The Astrophysical Journal, 877:140 (21pp), 2019 June 1 https://doi.org/10.3847/1538-4357/ab1d5a

An Older, More Quiescent Universe from Panchromatic SED Fitting of the 3D-*HST* Survey

Joel Leja[1,8], Benjamin D. Johnson[1], Charlie Conroy[1], Pieter van Dokkum[2], Joshua S. Speagle[1], Gabriel Brammer[3], Ivelina Momcheva[4], Rosalind Skelton[5], Katherine E. Whitaker[6], Marijn Franx[7], and Erica J. Nelson[1]
[1] Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA; joel.leja@cfa.harvard.edu
[2] Department of Astronomy, Yale University, New Haven, CT 06511, USA
[3] University of Copenhagen, Nørregade 10, DK-1165 København, Denmark
[4] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21211, USA
[5] South African Astronomical Observatory, PO Box 9, Observatory, Cape Town 7935, South Africa
[6] Department of Physics, University of Connecticut, 2152 Hillside Road, Unit 3046, Storrs, CT 06269, USA
[7] Leiden Observatory, Leiden University, NL-2300 RA Leiden, The Netherlands
Received 2018 December 15; revised 2019 April 16; accepted 2019 April 26; published 2019 June 5

Abstract

Galaxy observations are influenced by many physical parameters: stellar masses, star formation rates (SFRs), star formation histories (SFHs), metallicities, dust, black hole activity, and more. As a result, inferring accurate physical parameters requires high-dimensional models that capture or marginalize over this complexity. Here we reassess inferences of galaxy stellar masses and SFRs using the 14-parameter physical model `Prospector-`α built in the `Prospector` Bayesian inference framework. We fit the photometry of 58,461 galaxies from the 3D-*HST* catalogs at $0.5 < z < 2.5$. The resulting stellar masses are ~0.1–0.3 dex larger than the fiducial masses while remaining consistent with dynamical constraints. This change is primarily due to the systematically older SFHs inferred with `Prospector`. The SFRs are ~0.1–1+ dex lower than UV+IR SFRs, with the largest offsets caused by emission from "old" ($t > 100$ Myr) stars. These new inferences lower the observed cosmic SFR density by ~0.2 dex and increase the observed stellar mass growth by ~0.1 dex, finally bringing these two quantities into agreement and implying an older, more quiescent universe than found by previous studies at these redshifts. We corroborate these results by showing that the `Prospector-`α SFHs are both more physically realistic and much better predictors of the evolution of the stellar mass function. Finally, we highlight examples of observational data that can break degeneracies in the current model; these observations can be incorporated into priors in future models to produce new and more accurate physical parameters.

Key words: galaxies: evolution – galaxies: fundamental parameters – galaxies: star formation

1. Introduction

The modern approach to galaxy spectral energy distributions (SEDs) with stellar population synthesis (SPS) models was pioneered by Sawicki & Yee (1998). These authors fit the rest-frame UV–optical broadband photometry of Lyman break galaxies with an exponentially declining τ-model star formation history (SFH), allowing variation in the start time, the duration of star formation (τ), the stellar metallicity, and a reddening factor. This basic formula of a four- to five-parameter model covering a simple functional SFH, a dust attenuation vector, and perhaps stellar metallicity has remained a robust feature in the literature over the past two decades (Brinchmann & Ellis 2000; Papovich et al. 2001; Shapley et al. 2001; Ilbert et al. 2006; Salim et al. 2007; Kriek et al. 2009; Maraston et al. 2010; Acquaviva et al. 2011; Skelton et al. 2014; Salmon et al. 2015).

These fits have been extraordinarily successful, as they provide a physical map from galaxy photometry to physical properties. The most widely used parameters from such fits are star formation rates (SFRs) and stellar masses (e.g., Shapley et al. 2001; Hopkins & Beacom 2006; Genel et al. 2014; Madau & Dickinson 2014; Speagle et al. 2014; Behroozi et al. 2019). Stellar masses are considered particularly robust owing to fortuitous degeneracies between dust, age, and metallicity, which means that there is a fairly tight relation between *M*/*L* ratio and color (Bell & de Jong 2001).

However, there are known uncertainties and systematic errors in this approach. There has remained a persistent and systematic factor-of-two uncertainty in stellar masses derived from SED-fitting codes (Papovich et al. 2001; Marchesini et al. 2009; Wuyts et al. 2009; Behroozi et al. 2010; Pforr et al. 2012; Conroy 2013; Mitchell et al. 2013; Leja et al. 2015, 2019; Mobasher et al. 2015; Santini et al. 2015; Tomczak et al. 2016; Carnall et al. 2018), while SFRs obtained via either monochromatic indicators or SED modeling are subject to similar 0.3–0.5 dex systematics (Wuyts et al. 2011a; Speagle et al. 2014; Carnall et al. 2018; Leja et al. 2018). These systematics are caused by a combination of (1) fundamental uncertainties in the input physics such as dust models, stellar evolution, initial mass function (IMF), and stellar spectral libraries and (2) observations that are at best weakly informative about the complexities of extragalactic stellar populations, resulting in strong model degeneracies. Examples of specific issues include differences in the underlying physics of SPS models (~0.1–0.2 dex), degeneracies from fundamental limitations such as the "outshining" of old stellar populations by young stars, the relative similarity of old stellar populations, and the age–dust–metallicity degeneracy (for a more complete list, see the review by Conroy 2013 and discussion therein). Due to the many confounding factors, solving any one of these problems in isolation is challenging and requires very carefully

---

[8] NSF Astronomy and Astrophysics Postdoctoral Fellow.

designed experiments (e.g., measuring contribution of TP-AGB stars to the near-IR fluxes; Kriek et al. 2010). As a result, the conventional wisdom has been that there is a nigh-unbreakable factor-of-two error in SED-fitting outputs. This has created little incentive to improve on the basic SED-fitting approach presented in Sawicki & Yee (1998), which is likely related to the persistence of this four- to five-parameter framework in the literature.

Fortunately, many big-picture questions in galaxy evolution are on order-of-magnitude scales and relatively insensitive to uncertainties at the factor-of-two level. For example, the cosmic SFR density is now known to peak at $z \sim 2$ (Madau & Dickinson 2014), the amount of stellar mass in the universe has increased by a factor of ~ 4 since $z \sim 2$ (Madau & Dickinson 2014), and galaxies likely reionized the universe around $z \sim 7$ (Schmidt et al. 2014; Mason et al. 2018).

However, our understanding of many other key aspects of galaxy formation is sensitive to factor-of-two systematics in stellar mass, SFRs, and other SED-fitting parameters. Massive galaxies are thought to approximately double their stellar mass from $z = 2$ to the present (van Dokkum et al. 2010; Patel et al. 2013b), while Milky Way–mass galaxies grow their mass by a factor of ~ 10 (Patel et al. 2013a; van Dokkum et al. 2013; Papovich et al. 2015). Both star-forming and quiescent galaxies approximately double their size at a fixed stellar mass from $z = 2.75$ (van der Wel et al. 2014). The stellar mass–metallicity relationship most likely evolves at fixed mass by a factor of ~ 2 from $z \sim 2$ to the present in observations (Erb et al. 2006a) and simulations (Torrey et al. 2019). A fundamental factor-of-two uncertainty in stellar mass means that even well-measured dynamical masses cannot be used to constrain the dark matter fraction in the inner regions of a galaxy (Cappellari et al. 2012; van de Sande et al. 2015; Wuyts et al. 2016). The slope of the star-forming sequence is quite sensitive to factor-of-two changes (e.g., Shivaei et al. 2017), meaning that relatively small changes in this slope can cause large changes in inferred galaxy formation histories (Leitner 2012; Leja et al. 2015) or that gas depletion times are no longer constant (Genzel et al. 2015). Systematic factor-of-two changes in SED-derived parameters can invalidate or inalterably change any or all of these conclusions. This presents a strong motivation to break the "factor-of-two barrier," the same motivation that has inspired our new approach to galaxy SED fitting.

Fortunately, many of the model improvements needed for this work have seen significant improvement over the past several decades. MAGPHYS was the first code to use energy balance to tie together UV–NIR and MIR–FIR photometry into a single physical model (da Cunha et al. 2008). More complex and more flexible SFH parameterizations have been explored, starting with SFH libraries with random bursts superimposed (Kauffmann et al. 2003; Gallazzi et al. 2005; da Cunha et al. 2008), to fits using multiple parametric SFHs (Iyer & Gawiser 2017; Lee et al. 2018), to nonparametric piecewise-constant SFHs (Cid Fernandes et al. 2005; Ocvirk et al. 2006; Tojeiro et al. 2007; Leja et al. 2017), to libraries of SFHs from simulations (Finlator et al. 2007; Pacifici et al. 2012). Spatially complex dust attenuation models have been developed that include extra attenuation toward younger star-forming regions (Charlot & Fall 2000) and flexible attenuation curves (Noll et al. 2009; Salmon et al. 2016; Leja et al. 2017; Salim et al. 2018). Emission from central active galactic nuclei (AGNs) is now built into many SED-fitting models (Berta et al. 2013; Ciesla et al. 2015; Calistro Rivera et al. 2016; Leja et al. 2018). Including the effect of nebular emission using photoionization models such as CLOUDY (Ferland et al. 1998, 2013) and MAPPINGS III (Groves et al. 2004) has become standard practice. Large uncertainties in the IR contribution of TP-AGB stars have largely been resolved (Maraston et al. 2006; Kriek et al. 2010), though other fundamental uncertainties in SPS techniques remain (e.g., the effect of binaries and rotation on the ionizing flux production rates of massive stars; Choi et al. 2017).

These new model components necessitate more robust statistical frameworks to properly constrain them. Bayesian forward-modeling techniques pioneered by Kauffmann et al. (2003), Burgarella et al. (2005), and Salim et al. (2007) help to constrain the complex, correlated parameter uncertainties typically present in galaxy models. Classic grid-based models grow exponentially in size with model dimensionality, but gridless "on-the-fly" models combined with Markov chain Monte Carlo (MCMC) algorithms can efficiently explore high-dimensional ($N \gtrsim 7$) spaces (Chevallard & Charlot 2016; Leja et al. 2017; Carnall et al. 2018). The computational time necessary for on-the-fly model exploration is both less expensive and more readily available than ever before.

By combining many of these advances into a single consistent framework, it may be possible to finally break the factor-of-two accuracy barrier in galaxy SED modeling. Here we take the first step toward this goal with the `Prospector-`α physical model built within the `Prospector` inference framework. `Prospector-`α has been cross-calibrated by fitting broadband photometry and using the posteriors to predict independent spectroscopic and spatially resolved data as an external check (Leja et al. 2017, 2018). These checks ensure that SED-fitting results are consistent with the overall picture of galaxy formation; given the lack of "ground truth" in SED modeling, such an approach is necessary to ensure accurate results. This necessitates an iterative cycle of refining the model, fitting new data, performing new predictive checks, and further refining the model. These new data could include large catalogs of photometry at longer wavelengths from, e.g., ALMA or *Herschel*, or intermediate-redshift information-rich spectroscopic surveys such as MOSDEF or KMOS-3D (Kriek et al. 2015; Wisnioski et al. 2015). This approach sets us on a long path, but it is the best path available to move the field forward.

This model is fit to the 3D-*HST* photometric catalogs. These are ideal data to investigate the population-wide 0.3 dex systematic errors in SED fitting: they provide rest-frame UV–IR photometry for $\sim$180,000 galaxies across $0.5 < z < 2.5$ and are complete in stellar mass down to $\sim 10^9\,M_\odot$ at $z = 2$ (Tal et al. 2014).

Section 2 describes the 3D-*HST* catalogs and how they are fit. Section 3 describes the SED model that is fit to these photometry. Section 4 details how the `Prospector-`α masses and SFRs differ from previous estimates. Section 5 performs model cross-validation tests to explore the accuracy of the inferred parameters and also shows the change in the cosmic star formation rate density (SFRD) as a result of the new measurements. The results and next steps are discussed in Section 6, and the conclusion is presented in Section 7. This work is done with a Chabrier (2003) IMF and a WMAP9 cosmology (Hinshaw et al. 2013). Unless otherwise noted, all parameters are reported as the median of the posterior probability distribution function (pdf).

2. Sample and Data

Galaxies are selected from the 3D-*HST* photometric catalogs (Skelton et al. 2014). The 3D-*HST* catalogs consist of state-of-the-art PSF-matched UV–IR photometry for hundreds of thousands of distant galaxies, covering ∼900 arcmin^2 in five well-studied extragalactic fields. Galaxies are identified in deep near-infrared *Hubble Space Telescope* (*HST*) imaging from the CANDELS survey (Grogin et al. 2011; Koekemoer et al. 2011) and have between 17 (the UDS field) and 44 (the COSMOS field) photometric bands spanning a range of 0.3–8 μm in the observed frame. The photometry is supplemented by *Spitzer*/MIPS 24 μm fluxes from Whitaker et al. (2014). The MIPS 24 μm coverage is critical because the rest-frame MIR wavelengths are dominated by warm dust emission, a key empirical proxy for obscured star formation (Kennicutt 1998). Obscured star formation is the dominant form of star formation for massive galaxies in this redshift range (Whitaker et al. 2017).

The 3D-*HST* catalogs contain additional stellar population parameters, including stellar masses from FAST (Kriek et al. 2009) and SFR$_{\rm UV+IR}$ (Whitaker et al. 2014). In this work these parameters are referred to as the 3D-*HST* catalog masses and SFRs. The photometry is complete in stellar mass to at least $M_* = 10^{9.3} M_\odot$ between $0.5 < z < 2.5$ (Tal et al. 2014). Redshifts are taken from, in order of reliability, (1) ground-based spectroscopic redshifts, (2) near-infrared grism redshifts from the 3D-*HST* survey (Momcheva et al. 2016), and (3) photometric redshifts from EAZY (Brammer et al. 2008; Skelton et al. 2014).

2.1. Sample Selection

There are 176,146 galaxies in v4.1 of the 3D-*HST* catalogs with usable photometry and derived stellar population parameters from FAST (Kriek et al. 2009). Due to computational constraints, we do not fit the entire sample in this work. Here we describe the subsample of galaxies that are fit. We also calculate the fraction of stellar mass density and SFR density (SFRD) covered by this sample in order to put measurements of the SFRD in Section 5.2 in proper context.

We fit all galaxies above the FAST stellar mass completeness limit from Tal et al. (2014) between $0.5 < z < 2.5$ that have usable photometry (i.e., 3D-*HST* *use_phot* = 1). We include a small fraction of galaxies that are below the mass limits but have high-quality data according to the following criteria:

1. S/N(F160W) > 10,
2. $0.5 < z < 2.5$,
3. $\sigma_z < 0.25$,
4. 3D-*HST* *use_phot* = 1.

These cuts result in 58,461 galaxies, of which 2702 (5%) have measured $z_{\rm spec}$, 12,513 (21%) use $z_{\rm grism}$, and the remaining 43,246 (74%) use $z_{\rm phot}$. The target sample is ∼33% of the total 3D-*HST* catalog by number but covers the majority of the observed SFR density ($\gtrsim$74%) and the stellar mass density ($\gtrsim$95%) at $0.5 < z < 2.5$ (Figure 1).

This subsample has reliable photometry and high signal-to-noise ratio (S/N) in the detection bands where it is most efficient to fit the computationally intensive `Prospector-`α model. The higher-S/N data provide stronger parameter constraints. Additionally, the redshift quality cuts ensure that systematic errors due to redshift uncertainties are minimized (future prospects for propagating redshift uncertainties to the SED parameters are discussed in Section 6.2.1). The galaxies removed by these cuts thus have either uncertain photometry, uncertain redshifts, or both.

The price of creating a computationally tractable sample is completeness: not every galaxy in the 3D-*HST* catalogs has an associated `Prospector` fit. The completeness of the target sample in FAST stellar mass and SFR$_{\rm UV+IR}$ is shown in Figure 1. Galaxies in the 3D-*HST* photometric catalog with *use_phot* = 1 are taken as the master sample against which this completeness is inferred. The fraction of the total stellar mass and total SFR covered by the target sample in each redshift window is indicated in the upper right corner of each panel. Here 95%–100% of the total stellar mass and 74%–91% of the total SFR is covered by our target sample.

In some cases, the incompleteness due to imaging depth becomes comparable to the incompleteness due to the subsampling of the catalog. The 90% completeness in FAST stellar mass is taken from Tal et al. (2014) and is derived by comparing object detection rates in the CANDELS deep fields with a recombined subset of the exposures that reach the depth of the CANDELS wide fields. The completeness in SFR$_{\rm UV+IR}$ is taken as the 3σ 24 μm depth calculated in Whitaker et al. (2014) and represents where the observable constraint on IR SFRs starts to become unreliable.

2.2. Treatment of Photometric Zero-points

The 3D-*HST* team self-consistently rederives zero-points for each instrument and filter. This is necessary to bring data from many telescopes and instruments onto a common flux scale. This procedure is described in detail in the appendix of Skelton et al. (2014). In brief, every galaxy is fit by the photometric redshift code EAZY, and the systematic residuals between the EAZY templates and the observed photometry are tabulated. In general, the systematic residuals are caused by a combination of template mismatch and zero-point errors. These two effects can be distinguished with sufficient quantities of high-quality data, as template mismatch occurs in the *rest* frame, while zero-point errors are in the *observed* frame. The resulting derived zero-point errors are used to correct the raw 3D-*HST* photometric fluxes to the fluxes reported in the catalog.

However, this process is imperfect: the "edges" of the wavelength coverage (IRAC 4 and *U* band) are more poorly calibrated, and effects such as redshift-dependent template mismatch may also be folded into the derived zero-point offsets. To avoid potentially imprinting any systematic offsets from this process into the photometry, we add the zero-point correction for each band of photometry to the flux errors in quadrature. This effect varies from 0% to 28% of the total flux, depending on the photometric band.

The *HST* zero-points are considerably more stable than the other bands and are therefore treated differently. For *HST* bands the zero-point corrections derived by the 3D-*HST* team are removed (these are typically near zero, though they can be up to 8% of the total flux), and no inflation of photometric errors is performed.

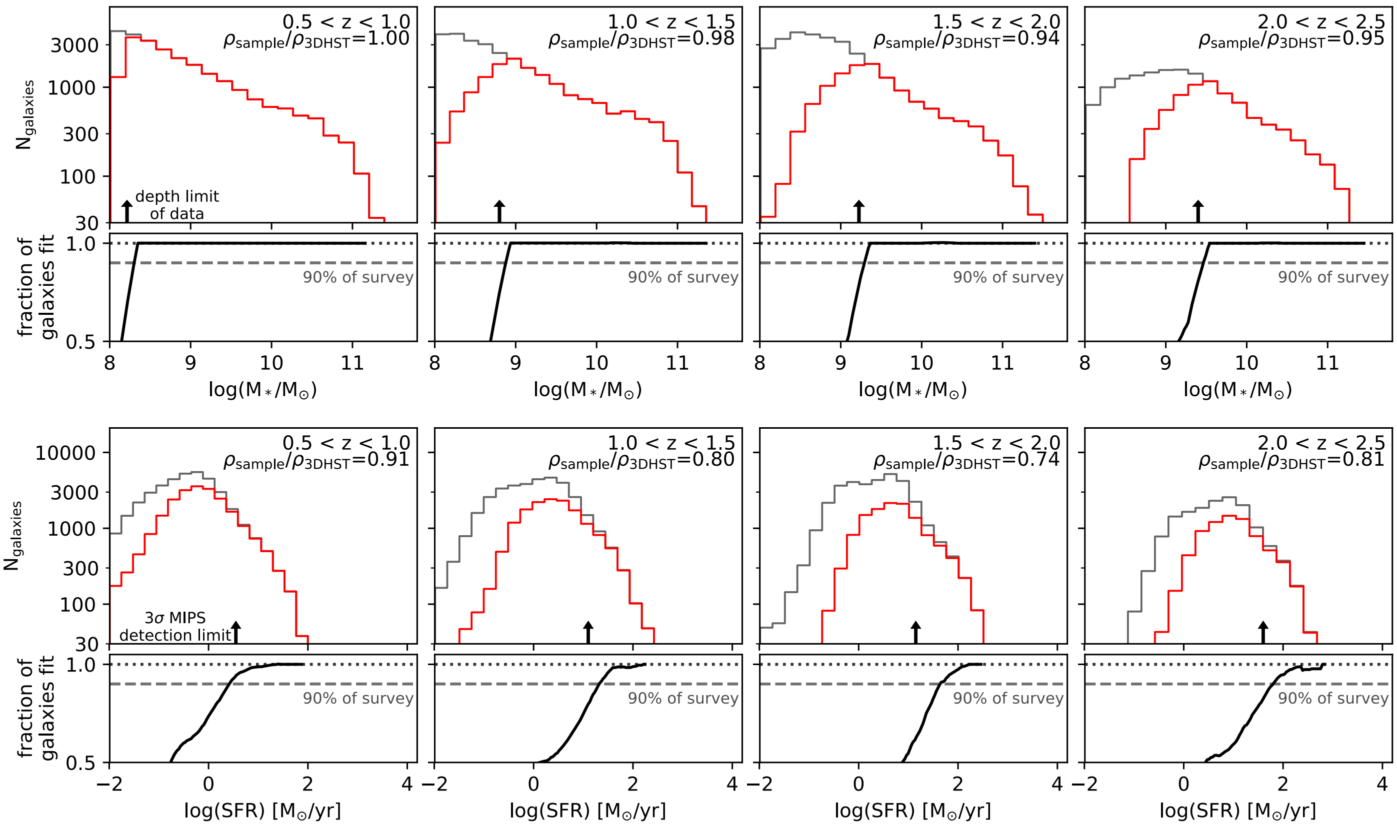

Figure 1. Sample selection completeness in stellar mass and SFR. The black lines in the histograms represent the number of galaxies in the full 3D-*HST* photometric catalogs, while the red lines represent the subset fit with the `Prospector`-α model. Stellar masses come from FAST, and the SFR shows $\mathrm{SFR}_{\mathrm{UV+IR}}$. The fraction of the (stellar mass density/SFRD) measured by the target sample is indicated in the upper right corner of each panel, where the total is taken to be the full 3D-*HST* sample in that redshift range. The 95% completeness limit is marked in gray.

After this process, a 5% minimum error is enforced for each band of photometry to allow for systematic errors in the physical models for stellar, gas, and dust emission.

3. SED Modeling

3.1. The `Prospector`-α Physical Model

We use the `Prospector` inference framework (Johnson & Leja 2017; Leja et al. 2017) to construct a galaxy SED model. `Prospector` adopts a Bayesian approach to forward-modeling galaxy SEDs.

`Prospector` uses the Flexible Stellar Population Synthesis (`FSPS`) code (Conroy et al. 2009) to generate the underlying physical model and `python-fsps` (Foreman-Mackey et al. 2014) to interface with `FSPS` in python. The physical model uses the MIST stellar evolutionary tracks and isochrones (Choi et al. 2016; Dotter 2016) based on MESA, an open-source stellar evolution package (Paxton et al. 2011, 2013, 2015).

Notably, MIST models include stellar rotation. This has several salient effects on massive star evolution: (i) it channels additional fuel into the stellar core; (ii) it causes distortion in the shape of the star, causing the poles to be hotter than the equator ("gravity darkening"); and (iii) it boosts the effect of mass loss. The net result is hotter, brighter, and longer-lived (by between 5% and 20%) massive stars (Choi et al. 2016). This in turn causes higher UV and ionizing photon production in stellar populations between 0 and 20 Myr (Choi et al. 2017), especially at subsolar metallicities. We note that models for rotation in stars are sensitive to implementation details and their predictions vary substantially; for example, the Geneva rotation model (Ekström et al. 2012) predicts both hotter and brighter stars than the MIST models adopted here (Figure 1 of Gossage et al. 2018). These models are currently difficult to test owing to both the relative lack of nearby star clusters with populated main-sequence turnoffs and the ongoing debate about the similar effects of age spread and rotation (Bastian & de Mink 2009; Goudfrooij et al. 2014; Piatti & Cole 2017; Gossage et al. 2018). While the physical mechanisms are distinct, the net effect on galaxy SEDs is similar to but weaker than the effect of stellar binaries (Eldridge et al. 2017).

In this study, we use an adapted version of the `Prospector`-α model framework described in Leja et al. (2017, 2018). The `Prospector`-α model includes a nonparametric SFH, a two-component dust attenuation model with a flexible attenuation curve, variable stellar metallicity, and dust emission powered via energy balance. Nebular line and continuum emission is generated self-consistently through use of CLOUDY (Ferland et al. 2013) model grids from Byler et al. (2017). Extensive calibration and testing of this model has been performed on local galaxies (Leja et al. 2017, 2018).

We make multiple changes to the `Prospector`-α model in order to reflect the different physics of galaxies at higher redshifts and to tailor the model more closely to the wavelength coverage and S/N of the 3D-*HST* photometry. The full set of priors and parameter ranges for the adjusted 14-parameter `Prospector`-α model are shown in Table 1. The salient changes are described below.

Table 1
Free Parameters and Their Associated Priors for the `Prospector-`α Physical Model

Parameter	Description	Prior
$\log(M/M_\odot)$	Total mass formed	Uniform: min = 7, max = 12.5
$\log(Z/Z_\odot)$	Stellar metallicity	Clipped normal: min = −1.98, max = 0.19, mean and σ from the Gallazzi et al. (2005) mass–metallicity relationship (see Section 3.1)
SFR ratios	Ratio of the SFRs in adjacent bins of the seven-bin nonparametric SFH (six parameters total)	Student's t-distribution with $\sigma = 0.3$ and $\nu = 2$
$\hat{\tau}_{\lambda,2}$	Diffuse dust optical depth	Clipped normal: min = 0, max = 4, mean = 0.3, $\sigma = 1$
$\hat{\tau}_{\lambda,1}$	Birth-cloud dust optical depth	Clipped normal in $(\hat{\tau}_{\lambda,1}/\hat{\tau}_{\lambda,2})$: min = 0, max = 2, mean = 1, $\sigma = 0.3$
n	Power-law modifier to the shape of the Calzetti et al. (2000) attenuation curve	Uniform: min = −1, max = 0.4
$\log(Z_{\rm gas}/Z_\odot)$	Gas-phase metallicity	Uniform: min = −2, max = 0.5
$f_{\rm AGN}$	AGN luminosity as a fraction of the galaxy bolometric luminosity	Log-uniform: min = 10^{-5}, max = 3
$\tau_{\rm AGN}$	Optical depth of AGN torus dust	Log-uniform: min = 5, max = 150

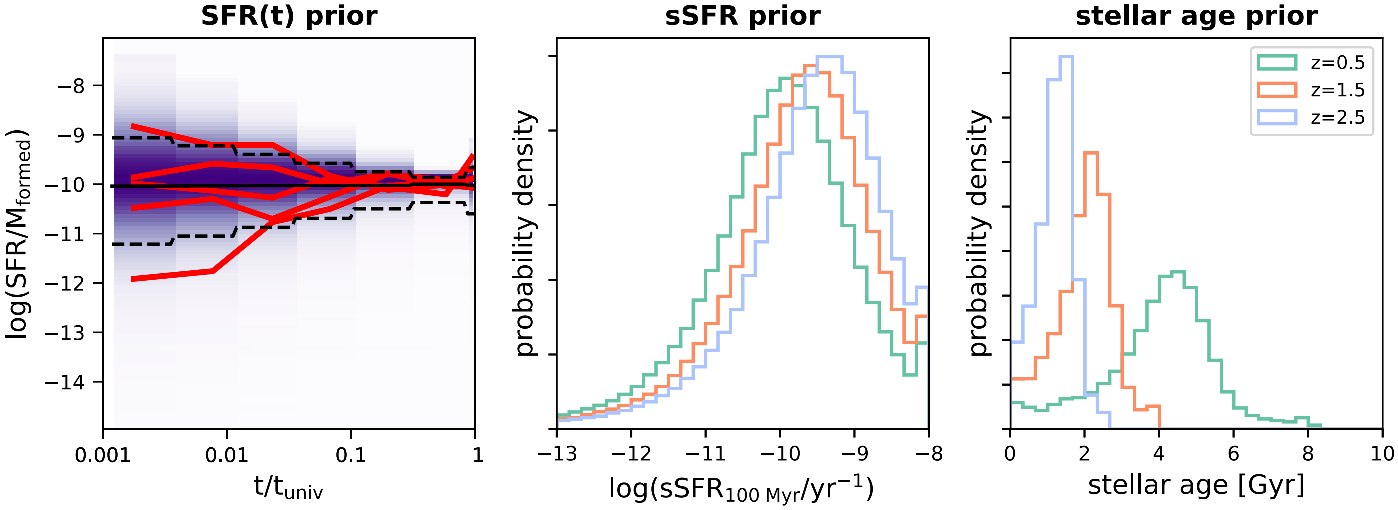

Figure 2. Continuity SFH prior adopted in the `Prospector-`α model. The left panel shows the prior density for SFR(t), with five random draws from the prior illustrated in red. The solid black line shows the median, while the dashed black lines show the 16th and 84th percentiles. The middle and right panels show the prior in sSFR(100 Myr) and mass-weighted age as a function of redshift, respectively. See Leja et al. (2019) for further details.

Nonparametric SFH prior: the continuity prior described in Leja et al. (2019) is taken as the prior for the nonparametric SFR(t). In brief, this prior weights against sharp transitions in SFR(t), similar to the regularization schemes from Ocvirk et al. (2006) and Tojeiro et al. (2007). The prior is tuned to allow similar transitions in SFR(t) to those of galaxies in the Illustris hydrodynamical simulations (Torrey et al. 2014; Vogelsberger et al. 2014a, 2014b; Diemer et al. 2017), though it is deliberately set to encompass broader behavior than is seen in these simulations. The resulting prior probability density for SFR(t), mass-weighted age, and sSFR$_{100\ \rm Myr}$ is shown in Figure 2.

Spacing of the nonparametric SFH bins: Seven time bins are used in the nonparametric SFH model. The bins are specified in lookback time. Two bins are fixed at 0–30 Myr and 30–100 Myr to capture variations in the recent SFH of galaxies. A third bin is placed at ($0.85t_{\rm univ}$–$t_{\rm univ}$), where $t_{\rm univ}$ is the age of the universe at the observed redshift, to model a "maximally old" population. The remaining four bins are spaced equally in logarithmic time between 100 Myr and $0.85t_{\rm univ}$.

Stellar mass–stellar metallicity prior: A single stellar metallicity is fit for all stars in a galaxy. A modified version of the stellar mass–stellar metallicity relationship from $z = 0$ Sloan Digital Sky Survey data (Gallazzi et al. 2005) is adopted as a prior. The relationship is modeled as a clipped normal distribution with limits of $-1.98 < \log(Z/Z_\odot) < 0.19$ set by the range of the MIST stellar evolution tracks. The standard deviation is taken as the 84th–16th percentile range from the Gallazzi et al. (2005) $z = 0$ relationship, i.e., twice the observed standard deviation of the $z = 0$ relationship. This wider relationship is adopted to allow potential redshift evolution in the stellar mass–stellar metallicity relationship.

A fixed IR SED: The rest-frame mid-infrared is poorly sampled by the 3D-*HST* photometric catalog, as the reddest two filters are *Spitzer*/IRAC channel 4 (7.8 μm) and *Spitzer*/MIPS 24 μm. This results in poor constraints on the shape of the IR SED (rest frame ∼4–1000 μm). Accordingly, we fix the shape of the IR SED in `Prospector-`α such that the *Spitzer*/MIPS 24 μm to $L_{\rm IR}$(8–1000 μm) conversion approximates that of the log-average of the Dale & Helou (2002) templates (Wuyts et al. 2008). Wuyts et al. (2011a) show that this luminosity-independent conversion produces $L_{\rm IR}$ estimates that are in agreement with observed *Herschel*/PACS photometry, though with significant scatter. Additionally, this choice of IR SED follows Whitaker et al. (2014), which facilitates direct comparisons with SFR$_{\rm UV+IR}$ from the 3D-*HST* catalog. Hot dust emission powered by an AGN of variable strength is also permitted in the `Prospector-`α model (Leja et al. 2018)—notably, this energy balance is performed separately from the rest of the IR SED, which is powered solely by stellar emission. Future potential for a more flexible IR SED model in `Prospector-`α is discussed in Section 6.2.2.

Altered nebular physics: Observations suggest that the gas in star-forming galaxies at higher redshifts experiences more extreme ionizing radiation fields and has metallicity abundances that may differ significantly from their stellar abundances (Shapley et al. 2015; Steidel et al. 2016). Accordingly, the ionization parameter for the nebular emission model is raised from $\log(U) = -2$ to $\log(U) = -1$, and gas-phase metallicity is decoupled from the stellar metallicity and allowed to vary between $-2 < \log(Z/Z_{\odot}) < 0.5$. This is a nuisance parameter for the majority of galaxies, as it typically is very poorly constrained by the photometry, though it can be important for very blue galaxies with high sSFRs (Cohn et al. 2018).

3.2. Posterior Sampling

The posteriors are sampled with the dynamic nested sampling code `dynesty` (Speagle 2019).[9] Nested sampling has a number of desirable properties over standard MCMC sampling, including well-defined stopping criteria, easier access to independent samples, more sophisticated treatment of multimodal solutions, and simultaneous estimation of the Bayesian evidence. Additionally, dynamic nested sampling can be performed such that samples are targeted adaptively during the fit to better sample specific areas of the posterior. Finally, internal testing with `Prospector` shows that `dynesty` requires ~2× fewer model calls to produce similar posteriors to MCMC methods, which translates to a ~50% decrease in run time. Each galaxy takes an average of ~25 CPU-hours to converge for our 14-parameter model, resulting in ~1.5 million CPU-hours[10] to analyze the whole sample.

Unless indicated otherwise, all reported parameters are the median of the marginalized posterior probability function, with 1σ error bars reported as half of the 84th–16th interquartile range. The `Prospector` parameter file for this version of the `Prospector-`α model is available online.[11]

3.3. Benchmark Models for SFR and Stellar Mass

The next section compares the stellar masses and SFRs derived from the `Prospector-`α fits to the fiducial inferences from the 3D-*HST* catalogs. The key physical assumptions made in the 3D-*HST* derivations are repeated here for completeness.

Stellar masses in the 3D-*HST* catalogs are calculated with FAST (Kriek et al. 2009), a grid-based χ^2 minimization code. Bruzual & Charlot (2003; BC03) SPS models are used with a Chabrier (2003) IMF, fixed solar metallicity, exponentially declining SFHs, and a single dust screen with a Calzetti et al. (2000) attenuation law. There is no nebular or dust emission; accordingly, regions of the SED with significant dust emission ($\lambda_{\rm rest} \gtrsim 3\,\mu$m) are heavily downweighted, and *Spitzer*/MIPS 24 μm photometry is not included in the fit. Only the best-fit parameters are reported. These are interchangeably called the 3D-*HST* catalog masses or the FAST masses in the text.

SFRs are calculated with the following relationship from Bell et al. (2005):

$$\mathrm{SFR}\ [M_{\odot}\ \mathrm{yr}^{-1}] = 1.09 \times 10^{-10}(L_{\rm IR} + 2.2L_{\rm UV})\ [L_{\odot}], \quad (1)$$

with $L_{\rm IR}$(8–1000 μm) estimated directly from the *Spitzer*/MIPS 24 μm flux and $L_{\rm UV}$(1216–3000 Å) determined from the best-fit EAZY template (Whitaker et al. 2014). This conversion does not include any additional information about the composition of the underlying stellar populations. These are interchangeably called the 3D-*HST* catalog SFRs or $\mathrm{SFR}_{\rm UV+IR}$ in the text.

4. Results

Stellar masses and SFRs are among the most basic and important outputs of galaxy SED-fitting codes and are therefore critical benchmarks for cross-code comparison. Here we compare the stellar masses and SFRs inferred from `Prospector-`α to the fiducial masses and SFRs in the 3D-*HST* catalogs. There are systematic offsets in this comparison such that `Prospector-`α masses are higher and the SFRs are lower. We demonstrate that the most significant causes of these offsets are older stellar populations and dust heating from old stars, respectively.

4.1. Revised Stellar Masses

Stellar mass is generally considered to be one of the most robust outputs of SED fitting, with typical systematic variations of ~0.2 dex between codes (e.g., Mobasher et al. 2015). Though robust when compared to other outputs, systematic uncertainties of 0.2 dex in stellar masses result in critical uncertainties when interpreting dynamical masses, measuring galaxy mass assembly rates, and calibrating simulations of galaxy formation.

Figure 3 shows the difference between the 3D-*HST* catalog masses and `Prospector` masses as a function of redshift. Specifically, the probability function for $\log(M_{\rm Prospector}/M_{\rm FAST})$ as a function of $\log(M_{\rm FAST})$ is created by summing the individual pdf's for all galaxies. The individual pdf's are calculated with the best-fit 3D-*HST* stellar masses and the full posterior distribution for the `Prospector-`α stellar masses. As the 3D-*HST* stellar masses do not include error estimates, they are assigned a Gaussian pdf with a uniform width of 0.1 dex. The stacked pdf's thus include both galaxy-to-galaxy scatter and measurement uncertainty.

The correlation of the offset with mass and redshift gives important clues as to the cause of the offsets. The median stellar mass difference is ~0.1–0.2 dex (~25%–60%), with a 68th percentile range between 0.2 and 0.4 dex. As stellar mass increases, the offset becomes smaller and the distribution becomes tighter. The offset also increases with decreasing redshift, with a larger increase at lower masses.

One potential cause of the mass offset is that FAST and `Prospector` use different SPS codes (BC03 versus FSPS, respectively). The modularity of `Prospector` makes it possible to build a physical model in the `Prospector` framework that mimics the FAST physical model, thereby isolating the effect of different SPS codes. This FAST-like model is fit to a fraction of the 3D-*HST* catalog (~2700

---

[9] https://github.com/joshspeagle/dynesty

[10] As a useful point of comparison, at the time of this writing 1.5 million CPU-hours costs approximately $20,000 on Amazon Web Services. This is ~40% of the cost of one observing night on the Keck telescopes.

[11] https://github.com/jrleja/prospector_alpha/blob/master/parameter_files/td_delta_params.py

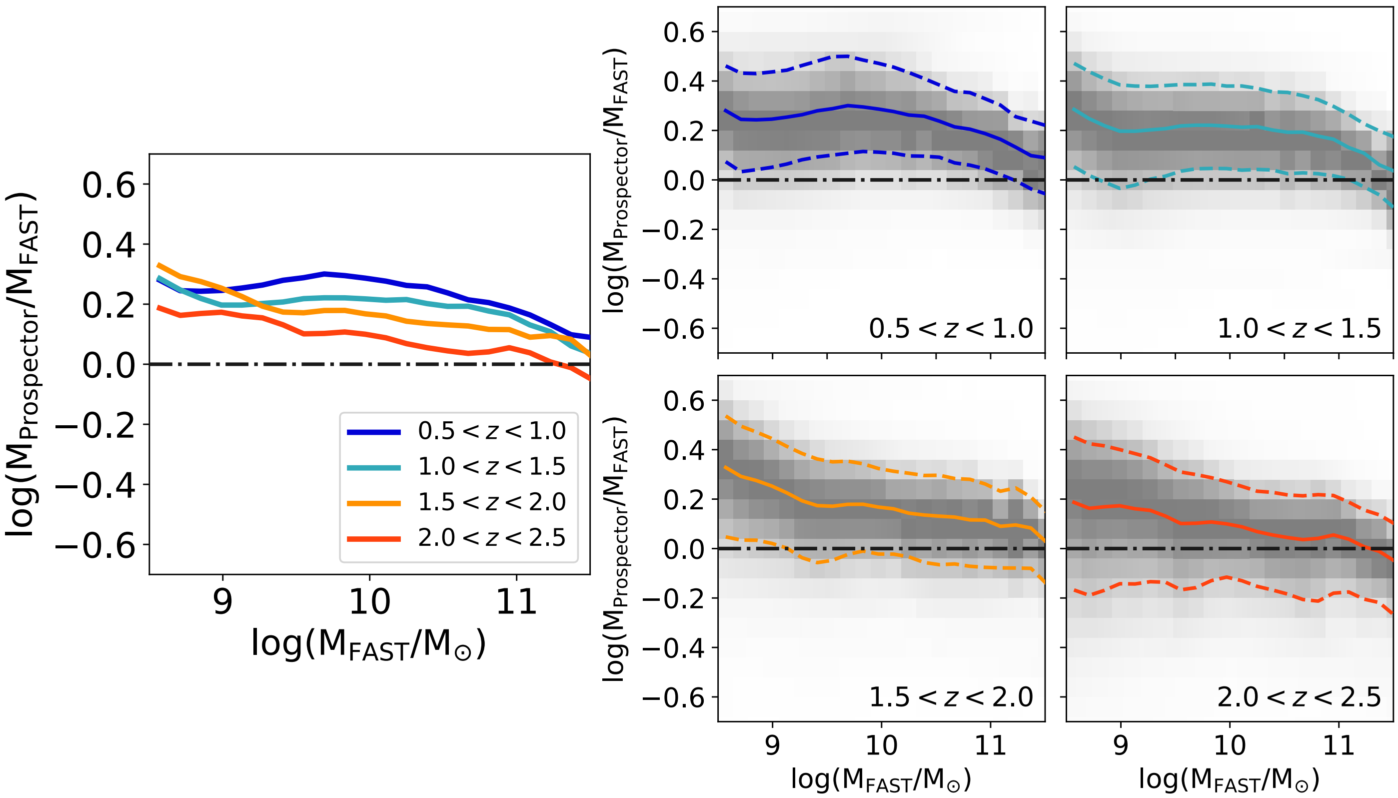

Figure 3. `Prospector-`α infers larger stellar masses than FAST. The right panels show the ratio of stellar masses in four discrete redshift windows, while the left panel shows the median from each redshift window. The offset increases with decreasing redshift and increases with decreasing stellar mass. The gray shading is proportional to the stacked pdf. The median is indicated by a colored solid line, and the 16th and 84th percentiles are indicated by colored dashed lines.

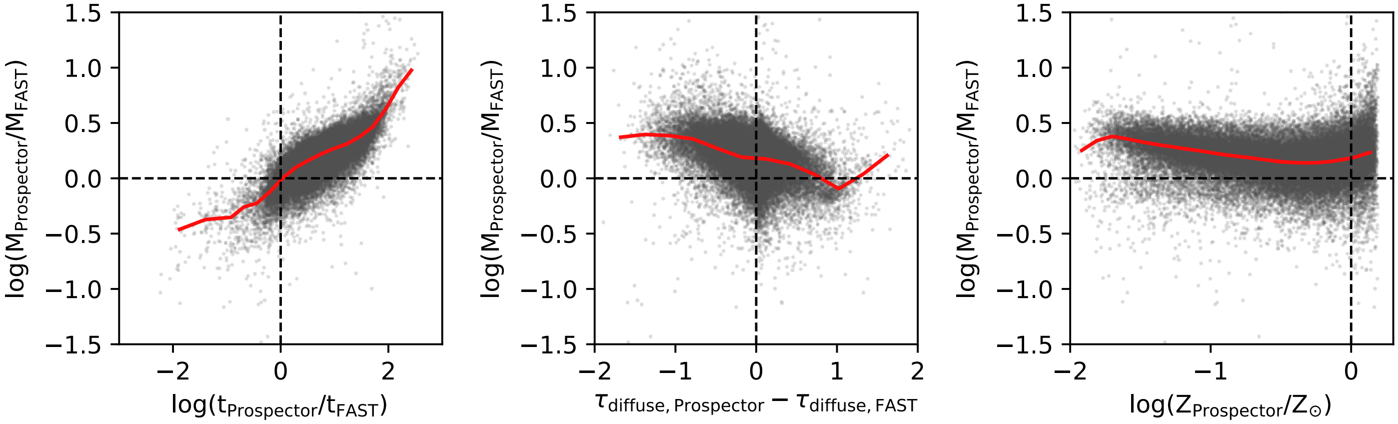

Figure 4. Correlations between the difference in stellar mass and properties derived from SED fitting. From left to right, the x-axis shows the difference in mass-weighted age, the optical depth from diffuse dust attenuation, and stellar metallicity. The running median is highlighted in red. Stellar age appears to be most closely associated with the stellar mass difference between the 3D-*HST* catalog values and `Prospector-`α and thus the likeliest cause of the offset.

galaxies), and the resulting mass offset is $\log(M_{\rm FSPS}/M_{\rm BC03}) \approx 0.05$ dex. This implies that different SPS codes contribute to, but do not dominate, the observed mass offset.

The bulk of the difference must then come from other differences in the SED models. Figure 4 explores three primary candidates: the mass-weighted stellar age, the stellar metallicity, and the dust optical depth. The FAST mass-weighted stellar ages are calculated from the best-fit FAST SFH, while the `Prospector` mass-weighted ages are calculated from samples of the SFH posterior. As the stellar metallicity is fixed to solar in the 3D-*HST* catalog fits, the variable `Prospector-`α metallicity is shown alone. The dust attenuation models have substantial differences: here we compare the V-band dust optical depth from the 3D-*HST* catalogs (computed with a fixed Calzetti et al. 2000 attenuation curve) to the `Prospector-`α V-band diffuse dust optical depth (computed with a flexible attenuation curve), which is only one component of the two-component Charlot & Fall (2000) dust model in `Prospector-`α. The relative difference in the V-band dust optical depths is a good proxy for the differential attenuation between each model.

Figure 4 makes it clear that, of the model differences considered, the age differences are the primary driver of the systematic offset in stellar mass. Indeed, older stellar ages provide a clean explanation for the trend in median mass offset with redshift and stellar mass. The trend with redshift comes from the dependence of age on $t_{\mathrm{universe}}(z)$: as redshift decreases, the upper limit on stellar age increases. This results in larger relative age differences permitted between `Prospector-α` and the 3D-*HST* catalog inferences. The offset increases with decreasing stellar mass because low-mass galaxies are primarily blue and star-forming: these galaxies display the most sensitivity to the SFH parameterization and priors (Leja et al. 2019).

Notably, the systematic mass differences suggest that `Prospector-α` will modify or break the tight relationship between mass-to-light ratio (*M/L*) and optical color (Bell & de Jong 2001). As may be expected, `Prospector-α` finds an increased *M/L* ratio at fixed optical color. It also finds greatly increased scatter in this relationship. This can broadly be attributed to the fact that a more complex physical model allows a wider range of physical properties at fixed optical color. This scatter in *M/L* is associated with variations in stellar age, metallicity, and the shape of the dust attenuation curve and will be explored further in future work.

4.2. Contrasting Pictures of Galaxy SFHs

The previous section demonstrated that differences in galaxy SFHs can cause systematics in inferred stellar masses. These differences in SFR(t) can be substantial: the mass-weighted ages inferred in the 3D-*HST* catalog and `Prospector-α` differ by factors of 3–5 for the majority of the galaxy population, despite being constrained by the exact same photometry. There are several reasons that SFHs are typically only weakly constrained by broadband photometry:

1. Younger stars ($t \lesssim 100$ Myr) dominate the observed SEDs of star-forming galaxies, greatly outshining older stars (Maraston et al. 2010).
2. Stellar isochrones evolve very little at late ages ($t \gtrsim 2$ Gyr), making it relatively difficult to distinguish between different age models for older galaxies (Conroy 2013).
3. Stellar age, stellar metallicity, and dust have similar effects on the UV–NIR SED that can result in significant parameter degeneracies (Bell & de Jong 2001).

When the data provide poor constraints, the prior for SFR(t) becomes very important in determining the output (Carnall et al. 2019; Leja et al. 2019). The prior on SFR(t) is determined both by the chosen SFH parameterization and by the priors on each parameter. Crucially, sensitivity to the prior is not specific to Bayesian analysis; classical methods implicitly set a uniform prior over the chosen SFH parameterization and range of the parameter grids. The continuity prior in `Prospector-α` is qualitatively very different from the exponentially declining SFH assumed in the 3D-*HST* analysis, so the difference in recovered SFHs is not surprising.

The SFHs inferred from these two analyses imply very different pictures of galaxy evolution. Figure 5 shows SFHs stacked across the star-forming sequence from both the 3D-*HST* analysis and the `Prospector-α` fits. These stacks are composed of galaxies split into four categories: above, on, and below the star-forming sequence, and quiescent. For consistency, SFRs from `Prospector-α` are used to sort galaxies in both stacks (the FAST SFRs are unreliable, as they do not include IR constraints). The locus of the star-forming main sequence is taken from Whitaker et al. (2014) and corrected downward by 0.3 dex to account for the typical difference between $\mathrm{SFR}_{\mathrm{Prospector}}$ and $\mathrm{SFR}_{\mathrm{UV+IR}}$ (see Section 4.3). The vertical divisions are taken to be 0.6 dex wide, or roughly twice the logarithmic scatter in the main sequence (Speagle et al. 2014). The SFH stacks are created by summing the individual pdf's for $\mathrm{SFR}(t)/M_{\mathrm{formed}}$[12] such that each galaxy in the stack is weighted equally.

The most striking result in Figure 5 is the contrast in average galaxy age. For example, the FAST fits infer that at $0.5 < z < 1$, galaxies above the star-forming sequence are ~200–300 Myr old, while galaxies on the star-forming sequence are ~1 Gyr old. In contrast, the `Prospector-α` SFHs infer galaxy ages of order a few gigayears regardless of their position on the star-forming sequence. These SFHs imply very different galaxy mass assembly histories. We demonstrate via a continuity analysis (Section 5.1) that the assembly histories implied by the 3D-*HST* fits are far too rapid to be consistent with the observed evolution of the stellar mass function.

There are also strikingly different descriptions of a galaxy's lifetime on the star-forming sequence. The `Prospector-α` SFHs find that galaxy ages show little correlation with their position relative to the star-forming sequence. Indeed, the `Prospector-α` SFHs are consistent with a galaxy's position on the star-forming sequence being a temporary status, lasting of order ~100–500 Myr before converging on long-term SFHs with similar trajectories. On the other hand, the 3D-*HST* fits imply that a galaxy's position relative to the star-forming sequence is strongly correlated with its lifetime, with galaxies above the main sequence having appeared between 300 and 500 Myr in the past and galaxies on the star-forming sequence having lifetimes of ~1 Gyr. This is almost a necessary conclusion when fitting exponentially declining SFHs, as the only way to generate relatively high sSFRs in such a framework is to have very young ages.

A common rationale for using exponentially declining SFHs is that the inferred τ and age are meant to characterize the bulk of the most recent star formation rather than representing an actual SFH. However, given that the actual SFH implied by these models directly affects the mass estimate, it is more useful in this comparison to take the SFHs at face value.

Beyond the cross-comparison, the `Prospector-α` SFHs in Figure 5 provide an interesting overview of galaxy formation and evolution over the critical period of $0.5 < z < 2.5$. The `Prospector-α` stacks show that at higher redshifts typical galaxies on and above the star-forming sequence have rising SFHs, while those below the star-forming sequence have flatter SFHs. Galaxies above the star-forming sequence at $2 < z < 2.5$ were on this sequence ~100 Myr in the past, while galaxies below the star-forming sequence have been off this sequence for three times longer. Quenched galaxies have falling SFHs and get older with decreasing redshift, and they have also been quenched for longer at lower redshifts.

Given the dominant role of the prior in the recovery of SFHs (Carnall et al. 2018; Leja et al. 2018), it will be important to establish which of the `Prospector-α` results are driven by the data and which are driven by the prior. For example, it is unlikely that the data are constraining the characteristic timescales on which SFRs change, as a broad range of

---

[12] Note that sSFR is calculated using *stellar* mass but SFR(t) is normalized by *total mass formed*. This causes some overlap in the youngest SFH bins, which would be strictly forbidden if the definitions of mass were the same.

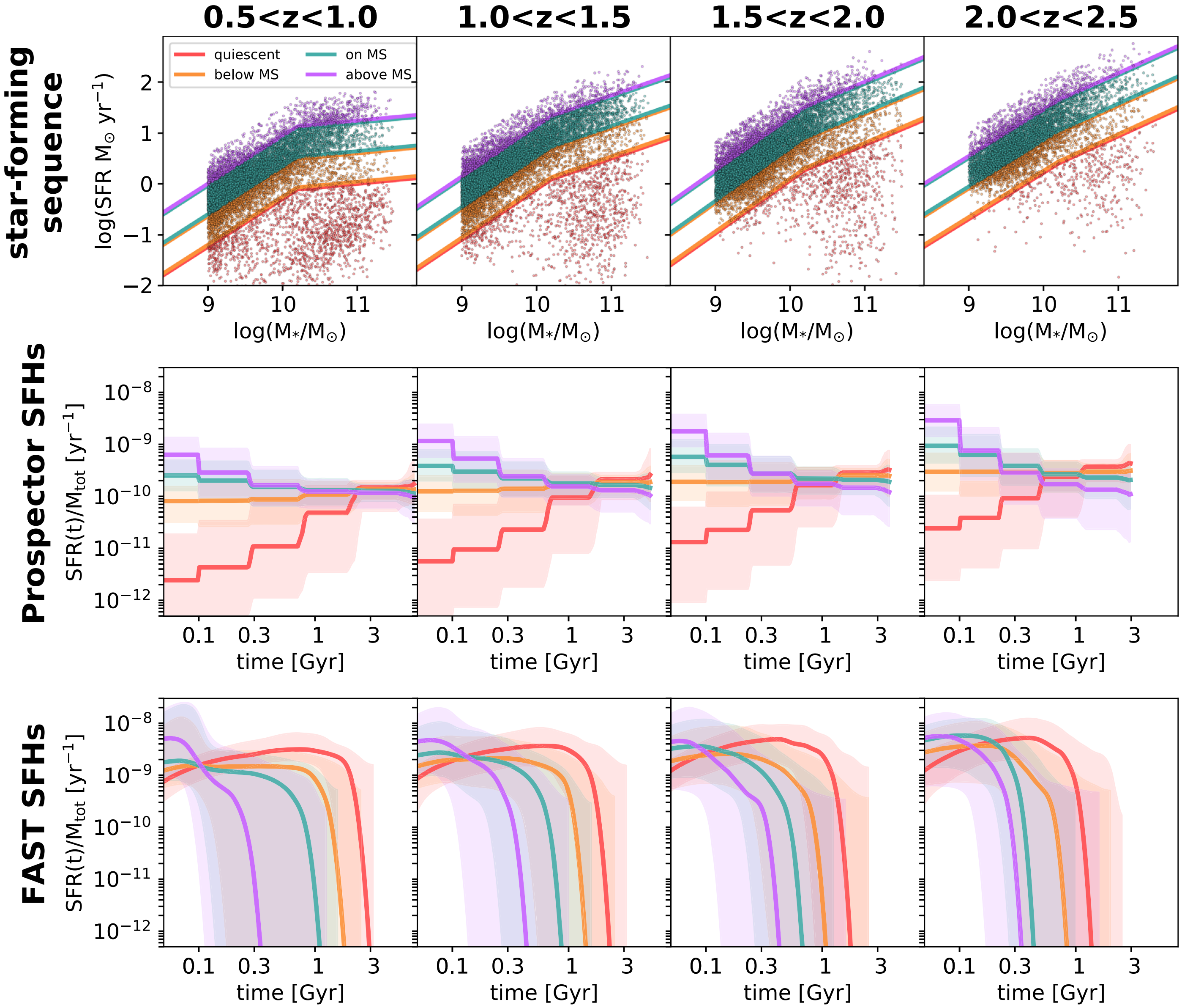

Figure 5. Stacked SFHs from `Prospector-`α and FAST as a function of SFR, stellar mass, and redshift. The top panels show the distribution of galaxies in the star-forming sequence. Galaxies are divided into above, on, and below the star-forming sequence and quiescent, and their SFHs are stacked separately. The middle and bottom panels show the median of the SFH stacks, and the shaded regions cover the 16th and 84th percentiles from both `Prospector-`α and the 3D-*HST* catalogs. The 3D-*HST* catalog SFHs produce stellar populations that are far younger (factors of 3–5 and more) than the `Prospector-`α SFHs.

characteristic timescales are often equally consistent with constraints from broadband photometry (Leja et al. 2018). A more detailed analysis of these trends is deferred to future work.

4.3. Revised SFRs

UV+IR SFRs are considered more reliable than those from SED-fitting codes such as FAST because they also include contributions from dusty star formation via the observed IR luminosities. However, these values do not include galaxy-to-galaxy variation in the underlying stellar populations properties, which is measured directly in SED fitting. Here we show that SFRs from panchromatic SED fitting are systematically lower than $\mathrm{SFR_{UV+IR}}$ and that this offset is largely due to energy emitted from older stellar populations. This includes energy observed directly in the UV and energy attenuated and reemitted by dust.

The 3D-*HST* catalogs provide $\mathrm{SFR_{UV+IR}}$ from Equation (1) following the methodology of Whitaker et al. (2014). L_{IR} is obtained in the 3D-*HST* analysis by converting the observed *Spitzer*/MIPS 24 μm flux directly into L_{IR} using a fixed template. However, the observed IR fluxes are not reliable for low-mass galaxies owing to confusion limits. To extend this comparison to low-mass galaxies, we instead calculate the *Spitzer*/MIPS 24 μm flux from model spectra drawn from the `Prospector-`α posteriors. These are combined with the log-average of the Dale & Helou (2002) templates to calculate L_{IR}. L_{UV} is measured directly from the `Prospector-`α model spectra.

To ensure that the resulting $\mathrm{SFR_{UV+IR}}$ values are not systematically biased by this approximation, we compare UV+IR SFRs calculated from the posteriors of the

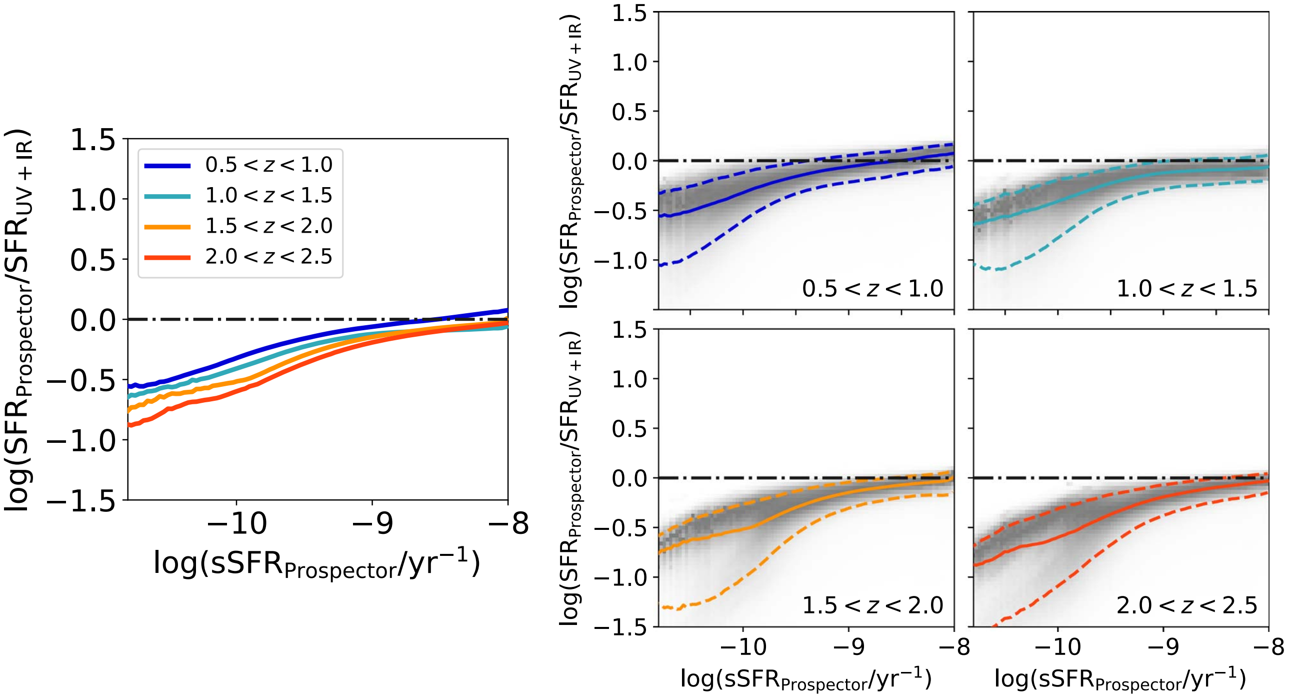

Figure 6. Offset between $\mathrm{SFR}_{\mathrm{UV+IR}}$ and $\mathrm{SFR}_{\mathrm{Prospector}}$ as a function of $\mathrm{sSFR}_{\mathrm{Prospector}}$. The right panels show four different redshift windows, with gray shading representing the stacked pdf. The median is a colored solid line, and the 16th and 84th percentiles are colored dashed lines. The left panel highlights the redshift evolution of the median. There is good agreement at high sSFR, but at lower sSFRs the `Prospector-`α SFRs are increasingly lower than $\mathrm{SFR}_{\mathrm{UV+IR}}$.

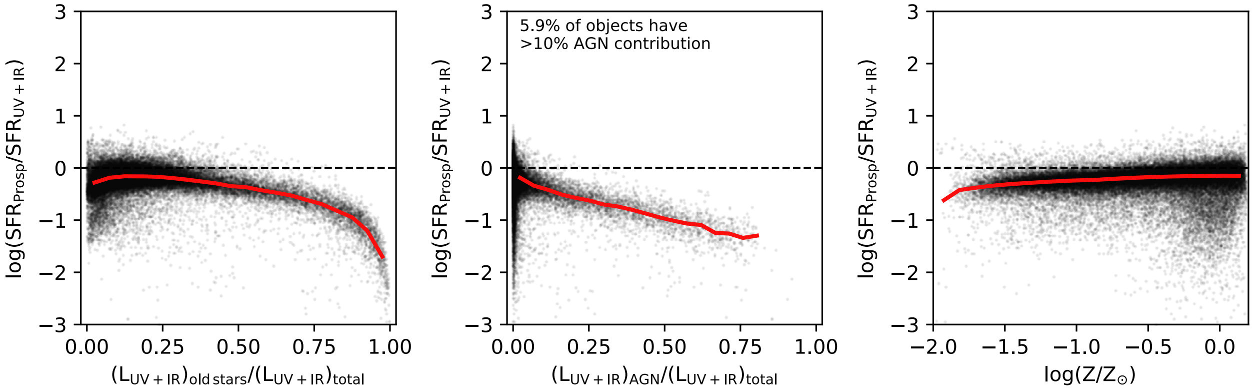

Figure 7. Correlations between $\mathrm{SFR}_{\mathrm{Prospector}}/\mathrm{SFR}_{\mathrm{UV+IR}}$ and derived galaxy properties. From left to right, the x-axis values are the fraction of $L_{\mathrm{IR}}+L_{\mathrm{UV}}$ emitted by old stars, the fraction of $L_{\mathrm{IR}}+L_{\mathrm{UV}}$ emitted by AGNs, and the stellar metallicity. While all three components are correlated with the offset, the offset correlates most clearly with heating by old stellar populations ("old" defined here as $t > 100$ Myr). We note that while AGNs are strongly associated with lower SFRs, they affect a relatively small proportion of the population (only 5.9% of galaxies have an AGN contribution of >10%).

`Prospector-`α model fits to the UV+IR SFRs from Whitaker et al. (2014). There is no measurable offset as a function of SFR, and there is a relatively low scatter of 0.24 dex, suggesting that the model $\mathrm{SFR}_{\mathrm{UV+IR}}$ are an acceptable approximation for the values in the 3D-*HST* catalog.

Figure 6 shows the stacked distribution of $\mathrm{SFR}_{\mathrm{UV+IR}}/\mathrm{SFR}_{\mathrm{Prospector}}$ as a function of $\mathrm{sSFR}_{\mathrm{Prospector}}$. This is created by summing the individual pdf's for all galaxies. The median offset ranges between 0 and 1 dex and is largest at low sSFRs. The central 68th percentile ranges from 0.2 to 0.8 dex and is also largest at low sSFRs.

Figure 7 explores potential physical causes of this offset: additional flux from "old" ($t > 100$ Myr) stellar populations, hot dust emission from AGN activity, and a nonsolar stellar metallicity. The x-axis of the left two panels shows the fractional change in $(L_{\mathrm{UV}} + L_{\mathrm{IR}})$ when old stars and AGNs are removed from the `Prospector-`α model, while the third panel simply shows $\log(Z/Z_{\odot})$.

The offsets show some correlation with all three parameters, suggesting that the overall change in inferred SFR cannot be simply associated with a single cause. However, the clearest correlation is with energy from old stars. This effect naturally explains the trend of increasing offset with decreasing sSFR: at lower sSFRs, a higher fractional contribution of total flux is emitted by old stars. This energy from old stars includes both energy emitted directly in the UV and energy that is attenuated from the UV, optical, and near-infrared and reemitted in the IR. Emission from buried AGNs also strongly affects the SFR of a small fraction of galaxies, while stellar metallicity has a more subtle effect for many galaxies below $Z = Z_{\odot}$.

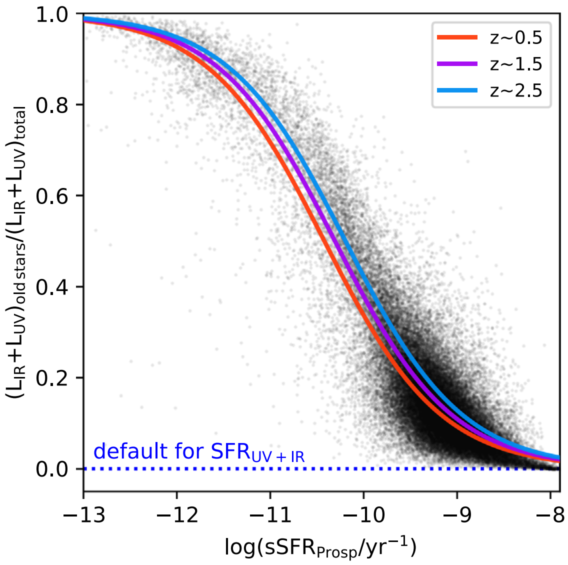

Figure 8. Relationship between the fraction of L_{UV}+L_{IR} emitted by old stars (t > 100 Myr) and the sSFR inferred from panchromatic SED modeling. The fit to this relationship from Equation (2) is shown in red, while the 16th–84th percentile range is shaded in red. As the sSFR decreases, more and more of the luminosity is emitted by old stars. A linear transformation between UV+IR luminosity and SFR can thus overestimate the SFR for galaxies with low sSFR.

4.4. Effect of Old Stellar Heating on SFR Estimates

Flux from old stars can have a strong effect on SFRs inferred only from L_{UV}+L_{IR}. It is therefore important to clarify both how the strength of this effect varies across the galaxy population and how robustly this effect can be modeled within `Prospector`.

Equation (1) for $\mathrm{SFR_{UV+IR}}$ was derived by creating a stellar population with a constant SFR over 100 Myr. The underlying principle is energy balance: if all the observed luminosity comes from young stars, inverting this will return the number of young stars (i.e., the SFR). This is a good assumption when young stars dominate the stellar energy budget. However, old stars (t > 100 Myr) also contribute to the observed UV emission and indirectly to the observed IR emission via dust attenuation. This heating is undoubtedly occurring at some level: the salient question is to what extent it is important in affecting the simple $\mathrm{SFR_{UV+IR}}$ estimates.

Figure 8 shows the fraction of L_{UV}+L_{IR} emission originating from stars older than 100 Myr in the `Prospector`-α model as a function of sSFR. The effect of old stellar heating on SFR estimates has been demonstrated at both low and high redshift for small samples (Cortese et al. 2008; De Looze et al. 2014; Utomo et al. 2014), but the measurement presented here is the first for a statistically significant sample of galaxies. The relationship in Figure 8 is fit with the equation

$$y = 0.5\tanh(a\log[\mathrm{sSFR/yr^{-1}}] + bz + c) + 1, \qquad (2)$$

where $y = (L_{\mathrm{UV+IR}})_{\mathrm{old\ stars}}/(L_{\mathrm{UV+IR}})_{\mathrm{total}}$, $a = -0.8$, $b = 0.09$, and $c = -8.4$.

As might be expected, galaxies with high sSFRs ($\gtrsim 10^{-9}\ \mathrm{yr^{-1}}$) experience negligible contribution from old stars, while galaxies with low sSFRs $\lesssim 10^{10.5}\ \mathrm{yr^{-1}}$ are dominated by emission from old stars. The point of equal contribution is at sSFR $\approx 10^{-10.3}\ \mathrm{yr^{-1}}$. For reference, a $10^{10.5}\ M_{\odot}$ galaxy on the star-forming sequence at $z = 0.75$ has an sSFR of $\sim 10^{-9.4}\ \mathrm{yr^{-1}}$ (Whitaker et al. 2014), and approximately 20% of the observed IR and UV luminosity in such a galaxy is expected to come from old stars. This effect decreases to <10% at $z = 2.25$.

Somewhat counterintuitively, the offset between $\mathrm{SFR_{Prospector}}$ and $\mathrm{SFR_{UV+IR}}$ increases with increasing redshift (Figure 6), implying that old stars make up a larger fraction of the observed $L_{\mathrm{UV+IR}}$ in higher-redshift galaxies. Figure 9 confirms this, showing the fractional contribution to the total $L_{\mathrm{UV+IR}}$ from stars as a function of age and galaxy redshift. The bulk of "old" stellar heating is performed by stars aged 0.1–1 Gyr at $z = 2.25$ and by stars aged 2–6 Gyr at $z = 0.75$. The old stellar populations in high-redshift galaxies are comparatively younger and brighter, contributing more to $L_{\mathrm{UV+IR}}$ at a fixed value of sSFR. The strength of old stellar heating thus increases with redshift because the old stellar populations at $z = 2$ are on average more luminous.

There is a good reason that this effect is not typically included in SFR estimates: it is technically challenging to include the effect of dust heating from old stars, as it requires that SFR, SFH, and dust attenuation be estimated from a single self-consistent model. In theory, it is possible to modify the assumed SFH assumed in calculating $\mathrm{SFR_{UV+IR}}$ to include more emission from old stars and reduce this bias (Kennicutt & Evans 2012). This is not a universal solution, though, as revising the recipe for $\mathrm{SFR_{UV+IR}}$ in this fashion will then necessarily underestimate SFRs in high-sSFR galaxies.

Using a sophisticated model such as `Prospector` to estimate SFRs is not necessarily a panacea either. The fractional amount of energy generated by old stars depends not only on accurate estimates of the long-term SFH but also on the spatial distribution of old and young stars relative to the dust. Thus, the size of the effect in Figure 8 is dependent on the adopted dust model. `Prospector`-α uses a two-component Charlot & Fall (2000) model wherein all stars are attenuated equally by a diffuse screen of dust, while younger stars experience extra attenuation. The variable shape of the dust attenuation curve adds more variance to the age-dependent attenuation, as wavelength-dependent attenuation translates into age-dependent attenuation owing to the different emission profiles of young and old stars.

Assumptions about the star–dust geometry can be explored by observing systems where the contribution of old stars to the integrated UV and IR emission of galaxies can be separated. For example, the bulge of Andromeda is composed almost entirely of old stars and constitutes 30% of the total stellar mass yet only contributes 5% of the IR luminosity. This may not be surprising, given that the bulge also contains only 0.5% of the total dust mass (Groves et al. 2012). The majority of the dust lives in star-forming regions in the spiral arms. The key question, then, is to what extent the IR emission from the dusty spiral arms is caused by old stars, both nearby and from roughly kiloparsec distances. This can be answered by spatially resolved modeling of mixed systems of old and young stars with a careful accounting of energy transfer between adjacent pixels. Studies that employ this approach find that a large fraction of the energy absorbed by dust in nearby spiral galaxies originates from the old stellar populations (e.g., 37% for M51, 91% for M31; De Looze et al. 2014; Viaene et al. 2017).

Spatially resolved modeling may also have the potential to yield relationships that can better calibrate the energy contribution of old stars in unresolved SED modeling. For example, it is well established that the dust temperature is

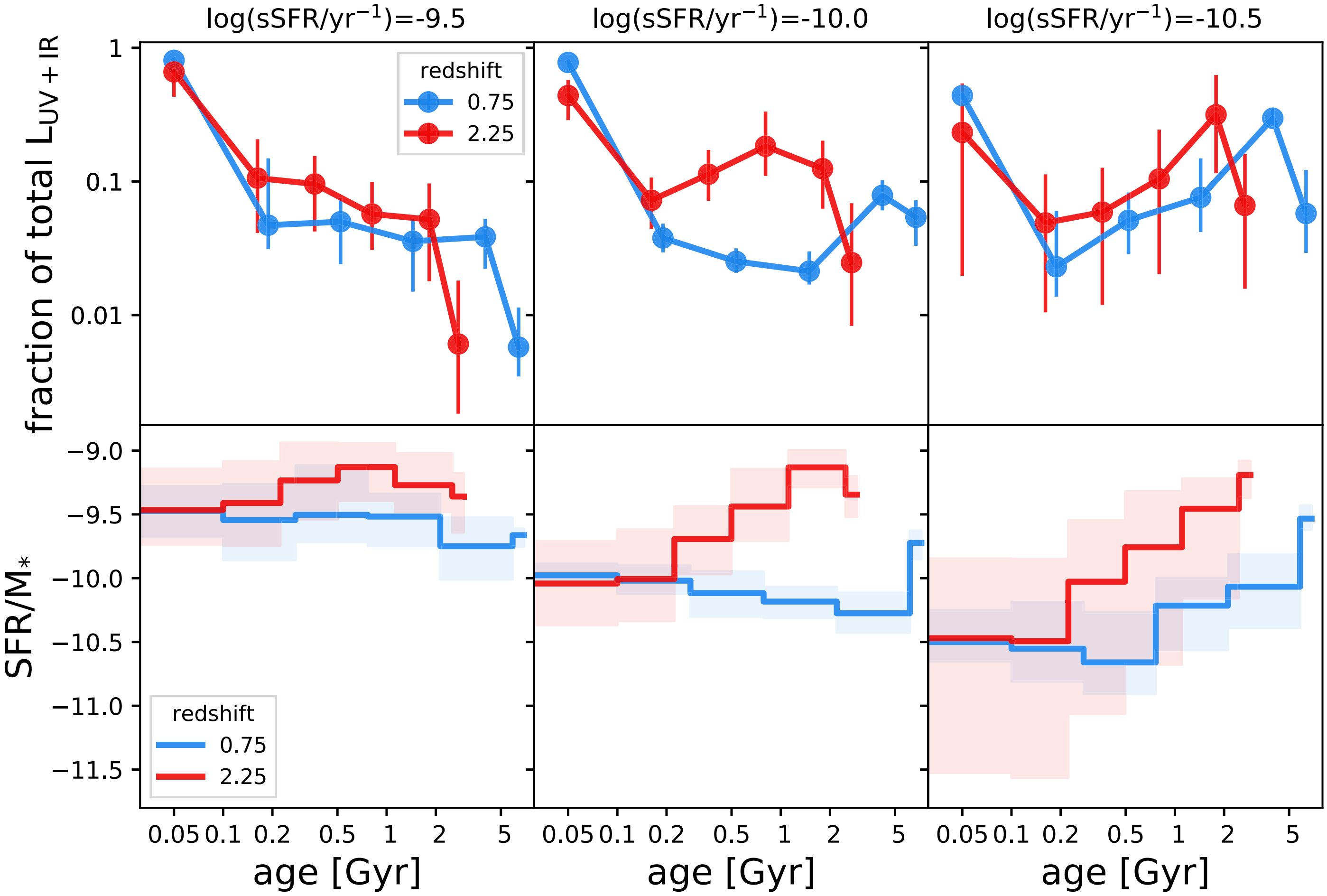

Figure 9. The bulk of "old" stellar heating is typically performed by stars aged 0.1–1 Gyr at $z = 2.25$ and by stars aged 2–6 Gyr at $z = 0.75$. The top panels show the relative contribution to $L_{\rm UV}+L_{\rm IR}$ from stellar populations of different ages, with the youngest bin encompassing 0–100 Myr. Each panel compares randomly selected galaxies from different redshifts at a fixed log(sSFR). In each case, the relative contribution of young (<100 Myr) stars decreases with increasing redshift. The bottom panels show the corresponding SFHs, highlighting the fact that the "old" stars in high-redshift galaxies are comparatively much younger and more luminous. This means that the strength of old stellar heating *increases* with redshift. Note that the consistent "dip" in the top panels for the contribution is caused by the SFH bin spacing. The oldest time bin is approximately four times smaller than the previous bin, meaning that there is consistently less mass in the oldest time bin.

closely related to the stellar density (e.g., Chanial et al. 2007; Rujopakarn et al. 2011), a relationship driven by the underlying relationship between dust temperature and the intensity of incident radiation. This means that systems that have different spatial distributions of young and old stars will show a wavelength-dependent infrared contribution from old stars. For example, direct *Herschel* observations of Andromeda show that optical light from old bulge stars heats dust to higher temperatures than star-forming regions do (Groves et al. 2012). Panchromatic radiative transfer models of Andromeda corroborate this picture, suggesting that dust heated only by old stars would peak at 150 μm, whereas younger stellar populations would cause it to peak around 200–250 μm (Viaene et al. 2017). The findings in Andromeda are generalizable to the nearby galaxy population: the KINGFISH survey (Skibba et al. 2011) finds that, at fixed sSFR (i.e., a fixed ratio of young to old stars), early-type galaxies have hotter dust temperatures on average than late-type galaxies. This means that old stars are heating the dust to higher temperatures because radiation density is extremely high in dense stellar regions such as bulges where old stars happen to live.

However, this relationship between dust temperature and stellar age can work in either direction depending on the relationship between stellar morphology and stellar age. Thus, galaxies that do not have a classic bulge-and-disk stellar morphology will likely behave differently: for example, Chanial et al. (2007) show that the overall dust temperature in galaxies is more closely related to the density of young stars. This relationship between young stars and hot dust emission is likely to become more dominant in shaping the overall SED of the galaxy at high redshift as stellar populations become younger (e.g., Imara et al. 2018). The relationship among stellar morphology, stellar age, redshift, and dust temperature is rich and complex and deserves much deeper exploration in future work.

5. Global Implications and Model Cross-validation

The `Prospector-`α model finds that, on average, galaxies in the distant universe are both more massive and more quiescent than suggested in previous studies. These effects are due to `Prospector-`α inferring older ages and including the effect of old stellar heating, respectively. In this section we examine the implications and the self-consistency of these results by cross-comparing different inferences of global quantities, including the evolution of the stellar mass function and the cosmic SFR density. We also indirectly test the

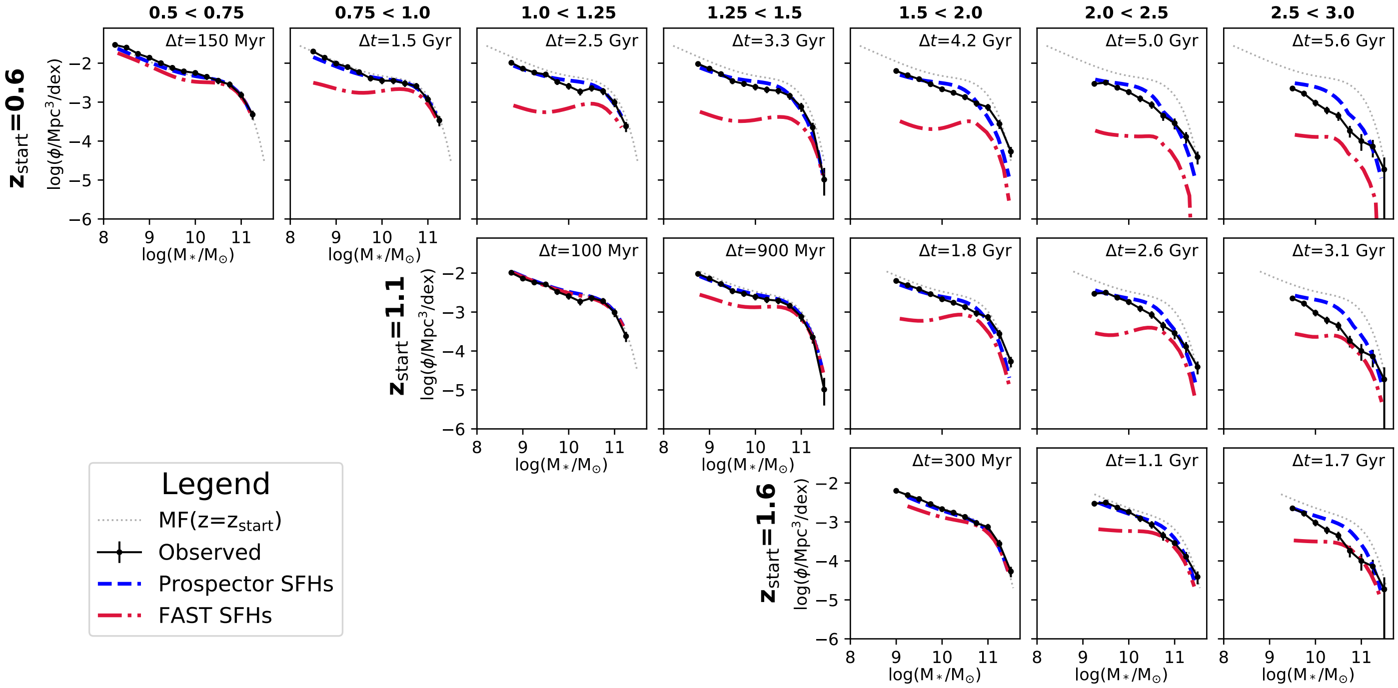

Figure 10. Evolving the observed mass functions backward in time according to the SFHs both from the FAST fits in the 3D-*HST* catalogs and from `Prospector-`α. Observed mass functions are taken from Tomczak et al. (2014) at three different redshifts and evolved to $z = 3$. Overall, `Prospector-`α provides much more realistic SFHs than FAST. The 3D-*HST* SFHs for low-mass galaxies are far too young at all redshifts, while `Prospector` SFHs are slightly too old.

accuracy of `Prospector-`α masses by comparing stellar and dynamical masses.

5.1. The Consistency between SFHs and the Growth of the Stellar Mass Function

SED fitting simultaneously infers both the current stellar mass $M_*(t = 0)$ and the past SFH, $dM/dt(t)$. In principle, this means that the galaxy stellar mass function $\phi(M, z)$ needs only to be observed at $z = 0$; the redshift dependence of this function can then be predicted by evolving each galaxy backward in time according to $dM/dt(t)$ while also accounting for the effect of galaxy mergers. In practice, the current stellar mass is a much more robust quantity than the SFH, and so the mass function is better constrained by measuring the current stellar mass for galaxy populations across a range of redshifts. This “redundant” measurement creates an opportunity to test the self-consistency of SED-fitting models. The inferred SFHs can be used to evolve the observed stellar mass function at a lower redshift $z_{\rm start}$ to some higher redshift $z_{\rm obs}$ and then compared with the observed stellar mass function at that redshift.

Here we perform this consistency check for the SFHs from `Prospector-`α and from the FAST fits in the 3D-*HST* catalogs. To do this properly, it would be necessary to couple a `Prospector`-derived stellar mass function with a full `Prospector` analysis of SFHs and SFRs derived from different data sets at redshifts. As this full analysis has not yet been performed, here we instead recast the `Prospector` growth rates in terms of FAST mass functions measured in previous studies. This exercise will give a sense of what the full analysis might reveal, and indeed it looks promising.

We take the observed mass functions from Tomczak et al. (2014), specifically adopting the smooth parameterizations of this mass function as a function of redshift from Leja et al. (2015) to ensure a monotonic evolution with redshift. The SED fitting in Tomczak et al. (2014) is performed using FAST, the same code used to generate the SED-fitting outputs in the 3D-*HST* catalog, which ensures that there is minimal systematic offset between the mass function masses and the 3D-*HST* catalog masses. Accordingly, for consistency, the `Prospector` growth rate function is also cast in terms of the 3D-*HST* catalog mass.

For three initial redshifts $z = (0.6,\ 1.1,\ 1.6)$, we select galaxies in a narrow range $\delta z = 0.1$ and transform their SFHs into the distribution of fractional change in total mass formed $\Delta M_{\rm formed}/M_{\rm formed}$, hereafter called the growth kernel $f_M(z, M_*)$. For the `Prospector` results the kernel is built by summing the full pdf’s; the 3D-*HST* results lack error estimates, so the kernel is composed of the distribution of best-fit SFHs. The growth kernel $f_M(z, M_*)$ is then smoothed in the mass direction, equivalent to assuming a smooth growth rate as a function of mass. Finally, the mass function at a higher redshift $z_{\rm obs}$ is predicted by convolving the mass function observed at $z_{\rm start}$ by the growth kernel $f_M(z_{\rm start}, M_*)$.

We additionally include a simple model for the effect of galaxy–galaxy mergers on the stellar mass function from Leja et al. (2015). In brief, this model includes effects from both the rate at which galaxies merge with more massive galaxies than themselves (i.e., the “destruction” rate) and the rate at which galaxies gain stellar mass from mergers (the “growth” rate) as a function of both stellar mass and redshift from the Guo et al. (2013) semi-analytical model of galaxy formation. For this work we first increase the number density according to the destruction rate integrated between $z_{\rm start}$ and $z_{\rm obs}$ and then remove mass from galaxies according to the growth rate as a function of mass.

The results of this exercise are shown in Figure 10. For all combinations of $z_{\rm start}$ and $z_{\rm obs}$, the FAST SFHs greatly underpredict the number density of low-mass galaxies

($\lesssim 10^{10}\,M_\odot$). This suggests that the exponentially declining SFHs assumed in FAST greatly underestimate the ages of low-mass galaxies, in agreement with the findings of Wuyts et al. (2011a), who use a similar methodology. Meanwhile, the predictions from the `Prospector-`α SFHs are in much better agreement with the observations, though there are hints that there is more rapid evolution at higher redshifts ($z > 2.5$) than predicted from the `Prospector-`α SFHs.

The story is more complex at the higher masses. The 3D-*HST* SFHs underpredict the ages of massive galaxies at lower redshifts ($z \sim 0.6$) but give much more accurate ages at $z \sim 1.1$ and 1.6. The `Prospector-`α SFHs accurately predict the evolution of very massive galaxies ($M_* > 10^{11}\,M_\odot$) but somewhat overpredict the ages of galaxies around the knee of the mass function ($10^{10} < M_*/M_\odot < 10^{11}$).

In summary, the `Prospector-`α SFHs present a remarkable improvement over the FAST SFHs, but there remain specific mass and redshift regimes that can be improved. The continuity prior appears to be a reasonable prior for some (even most) combinations of redshift and mass but perhaps can be improved on for galaxies around the knee of the mass function. A hierarchical Bayesian model would be a logical next step to craft an SFH prior that is simultaneously consistent with the observed SEDs and with observations such as the evolution of the stellar mass function with time. Thorough comparisons of SFHs derived from integrated light to those derived from resolved (e.g., Johnson et al. 2013) and semi-resolved (Cook et al. 2019) stellar populations in local galaxies are also a promising way forward.

5.2. A New Consistency between Independent Inferences of the Cosmic SFR Density

The cosmic SFR density is the rate of new stars produced per unit volume and unit time. In principle, this quantity can be inferred with SED modeling in two ways: (1) by summing the instantaneous SFR for all galaxies in a fixed volume, or (2) by measuring the change in total stellar mass in the galaxy population as a function of time. Previous work has demonstrated that these two methods are inconsistent with one another at roughly the 0.3 dex level (Madau & Dickinson 2014; Leja et al. 2015; Tomczak et al. 2016). While this offset is improved from the 0.6 dex discrepancy measured just a decade ago (Wilkins et al. 2008), it remains a serious concern, as it implies systematic, across-the-board errors in inferred stellar masses and SFRs at the factor-of-two level. Here we show that the new masses and SFRs estimated with `Prospector-`α resolve this tension.

We estimate $\rho_{\rm SFR}(z)$ (i.e., the SFRD) by again using the phenomenological description of the Tomczak et al. (2014) mass functions from Leja et al. (2015) as an intermediate step. This mass function is multiplied by $\mathrm{SFR}_{\rm Prospector}(M_{\rm FAST})$ to produce the number density of galaxies as a function of SFR. The average value of $\mathrm{SFR}_{\rm Prospector}(M_{\rm FAST})$ is calculated by stacking individual galaxy posterior pdf's for this quantity. This produces the number density of galaxies as a function of SFR, which is then integrated numerically to produce the SFR density $\rho_{\rm SFR}(z)$. This calculation is performed in small δz steps between $0.5 < z < 2.5$. This procedure is repeated for $\mathrm{SFR}_{\rm UV+IR}$. The integration is performed at a fixed mass range of $9 < \log(M_{\rm FAST}/M_\odot) < 13$ for all redshifts.

To estimate $\dot{\rho}_{\rm mass}(z)$ (i.e., the SFRD from stellar mass growth), we take Equation (5) from Tomczak et al. (2014) describing the growth of stellar mass density from FAST:

$$\log(\rho_{\rm mass}) = a(1+z) + b, \tag{3}$$

with $\rho_{\rm mass}$ the total mass density in $M_\odot/\mathrm{Mpc}^3$, $a = -0.33$, and $b = 8.75$. The `Prospector-`α stellar mass density is calculated using a correction to this equation estimated from $M_{\rm Prospector}(M_{\rm FAST})$ and the Tomczak et al. (2014) stellar mass functions. The stellar mass density $\rho_{\rm mass}(z)$ is then converted into $\dot{\rho}_{\rm mass}(z)$ by numerically estimating $d\rho_{\rm mass}/dt$ between time steps and multiplying by $1 - R$, where R is the fraction of mass ejected from a stellar population during the course of passive stellar evolution. This mass loss is assumed to occur instantaneously. For a Chabrier (2003) IMF, $R = 0.36$ (Leja et al. 2015).

This exercise produces $\dot{\rho}_{\rm mass}$ and $\rho_{\rm SFR}$ at $0.5 < z < 2.5$ both from `Prospector-`α and from the combination of FAST stellar masses and $\mathrm{SFR}_{\rm UV+IR}$. In principle, $\dot{\rho}_{\rm mass}$ and $\rho_{\rm SFR}$ may disagree when using a fixed mass selection as done in this work because both mass growth and star formation occur in lower-mass galaxies.

To correct for this effect, we estimate the fraction of $\dot{\rho}_{\rm mass}$ and $\rho_{\rm SFR}$ occurring below this mass limit using the Universe Machine (Behroozi et al. 2019), a semi-empirical model that generates self-consistent estimates of the mass assembly history of the galaxy population. We caution that the estimated mass and SFR completeness estimated by comparing our sample selection criteria to the full 3D-*HST* catalog (Figure 1) are slightly larger (<10%) than those estimated from the Universe Machine. These completeness corrections have a limited effect on the comparison with previous SFR and mass density measurements, which is the key result of this paper, and so they are not explored further in this work. However, completeness corrections are essential to accurately measuring the total SFR density and will be derived self-consistently in future work.

The values of $\rho_{\rm SFR}/\dot{\rho}_{\rm mass}$ from these two procedures are shown in Figure 11. The combination of FAST *dM/dt* and $\mathrm{SFR}_{\rm UV+IR}$ recovers the inconsistency in SFRD inferences observed in previous work (e.g., Madau & Dickinson 2014; Leja et al. 2015; Tomczak et al. 2016): a ~0.3 dex gap between the observed SFRD and the SFRD implied by the mass function. Indeed, many galaxy formation models have long been in tension with the observed SFRs at $1 < z < 3$, roughly at the factor-of-two level (Bouché et al. 2010; Firmani et al. 2010; Davé et al. 2011; Lilly et al. 2013; Dekel & Burkert 2014; Genel et al. 2014; Mitchell et al. 2014). Given that models of galaxy formation often calibrate themselves to the evolution of the stellar mass function, this tension is not unexpected (Leja et al. 2015).

This tension disappears with the new stellar masses and SFRs from the `Prospector-`α model. Internally, the SFR density decreases by ~0.2 dex compared to $\mathrm{SFR}_{\rm UV+IR}$, while the observed growth of stellar mass increases by ~0.1 dex compared to FAST stellar masses. The new estimates are internally consistent to within $\lesssim$0.1 dex.

It is worth emphasizing that `Prospector` infers masses and SFRs using the same physical model. This is in contrast to the 3D-*HST* catalog masses and SFRs that are estimated from models with different and conflicting physical assumptions. It is better to use self-consistent estimates of mass and SFR when possible (e.g., Driver et al. 2018). Despite the internal

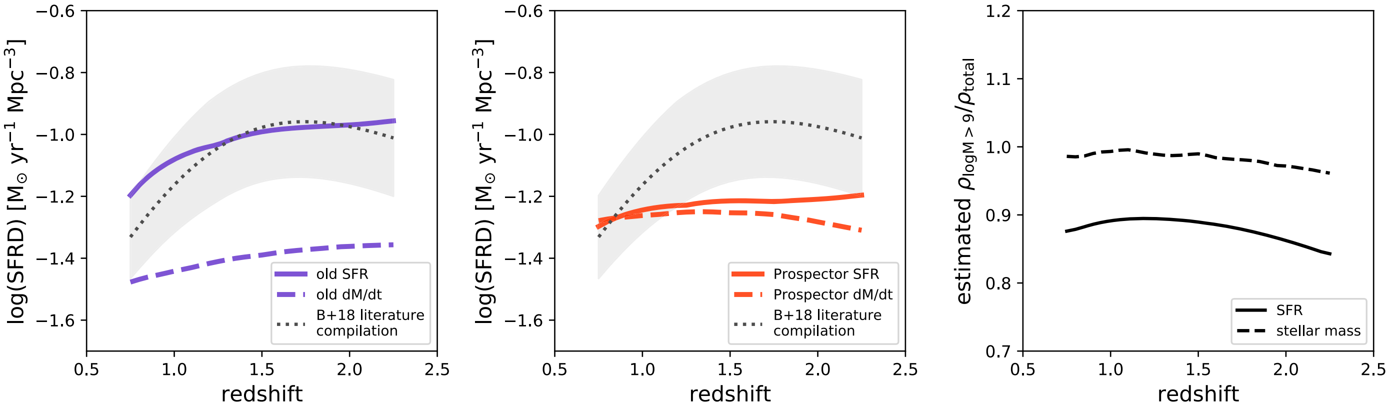

Figure 11. Comparison between the observed cosmic SFRD and the cosmic SFRD implied by the observed growth of stellar masses. The canonical values from the FAST SED-fitting code and $\mathrm{SFR_{UV+IR}}$ (left panel) disagree such that there is too much observed star formation by ∼0.2–0.4 dex. The revised estimates from `Prospector-`α (middle panel) largely remove this offset, due to a combination of lower SFRs (∼0.2 dex) and higher stellar masses (∼0.1 dex). As only galaxies with $\log(M_{\mathrm{FAST}}/M_{\odot}) > 9$ are modeled, a correction is applied for the stellar mass growth and star formation that occurs in galaxies with $\log(M_*) < 9$. This correction factor is measured from the Universe Machine (Behroozi et al. 2019) and shown in the right panel.

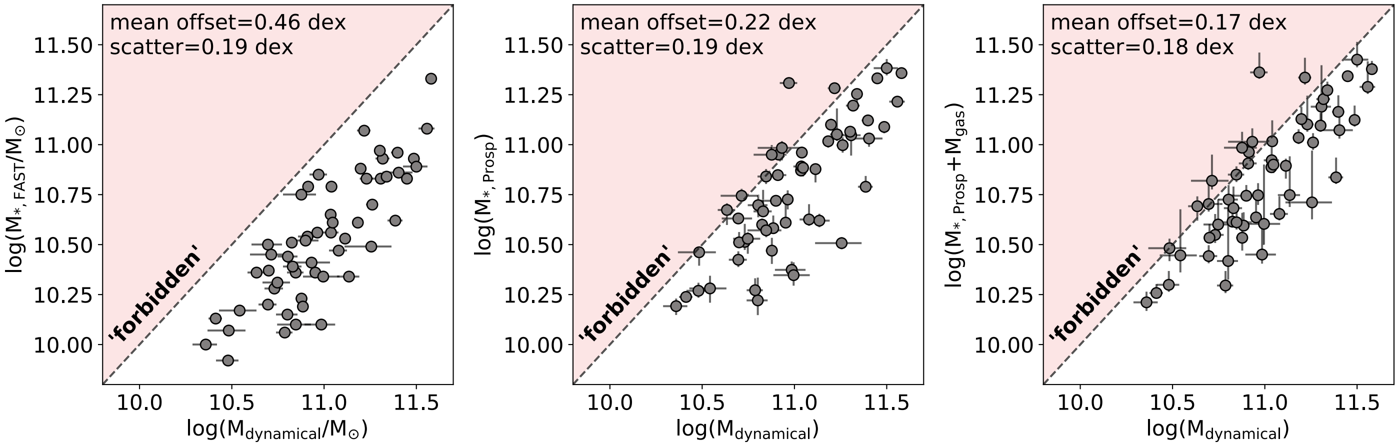

Figure 12. Comparison between stellar and dynamical masses. The left panel shows stellar masses from FAST, while the middle panel shows stellar masses from `Prospector-`α. The scatter is similar, though the offset decreases by ∼0.25 dex. The right panel includes molecular gas masses estimated from the scaling relationships of Tacconi et al. (2018). The outlier in the "forbidden" region of the middle panel has a poorly determined stellar mass and is consistent with the dynamical constraint at the 3σ level.

consistency enforced in `Prospector-`α, there is no guarantee that the global average of the stellar mass growth and SFR will agree. This makes the global $\lesssim$0.1 dex agreement quite remarkable.

5.3. Comparison to Dynamical Masses

Galaxy dynamical masses are an independent constraint on stellar masses. More specifically, since the total galaxy mass budget is composed of gaseous, stellar, and dark matter components, dynamical mass can be thought of as an "upper limit" to the stellar mass. Given that the `Prospector-`α model increases stellar masses by an average of ∼0.2 dex, it is important to ensure that the higher stellar masses do not violate dynamical constraints.

We test this with dynamical masses measured from deep Keck-DEIMOS spectra of star-forming and quiescent galaxies at $z \sim 0.7$ (Bezanson et al. 2015a). We adopt the structure-corrected dynamical masses calculated with the Sérsic-dependent virial constant from Cappellari et al. (2006). The dynamical masses are measured within the effective radius for each galaxy. We match 56 galaxies in the Bezanson et al. (2015a) sample to the 3D-*HST* photometric catalogs and fit these galaxies with `Prospector-`α using the spectroscopic redshifts from Bezanson et al. (2015a).

Figure 12 compares the measured dynamical masses to FAST stellar masses from Bezanson et al. (2015a) and to `Prospector-`α stellar masses. The mean $\log(M_{\mathrm{dyn}}/M_*)$ is 0.46 dex for FAST stellar masses and 0.22 dex for `Prospector-`α stellar masses. The final panel includes molecular gas masses estimated with the scaling relationships from Tacconi et al. (2018); this has a small overall effect, as most of the galaxies in this sample are quiescent. Crucially, this figure demonstrates that the distribution of `Prospector-`α masses does not violate the observed dynamical constraints. There is one object that is more massive than the dynamical constraints by ∼0.35 dex; however, it is consistent with the dynamical mass at the 3σ level owing to a long, non-Gaussian tail in the stellar mass posterior pdf.

While there is considerable scatter in $M_{\mathrm{Prospector}}/M_{\mathrm{FAST}}$, this scatter is not applied randomly, as it seems to respect the dynamical constraints. This is unlikely to occur as a result of chance: using the observed distribution of $M_{\mathrm{Prospector}}/M_{\mathrm{FAST}}$ and applying these offsets randomly to M_{FAST} shows that 98% of the time there should be more critical outliers (>0.1 dex mass discrepancy) than the single one observed here. This implies that the increased stellar mass inferred by `Prospector-`α is not added randomly, but instead is likely to reflect real variations in the underlying physical properties of these galaxies.

Overall, these results demonstrate that the new `Prospector-`α stellar masses are consistent with the direct dynamical constraints. The new masses do leave less room on average for additional massive components such as dark matter or a more bottom-heavy IMF. A key question is whether the maximal allowed dark matter fractions are "reasonable" compared to hydrodynamical simulations of ellipticals and spirals. At these redshifts and masses, the Illustris TNG simulation suggests that dark matter should constitute about 50% of the total matter within the effective radius (Lovell et al. 2018). This is closer to the revised stellar masses than the old stellar masses. Observational estimates of dark matter fractions necessarily rely on other methods to estimate stellar masses and in general create mixed expectations of the amount of dark matter within the effective radius. For example, Genzel et al. (2017) find that star-forming galaxies at $0.9 < z < 2.4$ have dark matter fractions of <0.22, but Tiley et al. (2019) argue that these should be considerably larger after correcting details of normalization prescription (they report dark matter fractions of $>60\%$ within six disk radii). Cappellari et al. (2013) use a variable IMF and measure dark matter fractions <0.4 in local early-type galaxies from the ATLAS-3D project. Ultimately, it is clear that the `Prospector-`α masses are consistent with the dynamical masses in the sense that the stellar mass alone does not violate the constraints; however, given uncertainties in dynamical masses and expected galaxy-to-galaxy scatter in dark matter fractions, it remains to be seen whether the `Prospector-`α masses are consistent with the full mass budget including dark matter.

6. Discussion

The accuracy of the updated physical parameters presented in this work is necessarily contingent on the accuracy of the 14-parameter `Prospector-`α model. Yet it can be challenging to perform hypothesis testing for high-redshift galaxy SED modeling owing to the large number of "unknowns" relative to "knowns." We first discuss the necessity of performing model cross-validation to further verify, dismantle, or alter the new picture presented in this work (Section 6.1). We then discuss potential future improvements in SED modeling that could further improve our interpretation of the observed galaxy photometry (Section 6.2).

6.1. Complex Models and Falsifiability

In this work we present new inferences of stellar masses and SFRs from a high-dimensional physical model for galaxy SEDs. This model pushes the field forward by allowing galaxy-to-galaxy variation for many components of galaxy formation that were fixed in previous work, such as the shape of the dust attenuation curve or the highly flexible step-function SFHs. This is possible because of advances in statistical and sampling methodologies, the ongoing and dramatic decrease in the price of computing time, and substantial improvements in SPS techniques.

The primary challenge in evaluating this model (or any such model) is that there is no "ground truth" with which to compare basic properties derived from galaxy SED fitting. Due to this lack of corroboration, there has been a long history of skepticism in the literature about the accuracy of galaxy SED modeling results (e.g., Papovich et al. 2001; Shapley et al. 2005; Conroy & Wechsler 2009; Wuyts et al. 2009; Behroozi et al. 2010; Taylor et al. 2011; Walcher et al. 2011; Mobasher et al. 2015; Santini et al. 2015).

Fitting simulated galaxies with galaxy SED models is a useful way to cross-examine their assumptions (e.g., Hayward & Smith 2015), as this is a scenario in which the ground truth is known. Simulations reproduce many key components of galaxy formation, including complex SFHs, physically motivated metallicity enrichment histories, and (for high-resolution simulations) complex spatial mixtures of stars and dust. However, such comparisons are only useful insofar as the physical conditions in simulated galaxies approximate those of real galaxies. It has been shown that the outputs of numerical simulations of galaxy formation are sensitive to the implementation of their subgrid physics (e.g., Crain et al. 2015). This is notable because different numerical simulations adopt different subgrid physics recipes (Somerville & Davé 2015). This means that the accuracy of simulation outputs varies according to the accuracy of their unique subgrid recipes, which are difficult to assess. Furthermore, it is only possible to use simulations to test SED-fitting ingredients that are *not* inputs to simulated galaxies. This forbids testing many basic components of galaxy SED models, including SPS assumptions, AGN emission models, and the subresolution behavior of dust and the interstellar medium (Smith & Hayward 2015; Nelson et al. 2018).

Given that a direct comparison between SED modeling results and ground truth is not possible, we suggest here that the next-best approach is to build a model that is, to the greatest extent possible, consistent with all other observations. This involves projecting the implications of galaxy SED models conditional on the observed data into the space of completely independent observables. Informative comparisons of this type can include comparing stellar masses to dynamical masses (Erb et al. 2006b; Taylor et al. 2010), predicting the strength of spectral features from fits to the photometry (Leja et al. 2017), and comparing SFHs of galaxies at low redshift to the observed SFRs and stellar masses of galaxies at higher redshift (Wuyts et al. 2011b). This approach is particularly fruitful for galaxy SED fitting: due to the covariance of basic parameters like age, dust, and metallicity, a simple change to the prior for one parameter can have ramifications for many other parameters of interest.

Figure 13 illustrates the potential for additional data to further constrain the parameters in the `Prospector-`α model. The top panel shows a model fit to photometry from the 3D-*HST* survey. The bottom panels show the joint pdf between key model parameters (sSFR, AGN strength, stellar metallicity, and stellar age) and potential future observables (Brγ emission equivalent width, Hδ and Fe λ5782 absorption equivalent width, and *WISE* rest-frame mid-infrared colors). The covariance between these parameters means that future observations can constrain key remaining uncertainties in the `Prospector-`α models. Notably, while these types of covariances are very common, the particular galaxy shown in Figure 13 is unusual in displaying strong covariances with all of these observables at once.

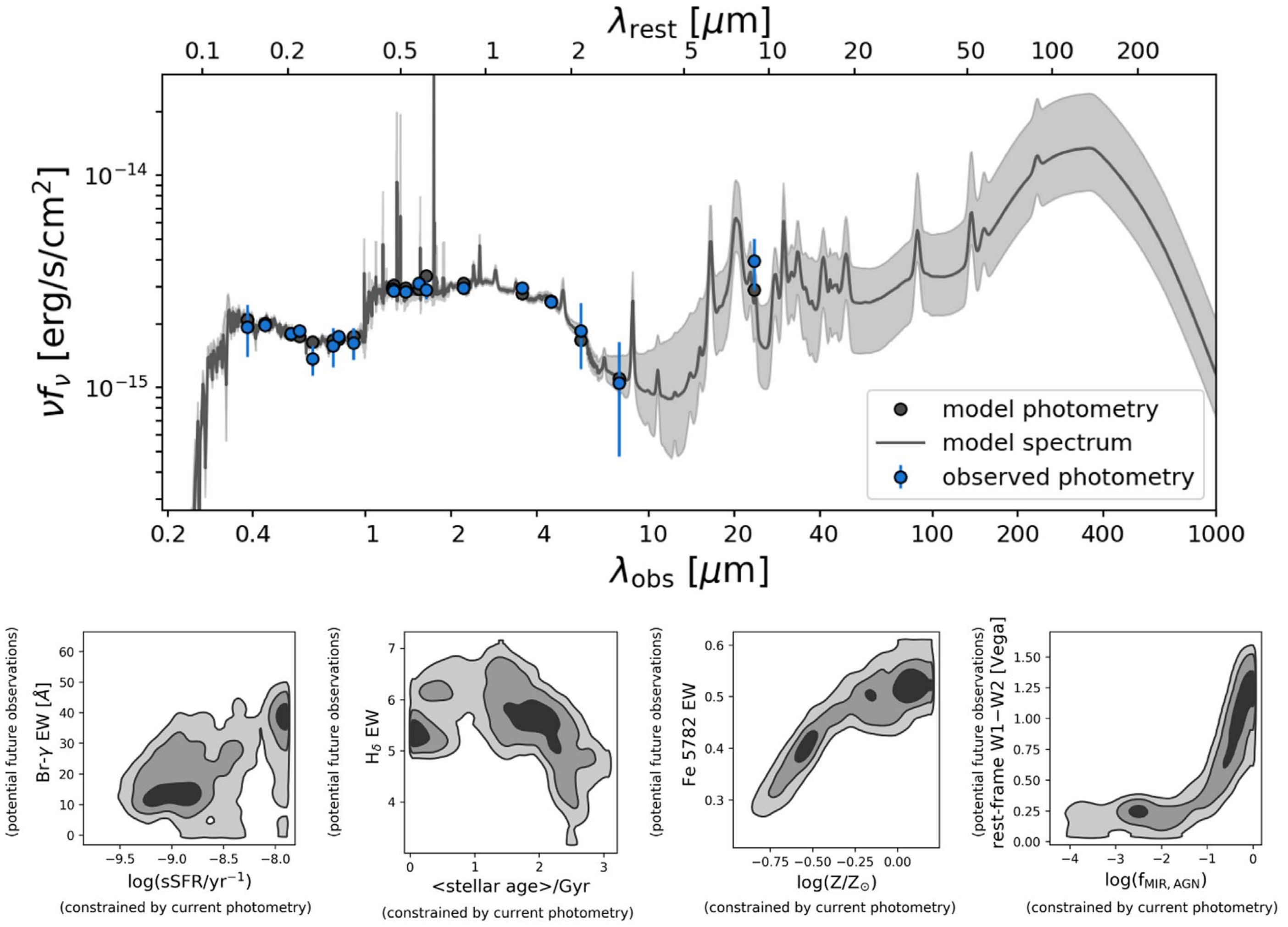

Figure 13. Future data have the potential to better constrain parameters in the `Prospector-`α model. The top panel shows the fit to the photometry of a galaxy from the 3D-*HST* catalogs, UDS 7610. The gray shaded region in the top panel represents the 1σ range of model spectra drawn from the posteriors. The bottom panels show predictions for future data that can constrain the major uncertainties in the `Prospector-`α posteriors. The shaded regions in the bottom panels correspond to 1σ, 2σ, and 3σ ranges.

6.2. Toward a More Accurate SED Model

One key improvement in `Prospector-`α is the large number of free parameters coupled with the statistical machinery to put realistic constraints on them. Allowing significant deviations from the "standard script" for galaxy formation permits more accurate properties to be inferred on a galaxy-by-galaxy basis.

However, there are still a number of key physical parameters that remain fixed. It is reasonable to think of models such as `Prospector-`α as one important step toward the ultimate goal, which is a fully flexible physical model for galaxy emission across all redshifts. One key step is to constrain the stellar attenuation as a function of age in order to confirm the global effect of old stellar heating as described in Section 4.4. Here we discuss several additional future steps on the path to this goal.

6.2.1. Propagation of Redshift Uncertainties

`Prospector-`α treats redshift as a fixed parameter. This approach explicitly neglects the effect of errors in distance determination on the resulting galaxy properties.

This assumption will affect some galaxy fits more than others. In the 3D-*HST* catalogs, redshift has been inferred independently from a combination of *HST* grism spectroscopy, ground-based spectroscopy, and photometric redshifts from EAZY. A fixed redshift is an excellent approximation for galaxies with solid spectroscopic or grism redshifts but is a less robust approximation for photometric redshifts. The reliability of photometric redshifts will also scale with the S/N of the photometry. For example, the scatter between photometric and spectroscopic redshifts for the entire 3D-*HST* survey is 0.0197, but for galaxies with $H_{\rm F160W}$ magnitude >26 this scatter increases to $\sim$0.05 (Bezanson et al. 2015b).

Redshift errors can have a strong effect on the physical parameters inferred from SED fitting. For example, Chevallard & Charlot (2016) use the Bayesian SED-fitting tool BEAGLE to fit two high-redshift galaxy SEDs simultaneously for redshift and stellar population parameters. The results show that redshift can have a complex interplay with the derived stellar population parameters: even moderate-redshift errors of $\sim$0.15 can affect individual stellar masses by a full order of magnitude or more. The systematic effect of redshift errors on global properties of the galaxy population—such as the stellar mass function or the cosmic SFR density—has yet to be characterized in a Bayesian framework.

One simple step forward is to use posteriors from photometric and grism redshift-fitting codes as priors for the redshift estimated in SED fitting. This is an imperfect solution, as it mixes multiple different assumptions about stellar populations. Ultimately, it would be ideal to use a single workflow to analyze all the available spectroscopic and photometric data and then simultaneously estimate redshifts and stellar population parameters; for a first step toward this, see Acquaviva et al. (2015).

6.2.2. A Flexible IR SED

In this work we adopt a fixed shape for the IR SED. This fixed shape is used to extrapolate the total infrared luminosity from the observed MIPS 24 μm photometry. The total infrared luminosity is a critical parameter, as it is closely related to the total SFR, though the exact relationship depends on the stellar properties as discussed in Section 4.4. Approximating the IR SED as fixed is helpful owing to the lack of MIR or FIR photometry at intermediate and high redshifts for the majority of the galaxy population. However, the IR SED shows significant variation on a galaxy-by-galaxy basis in the local universe (e.g., Dale et al. 2005, 2012), and this variation is likely to persist at higher redshifts.

Observations of variations in IR SED shape at higher redshifts are limited by the depth of available *Herschel* photometry. The handful of galaxies with individual detections in *Herschel* IR photometry show that the $L_{8\,\mu\mathrm{m}}/L_{\mathrm{IR}}$ ratio has a scatter of a factor of $\sim$2, with a tail toward higher values of $L_{8\,\mu\mathrm{m}}/L_{\mathrm{IR}}$ in systems with SFR $\gtrsim 100\,M_{\odot}\,\mathrm{yr}^{-1}$ (Elbaz et al. 2011; Wuyts et al. 2011a). Lower flux limits can be reached with stacking analysis. Shivaei et al. (2017) show that the $L_{8\,\mu\mathrm{m}}/L_{\mathrm{IR}}$ conversion is likely a strong function of stellar mass as well, varying *systematically* by a factor of 2 when comparing massive galaxies to galaxies with $\log(M/M_{\odot}) < 10$. $L_{8\,\mu\mathrm{m}}/L_{\mathrm{IR}}$ also shows significant redshift evolution (Whitaker et al. 2017).

A comprehensive study of the variation of the IR SED at $z > 0.5$ with galaxy properties has not yet been performed owing to the shallow limits of the available MIR and FIR imaging. Stacking or deblending *Herschel* photometry, combined with accurate galaxy properties from SED modeling, is one potential way to address this issue. It would be straightforward to incorporate these results into galaxy SED-fitting models via priors. Systematic change in the IR SED with galaxy properties has the potential to alter important galaxy scaling relationships such as the low-mass slope of the star-forming sequence in a mass-dependent fashion and correspondingly alter the cosmic SFR density (Whitaker et al. 2014; Leja et al. 2015; Shivaei et al. 2017).

6.2.3. α-element Abundances

Galaxy SED models currently assume a solar abundance pattern by necessity. However, there is clear evidence from high-resolution spectra of quiescent galaxies that the α-element abundance varies systematically with galaxy properties. This correlation is apparent in the nearby universe, where massive galaxies have $[\alpha/\mathrm{Fe}] \sim +0.23$ (Thomas et al. 2005; Conroy et al. 2014). This trend strengthens at intermediate redshifts ($0.5 < z < 2$), where small samples of massive galaxies have $[\alpha/\mathrm{Fe}] \sim +0.3$ (Choi et al. 2014; Onodera et al. 2015). More extreme individual causes have been detected, including $[\alpha/\mathrm{Fe}] > 0.4$ (Lonoce et al. 2015) and $[\alpha/\mathrm{Fe}] = +0.6$ (Kriek et al. 2016). It is more difficult to infer elemental abundance patterns in young galaxies owing to the lack of strong absorption lines, but simulations predict that star-forming galaxies have trends in α-element abundance patterns with mass, redshift, and SFR (Matthee & Schaye 2018). These can be $[\alpha/\mathrm{Fe}] = +0.6$ or higher in highly star-forming galaxies at $z = 2$ and above, consistent with observed nebular abundances (Steidel et al. 2016).

These trends in abundance patterns have ramifications for the integrated photometry of galaxies. Vazdekis et al. (2015) generate α-enhanced models with $[\alpha/\mathrm{Fe}] = +0.4$ and show that the resulting optical fluxes change by 10%–40% and the optical colors change by $\sim$0.1 mag, depending on the age and metallicity of the stars. This suggests that variations in α-element patterns should be included when fitting galaxy photometry: for example, α abundance patterns could be important in explaining the *ugr* colors of massive ellipticals, which have been too red in models for many years (e.g., Conroy & Gunn 2010; Vazdekis et al. 2015). Choi et al. (2019) show explicitly that synthesizing *ugriz* fluxes from the best-fit spectrum with individual elemental abundances allowed to vary will reproduce the observed colors to within <0.03 mag, while using solar-scaled abundances results in larger residuals (up to 0.1 mag for the oldest systems).

It remains unclear how much variation in α-element abundance will affect SED modeling at higher redshifts. On one hand, the α-element abundance patterns are more extreme at higher redshifts, but on the other hand, galaxies are younger on average and therefore less sensitive to α-element variations. Future versions of `FSPS` will include variation in the α-abundance pattern, providing a straightforward way to include the effect of variations in α-enhancement on galaxy properties derived from SED modeling.

6.2.4. IMF Variations

The shape of the stellar IMF is a critical assumption in galaxy SED modeling. Changing the IMF below $\sim 0.8\,M_{\odot}$ substantially changes inferred stellar masses and SFRs without significantly changing the predicted SED. Such a change does not affect the global agreement between the SFRD and the growth of the stellar mass density, as both SFR and mass are changed proportionally (Leja et al. 2015). Changing the high-mass end of the IMF will substantially change the inferred SFRs again without much consequence for the predicted SEDs, though this change would alter the global agreement between mass and SFR.

Recent work has provided solid evidence in nearby galaxies for long-suspected systematic variations in the stellar IMF between galaxies. IR spectroscopy (van Dokkum & Conroy 2010; Conroy & van Dokkum 2012), dynamical modeling (Cappellari et al. 2013), and gravitational lens analysis (Treu et al. 2010) all independently suggest that ellipticals with higher velocity dispersions have increasingly “bottom-heavy” IMFs, though there remains some tension in the exact agreement between these techniques (Newman et al. 2017). Star counts in ultra-faint dwarf galaxies find that these galaxies are deficient in low-mass stars (“bottom-light”; Geha et al. 2013). These results taken together are qualitatively consistent with a continuous variation in the IMF from low-mass to high-mass galaxies.

The ramifications of a variable IMF for the $z \gtrsim 1$ galaxy population have not been fully explored. This is at least in part because of the paucity of observables that directly correlate with the IMF (van Dokkum & Conroy 2012). There is also recent evidence that bottom-heavy IMFs might only be confined to the very central regions (Conroy et al. 2017; van Dokkum et al. 2017; Sarzi et al. 2018; Zhou et al. 2019), making comparisons to global quantities challenging. This results in a greater emphasis on indirect methods such as comparing the inferred stellar and dynamical masses. These comparisons typically assume canonical IMFs and find that the stellar mass takes up an increasing fraction of the total mass budget at higher redshift (van de Sande et al. 2013; Belli et al. 2017), or even exceeds the total mass budget (Price et al. 2019). This presents a difficult conundrum: if old galaxies in the local universe show bottom-heavy IMFs, why are these IMFs seemingly incompatible with dynamical measurements of their putative progenitor galaxies at $z \sim 2$? One intriguing possibility is that the star-forming progenitors of the cores of local elliptical galaxies have yet to be found owing to high levels of dust attenuation (Nelson et al. 2014).

Making progress on this issue will require careful simultaneous dynamical and SED modeling in order to satisfy both dynamical constraints and the observed photometry. Such work will be crucial to ensuring the absolute accuracy of SED-derived quantities.

7. Conclusions

In this work we present a revised estimate on the rate of galaxy stellar mass assembly at $0.5 < z < 2.5$ using the `Prospector-`α galaxy physical model. The primary advance over previous work is the much larger number of physical parameters that are modeled within `Prospector` ($N = 14$, compared to $N \sim 4$–7). This high dimensionality permits modeling the effect of a number of second-order physical effects on both stellar mass and SFR estimates on an object-by-object basis. These new high-dimensional SED models are possible owing to a number of technical improvements: the nested sampling routine `dynesty`, on-the-fly model generation with `FSPS`, and the `Prospector` Bayesian inference framework.

We fit a version of the `Prospector-`α physical model from Leja et al. (2017, 2018) modified for high-redshift galaxies. This model makes use of the wide range of physics available in `FSPS` and has a total of 14 free parameters. These physics include a flexible six-parameter nonparametric SFH, state-of-the-art MIST stellar isochrones, a broad range of stellar metallicities, a two-component dust attenuation model with a flexible dust attenuation curve, dust emission via energy balance, nebular line and continuum emission, and a model for the MIR emission of dusty AGN torii.

The `Prospector-`α model is fit to rest-frame UV–MIR photometry of 58,461 galaxies from the 3D-*HST* survey in the redshift range $0.5 < z < 2.5$. These catalogs provide an immense amount of information: there are between 17 and 44 bands of aperture-matched photometry available across five distinct extragalactic fields. These photometric data are coupled with redshifts inferred from a combination of ground-based spectroscopy, the *HST* G141 grism, and photometric redshifts from EAZY. After fitting these data, we present the following conclusions:

> The `Prospector-`α stellar masses are systematically 0.1–0.3 dex higher than stellar masses from the 3D-*HST* catalogs inferred with the FAST SED-fitting code. This offset correlates with stellar mass and, more weakly, with redshift.
> While multiple effects contribute at a low level, the primary cause of the offset is the older stellar ages inferred with `Prospector-`α. Comparing stacked SFHs inferred from the 3D-*HST* SED and the `Prospector-`α model shows that these differences can be dramatic: highly star-forming galaxies are older by a factor of ∼10 and galaxies on the star-forming sequence are older by a factor of ∼5.
> The `Prospector-`α SFRs match state-of-the-art UV+IR SFRs at high sSFRs ($\log(\mathrm{sSFR}/\mathrm{yr}^{-1}) \approx 8$). They are increasingly lower than $\mathrm{SFR}_{\mathrm{UV+IR}}$ with decreasing sSFR such that by $\log(\mathrm{sSFR}/\mathrm{yr}^{-1}) \approx -10.5$ there is an offset of 0.75–1 dex.
> While again multiple effects contribute, the largest cause of this offset is the emission from old stars. This is neglected in $\mathrm{SFR}_{\mathrm{UV+IR}}$ but self-consistently estimated in the `Prospector-`α model. The fraction of L_{IR}+L_{UV} powered by emission from "old" ($t > 100$ Myr) stars as a function of sSFR is derived, and an equation to estimate this effect is presented.

We explore the global implications of these new inferences with several model cross-validation techniques:

i. The global SFR density is estimated from the SED fits using both dM_*/dt and SFR(t). These two estimators are inconsistent when estimated with FAST stellar masses and $\mathrm{SFR}_{\mathrm{UV+IR}}$ in the sense that ρ_{SFR} is higher than $\dot{\rho}_{\mathrm{mass}}$ by ∼0.3 dex, in agreement with other studies in the literature. The `Prospector-`α estimates bring ρ_{SFR} down by ∼0.2 dex and $\dot{\rho}_{\mathrm{mass}}$ up by ∼0.1 dex such that there is now consistency in the inferred SFRD. This is a notable finding, as there is no guarantee of self-consistency in the cosmic sum of these values.

ii. The `Prospector-`α SFHs are much better predictors of the redshift evolution of the stellar mass function. This is demonstrated by using observed SFHs coupled with a merger model to wind the observed stellar mass function back in time. This model mass function is compared to the observed stellar mass functions to test the consistency of the SFHs. The `Prospector-`α SFHs are older on average and better describe the observations than the 3D-*HST* SFHs across most combinations of mass and redshift, though galaxies in the knee of the mass function ($10 < \log(M/M_\odot) < 11$) are likely too old within the `Prospector-`α model.

iii. The new stellar masses from `Prospector-`α are consistent with observed dynamical constraints, with the average offset between stellar and dynamical mass decreasing from ∼0.46 to ∼0.22 dex (0.17 dex when including gas), though the new masses do leave less room on average for additional components such as dark matter or a more bottom-heavy IMF.

The primary goal of this work is to build a model for galaxy properties that is, to the greatest extent possible, consistent with all observations. We take the first steps in this direction by performing cross-validation both within `Prospector-`α and with external data sets and by highlighting future observations that will provide deeper constraints for the `Prospector-`α

physical model. Such future data will lead to updates of model priors used in SED fitting. Due to the covariance of basic galaxy parameters, a change to the prior for one parameter will have ramifications for other parameters of interest: in this way such updates will create “evolving results.” It is hoped that this methodology can be used to converge toward the truth.

J.L. is supported by an NSF Astronomy and Astrophysics Postdoctoral Fellowship under award AST-1701487. We thank Sandra Faber, Sandro Tacchella, Maarten Baes, and Rohan Naidu for fruitful discussions. The computations in this paper were run on the Odyssey cluster supported by the FAS Division of Science, Research Computing Group at Harvard University. This research made use of Astropy,[13] a community-developed core Python package for Astronomy (Astropy Collaboration et al. 2013, 2018).

Software: `Prospector` (Johnson & Leja 2017), `python-fsps` (Foreman-Mackey et al. 2014), `Astropy` (Astropy Collaboration et al. 2013, 2018), `FSPS` (Conroy et al. 2009), `matplotlib` (Caswell et al. 2018), `scipy` (Jones et al. 2001), `ipython` (Pérez & Granger 2007), `numpy` (Walt et al. 2011).

ORCID iDs

Joel Leja https://orcid.org/0000-0001-6755-1315
Benjamin D. Johnson https://orcid.org/0000-0002-9280-7594
Charlie Conroy https://orcid.org/0000-0002-1590-8551
Pieter van Dokkum https://orcid.org/0000-0002-8282-9888
Joshua S. Speagle https://orcid.org/0000-0003-2573-9832
Gabriel Brammer https://orcid.org/0000-0003-2680-005X
Ivelina Momcheva https://orcid.org/0000-0003-1665-2073
Rosalind Skelton https://orcid.org/0000-0001-7393-3336
Katherine E. Whitaker https://orcid.org/0000-0001-7160-3632
Marijn Franx https://orcid.org/0000-0002-8871-3026
Erica J. Nelson https://orcid.org/0000-0002-7524-374X

References

Acquaviva, V., Gawiser, E., & Guaita, L. 2011, ApJ, 737, 47
Acquaviva, V., Raichoor, A., & Gawiser, E. 2015, ApJ, 804, 8
Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Bastian, N., & de Mink, S. E. 2009, MNRAS, 398, L11
Behroozi, P., Wechsler, R., Hearin, A., & Conroy, C. 2019, MNRAS, in press
Behroozi, P. S., Conroy, C., & Wechsler, R. H. 2010, ApJ, 717, 379
Bell, E. F., & de Jong, R. S. 2001, ApJ, 550, 212
Bell, E. F., Papovich, C., Wolf, C., et al. 2005, ApJ, 625, 23
Belli, S., Genzel, R., Förster Schreiber, N. M., et al. 2017, ApJL, 841, L6
Berta, S., Lutz, D., Santini, P., et al. 2013, A&A, 551, A100
Bezanson, R., Franx, M., & van Dokkum, P. G. 2015a, ApJ, 799, 148
Bezanson, R., Wake, D. A., Brammer, G. B., et al. 2015b, arXiv:1510.07049
Bouché, N., Dekel, A., Genzel, R., et al. 2010, ApJ, 718, 1001
Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503
Brinchmann, J., & Ellis, R. S. 2000, ApJL, 536, L77
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Burgarella, D., Buat, V., & Iglesias-Páramo, J. 2005, MNRAS, 360, 1413
Byler, N., Dalcanton, J. J., Conroy, C., & Johnson, B. D. 2017, ApJ, 840, 44
Calistro Rivera, G., Lusso, E., Hennawi, J. F., & Hogg, D. W. 2016, ApJ, 833, 98
Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682
Cappellari, M., Bacon, R., Bureau, M., et al. 2006, MNRAS, 366, 1126
Cappellari, M., McDermid, R. M., Alatalo, K., et al. 2012, Natur, 484, 485
Cappellari, M., Scott, N., Alatalo, K., et al. 2013, MNRAS, 432, 1709
Carnall, A. C., Leja, J., Johnson, B. D., et al. 2019, ApJ, 873, 44
Carnall, A. C., McLure, R. J., Dunlop, J. S., & Davé, R. 2018, MNRAS, 480, 4379
Caswell, T. A., Droettboom, M., Hunter, J., et al. 2018, matplotlib/matplotlib, v3.0.0, Zenodo, doi:10.5281/zenodo.1420605
Chabrier, G. 2003, PASP, 115, 763
Chanial, P., Flores, H., Guiderdoni, B., et al. 2007, A&A, 462, 81
Charlot, S., & Fall, S. M. 2000, ApJ, 539, 718
Chevallard, J., & Charlot, S. 2016, MNRAS, 462, 1415
Choi, J., Conroy, C., & Byler, N. 2017, ApJ, 838, 159
Choi, J., Conroy, C., & Johnson, B. D. 2019, ApJ, 872, 136
Choi, J., Conroy, C., Moustakas, J., et al. 2014, ApJ, 792, 95
Choi, J., Dotter, A., Conroy, C., et al. 2016, ApJ, 823, 102
Cid Fernandes, R., Mateus, A., Sodré, L., Stasińska, G., & Gomes, J. M. 2005, MNRAS, 358, 363
Ciesla, L., Charmandaris, V., Georgakakis, A., et al. 2015, A&A, 576, A10
Cohn, J. H., Leja, J., Tran, K.-V. H., et al. 2018, ApJ, 869, 141
Conroy, C. 2013, ARA&A, 51, 393
Conroy, C., Graves, G. J., & van Dokkum, P. G. 2014, ApJ, 780, 33
Conroy, C., & Gunn, J. E. 2010, ApJ, 712, 833
Conroy, C., Gunn, J. E., & White, M. 2009, ApJ, 699, 486
Conroy, C., & van Dokkum, P. G. 2012, ApJ, 760, 71
Conroy, C., van Dokkum, P. G., & Villaume, A. 2017, ApJ, 837, 166
Conroy, C., & Wechsler, R. H. 2009, ApJ, 696, 620
Cook, B. A., Conroy, C., van Dokkum, P., & Speagle, J. S. 2019, ApJ, 876, 78
Cortese, L., Boselli, A., Franzetti, P., et al. 2008, MNRAS, 386, 1157
Crain, R. A., Schaye, J., Bower, R. G., et al. 2015, MNRAS, 450, 1937
da Cunha, E., Charlot, S., & Elbaz, D. 2008, MNRAS, 388, 1595
Dale, D. A., Aniano, G., Engelbracht, C. W., et al. 2012, ApJ, 745, 95
Dale, D. A., Bendo, G. J., Engelbracht, C. W., et al. 2005, ApJ, 633, 857
Dale, D. A., & Helou, G. 2002, ApJ, 576, 159
Davé, R., Oppenheimer, B. D., & Finlator, K. 2011, MNRAS, 415, 11
Dekel, A., & Burkert, A. 2014, MNRAS, 438, 1870
De Looze, I., Fritz, J., Baes, M., et al. 2014, A&A, 571, A69
Diemer, B., Sparre, M., Abramson, L. E., & Torrey, P. 2017, ApJ, 839, 26
Dotter, A. 2016, ApJS, 222, 8
Driver, S. P., Andrews, S. K., da Cunha, E., et al. 2018, MNRAS, 475, 2891
Ekström, S., Georgy, C., Eggenberger, P., et al. 2012, A&A, 537, A146
Elbaz, D., Dickinson, M., Hwang, H. S., et al. 2011, A&A, 533, A119
Eldridge, J. J., Stanway, E. R., Xiao, L., et al. 2017, PASA, 34, e058
Erb, D. K., Shapley, A. E., Pettini, M., et al. 2006a, ApJ, 644, 813
Erb, D. K., Steidel, C. C., Shapley, A. E., et al. 2006b, ApJ, 646, 107
Ferland, G. J., Korista, K. T., Verner, D. A., et al. 1998, PASP, 110, 761
Ferland, G. J., Porter, R. L., van Hoof, P. A. M., et al. 2013, RMxAA, 49, 137
Finlator, K., Davé, R., & Oppenheimer, B. D. 2007, MNRAS, 376, 1861
Firmani, C., Avila-Reese, V., & Rodríguez-Puebla, A. 2010, MNRAS, 404, 1100
Foreman-Mackey, D., Sick, J., & Johnson, B. 2014, python-fsps: Python bindings to FSPS, v0.1.1, Zenodo, doi:10.5281/zenodo.12157
Gallazzi, A., Charlot, S., Brinchmann, J., White, S. D. M., & Tremonti, C. A. 2005, MNRAS, 362, 41
Geha, M., Brown, T. M., Tumlinson, J., et al. 2013, ApJ, 771, 29
Genel, S., Vogelsberger, M., Springel, V., et al. 2014, MNRAS, 445, 175
Genzel, R., Schreiber, N. M. F., Übler, H., et al. 2017, Natur, 543, 397
Genzel, R., Tacconi, L. J., Lutz, D., et al. 2015, ApJ, 800, 20
Gossage, S., Conroy, C., Dotter, A., et al. 2018, ApJ, 863, 67
Goudfrooij, P., Girardi, L., Kozhurina-Platais, V., et al. 2014, ApJ, 797, 35
Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, ApJS, 197, 35
Groves, B., Krause, O., Sandstrom, K., et al. 2012, MNRAS, 426, 892
Groves, B. A., Dopita, M. A., & Sutherland, R. S. 2004, ApJS, 153, 9
Guo, K., Zheng, X. Z., & Fu, H. 2013, ApJ, 778, 23
Hayward, C. C., & Smith, D. J. B. 2015, MNRAS, 446, 1512
Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, ApJS, 208, 19
Hopkins, A. M., & Beacom, J. F. 2006, ApJ, 651, 142
Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, A&A, 457, 841
Imara, N., Loeb, A., Johnson, B. D., Conroy, C., & Behroozi, P. 2018, ApJ, 854, 36
Iyer, K., & Gawiser, E. 2017, ApJ, 838, 127
Johnson, B., & Leja, J. 2017, bd-j/prospector: Initial release, Zenodo, doi: 10.5281/zenodo.1116491
Johnson, B. D., Weisz, D. R., Dalcanton, J. J., et al. 2013, ApJ, 772, 8
Jones, E., Oliphant, T., Peterson, P., et al. 2001, SciPy: Open Source Scientific Tools for Python, http://www.scipy.org/

---

[13] http://www.astropy.org

Kauffmann, G., Heckman, T. M., White, S. D. M., et al. 2003, MNRAS, 341, 33

Kennicutt, R. C., Jr. 1998, ARA&A, 36, 189

Kennicutt, R. C., & Evans, N. J. 2012, ARA&A, 50, 531

Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, ApJS, 197, 36

Kriek, M., Conroy, C., van Dokkum, P. G., et al. 2016, Natur, 540, 248

Kriek, M., Labbé, I., Conroy, C., et al. 2010, ApJL, 722, L64

Kriek, M., Shapley, A. E., Reddy, N. A., et al. 2015, ApJS, 218, 15

Kriek, M., van Dokkum, P. G., Labbé, I., et al. 2009, ApJ, 700, 221

Lee, B., Giavalisco, M., Whitaker, K., et al. 2018, ApJ, 853, 131

Leitner, S. N. 2012, ApJ, 745, 149

Leja, J., Carnall, A. C., Johnson, B. D., Conroy, C., & Speagle, J. S. 2019, ApJ, 876, 3

Leja, J., Johnson, B. D., Conroy, C., & van Dokkum, P. 2018, ApJ, 854, 62

Leja, J., Johnson, B. D., Conroy, C., van Dokkum, P. G., & Byler, N. 2017, ApJ, 837, 170

Leja, J., van Dokkum, P. G., Franx, M., & Whitaker, K. E. 2015, ApJ, 798, 115

Lilly, S. J., Carollo, C. M., Pipino, A., Renzini, A., & Peng, Y. 2013, ApJ, 772, 119

Lonoce, I., Longhetti, M., Maraston, C., et al. 2015, MNRAS, 454, 3912

Lovell, M. R., Pillepich, A., Genel, S., et al. 2018, MNRAS, 481, 1950

Madau, P., & Dickinson, M. 2014, ARA&A, 52, 415

Maraston, C., Daddi, E., Renzini, A., et al. 2006, ApJ, 652, 85

Maraston, C., Pforr, J., Renzini, A., et al. 2010, MNRAS, 407, 830

Marchesini, D., van Dokkum, P. G., Förster Schreiber, N. M., et al. 2009, ApJ, 701, 1765

Mason, C. A., Treu, T., Dijkstra, M., et al. 2018, ApJ, 856, 2

Matthee, J., & Schaye, J. 2018, MNRAS, 479, L34

Mitchell, P. D., Lacey, C. G., Baugh, C. M., & Cole, S. 2013, MNRAS, 435, 87

Mitchell, P. D., Lacey, C. G., Cole, S., & Baugh, C. M. 2014, MNRAS, 444, 2637

Mobasher, B., Dahlen, T., Ferguson, H. C., et al. 2015, ApJ, 808, 101

Momcheva, I. G., Brammer, G. B., van Dokkum, P. G., et al. 2016, ApJS, 225, 27

Nelson, D., Pillepich, A., Springel, V., et al. 2018, MNRAS, 475, 624

Nelson, E., van Dokkum, P., Franx, M., et al. 2014, Natur, 513, 394

Newman, A. B., Smith, R. J., Conroy, C., Villaume, A., & van Dokkum, P. 2017, ApJ, 845, 157

Noll, S., Burgarella, D., Giovannoli, E., et al. 2009, A&A, 507, 1793

Ocvirk, P., Pichon, C., Lançon, A., & Thiébaut, E. 2006, MNRAS, 365, 46

Onodera, M., Carollo, C. M., Renzini, A., et al. 2015, ApJ, 808, 161

Pacifici, C., Charlot, S., Blaizot, J., & Brinchmann, J. 2012, MNRAS, 421, 2002

Papovich, C., Dickinson, M., & Ferguson, H. C. 2001, ApJ, 559, 620

Papovich, C., Labbé, I., Quadri, R., et al. 2015, ApJ, 803, 26

Patel, S. G., Fumagalli, M., Franx, M., et al. 2013a, ApJ, 778, 115

Patel, S. G., van Dokkum, P. G., Franx, M., et al. 2013b, ApJ, 766, 15

Paxton, B., Bildsten, L., Dotter, A., et al. 2011, ApJS, 192, 3

Paxton, B., Cantiello, M., Arras, P., et al. 2013, ApJS, 208, 4

Paxton, B., Marchant, P., Schwab, J., et al. 2015, ApJS, 220, 15

Pérez, F., & Granger, B. E. 2007, CSE, 9, 21

Pforr, J., Maraston, C., & Tonini, C. 2012, MNRAS, 422, 3285

Piatti, A. E., & Cole, A. 2017, MNRAS, 470, L77

Price, S. H., Kriek, M., Barro, G., et al. 2019, arXiv:1902.09554

Rujopakarn, W., Rieke, G. H., Eisenstein, D. J., & Juneau, S. 2011, ApJ, 726, 93

Salim, S., Boquien, M., & Lee, J. C. 2018, ApJ, 859, 11

Salim, S., Rich, R. M., Charlot, S., et al. 2007, ApJS, 173, 267

Salmon, B., Papovich, C., Finkelstein, S. L., et al. 2015, ApJ, 799, 183

Salmon, B., Papovich, C., Long, J., et al. 2016, ApJ, 827, 20

Santini, P., Ferguson, H. C., Fontana, A., et al. 2015, ApJ, 801, 97

Sarzi, M., Spiniello, C., La Barbera, F., Krajnović, D., & van den Bosch, R. 2018, MNRAS, 478, 4084

Sawicki, M., & Yee, H. K. C. 1998, AJ, 115, 1329

Schmidt, K. B., Treu, T., Trenti, M., et al. 2014, ApJ, 786, 57

Shapley, A. E., Reddy, N. A., Kriek, M., et al. 2015, ApJ, 801, 88

Shapley, A. E., Steidel, C. C., Adelberger, K. L., et al. 2001, ApJ, 562, 95

Shapley, A. E., Steidel, C. C., Erb, D. K., et al. 2005, ApJ, 626, 698

Shivaei, I., Reddy, N. A., Shapley, A. E., et al. 2017, ApJ, 837, 157

Skelton, R. E., Whitaker, K. E., Momcheva, I. G., et al. 2014, ApJS, 214, 24

Skibba, R. A., Engelbracht, C. W., Dale, D., et al. 2011, ApJ, 738, 89

Smith, D. J. B., & Hayward, C. C. 2015, MNRAS, 453, 1597

Somerville, R. S., & Davé, R. 2015, ARA&A, 53, 51

Speagle, J. S. 2019, arXiv:1904.02180

Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, ApJS, 214, 15

Steidel, C. C., Strom, A. L., Pettini, M., et al. 2016, ApJ, 826, 159

Tacconi, L. J., Genzel, R., Saintonge, A., et al. 2018, ApJ, 853, 179

Tal, T., Dekel, A., Oesch, P., et al. 2014, ApJ, 789, 164

Taylor, E. N., Franx, M., Brinchmann, J., van der Wel, A., & van Dokkum, P. G. 2010, ApJ, 722, 1

Taylor, E. N., Hopkins, A. M., Baldry, I. K., et al. 2011, MNRAS, 418, 1587

Thomas, D., Maraston, C., Bender, R., & Mendes de Oliveira, C. 2005, ApJ, 621, 673

Tiley, A. L., Swinbank, A. M., Harrison, C. M., et al. 2019, MNRAS, 485, 934

Tojeiro, R., Heavens, A. F., Jimenez, R., & Panter, B. 2007, MNRAS, 381, 1252

Tomczak, A. R., Quadri, R. F., Tran, K.-V. H., et al. 2014, ApJ, 783, 85

Tomczak, A. R., Quadri, R. F., Tran, K.-V. H., et al. 2016, ApJ, 817, 118

Torrey, P., Vogelsberger, M., Genel, S., et al. 2014, MNRAS, 438, 1985

Torrey, P., Vogelsberger, M., Marinacci, F., et al. 2019, MNRAS, 484, 5587

Treu, T., Auger, M. W., Koopmans, L. V. E., et al. 2010, ApJ, 709, 1195

Utomo, D., Kriek, M., Labbé, I., Conroy, C., & Fumagalli, M. 2014, ApJL, 783, L30

van de Sande, J., Kriek, M., Franx, M., et al. 2013, ApJ, 771, 85

van de Sande, J., Kriek, M., Franx, M., Bezanson, R., & van Dokkum, P. G. 2015, ApJ, 799, 125

van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2014, ApJ, 788, 28

van Dokkum, P., Conroy, C., Villaume, A., Brodie, J., & Romanowsky, A. J. 2017, ApJ, 841, 68

van Dokkum, P. G., & Conroy, C. 2010, Natur, 468, 940

van Dokkum, P. G., & Conroy, C. 2012, ApJ, 760, 70

van Dokkum, P. G., Leja, J., Nelson, E. J., et al. 2013, ApJL, 771, L35

van Dokkum, P. G., Whitaker, K. E., Brammer, G., et al. 2010, ApJ, 709, 1018

Vazdekis, A., Coelho, P., Cassisi, S., et al. 2015, MNRAS, 449, 1177

Viaene, S., Baes, M., Tamm, A., et al. 2017, A&A, 599, A64

Vogelsberger, M., Genel, S., Springel, V., et al. 2014a, MNRAS, 444, 1518

Vogelsberger, M., Genel, S., Springel, V., et al. 2014b, Natur, 509, 177

Walcher, J., Groves, B., Budavári, T., & Dale, D. 2011, Ap&SS, 331, 1

Walt, S. v. d., Colbert, S. C., & Varoquaux, G. 2011, CSE, 13, 22

Whitaker, K. E., Franx, M., Leja, J., et al. 2014, ApJ, 795, 104

Whitaker, K. E., Pope, A., Cybulski, R., et al. 2017, ApJ, 850, 208

Wilkins, S. M., Trentham, N., & Hopkins, A. M. 2008, MNRAS, 385, 687

Wisnioski, E., Förster Schreiber, N. M., Wuyts, S., et al. 2015, ApJ, 799, 209

Wuyts, E., Wisnioski, E., Fossati, M., et al. 2016, ApJ, 827, 74

Wuyts, S., Förster Schreiber, N. M., Lutz, D., et al. 2011a, ApJ, 738, 106

Wuyts, S., Förster Schreiber, N. M., van der Wel, A., et al. 2011b, ApJ, 742, 96

Wuyts, S., Franx, M., Cox, T. J., et al. 2009, ApJ, 696, 348

Wuyts, S., Labbé, I., Förster Schreiber, N. M., et al. 2008, ApJ, 682, 985

Zhou, S., Mo, H. J., Li, C., et al. 2019, MNRAS, 485, 5256