











# Missing Giants: Predictions on Dust-obscured Galaxy Stellar Mass Assembly Throughout Cosmic Time

Arianna S. Long<sup>1,9</sup> , Caitlin M. Casey<sup>1</sup> , Claudia del P. Lagos<sup>2,3</sup> , Erini L. Lambrides<sup>4,10</sup> , Jorge A. Zavala<sup>5</sup> ,  
Jaclyn Champagne<sup>6</sup> , Olivia R. Cooper<sup>7,11</sup> , and Asantha R. Cooray<sup>8</sup> 

<sup>1</sup> Department of Astronomy, The University of Texas at Austin, Austin, TX, USA; [arianna.sage.long@gmail.com](mailto:arianna.sage.long@gmail.com)

<sup>2</sup> International Centre for Radio Astronomy Research (ICRAR), M468, University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia

<sup>3</sup> ARC of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Australia

<sup>4</sup> NASA Goddard Space Flight Center, 8800 Greenbelt Road Greenbelt, MD 20771, USA

<sup>5</sup> National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan

<sup>6</sup> Steward Observatory, University of Arizona, 933 N Cherry Avenue, Tucson, AZ 85721, USA

<sup>7</sup> The University of Texas at Austin, Department of Astronomy, Austin, TX, USA

<sup>8</sup> Department of Physics & Astronomy, University of California, Irvine, 4129 Reines Hall, Irvine, CA 92697, USA

Received 2022 November 3; revised 2023 May 28; accepted 2023 June 11; published 2023 July 28

## Abstract

Due to their extremely dust-obscured nature, much uncertainty still exists surrounding the stellar mass growth and content in dusty, star-forming galaxies (DSFGs) at  $z > 1$ . In this work, we present a numerical model built using empirical data on DSFGs to estimate their stellar mass contributions across the first  $\sim 10$  Gyr of cosmic time. We generate a dust-obscured stellar mass function that extends beyond the mass limit of star-forming stellar mass functions in the literature, and predict that massive DSFGs constitute as much as 50%–100% of all star-forming galaxies with  $M \geq 10^{11} M_{\odot}$  at  $z > 1$ . We predict the number density of massive DSFGs and find general agreement with observations, although more data is needed to narrow wide observational uncertainties. We forward-model mock massive DSFGs to their quiescent descendants and find remarkable agreement with observations from the literature demonstrating that, to first order, massive DSFGs are a sufficient ancestral population to describe the prevalence of massive quiescent galaxies at  $z > 1$ . We predict that massive DSFGs and their descendants contribute as much as 25%–60% to the cosmic stellar mass density during the peak of cosmic star formation, and predict an intense epoch of population growth during the  $\sim 1$  Gyr from  $z = 6$  to 3 during which the majority of the most massive galaxies at high- $z$  grow and then quench. Future studies seeking to understand massive galaxy growth and evolution in the early universe should strategize synergies with data from the latest observatories (e.g., JWST and the Atacama Large Millimeter/submillimeter Array) to better include the heavily dust-obscured galaxy population.

*Unified Astronomy Thesaurus concepts:* [Galaxy evolution \(594\)](#); [Luminous infrared galaxies \(946\)](#); [High-redshift galaxies \(734\)](#)

## 1. Introduction

Most of what we understand about galaxy evolution in the first 6 Gyr of the universe is largely based on rest-frame ultraviolet-to-optical studies. This includes, but is not limited to: the evolution of the galaxy stellar mass function (SMF; e.g., Peng et al. 2010; Muzzin et al. 2013; Davidzon et al. 2017; Thorne et al. 2021) used to describe the growth of galaxies over time; the large-scale structure of galaxies (e.g., Eisenstein et al. 2011; Zehavi et al. 2011) that demonstrates how galaxies are distributed throughout space as they form and grow; and the star-forming main sequence (e.g., Whitaker et al. 2014 and Table 3 in Speagle et al. 2014), establishing that the majority of galaxies form their stars steadily, over long periods of time ( $\sim 1$  Gyr), at a rate proportional to their stellar mass. These hereinafter referred to as “UV/OP-based” relationships are often perceived as fundamental truths that are either fed into or used as benchmarks for success in testing our best

cosmological models (e.g., Somerville & Davé 2015; Wechsler & Tinker 2018). In other words, our primary understanding of galaxy evolution—from observations to simulations—is deeply and intricately rooted in UV/optical astronomy, including all of its strengths and limitations.

It is now well known that the rest-frame UV/optical spectrum of the universe contains only a piece of the larger picture on star formation in the cosmos. Nearly half of all cosmic starlight in the universe is obscured by dust (Driver et al. 2008), which re-radiates at longer infrared wavelengths. This phenomenon is most evident in the cosmic star formation rate (SFR) density where, out to  $z \sim 3$ , the majority of stellar mass is built in regions obscured by dust and thus primarily detected via infrared (IR) observations (see Madau & Dickinson 2014, for a review). This means that studies using only UV/optical to chart stellar mass assembly throughout cosmic time are incomplete and biased against significantly dusty systems.

Over the last two decades, wide-field far-IR (FIR) and submillimeter surveys ( $\lambda > 40 \mu\text{m}$ ) have discovered significant populations of dusty, star-forming galaxies (DSFGs; see Casey et al. 2014 for a review). Though rare in the local universe, DSFGs are 1000 times more populous at  $z \sim 1$ –3 (e.g., Simpson et al. 2014; Zavala et al. 2021), dominating cosmic star formation during these times. Their extreme rates of star formation ( $\sim 10^{2-3} M_{\odot} \text{ yr}^{-1}$ ) generate massive reservoirs of

<sup>9</sup> NASA Hubble Fellow.

<sup>10</sup> NPP Fellow.

<sup>11</sup> NSF Graduate Fellow.



dust that obscures the starlight, making some DSFGs nearly invisible to even the deepest UV/optical/near-IR observations, but bright like beacons at FIR and (sub)millimeter wavelengths (e.g., da Cunha et al. 2015; Williams et al. 2019; Manning et al. 2022).

While some DSFGs are detected in the deepest of UV/optical/near-IR surveys, they are generally left out of large surveys broadly describing galaxy evolution. This usually happens through e.g., stringent requirements for photometric coverage over the rest-frame UV/optical spectrum (e.g., requiring  $r$ -,  $i$ -, and  $z$ -band detections; Sherman et al. 2020), and/or “redder” selection bands with observations that are still insufficiently deep to capture the heavily obscured (and redshifted) starlight (e.g.,  $K_s$  band  $< 24$  mag; Lee et al. 2015). While often the best available tool for such studies, these requirements are biased against sources with significant attenuation in their UV/optical spectral energy distributions (SEDs); and, if DSFGs are captured, they are often treated as contaminants to the broader goals of the study, to be either removed or ignored as an insignificantly small portion of the broader galaxy population (e.g.,  $< 10\%$  at its peak fraction; Sargent et al. 2012; Hwang et al. 2021). Thus, it is likely that several of the fundamental UV/OP-based scaling relations used to describe galaxy evolution and inform our finest models and simulations are lacking a proper accounting of the DSFG population—a population that drives and dominates stellar mass assembly in the first  $\sim 6$  Gyr of the cosmos.

Indeed, the astrophysical community has been struggling to properly model and reproduce DSFG populations—and their measured physical characteristics—since their discovery (e.g., Granato et al. 2000; Baugh et al. 2005; Hopkins et al. 2010; Hayward et al. 2013; Aoyama et al. 2019; McAlpine et al. 2019). This may have to do with a need for more sophisticated dust radiative transfer treatments, degeneracies in constraining DSFG stellar initial mass functions (IMFs), complex dust and star geometries, and/or a general lack of constraints on the physical origins/triggering mechanisms of DSFGs (see, e.g., Hayward et al. 2021 for an in-depth discussion). Many of the same simulations struggling to properly reproduce DSFG population observations also struggle to produce sufficient populations of massive ( $M \sim 10^{11} M_\odot$ ) quiescent galaxies at early times (e.g., Brennan et al. 2015; Merlin et al. 2019). It is possible that these mysteries and shortcomings are related because it is also likely that these galaxy subpopulations are related.

Several lines of evidence suggest that the brightest DSFGs constitute the primary ancestral population of the most massive quiescent galaxies in the universe. This includes comparable comoving number densities ( $n \sim 10^{-5} - 10^{-6} \text{ Mpc}^{-3}$ ; e.g., Simpson et al. 2014; Toft et al. 2014), dark matter halo masses and clustering properties ( $M \sim 10^{12-13} M_\odot$ ; e.g., Hickox et al. 2012; Lim et al. 2020), physical shapes and effective radii ( $r_{\text{eff}} \sim 2 \text{ kpc}$ ; e.g., Hodge et al. 2016), star formation histories (strong and bursty with SFRs  $\gtrsim 10^{2-3} M_\odot \text{ yr}^{-1}$  over 50–100 Myr; e.g., Valentino et al. 2020), and stellar masses ( $\sim 10^{11} M_\odot$ ; e.g., da Cunha et al. 2015; Miettinen et al. 2017). Prior studies have also demonstrated a clear relationship between stellar mass and dustiness, where the most massive galaxies are more dust-obscured across cosmic time (van Dokkum et al. 2006; Pannella et al. 2009; Wuyts et al. 2011; Whitaker et al. 2012, 2017; Deshmukh et al. 2018; Algera et al. 2023). Recent discoveries of massive,  $z > 3$  quiescent galaxies

demand short, bursty star formation histories in line with those exhibited by DSFGs (e.g., Glazebrook et al. 2017; Schreiber et al. 2018a). And, finally, the SHARK semianalytical model (which is largely successful in modeling on-sky DSFG populations) suggests that the majority of the most massive galaxies ( $M > 10^{11} M_\odot$ ) at  $z \approx 1$  undergo a (sub)millimeter-bright phase at  $z > 1$  (e.g., Lagos et al. 2020). Thus, in order for us to understand how the most massive galaxies form and quench during the first half of the cosmic time, it is of the utmost importance that we are dogged in our pursuit toward understanding massive, dust-obscured galaxies and placing them into the wider context of massive galaxy evolution at  $z > 1$ .

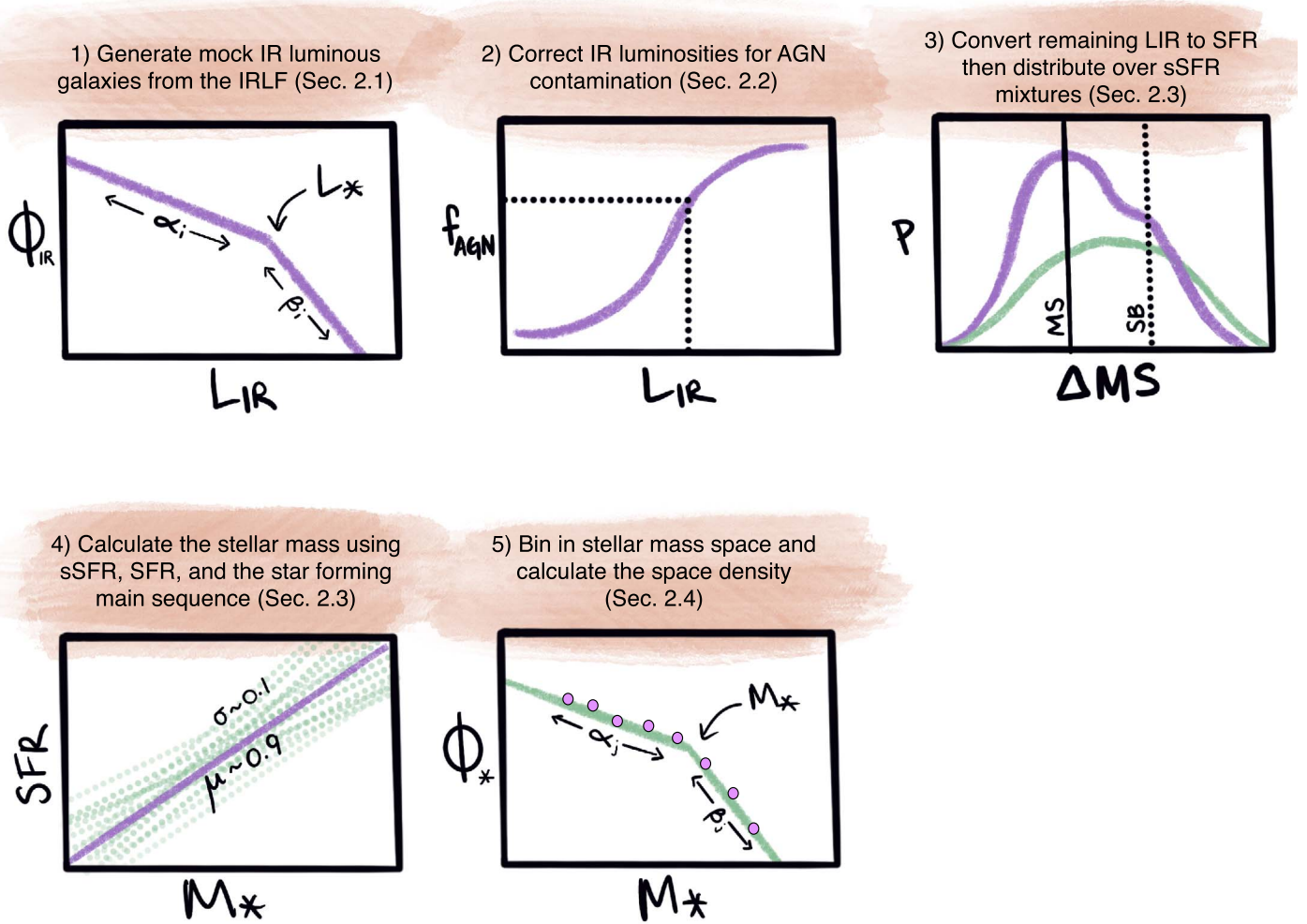
The objectives of this work are twofold. The first is to quantify, to first order, the significance of missing DSFG populations at the massive end of the star-forming galaxy SMF. Star-forming SMFs are biased against dusty galaxies as they are determined using UV/optical tracers. Infrared luminosity functions (IRLFs), however, primarily trace dust emission from the star-forming regions of galaxies. Combined, the UV/OP and IRLFs make up the cosmic SFR density known today (Madau & Dickinson 2014). Still missing from the picture, however, is an evolutionary depiction of the stellar mass contributions from dust-obscured star-forming galaxies. Using the IRLF as a nearly complete consensus of DSFGs out to  $z \sim 3-4$ , we seek to quantify what the dust-obscured SMF might look like, and posit that it would generate more numerous massive galaxies than previously measured in UV/OP-based SMFs.

The second objective is to test the hypothesis that, despite the heterogeneity of the DSFG population, it is sufficient and complete enough to describe and model the assembly of massive quiescent galaxies in the early universe. Using a suite of simple assumptions to forward evolve mock populations of DSFGs, we are able to reproduce the observed evolution of massive, passive galaxies—without needing to invoke any complex assumptions on merger histories, feedback/quenching mechanisms, or even contributions from outside populations (e.g., blue nuggets; Miettinen et al. 2017).

This paper is organized as follows: in Section 2 we describe the basis assumptions and empirical data used to build the model used in our analysis; in Section 3.1 we describe the resulting dust-obscured SMF from our model and how we forward evolve mock galaxies to create their quiescent descendants (Section 4.3.1); in Section 4 we present the number and stellar mass density predictions and compare to those in the literature; in Section 5, we discuss some additional implications of this model, and provide a summary in Section 6. Throughout this work, we adopt a Planck cosmology, where  $H_0 = 67.7 \text{ km s}^{-1} \text{ Mpc}^{-1}$  and  $\Omega_\Lambda = 0.692$  (Planck Collaboration et al. 2016); where relevant, we adopt a Chabrier IMF (Chabrier 2003).

## 2. Model Ingredients

Recent theoretical and observational studies have shown that the most massive star-forming galaxies in the  $z > 1$  universe are heavily obscured by dust (van Dokkum et al. 2006; Pannella et al. 2009; Wuyts et al. 2011; Whitaker et al. 2012, 2017; Deshmukh et al. 2018; Algera et al. 2023). This would mean that a significant portion of cosmic stellar mass is enshrouded in dust, and therefore difficult to detect with traditional photometric methods in the rest-frame UV/optical spectrum.



**Figure 1.** A cartoon of the model presented in this work. Several of the parameters highlighted in this cartoon are randomly sampled to account for uncertainties, variation, and unknowns across the literature. (1) The model begins by generating mock populations of IR luminous galaxies according to the IRLF defined in Casey et al. (2018a) and Zavala et al. (2021); see also Section 2.1. (2) Then, mock galaxy IR luminosities are corrected to remove AGN contributions in the infrared spectrum (Section 2.2). (3) We then convert the corrected IR luminosities to SFRs and use these SFRs to randomly distribute galaxies in specific SFR (sSFR) space according to various Gaussian mixtures of main-sequence vs. starburst galaxies (a.k.a. distance from the star-forming main sequence, Section 2.3.1; see also Figure 2). (4) Next, we generate a star-forming main sequence at the coeval redshift of interest (Section 2.3) and use this relationship, combined with the SFR and sSFR, to back out the galaxy stellar masses. (5) Finally, we re-bin the galaxies in stellar mass space to derive space densities (Section 2.4). This process is repeated 1000 times per  $1 \text{ deg}^2$  redshift slice (i.e.,  $\delta z$ ) for a given main sequence vs. starburst ratio, taking care to randomly sample the relevant parameter space (see Table 1), then combined to derive an overall SMF for dust-obscured galaxies (see Section 3).

Our goal in this work is to use existing empirical data on DSFGs to test this hypothesis and quantify the impact of DSFGs on the massive end of galaxy evolution. In this section, we describe the empirical data and assumptions used to build our numerical model, as well as the model itself. There are four main steps to populate the model, each described and motivated in greater detail in the proceeding subsections. Specifically, over  $6 \times 10^3$  realizations (i.e., light cones) per redshift bin, we:

1. Use the IRLF to derive a mock population of DSFGs (Section 2.1).
2. Correct IR luminosities for active galactic nucleus (AGN) contamination so that  $L_{IR}$  accounts for only star formation emission (Section 2.2).
3. Convert  $L_{IR}$  to SFR and distribute SFRs over the star-forming main sequence (i.e., the  $M_*$ -SFR relation) to derive stellar masses for the mock DSFGs (Section 2.3).

4. Bin in stellar mass space and compute the volume-averaged stellar mass density (Section 2.4).

In Figure 1, we share a cartoon depiction of our model. The model is realized in a light cone of solid angle  $1 \text{ deg}^2$  over a given redshift interval; the redshift windows cover  $0.5 < z < 6$  and are set to match existing SMFs derived from deep and wide surveys (e.g., Muzzin et al. 2013; Davidzon et al. 2017). All four steps occur within each of the  $6 \times 10^3$  realizations, with each realization drawing from probability density functions that mimic uncertainties based on empirical measurements. In Table 1, we summarize the parameters and their probability distributions that we sample in this model. See the specific subsections that follow for details and motivation on the chosen parameters and distributions, and note that the factor of 6 in the number of light cone realizations hails from the range of distributions for assigning stellar mass to mock galaxies, as detailed in Section 2.3.1.

**Table 1**  
Physical Relationships and Respective Parameter Spaces Used in This Model

Physical Relationship	References	Section	Parameter Space and Modeled Distribution
Infrared Luminosity Function	Casey et al. (2018b) and Zavala et al. (2021)	2.1	Same as in Casey et al. (2018b, Table 3 therein), except for $\psi(z > z_{\text{turn}})$ and $\alpha_{\text{LF}}$ , reported in Zavala et al. (2021) as $-6.5^{+0.8}_{-1.8}$ and $-0.42^{+0.02}_{-0.04}$ , respectively. We employ skewed Gaussians for both parameters.
AGN Host Selection	Kartaltepe et al. (2010)	2.2	$f_{\text{AGN Hosts}} = 4\% - 100\%$ over an S-shaped curve increasing as a function of $L_{\text{IR}}$ (defined in Figure 30, therein). We interpolate over the S-shaped curve in each $L_{\text{IR}}$ bin.
$L_{\text{IR}}$ Correction for AGN Hosts	Kirkpatrick et al. (2015)	2.2	We sample the cumulative distribution function of composite and AGN hosts in Figure 3 therein to randomly assign $f_{\text{AGN MIR}}$ . Then, we use the quadratic equation (Equation (5) therein) to assign $f_{\text{AGN TIR}}$ . We employ Gaussians for the quadratic equation variables.
Star-forming Main Sequence	Speagle et al. (2014)	2.3	For the slope, we sample a Gaussian with $\mu = 0.9$ and $\sigma = 0.1$ . We keep the time evolution components (Equation (28), therein), employing Gaussians for the parameters related to time. The starburst region is fixed to 0.6 dex above the main sequence at a given stellar mass. We fix $\sigma_{\text{MS}} = \sigma_{\text{SB}} = 0.3$ dex.
Main Sequence versus Starburst Decomposition	S12, D15, S15, M17, P19, B18, D20	2.3.1	We step through six double Gaussian distributions, each with one mean centered on the main sequence and the other centered on the starburst regime (+0.6 dex). The distributions match the following starburst fractions: 50%, 40%, 30%, 20%, 10%, and 0%.

Finally, Section 3.1 describes how we combine these realizations to create our resulting SMF, and provides the parameters describing the shape of this dust-obscured SMF.

### 2.1. The Infrared Luminosity Function

We seed our model by generating mock populations of IR luminous galaxies according to the IRLF, with the structure described by Casey et al. (2018a, 2018b) and updated parameters measured in Zavala et al. (2021). This IRLF is empirically constrained at  $z < 3$  by well-known DSFG SEDs and physical properties derived using Spitzer and Herschel data (Casey et al. 2018b), and at  $z > 3$  using Atacama Large Millimeter/submillimeter Array (ALMA) observations (Zavala et al. 2021). Below, we briefly describe the IRLF parameterization and constraints used in this work, and refer the reader to Casey et al. (2018a, 2018b) for more details.

The IRLF in this model takes the form of a double power law such that:

$$\Phi(L, z) = \begin{cases} \Phi_*(z) \left( \frac{L}{L_*(z)} \right)^{\alpha_{\text{LF}}(z)} & \text{if } L < L_*(z) \\ \Phi_*(z) \left( \frac{L}{L_*(z)} \right)^{\beta_{\text{LF}}(z)} & \text{if } L \geq L_*(z) \end{cases} \quad (1)$$

where  $L_*$  (in  $L_{\odot}$ ) represents the evolving characteristic “knee” of the luminosity function;  $\alpha_{\text{LF}}$  represents the faint-end slope of the IRLF (below  $L_*$ ),  $\beta_{\text{LF}}$  represents the bright-end slope of the IRLF (above  $L_*$ ), and  $\Phi_*$  (in  $\text{Mpc}^{-3} \text{dex}^{-1}$ ) represents the evolving characteristic number density of the luminosity function.

Both  $L_*$  and  $\Phi_*$  evolve with redshift, reaching a “turning point” at  $z_{\text{turn}} \sim 2$  (corresponding to the peak of the CSFRD) such that, at  $z > z_{\text{turn}}$ ,  $L_*$  and  $\Phi_*$  evolve with a different slope than at  $z < z_{\text{turn}}$ . In other words,  $L_*$  and  $\Phi_*$  evolve as  $(1+z)^{\gamma}$  and  $(1+z)^{\psi}$ , respectively, with  $\gamma$  and  $\psi$  changing values at  $z \approx z_{\text{turn}}$ .

In general, the parameters describing the IRLF at  $z \lesssim z_{\text{turn}}$  are observationally well constrained. We use the values listed in

Table 3 of Casey et al. (2018b), with the exception of  $\psi$  at  $z > 2$  for which we use the value measured in Zavala et al. (2021), and refer the reader to these works for more details on how these values were determined. Specifically, we fix the bright-end slope to  $\beta_{\text{LF}} = -3.0$ , the redshift evolution of  $L_*$  to  $\gamma = 2.8$ , and the redshift evolution of  $\Phi_*$  to  $\psi = 0.0$ . Together, these values successfully reproduce the measured IRLF,  $L_*$ ,  $\Phi_*$ , and the IR luminous galaxy contributions to the CSFRD at  $0 < z \lesssim 2$  (see Appendix A.1 in Casey et al. 2018b).

At  $z > z_{\text{turn}}$ , there is much less information on the evolution of  $L_*$  and  $\Phi_*$ . For  $L_*$ , studies report  $\gamma$ -values as steep as  $\sim 1.6$  (Gruppioni et al. 2013) and as flat as 0.2 (Koprowski et al. 2017); this IRLF assumes a redshift evolution of  $\gamma = 1$ , which is in line with the assumption that  $L_*$  evolves toward brighter luminosities at higher redshifts—as seen in several observational studies (e.g., Gruppioni et al. 2013; Koprowski et al. 2017; Lim et al. 2020). Higher values (e.g.,  $\gamma > 1.5$ ) would create unphysically bright IR luminous galaxies at  $z > 4$  that are not known to exist.

For the evolution of  $\Phi_*$  beyond  $z_{\text{turn}}$ , studies report  $\psi$  values from  $\lesssim -10$  (Koprowski et al. 2017) to  $-0.5$  (Lim et al. 2020), where more positive values correspond to a dust-rich early universe, and more negative values correspond to a dust-poor early universe. In other words, a dust-poor early universe would manifest as a CSFRD that is not dominated by obscured star formation at  $z > z_{\text{turn}}$ , and there would be significantly fewer  $z > 3$  DSFGs in blank-field submillimeter surveys than there are at  $z \sim 2$ . In this work, we adopt the value presented in Zavala et al. (2021) where  $\psi = -6.5^{+0.8}_{-1.8}$ , corresponding to a more dust-poor early universe. This  $\psi$ -value reproduces the most recent ALMA number counts at 1, 2, and 3 mm, as well as number counts from single-dish telescopes at  $700 \mu\text{m}$  to 1.1 mm.

The average faint-end slope collated in Casey et al. (2018b) is  $\alpha_{\text{LF}} = -0.6$ . However, recent works suggest a potentially shallower slope of  $\alpha_{\text{LF}} = [-0.2, -0.5]$  (Koprowski et al. 2017; Zavala et al. 2018; Gruppioni et al. 2020; Lim et al. 2020; Zavala et al. 2021). For this work, we adopt the value reported in Zavala et al. (2021) where  $\alpha_{\text{LF}} = -0.42^{+0.02}_{-0.04}$ , this value is

**Table 2**  
Infrared Luminosity Function Parameter Values Adopted in This Model

Parameter	Value	References
$z \lesssim z_{\text{turn}}$		
$\alpha_{\text{LF}}$	-0.6	Casey et al. (2018b)
$\beta_{\text{LF}}$	-3.0	Casey et al. (2018b)
$\gamma$	2.8	Casey et al. (2018b)
$\psi$	0.0	Casey et al. (2018b)
$z > z_{\text{turn}}$		
$\alpha_{\text{LF}}$	$-0.42^{+0.02}_{-0.04}$	Zavala et al. (2021)
$\beta_{\text{LF}}$	-3.0	Casey et al. (2018b)
$\gamma$	1.0	Casey et al. (2018b)
$\psi$	$-6.5^{+0.8}_{-1.8}$	Zavala et al. (2021)

aligned with the best-fit faint-end slopes in Koprowski et al. (2017), Zavala et al. (2018), and Lim et al. (2020).

For this analysis, we randomly sample the parameter space for  $\psi$  and  $\alpha_{\text{LF}}$  as listed above and defined in Zavala et al. (2021). The parameters are modeled as skewed Gaussians, with values drawn for every realization.

### 2.1.1. Summary of Mock IR Galaxy Catalog Generation

In review, for each of the  $10^3$  model realizations (i.e.,  $10^3$  light cones), we first randomly sample IRLF parameter values from their aforementioned distributions. Specifically, we sample from a skewed Gaussian built using the values and uncertainties for  $\alpha_{\text{LF}}$  and  $\psi$  found in Zavala et al. (2021), while the remaining parameters remain fixed. These values are summarized in Table 2.

We then use the IRLF to generate a mock IR luminous galaxy population over a  $1 \text{ deg}^2$  area. To determine the number of IR luminous galaxies in a light cone slice over  $1 \text{ deg}^2$  ( $n_{L_{\text{IR}}}$ ), we integrate the IRLF in bins of width  $\delta L_{\text{IR}} = 0.25 \text{ dex}$  down to an IR luminosity of  $10^9 L_{\odot}$ . To assign IR luminosities in a way that accurately reflects the shape of the IRLF, we then use the IRLF to generate a cumulative distribution function (CDF) for each  $L_{\text{IR}}$  bin. Then, we draw  $n_{\delta L_{\text{IR}}}$  galaxies from a random uniform distribution applied to the CDF ordinate. This method accurately reproduces the IRLF with a mock catalog. We note that using wider  $L_{\text{IR}}$  bins for catalog generation yields unphysically large populations of  $L_{\text{IR}} > 10^{10} L_{\odot}$  galaxies as the more numerous low-luminosity galaxies up-scatter into higher-luminosity bins.

We use the generated IR luminosities to calculate corresponding SFRs according to Equation (12) in Kennicutt & Evans (2012; originally derived in Murphy et al. 2011) using the variables listed in Table 1 therein for a total IR (TIR) to SFR conversion. We note that not all IR luminosity in a DSFG can be attributed to star-forming processes; some likely hails from AGN accretion activity. In the following section, we detail how we model potential AGN hosts and adjust their IR luminosities accordingly.

## 2.2. AGN Contributions to Total Infrared Luminosity

The integrated IR luminosity of galaxies encapsulates dust emission heated by both star formation activity and active black hole growth (i.e., AGNs). Young stars embedded in dusty nebulae emit primarily UV and optical light that's absorbed by the enshrouding dust. This warmed dust re-emits in the infrared

spectrum. Similarly, actively growing coeval black holes are often surrounded by obscuring clumps of dust and gas that are heated by high-energy X-ray, UV, and optical photons, and also re-emit in the infrared. Theoretical and observational studies show that the bright end of the IRLF ( $L_{\text{IR}} > 0.1\text{--}1 \times 10^{13} L_{\odot}$ , depending on the epoch) is likely dominated by AGN-powered IR emission (Fu et al. 2010; Hopkins et al. 2010; Gruppioni et al. 2013; Symeonidis & Page 2018, 2021), though AGN-driven IR emission is probably only significant at rest-frame  $\lambda = 8\text{--}30 \mu\text{m}$  wavelengths (Brown et al. 2019). Similarly, studies on samples of IR luminous galaxies demonstrate that the higher the  $L_{\text{IR}}$ , the more likely an AGN is present (Kartaltepe et al. 2010; Juneau et al. 2013; Lemaux et al. 2014)—with an occurrence rate of  $>50\%$  at  $L_{\text{IR}} \gtrsim 10^{12} L_{\odot}$  at  $0 < z < 4$ . Together, these lines of evidence indicate that we cannot assume that all IR emission from our mock galaxies hails only from star formation processes. Since we wish to use IR luminosity as an indicator for star formation activity, we must correct for AGN contamination in the total infrared spectrum ( $8\text{--}1000 \mu\text{m}$ ).

To create a realistic correction effect, we must first define what fraction of IR luminous galaxies likely host an active supermassive black hole (regardless of how IR luminous the AGN is). For this step, we use the AGN population fraction in  $70 \mu\text{m}$  bright sources as a function of IR luminosity reported in Kartaltepe et al. (2010, Figure 30 therein). This analysis included AGN selected through myriad techniques—X-ray detections, radio-loudness, IRAC colors, and mid-IR power laws. At IR luminosities  $< 10^{12} L_{\odot}$ , less than 20% of the IR luminous galaxies host significantly bright AGNs; but between  $10^{12\text{--}14} L_{\odot}$ , the occurrence rate of AGNs jumps from 60% to nearly 100%. We note that other, more recent studies on AGN-SF host galaxy fractions report smaller AGN occurrence rates as a function of IR luminosity (e.g., Lemaux et al. 2014; Symeonidis & Page 2021). Since we are using TIR luminosity as an indicator for star formation, which in turn will be used to later derive stellar masses (see Section 2.3), smaller AGN occurrence rates will result in larger stellar masses in this model. Thus, by using the larger AGN population fractions in Kartaltepe et al. (2010), we are choosing a more conservative approach that is less likely to predict an unphysical over-abundance of massive star-forming galaxies.

After selecting a subset of mock IR luminous galaxies to host AGNs, we must then decide exactly how much of the IR radiation comes from the AGN. For this step, we use the detailed analysis reporting mid-IR and TIR contributions by AGNs in ULIRGs at  $z = 0.3\text{--}2.8$  by Kirkpatrick et al. (2015). We use the distribution of mid-IR fractions (determined by Spitzer IRS mid-IR spectral decomposition, Figure 3 therein) for designated composite and AGN-dominated galaxies as a probability distribution, and draw from this probability distribution to assign mid-IR AGN fractions to AGN hosts. Then, we use the quadratic relationship between mid-IR AGN fractions and TIR AGN fractions to determine what percent of the TIR emission can be assigned to star formation (Equation (5) and Figure 14, therein). The majority ( $\sim 70\%$ ) of AGN hosts in Kirkpatrick et al. (2015) exhibit mid-IR spectral features indicative of AGN mid-IR dominance (where  $f_{\text{AGN, mid-IR}} \geq 60\%$ ); it is only at this mid-IR fraction that AGN contamination to the TIR luminosity ( $8\text{--}1000 \mu\text{m}$ ) becomes significant ( $\geq 20\%$ ).

In general, this procedure has the most significant impact on our most luminous galaxies of  $L_{\text{IR}} \geq 10^{13} L_{\odot}$ —which often correspond to the most massive galaxies—with an average AGN TIR fraction of  $\sim 50\%$ . Note that the Kirkpatrick et al. (2015) relationship between mid-IR and TIR AGN dominance maxes out at  $f_{\text{AGN, mid-IR}} = 100\%$ , corresponding to  $f_{\text{AGN, total IR}} \approx 60\%$ . Some studies argue that AGNs can be powerful enough to dominate the entire IR spectrum (e.g., Tsai et al. 2015; Symeonidis et al. 2016; Symeonidis & Page 2018) while other studies argue AGN contributions are predominantly limited to the mid-IR (e.g., Brown et al. 2019). Additionally, recent results demonstrate the existence of a population of moderate to powerful AGNs that were previously missing from these major AGN-galaxy coevolution studies as they were assumed to be weak AGNs but are instead likely heavily obscured (Lambrides et al. 2020). Together, these studies indicate that the true underlying relationship between galaxy IR luminosities and AGN processes is still unknown.

In summation: we first define a subsample of mock IR luminous galaxies to host AGNs using results presented in Kartaltepe et al. (2010); then, for that subsample, we correct their IR luminosities to remove a fraction of the IR emission that we attribute to AGN heating processes using relationships presented in Kirkpatrick et al. (2015). This suite of simple assumptions corrects only a modest fraction of IR luminous galaxies across our mock sample. The corrected IR luminosities are then used to calculate SFRs (as detailed at the end of Section 2.1). In future work, we will explore more detailed treatments and characterization of AGNs in massive dusty galaxies—and the consequences of these treatments and assumptions—in the context of this model.

### 2.3. The Star-forming Main Sequence

One of the most ubiquitous and well-studied relationships describing galaxy evolution is the correlation between a galaxy’s SFR and stellar mass. This tight relationship, known as the star-forming main sequence (SFMS), suggests that most star-forming galaxies sustain steady (i.e., non-bursty) SFRs over several billion years (Noeske et al. 2007a, 2007b; Daddi et al. 2007), at the end of which a galaxy exhausts its gas supply and moves off the main sequence toward quiescence. The correlation exists unambiguously across several orders of magnitude in stellar mass throughout cosmic time—from  $z \sim 0$  (e.g., Popesso et al. 2019b; Leroy et al. 2019) to as high as  $z \sim 4$  (e.g., Santini et al. 2017; Schreiber et al. 2017; Pearson et al. 2018; Thorne et al. 2021), theoretically making the SFMS useful in predicting stellar masses for a population galaxies with known SFRs, or vice versa.

The exact shape of the SF main sequence and its evolution through time is still the subject of intense debate. There is general agreement that the main sequence takes the form of a power law (i.e.,  $\text{SFR} \propto M_*^{\alpha}$ ) with a tight intrinsic scatter of 0.2–0.35 dex at  $M_* \lesssim 10^{10} M_{\odot}$  (e.g., Speagle et al. 2014; Schreiber et al. 2015; Matthee & Schaye 2019). In this same stellar mass regime, recent studies report that the SFMS slope,  $\alpha$ , lies between 0.8 and 1 (e.g., Speagle et al. 2014; Lee et al. 2015; Schreiber et al. 2015; Tomczak et al. 2016; Popesso et al. 2023). However, there is less consensus for galaxies with stellar masses  $>10^{10} M_{\odot}$ —which is the range of stellar mass that we are primarily interested in for this work.

Several studies describe a flattening of the main sequence above a certain “turnover” mass, typically reported at

$\sim 1\text{--}3 \times 10^{10} M_{\odot}$  at  $z = 0.5\text{--}4$ , with possible evolution toward higher values during earlier times (e.g., Whitaker et al. 2014; Lee et al. 2015; Schreiber et al. 2015; Tasca et al. 2015; Tomczak et al. 2016; Popesso et al. 2023). Above this mass, the main sequence is typically parameterized as a power law with a slope that is more shallow than the slope in the low mass regime, effectively saturating in SFRs at high stellar mass. This bending could be driven by physical phenomena such as star formation suppression at high masses or increased bulge contributions (e.g., Whitaker et al. 2015; Pan et al. 2017; Leslie et al. 2020; Thorne et al. 2021), and/or by observational effects such as contamination by passive galaxies (e.g., Johnston et al. 2015), biased/incomplete SFR indicators (e.g., Dunlop et al. 2017), or biased/incomplete SF galaxy sample selection techniques (e.g., Rodighiero et al. 2014).

For this work, we assume that the turnover at high masses is driven by observational effects, and that the true SFMS is a singular power law. The majority of studies reporting a high mass turnover selected their star-forming galaxy samples using optical/near-IR color cuts (Whitaker et al. 2014; Lee et al. 2015; Schreiber et al. 2015; Tomczak et al. 2016), such as the UVJ technique (Wuyts et al. 2007; Whitaker et al. 2011). However, color diagnostics centered in the ultraviolet to near-IR regime, while often the best available tool for such studies, are imperfect identifiers of dust-obscured star-forming galaxies. Depending on colors and sample characteristics, anywhere from 5%–50% of  $M_* > 10^{10} M_{\odot}$  dust-obscured galaxies could be misclassified as quiescent due to their dust-reddened spectra (Schreiber et al. 2015; Domínguez Sánchez et al. 2016; Popesso et al. 2019a; Patil et al. 2019; Santini et al. 2019; Hwang et al. 2021). This misclassification effectively removes some of the most dust obscured and, therefore, star-forming, massive galaxies from the analysis, whose SFRs would likely increase the high mass SFMS slope. Furthermore, the high mass slope can also be analytically flattened by the misclassification of passive galaxies as star-forming (Johnston et al. 2015).

We note that if we do include a turnover component, extremely IR luminous galaxies in this model will scatter to unphysically high stellar masses ( $>10^{13} M_{\odot}$ ). As described later in this section, we use IR-derived SFRs to pivot over the SF main sequence to assign stellar mass. A more IR luminous galaxy translates to a more star-forming galaxy, which translates (generally) to a more massive galaxy. However, with the turnover, which is typically placed around  $\sim 1\text{--}3 \times 10^{10} M_{\odot}$  at  $z = 0.5\text{--}4$ , most galaxies at and beyond ULIRG luminosities ( $L_{\text{IR}} > 10^{12} L_{\odot}$ ) have SFRs beyond the SFR that corresponds to the turnover mass. In other words, the SFRs exhibited by some of the most extreme DSFGs are well beyond the turnover SFR that corresponds to the turnover mass and therefore have no “anchor” to pivot over. Thus, in this model, employing a turnover would require forcing highly star-forming DSFGs to scatter around the asymptotic turn over and thereby generate an unphysically abundant population of incredibly massive galaxies.

Studies that sought specifically to constrain the stellar properties of DSFGs found these galaxies sitting on or above a proportionately more massive and more star-forming extension of a singular power-law SFMS (e.g., da Cunha et al. 2015; Martis et al. 2016; Michałowski et al. 2017; Pearson et al. 2018). Recently, a growing body of literature is starting to establish the existence of massive, optically dark dust-obscured

galaxy populations that were previously undetected in deep UV/optical/near-IR surveys (e.g., Spitler et al. 2014; Casey et al. 2019; Wang et al. 2019; Williams et al. 2019; Toba et al. 2020; Smail et al. 2021; Manning et al. 2022)—the same surveys used to establish a global SFMS. It is therefore likely that the true underlying SFMS—one that includes massive dust-obscured galaxies—takes the form of a power law all the way through the high stellar mass regime.

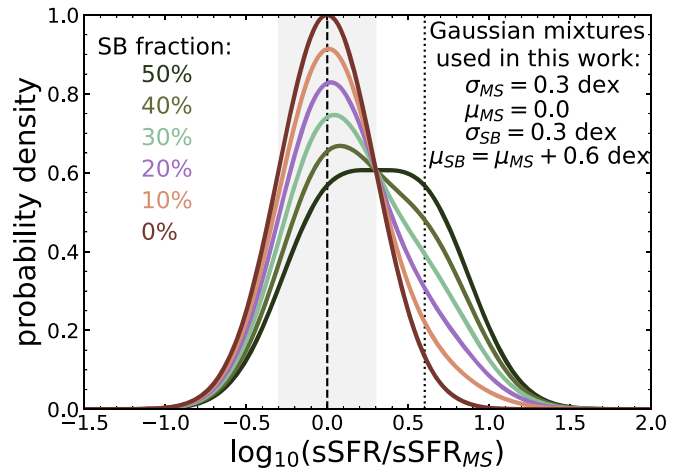
For this analysis, we randomly sample the power-law slope of the SFMS from a Gaussian distribution centered at  $\mu_\alpha = 0.9$  with a dispersion of  $\sigma_\alpha = 0.1$ . Using these slopes as opposed to more shallow ones is a more conservative approach; since we pivot SFRs over the main-sequence power law to derive stellar masses, employing shallower slopes (e.g.,  $\alpha = 0.4\text{--}0.6$ ) generates large, unphysical excesses of massive dust-obscured galaxies. Furthermore, this approach allows us to marginalize our results over the different competing main-sequence slopes established in more recent literature (e.g., Lee et al. 2015; Schreiber et al. 2015; Tomczak et al. 2016; Bisigello et al. 2018). We use the evolving normalization ( $\beta$ ) from the Speagle et al. (2014) SFMS (Equation (28), therein); this form is generally consistent across several studies using different populations of galaxies to determine SFMS properties (e.g., Johnston et al. 2015; Schreiber et al. 2015; Kurczynski et al. 2016; Tomczak et al. 2016; Pearson et al. 2018; Popesso et al. 2019a).

### 2.3.1. Main Sequence versus Starburst Decomposition

In order to derive stellar masses for our mock IR luminous galaxies, we distribute them over the star-forming main-sequence space. However, there is currently no consensus on exactly where DSFGs as a population lie in main-sequence space. Some studies suggest these objects are primarily starburst galaxies, sitting  $\sim 4\times$  above the main sequence at a given stellar mass, with such extreme SFRs driven potentially by mergers (e.g., Daddi et al. 2010; Hainline et al. 2011; Lee et al. 2013; Magnelli et al. 2014). In this work, we adopt the same starburst galaxy definition.<sup>12</sup> As a whole, the starburst population is a relatively small portion of the overall star-forming galaxy population (e.g.,  $\sim 3\%\text{--}15\%$  at  $z = 0\text{--}3$ ; Sargent et al. 2012; Schreiber et al. 2015; Bisigello et al. 2018; Popesso et al. 2019a), but DSFGs are sometimes believed to be a “starburst-dominated” subpopulation of galaxies. Starburst-dominated populations could stem from sensitivity limits by the same surveys, where only the most extreme sources at  $z > 1$  will be detected. Such a designation is likely also driven by catastrophic projection / blending effects stemming from the large angular resolutions in most canonical wide-field submillimeter surveys built by both space-based (e.g., Herschel, Béthermin et al. 2017) and ground-based telescopes (e.g., SCUBA; Cowley et al. 2015).

Recent studies with better angular resolutions (using, e.g., ALMA) reveal that significant percentages (50%–100%) of DSFGs are proportionately more massive and star-forming extensions of coeval main-sequence galaxies (da Cunha et al. 2015; Michałowski et al. 2017; Wang et al. 2019; Dudzevičiūtė et al. 2020). This fraction may evolve with redshift, as shown

<sup>12</sup> The traditional definition of starburst galaxies was based not on their distance from the SFMS, but on their rates of gas consumption (i.e.,  $M_{\text{H}_2}/\text{SFR}$ , also known as the integrated Kennicutt–Schmidt relationship or star formation efficiency; e.g., Sargent et al. 2014). For the purposes of this paper, we use the distance to the main sequence as the definition.



**Figure 2.** The main-sequence and starburst population distributions used to scatter IR luminous galaxies over stellar mass space. Recent studies using more resolved observations with, e.g., ALMA (e.g., da Cunha et al. 2015; Michałowski et al. 2017; Dudzevičiūtė et al. 2020) found that 50%–100% of DSFGs are proportionately more massive and more star-forming extensions of main-sequence galaxies. Since this exact distribution is not well known, we instead step through different Gaussian mixtures that span these observed distributions. For example, the dark green line shows a distribution where half of DSFGs are starbursts and half are main-sequence galaxies, and the dark red line shows a distribution where all DSFGs are main-sequence galaxies and none are starbursts.

in da Cunha et al. (2015) and Michałowski et al. (2017), but to generalize these fractions as evolutionary trends is problematic as they are dependent on the juxtaposed main-sequence parameterization and the FIR/submillimeter selection limits; there is also a dearth of sufficiently large samples at epochs beyond  $z = 2$  to derive a statistically relevant evolutionary picture. Thus, the true starburst fraction of DSFGs (at least the particularly massive and IR luminous ones) is still unknown, with the true value likely lying somewhere in between 0% and 50%.

Instead of forcing a singular underlying distribution as the truth, we sample the model over several steps within the observed 50%–100% range. Specifically, our model steps through DSFGs-as-starburst fractions of 0%, 10%, 20%, 30%, 40%, and 50%, with the remainder belonging to the main-sequence regime. The distribution is modeled as a double Gaussian over main-sequence space, with one Gaussian centered over the main sequence and the other centered  $+0.6$  dex away ( $4\times$  above the main sequence; Rodighiero et al. 2011). Both distributions have fixed widths of  $\sigma_{\text{MS}} = \sigma_{\text{SB}} = 0.3$  dex, similar to those measured in many main-sequence studies (e.g., Speagle et al. 2014). We tested the impact of these chosen widths by running the model with both smaller and larger dispersions ( $\sigma = 0.2$  and  $0.4$ , respectively) and found no significant difference ( $< 3\%$ ) in our results. These distributions are shown in Figure 2.

Similar to Section 2.1, we wish to randomly assign stellar masses in a way that accounts for Eddington bias: if we were to simply draw from the starburst versus main-sequence Gaussian mixtures as a way to assign stellar mass, more numerous low star-forming galaxies would up-scatter into higher mass bins than high SFR galaxies into lower mass bins. Instead, we model the main-sequence versus starburst mixtures as CDFs, from which we sample a uniform distribution of stellar masses for our mock galaxy populations. This method ensures that

**Table 3**  
Parameters for the Resulting Broken Power-law Functions That Model the Dust-obscured Stellar Mass Function

Redshift	Dusty Star-forming Galaxies				Quiescent Descendant Galaxies			
	$\log_{10}(\Phi_*)$	$\alpha$	$\log_{10}(M_*)$	$\beta$	$\log_{10}(\Phi_*)$	$\alpha$	$\log_{10}(M_*)$	$\beta$
$0.5 < z < 0.8$	$-3.77^{+0.17}_{-0.17}$	-0.5	10.8	$-1.30^{+0.41}_{-0.43}$	$-3.61^{+0.16}_{-0.17}$	-0.5	10.8	$-1.44^{+0.35}_{-0.38}$
$0.8 < z < 1.1$	$-3.73^{+0.15}_{-0.19}$	-0.5	10.8	$-1.33^{+0.43}_{-0.41}$	$-3.55^{+0.2}_{-0.19}$	-0.5	10.8	$-1.44^{+0.39}_{-0.41}$
$1.1 < z < 1.5$	$-3.70^{+0.16}_{-0.16}$	-0.5	10.8	$-1.31^{+0.39}_{-0.37}$	$-3.46^{+0.16}_{-0.17}$	-0.5	10.8	$-1.42^{+0.35}_{-0.35}$
$1.5 < z < 2.0$	$-3.75^{+0.17}_{-0.15}$	-0.5	10.8	$-1.28^{+0.36}_{-0.33}$	$-3.5^{+0.18}_{-0.19}$	-0.5	10.8	$-1.37^{+0.35}_{-0.34}$
$2.0 < z < 2.5$	$-3.98^{+0.16}_{-0.14}$	-0.5	10.8	$-1.20^{+0.33}_{-0.33}$	$-3.83^{+0.19}_{-0.29}$	-0.5	10.8	$-1.16^{+0.42}_{-0.3}$
$2.5 < z < 3.0$	$-4.28^{+0.13}_{-0.14}$	-0.5	10.8	$-1.20^{+0.31}_{-0.3}$	$-4.11^{+0.17}_{-0.17}$	-0.5	10.8	$-1.25^{+0.31}_{-0.27}$
$3.0 < z < 3.5$	$-4.63^{+0.14}_{-0.14}$	-0.5	10.8	$-1.12^{+0.32}_{-0.3}$	$-4.45^{+0.15}_{-0.15}$	-0.5	10.8	$-1.21^{+0.28}_{-0.26}$
$3.5 < z < 4.0$	$-4.95^{+0.15}_{-0.17}$	-0.5	10.8	$-1.10^{+0.25}_{-0.26}$	$-4.95^{+0.13}_{-0.13}$	-0.5	10.8	$-1.09^{+0.27}_{-0.26}$
$4.0 < z < 6.0$	$-5.56^{+0.1}_{-0.12}$	-0.5	10.8	$-1.25^{+0.37}_{-0.40}$	$-5.43^{+0.07}_{-0.07}$	-0.5	10.8	$-0.97^{+0.18}_{-0.2}$

galaxies with higher SFRs are more likely to have larger stellar masses, in alignment with observations.

#### 2.4. Stellar Mass Space Density

After using the star-forming main sequence to generate stellar masses for the mock IR luminous galaxies, we can then bin in stellar mass space to determine the three-dimensional stellar mass space density. Galaxies are binned based on their stellar mass in bins of size  $\delta_m = 0.3$  dex—corresponding to the average spread in stellar mass in the SFMS (Speagle et al. 2014). The stellar mass histograms are then divided by the same respective light cone volumes used to generate mock galaxies from the IRLF.

### 3. The Dust-obscured Stellar Mass Function

#### 3.1. Model Fitting Procedure

To briefly review the model thus far (Sections 2.1–2.4), we use the IRLF to generate a mock catalog of dust-obscured star-forming galaxies in a given redshift range over  $1 \text{ deg}^2$ . Redshift bins (listed in Table 3) were designed to match SMF studies in the literature for ease of comparison. We correct the IR luminosities for AGN contributions, then convert the corrected IR luminosities to SFRs. Then, we use the SFRs to pivot the mock galaxies over the star-forming main sequence to generate mock stellar masses. Finally, we bin the mock galaxies in stellar mass space and divide by the light cone volume to calculate stellar mass space densities across cosmic time (as a function of stellar mass).

This entire procedure is repeated  $6 \times 10^3$  times per redshift bin: 1000 realizations for each of the six different main-sequence versus starburst distributions (described in Section 2.3.1 and shown in Figure 2). Median values of the space densities within each of the six runs are calculated (across the abscissa, i.e., stellar mass bins), and uncertainties are determined from the inner 68% confidence interval. This allows us to directly gauge the impact of varying main-sequence versus starburst Gaussian mixtures on the resulting SMF.

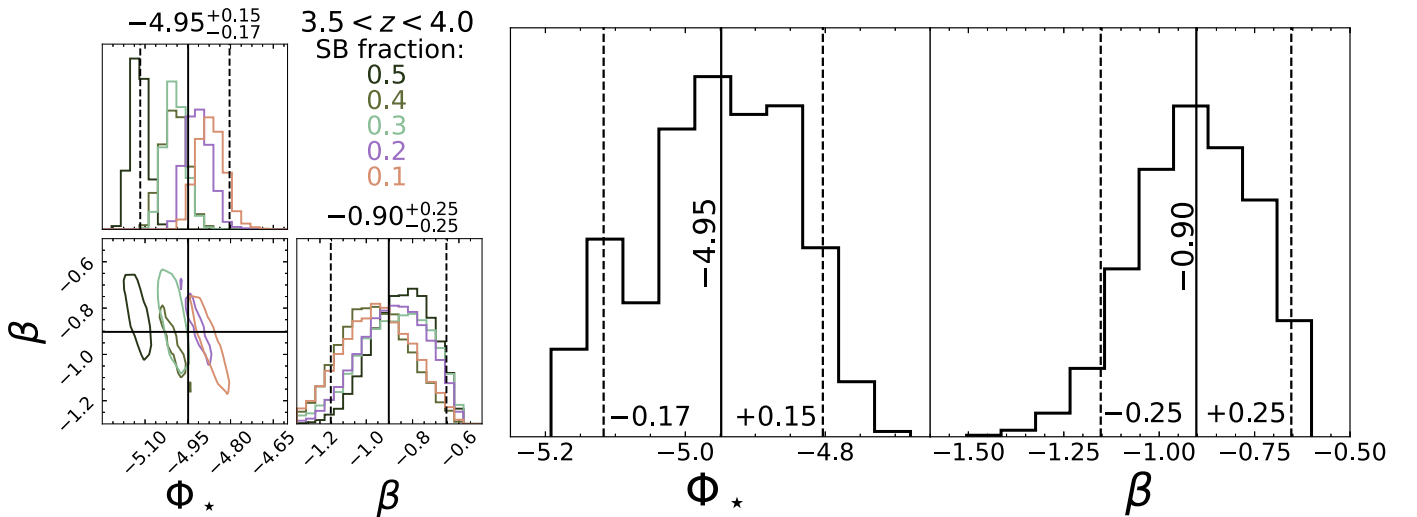
Then, within each of the six starburst versus main-sequence mixtures and for each respective redshift window, we fit a broken power law to the stellar mass space densities. Historically, SMFs are modeled as Schechter functions (Schechter 1976), but the underlying distribution that our model draws from is a broken power law (the IRLF; Casey et al. 2018b). Still, we computed the Bayesian information

criterion (BIC; Schwarz 1978) to determine whether the Schechter function or a broken power law better models our resulting SMF. The BIC is a tool used to choose between two or more models, where models are penalized for the number of parameters included in the fit. A smaller BIC corresponds to a better fit, with a difference of  $\delta \text{BIC} \geq 10$  indicating a very strong case for the best model. Across all mixtures/redshifts, we found that the broken power law resulted in a stronger model (with  $\delta \text{BIC} \geq 7$ –10 in nearly all cases). Thus, we adopt the broken power-law form to define the number density of dust-obscured galaxies,  $\Phi(M)$ , as:

$$\Phi(M, z) = \begin{cases} \Phi_*(z) \left( \frac{M}{M_*(z)} \right)^{\alpha(z)} & \text{if } M < M_*(z) \\ \Phi_*(z) \left( \frac{M}{M_*(z)} \right)^{\beta(z)} & \text{if } M \geq M_*(z) \end{cases} \quad (2)$$

where  $\Phi_*$  (in  $\text{Mpc}^{-3} \text{ dex}^{-1}$ ) is the number density normalization,  $M = M_{\text{star}}/M_{\odot}$ ,  $\alpha$  is the low mass slope,  $M_*$  is the characteristic mass or “knee” of the SMF, and  $\beta$  is the high mass slope (beyond the knee). We visually inspect the stellar mass space densities in each mixture-redshift combination and find that the low mass slope ( $\alpha$ ) and the characteristic mass ( $M_*$ ) do not significantly evolve across mixtures or redshifts ( $< 5\%$  difference between all mixture-redshift combinations). Therefore, to reduce the degrees of freedom, we fix  $\alpha = -0.5$  and  $M_* = 10^{10.8} M_{\odot}$  universally across all epochs/mixtures. Upon visual inspection, both of these values combined well-reproduced the low mass regime of the DSFG SMF for all mixture-redshift combinations. For posterity, we also tested a smaller  $M_* = 10^{10.5} M_{\odot}$  and a steeper  $\alpha = -0.7$  (more in line with UV/OP-based studies). This resulted in a  $< 3\%$  difference in the derivations of  $\Phi_*$  and  $\beta$ , the overall  $\chi^2$  values were poorer, indicating a preference for a more massive knee and shallower low mass slope.

Then, still within the individual mixture-redshift combinations, we run  $10^4$  realizations of a Metropolis-Hastings (i.e., Markov Chain Monte Carlo, MCMC) algorithm to find the best-fit values for the number density normalization,  $\Phi_*$ , and massive-end slope,  $\beta$ . We allow  $\log_{10}(\Phi_*)$  to range from  $[-6, -3]$  and  $\beta$  from  $[-2.5, -0.6]$ . These parameter ranges cover sufficiently extreme scenarios of either incredibly populous massive, dusty galaxies, or a universe that has a severe dearth of massive, dusty galaxies. The quality of each broken power-law fit is determined by computing the chi-square goodness of



**Figure 3.** Left: corner (correlation) plots for the SMF parameters ( $\Phi_*$ ,  $\beta$ ) derived in the  $3.5 < z < 4.0$  redshift bin. Each color represents the best-fit parameters for the six distinct main-sequence vs. starburst mixtures, while the solid line represents the median value, and the dashed lines represent the 68% confidence interval (matching the values listed at the top of each histogram). There is some migration of  $\Phi_*$  toward smaller values with increasing starburst fraction, though the differences for both parameters across mixtures are small ( $< 5\%$ ) and encapsulated within the  $1\sigma$  uncertainties of the final combined value (shown in the histograms on the right). Right: combined distributions for  $\Phi_*$  and  $\beta$  in the  $3.5 < z < 4$  redshift bin. To create these combined distributions, we sampled  $10^3$  successful pairs of  $\Phi_*$  and  $\beta$  from each mixture (for a total of  $6 \times 10^3$  samples, 1000 per mixture), then adopted the median value (shown as the black solid lines) as the true value for the redshift bin. The black dashed lines show the  $1\sigma$  dispersion plus an additional 10% uncertainty added due to uncertainty in the underlying IRLF; these values are also listed at the top of each parameter column in the corner plot, and at the bottom of the histograms to the right and left of the median value. The ultimate values used in this work are reported in Table 3.

fit against the mock stellar mass space densities; we only compute the  $\chi^2$  using stellar mass space densities with 68% confidence intervals that are less than an order of magnitude away from the median value (i.e.,  $\log_{10}(\sigma) < 1$ ). This primarily affects the highest mass bins ( $M > 10^{12} M_\odot$ ) where mock IR luminous galaxies appear rarely and therefore are not produced consistently in the model.

In Figure 3, we show the resulting constraints on  $\Phi_*$  and  $\beta$  for an example redshift bin. There is some marginal covariance with starburst fraction and  $\Phi_*$ , such that increasing starburst fractions causes a smaller overall number density of massive IR luminous galaxies (as expected). However, the differences between  $\Phi_*$  across mixtures is small ( $\lesssim 0.25$  dex at most) and, moreover, the median  $\Phi_*$  in each mixture is captured within the final uncertainty derived for this redshift bin. Thus, this covariance is insignificant for the purposes of this work. Additionally, there is no clear trend between the high mass slope  $\beta$  and starburst fractions across all epochs. One might expect a shallower  $\beta$  to arise from a smaller starburst fraction because with low starburst fractions, more galaxies are deemed main-sequence-like and thus scatter to higher stellar masses; this effect in theory creates more numerous massive galaxies than model realizations with “high” starburst fractions, and causes the population density near/beyond  $M_*$  to increase. However, in this case, it is likely that the most extreme IR luminous/dust-obscured star-forming galaxies are so rare that a consistently small population emerges at masses  $> M_*$ , regardless of the employed main-sequence versus starburst mixture.

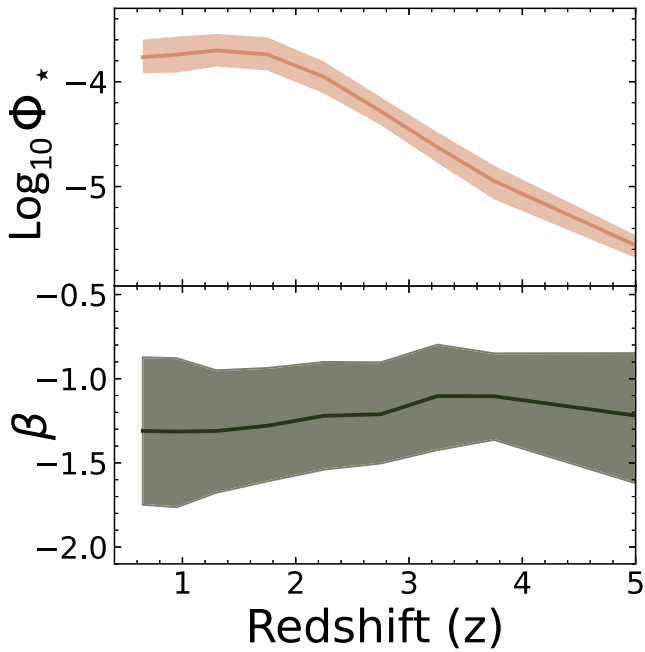
To derive a singular SMF from the six different main-sequence versus starburst mixtures, we randomly draw  $10^3$  pairs of  $(\Phi_*, \beta)$  from the converged MCMCs of each mixture. We collate the  $6 \times 10^3$  pairs of  $(\Phi_*, \beta)$  and determine the median and 68% confidence interval. This process is completed separately for each redshift window. We show example

collated histograms for  $\Phi_*$  and  $\beta$  in the  $3.5 < z < 4$  redshift bin in Figure 3.

It is important to note that the main basis of this model hinges upon the accuracy of the modeled IRLF. As reviewed in Section 2.1, the evolution of the IRLF beyond  $z \sim 2$  is still under debate (e.g., Gruppioni et al. 2020). While we marginalize over several parameters to account for these  $z > 2$  unknowns, there are still some fixed assumptions in this work, such as the evolution of  $L_*$  at  $z > z_{\text{turn}}$ , or the lack thereof for the faint-end slope,  $\alpha_{\text{LF}}$ . We find that a steeper  $\alpha_{\text{LF}} = -0.7$  results in a less-than-3% difference in the final derivations of the parameters describing the dust-obscured SMF; a flatter evolution of  $L_*$  with  $\psi = 0.5$  results in a less-than-1% difference in the resulting SMF parameters. Thus, to be conservative and account for the impacts of our model’s main assumptions, we add an additional 10% uncertainty on the ultimate SMF parameters. This additional uncertainty is included in Figure 3 and in Table 3, where the ultimate SMF parameters for this analysis are listed.

### 3.2. Comparison to Literature SMFs

In Figure 4, we show the redshift evolution of  $\Phi_*$  and  $\beta$  for our final dust-obscured SMF across time. The number density normalization stays relatively constant at  $\log_{10}(\Phi_*/(\text{Mpc}^{-3} \text{dex}^{-1})) \approx -3.7$  in the 5 Gyr between  $z = 2$  and  $z = 0.5$ , but experiences nearly an order of magnitude increase in the 1.5 Gyr between  $z = 3.5$  and  $z = 2$  and a similar strength in growth in the 1 Gyr between  $z = 6$  and  $z = 3$ . This points to an era of rapid stellar mass growth in the  $z > 3$  universe for massive dust-obscured galaxies that is unmatched at later times. Interestingly, however, we do not see any significant evolution in the massive-end slope of the SMF,  $\beta$ , which has an average value of  $-1.23 \pm 0.08$  across all epochs. Combined, the evolutionary properties for these two parameters paint a picture in line with the cosmic downsizing scenario, where the most



**Figure 4.** The redshift evolution of  $\Phi_*$  and  $\beta$ —the broken power-law parameters used to model the DSFG SMFs in this work, also listed in Table 3. The line marks the median value found when combining  $10^3$  pairs of  $(\Phi_*, \beta)$  from the converged MCMCs across all starburst vs. main-sequence mixtures, and the shaded area corresponds to the 68% confidence interval plus an additional 10% uncertainty. In the top panel, we show the evolution of  $\Phi_*$ , the normalization of the space density, decreases at higher- $z$ , which is expected within the context of galaxy growth and assembly. In the bottom panel, we show the evolution of the high mass slope,  $\beta$ , does not evolve significantly with redshift, indicating that massive galaxies grow and then quench out of the massive end of the star-forming SMF over timescales shorter than 1 Gyr. See Section 3.2 for more discussion on these resulting parameters.

massive star-forming galaxies built the majority of their stellar mass in the first 1–2 Gyr, and quench by  $z \sim 2$ . Furthermore, the rate of growth and quenching in dust-obscured galaxies must occur on short timescales (at least shorter than the  $\sim 1$  Gyr-wide redshift windows); in other words, if  $\beta$  remains relatively stable across time, then dust-obscured galaxies likely grow into and then quench out of the massive end of the star-forming SMF at a nearly equivalent rate.

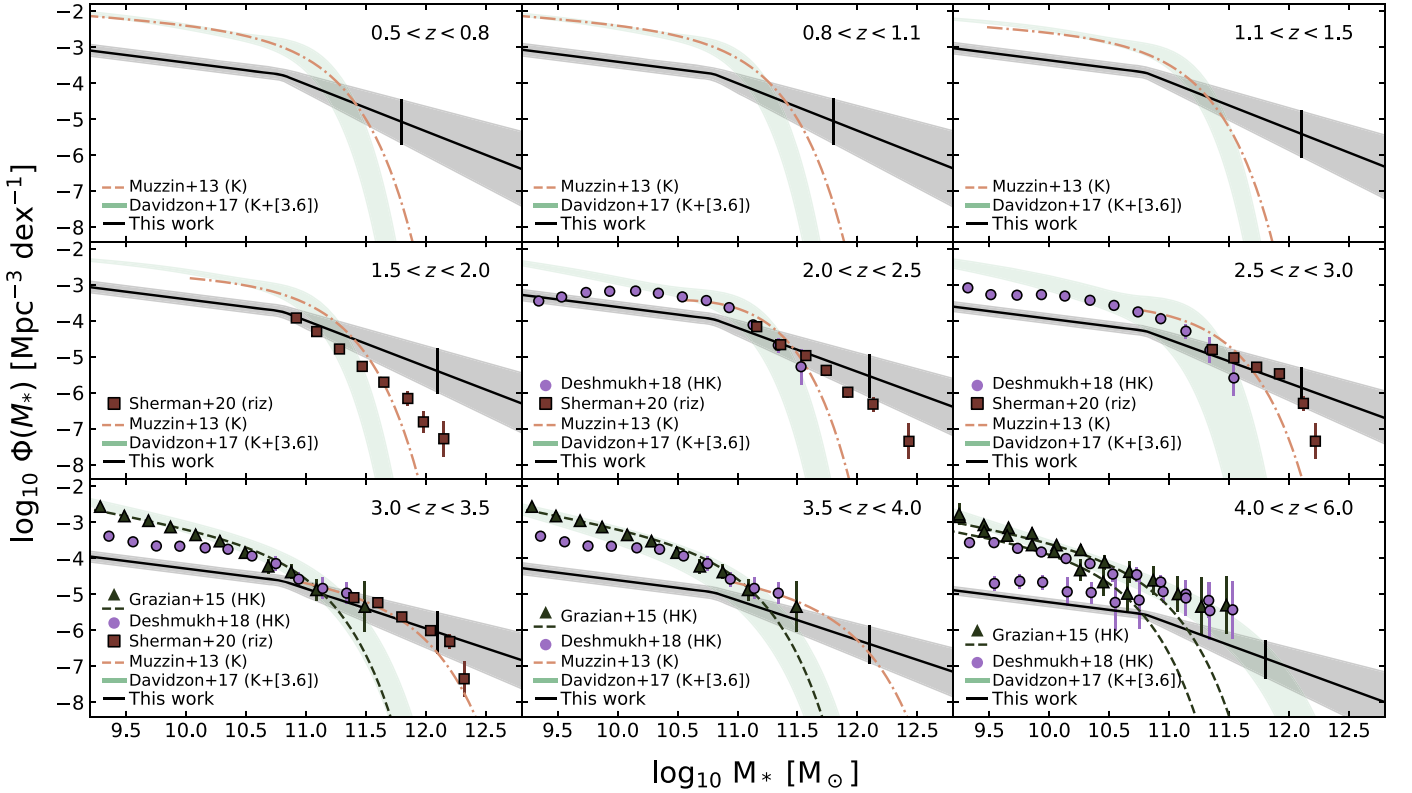
In Figure 5, we show the final star-forming SMFs derived from IR luminous galaxies across  $0.5 < z < 6$ . We remind the reader that the IRLF does not account for all star-forming galaxies, and therefore agreement between rest-frame UV/OP-based SMFs and the dust-obscured SMF generated in this work is not necessarily expected. Moreover, since our SMF takes a different functional form than those presented in the literature, we cannot quantitatively speak to differences in shape and parameterization. We can, however, discuss qualitative and integrated differences (the latter of which is discussed in detail in Sections 4.2 and beyond). Furthermore, due to the scientific purpose of this work, we focus the majority of this discussion on the more massive end of galaxy evolution at  $M > 3 \times 10^{10} M_\odot$ , but show the less massive end of this and literature SMFs so that the reader can understand the overall extrapolation of this model’s results to low masses. Where necessary, all literature SMFs are scaled to a Chabrier (2003) IMF.

In general, the dust-obscured SMF produced by this model is consistently below that of UV/OP-based predictions for star-forming galaxies at masses  $\lesssim 10^{11} M_\odot$ . It is around this mass that there appears to be a convergence between UV/OP and

dusty SMFs, which then turns into a dust-obscured dominated SMF at high masses. As mentioned before, this is likely because low mass galaxies are not accounted for in the IRLF, as they are less obscured by dust. The discrepancies between the observed SMFs and our model are greatest at  $z = 1-4$ , where our model posits a significant population of  $M > 10^{11.5-12} M_\odot$  galaxies, while the UV/OP-based estimates see none (see Figure 6). At  $z = 2-3.5$ , our model shows some agreement with that of the *riz*-based SMF derived in Sherman et al. (2020). We believe this agreement to be artificial as it could be argued that there are high rates of AGN contamination in the Sherman et al. (2020) sample. This is because AGNs are removed from this sample using SDSS data, which likely does not sufficiently capture highredshift and/or obscured AGNs. At  $z \geq 3$ , we also see some partial agreement between our model, the Muzzin et al. (2013) results, and the massive, dusty galaxy sample within Deshmukh et al. (2018) out to  $M \sim 10^{11.5-12} M_\odot$ . Both of these works select their galaxy samples using *H*- and/or *K*-band images with observational depths near the limits of known massive, dust-obscured galaxies (e.g.,  $m_{K_s} \approx 23-24.5$ ; Simpson et al. 2014; da Cunha et al. 2015; Cowie et al. 2018; Dudzevičiūtė et al. 2020), so it is likely some of the massive dust-obscured population is captured in these works. However, at these *H*- and *K*-band depths, only 50%–75% of DSFGs are typically detected, leaving a significant population of optically faint galaxies (also known as “OIR-dark”; Casey et al. 2019; Wang et al. 2019; Williams et al. 2019; Manning et al. 2022)—that are likely some of the most massive and heavily dust-obscured galaxies in the cosmos (Pannella et al. 2009; Dunlop et al. 2017; Whitaker et al. 2017; Fudamoto et al. 2020)—unaccounted for in the massive end of these SMFs.

It is also interesting that our results are in more severe disagreement with the Davidzon et al. (2017) analysis, as the authors introduce a deep Spitzer IRAC [ $3.6 \mu\text{m}$ ] image to help capture more dust-obscured and quiescent galaxies than the typical rest-frame UV/OP-selected SMFs. Yet, the Davidzon et al. (2017) SMF yields fewer star-forming galaxies beyond  $M > 10^{11} M_\odot$  than any other SMF discussed in this section. This could be due to a variety of factors. For example, the Davidzon et al. (2017) work required detection in a combined *zYJHK<sub>s</sub>* image; the  $3\sigma$  detection limit of the reddest band ( $K_s = 24$ ) could miss 40% or more of dusty galaxies due to their faintness (Simpson et al. 2014; Brisbin et al. 2017; Dudzevičiūtė et al. 2020). Another factor could be the fact that galaxy stellar masses were derived solely from rest-frame UV to optical wavelengths—no FIR (i.e., cold dust) information was included in SED fitting. If there is heavy obscuration of the stellar spectrum, and no energy coupling between optical/mid-IR and FIR observations, stellar masses can be underestimated by a factor of 2–3 due to faintness (Buat et al. 2019), and/or a mildly quiescent galaxy spectrum could be assumed (especially with limited photometry). Both of these employed methods can remove massive, dust-obscured galaxies from the massive end of the star-forming SMF.

At  $4 < z < 6$ , the dust-obscured SMF is below UV/OP-based estimates at similar redshifts, except where they meet at the most massive point around  $M \sim 10^{12} M_\odot$ . While there is still large uncertainty in the prevalence of dust-obscured galaxies at  $z > 4$  (see, e.g., Zavala et al. 2021; Gruppioni et al. 2020), it appears from this model and from UV/OP-based estimates that such massive galaxies are exceedingly rare at these epochs, and their population density really begins to build after this epoch.



**Figure 5.** The resulting SMF of the IR luminous galaxy population as derived from this model. Our results are shown as the black line, with  $1\sigma$  uncertainty shaded in gray. The vertical black line marks the limit where the uncertainty of the stellar mass space densities is  $<1$  order of magnitude; beyond this mass, massive IR luminous galaxies were not consistently produced in the model realizations. Shown for comparison are star-forming SMFs from the literature, with a particular focus on those using the deepest available observational data: Muzzin et al. (2013, pink dashed–dotted line), Davidzon et al. (2017, green shaded curve), Sherman et al. (2020, red squares), Deshmukh et al. (2018, purple circles), and Grazian et al. (2015, black triangles and black dashed line). For Grazian et al. (2015) and Deshmukh et al. (2018) at  $4 < z < 6$ , we include two SMFs: one for each of their  $z = 4\text{--}5$  and  $z = 5\text{--}6$  bins. In this work, UV/OP-bright galaxies tend to dominate the star-forming population at lower stellar masses ( $M \lesssim 10^{10.5-11} M_{\odot}$ ) by  $\sim$  an order of magnitude, but become comparatively or less abundant at high stellar masses, particularly at  $1 < z < 4$ . See Section 3.2 for more details.

In Figure 6, we show the fraction of massive, dust-obscured galaxies among all star-forming galaxies using the Davidzon et al. (2017) SMF as a proxy for UV/OP-bright star-forming galaxies (i.e., the denominator is derived from Davidzon et al. 2017, plus the estimate from this work). We show that from  $0.5 < z < 4$ , dust-obscured galaxies become more prominent at higher stellar masses. Though there is large uncertainty, we predict that 20%–40% of galaxies at  $M > 10^{10.5} M_{\odot}$  are dust obscured, which is similar to the fraction of  $K$ -band undetected SMGs in Dudzevičiūtė et al. (2020), the  $H$ -band dropout sample in Wang et al. (2019), and the dust-reddened sample in Marsan et al. (2022). The percentage increases to 50%–60% at  $M > 10^{11} M_{\odot}$ , and finally maxes out to nearly 100% at  $M > 10^{11.5} M_{\odot}$ . Interestingly, these results are in general agreement with the Whitaker et al. (2017) study, which arrived at a dust-obscured fraction of massive galaxies using a completely different method—quantifying the fraction of dust-obscured versus UV/OP-bright star formation in individual massive star-forming galaxies at  $0 < z < 2$ . However, in Whitaker et al. (2017),  $>90\%$  of star formation in  $M > 10^{10.5} M_{\odot}$  galaxies at  $z = 0.5\text{--}2$  is obscured by dust (where we predict roughly 40% of massive SF galaxies at this epoch are dust obscured). We note that this estimate is significantly higher than our derived massive, dusty galaxy fraction at similar masses, but more in agreement at the most extreme end of  $M > 10^{11.5} M_{\odot}$ . At  $z > 4$ , we caution interpretation of the

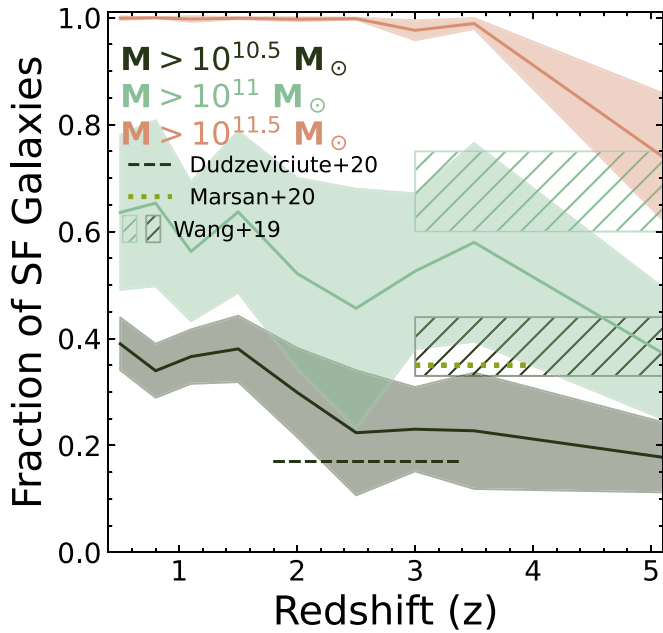
$M > 10^{11.5} M_{\odot}$  DSFG fractions, as galaxies of this mass were so rare that they were not consistently produced in this model.

We note that there may be some massive DSFGs with sufficient escaping stellar emission to be captured in the UV/OP-based estimates, and therefore there may be some “double counting” of massive SF galaxies in this space. Understanding the significance of that effect is beyond the scope of this paper, and likely unfeasible until sufficient wide-field JWST near-IR surveys are completed.

#### 4. Number and Stellar Mass Densities

For the remainder of this work, we focus on the massive, dust-obscured population of galaxies predicted by our model, defined as all objects with  $M \geq 10^{10.5} M_{\odot}$  and  $L_{\text{IR}} \geq 10^{11} L_{\odot}$ . We emphasize that the latter criterion emerged naturally in this model: nearly all ( $\geq 99\%$ )  $M \geq 10^{10.5} M_{\odot}$  star-forming galaxies produced by our model, across the redshift space probed here, have LIRGs or greater IR luminosities; moreover,  $>75\%$  of these massive star-forming galaxies have ULIRG ( $L_{\text{IR}} \geq 10^{12} L_{\odot}$ ) luminosities or higher. We hereafter refer to these IR luminous massive galaxies as massive DSFGs.

For both massive DSFGs and their quiescent descendants (the latter of which is described further in Section 4.3), we calculate the number and stellar mass densities over the same redshift slices/light cones used to generate the dust-obscured SMF (also presented in Tables 3 and 4). For the massive



**Figure 6.** Fraction of massive star-forming galaxies that are dust-obscured, defined by LIRG luminosities or greater. The denominator is calculated by adding number counts from the Davidzon et al. (2017) SMF to the estimates from our dust-obscured SMF. Fractional uncertainties from both SMFs are added in quadrature, with  $1\sigma$  uncertainties shown as the lighter contours. The dark green, light green, and pink bands correspond to lower stellar mass limits of  $\log_{10}(M) > 10.5$ , 11, and  $11.5 M_{\odot}$ , respectively. Overlaid are comparable results of massive dust-obscured galaxy fractions from Wang et al. (2019), Dudzeviciute et al. (2020), and Marsan et al. (2022). Consistent with the literature, we find that the fraction of dust-obscured SF galaxies increases with increasing stellar mass, reaching totality ( $\sim 100\%$ ) at the highest and rarest limit of  $M \sim 10^{11.5} M_{\odot}$ .

DSFGs, we convert the SMF to a number density by integrating at  $M \geq 10^{10.5} M_{\odot}$ , and uncertainties are derived by sampling the SMF parameter uncertainty (presented in Table 3). For the quiescent descendants, since their generation is more complex and directly dependent on their progenitor galaxy, we instead calculate the number and stellar mass densities directly from the mock catalogs (as opposed to an integrated SMF). These quiescent descendant density measurements are calculated directly in each of the  $10^3$  realizations per light cone, then the median value and inner 68% confidence interval is reported as the uncertainty.

#### 4.1. Duty Cycle Correction

Gas depletion timescales measured in high- $z$  DSFGs suggest that the IR luminous phase of star formation is short lived (on the order of tens to a few hundred megayears; Tacconi et al. 2006; Barro et al. 2014; Aravena et al. 2016; Yang et al. 2017; D’Eugenio et al. 2020; Zavala et al. 2022). This means that massive DSFGs are only traceable via their incredible IR emission for a small window of time, and therefore IR/submillimeter surveys are likely incomplete in capturing the DSFG population over cosmic timescales significantly wider than a few hundred megayears. If we assume that massive DSFGs are the primary progenitor population to massive quiescent galaxies at later times, we can derive a “duty cycle correction factor” that forces the two population densities to match over respective redshift windows (as seen in, e.g., Toft et al. 2014; Valentino et al. 2020).

**Table 4**

Number Densities and Stellar Mass Densities as Shown in Figures 7 and 9 for Galaxies of Stellar Mass  $M > 10^{10.5} M_{\odot}$

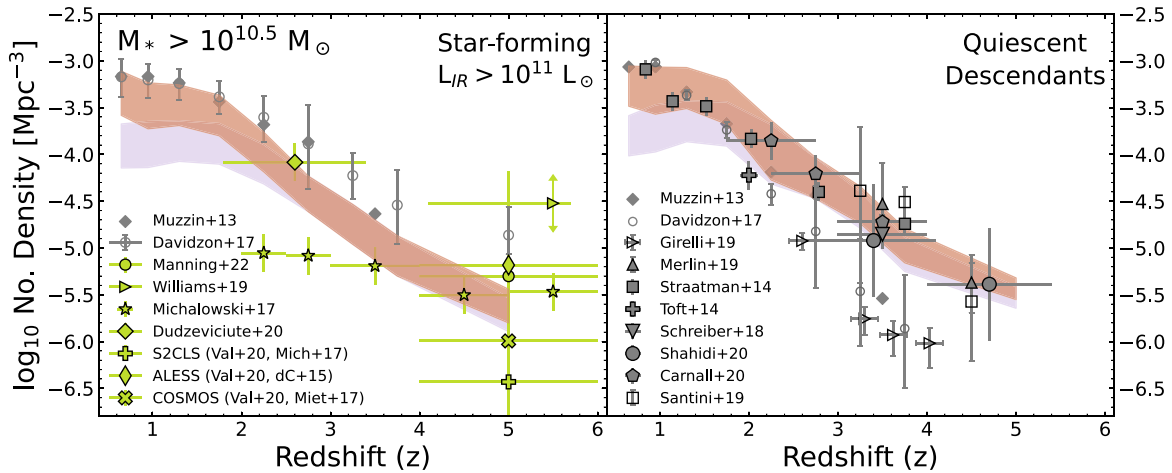
Redshift	Massive DSFGs		Quiescent Descendants	
	$\log_{10}(n)$	$\log_{10}(\rho)$	$\log_{10}(n)$	$\log_{10}(\rho)$
$0.5 < z < 0.8$	$-3.37^{+0.26}_{-0.21}$	$7.23^{+0.69}_{-0.42}$	$-3.23^{+0.22}_{-0.21}$	$7.28^{+0.48}_{-0.36}$
$0.8 < z < 1.1$	$-3.49^{+0.26}_{-0.23}$	$7.25^{+0.69}_{-0.44}$	$-3.32^{+0.26}_{-0.24}$	$7.35^{+0.58}_{-0.39}$
$1.1 < z < 1.5$	$-3.48^{+0.23}_{-0.20}$	$7.30^{+0.57}_{-0.39}$	$-3.26^{+0.22}_{-0.21}$	$7.45^{+0.49}_{-0.36}$
$1.5 < z < 2.0$	$-3.60^{+0.23}_{-0.20}$	$7.29^{+0.56}_{-0.38}$	$-3.37^{+0.24}_{-0.23}$	$7.44^{+0.54}_{-0.38}$
$2.0 < z < 2.5$	$-3.97^{+0.22}_{-0.21}$	$7.13^{+0.54}_{-0.41}$	$-3.84^{+0.30}_{-0.34}$	$7.31^{+0.78}_{-0.55}$
$2.5 < z < 3.0$	$-4.42^{+0.20}_{-0.18}$	$6.81^{+0.52}_{-0.37}$	$-4.26^{+0.23}_{-0.21}$	$6.95^{+0.53}_{-0.38}$
$3.0 < z < 3.5$	$-4.75^{+0.22}_{-0.20}$	$6.58^{+0.57}_{-0.45}$	$-4.60^{+0.21}_{-0.18}$	$6.64^{+0.49}_{-0.36}$
$3.5 < z < 4.0$	$-5.07^{+0.21}_{-0.22}$	$6.26^{+0.49}_{-0.43}$	$-5.07^{+0.19}_{-0.17}$	$6.27^{+0.51}_{-0.38}$
$4.0 < z < 6.0$	$-5.62^{+0.18}_{-0.18}$	$5.53^{+0.57}_{-0.41}$	$-5.43^{+0.12}_{-0.11}$	$5.95^{+0.34}_{-0.31}$

We iteratively tested a range of duty cycles from 50–1000 Myr and found that a duty cycle correction factor of  $\sim 500$  Myr was most effective in matching our predicted number densities to those of massive quiescent galaxies observed in the literature. Duty cycles on the order of a few hundred megayears or less result in a significant over production ( $10\times$  more) of observed quiescent descendants across all epochs, while cycles closer to 1 Gyr underproduce the descendant population by 0.5 dex at  $z < 2$ . We note that a duty cycle correction of  $\sim 500$  Myr only significantly impacts the model results at  $z < 2$ ; all results at  $z > 2$  remain roughly the same in both the pre- and post-duty-cycle-corrected model, as the widths of the redshift windows are approximately equivalent to the duty cycle correction factor. We show the pre- and post-duty-cycle-correction model predictions in Figure 7 as purple and red bands, respectively, and discuss further potential implications of this value in Section 5.2.

#### 4.2. Massive DSFG Number Densities

In Figure 7, we show the redshift evolution of the number density of massive DSFGs (left), and of their quiescent descendants (right; we discuss in the proceeding section how quiescent descendants are generated.) During the  $\sim 1$  Gyr between  $z=6$  and 3, we predict an order of magnitude increase in the number density of  $M > 10^{10.5} M_{\odot}$  dust-obscured SF galaxies. This growth rate tapers off at later times, taking another  $\sim 2$  Gyr (from  $z=3$  to 1.5) to achieve another order of magnitude of growth in population density. This  $z > 3$  evolution is steeper than that exhibited by UV/OP-based estimates from, e.g., Davidzon et al. (2017), where it takes an additional  $\sim 0.5$  Gyr to achieve similar growth. In general, across  $0.5 < z < 4$ , UV/OP-bright massive SF galaxies are  $\sim 6\text{--}7\times$  more populous, except at  $z=4\text{--}6$ , where we find an order of magnitude more UV/OP-bright galaxies than massive DSFGs. (However, as shown in Section 3.2 and Figure 6, the overall fraction of massive DSFGs compared to UV/OP-bright galaxies is larger at higher stellar mass thresholds.)

One of the primary pieces of evidence that pins  $z > 3$  DSFGs as ancestors to giant ellipticals at later times is the similar comoving number densities between the two populations (e.g., Toft et al. 2014). However, the strength of this evolutionary picture is discrepant depending on the sample size, depths, and sensitivities of FIR and submillimeter surveys used to characterize the massive, dust-obscured population. For example, Herschel discovered DSFGs at  $z > 2$  often comprise



**Figure 7.** The evolution of the number density of massive galaxies. Our model is presented as the red band, with the non-duty-cycle-corrected values represented by the purple band. Left: observations of similar (but not necessarily equivalent; see Section 4.2) DSGF populations are shown as green points (Michałowski et al. 2017; Williams et al. 2019; Dudzevičiūtė et al. 2020; Valentino et al. 2020; Manning et al. 2022), while UV/OP-based estimates are shown in gray (Muzzin et al. 2013; Davidzon et al. 2017). It is important to note that, for the observational data in green, the uncertainties on the abscissa demarcate the redshift bins of each observational sample. For example, when integrating this model over a similar redshift bin as the Dudzevičiūtė et al. (2020) sample, we achieve a similar number density estimate (demonstrated by the overlap between the  $z \sim 2$  portion of our model and the low redshift end of the Dudzevičiūtė et al. 2020 sample). Right: massive, passive galaxy number density constraints from the literature are shown as gray points (Muzzin et al. 2013; Straatman et al. 2014; Toft et al. 2014; Davidzon et al. 2017; Schreiber et al. 2018a; Girelli et al. 2019; Merlin et al. 2019; Santini et al. 2019; Carnall et al. 2020; Shahidi et al. 2020). Note that the selection functions for the comparing quiescent galaxy samples vary significantly across these studies. Our results appear to align best with massive post-starburst (sSFR  $< 10^{-10} \text{ yr}^{-1}$ ) galaxies. See Section 4.3 for more discussion.

a rare and extreme star-forming population (with SFRs up to and exceeding  $1000 M_{\odot} \text{ yr}^{-1}$ ); their number densities are insufficient to match those estimated for massive quiescent galaxies at  $z > 2$  (Dowell et al. 2014; Ivison et al. 2016; Duivenvoorden et al. 2018; Ma et al. 2019; Montaña et al. 2021). Instead, the recently discovered  $H$ -band dropouts, or alternatively the larger umbrella of galaxies that are optically “invisible” but sufficiently bright in the FIR/submillimeter regime ( $S_{850\mu\text{m}} \gtrsim 1 \text{ mJy}$ ), may make up the bulk of the massive, star-forming galaxy population at  $z > 3$  (Wang et al. 2019; Williams et al. 2019; Lagos et al. 2020; Manning et al. 2022). Thus, it is perhaps more important to successfully model these more common, less-extreme DSGFs.

At  $z = 2\text{--}3$ , when integrated over similar redshift ranges, our model successfully predicts similar space densities of the massive ( $M_{*} \gtrsim 3 \times 10^{10} M_{\odot}$ ) submillimeter galaxy populations presented in Dudzevičiūtė et al. (2023;  $\sim 10^{-4} \text{ Mpc}^3$  at  $z = 1.8\text{--}3.4$ ), but predicts  $5\text{--}10 \times$  more massive DSGFs than what is seen in the Michałowski et al. (2017) SCUBA-2-detected sample in a similar redshift range. This difference is likely due to the employed SFR cutoff in the latter work, which only includes galaxies with SFRs  $> 300 M_{\odot} \text{ yr}^{-1}$ . In our model, 20%–40% of  $M \gtrsim 3 \times 10^{10} M_{\odot}$  galaxies at  $z \sim 2$  have SFRs  $< 300 M_{\odot} \text{ yr}^{-1}$  (with the percent decreasing with increasing redshift)—this difference in population properties is sufficient to match the gap between our  $z \sim 2$  results and those seen in Michałowski et al. (2017), within uncertainty.

At  $z > 3$ , the model’s success is difficult to quantify due to lack of strong observational constraints, including insufficient sample sizes of DSGFs and highly uncertain redshifts. Still, we find general agreement between the model and observations out to  $z \sim 5$ , though observational uncertainties span 2 orders of magnitude. At  $z = 3\text{--}4$ , the model predicts massive DSGF number densities in agreement (within uncertainty) with those of the Michałowski et al. (2017) sample ( $n \sim 1\text{--}2 \times 10^{-5} \text{ Mpc}^3$ ). At  $4 < z < 6$ , we compare to the ALESS (da Cunha et al. 2015),

COSMOS (Miettinen et al. 2017), and S2CLS (Michałowski et al. 2017) samples as calculated in Valentino et al. (2020; pre-duty cycle correction). To derive number densities for these surveys, Valentino et al. (2020) took strict photometric redshift constraints, requiring that both upper and lower limits on the photoz sit at or above  $z = 4$ . As seen with the S2CLS derivations between Michałowski et al. (2017) and Valentino et al. (2020), this requirement can reduce the number densities by an entire order of magnitude, thereby demonstrating to first order one of the main difficulties in constraining DSGF number densities at  $z > 3$ . Across these three surveys with conservative photoz restrictions, our model estimates lie between those derived from the COSMOS Miettinen et al. (2017) and ALESS da Cunha et al. (2015) samples, where  $n \sim 2 \times 10^{-6} \text{ Mpc}^3$  at  $4 < z < 6$ , which is also roughly in alignment with Michałowski et al. (2017).

Comparing to the observed number densities of the (potentially more populous) optically dark dust-obscured galaxies at  $4 < z < 6$  is particularly difficult as observational constraints suffer either from small sample sizes ( $n = 1\text{--}4$ ; e.g., Williams et al. 2019; Manning et al. 2022), or are averaged over extremely large redshift ranges (e.g.,  $z = 3\text{--}6$  in Wang et al. 2019). When comparing to estimates derived from a single exceptional object in Williams et al. (2019), our model predicts nearly  $30 \times$  smaller populations of massive DSGFs. While the Wang et al. (2019) sample of  $H$ -band dropouts includes a larger sample size ( $n = 39$ ) over a larger survey area ( $\sim 600 \text{ arcmin}^2$ ), these objects span a (photometric redshift) range of  $3 < z < 8$ , with over half of the sample exhibiting redshift errors  $\delta z \geq 1$ . Still, when integrating our model across  $3 < z < 5$ , we find a similar space density as the authors (reported as  $n \sim 2 \times 10^{-5} \text{ Mpc}^3$ ) as well as the entire sample of 2 mm detected galaxies in Manning et al. (2022). When including only galaxies with redshifts estimated at  $4 < z < 6$ , both the Wang et al. (2019) and Manning et al. (2022) estimates drop to  $n \sim 5\text{--}6 \times 10^{-6} \text{ Mpc}^3$ , which is closer to our

model predictions. Unfortunately, due to the large uncertainties in the observational data, we limit and conclude our interpretation of the discrepancies between this work and the aforementioned optically dim sources until further, more robust surveys are completed.

### 4.3. Quiescent Descendants

#### 4.3.1. Generating Quiescent Descendants

The discovery and follow-up observations of massive galaxies within the first 1–2 Gyr of the universe suggest that they form via extremely efficient and rapid star formation mechanisms during the early cosmos (e.g., Glazebrook et al. 2017; Schreiber et al. 2018a, 2018b; D’Eugenio et al. 2020; Valentino et al. 2020; Marsan et al. 2022). This is likely in the form of bursts of extreme star formation ( $300 - 1000 M_{\odot} \text{ yr}^{-1}$ ) lasting  $\sim 100\text{--}300$  Myr (e.g., Tacconi et al. 2006; Barro et al. 2014; Aravena et al. 2016; Yang et al. 2017; D’Eugenio et al. 2020; Zavala et al. 2022), after which the cold gas reservoir is depleted, heated, and/or blown out.

In our model, we make the following simplistic assumptions to derive descendant quiescent galaxy masses to compare with observations and models in the literature. We assume that the current SFR of a mock galaxy is maintained for an additional 50–300 Myr, in alignment with the gas depletion timescales seen in the aforementioned observational studies. These timescales are randomly sampled from a uniform grid with steps of  $\delta_{\text{time}} = 5$  Myr. Then, we forward evolve the galaxy, adding the additional stellar mass built assuming a continuous SFR over the given timescale; to determine the redshift of the “quenched” galaxy, we add the same timescale to the galaxy’s star-forming redshift (which was assigned in the initial model).

There are caveats to our simple forward evolution model: we do not include any prescriptions for AGN feedback, merging activity, additional gas accretion, or any other potential phenomenon that could theoretically slow down or speed up the SFRs/gas depletion timescales assigned to our mock galaxies. We see this as a strength of this model: there is no need to invoke complex assumptions on, e.g., available gas mass reservoirs and potential heating mechanisms in order to successfully reproduce observed massive galaxy population densities (as shown in Section 4.2). Furthermore, it is entirely possible that these additional second-order corrections can be accounted for in the marginalization/uncertainty propagation process. For example, minor bumps or dips in the IRLF due to major mergers triggering IR luminous star formation would be accounted for in the uncertainty in the IRLF parameter space. Thus, while a more sophisticated model may be more “complete” in its physical assumptions, it does not appear necessary to describe massive galaxy evolution to a first-order approximation.

#### 4.3.2. Quiescent Descendant Densities

In the right panel of Figure 7, we show the expected evolution in the number density of massive DSFG quiescent descendants derived in Section 4.3.1. In general, this model does an excellent job at modeling observed number counts in heterogeneous samples of massive, passive galaxies across  $z = 2\text{--}6$  (Muzzin et al. 2013; Straatman et al. 2014; Toft et al. 2014; Davidzon et al. 2017; Schreiber et al. 2018a; Merlin et al. 2019; Santini et al. 2019; Carnall et al. 2020; Shahidi et al. 2020). This is particularly promising because of the

heterogeneity across these samples: each of the aforementioned bodies of work has their own methods/thresholds of selection and characterization of passive galaxies, yet we are able to successfully reproduce similar estimates across such diversity. Some have slightly lower stellar mass thresholds (e.g.,  $M > 10^{10} M_{\odot}$ ; Carnall et al. 2020; Shahidi et al. 2020), or are only complete at higher stellar masses than those explored in this work (e.g.,  $M > 10^{11} M_{\odot}$  at  $z = 3\text{--}4$ ; Muzzin et al. 2013). Some employ specific SFR (sSFR =  $\text{SFR}/M_{\star}$ ) cuts that capture post-starbursts (e.g.,  $\log(\text{sSFR yr}^{-1}) < -10$ ; Straatman et al. 2014), which are likely more in line with our model’s assumptions/definition of massive DSFG descendants, while others focus strictly on fully quiescent galaxies (e.g.,  $\log(\text{sSFR yr}^{-1}) < -11$ ; Davidzon et al. 2017).

Our model most disagrees with the work of Girelli et al. (2019), where we predict nearly an order of magnitude more passive galaxies at  $3 < z < 4$ . While Girelli et al. (2019) placed similar stellar mass and sSFR cuts on their sample as some of the aforementioned studies, their quiescent galaxy selection technique is unique. They develop a color selection criteria based on mock quiescent galaxy colors generated from stellar population synthesis models. They employed an exponentially declining star formation history with a short  $e$ -folding time (100–300 Myr) such that the galaxies are quiescent at the time of observation ( $\sim 1\text{--}2$  Gyr into the universe). In effect, this could mean that post-starburst galaxies are not captured in this color space, as their colors are likely bluer than the objects discovered in Girelli et al. (2019). As mentioned before, it is likely that our model quiescent descendants fall into the post-starburst category, and therefore would be missed in the Girelli et al. (2019) study, possibly accounting for such a discrepancy in number densities.

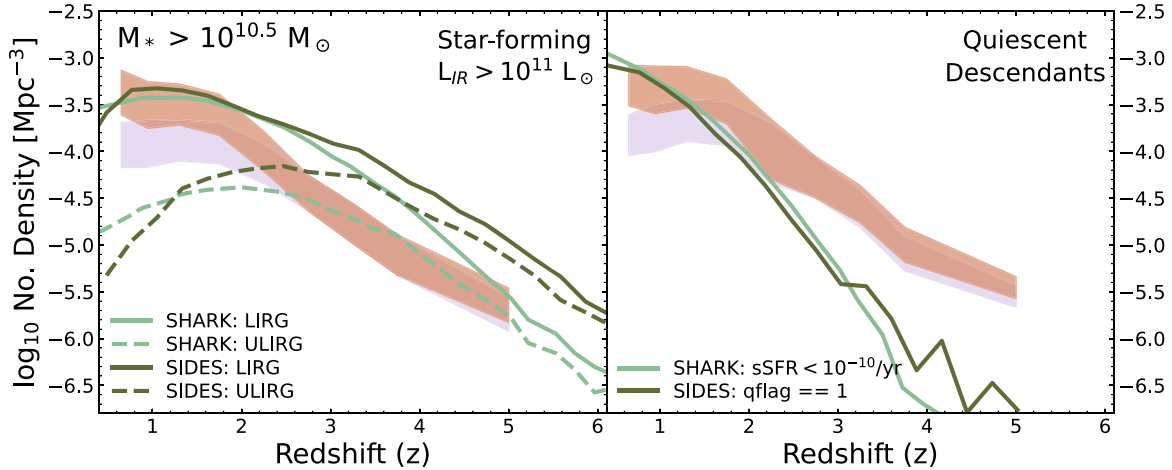
Like the ancestral massive DSFG population, this model predicts an order of magnitude increase in the number density of passive galaxies from  $z = 6$  to 3 (i.e., over  $\sim 1$  Gyr), a growth rate similar to that seen in Santini et al. (2019) and Merlin et al. (2019), and another order of magnitude over nearly twice the amount of time from  $z = 3$  to 1.5. This is unsurprising given the relatively short gas-to-stars duty cycles used in this work and exhibited by DSFGs at high redshifts (e.g., Tacconi et al. 2006; Aravena et al. 2016).

### 4.4. Comparison to Models

In Figure 8, we compare our results to other models relevant to this work as evident by their established success in reproducing properties and/or observations of submillimeter-bright galaxies. First is the SHARK semianalytical model (Lagos et al. 2018, 2020), which well-reproduced on-sky number counts at  $\lambda = 600 \mu\text{m}$  to 2 mm across  $1 < z < 4$ . SHARK is a physical model that follows the formation of galaxies in a  $\Lambda$  cold dark matter universe using a large suite of physical models, e.g., star formation, as well as stellar and AGN feedback. We also employ the Simulated Infrared Dusty Extragalactic Sky (SIDES) empirical model (B  thermin et al. 2017), which is successful in reproducing number on-sky counts of single-dish (i.e., confusion-limited) instruments, e.g., Herschel PACS and SPIRE observations at  $\lambda = 70\text{--}500 \mu\text{m}$ . SIDES was built using compiled observational results on the SFMS and stellar-to-halo mass relationships.

To compare the SIDES model, we use the publicly available catalogs,<sup>13</sup> taking all galaxies with stellar mass  $> 10^{10.5} M_{\odot}$  in

<sup>13</sup> The SIDES catalogs can be found here: <http://cesam.lam.fr/sides>.



**Figure 8.** The evolution of the number density of massive galaxies: our model vs. other models. Our model is presented as the red band, with the non-duty-cycle-corrected values represented by the purple band. The SHARK model (Lagos et al. 2018) is represented by the light green lines, while the SIDES model (B  thermin et al. 2017) corresponds to the dark green lines. Left: solid lines represent the massive LIRG ( $L_{\text{IR}} \geq 10^{11} L_{\odot}$ ) population, while dashed lines correspond to the massive ULIRG ( $L_{\text{IR}} \geq 10^{12} L_{\odot}$ ). Right: the SHARK model quiescent galaxies are selected primarily as a post-starburst ( $\text{sSFR} < 10^{-10} \text{yr}^{-1}$ ) population as this is likely most in line with our mock descendant population, whereas the SIDES galaxies are selected from a publicly available catalog and categorized as strictly passive. See Sections 4.2–4.4 for more details.

the same redshift bins as defined in Table 4. We use the passive galaxies flag ( $\text{qflag}=1$ ) to separate star-forming and quiescent galaxies, and further divide the star-forming sample into LIRGs and ULIRGs ( $L_{\text{IR}} > 10^{12} L_{\odot}$ ) by converting the provided SFRs into IR luminosities using the same conversion employed in our model (Murphy et al. 2011; Kennicutt & Evans 2012; see Section 2.1). To pull an analogous massive DSFG population from the SHARK model (C. Lagos, private communication), we use the same stellar mass, redshift, and IR luminosity criteria; for a comparative quiescent galaxy sample, we include post-starburst/near-quenching galaxies in the SHARK passive galaxy sample, defined as galaxies with  $\log(\text{sSFR yr}^{-1}) < -10$ . Galaxy properties (i.e., SFR and stellar mass) in both populations were convolved with random Gaussian noise of width  $\sim 0.2$  dex prior to calculating number and stellar mass densities.

At  $z < 2$ , our (duty-corrected) model is in agreement with the massive LIRG (aka massive DSFG) estimates in both SHARK and SIDES. Beyond this redshift, the SIDES model predicts roughly three times as many massive DSFGs at all epochs, while the SHARK model predicts only twice as many up until  $z \sim 4$  when our results drop back into alignment.

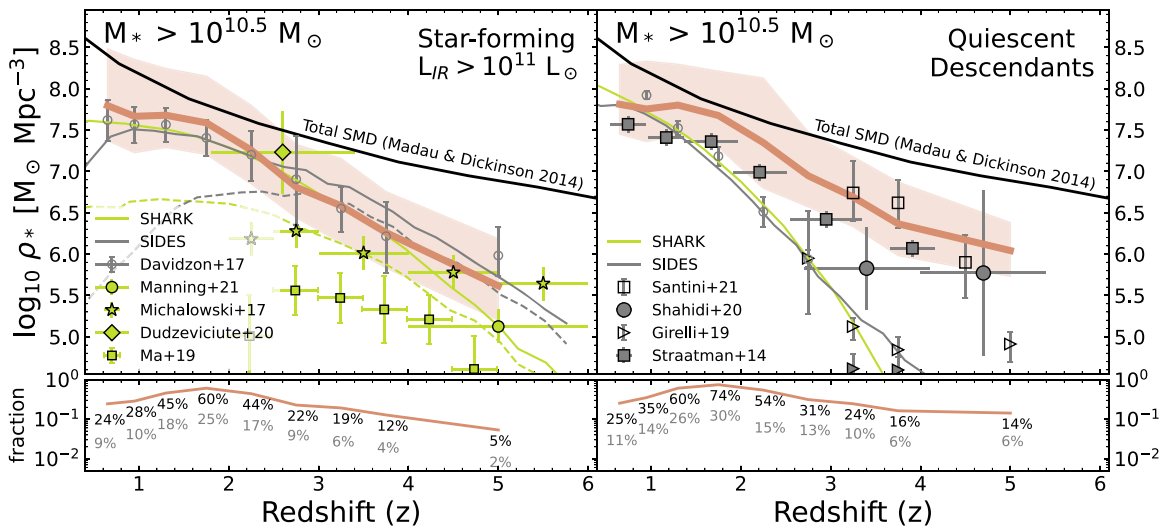
While all mock massive DSFGs in this work have IR luminosities of LIRG status or brighter across  $0.5 < z < 6$ , the majority (75%–99%) of massive DSFGs exhibit even brighter ULIRG-like luminosities at  $z \gtrsim 2$ . Thus, a more appropriate population comparison might be the ULIRG population in both the SIDES and SHARK models. Indeed, when comparing to the ULIRG population in both models, our results are significantly less discrepant, with a near-perfect agreement in predictions from the SHARK massive ULIRG model population at  $z > 2$ . Consistency across the SHARK model and this work may be due to their careful and considerate treatment in relating dust surface densities with attenuation in the stellar continuum, including an energy balance in the FIR. One potential reason that the SIDES model, which is empirically based and similar in build to this work, produces a higher density of massive IR luminous galaxies is due to differing SFR calculations. In SIDES,  $L_{\text{IR}}$  is converted to SFR using the Kennicutt (1998) equation, while in this work we use the

relationship from Kennicutt & Evans (2012), which produces SFRs roughly 24% lower than the former. We tested our model using the same  $L_{\text{IR}}$ -to-SFR conversion as in SIDES and found that this SFR prescription brings our models into agreement for the massive ULIRG population, but our quiescent descendant population is overproduced by up to 0.5 dex when compared to observations.

As shown in the right side of Figure 8, both the SIDES and SHARK models produce significantly fewer massive quiescent galaxies at  $z > 2$  than this model. The SIDES team used SMF results from Davidzon et al. (2017) to populate their model; this limits their model quiescent sample to strictly passive galaxies ( $\log_{10}(\text{sSFR yr}^{-1}) < -11$ ), while many of our model quiescent descendants are more likely post-starbursts ( $\log_{10}(\text{sSFR yr}^{-1}) < -10$ ). Thus, it is unsurprising that the SIDES quiescent number densities, like those reported in Davidzon et al. (2017), are up to a decade less numerous in population density than our model predicts at  $z > 2$ . For the SHARK model, we capture post-starburst and strictly passive galaxies, yet see predictions as discrepant as those in the SIDES model. The massive end of the total galaxy SMF in SHARK appears most sensitive to AGN feedback efficiency, but it is unclear if weaker AGN feedback would (or would not) produce sufficient populations of massive quiescent galaxies to match our predictions at earlier times.

#### 4.5. Stellar Mass Density

Cosmic volume-averaged studies over the last decade show how instrumental the DSFG population is in understanding stellar mass buildup over cosmic time: their intense dust-obscured SFRs dominate over unobscured (UV/OP-bright) star formation at  $1 \lesssim z \lesssim 4$  (Casey et al. 2014; Madau & Dickinson 2014; Zavala et al. 2021). However, the integrated cosmic star formation rate density (cSFRD) appears roughly  $2 \times (0.2\text{--}0.4 \text{ dex})$  greater than the measured cosmic stellar mass density out to  $z \sim 3$  (Madau & Dickinson 2014; Leja et al. 2015), indicating an issue with “missing” stellar mass in direct measurements. Most of the potential explored drivers of this discrepancy are systemics such as overestimated SFRs



**Figure 9.** The evolution of the stellar mass density contributions by massive galaxies. Our model is represented by the red line, with  $1\sigma$  uncertainties shown as the paler contours. The black line represents the integrated SFR density from Madau & Dickinson (2014). Left: green points represent observations of massive DSFGs from the literature (Michalowski et al. 2017; Ma et al. 2019; Dudzevičiūtė et al. 2020; Manning et al. 2022); note that several of these are incomplete in stellar mass constraints for their DSFGs, as the majority are insufficiently detected at rest-frame UV/optical wavelengths. Uncertainties on the abscissa represent redshift ranges. We also compare to UV/OP-bright galaxies derived in Davidzon et al. (2017; gray points), and to massive LIRGs (solid) and ULIRGs (dashed) in the SHARK (green; Lagos et al. 2020) and SIDES (gray; Béthermin et al. 2017) models. Right: gray points represent observations from the literature. SHARK (green) and SIDES model galaxies are selected in the same way as in Figure 8. Bottom: the fraction of the total SMD (derived from the integrated SFRD) that our model massive DSFGs make up. The red line and black percentages correspond to the red band in the upper panels, and gray percentages represent the lower limits.

(Tomczak et al. 2016), underestimated stellar masses due to inflexible SED analysis (Leja et al. 2015, 2020), and photometry limitations (Madau & Dickinson 2014). Missing from this conversation is a potential key component: it is the incomplete accounting of stellar mass from the massive dust-obscured population (see Ma et al. 2019 for one of the only attempts). Owing to their heavily obscured nature, the field is still lacking in sufficient stellar continua observations for complete samples of DSFGs. One may argue that DSFGs are extremely rare and therefore not as significant in this discussion; however, as shown by the cSFRD, extreme properties of rare galaxies can be major—or even dominant—contributors when computing cosmic volume-averaged quantities.

In Figure 9, we show the predicted contributions of massive DSFGs and their descendants to the cosmic stellar mass density (SMD). At  $z = 0.5$ , this model predicts that massive DSFGs hold from one-tenth to a quarter of the total SMD derived from the integrated cSFRD (Madau & Dickinson 2014). This fraction increases to a peak around  $z \sim 2$ , where massive DSFGs may contribute as much as 60% of the SMD (and as little as 25%). Such high fractions translate to a 0.1–0.3 dex increase on the observed SMD, which would be sufficient to bring the observed and integrated SMDs into alignment. Beyond  $z \sim 2$ , the fraction steadily declines to 2%–5% at  $z = 4$ –6, demonstrating how this population is likely less significant in cosmic volume studies in the high- $z$  universe (possibly because this population may still be growing to sufficient sizes during these epochs). We see a similar rise and fall in the quiescent descendant population, reaching a maximum fraction of 30%–74% at  $z \sim 2$ .

When compared to other studies, we predict that the massive dust-obscured population is an equally important contributor to the SMD as are the massive UV/OP-bright star-forming galaxies captured in Davidzon et al. (2017). Furthermore, this model is in rough agreement with the handful of existing

constraints on observed massive DSFGs (e.g., Michałowski et al. 2017; Dudzevičiūtė et al. 2020; Manning et al. 2022), with the exception of those in Ma et al. (2019), which probe a much more extreme and rare subpopulation of massive DSFGs. We find that the lower SMD estimates of our quiescent descendant population are very similar to the Straatman et al. (2014) ZFOURGE mass-limited ( $M > 10^{10.6} M_{\odot}$ ) sample, while our median prediction is more in agreement with Santini et al. (2019). Interestingly, our model results agree most with SMD predictions for the massive LIRG population in both the SIDES and SHARK models, whereas our predicted number density evolution was more in line with massive ULIRGs in both simulations. This likely means that the intrinsic stellar mass distributions in our model are skewed higher. In other words, the difference in SMDs is likely driven not by the number density evolution ( $\Phi$ ) but by the shape of the high mass end of the SMF; our shallow broken power-law extension beyond the knee likely yields a higher abundance of extremely massive galaxies.

## 5. Other Model Implications

### 5.1. AGN $L_{IR}$ Correction Is Necessary

In Section 2.2, we described how we corrected the mock galaxies’ IR luminosities for AGN contamination so we could use  $L_{IR}$  to derive SFRs, which were in turn used to derive stellar masses (via distributing SFRs over the star-forming main sequence). We tested the model results without the AGN correction procedure and found that the resulting SMF parameters change by an average of  $\sim 5\%$ . While seemingly insignificant, this propagates into an overestimation of the number density of massive DSFGs and their descendants by roughly 0.5 dex at  $z > 2$ . This is particularly relevant for studies with galaxy populations selected via IR emission at  $\lambda = 8$ –50  $\mu\text{m}$  wavelengths without sufficient multiwavelength data to rule out major AGN contributions from  $\lambda_{\text{rest}} = 1$ –1000  $\mu\text{m}$ . For these

studies, it is possible that some of the selected galaxies have IR luminosities driven by AGN heating rather than by star formation (Ciesla et al. 2015).

### 5.2. The Duty Cycle and Multiple Star Formation Bursts

In Section 4.3.1, we discussed how we create the mock quiescent descendant population by sampling gas depletion timescales of 50–300 Myr, and in Section 4.1, we discussed how we correct the model number densities to match observed number counts using a duty cycle correction factor of 500 Myr. Why do these values not match?

The range of gas depletion timescales used in this work are derived from direct measurements in the literature. They are generated using a closed-box assumption such that, given a galaxy’s measured SFR and molecular gas mass, one can compute how long the galaxy could sustain such an SFR before depleting its molecular gas reservoir. Yet, observations and simulations show that the star formation history of massive DSFGs is stochastic, with several bursts over short timescales due to a mix of mergers, sustained gas flows, and/or violent disk instabilities (Narayanan et al. 2015; Tacchella et al. 2016; Cowie et al. 2018; Hodge et al. 2019; McAlpine et al. 2019; Lagos et al. 2020). It can take up to 1 Gyr for the halo to become hot enough (through, e.g., AGN feedback or halo shock heating) such that galaxy-wide star formation finally abates.

Thus, in the context of this model, we interpret these different timescales as referring to two distinct events. The first—the gas depletion timescale—refers to the current burst of star formation, while the duty cycle correction refers to the overall “average” time it takes a massive DSFG to finally reach quiescence after several bursts.

### 5.3. Relation to “Optically Dark” Galaxies

One of the main goals of the current study was to predict the “true” massive DSFG number density at  $z > 0.5$ , as the majority of DSFGs do not have sufficient rest-frame UV/optical detections to constrain their total stellar mass. A smaller portion of the massive DSFG population is nearly invisible at these wavelengths and thus sometimes referred to as “optically dark” galaxies (Wang et al. 2019; Williams et al. 2019; Talia et al. 2021; Manning et al. 2022). This population is believed to be extremely dust-obscured, gas-rich, and already massive ( $M \gtrsim 10^{10.5} M_{\odot}$ ) at  $z = 3$ –6. Furthermore, these studies posit that optically dark galaxies constitute the majority of the massive galaxy population in the first 2 Gyr of the universe, dominating in stellar mass density and in SFR density.

Consistent with observations, our model also suggests a strong, if not dominating, prevalence of massive, heavily dust-obscured galaxies at similar epochs. It is interesting that we arrive at a similar conclusion using relatively modest assumptions on the properties and evolution of massive DSFGs. Moreover, our mock quiescent descendant population requires that some of these descendants are observed as high- $z$  post-starbursts (as opposed to strictly passive). This would manifest observationally as a significant population of massive gas-poor galaxies with young ages from recent bursts, yet strong Balmer breaks from evolved stellar populations that developed during earlier bursts. Indeed, recent studies on massive quiescent galaxies at  $z \gtrsim 3$  broadly corroborate these implications (Alcalde Pampliega et al. 2019; Carnall et al. 2020; D’Eugenio

et al. 2020; Forrest et al. 2020), suggesting a strong connection between massive DSFGs and quiescent galaxies in the early cosmos. Scheduled wide-field JWST surveys with sufficient near-IR depths to capture these faint and rare populations such as the COSMOS-Web Survey (Casey et al. 2022) will be instrumental in constraining the stellar properties and number densities of these galaxy subtypes, and in more firmly investigating their connections through cosmic time.

## 6. Conclusions

Owing to their extreme dust obscuration, much uncertainty still exists surrounding the stellar mass growth and content in DSFGs at  $z > 1$ . Observations and simulations establish strong links between DSFGs and massive quiescent galaxies at high- $z$ , but a full characterization of the massive, dust-obscured galaxy population is still missing and critical to understanding massive galaxy evolution in the early cosmos. In this work, we present a novel numerical model that uses empirical data on DSFGs to model their stellar mass contributions across the first  $\sim 10$  Gyr of cosmic time. This method allows us to circumvent the observational limitations and restrictions that typically prevent galaxy evolution studies from fully capturing this extreme population. It also allows us to marginalize over the diverse range of properties observed in DSFG populations (e.g., position on the star-forming main sequence, gas depletion timescales), and to carefully fold in uncertainties in the underlying model assumptions (e.g., uncertainties in the parameters describing the IRLF). Where possible, we take the most conservative and/or simple assumptions.

With this model, we generate a dust-obscured SMF with a high mass power-law extension that reaches beyond the “knee” (i.e., characteristic mass) of UV/OP-based SMFs in the literature. We find that two SMF parameters—the number density evolution ( $\Phi_{\star}$ ) and the slope of the high mass end ( $\beta$ )—are not significantly dependent on whether more DSFGs are starbursts (and therefore lower in stellar mass) or main-sequence galaxies (and therefore higher in stellar mass). Almost all ( $\sim 99\%$ ) of the galaxies at the massive end of our dust-obscured SMF (at  $M \geq 10^{10.5} M_{\odot}$ ) have LIRG-like luminosities ( $L_{\text{IR}} \geq 10^{11} L_{\odot}$ ), with the majority ( $> 75\%$ ) exhibiting ULIRG luminosities or greater ( $L_{\text{IR}} \geq 10^{12} L_{\odot}$ ). When compared to UV/OP-based estimates, we predict massive DSFGs are equally if not more prevalent in number at the massive end of galaxy evolution at  $z > 1$ , constituting as much as 50%–100% of all massive star-forming galaxies at  $M \geq 10^{11} M_{\odot}$ .

We predict the volume-averaged number density of massive DSFGs and find general agreement with observations, but the lack of sufficient measurements on real, observed DSFG stellar masses leave up to 2 orders of magnitude in uncertainty at  $z > 4$ . What is perhaps more striking is that, after forward-modeling mock massive DSFGs to their quiescent descendants, we find remarkable agreement with our predictions and the number densities of observed massive passive galaxies at high- $z$ . This means that, to first order, massive dust-obscured galaxies with ULIRG-like luminosities or greater are a sufficient ancestral population to describe the prevalence of massive quiescent galaxies at  $z > 1$ . Between the two populations (massive DSFGs and their quiescent descendants), we predict an intense epoch of growth during the  $\sim 1$  Gyr from  $z = 6$  to 3 during which the majority of the most massive galaxies at high- $z$  grow and then quench. We also show that

massive DSFGs and their quiescent descendants are significant contributors to the cosmic stellar mass density, constituting as much as 25%–60% during the peak of cosmic star formation at  $z \sim 2$ ; this could be sufficient in closing the gap between the integrated cSFRD and the directly measured stellar mass density.

In order to close the uncertainty on this model's implications and better understand massive galaxy evolution in the first 2 Gyr of the cosmos, future studies should focus on collating large samples of DSFGs (perhaps complete down to ULIRG-like luminosities) with both strong redshift constraints and deep near-IR coverage. What drives the era of massive galaxy growth between  $z = 3$  and 6? What astrophysical processes precipitate extreme rates of star formation (and dust production) so early on? And what causes such massive galaxies to rapidly quench during an epoch where cold molecular gas is so readily available for most other galaxies? We can only hope to answer these questions through a coordinated symbiosis between observatories across the electromagnetic spectrum. Specifically, we will be able to test, explore, and constrain the implications of this work by coupling deep near-IR observations afforded by JWST with sensitive, high-resolution submillimeter interferometers such as ALMA.

### Acknowledgments

A.S.L. acknowledges support for this work provided by NASA through the NASA Hubble Fellowship Program grant No. HST-HF2-51511.001-A, awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. A.S.L. also acknowledges support from the Ford Foundation. C.M.C. thanks the National Science Foundation for support through grants AST-1814034 and AST-2009577 as well as the University of Texas at Austin College of Natural Sciences for support. C.M.C. also acknowledges support from the Research Corporation for Science Advancement from a 2019 Cottrell Scholar Award sponsored by IF/THEN, an initiative of Lyda Hill Philanthropies. This research made use of Astropy,<sup>14</sup> a community-developed core Python package for Astronomy (Astropy Collaboration et al. 2013, 2018, 2022).


### ORCID iDs

Arianna S. Long  <https://orcid.org/0000-0002-7530-8857>

Caitlin M. Casey  <https://orcid.org/0000-0002-0930-6466>

Claudia del P. Lagos  <https://orcid.org/0000-0003-3021-8564>

Erini L. Lambrides  <https://orcid.org/0000-0003-3216-7190>

Jorge A. Zavala  <https://orcid.org/0000-0002-7051-1100>

Jaclyn Champagne  <https://orcid.org/0000-0002-6184-9097>

Olivia R. Cooper  <https://orcid.org/0000-0003-3881-1397>

Asantha R. Cooray  <https://orcid.org/0000-0002-3892-0190>

### References

- Alcalde Pampliega, B., Pérez-González, P. G., Barro, G., et al. 2019, *ApJ*, 876, 135
- Algera, H. S. B., Inami, H., Oesch, P. A., et al. 2023, *MNRAS*, 518, 6142
- Aoyama, S., Hirashita, H., Lim, C.-F., et al. 2019, *MNRAS*, 484, 1852
- Aravena, M., Spilker, J. S., Bethermin, M., et al. 2016, *MNRAS*, 457, 4406
- Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., et al. 2022, *ApJ*, 935, 167
- 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
- Barro, G., Trump, J. R., Koo, D. C., et al. 2014, *ApJ*, 795, 145
- Baugh, C. M., Lacey, C. G., Frenk, C. S., et al. 2005, *MNRAS*, 356, 1191
- Bethermin, M., Wu, H.-Y., Lagache, G., et al. 2017, *A&A*, 607, A89
- Bisigello, L., Caputi, K. I., Grogin, N., & Koekemoer, A. 2018, *A&A*, 609, A82
- Brennan, R., Pandya, V., Somerville, R. S., et al. 2015, *MNRAS*, 451, 2933
- Brisbin, D., Miettinen, O., Aravena, M., et al. 2017, *A&A*, 608, A15
- Brown, A., Nayyeri, H., Cooray, A., et al. 2019, *ApJ*, 871, 87
- Buat, V., Ciesla, L., Boquien, M., Malek, K., & Burgarella, D. 2019, *A&A*, 632, A79
- Carnall, A. C., Walker, S., McLure, R. J., et al. 2020, *MNRAS*, 496, 695
- Casey, C. M., Hodge, J., Zavala, J. A., et al. 2018a, *ApJ*, 862, 78
- Casey, C. M., Kartaltepe, J. S., & Drakos, N. E. 2022, arXiv:2211.07865
- Casey, C. M., Narayanan, D., & Cooray, A. 2014, *PhR*, 541, 45
- Casey, C. M., Zavala, J. A., Aravena, M., et al. 2019, *ApJ*, 887, 55
- Casey, C. M., Zavala, J. A., Spilker, J., et al. 2018b, *ApJ*, 862, 77
- Chabrier, G. 2003, *PASP*, 115, 763
- Ciesla, L., Charmandaris, V., Georgakakos, A., et al. 2015, *A&A*, 576, A10
- Cowie, L. L., González-López, J., Barger, A. J., et al. 2018, *ApJ*, 865, 106
- Cowley, W. I., Lacey, C. G., Baugh, C. M., & Cole, S. 2015, *MNRAS*, 446, 1784
- da Cunha, E., Walter, F., Smail, I. R., et al. 2015, *ApJ*, 806, 110
- Daddi, E., Dickinson, M., Morrison, G., et al. 2007, *ApJ*, 670, 156
- Daddi, E., Elbaz, D., Walter, F., et al. 2010, *ApJL*, 714, L118
- Davidzon, I., Ilbert, O., Laigle, C., et al. 2017, *A&A*, 605, A70
- Deshmukh, S., Caputi, K. I., Ashby, M. L. N., et al. 2018, *ApJ*, 864, 166
- D'Eugenio, C., Daddi, E., Gobat, R., et al. 2020, *ApJL*, 892, L2
- Domínguez Sánchez, H., Pérez-González, P. G., Esquej, P., et al. 2016, *MNRAS*, 457, 3743
- Dowell, C. D., Conley, A., Glenn, J., et al. 2014, *ApJ*, 780, 75
- Driver, S. P., Popescu, C. C., Tuffs, R. J., et al. 2008, *ApJL*, 678, L101
- Dudzevičiūtė, U., Smail, I., Swinbank, A. M., et al. 2020, *MNRAS*, 494, 3828
- Duivendoorn, S., Oliver, S., Scudder, J. M., et al. 2018, *MNRAS*, 477, 1099
- Dunlop, J. S., McLure, R. J., Biggs, A. D., et al. 2017, *MNRAS*, 466, 861
- Eisenstein, D. J., Weinberg, D. H., Agol, E., et al. 2011, *AJ*, 142, 72
- Forrest, B., Marsan, Z. C., Annunziatella, M., et al. 2020, *ApJ*, 903, 47
- Fu, H., Yan, L., Scoville, N. Z., et al. 2010, *ApJ*, 722, 653
- Fudamoto, Y., Oesch, P. A., Magnelli, B., et al. 2020, *MNRAS*, 491, 4724
- Girelli, G., Bolzonella, M., & Cimatti, A. 2019, *A&A*, 632, A80
- Glazebrook, K., Schreiber, C., Labbé, I., et al. 2017, *Natur*, 544, 71
- Granato, G. L., Lacey, C. G., Silva, L., et al. 2000, *ApJ*, 542, 710
- Grazian, A., Fontana, A., Santini, P., et al. 2015, *A&A*, 575, A96
- Gruppioni, C., Béthermin, M., Loiacono, F., et al. 2020, *A&A*, 643, A8
- Gruppioni, C., Pozzi, F., Rodighiero, G., et al. 2013, *MNRAS*, 432, 23
- Hainline, L. J., Blain, A. W., Smail, I., et al. 2011, *ApJ*, 740, 96
- Hayward, C. C., Narayanan, D., Kereš, D., et al. 2013, *MNRAS*, 428, 2529
- Hayward, C. C., Sparre, M., Chapman, S. C., et al. 2021, *MNRAS*, 502, 2922
- Hickox, R. C., Wardlow, J. L., Smail, I., et al. 2012, *MNRAS*, 421, 284
- Hodge, J. A., Smail, I., Walter, F., et al. 2019, *ApJ*, 876, 130
- Hodge, J. A., Swinbank, A. M., Simpson, J. M., et al. 2016, *ApJ*, 833, 103
- Hopkins, P. F., Younger, J. D., Hayward, C. C., Narayanan, D., & Hernquist, L. 2010, *MNRAS*, 402, 1693
- Hwang, Y.-H., Wang, W.-H., Chang, Y.-Y., et al. 2021, *ApJ*, 913, 6
- Ivion, R. J., Lewis, A. J. R., Weiss, A., et al. 2016, *ApJ*, 832, 78
- Johnston, R., Vaccari, M., Jarvis, M., et al. 2015, *MNRAS*, 453, 2540
- Juneau, S., Dickinson, M., Bournaud, F., et al. 2013, *ApJ*, 764, 176
- Kartaltepe, J. S., Sanders, D. B., Le Floch, E., et al. 2010, *ApJ*, 709, 572
- Kennicutt, R. C., & Evans, N. J. 2012, *ARA&A*, 50, 531
- Kennicutt, R. C. J. 1998, *ApJ*, 498, 541
- Kirkpatrick, A., Pope, A., Sajina, A., et al. 2015, *ApJ*, 814, 9
- Koprowski, M. P., Dunlop, J. S., Michałowski, M. J., et al. 2017, *MNRAS*, 471, 4155
- Kurczynski, P., Gawiser, E., Acquaviva, V., et al. 2016, *ApJL*, 820, L1
- Lagos, C. d. P., da Cunha, E., Robotham, A. S. G., et al. 2020, *MNRAS*, 499, 1948
- Lagos, C. d. P., Tobar, R. J., Robotham, A. S. G., et al. 2018, *MNRAS*, 481, 3573
- Lambrides, E. L., Chiaberge, M., Heckman, T., et al. 2020, *ApJ*, 897, 160
- Lee, N., Sanders, D. B., Casey, C. M., et al. 2013, *ApJ*, 778, 131
- Lee, N., Sanders, D. B., Casey, C. M., et al. 2015, *ApJ*, 801, 80
- Leja, J., Speagle, J. S., Johnson, B. D., et al. 2020, *ApJ*, 893, 111

<sup>14</sup> <http://www.astropy.org>

- Leja, J., van Dokkum, P. G., Franx, M., & Whitaker, K. E. 2015, *ApJ*, 798, 115
- Lemaux, B. C., Le Floc'h, E., Le Fèvre, O., et al. 2014, *A&A*, 572, A90
- Leroy, A. K., Sandstrom, K. M., Lang, D., et al. 2019, *ApJS*, 244, 24
- Leslie, S. K., Schinnerer, E., Liu, D., et al. 2020, *ApJ*, 899, 58
- Lim, C.-F., Wang, W.-H., Smail, I., et al. 2020, *ApJ*, 889, 80
- Ma, J., Cooray, A., Nayyeri, H., et al. 2019, *ApJS*, 244, 30
- Madau, P., & Dickinson, M. 2014, *ARA&A*, 52, 415
- Magnelli, B., Lutz, D., Saintonge, A., et al. 2014, *A&A*, 561, A86
- Manning, S. M., Casey, C. M., Zavala, J. A., et al. 2022, *ApJ*, 925, 23
- Marsan, Z. C., Muzzin, A., Marchesini, D., et al. 2022, *ApJ*, 924, 25
- Martis, N. S., Marchesini, D., Brammer, G. B., et al. 2016, *ApJL*, 827, L25
- Matthee, J., & Schaye, J. 2019, *MNRAS*, 484, 915
- McAlpine, S., Smail, I., Bower, R. G., et al. 2019, *MNRAS*, 488, 2440
- Merlin, E., Fortuni, F., Torelli, M., et al. 2019, *MNRAS*, 490, 3309
- Michałowski, M. J., Dunlop, J. S., Koprowski, M. P., et al. 2017, *MNRAS*, 469, 492
- Miettinen, O., Delvecchio, I., Smolčić, V., et al. 2017, *A&A*, 606, A17
- Montaña, A., Zavala, J. A., Aretxaga, I., et al. 2021, *MNRAS*, 505, 5260
- Murphy, E. J., Condon, J. J., Schinnerer, E., et al. 2011, *ApJ*, 737, 67
- Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013, *ApJ*, 777, 18
- Narayanan, D., Turk, M., Feldmann, R., et al. 2015, *Natur*, 525, 496
- Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007a, *ApJL*, 660, L43
- Noeske, K. G., Faber, S. M., Weiner, B. J., et al. 2007b, *ApJL*, 660, L47
- Pan, Z., Zheng, X., & Kong, X. 2017, *ApJ*, 834, 39
- Pannella, M., Carilli, C. L., Daddi, E., et al. 2009, *ApJL*, 698, L116
- Patil, P., Nyland, K., Lacy, M., et al. 2019, *ApJ*, 871, 109
- Pearson, W. J., Wang, L., Hurley, P. D., et al. 2018, *A&A*, 615, A146
- Peng, Y.-j., Lilly, S. J., Kovač, K., et al. 2010, *ApJ*, 721, 193
- Popesso, P., Concas, A., Cresci, G., et al. 2023, *MNRAS*, 519, 1526
- Popesso, P., Morselli, L., Concas, A., et al. 2019a, *MNRAS*, 490, 5285
- Popesso, P., Concas, A., Morselli, L., et al. 2019b, *MNRAS*, 483, 3213
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, *A&A*, 594, A13
- Rodighiero, G., Daddi, E., Baronchelli, I., et al. 2011, *ApJL*, 739, L40
- Rodighiero, G., Renzini, A., Daddi, E., et al. 2014, *MNRAS*, 443, 19
- Santini, P., Fontana, A., Castellano, M., et al. 2017, *ApJ*, 847, 76
- Santini, P., Merlin, E., Fontana, A., et al. 2019, *MNRAS*, 486, 560
- Sargent, M. T., Béthermin, M., Daddi, E., & Elbaz, D. 2012, *ApJL*, 747, L31
- Sargent, M. T., Daddi, E., Béthermin, M., et al. 2014, *ApJ*, 793, 19
- Schechter, P. 1976, *ApJ*, 203, 297
- Schreiber, C., Glazebrook, K., Nanayakkara, T., et al. 2018a, *A&A*, 618, A85
- Schreiber, C., Labbé, I., Glazebrook, K., et al. 2018b, *A&A*, 611, A22
- Schreiber, C., Pannella, M., Elbaz, D., et al. 2015, *A&A*, 575, A74
- Schreiber, C., Pannella, M., Leiton, R., et al. 2017, *A&A*, 599, A134
- Schwarz, G. 1978, *AnSta*, 6, 461
- Shahidi, A., Mobasher, B., Nayyeri, H., et al. 2020, *ApJ*, 897, 44
- Sherman, S., Jogee, S., Florez, J., et al. 2020, *MNRAS*, 491, 3318
- Simpson, J. M., Swinbank, A. M., Smail, I., et al. 2014, *ApJ*, 788, 125
- Smail, I., Dudzevičiūtė, U., Stach, S. M., et al. 2021, *MNRAS*, 502, 3426
- Somerville, R. S., & Davé, R. 2015, *ARA&A*, 53, 51
- Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, *ApJS*, 214, 15
- Spitler, L. R., Straatman, C. M. S., Labbé, I., et al. 2014, *ApJL*, 787, L36
- Straatman, C. M. S., Labbé, I., Spitler, L. R., et al. 2014, *ApJL*, 783, L14
- Symeonidis, M., Giblin, B. M., Page, M. J., et al. 2016, *MNRAS*, 459, 257
- Symeonidis, M., & Page, M. J. 2018, *MNRAS Lett.*, 479, L91
- Symeonidis, M., & Page, M. J. 2021, *MNRAS*, 503, 3992
- Tacchella, S., Dekel, A., Carollo, C. M., et al. 2016, *MNRAS*, 457, 2790
- Tacconi, L. J., Neri, R., Chapman, S. C., et al. 2006, *ApJ*, 640, 228
- Talia, M., Cimatti, A., Giulietti, M., et al. 2021, *ApJ*, 909, 23
- Tasca, L. A. M., Le Fèvre, O., Hathi, N. P., et al. 2015, *A&A*, 581, A54
- Thorne, J. E., Robotham, A. S. G., Davies, L. J. M., et al. 2021, *MNRAS*, 505, 540
- Toba, Y., Goto, T., Oi, N., et al. 2020, *ApJ*, 899, 35
- Toft, S., Smolčić, V., Magnelli, B., et al. 2014, *ApJ*, 782, 68
- Tomczak, A. R., Quadri, R. F., Tran, K.-V. H., et al. 2016, *ApJ*, 817, 118
- Tsai, C.-W., Eisenhardt, P. R. M., Wu, J., et al. 2015, *ApJ*, 805, 90
- Valentino, F., Tanaka, M., Davidzon, I., et al. 2020, *ApJ*, 889, 93
- van Dokkum, P. G., Quadri, R., Marchesini, D., et al. 2006, *ApJL*, 638, L59
- Wang, T., Schreiber, C., Elbaz, D., et al. 2019, *Natur*, 572, 211
- Wechsler, R. H., & Tinker, J. L. 2018, *ARA&A*, 56, 435
- Whitaker, K. E., Franx, M., Bezanson, R., et al. 2015, *ApJL*, 811, L12
- Whitaker, K. E., Franx, M., Leja, J., et al. 2014, *ApJ*, 795, 104
- Whitaker, K. E., Labbé, I., van Dokkum, P. G., et al. 2011, *ApJ*, 735, 86
- Whitaker, K. E., Pope, A., Cybulski, R., et al. 2017, *ApJ*, 850, 208
- Whitaker, K. E., van Dokkum, P. G., Brammer, G., & Franx, M. 2012, *ApJL*, 754, L29
- Williams, C. C., Labbe, I., Spilker, J., et al. 2019, *ApJ*, 884, 154
- Wuyts, S., Förster Schreiber, N. M., van der Wel, A., et al. 2011, *ApJ*, 742, 96
- Wuyts, S., Labbé, I., Franx, M., et al. 2007, *ApJ*, 655, 51
- Yang, C., Omont, A., Beelen, A., et al. 2017, *A&A*, 608, A144
- Zavala, J. A., Casey, C. M., da Cunha, E., et al. 2018, *ApJ*, 869, 71
- Zavala, J. A., Casey, C. M., Manning, S. M., et al. 2021, *ApJ*, 909, 165
- Zavala, J. A., Casey, C. M., Spilker, J., et al. 2022, *ApJ*, 933, 242
- Zehavi, I., Zheng, Z., Weinberg, D. H., et al. 2011, *ApJ*, 736, 59