of physical quantities at a particular epoch from a small subset of high-S/N observations, scaling relations, and mass functions, as well as semi-analytic or empirical models that encode the physics that lead to these observables, explicitly quantifying the covariance between star formation, chemical attenuation, and dust enrichment and destruction histories.

### 6.2. SFHs as a Probe to Higher Redshifts

The SFHs of galaxies allow us to probe the behavior of mass functions and scaling relations out to higher redshifts than is currently possible. At low to intermediate redshifts, this can be used as a consistency check, to ensure that the reconstructed SFHs are not biased due to noise or prior assumptions. At high redshifts, this can be a powerful tool to increase observational statistics and push measurements out to redshifts higher than those directly accessible through observations. Because we can only measure the SFH summed over all progenitors, it is important to consider the effects of mergers while propagating galaxies backward in time for periods longer than the typical merger timescale (Mantha et al. 2017; Duncan et al. 2019) at a given redshift and stellar mass.

Iyer et al. (2018) propagated galaxies backward in time along their SFHs in the form of trajectories in SFR– $M_*$  space, to probe the high-redshift low-stellar-mass regime of the SFR– $M_*$  correlation. They found that the projected correlation at intermediate redshifts matches the observed distribution well, and extended it by nearly two orders of magnitude, out to  $z \sim 6$ , where observations are extremely faint. C. Pacifici et al. (2019, in preparation) implements a validation test by reconstructing the stellar mass function using galaxies at lower redshifts and comparing them to the stellar mass function obtained through direct fits. Leja et al. (2019b) use SFHs to probe the cosmic SFRD using galaxies at low redshifts, finding that the apparent mismatch between the mass functions and star formation rate functions is alleviated by using nonparametric SFHs.

### 6.3. Galaxy Evolution Studies Enabled by SFH Reconstruction

The smooth, nonparametric SFHs obtained with the improved Dense Basis method offer a window into the pasts of different galaxy populations. While interesting in itself, this insight has the potential to be combined with a variety of

ancillary data to probe a wide range of previously inaccessible quantities, some of which we briefly describe below.

1. Higher-S/N observations or spectrophotometric data would help obtain better SFH constraints, allowing us to better constrain the number of major episodes of star formation and the timescales on which rejuvenation, starbursts, and quiescence occur at different epochs. While current observations studying the separation between multiple major episodes or transition timescales (this work; Abramson et al. 2016; Pandya et al. 2017) show a flat trend with mass and a linear one with the age of the universe, it is an interesting problem to understand the physical mechanisms responsible for this trend and the dispersion of  $\sim 0.25t_{\text{univ}}$  Gyr using simulations.
2. Spatially resolved SFHs computed using IFU data from surveys like SDSS-IV MaNGA and CALIFA allow us to better understand the correlation between the SFH and morphology and discriminate between inside-out versus outside-in scenarios for galaxy growth and quenching (Goddard et al. 2016), better examine the connection between the physical properties of individual regions within galaxies and their SFHs (Rowlands et al. 2018), and test scaling relations at different regimes (Hsieh et al. 2017). Care needs to be exercised in interpreting these results because we only see where the stellar populations are today. Additional kinematic information would help alleviate this problem to a certain extent.
3. Correlating SFHs with environment, size, kinematics, central density, and morphology in addition to stellar mass, SFR, and redshift could help build a unified picture of how galaxies evolve, with the SFHs providing a link between galaxy populations of different types and their earlier progenitors. Although this approach is similar to empirical models (Behroozi et al. 2018; Moster et al. 2018), it has the advantage of much richer observational constraints from the individual SFHs of galaxies. A comparison of SFH distributions between simulations and observations would allow us to qualify additional factors that are not directly accessible.
4. The Gaussian process–based parameterization can also be used as a general compression method for compressing and storing PDFs from all kinds of codes, similar to Malz et al. (2018).

#### 6.4. Caveats in SED Fitting and SFH Reconstruction

While we have performed an extensive range of validation tests (Section 3) to ensure that all the quantities reported in this work are robust, there are some caveats to keep in mind while extending the SED fitting to different data sets or the analysis beyond what is performed here.

1. SFHs are not mass accretion histories. The SFH is a record of when the stars present in a galaxy at the time of observation were formed, as opposed to the mass accretion history, which is a record of when those stars entered the galaxy. These two quantities are the same for stars formed in situ, but differ when the stars were brought in through mergers. This needs to be taken into account in certain kinds of analysis; for example, when using a mass- and redshift-dependent merger fraction to

correct for mass functions calculated by propagating galaxies backward in time along their SFHs.

2. Lack of sensitivity to the shortest timescales. The smooth SFHs reconstructed using SED fitting in this work cannot capture starbursts that can happen on extremely short timescales of  $\sim \mathcal{O}(10)$  Myr. While fitting galaxies from the semi-analytic model that contain such starbursts, we generally find that the overall stellar mass is well-recovered, but the starburst is smeared out over larger timescales, depending on when it occurred. This needs to be accounted for in the uncertainty budget; for example, while calculating the scatter along the SFR– $M_*$  correlation using galaxies propagated backward in time along their SFR– $M_*$  trajectories.
3. Non-uniform sensitivity to variations in SFH. SED fitting is more sensitive to recent star formation than it is to star formation older than a few Gyr. As we go back in time, our SFH reconstruction transitions from being likelihood-dominated to being prior-dominated; while we show that this does not cause biases in our SFH reconstruction in Section 3, it does mean that we are less sensitive to sharp variations in SFH at large lookback times than when closer to the time of observation.
4. Correlation with chemical enrichment histories: While we have considered the problem of estimating the SFHs of galaxies in this work, in practice they are highly correlated with the chemical enrichment histories of galaxies. Although metallicity is poorly constrained in our current observations, when working with higher-S/N data or spectra, the analysis should include a joint model for SFR( $t$ ) and  $Z(t)$ , which can be achieved through joint priors on the metallicity given by  $Z(M_*, \{t_x\})$ , informed using simulations.

## 7. Conclusions

Studying the SFHs of galaxies lets us better understand the timescales on which different physical processes shape galaxy growth. High-S/N multiwavelength observations from current and upcoming galaxy surveys make it possible to reconstruct the SFHs for large ensembles of galaxies with suitably sophisticated analysis techniques.

We update the Dense Basis SED-fitting method (Iyer & Gawiser 2017) using a flexible SFH parameterization described by the tuple  $(M_*, \text{SFR}, \{t_x\})$ , where  $M_*$  is the stellar mass, SFR is the star formation rate averaged over the past 100 Myr, and the set  $\{t_x\}$  contains the lookback times at which a galaxy formed  $N$  equally spaced quantiles of its stellar mass. These parameters, representing a set of integral constraints and SFHs corresponding to a particular tuple, are constructed using GPR in cumulative mass versus time, which creates smooth curves that satisfy these constraints and are completely independent of the choice of a functional form. We reconstruct the SFHs of galaxies with uncertainties using a brute-force Bayesian approach with a large pregrid of model SEDs. To make the method fully nonparametric, we determine  $N$  on an SED-to-SED basis using a BIC-based selection. Using the reconstructed SFHs and a peak-finding algorithm, we determine the number of major episodes of star formation in a galaxy’s past.

Our method provides the following advantages.

1. This method encodes a maximal amount of SFH information in a minimal number of parameters.

2. Being independent of the choice of a functional form, it does not suffer from the traditional biases associated with simple parametric assumptions for the SFH shape.
3. This method also circumvents the pitfalls associated with the traditional nonparametric approach of describing SFHs as fixed bins in lookback time with constant SFR within a bin, such as artifacts due to bin edges, and reduces uneven S/N distribution across different parameters.
4. The parameters used to describe SFHs are physically interpretable and allow easy comparison between different data sets from observations and simulations.
5. Informative priors can be constructed by studying these parameters in cosmological simulations, which can be used while fitting low-S/N data or SEDs with partial wavelength coverage.
6. The method is computationally fast, able to fit  $\sim 34$  galaxies/minute/core on a 2.9 GHz Intel processor, and is capable of being adapted to most data compression problems.

We apply the method to a sample of 48,791 galaxies across the five CANDELS fields with  $HST/WFC3 F160W < 25$  and  $0.5 < z < 3.0$ .

We use the reconstructed SFHs to study galaxy evolution across stellar mass and redshift, and to quantify the fraction of galaxies at each epoch that have multiple strong episodes of star formation. For the galaxies that show multiple strong episodes of star formation, we find that the timescale separating two peaks in the SFH is roughly constant with mass, and increases linearly with the age of the universe as  $t_{\text{peak-to-peak}} \sim 0.42_{-0.10}^{+0.15} t_{\text{univ}}$  Gyr. We also find that classifying galaxies by  $t_{50}$  is a robust way of selecting for star-forming galaxies at a given epoch.

Using the Kartaltepe et al. (2015) morphology catalog, we can examine the SFHs for subsets of galaxies with particular morphological features, finding the expected correlation between the SFHs of galaxies and morphological features. In addition, we quantify the timescale on which the SFH declines as a function of morphology, finding that this increases from  $\sim 0.60_{-0.54}^{+1.54}$  Gyr for galaxies with spiral arms to  $2.50_{-2.25}^{+1.60}$  Gyr for spheroids.

The SFH formalism presented here is broad in scope and can be incorporated into any SED fitting code, can be used to compress and store SFHs in simulations, and can be used as a common parameterization to compare SFHs across different observations and simulations.

The authors would like to thank the anonymous referee for a thorough perusal of the paper and several discerning suggestions that have helped improve the robustness of the method. The authors would like to thank Louis Abramson and Steven Finkelstein for their insightful comments and suggestions. K.I. would like to thank Boris Leistedt for a wonderful talk introducing Gaussian processes. The authors acknowledge Dritan Kodra, Jeff Newman, Steve Finkelstein, Adriano Fontana, Janine Pforr, Mara Salvato, Tommy Wiklund, and Stijn Wuyts for generating the photo- $z$  PDFs for the compilation of  $z_{\text{best}}$  in the v2 CANDELS photo- $z$  catalog, Guillermo Barro for compiling the GOODS-North photometric catalog, and Viraj Pandya for compiling the Santa Cruz semi-analytic model SFHs used in this work. K.I. and E.G. gratefully acknowledge support from Rutgers University. We gratefully

acknowledge the `george` (Ambikasaran et al. 2015; Foreman-Mackey 2015), `FSPS` (Conroy et al. 2009; Conroy & Gunn 2010; Foreman-Mackey et al. 2014), `corner` (Foreman-Mackey 2016), and `astropy` (Astropy Collaboration et al. 2013; Price-Whelan et al. 2018) packages for helping develop the code introduced in this paper. This work used resources from the Rutgers Discovery Informatics Institute, supported by Rutgers and the State of New Jersey. The Flatiron Institute is supported by the Simons Foundation. Support for Program number *HST*-AR-14564.001-A and GO-12060 was provided by NASA through a grant from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Incorporated, under NASA contract NAS5-26555.

## ORCID iDs

Kartheik G. Iyer  <https://orcid.org/0000-0001-9298-3523>  
 Eric Gawiser  <https://orcid.org/0000-0003-1530-8713>  
 Henry C. Ferguson  <https://orcid.org/0000-0001-7113-2738>  
 Jeyhan Kartaltepe  <https://orcid.org/0000-0001-9187-3605>  
 Anton M. Koekemoer  <https://orcid.org/0000-0002-6610-2048>  
 Camilla Pacifici  <https://orcid.org/0000-0003-4196-0617>

## References

- Abramson, L., Newman, A., Treu, T., et al. 2018, *AJ*, 156, 29  
 Abramson, L. E., Gladders, M. D., Dressler, A., et al. 2016, *ApJ*, 832, 7  
 Acquaviva, V., Gawiser, E., & Guaita, L. 2012, in IAU Symp. 284, The Spectral Energy Distribution of Galaxies—SED 2011, ed. R. J. Tuffs & C. C. Popescu (Cambridge: Cambridge Univ. Press), 42  
 Ambikasaran, S., Foreman-Mackey, D., Greengard, L., Hogg, D. W., & O’Neil, M. 2015, *ITPAM*, 38, 2  
 Anglés-Alcázar, D., Faucher-Giguère, C.-A., Quataert, E., et al. 2017, *MNRAS Letters*, 472, L109  
 Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, 558, A33  
 Barro, G., Perez-Gonzalez, P. G., Cava, A., et al. 2019, *ApJS*, in press  
 Behroozi, P., Wechsler, R., Hearin, A., & Conroy, C. 2018, *MNRAS*, submitted arXiv:1806.07893  
 Belfiore, F., Maiolino, R., Maraston, C., et al. 2016, *MNRAS*, 466, 2570  
 Bell, E. F., Zheng, X. Z., Papovich, C., et al. 2007, *ApJ*, 663, 834  
 Boogaard, L. A., Brinchmann, J., Bouché, N., et al. 2018, *A&A*, 619, A27  
 Boselli, A., & Gavazzi, G. 2006, *PASP*, 118, 517  
 Brinchmann, J., Charlot, S., White, S., et al. 2004, *MNRAS*, 351, 1151  
 Broussard, A., Gawiser, E., Iyer, K., et al. 2019, *ApJ*, 873, 74  
 Bruzual, G., & Charlot, S. 2003, *MNRAS*, 344, 1000  
 Bundy, K., Ellis, R. S., & Conselice, C. J. 2005, *ApJ*, 625, 621  
 Butt, S., & Brodrie, K. 1993, *Computers & Graphics*, 17, 55  
 Calzetti, D. 2001, *PASP*, 113, 1449  
 Caplar, N., & Tacchella, S. 2019, *MNRAS*, 487, 3845  
 Carnall, A., McLure, R., Dunlop, J., & Davé, R. 2018, *MNRAS*, 480, 4379  
 Carnall, A. C., Leja, J., Johnson, B. D., et al. 2019, *ApJ*, 873, 44  
 Chabrier, G. 2003, *PASP*, 115, 763  
 Chevallard, J., & Charlot, S. 2016, *MNRAS*, 462, 1415  
 Chiles, J.-P., & Delfiner, P. 2009, *Geostatistics: Modeling Spatial Uncertainty*, Vol. 497 (New York: Wiley)  
 Ciesla, L., Elbaz, D., & Fensch, J. 2017, *A&A*, 608, A41  
 Conroy, C., & Gunn, J. E. 2010, *ApJ*, 712, 833  
 Conroy, C., Gunn, J. E., & White, M. 2009, *ApJ*, 699, 486  
 Conselice, C. J. 2014, *ARA&A*, 52, 291  
 Da Cunha, E., Charlot, S., & Elbaz, D. 2008, *MNRAS*, 388, 1595  
 Daddi, E., Dickinson, M., Morrison, G., et al. 2007, *ApJ*, 670, 156  
 Davé, R. 2008, *MNRAS*, 385, 147  
 Davé, R., Thompson, R., & Hopkins, P. F. 2016, *MNRAS*, 462, 3265  
 De Boor, C., De Boor, C., Mathématicien, E.-U., De Boor, C., & De Boor, C. 1978, *A Practical Guide to Splines*, Vol. 27 (New York: Springer)  
 Delgado, R. G., Pérez, E., Fernandes, R. C., et al. 2014, *A&A*, 562, A47  
 De Lucia, G., & Blaizot, J. 2007, *MNRAS*, 375, 2

- Domínguez Sánchez, H., Pérez-González, P. G., Esquej, P., et al. 2016, *MNRAS*, **457**, 3743
- Donnari, M., Pillepich, A., Nelson, D., et al. 2019, *MNRAS*, **485**, 4817
- Dressler, A., & Abramson, L. 2014, in Proceedings of the International Astronomical Union, 10, 140
- Dressler, A., Kelson, D. D., Abramson, L. E., et al. 2016, *ApJ*, **833**, 251
- Duncan, K., Conselice, C. J., Mundy, C., et al. 2019, *ApJ*, **876**, 110
- Dye, S. 2008, *MNRAS*, **389**, 1293
- Elbaz, D., Daddi, E., Le Borgne, D., et al. 2007, *A&A*, **468**, 33
- Emami, N., Siana, B., Weisz, D. R., & Johnson, B. D. 2018, *ApJ*, in press arXiv:1809.06380
- Fang, J. J., Faber, S., Salim, S., Graves, G. J., & Rich, R. M. 2012, *ApJ*, **761**, 23
- Foreman-Mackey, D. 2015, George: Gaussian Process regression, Astrophysics Source Code Library, ascl: 1511.015
- Foreman-Mackey, D. 2016, *JOSS*, **1**, 24
- Foreman-Mackey, D., Sick, J., & Johnson, B. 2014, python-fsps: Python bindings to FSPS, v0.1.1, doi:10.5281/zenodo.12157
- Galametz, A., Grazian, A., Fontana, A., et al. 2013, *ApJS*, **206**, 10
- Gebhardt, K., Bender, R., Bower, G., et al. 2000, *ApJL*, **539**, L13
- Genel, S., Bryan, G. L., Springel, V., et al. 2019, *ApJ*, **871**, 21
- Goddard, D., Thomas, D., Maraston, C., et al. 2016, *MNRAS*, **466**, 4731
- Grogin, N. A., Kocevski, D. D., Faber, S., et al. 2011, *ApJS*, **197**, 35
- Guo, Y., Ferguson, H. C., Giavalisco, M., et al. 2013, *ApJS*, **207**, 24
- Guo, Y., Rafelski, M., Faber, S., et al. 2016, *ApJ*, **833**, 37
- Hahn, C., Starkeburg, T. K., Choi, E., et al. 2019, *ApJ*, **872**, 160
- Han, Y., & Han, Z. 2019, *ApJS*, **240**, 3
- Heavens, A. F., Jimenez, R., & Lahav, O. 2000, *MNRAS*, **317**, 965
- Hopkins, P. F. 2015, *MNRAS*, **450**, 53
- Hopkins, P. F., Cox, T. J., Kereš, D., & Hernquist, L. 2008, *ApJS*, **175**, 390
- Hopkins, P. F., Kereš, D., Oñorbe, J., et al. 2014, *MNRAS*, **445**, 581
- Hsieh, B., Lin, L., Lin, J., et al. 2017, *ApJL*, **851**, L24
- Inutsuka, S.-i., Inoue, T., Iwasaki, K., & Hosokawa, T. 2015, *A&A*, **580**, A49
- Ivezic, Z., Connolly, A. J., VanderPlas, J. T., & Gray, A. 2014, Statistics, Data Mining, and Machine Learning in Astronomy: A Practical Python Guide for the Analysis of Survey Data, Vol. 1 (Princeton, NJ: Princeton Univ. Press)
- Iyer, K., & Gawiser, E. 2017, *ApJ*, **838**, 127
- Iyer, K., Gawiser, E., Davé, R., et al. 2018, *ApJ*, **866**, 120
- Kartalteppe, J. S., Mozena, M., Kocevski, D., et al. 2015, *ApJS*, **221**, 11
- Kass, R. E., & Wasserman, L. 1995, *J. Am. Sta. Assoc.*, **90**, 928
- Kauffmann, G., Heckman, T. M., White, S. D., et al. 2003, *MNRAS*, **341**, 33
- Kelson, D. D. 2014, arXiv:1406.5191
- Klypin, A. A., Trujillo-Gomez, S., & Primack, J. 2011, *ApJ*, **740**, 102
- Koekemoer, A. M., Faber, S., Ferguson, H. C., et al. 2011, *ApJS*, **197**, 36
- Kovač, K., Lilly, S., Knobel, C., et al. 2010, *ApJ*, **718**, 86
- Krivoruchko, K. 2011, Spatial Statistical Data Analysis for GIS Users (Redlands, CA: Esri Press)
- Kurczynski, P., Gawiser, E., Acquaviva, V., et al. 2016, *ApJL*, **820**, L1
- Laidler, V. G., Papovich, C., Grogin, N. A., et al. 2007, *PASP*, **119**, 1325
- Lee, B., Giavalisco, M., Whitaker, K., et al. 2018, *ApJ*, **853**, 131
- Lee, S.-K., Ferguson, H. C., Somerville, R. S., Wiklund, T., & Giavalisco, M. 2010, *ApJ*, **725**, 1644
- Leistedt, B., & Hogg, D. W. 2017, *ApJ*, **838**, 5
- Leitner, S. N. 2012, *ApJ*, **745**, 149
- 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
- Levasseur, L. P., Hezaveh, Y. D., & Wechsler, R. H. 2017, *ApJL*, **850**, L7
- Liddle, A. R. 2007, *MNRAS*, **377**, L74
- Malz, A., Marshall, P., DeRose, J., et al. 2018, *AJ*, **156**, 35
- Mantha, K. B., McIntosh, D. H., Brennan, R., et al. 2017, *MNRAS*, **475**, 1549
- Matthee, J., & Schaye, J. 2019, *MNRAS*, **484**, 915
- Mobasher, B., Dahlen, T., Ferguson, H. C., et al. 2015, *ApJ*, **808**, 101
- Morishita, T., Abramson, L., Treu, T., et al. 2018, *ApJ*, **877**, 141
- Moster, B. P., Naab, T., & White, S. D. 2018, *MNRAS*, **477**, 1822
- Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013, *ApJ*, **777**, 18
- Muzzin, A., Van der Burg, R., McGee, S. L., et al. 2014, *ApJ*, **796**, 65
- Naab, T., & Ostriker, J. P. 2017, *ARA&A*, **55**, 59
- Nayyeri, H., Hemmati, S., Mobasher, B., et al. 2017, *ApJS*, **228**, 7
- Noeske, K., Weiner, B., Faber, S., et al. 2007, *ApJL*, **660**, L43
- Noll, S., Burgarella, D., Giovannoli, E., et al. 2009, *A&A*, **507**, 1793
- Ocvirk, P., Pichon, C., Lançon, A., & Thiébaud, E. 2006, *MNRAS*, **365**, 74
- Pacifici, C., Kassin, S. A., Weiner, B., Charlot, S., & Gardner, J. P. 2012, *ApJL*, **762**, L15
- Pacifici, C., Oh, S., Oh, K., Lee, J., & Yi, S. K. 2016, *ApJ*, **824**, 45
- Pandya, V., Brennan, R., Somerville, R. S., et al. 2017, *MNRAS*, **472**, 2054
- Peng, Y.-j., Lilly, S. J., Kovač, K., et al. 2010, *ApJ*, **721**, 193
- Pérez-González, P. G., Rieke, G. H., Villar, V., et al. 2008, *ApJ*, **675**, 234
- Porter, L. A., Somerville, R. S., Primack, J. R., & Johansson, P. H. 2014, *MNRAS*, **444**, 942
- Price-Whelan, A. M., Sipőcz, B. M., Günther, H. M., et al. 2018, *AJ*, **156**, 123
- Rasmussen, C. E., & Williams, C. K. 2006, Gaussian Processes for Machine Learning (Cambridge, MA: MIT Press)
- Reichardt, C., Jimenez, R., & Heavens, A. F. 2001, *MNRAS*, **327**, 849
- Rowlands, K., Heckman, T., Wild, V., et al. 2018, *MNRAS*, **480**, 2544
- Schwarz, G. 1978, *AnSta*, **6**, 461
- Searle, L., Sargent, W., & Bagnuolo, W. 1973, *ApJ*, **179**, 427
- Seeger, M. 2004, *IJNS*, **14**, 69
- Shivaei, I., Reddy, N. A., Steidel, C. C., & Shapley, A. E. 2015, *ApJ*, **804**, 149
- Smith, D. J., & Hayward, C. C. 2015, *MNRAS*, **453**, 1597
- Somerville, R. S., & Davé, R. 2015, *ARA&A*, **53**, 51
- Somerville, R. S., Hopkins, P. F., Cox, T. J., Robertson, B. E., & Hernquist, L. 2008, *MNRAS*, **391**, 481
- Somerville, R. S., Popping, G., & Trager, S. C. 2015, *MNRAS*, **453**, 4337
- Sparre, M., Hayward, C. C., Springel, V., et al. 2015, *MNRAS*, **447**, 3548
- Stefanon, M., Yan, H., Mobasher, B., et al. 2017, *ApJS*, **229**, 32
- Szydlowski, M., Krawiec, A., Kurek, A., & Kamionka, M. 2015, *EPJ C*, **75**, 5
- Tacchella, S., Bose, S., Conroy, C., Eisenstein, D. J., & Johnson, B. D. 2018, *ApJ*, **868**, 92
- Tacchella, S., Dekel, A., Carollo, C. M., et al. 2016, *MNRAS*, **457**, 2790
- Tinsley, B. M. 1968, *ApJ*, **151**, 547
- Tojeiro, R., Heavens, A. F., Jimenez, R., & Panter, B. 2007, *MNRAS*, **381**, 1252
- Torrey, P., Vogelsberger, M., Hernquist, L., et al. 2018, *MNRAS*, **477**, L16
- Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, *ApJ*, **613**, 898
- Tully, R. B., & Fisher, J. R. 1977, *A&A*, **54**, 661
- Wackernagel, H. 2003, Multivariate Geostatistics: An Introduction with Applications (Berlin: Springer)
- Weinberger, R., Springel, V., Pakmor, R., et al. 2018, *MNRAS*, **479**, 4056
- Weisz, D. R., Dalcanton, J. J., Williams, B. F., et al. 2011, *ApJ*, **739**, 5
- Whitaker, K. E., Franx, M., Bezanson, R., et al. 2015, *ApJL*, **811**, L12
- Whitaker, K. E., Franx, M., Leja, J., et al. 2014, *ApJ*, **795**, 104
- White, S. D., & Rees, M. J. 1978, *MNRAS*, **183**, 341
- Wright, A. C., Brooks, A. M., Weisz, D. R., & Christensen, C. R. 2019, *MNRAS*, **482**, 1176
- Zolotov, A., Dekel, A., Mandelker, N., et al. 2015, *MNRAS*, **450**, 2327