

THE MOST MASSIVE GALAXIES AT $3.0 \leq z < 4.0$ IN THE NEWFIRM MEDIUM-BAND SURVEY: PROPERTIES AND IMPROVED CONSTRAINTS ON THE STELLAR MASS FUNCTION

DANILO MARCHESINI^{1,9}, KATHERINE E. WHITAKER^{2,9}, GABRIEL BRAMMER^{2,9}, PIETER G. VAN DOKKUM^{2,9}, IVO LABBÉ^{3,9},
 ADAM MUZZIN^{2,9}, RYAN F. QUADRI^{4,9}, MARISKA KRIEK^{5,9}, KYOUNG-SOO LEE^{6,9}, GREGORY RUDNICK^{7,9}, MARIJN FRANX⁴,

GARTH D. ILLINGWORTH⁸, AND DAVID WAKE²

¹ Department of Physics and Astronomy, Tufts University, Medford, MA 02155, USA

² Department of Astronomy, Yale University, New Haven, CT 06520-8101, USA

³ Carnegie Observatories, Pasadena, CA 91101, USA

⁴ Sterrewacht Leiden, Leiden University, NL-2300 RA Leiden, The Netherlands

⁵ Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA

⁶ Yale Center for Astronomy and Astrophysics, Departments of Physics and Astronomy, Yale University, New Haven, CT 06520, USA

⁷ Department of Physics and Astronomy, University of Kansas, Lawrence, KS 66045, USA

⁸ UCO/Lick Observatory, University of California, Santa Cruz, CA 95064, USA

Received 2010 April 9; accepted 2010 September 27; published 2010 November 24

ABSTRACT

We use the optical to mid-infrared coverage of the NEWFIRM Medium-Band Survey (NMBS) to characterize, for the first time, the properties of a mass-complete sample of 14 galaxies at $3.0 \leq z < 4.0$ with $M_{\text{star}} > 2.5 \times 10^{11} M_{\odot}$, and to derive significantly more accurate measurements of the high-mass end of the stellar mass function (SMF) of galaxies at $3.0 \leq z < 4.0$. The accurate photometric redshifts and well-sampled spectral energy distributions (SEDs) provided by the NMBS combined with the large surveyed area result in significantly reduced contributions from photometric redshift errors and cosmic variance to the total error budget of the SMF. The typical very massive galaxy at $3.0 \leq z < 4.0$ is red and faint in the observer’s optical, with a median r -band magnitude of $\langle r_{\text{tot}} \rangle = 26.1$, and median rest-frame $U - V$ colors of $\langle U - V \rangle = 1.6$. About 60% of the mass-complete sample has optical colors satisfying either the U - or the B -dropout color criteria, although $\sim 50\%$ of these galaxies has $r > 25.5$. We find that $\sim 30\%$ of the sample has star formation rates (SFRs) from SED modeling consistent with zero, although SFRs of up to $\sim 1\text{--}18 M_{\odot} \text{ yr}^{-1}$ are also allowed within 1σ . However, $>80\%$ of the sample is detected at $24 \mu\text{m}$, resulting in total infrared luminosities in the range $(0.5\text{--}4.0) \times 10^{13} L_{\odot}$. This implies the presence of either dust-enshrouded starburst activity (with SFRs of $600\text{--}4300 M_{\odot} \text{ yr}^{-1}$) and/or highly obscured active galactic nuclei (AGNs). The contribution of galaxies with $M_{\text{star}} > 2.5 \times 10^{11} M_{\odot}$ to the total stellar mass budget at $3.0 \leq z < 4.0$ is $\sim 8_{-3}^{+13}\%$. Compared to recent estimates of the stellar mass density in galaxies with $M_{\text{star}} \approx 10^9\text{--}10^{11} M_{\odot}$ at $z \sim 5$ and $z \sim 6$, we find an evolution by a factor of 2–7 and 3–22 from $z \sim 5$ and $z \sim 6$, respectively, to $z = 3.5$. The previously found disagreement at the high-mass end between observed and model-predicted SMFs is now significant at the 3σ level when only random uncertainties are considered. However, systematic uncertainties dominate the total error budget, with errors up to a factor of ~ 8 in the densities at the high-mass end, bringing the observed SMF in marginal agreement with the predicted SMF. Additional systematic uncertainties on the high-mass end could be potentially introduced by either (1) the intense star formation and/or the very common AGN activities as inferred from the MIPS $24 \mu\text{m}$ detections, and/or (2) contamination by a significant population of massive, old, and dusty galaxies at $z \sim 2.6$.

Key words: cosmology: observations – galaxies: evolution – galaxies: formation – galaxies: fundamental parameters – galaxies: high-redshift – galaxies: luminosity function, mass function – galaxies: stellar content – infrared: galaxies

Online-only material: color figures

1. INTRODUCTION

Understanding the formation mechanisms and evolution with cosmic time of galaxies is one of the major goals of observational cosmology. An effective approach to understand the physical processes governing the assembly of galaxies (and their relative importance as a function of cosmic time) is to directly measure the growth of the stellar mass in galaxies. Galaxies can build their stellar mass both from in situ star formation and/or merger events. The mean space density of galaxies per unit stellar mass, or stellar mass function (SMF), is one of the most fundamental

of all cosmological observables. The shape of the SMF retains the imprint of galaxy formation and evolution processes. Therefore, the SMF and its evolution with cosmic time represent a powerful tool to constrain the physical mechanisms regulating the assembly and the evolution of galaxies (Baugh 2006).

In the past decade, significant observational progress has been made in the measurement of the SMF of galaxies and its evolution with redshift. Using photometric redshifts derived from multi-waveband imaging surveys, measurements of the SMF of galaxies are now routinely performed out to $z \sim 5$ (e.g., Dickinson et al. 2003; Conselice et al. 2005; Drory et al. 2005; Fontana et al. 2006; Elsner et al. 2008; Pérez-González et al. 2008; Kajisawa et al. 2009; Marchesini et al. 2009). The general consensus is that at $z > 1$ the stellar mass assembly proceeds much more quickly than at lower redshifts. In particular, very

⁹ Visiting Astronomer, Kitt Peak National Observatory, National Optical Astronomy Observatory, which is operated by the Association of Universities for Research in Astronomy (AURA) under cooperative agreement with the National Science Foundation.

recent measurements at $z < 4$ show a dramatic evolution of the SMF of galaxies with redshift as well as evidence of mass-dependent evolution of the SMF, with the low-mass end evolving more rapidly than the high-mass end (i.e., Pérez-González et al. 2008; Marchesini et al. 2009).

Measurements of the SMF have also been extended to even larger redshifts ($z \sim 4-7$; e.g., McLure et al. 2009; Stark et al. 2009), providing estimates of the stellar mass content of the universe when it was only ~ 800 Myr old. However, most of these studies only target Lyman break galaxies (LBGs), hence resulting in a potentially biased view of the universe against massive and evolved galaxies at such high redshifts. Whereas the discovery of a population of very massive and evolved galaxies at $z \gtrsim 5$ has now been claimed by several groups (e.g., Yan et al. 2006; Rodighiero et al. 2007; Wiklind et al. 2008; Mancini et al. 2009), convincing evidence for the existence of galaxies with $M_{\text{star}} > 3 \times 10^{11} M_{\odot}$ at $z > 4$ is still missing (e.g., Dunlop et al. 2007).

Closely related to this issue is the very intriguing finding that the number density of the most massive galaxies ($M_{\text{star}} > 3 \times 10^{11} M_{\odot}$) seems to evolve very little from $z \sim 4$ to $z \sim 1.5$ (Marchesini et al. 2009), suggesting that the most massive galaxies in the universe were mostly in place already at $z \sim 3.5$, and implying potentially severe disagreements with the predictions from the latest generations of semi-analytic models of galaxy formation.

However, uncertainties on the observed SMF are still large, especially at the high-mass end and at high redshifts (Marchesini et al. 2009). At $z \lesssim 3$, the SMF error budget is now almost entirely dominated by systematic uncertainties caused by the different spectral energy distribution (SED)-modeling assumptions adopted to derive stellar masses (e.g., stellar population synthesis models or initial mass function, IMF). Progress in reducing the impact of these systematic uncertainties necessarily requires better calibrations of the stellar mass estimates (e.g., through measurements of the dynamical masses from studies of their kinematics). At $3 < z < 4$, instead, the contributions of photometric redshift errors, small-number statistics, and sample variance (due to the relatively small probed volumes) are still significant, and dominate the total error budget at the high-mass end of the SMF.

In this paper, we take advantage of the high-quality data from the ultraviolet to the mid-infrared (MIR) available through the NEWFIRM Medium-Band Survey (NMBS; van Dokkum et al. 2009) to derive more accurate measurements of the high-mass end of the SMF of galaxies at $3.0 \leq z < 4.0$ by significantly reducing the impact of random uncertainties and to characterize, for the first time, the observed and rest-frame properties of a mass-selected sample of galaxies at $3.0 \leq z < 4.0$. These results are made possible by the combination of accurate photometric redshifts and well-sampled SEDs of K -selected galaxies at $z > 1.5$ delivered by the medium-band near-infrared (NIR) filters of the NMBS, as well as by its large surveyed area ($\sim 0.5 \text{ deg}^2$).

This paper is structured as follows. In Section 2, we present the mass-selected ($M_{\text{star}} > 2.5 \times 10^{11} M_{\odot}$) sample used to measure the SMF of galaxies at $3.0 \leq z < 4.0$; in Section 3 we present the observed and rest-frame properties of the galaxies in the mass-selected sample. The stellar mass function and densities of massive galaxies at $3.0 \leq z < 4.0$ are presented in Section 4, while in Section 5 the systematic effects caused by systematic uncertainties in the photometric redshift estimates are quantified. The results are summarized in Section 6. We

assume $\Omega_M = 0.3$, $\Omega_{\Lambda} = 0.7$, and $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ throughout the paper. All magnitudes are on the AB system.

2. SAMPLE SELECTION

2.1. The NEWFIRM Medium-band Survey

The sample is selected from the NMBS, a moderately wide, moderately deep near-infrared imaging survey (van Dokkum et al. 2009). The survey uses the NEWFIRM camera on the Kitt Peak 4 m telescope. The camera images a $28' \times 28'$ field with four arrays with a native pixel size of $0''.4$. We developed a custom NIR filter system for NEWFIRM, comprised of five medium bandwidth filters in the wavelength range $1-1.7 \mu\text{m}$. As shown in van Dokkum et al. (2009), these filters pinpoint the Balmer/4000 Å breaks of galaxies at $1.5 < z \lesssim 3.5$, providing accurate photometric redshifts and improved measurements of the stellar population parameters. The survey targets two $28' \times 28'$ fields: a subsection of the COSMOS field (Scoville et al. 2007), and a field containing part of the AEGIS strip (Davis et al. 2007). Field positions and other information are given in van Dokkum et al. (2009). Both fields have excellent supporting data, including extremely deep optical *ugriz* data from the CFHT Legacy Survey¹⁰ and deep *Spitzer*-IRAC and MIPS imaging (Barmby et al. 2008; Sanders 2007). *Spitzer*-IRAC and MIPS photometry have been added following the procedure described in Wuyts et al. (2007), which uses a source-fitting algorithm developed by I. Labbé et al. (2010, in preparation) especially suited for heavily confused images for which a higher resolution prior (in this case the K -band image) is available.¹¹ Reduced CFHT mosaics were kindly provided to us by the CARS team (Erben et al. 2009; Hildebrandt et al. 2009). Additionally, in the COSMOS field, we include deep Subaru images with the $B_J V_{Jr^+i^+z^+}$ broadband filters (Capak et al. 2007), Subaru images with 12 intermediate-band filters from 427 to 827 nm, and the CFHT K_S -band image (Ilbert et al. 2009). In both the COSMOS and AEGIS fields, *Galaxy Evolution Explorer* photometry in the FUV (150 nm) and NUV (225 nm) passbands were added. The NMBS adds six filters: J_1 , J_2 , J_3 , H_1 , H_2 , and K . The filter characteristics of the five medium-band filters are given in van Dokkum et al. (2009).

The NMBS is an NOAO Survey Program, with 45 nights allocated over three semesters (2008A, 2008B, 2009A). An additional 30 nights were allocated through a Yale–NOAO time trade. The full details of the reduction, source detection, and generation of the photometric catalogs are described in Whitaker et al. (2010b). In the present analysis, we use a K -selected catalog based on all the data obtained over the three semesters. The AEGIS catalog contains 17 filters and the COSMOS catalog contains 35 filters (FUV– $8 \mu\text{m}$). The images were convolved to the same point-spread function (PSF) before performing aperture photometry, so as to limit any bandpass-dependent effects. Following previous studies (Labbé et al. 2003; Quadri et al. 2007), the photometry was performed with SExtractor (Bertin & Arnouts 1996) in relatively small “color” apertures which optimize the S/N ratio. Total magnitudes in each band were determined from an aperture correction computed from the K band. The aperture correction is a combination of the ratio of the flux in SExtractor’s AUTO aperture to the flux in

¹⁰ <http://www.cfht.hawaii.edu/Science/CFHTLS/>.

¹¹ The IRAC fluxes measured in this work have been compared with the publicly available IRAC photometry over COSMOS (<http://irsa.ipac.caltech.edu/data/COSMOS/tables/scosmos/>; Ilbert et al. 2009) and AEGIS (<http://www.cfa.harvard.edu/irac/egs/>; Barmby et al. 2008). The agreement is excellent, with systematic differences of $\sim 2\%$.

Table 1
Mass-selected Sample of $3.0 \leq z < 4.0$ Galaxies

ID	r (mag)	H (mag)	K (mag)	z	$\log M_{\text{star}}$ (M_{\odot})	SFR ($M_{\odot} \text{ yr}^{-1}$)	A_V (mag)	$\log \text{Age}$ (yr)
C1-4890	26.27	24.90	22.39 ± 0.16	$3.47^{+0.34}_{-0.50}$	$11.70^{+0.17}_{-0.04}$ ($11.67^{+0.07}_{-0.24}$)	$40.7^{+188}_{-25.3}$ ($33.9^{+313}_{-20.1}$)	$1.4^{+0.6}_{-0.4}$ ($1.2^{+1.2}_{-0.4}$)	$9.1^{+0.1}_{-0.2}$ ($9.2^{+0.1}_{-0.1}$)
C1-6110	> 28.4	24.91	23.11 ± 0.16	$3.58^{+0.32}_{-0.25}$	$11.41^{+0.07}_{-0.13}$ ($11.30^{+0.08}_{-0.19}$)	$0.0^{+0.1}_{-0.0}$ ($0.0^{+0.1}_{-0.0}$)	$1.0^{+0.6}_{-0.2}$ ($0.8^{+0.4}_{-0.5}$)	$9.2^{+0.0}_{-0.3}$ ($9.2^{+0.0}_{-0.1}$)
C1-7340	28.87	24.32	23.07 ± 0.15	$3.41^{+0.36}_{-0.27}$	$11.46^{+0.07}_{-0.11}$ ($11.31^{+0.07}_{-0.11}$)	$8.3^{+11.6}_{-6.5}$ ($2.3^{+4.1}_{-2.3}$)	$1.2^{+0.4}_{-0.4}$ ($0.7^{+0.5}_{-0.2}$)	$9.2^{+0.1}_{-0.1}$ ($9.2^{+0.1}_{-0.1}$)
C1-15182	25.88	22.97	21.62 ± 0.09	$3.56^{+0.11}_{-0.11}$	$11.54^{+0.07}_{-0.05}$ ($11.45^{+0.05}_{-0.09}$)	$0.0^{+0.1}_{-0.0}$ ($0.0^{+0.1}_{-0.0}$)	$1.6^{+0.2}_{-0.4}$ ($1.3^{+0.4}_{-0.3}$)	$8.3^{+0.2}_{-0.1}$ ($8.3^{+0.2}_{-0.2}$)
C1-15367	28.49	23.86	23.07 ± 0.19	$3.73^{+0.22}_{-0.06}$	$11.71^{+0.05}_{-0.30}$ ($11.51^{+0.12}_{-0.02}$)	$14.8^{+74.3}_{-13.3}$ ($13.2^{+18.4}_{-9.2}$)	$1.4^{+1.2}_{-0.2}$ ($1.1^{+0.4}_{-0.3}$)	$9.2^{+0.0}_{-1.2}$ ($9.1^{+0.1}_{-0.1}$)
C1-18825	27.18	23.63	22.62 ± 0.12	$3.05^{+0.19}_{-0.19}$	$11.40^{+0.04}_{-0.01}$ ($11.26^{+0.10}_{-0.22}$)	$1.9^{+0.2}_{-0.1}$ ($2.0^{+4.0}_{-0.5}$)	$0.8^{+0.2}_{-0.2}$ ($0.8^{+0.5}_{-0.2}$)	$9.3^{+0.0}_{-0.1}$ ($9.2^{+0.1}_{-0.1}$)
C1-19536	25.49	22.66	21.65 ± 0.06	$3.19^{+0.07}_{-0.08}$	$11.55^{+0.03}_{-0.03}$ ($11.22^{+0.09}_{-0.03}$)	$28.2^{+2.9}_{-1.9}$ ($2.9^{+7.8}_{-0.2}$)	$1.0^{+0.2}_{-0.2}$ ($0.1^{+0.5}_{-0.1}$)	$9.1^{+0.1}_{-0.1}$ ($9.0^{+0.1}_{-0.1}$)
C1-21316	25.60	23.79	22.29 ± 0.16	$3.68^{+0.12}_{-0.11}$	$11.52^{+0.01}_{-0.61}$ ($11.39^{+0.11}_{-0.58}$)	316^{+1233}_{-240} (251^{+1527}_{-179})	$1.8^{+0.6}_{-0.2}$ ($1.6^{+0.8}_{-0.5}$)	$9.1^{+0.1}_{-1.4}$ ($9.1^{+0.1}_{-1.5}$)
C1-22857	> 28.3	24.64	23.09 ± 0.19	$3.54^{+0.20}_{-0.17}$	$11.42^{+0.02}_{-0.08}$ ($11.25^{+0.05}_{-0.04}$)	$0.0^{+0.1}_{-0.0}$ ($0.0^{+0.1}_{-0.0}$)	$0.8^{+0.2}_{-0.2}$ ($0.4^{+0.3}_{-0.3}$)	$9.2^{+0.1}_{-0.2}$ ($9.2^{+0.1}_{-0.1}$)
C1-23152	23.22	20.96	20.31 ± 0.02	$3.29^{+0.06}_{-0.06}$	$11.42^{+0.01}_{-0.01}$ ($11.37^{+0.02}_{-0.01}$)	$0.1^{+1.4}_{-0.1}$ ($0.9^{+1.3}_{-0.9}$)	$0.8^{+0.2}_{-0.2}$ ($0.7^{+0.1}_{-0.1}$)	$8.1^{+0.1}_{-0.1}$ ($8.0^{+0.1}_{-0.1}$)
A2-6835	26.19	23.61	22.28 ± 0.13	$3.07^{+0.54}_{-0.70}$	$11.48^{+0.24}_{-0.19}$ ($11.48^{+0.38}_{-0.66}$)	178^{+385}_{-172} (112^{+935}_{-112})	$2.0^{+0.2}_{-0.2}$ ($1.7^{+0.2}_{-1.6}$)	$9.3^{+0.0}_{-0.7}$ ($9.3^{+0.0}_{-0.7}$)
A2-15753	25.12	22.79	22.25 ± 0.06	$3.14^{+0.10}_{-0.09}$	$11.40^{+0.02}_{-0.09}$ ($11.09^{+0.16}_{-0.06}$)	$148^{+56.3}_{-13.0}$ ($60.3^{+180}_{-16.6}$)	$1.4^{+0.8}_{-0.2}$ ($1.0^{+0.2}_{-0.2}$)	$9.3^{+0.0}_{-0.2}$ ($8.8^{+0.5}_{-0.1}$)
A2-18070	25.35	23.04	22.33 ± 0.11	$3.08^{+0.16}_{-0.15}$	$11.44^{+0.03}_{-0.01}$ ($11.30^{+0.12}_{-0.17}$)	$166^{+7.8}_{-3.9}$ ($151^{+77.7}_{-107}$)	$1.6^{+0.6}_{-0.2}$ ($1.5^{+0.2}_{-0.2}$)	$9.3^{+0.0}_{-0.1}$ ($9.2^{+0.1}_{-0.4}$)
A2-24511	25.32	22.97	21.89 ± 0.09	$3.76^{+0.17}_{-0.12}$	$11.68^{+0.11}_{-0.10}$ ($11.38^{+0.28}_{-0.36}$)	170^{+298}_{-125} ($97.7^{+594}_{-97.7}$)	$1.4^{+0.2}_{-0.2}$ ($1.3^{+0.4}_{-0.6}$)	$8.9^{+0.3}_{-0.2}$ ($8.4^{+0.4}_{-0.7}$)

Notes. “C1” and “A2” refer to the COSMOS and AEGIS fields, respectively. The listed redshift is the adopted best-fit EAZY redshift z_{peak} . The stellar population parameters were derived using a pseudo-Kroupa (2001) IMF, Bruzual & Charlot (2003) stellar population synthesis models, and a Calzetti et al. (2000) extinction law (see Section 2.3). Quoted errors are the 1σ confidence intervals output by FAST (see Kriek et al. 2009 for a detailed description of the method adopted in FAST to estimate confidence intervals). The values in the parentheses correspond to the best-fit stellar population parameters assuming a Kroupa (2001) IMF, Maraston (2005) stellar population synthesis models, exponentially declining SFHs, and a Calzetti et al. (2000) extinction law (see Section 2.3). H -band magnitudes are derived by averaging the $H1$ - and $H2$ -band magnitudes.

the color aperture and a point-source-based correction for flux outside of the AUTO aperture, thereby enabling us to calculate total magnitudes (see, e.g., Labbé et al. 2003). The K -band completeness limit of the NMBS catalog adopted in this work is $K = 23.15$ mag. Stars are flagged based on their observed $U - J1$ and $J1 - K$ colors, where the stellar sequence cleanly separates from the bulk of galaxies in color space (see Whitaker et al. 2010b for more details). The total number of objects in the K -selected sample is 52259, 27520 of which are in the COSMOS field.

2.2. Photometric Redshifts

Photometric redshifts were determined with the EAZY code (Brammer et al. 2008), using the full FUV–8 μm SEDs (FUV– K for objects in the $\sim 50\%$ of our AEGIS field that does not have *Spitzer* coverage) and $z_{\text{max}} = 6.0$ (the maximum allowed redshift within EAZY). For this study, we have used the photometric redshift z_{peak} output by EAZY.¹² Publicly available redshifts in the COSMOS and AEGIS fields indicate that the redshift errors are very small at $\sigma_z/(1+z) < 0.02$ at $z_{\text{spec}} < 1.5$. Specifically, the photometric redshifts in COSMOS are in excellent agreement with the spectroscopic redshifts made publicly available by the zCOSMOS survey (Lilly et al. 2007), with $\sigma_z/(1+z) = 0.008$ for 1444 objects at $z_{\text{spec}} < 1.5$. We also find excellent agreement between the photometric and spectroscopic redshifts for a larger sample of 2313 objects at $z_{\text{spec}} < 1.5$ in AEGIS from the DEEP2 survey (Davis et al. 2003), with $\sigma_z/(1+z) = 0.017$. Both fields have very few catastrophic failures, with only $3\% > 5\sigma$ outliers. Although there are very few spectroscopic redshifts of optically

faint K -selected galaxies in these fields, we note that we found a similarly small scatter ($\sigma_z/(1+z) < 0.02$) in a pilot program targeting galaxies from the Kriek et al. (2008) near-IR spectroscopic sample (see van Dokkum et al. 2009). Spectroscopic redshifts also exist for 125 LBGs at $z \sim 3$ within the AEGIS field from Steidel et al. (2003), for which we find $\sigma_z/(1+z) = 0.045$, with $10\% > 5\sigma$ outliers. From the formal EAZY errors listed in Table 1, we find typical $\sigma_z/(1+z) = 0.04$, perfectly consistent with the scatter $\sigma_z/(1+z)$ found for LBGs. We conclude that, in the regime of interest in this paper, the errors of the photometric redshifts are larger than at $z_{\text{spec}} < 1.5$, as they are dominated by random errors in the photometry.

The observed SEDs with best-fit EAZY templates overplotted are shown in Figure 1, together with the EAZY redshift probability distributions. As shown in Figure 1, the medium-band filters H_1 and H_2 allow us to identify the redshifted Balmer/4000 Å breaks within the H band, improving the accuracy of the photometric redshift estimates with respect to previous analysis with only broadband photometry.

Rest-frame colors were measured using the best-fit EAZY templates, as described in Brammer et al. (2009) and, in particular, in Whitaker et al. (2010a). Briefly, from the best-fit EAZY template, we computed the rest-frame $U - V$ colors following the method used by Wolf et al. (2003) in the COMBO-17 survey. We used the Maiz Apellániz (2006) filter definitions and used the direct template fluxes to determine $U - V$. When using closely spaced medium-band observed filters, the template fluxes are found to be more robust than interpolating between observed filters (see G. Brammer et al. 2010, in preparation).

2.3. SED Modeling

Stellar masses and other stellar population parameters were determined with Fitting and Assessment of Synthetic Templates (FAST; Kriek et al. 2009), fixing the redshift to the EAZY output (or the spectroscopic redshift when available). For consistency

¹² The default template set used in this work consists of seven templates: the six templates taken from the optimized template set of EAZY, but augmented with emission lines, and a template of a 12.5 Gyr old single stellar population. In Section 5 we consider the case of an additional template, consisting of a dust-obscured ($A_V = 3$ mag), old (1 Gyr) population.

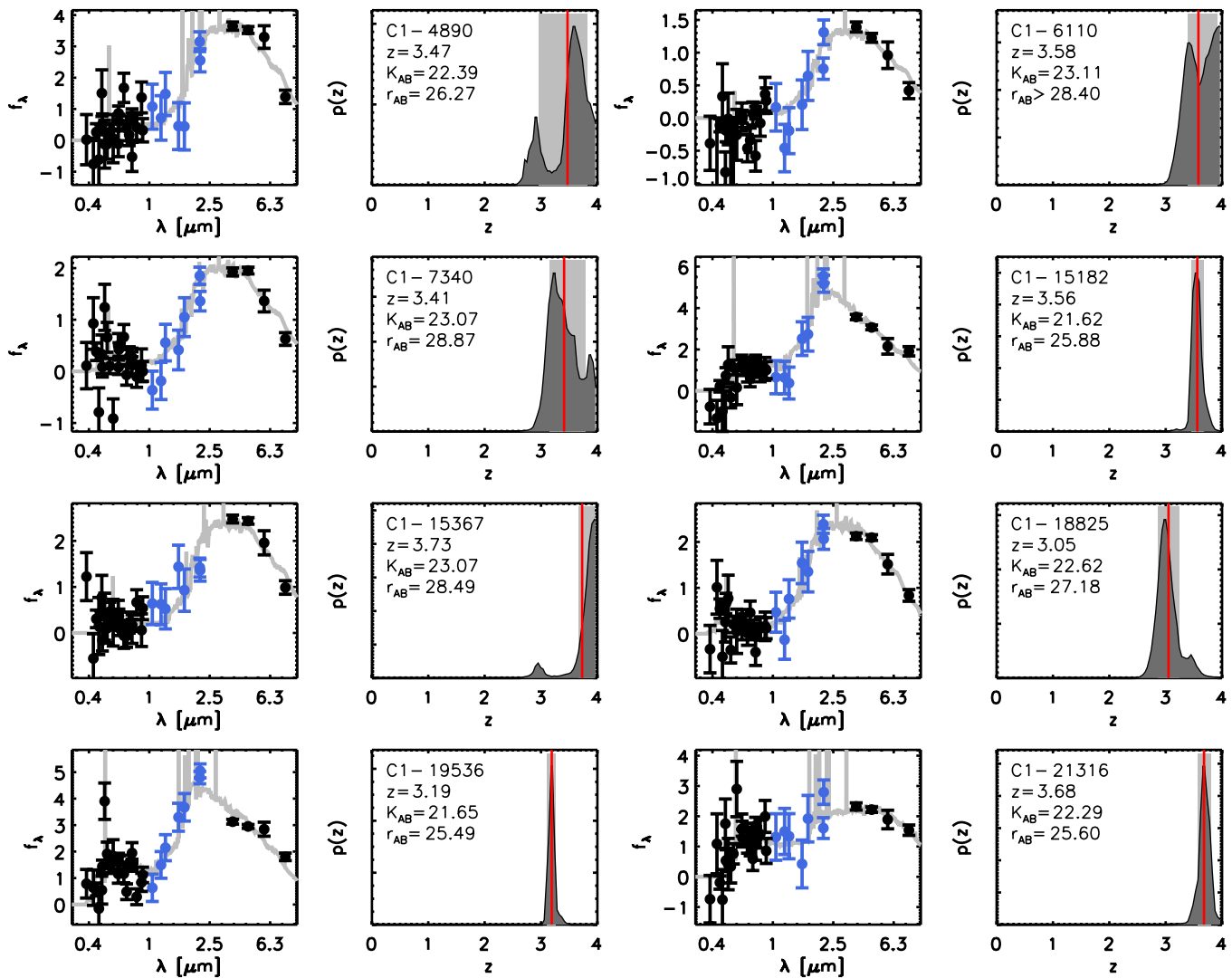


Figure 1. Observed SEDs of the mass-selected sample at $3.0 \leq z < 4.0$. Filled circles are the observed fluxes, in units of $10^{-19} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ \AA}^{-1}$, with corresponding 1σ errors. The blue symbols are the photometric points from NMBS. The solid gray curves represent the best-fit EAZY templates. The dark-gray-filled regions represent the EAZY redshift probability functions. The vertical red line is the adopted redshift from EAZY (z_{peak} , as specified in Section 2.2), while the shaded gray regions are the 1σ allowed values for the photometric redshifts. Also listed are the NMBS identification number, the adopted photometric redshift from EAZY, and the observed total K - and r -band magnitudes.

(A color version of this figure is available in the online journal.)

with previously published SMF measurements and straightforward comparisons, we assumed the default SED-modeling assumptions of Marchesini et al. (2009), i.e., stellar population synthesis models of Bruzual & Charlot (2003), the Calzetti et al. (2000) reddening law with A_V values ranging from 0 to 4 in steps of 0.2 mag, solar metallicity, pseudo-Kroupa (2001)¹³ IMF, and three star formation histories (SFHs): a single stellar population (SSP), a constant star formation history (CSF), and an exponentially declining SFH with an e -folding timescale of 300 Myr (τ_{300}). In order to quantify the systematic uncertainties due to different SED-modeling assumptions on the derived stellar population properties (i.e., M_{star} , age, SFR, and A_V) of the $3.0 \leq z < 4.0$ sample, we have also assumed the stellar population synthesis models of Maraston (2005) with a Kroupa (2001) IMF and exponentially declining SFHs with values of the

e -folding timescale ranging from 100 Myr to 10 Gyr in steps of 0.2 dex. We refer to Marchesini et al. (2009) for a detailed analysis of the systematic uncertainties on the SMF measurements due to the different SED-modeling assumptions. Figure 2 shows the observed SEDs of the mass-selected sample at $3.0 \leq z < 4.0$ together with the best-fit stellar population models from FAST for our two sets of SED-modeling assumptions.

2.4. The $3.0 \leq z < 4.0$ Mass-selected Sample

We constructed a mass-selected sample of galaxies at $3.0 \leq z < 4.0$ to study their observed and rest-frame properties, as well as to derive more accurate measurements of the high-mass end of the SMF of galaxies at $3.0 \leq z < 4.0$.

The redshift-dependent completeness limit in stellar mass has been estimated following the approach described in detail in Marchesini et al. (2009), which exploits the availability of samples with different depths. The completeness of a sample is estimated empirically from the available deeper samples, namely, the FIRES (Labbé et al. 2003; Förster Schreiber

¹³ SED modeling was performed using a Salpeter (1955) IMF with lower and upper mass cutoffs of $0.1 M_{\odot}$ and $100 M_{\odot}$, and the derived stellar masses and star formation rates (SFRs) were scaled to a pseudo-Kroupa (2001) IMF by dividing by a factor of 1.6.

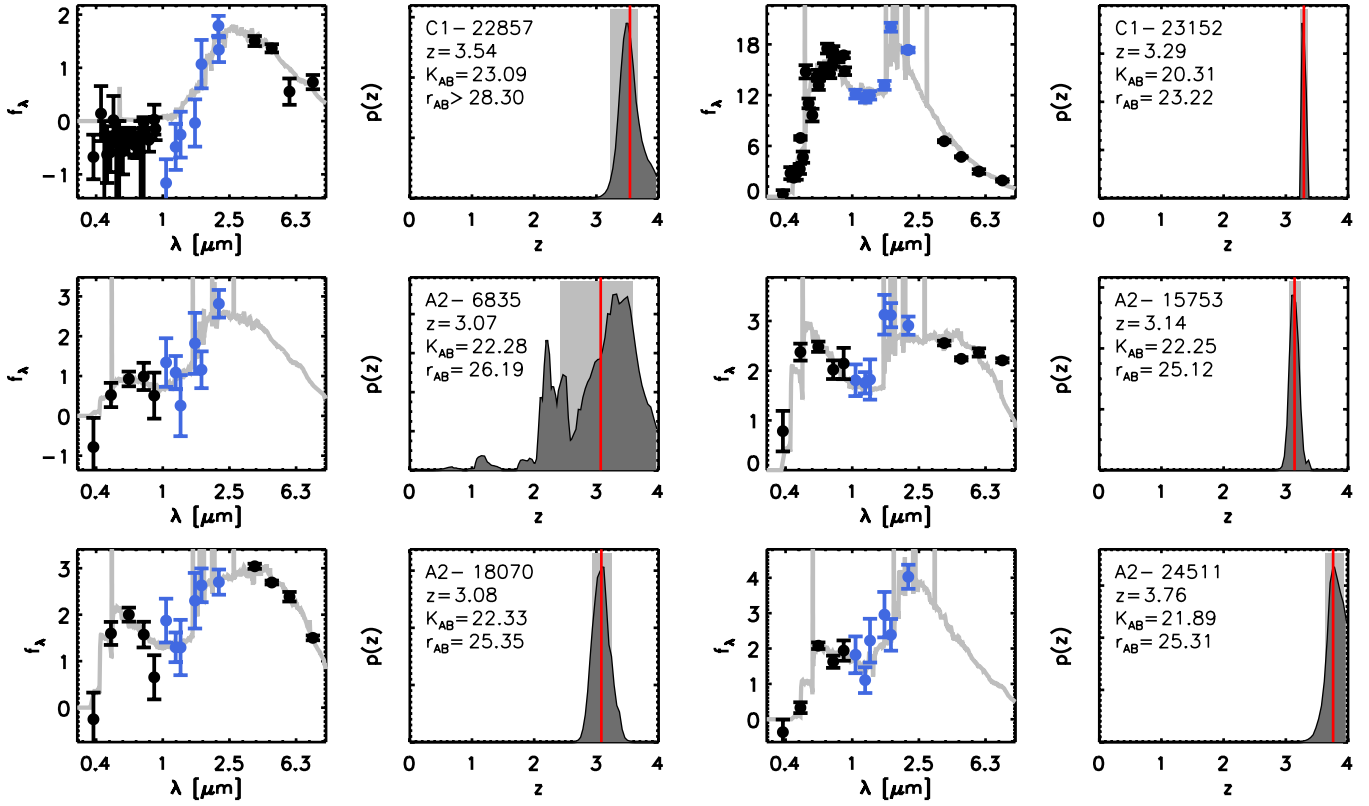


Figure 1. (Continued)

et al. 2006) and the FIREWORKS (Wuyts et al. 2008) catalogs. Briefly, to estimate the redshift-dependent stellar mass completeness limit of the NMBS sample, we first selected galaxies belonging to the available deeper samples. Then, we scaled their fluxes and stellar masses to match the K -band completeness limit of the NMBS sample. The upper envelope of points in the $(M_{\text{star, scaled}} - z)$ space, encompassing 95% of the points, represents the most massive galaxies at the flux limit considered ($K = 23.15$ for the NMBS catalog adopted in this work), and so provides a redshift-dependent stellar mass completeness limit for the NMBS sample. We refer to Marchesini et al. (2009) for a detailed description of this method. The resulting completeness in mass of the NMBS catalog used in this work is $M_{\text{star}} = 10^{11.40} M_{\odot} \approx 2.5 \times 10^{11} M_{\odot}$ over the targeted redshift range $3.0 \leq z < 4.0$.

The resulting mass-selected sample of galaxies at $3.0 \leq z < 4.0$ contains 14 sources with $M_{\text{star}} \geq 10^{11.40} M_{\odot}$ (10 from the COSMOS field and four from the AEGIS field) over an effective area of 0.44 deg^2 . The sample is listed in Table 1, along with the observed r -, H -, and K -band total magnitudes, adopted EAZY best-fit redshifts and 1σ errors, and FAST best-fit M_{star} , SFR, A_V , and age with corresponding 1σ errors. Table 2 lists the coordinates of 14 sources.

As shown in Table 1, the typical random error on the estimated stellar masses of the mass-selected sample is ~ 0.1 dex for the default set of SED-modeling assumptions (which uses the Bruzual & Charlot 2003 models), and ~ 0.16 dex for the other set (which adopts the Maraston 2005 models). These errors are in good agreement with the errors on stellar mass due to photometric redshift uncertainties estimated by Taylor et al. (2009), with a typical error on the stellar mass of ~ 0.1 dex for photometric redshift errors of $\sigma_z/(1+z) = 0.035$ at $z < 1.5$. As shown by Taylor et al. (2009), in a photometric

Table 2
Coordinates of the Massive Galaxies at $3.0 \leq z < 4.0$

ID	α (J2000)	δ (J2000)
C1-4890	$10^{\text{h}}00^{\text{m}}01^{\text{s}}55$	$+02^{\circ}15'00''.3$
C1-6110	$10^{\text{h}}00^{\text{m}}02^{\text{s}}50$	$+02^{\circ}16'22''.3$
C1-7340	$10^{\text{h}}00^{\text{m}}13^{\text{s}}69$	$+02^{\circ}17'34''.5$
C1-15182	$09^{\text{h}}59^{\text{m}}24^{\text{s}}39$	$+02^{\circ}25'36''.5$
C1-15367	$10^{\text{h}}00^{\text{m}}33^{\text{s}}42$	$+02^{\circ}25'55''.1$
C1-18825	$10^{\text{h}}00^{\text{m}}11^{\text{s}}83$	$+02^{\circ}29'35''.5$
C1-19536	$09^{\text{h}}59^{\text{m}}31^{\text{s}}82$	$+02^{\circ}30'18''.2$
C1-21316	$10^{\text{h}}00^{\text{m}}19^{\text{s}}74$	$+02^{\circ}32'04''.3$
C1-22857	$10^{\text{h}}00^{\text{m}}27^{\text{s}}72$	$+02^{\circ}33'45''.4$
C1-23152	$10^{\text{h}}00^{\text{m}}27^{\text{s}}81$	$+02^{\circ}33'49''.3$
A2-6835	$14^{\text{h}}18^{\text{m}}44^{\text{s}}82$	$+52^{\circ}30'06''.8$
A2-15753	$14^{\text{h}}18^{\text{m}}30^{\text{s}}83$	$+52^{\circ}40'24''.6$
A2-18070	$14^{\text{h}}18^{\text{m}}33^{\text{s}}58$	$+52^{\circ}42'57''.6$
A2-24511	$14^{\text{h}}17^{\text{m}}03^{\text{s}}34$	$+52^{\circ}49'44''.9$

redshift survey, the stellar mass estimates are relatively robust to random photometric redshift errors as a result of the similar (but opposite) systematic effects on luminosities and stellar mass-to-light ratios caused by random photometric redshift errors.

Figure 3 shows the images of the mass-selected sample at $3.0 \leq z < 4.0$ in the different filters, from the u band to the $24 \mu\text{m}$ *Spitzer*-MIPS channel.

In order to exclude contamination of the mass-selected sample due to blending, we have also inspected the higher spatial resolution images available over COSMOS and AEGIS. For COSMOS, we have used the CFHT-WIRCAM K_s -band image (FWHM $\sim 0''.8$) and the *Hubble Space Telescope*-Advanced Camera for Surveys (*HST*-ACS) I_{F814W} -band images.¹⁴ For

¹⁴ Available at <http://irsa.ipac.caltech.edu/data/COSMOS/datasets.html>.

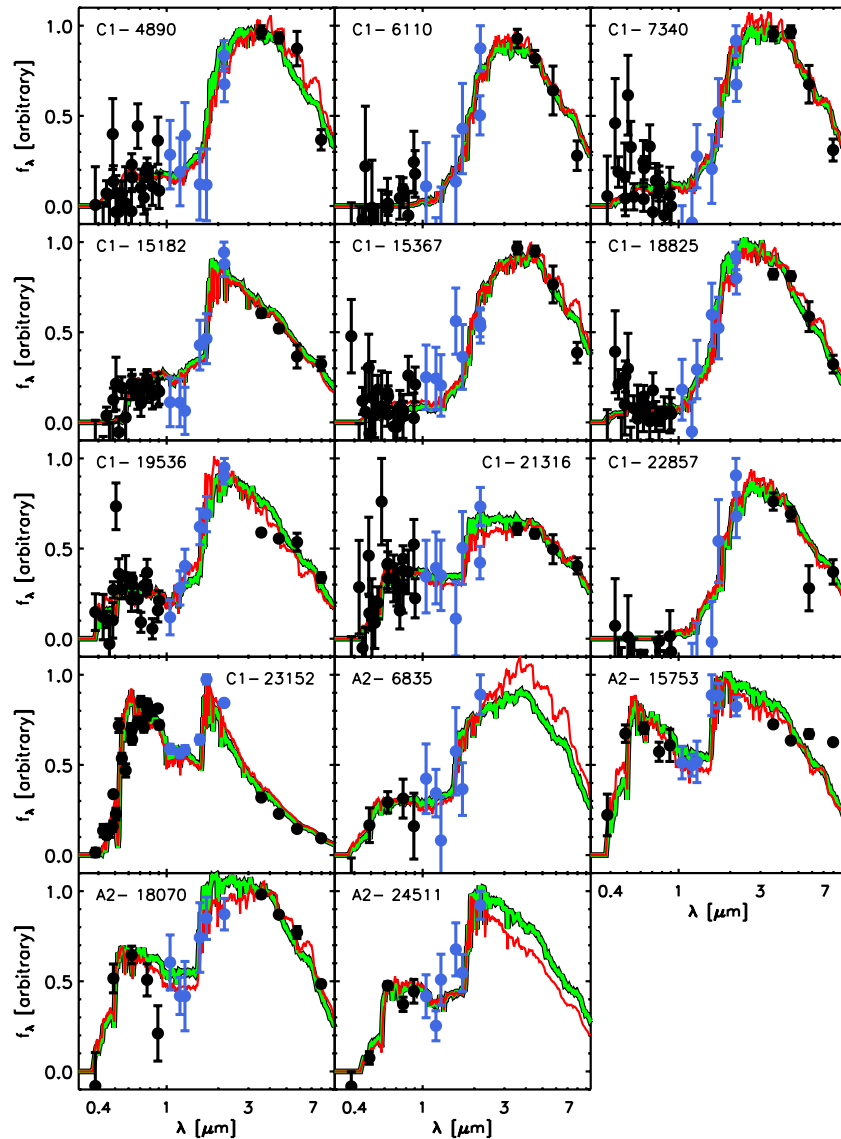


Figure 2. Observed SEDs of the mass-selected sample of galaxies at $3.0 \leq z < 4.0$. Filled circles are the observed fluxes in arbitrary units, with corresponding 1σ errors. The blue symbols are the photometric points from NMBS. The solid green curves represent the best-fit FAST templates using the stellar population synthesis models of Bruzual & Charlot (2003). The solid red curves represent the best-fit FAST templates using the stellar population synthesis models of Maraston (2005). Both stellar population synthesis models provide generally good fits to the observed SEDs.

(A color version of this figure is available in the online journal.)

AEGIS, we have used the *HST*-ACS I_{F814W} -band images.¹⁵ In the *HST* images, only C1-19536, C1-23152, and A2-15753 are detected, whereas the other sources do not show any obvious detection. All three detected sources appear to be resolved in the ACS images, indicating that the *I*-band fluxes are not dominated by a point-like component. Specifically, C1-19536 and A2-15753 are extended and quite elongated. All the sources but C1-23152 appear very isolated in the ACS images, consistent with the ground-based images. The *HST*-ACS I_{F814W} -band image of C1-23152 reveals two fainter knots at a distance of $\sim 1''.1$. The two knots contribute about 11% to its total flux in the ACS image. Inspection of the K_S -band image over COSMOS reveals that all selected massive galaxies are single, isolated objects, including C1-23152, showing no obvious signature of the two knots. If the photometry of C1-23152 is equally affected by the two knots at all wavelengths, the shape of its SED would not

be affected, and the resulting stellar mass would be smaller by ~ 0.05 dex, not changing the results of this paper. In contrast, if the contribution of the two knots changes as a function of wavelength, the observed SED would change accordingly, making it harder to predict how the derived stellar mass would be affected. A rough estimate of this effect was derived by re-fitting the observed SED of C1-23152 after assuming that only the optical fluxes are affected by the two knots. The resulting stellar mass is only ~ 0.03 dex smaller than that estimated with the current photometry, implying that the derived stellar mass for C1-23152 is not likely to be significantly biased by the two knots. We therefore conclude that none of the observed objects seems to be affected by blending issues, which might have resulted in systematically biased stellar mass estimates. Higher spatial-resolution NIR imaging is, however, required to confirm this.

Finally, we note that no a priori exclusion of active galactic nuclei (AGNs) has been performed in our mass-selected sample.

¹⁵ Available at http://aegis.ucolick.org/acs_data_descrip.html.

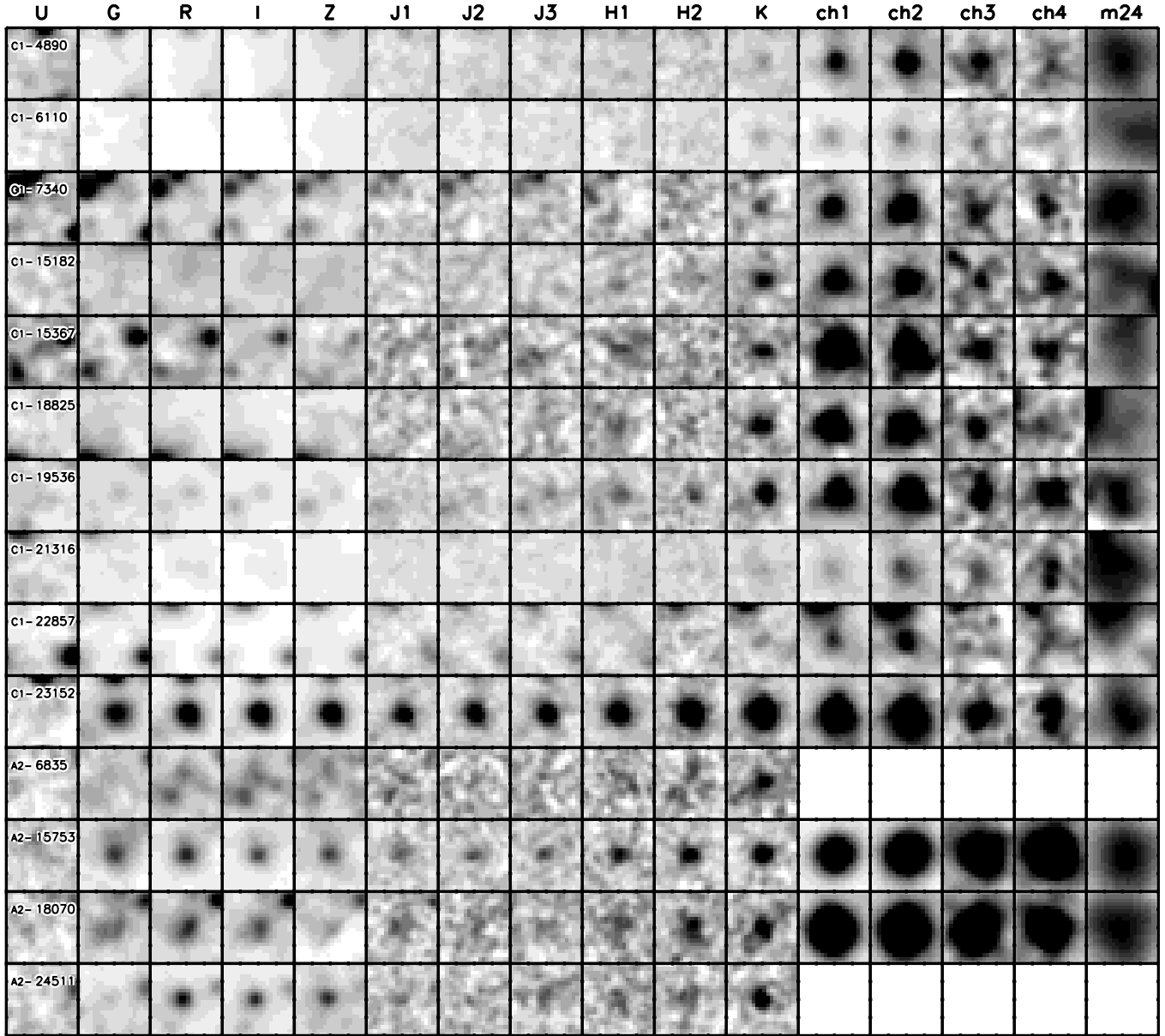


Figure 3. Images of the mass-selected sample of galaxies at $3.0 \leq z < 4.0$. From left to right, the columns show the CFHTLS u , g , r , i , and z images, the NMBS $J1$, $J2$, $J3$, $H1$, $H2$, and K images, the *Spitzer*-IRAC 3.6 μm , 4.5 μm , 5.8 μm , and 8.0 μm images, and the *Spitzer*-MIPS 24 μm image. Each cutout is $6'' \times 6''$ on a side.

3. PROPERTIES OF VERY MASSIVE GALAXIES AT $3.0 \leq z < 4.0$

We use our mass-selected sample of 14 galaxies to determine the median and dispersion in observed and rest-frame properties of the most massive galaxies ($M_{\text{star}} \gtrsim 2.5 \times 10^{11} M_{\odot}$) at $3.0 \leq z < 4.0$. Table 3 lists the median and 25th/75th percentiles of the distributions of observed r -band magnitude and $H - K$ color, the rest-frame V -band magnitude and $U - V$ color, and rest-frame UV slopes, parameterized by $F_{\lambda} \propto \lambda^{\beta}$. Rest-frame UV slopes β were determined from the best-fitting SEDs, following the Calzetti et al. (1994) method of fitting to the 10 rest-frame UV bins defined by those authors.

Figure 4 shows the distributions of the observed $H - K$ color (top panel), rest-frame $U - V$ color (middle panel), and rest-frame UV slopes of the mass-selected sample (bottom panel). For comparison, we have also plotted (1) the distribution of rest-frame $U - V$ colors and UV slopes of a mass-selected sample of galaxies at $2 < z < 3$ with $M_{\text{star}} > 6 \times 10^{10} M_{\odot}$ from van

Table 3
Observed and Rest-frame Properties of the $3.0 \leq z < 4.0$
Mass-selected Sample

Quantity	25%	Median	75%
r_{tot} (obs)	25.3	26.1	27.8
$H - K$ (obs)	0.75	1.16	1.42
V_{tot} (rest)	-24.2	-23.5	-23.3
$U - V$ (rest)	1.26	1.64	1.90
β (rest)	-0.56	-0.36	0.07

Dokkum et al. (2006, orange histogram), (2) the distribution of rest-frame $U - V$ colors and UV slopes of the galaxies that would be selected as LBGs from the sample of van Dokkum et al. (2006, purple histogram), and (3) the distribution of rest-frame UV slopes of a $z \sim 3.7$ sample of galaxies from Brammer & van Dokkum (2007) selected with the color criterion $H - K > 0.9$ to have prominent Balmer/4000 Å breaks between the H and K bands (cyan histogram).

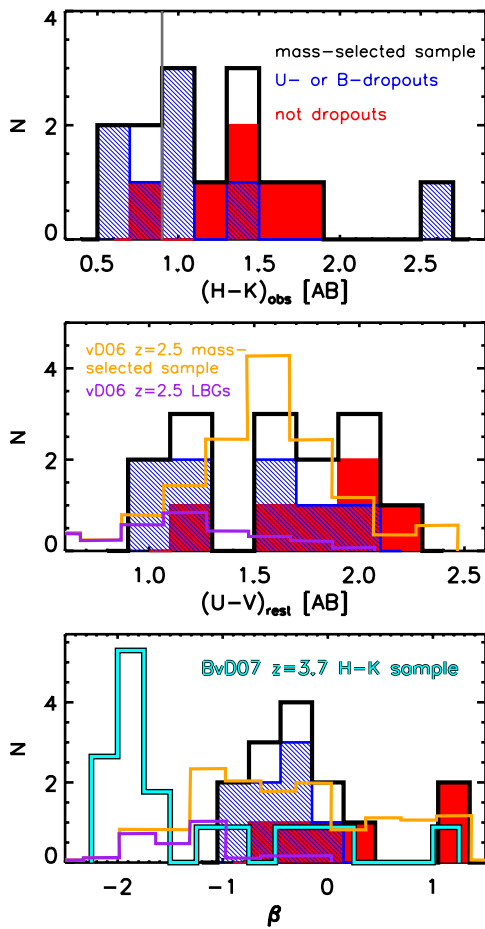


Figure 4. Top panel: the distribution of observed $H-K$ colors of the mass-selected sample of galaxies at $3.0 \leq z < 4.0$ (solid black line); the blue hatched area represents the distribution for those galaxies that would be selected as either U - or B -dropout galaxies based on their observed optical colors; the red filled area represents the distribution for those galaxies that would not be selected as either U - or B -dropout galaxies based on their observed optical colors; the vertical gray line represents the $H-K$ criterion adopted in Brammer & van Dokkum (2007) to select galaxies at $z \sim 3.7$. Middle panel: the distribution of rest-frame $U-V$ colors of the mass-selected sample; colors as in the top panel; the orange solid line represents the distribution of rest-frame $U-V$ colors of the mass-selected sample at $2 < z < 3$ from van Dokkum et al. (2006); the purple solid line represents the distribution of rest-frame $U-V$ colors of the galaxies in the mass-selected sample at $2 < z < 3$ from van Dokkum et al. (2006) that would be selected as LBGs. Bottom panel: the distribution of rest-frame UV slopes of the mass-selected sample; colors as in the middle panel; the solid cyan line represents the distribution of UV slopes of the $H-K$ -selected sample at $z \sim 3.7$ of Brammer & van Dokkum (2007).

(A color version of this figure is available in the online journal.)

The typical very massive galaxy at $3.0 \leq z < 4.0$ (median stellar mass $\langle M_{\text{star}} \rangle \sim 3 \times 10^{11} M_{\odot}$) is red and faint in the observer's optical, with $\langle r_{\text{tot}} \rangle = 26.1$. Most galaxies (10 out of 14) would be selected as $H-K$ galaxies with $H-K > 0.9$ (Brammer & van Dokkum 2007). Of the four galaxies with $H-K < 0.9$, only two galaxies have $H-K$ color significantly smaller than the $H-K$ criterion. This highlights the efficiency of this color technique in selecting galaxies at $z > 3$ with prominent breaks in the rest-frame optical, although the fraction of interlopers selected by this color technique remains uncertain.

From Table 1, 40%–60% of the very massive galaxies is characterized by ages consistent with the age of the universe at the targeted redshifts (~ 1.6 – 2.1×10^9 yr). About 30% of the very massive galaxies, namely C1-6110, C1-15182, C1-22857, and C1-23152, have SFR estimates from SED modeling consistent

with no star formation activity to within 1σ , independent of the specific SED-modeling assumptions adopted in FAST. Of the remaining galaxies, four have very large SFRs, of the order of a few hundreds solar masses per year. We stress that the estimated ages and SFRs from SED modeling are quite uncertain, even with the high-quality data set used in this work (e.g., Muzzin et al. 2009).

3.1. Rest-frame UV

The rest-frame $U-V$ colors range from $U-V = 1.01$, typical of nearby irregular galaxies, to $U-V = 2.2$, typical of local elliptical galaxies (e.g., Fukugita et al. 1995). The median $\langle U-V \rangle = 1.64$ is similar to local Sb spiral galaxies. As shown in the middle panel of Figure 4, the mass-selected sample of van Dokkum et al. (2006) at $z = 2.5$, which is complete in stellar mass down to $\sim 6 \times 10^{10} M_{\odot}$ (a factor of ~ 5 less than our sample), encompasses the range in $U-V$ colors spanned by our $z = 3.5$ sample, with a median $U-V$ color bluer by ~ 0.1 mag with respect to our mass-selected sample.

The median UV slope β is $\langle \beta \rangle = -0.36$, indicating a relatively flat spectrum in F_{λ} . The distribution of β , ranging from $\beta = -0.95$ to $\beta = 1.10$, is broadly consistent with the distribution of massive galaxies at $2 < z < 3$ from van Dokkum et al. (2006). As shown in the bottom panel of Figure 4, the distribution of β is instead very different from the distribution seen for $H-K > 0.9$ galaxies at $z \sim 3.7$, which shows a peak at $\beta \sim -2$ (Brammer & van Dokkum 2007). The observed distribution of β for the mass-selected sample at $3.0 \leq z < 4.0$ is also very different from the distributions seen for UV-selected galaxies at $z \sim 2.5$ and ~ 4 , which peak at $\beta \sim -1.6$ and ~ -1.8 , respectively (e.g., Adelberger & Steidel 2000; Ouchi et al. 2004; Hathi et al. 2008; Bouwens et al. 2009).

The intrinsically different rest-frame UV properties of the mass-selected sample at $3.0 \leq z < 4.0$ studied in this work and the typical UV-selected galaxies at these redshifts (i.e., U - and B -dropout galaxies) are also clear from Figure 5, which shows the location of the massive galaxies at $3.0 \leq z < 4.0$ in the U_nGR and $B_{435}V_{606z_{850}}$ diagrams commonly used to select U -dropout galaxies (i.e., LBGs; Steidel et al. 2003) and B -dropout galaxies (Giavalisco et al. 2004; Bouwens et al. 2009), respectively. The colors plotted in Figure 5 are synthetic colors integrated from the best-fit FAST templates. About 57% of the galaxies in the $3.0 \leq z < 4.0$ mass-selected sample have colors that satisfy either the U - or the B -dropout color criteria (gray shaded area in Figure 5). Of these, three would be selected as U -dropouts, and five as B -dropouts, based on their observed optical colors. However, $\sim 50\%$ of these UV-selected galaxies is fainter than $r_{\text{tot}} = 25.5$, which is the observed optical limit of typical spectroscopic samples of LBGs. While the $r_{\text{tot}} = 25.5$ cut is not relevant to the inclusion of our galaxies in the photometric window, it is relevant when considering our objects in the context of pre-existing LBG surveys.

The rest-frame SEDs of the mass-selected sample at $3.0 \leq z < 4.0$ are shown in Figure 6, together with the median rest-frame SED from the data (solid blue curve) and the median best-fit templates from FAST (green and red solid curves). Figure 6 clearly shows the strongly suppressed emission and the flatness of the spectrum in F_{λ} in the rest-frame UV, as well as the strong Balmer/4000 Å breaks in the rest-frame optical for the typical very massive galaxy at $z \sim 3.5$. Also plotted is the median rest-frame SED of the $H-K$ -selected sample at $z \sim 3.7$ from Brammer & van Dokkum (2007). Figure 6 clearly shows that the rest-frame optical SEDs of the $H-K$ -selected

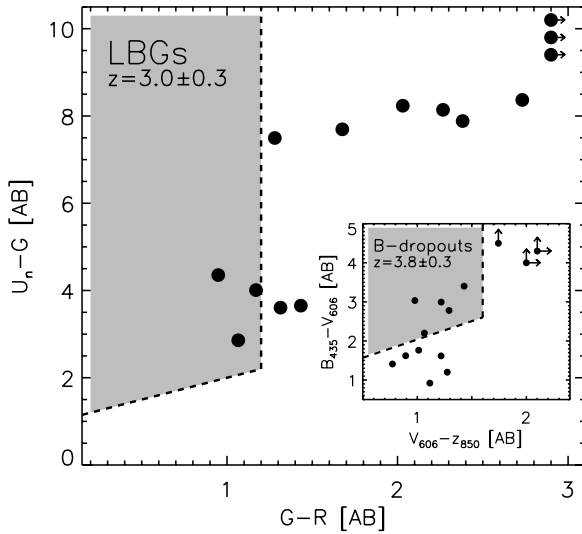


Figure 5. Color selection of LBGs at $z \sim 3.0 \pm 0.3$ (Steidel et al. 2003) and, in the inset, of B -dropout galaxies at $z \sim 3.8 \pm 0.3$ (Bouwens et al. 2009). Objects falling in the gray-shaded regions would be selected as LBGs or B -dropout galaxies. Out of the 14 very massive galaxies at $3.0 \leq z < 4.0$, three would be selected as LBGs and five as B -dropout galaxies based on their colors. However, only four out of eight of the dropout galaxies have $r \leq 25.5$.

sample and our mass-selected sample are very similar, characterized by strong rest-frame optical breaks. In contrast, their rest-frame UV SEDs are very different. The $H-K$ -selected galaxies are characterized by very blue rest-frame UV-optical colors. In contrast, our mass-selected galaxies are generally red also in the rest-frame UV.

The significant differences in observed and rest-frame properties between the $H-K$ -selected galaxies at $z \sim 3.7$ from Brammer & van Dokkum (2007) and the very massive galaxies at $3.0 \leq z < 4.0$ selected in our study are very interesting, as most of our galaxies would also be selected as $H-K$ galaxies. The simplest explanation for the observed differences is the very different regime in stellar masses probed by the two samples. The $H-K$ galaxies from Brammer & van Dokkum (2007) were selected from the FIRES survey over an area of $\sim 31 \text{ arcmin}^2$, with a median stellar mass a factor of ~ 15 smaller than the median stellar mass of our mass-selected sample. The lack of very massive galaxies in FIRES is simply caused by its very small surveyed area, as an area of $> 100 \text{ arcmin}^2$ is required to find one single object as massive as our mass-selected galaxies. The different stellar mass regimes probed by the two samples suggest that lower mass galaxies are characterized by much bluer rest-frame UV-optical colors than the most massive galaxies at these redshifts. This is also supported by the different rest-frame UV properties between our mass-selected sample and typical UV-selected B - and V -dropout samples, which have stellar masses in the range $10^9\text{--}10^{11} M_\odot$ (e.g., Shapley et al. 2001; Magdis et al. 2010). From the SED-modeling analysis, it appears that the differences in the rest-frame UV properties between the $H-K$ -selected galaxies and our mass-selected sample could be primarily due to the different amounts of dust in the two samples, as their typical stellar ages ($\sim 1\text{--}1.5 \text{ Gyr}$) are broadly consistent within the uncertainties. More specifically, the $H-K$ -selected galaxies from Brammer & van Dokkum (2007) are characterized by a median dust extinction of $\langle A_V \rangle \sim 0.2$, a factor of ~ 7 smaller than the median extinction of our mass-selected sample ($\langle A_V \rangle = 1.4 \text{ mag}$). Indeed, a significant amount of dust seems to be quite ubiquitous in massive galaxies at $3.0 \leq z < 4.0$.

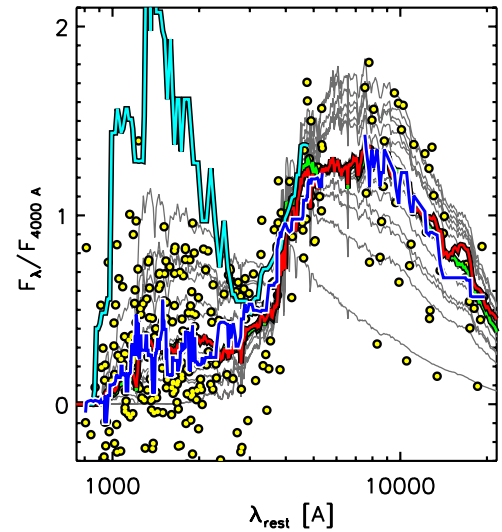


Figure 6. Rest-frame SEDs of the mass-selected sample at $3.0 \leq z < 4.0$ (yellow-filled circles). The SEDs are normalized to the flux at $\lambda_{\text{rest}} = 4000 \text{ \AA}$. The solid blue curve represents the running median of the 15 neighboring points. The best-fit FAST templates adopting the Bruzual & Charlot (2003) models are shown as dark gray solid curves. The green and red solid curves represent the median best-fit FAST templates adopting the Bruzual & Charlot (2003) and the Maraston (2005) models, respectively. The cyan solid curve represents the median SED of the $H-K$ galaxies at $z \sim 3.7$ from Brammer & van Dokkum (2007).

(A color version of this figure is available in the online journal.)

The amount of dust in our mass-selected sample is significantly higher than the amount of dust in local massive galaxies and in massive galaxies at $z \sim 1\text{--}2$ ($A_V \approx 0.2\text{--}0.3 \text{ mag}$, e.g., Whitaker et al. 2010a), further contributing to suppressing the rest-frame UV light in these massive galaxies.¹⁶

3.2. Spitzer-MIPS 24 μm Data

We have used the publicly available observations at $24 \mu\text{m}$ from *Spitzer*-MIPS provided by S-COSMOS and FIDEL¹⁷ to further constrain the activity in the most massive galaxies at $3.0 \leq z < 4.0$. The measured $24 \mu\text{m}$ fluxes with the corresponding 1σ errors are listed in Table 4. For two galaxies over AEGIS, namely A2-6835 and A2-24511, no MIPS $24 \mu\text{m}$ data are available. The $24 \mu\text{m}$ flux of C1-18825 is not reliable due to blending issues from a nearby very bright $24 \mu\text{m}$ source. The MIPS cutouts are shown in Figure 3.

In the sample of 11 massive galaxies at $3.0 \leq z < 4.0$ with MIPS coverage and no blending issues, $\sim 80\%$ has a MIPS $24 \mu\text{m}$ flux significant at $> 3\sigma$. This is broadly consistent with the high fraction of MIPS-detected sources in the sample of IRAC-selected massive ($M_{\text{star}} \sim 10^{10}\text{--}10^{11} M_\odot$) galaxies at $z > 3.5$ over GOODS-North (Mancini et al. 2009). The fraction of MIPS-detected massive galaxies in our sample increases up to $\sim 90\%$ for a $> 1\sigma$ detection. Only C1-22857 is undetected at $24 \mu\text{m}$.

For the redshift range targeted here, the $24 \mu\text{m}$ band probes the rest-frame wavelengths from $\sim 4.8 \mu\text{m}$ to $\sim 7.1 \mu\text{m}$, which includes the $5.27 \mu\text{m}$, $5.7 \mu\text{m}$, and $6.2 \mu\text{m}$ emission features

¹⁶ We note that performing the SED-modeling forcing $A_V = 0$ results in stellar masses smaller by $0\text{--}0.2 \text{ dex}$. However, the resulting χ^2 of the best-fit models are significantly worse than the χ^2 values corresponding to the modeling allowing for dust. Moreover, the MIPS detections strongly suggest the presence of a significant amount of dust. Therefore, we conclude that the dust-free assumption is not a realistic assumption.

¹⁷ <http://irsa.ipac.caltech.edu/data/SPITZER/FIDEL/>.

Table 4
Spitzer-MIPS 24 μ m Fluxes and Derived Properties of the $3.0 \leq z < 4.0$
 Mass-selected Sample

ID	S_{24} (μ Jy)	L_{IR} ($10^{13} L_{\odot}$)	SFR ($M_{\odot} \text{ yr}^{-1}$)
C1-4890	206.2 ± 32.8	$3.3 \pm 0.5^{(+2.1}_{-2.3)}$	$3611 \pm 574^{(+2280}_{-2500)}$
C1-6110	116.4 ± 28.0	$2.3 \pm 0.5^{(+1.1}_{-0.9)}$	$2458 \pm 591^{(+1142}_{-1030)}$
C1-7340	133.0 ± 28.9	$1.9 \pm 0.4^{(+1.5}_{-0.9)}$	$2078 \pm 451^{(+1682}_{-1025)}$
C1-15182	78.6 ± 25.7	$1.5 \pm 0.5^{(+0.6}_{-0.6)}$	$1614 \pm 528^{(+694}_{-663)}$
C1-15367	167.3 ± 30.8	$3.9 \pm 0.7^{(+1.2}_{-0.8)}$	$4275 \pm 787^{(+1226}_{-909)}$
C1-18825 ^a	blended
C1-19536	61.6 ± 24.6	$0.5 \pm 0.2^{(+0.3}_{-0.2)}$	$584 \pm 233^{(+300}_{-282)}$
C1-21316	177.1 ± 31.3	$4.0 \pm 0.7^{(+1.0}_{-1.0)}$	$4330 \pm 765^{(+1064}_{-1095)}$
C1-22857	< 19.9	< 0.4	< 395
C1-23152	110.5 ± 27.6	$1.2 \pm 0.3^{(+0.4}_{-0.4)}$	$1331 \pm 333^{(+424}_{-408)}$
A2-6835 ^b
A2-15753	165.7 ± 22.8	$1.3 \pm 0.3^{(+0.5}_{-0.5)}$	$1371 \pm 378^{(+557}_{-476)}$
A2-18070	127.0 ± 21.3	$0.9 \pm 0.3^{(+0.5}_{-0.4)}$	$918 \pm 308^{(+582}_{-420)}$
A2-24511 ^b

Notes.

^a No reliable MIPS 24 μ m flux could be obtained for C1-18825 due to blending issues.

^b No MIPS 24 μ m data available for A2-6835 and A2-23152. The errors listed for L_{IR} and SFR are computed using just the 24 μ m photometric errors (values not in parentheses) and the combination of the 24 μ m photometric errors and the photometric redshift errors (values in parentheses).

from polycyclic aromatic hydrocarbons (PAHs; Draine & Li 2007). MIR emission at these wavelengths could arise from warm/hot dust and PAH molecules, heated by either dust-enshrouded star formation or AGNs. The 24 μ m emission is widely used to estimate the SFRs in high-redshift galaxies (e.g., Rigby et al. 2008; Papovich et al. 2007). The dust-enshrouded SFRs can be estimated by transforming total infrared luminosities ($L_{\text{IR}} \equiv L(8\text{--}1000 \mu\text{m})$) into SFRs (Kennicutt 1998).

To convert the 24 μ m emission to a total IR luminosity we followed the approach presented in Wuyts et al. (2008). Specifically, we used the infrared SEDs of star-forming galaxies provided by Dale & Helou (2002). The template set allows us to quantify the IR/MIR flux ratio for different heating levels of the interstellar environment, parameterized by $dM(U) \sim U^{-\alpha} dU$, where $M(U)$ represents the dust mass heated by an intensity U of the interstellar field. We computed the total infrared luminosity $L_{\text{IR},\alpha}$ for each object for all Dale & Helou (2002) templates within the reasonable range from $\alpha = 1$ for active galaxies to $\alpha = 2.5$ for quiescent galaxies. The mean of the resulting $\log L_{\text{IR},\alpha=1,\dots,2.5}$ was adopted as a best estimate for the IR luminosity. Table 4 lists the estimated total IR luminosities, L_{IR} , with the corresponding 1σ errors, with and without the uncertainties due to random photometric redshift errors. The approach adopted to estimate L_{IR} from 24 μ m fluxes has been shown to produce SFRs that are in better agreement with the SFRs determined from SED modeling (Franx et al. 2008; Wuyts et al. 2008) and from dust-corrected H α line fluxes (Muzzin et al. 2010), with respect to the often used local luminosity-dependent approaches, which can systematically overestimate SFRs by a factor of 4–6 (Papovich et al. 2007; Murphy et al. 2009; Muzzin et al. 2010). More importantly, our approach adopted to estimate L_{IR} from 24 μ m fluxes is further supported by the detailed analysis of the far-IR SED (from 24 μ m to 870 μ m) performed

in Muzzin et al. (2010) on two ultra-luminous infrared galaxies (ULIRGs) at $z \sim 2$.

We convert the L_{IR} to SFRs using the Kennicutt (1998) relation, $\text{SFR}(L_{\text{IR}}) = 0.63 \cdot L_{\text{IR}}/5.8 \times 10^9 L_{\odot}$, where the factor of 0.63 is to convert to a pseudo-Kroupa (2001) IMF. The estimated SFRs are listed in Table 4, along with the 1σ errors, with and without the uncertainties due to random photometric redshift uncertainties. As shown in Table 4, the random uncertainties on L_{IR} and the SFRs are dominated by the contribution from random photometric redshift errors. Systematic template uncertainty (not included in the errors in Table 4) can contribute additionally to the total error budget, with ± 0.45 dex variation from $\log L_{\text{IR},\alpha=2.5}$ to $\log L_{\text{IR},\alpha=1}$ (Wuyts et al. 2008).

The estimated total IR luminosities of the MIPS-detected sources range from $\sim 5.0 \times 10^{12} L_{\odot}$ to $\sim 4.0 \times 10^{13} L_{\odot}$, with 80% of them having $L_{\text{IR}} \geq 10^{13} L_{\odot}$, typical of hyper-luminous infrared galaxies (HLIRGs), and the remaining being ULIRGs. The fraction of HLIRGs in the most massive galaxies at $3.0 \leq z < 4.0$ is larger by a factor ~ 10 with respect to the fraction of HLIRGs in the Kriek et al. (2008) sample of massive ($M_{\text{star}} \approx 10^{11} M_{\odot}$) galaxies at $2.0 < z < 2.7$ with spectroscopic redshifts (Muzzin et al. 2010). Whereas the sample at $2.0 < z < 2.7$ is less massive than our mass-selected sample at $3.0 \leq z < 4.0$ by a factor of ~ 3 , the large difference in the fraction of HLIRGs seems to suggest a large evolution in the number density of massive HLIRGs from $z = 3.5$ to $z = 2.3$.

As previously noted, the observed 24 μ m emission could arise from warm/hot dust and PAH molecules, heated by either dust-enshrouded star formation or highly obscured AGNs. If all the IR luminosity is associated with dust-enshrouded star formation, the resulting SFRs range between ~ 600 and $\sim 4300 M_{\odot} \text{ yr}^{-1}$, with the exclusion of C1-22857, for which only an upper limit was derived. These values are tens to several hundreds of times larger than the SFRs estimated from SED modeling. On average, the SFRs estimated from the 24 μ m fluxes are a factor of ~ 200 larger than the SFRs estimated from SED modeling. Moreover, three galaxies (C1-6110, C1-15182, and C1-23152) have MIPS-derived SFRs of the order of $1.3\text{--}2.5 \times 10^3 M_{\odot} \text{ yr}^{-1}$, whereas the FUV-to-8 μ m observed SEDs are consistent with zero star formation. This suggests that if the 24 μ m flux is from star formation, most of it must be completely obscured by dust. The MIPS-estimated SFRs translate to specific star formation rates $\text{sSFR} \approx 10^{-8.8}$ to $10^{-7.8} \text{ yr}^{-1}$, which would imply that the most massive galaxies at $z = 3.5$ are extremely actively star-forming systems that would double their stellar mass in $(0.6\text{--}7) \times 10^8 \text{ yr}$ if the derived SFRs were to be sustained at the current levels. However, little evolution seems to have been found in the number density of the most massive galaxies from $z = 3.5$ to 1.6 (Marchesini et al. 2009), which would imply a growth in stellar mass in the most massive galaxies over this redshift range by $\sim 30\%$ (>3 times smaller than the implied growth from $z = 3.5$ to 2.5 from the MIPS-derived SFRs), although larger evolution would be allowed once systematic uncertainties are taken into account. Therefore, either the very large star-forming activity indicated by the observed 24 μ m emission has to be very quickly quenched in the majority of the most massive galaxies at $3.0 \leq z < 4.0$, or the MIPS-derived SFRs are systematically biased by, e.g., contamination from AGN activity.

Indeed, in the local universe, an AGN is thought to be the dominant source of radiation responsible for the far- and mid-IR SEDs of galaxies with $L_{\text{IR}} \sim 10^{13} L_{\odot}$ (e.g., Genzel et al.

1998; Lutz et al. 1998). If the MIR emission is dominated by AGN-heated dust, the large fraction of very massive galaxies at $3.0 \leq z < 4.0$ with MIPS detection suggests that AGNs are very common ($\gtrsim 80\%$) in the most massive galaxies at these redshifts. While the fraction of AGNs in dropout galaxies at $z > 3$ is generally estimated to be low ($\sim 3\%$ – 7% ; Steidel et al. 2002; Laird et al. 2006; Reddy et al. 2006), estimates of the AGN fraction in massive galaxies at $3.0 \leq z < 4.0$ are still very uncertain. Common AGN activity in massive galaxies has been found at lower redshifts, with an AGN fraction of $\sim 30\%$ at $z \sim 2.5$ (Papovich et al. 2006; Kriek et al. 2007). Evidence for an increasing fraction of AGN as a function of stellar mass has also been shown by Kriek et al. (2007), with an AGN fraction that could reach as high as 70% for the most massive galaxies at $2.0 < z < 2.7$. If the observed $24 \mu\text{m}$ emission represents a signature of AGN-heated dust, then our results represent further supporting evidence for higher AGN fractions at high z and in the most massive galaxies.

With the currently available data it is not possible to discriminate between dust-enshrouded starburst or highly obscured AGNs as the dominant source responsible for heating the dust in our sample of very massive galaxies. Significant contributions from both sources to the observed MIPS $24 \mu\text{m}$ fluxes cannot be excluded, and their relative importance will certainly vary among our sample. However, if most of the massive galaxies in our sample have extreme SFRs, as derived from the MIPS data, then it is unlikely that we are witnessing short-lived bursts, as the duty cycle of the star formation has to be long to account for the observed large fraction of MIPS-detected sources. This seems to be in contradiction with the need for the extreme MIPS-derived star formation activity to be rapidly ($< 10^8 \text{ yr}$) quenched to avoid overprediction of the high-mass end of the SMF of galaxies at $z < 3$. Moreover, the estimated L_{IR} are typical of HLIRGs, for which the AGN is generally thought to be a significant (if not dominant) source of radiation responsible for the very large IR luminosities. Finally, for the targeted redshift range, the MIPS $24 \mu\text{m}$ band probes rest-frame wavelengths $\sim 5.5 \mu\text{m}$, where hot dust dominates the MIR emission, and the contribution from the AGN as the source of the radiation field heating the dust becomes increasingly more likely. For all these reasons, the very high MIPS-estimated SFRs are unlikely, and we therefore favor the AGN instead of starburst activity as the dominant source of the observed MIPS $24 \mu\text{m}$ emission.

Whatever the source of radiation responsible for heating the dust (dust-enshrouded star formation and/or AGN), the very large IR luminosities estimated in our sample of very massive galaxies at $3.0 \leq z < 4.0$ show that, despite already very large stellar masses, there is still plenty of gas and dust around either the supermassive black holes or the star-forming regions.

3.2.1. X-ray Emission

AGNs can efficiently be identified by their X-ray emission, which is thought to be due to up-scattered UV photons from the accretion disk. AGN-induced X-ray emission can be distinguished from that induced by star formation by the hardness ratio and (particularly) the luminosity. *Chandra* X-ray data are available over both the COSMOS and AEGIS fields. We have used the publicly available X-ray catalogs (Laird et al. 2009 and Elvis et al. 2009 for the AEGIS and COSMOS fields, respectively) to search for X-ray detections within our mass-selected sample at $3.0 \leq z < 4.0$.

Three sources, namely C1-15182, C1-19536, and A2-15753, are detected in the *Chandra* images, with hard-band ($2\text{--}7 \text{ keV}$) fluxes of $(3.8 \pm 1.0) \times 10^{-15}$, $(1.3 \pm 0.2) \times 10^{-14}$, and $(7.6 \pm 1.2) \times 10^{-15} \text{ erg s}^{-1} \text{ cm}^{-2}$, respectively. Assuming a power-law photon index $\Gamma = 1.9$ (Nandra & Pounds 1994), these fluxes correspond to X-ray luminosities $L_{2-7 \text{ keV}}$ of $(3.8 \pm 1.3) \times 10^{44}$, $(1.0 \pm 0.2) \times 10^{45}$, and $(5.7 \pm 1.5) \times 10^{44} \text{ erg s}^{-1}$, respectively, typical of high-luminosity AGNs ($L_{2-7 \text{ keV}} > 3 \times 10^{43} \text{ erg s}^{-1}$).¹⁸ Their hardness ratios are 0.17 ± 0.3 , -0.24 ± 0.11 , and 0.31 ± 0.13 , typical of narrow-line and high- z -obscured AGNs (Brusa et al. 2009). Using Figure 5 in Treister et al. (2009), which plots the neutral hydrogen column density (N_{H}) as a function of hardness ratio for 15 high-redshift X-ray sources, we estimated $N_{\text{H}} = (4_{-2}^{+4}) \times 10^{23}$, $(1_{-0.5}^{+1.0}) \times 10^{23}$, and $(6 \pm 2) \times 10^{23} \text{ cm}^{-2}$, characteristic of moderately obscured AGNs.

We therefore conclude that the X-ray data for the three detected sources support the presence of powerful AGNs in all three sources, providing further evidence for AGN radiation being the dominant source for heating the dust and for the large MIPS fluxes. Note, however, that only high-luminosity AGNs with $L_{2-7 \text{ keV}} > 10^{43.7-43.9} \text{ erg s}^{-1}$ could have been detected for sources at $z > 3$ given the limiting source detection depths of the X-ray data over AEGIS and COSMOS. In other words, the lack of X-ray detection does not provide information on the presence, or lack thereof, of high-luminosity AGNs in the remaining 11 sources of our mass-selected sample.

3.2.2. Continuum Emission from the AGN

In the previous section, we found that the AGN is likely the dominant source of radiation responsible for heating the dust and for the large IR luminosities. Moreover, three galaxies have X-ray detections with hardness ratios and estimated X-ray luminosities typical of obscured high-luminosity AGNs. Therefore, the AGN emission could potentially contribute to the observed SED, biasing the derived stellar masses. We investigate this by subtracting the AGN contribution from the observed SED and by re-fitting the corrected SED. Specifically, we assume a power-law SED for the AGN, with $F_{\nu} \propto \nu^{\alpha}$. The value of α has been derived by fitting the rest-frame UV and the MIPS $24 \mu\text{m}$ photometry, with values of α found in the range $-2.9 < \alpha < -1.7$. The maximum AGN contribution is then set by the rest-frame UV fluxes in combination with the $24 \mu\text{m}$ band, and subtracted from the observed SED. The resulting SEDs are finally re-modeled using FAST to derive stellar masses. We find that the derived stellar masses are smaller by typically $\sim 0.08 \text{ dex}$. For two of the three galaxies with X-ray detection, this analysis results in stellar masses smaller by only $0.05\text{--}0.08 \text{ dex}$, slightly larger than the random errors on M_{star} . For the third galaxy with X-ray detection (A2-15753), the AGN contribution is 0.18 dex , the largest in our mass-selected sample (although still much smaller than the systematic uncertainties due to different SED-modeling assumptions).

We stress that the estimated systematic effects caused by the AGN contributions should be regarded strictly as upper limits, as our approach maximizes, by construction, the contribution of the AGN to the observed SED. We therefore conclude that these systematic effects are in general small, and certainly much smaller than the systematic uncertainties caused by the different SED-modeling assumptions and/or by potential systematic errors in the photometric redshift estimates (see Section 5). We

¹⁸ The quoted errors on $L_{2-7 \text{ keV}}$ include the error due to photometric redshift random uncertainties.

note that additional contamination of some of the medium-band and broadband filter fluxes could potentially be caused by the presence of strong AGN line emission.

4. THE STELLAR MASS FUNCTION AND DENSITY

4.1. Methodology

We used the mass-selected sample defined in Section 2 to derive more accurate measurements of the high-mass end of the SMF of galaxies at $3.0 \leq z < 4.0$. To estimate the observed SMF, we have followed the analysis in Marchesini et al. (2009), which we refer to for detailed descriptions of the methods used in this work. Briefly, we have derived the SMF using two methods, the $1/V_{\max}$ estimator and a parametric maximum-likelihood method.

For the $1/V_{\max}$ estimator, we used the extended version as defined by Avni & Bahcall (1980). The Poisson error in each stellar mass bin was computed adopting the recipe of Gehrels (1986). As extensively discussed in the literature, the $1/V_{\max}$ estimator has the advantages of simplicity and no a priori assumption of a functional form for the stellar mass distribution; it also yields a fully normalized solution. However, it can be affected by the presence of clustering in the sample. Field-to-field variation represents a significant source of uncertainty in deep surveys, since they are characterized by small areas and hence small probed volumes. The contribution to the total error budget due to cosmic variance is quantified in Section 4.2.

We also measured the observed SMF using the STY method (Sandage et al. 1979), which is a parametric maximum-likelihood estimator. The STY method has been shown to be unbiased with respect to density inhomogeneities (e.g., Efstathiou et al. 1988), it has well-defined asymptotic error properties (e.g., Kendall & Stuart 1961), and does not require binning. We have assumed that the number density $\Phi(M)$ of galaxies is described by a Schechter (1976) function,

$$\Phi(M) = (\ln 10)\Phi^*[10^{(M-M^*)(1+\alpha)}] \times \exp[-10^{(M-M^*)}], \quad (1)$$

where $M = \log(M_{\text{star}}/M_{\odot})$, α is the low-mass end slope, $M^* = \log(M_{\text{star}}^*/M_{\odot})$ is the characteristic stellar mass at which the SMF exhibits a rapid change in the slope, and Φ^* is the normalization. Following Marchesini et al. (2009), the best-fit solution is obtained by maximizing the likelihood Λ with respect to the parameters α and M^* . The value of Φ^* is then obtained by imposing a normalization on the best-fit SMF such that the total number of observed galaxies in the sample is reproduced.

4.2. Uncertainties on the Stellar Mass Function

As discussed in Marchesini et al. (2009), uncertainties due to small-number statistics, photometric redshift errors, cosmic variance, and different SED-modeling assumptions contribute to the total error budget of the high-mass end of the high-redshift SMF.

The uncertainties on the SMF due to random photometric redshift errors have been estimated following the recipe in Marchesini et al. (2009).¹⁹ Specifically, for each galaxy in the K -selected sample, a set of 200 mock SEDs was created by

¹⁹ In Marchesini et al. (2009), systematic photometric redshift errors were estimated by adopting different template sets to derive the photometric redshifts. Here we decided not to use the other template sets distributed with EAZY due to the resulting $z_{\text{phot}} - z_{\text{spec}}$ comparisons being significantly worse, whereas in Marchesini et al. (2009) they resulted in $z_{\text{phot}} - z_{\text{spec}}$ comparisons of similar quality, or only slightly worse.

Table 5
SMF at $3 \leq z < 4.0$ Derived with the $1/V_{\max}$ Method

$\log M_{\text{star}}$ (M_{\odot})	$\log \Phi$ ($\text{Mpc}^{-3} \text{ dex}^{-1}$)	σ	σ_{Poi}	$\sigma_{z,\text{ran}}$	σ_{cv}	$\sigma_{\text{MA05,sys}}$
11.63	-5.282	+0.390 -0.409	+0.253 -0.282	0.151	0.256	+0.0 -0.6
11.48	-4.842	+0.240 -0.244	+0.146 -0.153	0.117	0.150	+0.0 -0.6

Notes. $\sigma = (\sigma_{\text{Poi}}^2 + \sigma_{\text{cv}}^2 + \sigma_{z,\text{ran}}^2)^{1/2}$ is the total 1σ random error, including the Poisson errors (σ_{Poi}), the errors due to random photometric redshift uncertainties ($\sigma_{z,\text{ran}}$), and the error due to cosmic variance (σ_{cv} ; see Section 4.2); $\sigma_{\text{MA05,sys}}$ is the systematic uncertainty due to the different SED-modeling assumptions, i.e., the Maraston (2005) stellar population synthesis models, Kroupa (2001) IMF, solar metallicity, Calzetti et al. (2000) extinction curve, and exponentially declining SFHs (see Sections 2.3 and 4.2).

perturbing each flux point according to its formal error bar. Second, we estimated the photometric redshift in the same way as described in Section 2.2. Third, we fitted the mock SEDs with FAST to estimate stellar masses as described in Section 2.3. Finally, we have derived SMFs of galaxies with the $1/V_{\max}$ and the maximum-likelihood analysis for each of the 200 Monte Carlo realizations of the K -selected sample. The contribution to the total error budget of the SMF derived using the $1/V_{\max}$ method due to random photometric redshift errors ($\sigma_{z,\text{ran}}$) is listed in Table 5, and is roughly 0.13 dex, about a factor of 1.7 smaller than in Marchesini et al. (2009). Similar to what was found in Marchesini et al. (2009), the contribution of random photometric redshift errors on the Schechter function parameters of the SMF at $3.0 \leq z < 4.0$ is instead negligible with respect to Poisson errors. In fact, Poisson errors largely dominate the random error budget of the Schechter function parameters due to the complete lack of constraint on the low-mass end slope α (see Table 6). Because the low-mass end slope is not constrained by the NMBS data set, we have also repeated the maximum-likelihood analysis fixing the value of the low-mass end slope at $\alpha = -1.0$ (corresponding to the value of the low-mass end slope of the SMF of galaxies at $1.3 \leq z < 2.0$ from Marchesini et al. 2009) and $\alpha = -1.75$ (corresponding to the value of the low-mass end slope of the SMF of galaxies at $2.5 < z < 3.5$ from Kajisawa et al. 2009).

To quantify the uncertainties due to field-to-field variations in the determination of the SMF, we proceeded as in Marchesini et al. (2007). Briefly, using the $1/V_{\max}$ method, we measured Φ^j , the galaxy number density in the stellar mass bin ΔM for the j th field. The contribution to the error budget from cosmic variance is estimated using $\sigma_{\text{cv}} = \text{rms}(\Phi^j)/\sqrt{2}$. The final 1σ random error associated with $\Phi(M)$ is then $\sigma = (\sigma_{\text{Poi}}^2 + \sigma_{\text{cv}}^2 + \sigma_{z,\text{ran}}^2)^{1/2}$, where σ_{Poi} is the Poisson error in each mass bin. The values of σ_{Poi} and σ_{cv} are also listed in Table 5. The contribution to the error budget from cosmic variance can also be estimated for a given population using predictions from cold dark matter theory and the galaxy bias. We have derived the cosmic variance following the cosmic variance cookbook presented by Moster et al. (2010) and using the parameters for our survey and for a population of massive galaxies with $M_{\text{star}} > 10^{11} M_{\odot}$, resulting in an uncertainty due to the cosmic variance of 0.18 dex, in very good agreement with our empirical estimate.

Whereas we refer to Marchesini et al. (2009) for a complete analysis and discussion of the systematic uncertainties due to different SED-modeling assumptions, we have repeated the whole analysis adopting the stellar population synthesis models of Maraston (2005) with a Kroupa (2001) IMF and exponentially

Table 6
Best-fit Schechter Function Parameters of the SMF

α	$\log M_{\text{star}}^*$ (M_{\odot})	Φ^* ($10^{-5} \text{ Mpc}^{-3} \text{ dex}^{-1}$)
$0.6^{+4.4}_{-5.4} (2.0^{+4.0}_{-6.0})$	$10.97^{+1.54}_{-0.38} (10.77^{+1.23}_{-0.34})$	$2.00^{+14.93}_{-1.90} (0.24^{+113.48}_{-0.23})$
$-1.0(-1.0)$	$11.17^{+0.18}_{-0.14} (11.05^{+0.31}_{-0.22})$	$3.73^{+5.35}_{-2.10} (2.58^{+14.2}_{-2.06})$
$-1.75(-1.75)$	$11.30^{+0.28}_{-0.18} (11.15^{+0.45}_{-0.25})$	$2.90^{+6.37}_{-2.13} (2.62^{+20.12}_{-2.38})$

Notes. The quoted error corresponds to the 1σ error estimated from the maximum-likelihood analysis as described in Section 4.2. The values in parentheses are the best-fit Schechter function parameters derived with different SED-modeling assumptions, i.e., the Maraston (2005) stellar population synthesis models, Kroupa (2001) IMF, solar metallicity, Calzetti et al. (2000) extinction curve, and exponentially declining SFHs (see Sections 2.3 and 4.2). Note that the low-mass end slope is completely unconstrained. The second and third rows list the best-fit Schechter function parameters obtained with fixed α .

declining SFHs with values of the e -folding timescale ranging from 100 Myr to 10 Gyr. The resulting systematic uncertainties on the SMF measured with the $1/V_{\text{max}}$ method are listed in Table 5, and the corresponding values of the Schechter function parameters measured with the maximum-likelihood analysis are listed in Table 6. The adoption of the Maraston (2005) models result, in general, in derived stellar masses smaller by ~ 0.15 dex. As previously shown in Marchesini et al. (2009) and Muzzin et al. (2009), different combinations of adopted metallicity and extinction curve also result in systematic effects on the derived stellar masses, although to a much smaller extent with respect to the biases introduced by different stellar population synthesis models and the specific choices of the adopted SFHs. In particular, different assumptions on the SFH with respect to those adopted in our work (e.g., two-component models of the SFH, or exponentially increasing SFH) can introduce additional systematic biases toward both larger and smaller stellar masses (Wuyts et al. 2007; Lee et al. 2009; Maraston et al. 2010). In Section 5, we consider the systematic uncertainties due to the inclusion of an additional template in the template set used to derive photometric redshifts with EAZY. This additional template consists of an old (1 Gyr) and dusty ($A_V = 3$ mag) single stellar population.

4.3. Stellar Mass Function

Figure 7 shows the SMFs of galaxies at redshift $3.0 \leq z < 4.0$ derived in this work (colored symbols) compared to the SMF of galaxies at $3.0 \leq z < 4.0$ derived in Marchesini et al. (2009, black and gray symbols). Points with error bars show the SMFs derived using the $1/V_{\text{max}}$ method. The solid curves show the SMFs derived with the maximum-likelihood analysis, while the shaded regions represent their 1σ uncertainties. The plotted uncertainties of the SMF measurements from Marchesini et al. (2009), the thick red error bars, and the yellow-shaded area represent the total 1σ random errors, including cosmic variance and photometric redshift errors as quantified in Section 4.2. The thin red error bars and the orange-shaded area also include the systematic uncertainties due to the different SED-modeling assumptions adopted in this work.

The large surveyed area (i.e., effective area of 0.44 deg^2 , a factor of ~ 3 larger than in Marchesini et al. 2009) and the accurate photometric redshift estimates allow for the determination of the number density of the most massive galaxies at $3.0 \leq z < 4.0$ with unprecedented accuracy, as clearly shown by Figure 7 from the comparison with the SMF previously derived by Marchesini et al. (2009). Figure 8 shows the comparison between the SMF

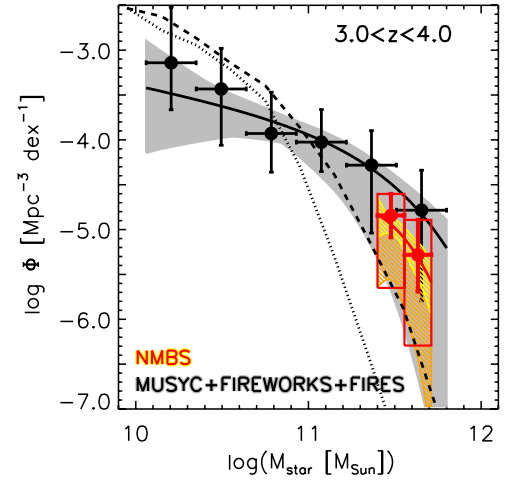


Figure 7. SMFs of galaxies at redshift $3.0 \leq z < 4.0$ from the NMBS (red, orange, and yellow symbols) and the analysis of Marchesini et al. (2009, black and gray symbols). The filled symbols represent the SMF derived with the $1/V_{\text{max}}$ method, with error bars showing the total 1σ random errors, including photometric redshift errors and field-to-field variations; the red boxes also include the systematic uncertainties due to the different SED-modeling assumptions adopted (see Section 2.3). The solid curves represent the SMFs derived with the maximum-likelihood analysis, with shaded regions representing the 1σ uncertainties. The black error bars and gray-shaded area include the systematic uncertainties due to different template sets in the photometric redshift estimate. The orange-hatched area also includes the systematic uncertainties due to the different SED-modeling assumptions adopted in our analysis. The dotted and dashed black curves represent the predicted SMFs from the semi-analytic model of Somerville et al. (2008), where the dashed curve is derived from the dotted curve after convolution with a normal distribution of standard deviation 0.25 dex. The NMBS allows us to derive more accurate measurements of the high-mass end of the SMF of galaxies at $3.0 \leq z < 4.0$.

(A color version of this figure is available in the online journal.)

derived in this work and previous measurements of the SMFs of galaxies at $z \sim 3.5$. The high-mass end of the SMF measured in our analysis is in good agreement with previous measurements.

Combined with the results from Marchesini et al. (2009), the number density of the most massive galaxies appears to have evolved by a factor of ~ 2 from $z = 3.5$ to 2.5 , and by a factor of ~ 3 from $z = 3.5$ to 1.6 . This is shown in Figure 9, where the ratio of the high- z SMFs and the local SMF from Cole et al. (2001) is plotted as a function of the stellar mass. However, due to the steepness of the high-mass end, the implied evolution of the number density translates to small growth in stellar mass of the most massive galaxies, by 30%–40% from $z = 3.5$ to 1.6 , and by $\sim 40\%$ from $z = 1.6$ to 0.1 , although systematic uncertainties allow for a larger evolution. Recently, van Dokkum et al. (2010) have estimated growth by a factor of ~ 2 in the stellar mass of massive galaxies from $z = 2$ to 0.1 , in apparent contradiction with our results. However, the selection of the sample of massive galaxies in van Dokkum et al. (2010) was very different to ours, as galaxies were selected at a constant number density of $n = 2 \times 10^{-4} \text{ Mpc}^{-3}$ over the targeted redshift range. This approach selects galaxies with stellar masses $\log M_{\text{star}} = 11.45 \pm 0.15$ at $z = 0.1$ and $\log M_{\text{star}} = 11.15 \pm 0.15$ at $z = 2$, a factor of ~ 1.2 and ~ 2.1 smaller than the typical galaxy in our mass-selected sample at $z = 3.5$. As a consequence of the mass-dependent evolution derived in Marchesini et al. (2009), smaller growth in the stellar mass of massive galaxies would therefore be derived if the selection were done at a value of the number density typical of galaxies with stellar masses $\log M_{\text{star}} = 11.5$ at $z = 0.1$ (the average mass of our mass-selected sample at $z = 3.5$). We also

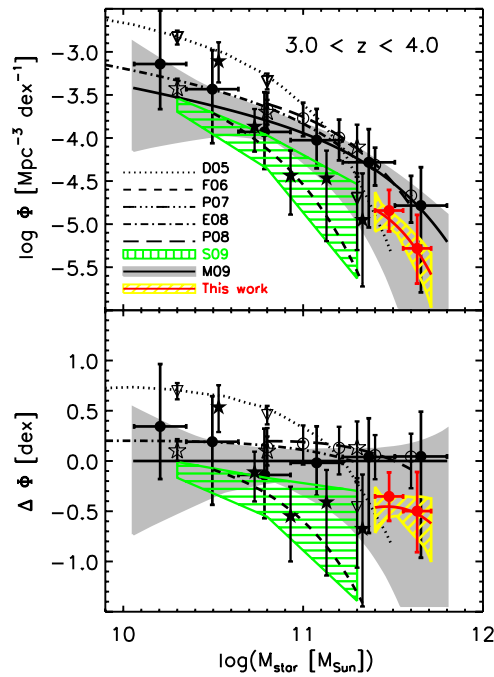


Figure 8. Top panel: comparison between the SMF at $3.0 \leq z < 4.0$ (yellow and red) derived from this work and previous measurements from the literature, including the measurement from Marchesini et al. (2009, gray and black). For the SMFs derived from this work and Marchesini et al. (2009), the filled circles represent the measurements using the $1/V_{\text{max}}$ method, while the solid curves represent the measurements using the maximum-likelihood analysis; the 1σ error bars of the $1/V_{\text{max}}$ measurements include Poisson errors, field-to-field variations, and uncertainties from photometric redshift uncertainties (both random and systematic). The same is true for the 1σ error of the maximum-likelihood measurements (yellow and gray regions). Previous works are plotted as filled stars and dashed curves (Fontana et al. 2006, F06); open circles and long-dashed curves (Pérez-González et al. 2008, P08); open stars and dot-dashed curves (Elsner et al. 2008, E08), open triangles and dotted curves (Drory et al. 2005, D05). The hatched green area shows the SMF of *B*-dropout galaxies from Stark et al. (2009). Bottom panel: symbols as in the top panel, but now the differences between the SMFs measurements shown in the top panel and the SMF from Marchesini et al. (2009), $\Delta \Phi = \log \Phi - \log \Phi_{\text{DM09}}$, are plotted as a function of stellar mass to highlight the differences.

(A color version of this figure is available in the online journal.)

note that systematic uncertainties in the SED modeling, as well as the choice of the $z \sim 0$ benchmark, can play an important role in the estimate of the evolution of the stellar content in massive galaxies. Therefore, we conclude that the results in van Dokkum et al. (2010) are broadly consistent with ours once the different selection and systematic uncertainties are taken into account.

Figure 7 also shows the SMF predicted from the semi-analytic model of Somerville et al. (2008, dotted black curve).²⁰ This model, built on the previous models described in Somerville & Primack (1999) and Somerville et al. (2001), presents several improvements, including, but not limited to, tracking of a diffuse stellar halo component built up of tidally destroyed satellites and stars scattered in mergers, galaxy-scale AGN-driven winds, fueling of black holes with hot gas via Bondi accretion, and heating by radio jets. The prediction from the Somerville et al. (2008) semi-analytic model is taken from their fiducial WMAP-3 model, which adopts a fraction $f_{\text{scatter}} = 0.4$ of the stars in merged satellite galaxies added to a diffuse

²⁰ In the comparison with the model predictions, we decided to use only the model of Somerville et al. (2008) as it was the model showing the smallest disagreements at the high-mass end in the comparison presented in Marchesini et al. (2009).

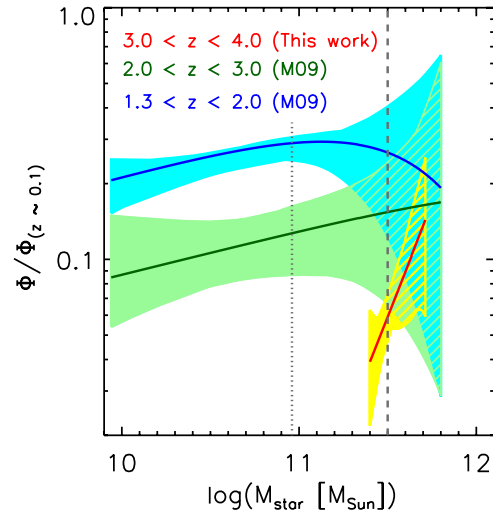


Figure 9. Ratio of the high- z SMFs (Φ) and the local SMF ($\Phi_{z \sim 0.1}$; Cole et al. 2001) plotted as a function of the stellar mass as measured from the maximum-likelihood analysis. The SMFs at $z \sim 1.6$ (blue) and $z \sim 2.5$ (green) are taken from Marchesini et al. (2009). The shaded regions represent the total 1σ random uncertainties. The vertical dashed and dotted lines represent the value of $3 \times 10^{11} M_{\odot}$ and the $z = 0.1$ characteristic stellar mass, $M_{\text{star}}^* \sim 10^{11} M_{\odot}$, respectively.

(A color version of this figure is available in the online journal.)

component distributed in a very extended halo or envelope. The dashed curve in Figure 7 represents the model-predicted SMF convolved with a normal distribution of standard deviation 0.25 dex, intended to represent measurement errors in $\log M_{\text{star}}$ (Fontanot et al. 2009).²¹

The comparison between the model-predicted and the NMBS-derived SMF provides further supporting evidence for the deficit of very massive galaxies at $3.0 \leq z < 4.0$ in the theoretical models of galaxy formation, a disagreement that was only marginal with the previously derived SMFs. Without the inclusion of the systematic uncertainties, the disagreement between the observed and the (convolved) model-predicted high-mass end of the SMF of galaxies at $3.0 \leq z < 4.0$ is significant at the 3σ level. The significance of the disagreement is reduced to only 1σ if we include the systematic uncertainties due to different SED-modeling assumptions as estimated in Section 4.2 (i.e., adopting the Maraston 2005 instead of the Bruzual & Charlot 2003 models, and different SFHs). We note that systematic uncertainties due to an evolving IMF can play an additional role in reducing the disagreement between observed and model-predicted SMFs.

4.4. Number and Stellar Mass Densities

The number density, η , and stellar mass density, ρ , in massive galaxies at $3.0 \leq z < 4.0$ have been estimated by integrating the SMF at $M_{\text{star}} > 10^{11.40} M_{\odot}$.²² These values are listed in Table 7 together with the corresponding total 1σ errors. Table 7 also lists the 1σ lower limits of the number and stellar mass densities of galaxies more massive than $10^8 M_{\odot}$, as well as the densities estimated with the second set of SED-modeling

²¹ We note that the typical random error on the stellar masses for the galaxies in our $3.0 \leq z < 4.0$ sample is smaller than 0.25 dex by a factor of ~ 2 , due to the combination of accurate photometric redshift estimates and well-sampled SEDs delivered by the NMBS.

²² The SMF has been integrated using $M_{\text{star}} = 10^{13} M_{\odot}$ as the upper limit of the integral. Due to the exponential behavior of the SMF at the high-mass end, the estimated stellar mass density does not depend significantly on the specific value of this limit.

Table 7
Number and Stellar Mass Densities at $3 \leq z < 4.0$

Densities	$M_{\text{star}} > 10^{11.40} M_{\odot}$	$M_{\text{star}} > 10^{8.0} M_{\odot}$
$\log(\eta \text{ [Mpc}^{-3}\text{)})$	$-5.55^{+0.19}_{-0.17} (-6.04^{+0.20}_{-0.18})$	$> -5.47 (> -6.00)$
$\log(\rho \text{ [} M_{\odot} \text{ Mpc}^{-3}\text{)})$	$6.00^{+0.30}_{-0.10} (5.49^{+0.36}_{-0.13})$	$> 6.09 (> 5.59)$
$\log(\eta \text{ [Mpc}^{-3}\text{)})$	$-5.55^{+0.17}_{-0.17} (-6.04^{+0.19}_{-0.17})$	$-3.61^{+0.71}_{-0.65} (-3.79^{+1.33}_{-1.03})$
$\log(\rho \text{ [} M_{\odot} \text{ Mpc}^{-3}\text{)})$	$6.00^{+0.12}_{-0.10} (5.48^{+0.15}_{-0.11})$	$6.74^{+0.31}_{-0.26} (6.45^{+0.62}_{-0.42})$
$\log(\eta \text{ [Mpc}^{-3}\text{)})$	$-5.55^{+0.18}_{-0.17} (-6.05^{+0.20}_{-0.17})$	$-1.90^{+0.71}_{-0.69} (-2.06^{+1.28}_{-1.03})$
$\log(\rho \text{ [} M_{\odot} \text{ Mpc}^{-3}\text{)})$	$6.01^{+0.14}_{-0.10} (5.48^{+0.18}_{-0.10})$	$7.25^{+0.37}_{-0.34} (7.04^{+0.70}_{-0.59})$

Notes. Number density, η , and stellar mass density, ρ , at $3.0 \leq z < 4.0$ estimated by integrating the best-fit Schechter SMF over the specified stellar mass range. The quoted 1σ errors include Poisson errors, errors due to photometric redshift uncertainties, and errors due to cosmic variance. The values in parentheses are the results corresponding to the different SED-modeling assumptions, i.e., the Maraston (2005) stellar population synthesis models, Kroupa (2001) IMF, solar metallicity, Calzetti et al. (2000) extinction curve, and exponentially declining SFHs (see Sections 2.3 and 4.2). The third and fifth rows list the number densities estimated fixing the low-mass end slope at $\alpha = -1.0$ and -1.75 , respectively; the fourth and sixth rows list the stellar mass densities estimated fixing the low-mass end slope at $\alpha = -1.0$ and -1.75 , respectively.

assumptions adopted (see Section 2.3). The number and stellar mass densities estimated by fixing the low-mass end slope at $\alpha = -1.0$ and $\alpha = -1.75$ are also listed in Table 7. Compared to the total stellar mass density in galaxies with $M_{\text{star}} > 10^{10} M_{\odot}$ at $3.0 \leq z < 4.0$ estimated by Marchesini et al. (2009), the contribution of the galaxies in our mass-selected sample is $\sim 8^{+13}_{-3}\%$ (not including systematic errors). Note, however, that this estimate depends strongly on the value of the low-mass end slope of the SMF derived in Marchesini et al. (2009), which is still very poorly constrained (e.g., Kajisawa et al. 2009; Marchesini et al. 2009; Reddy & Steidel 2009).

In recent years, there have been several claims of the existence of a population of massive (and evolved) galaxies at even larger redshifts, i.e., $z \gtrsim 4-5$ (e.g., Yan et al. 2006; Wiklind et al. 2008; Mancini et al. 2009).

In particular, Wiklind et al. (2008) reported a significant population of massive galaxies at $4.9 \leq z < 6.5$ found in the 125 arcmin² GOODS-South field dominated by old stellar populations with $M_{\text{star}} = (0.3-3) \times 10^{11} M_{\odot}$. In their sample, there are only two objects with $M_{\text{star}} > 2.5 \times 10^{11} M_{\odot}$ (one already identified by Mobasher et al. 2005 as a candidate for a massive, evolved galaxy at $z \sim 6.5$), implying a stellar mass density at $z \sim 5.7$ of $\rho(M_{\text{star}} > 10^{11.40} M_{\odot}) \approx 2 \times 10^6 M_{\odot} \text{ Mpc}^{-3}$, and suggesting no evolution of the stellar mass density in the most massive galaxies over the ~ 800 Myr interval from $z \sim 5.7$ to 3.5. However, Dunlop et al. (2007) have concluded that there is no convincing evidence for any galaxy with stellar mass $M_{\text{star}} > 2 \times 10^{11} M_{\odot}$ and $z > 4$ in the GOODS-South field, which implies a much stronger evolution of the stellar mass density in very massive galaxies in the first 1.5 Gyr of the universe. Wiklind et al. (2008) also estimated a stellar mass density in galaxies with $M_{\text{star}} > 10^{10.8} M_{\odot}$ of $\rho > 5 \times 10^6 M_{\odot} \text{ Mpc}^{-3}$ (after correction for the different IMF). Combining the results from our analysis and Marchesini et al. (2009), this implies an increase of the stellar mass density in massive galaxies by a factor of ~ 1.6 from $z \sim 5.7$ to 3.5.

SMFs and stellar mass densities have also been estimated for V- and i-dropout galaxies at $z \sim 5$ and ~ 6 , respectively (Yan et al. 2006; Eyles et al. 2007; Stark et al. 2007; McLure

et al. 2009; Stark et al. 2009). A comparison with these results is not straightforward. First, these are all optically selected samples, which can be potentially biased against massive and evolved galaxies. In contrast, our sample is a mass-complete sample constructed from a K-selected catalog. Second, the stellar mass ranges probed by these studies are very different from ours. Our mass-selected sample probes the most massive galaxies, i.e., those with $M_{\text{star}} > 2.5 \times 10^{11} M_{\odot}$, whereas all the above works probe galaxies with typically much smaller stellar masses, i.e., in the range $10^9-10^{11} M_{\odot}$. A rough estimate of the evolution of the stellar mass density in galaxies more massive than $2 \times 10^9 M_{\odot}$ can be performed by comparing the stellar mass densities derived by the above studies with the stellar mass density obtained by combining our results with those from Marchesini et al. (2009) at lower stellar masses, and extrapolating the Schechter function to the stellar mass limit probed by the above studies. After correcting for the different IMFs, the stellar mass density in galaxies more massive than $2 \times 10^9 M_{\odot}$ at $z \sim 5$ and ~ 6 is $\rho = (2.3-6.3) \times 10^6 M_{\odot} \text{ Mpc}^{-3}$ (McLure et al. 2009; Stark et al. 2007; Stark et al. 2009) and $\rho = (0.7-4.3) \times 10^6 M_{\odot} \text{ Mpc}^{-3}$ (Yan et al. 2006; Eyles et al. 2007; McLure et al. 2009; Stark et al. 2009), respectively. At $z = 3.5$, we estimate $\rho = 1.5 \times 10^7 M_{\odot} \text{ Mpc}^{-3}$, which implies an evolution of the stellar mass density in galaxies with $M_{\text{star}} > 2 \times 10^9 M_{\odot}$ by a factor of $\sim 2-7$ and $\sim 3-22$ from $z \sim 5$ and $z \sim 6$ to 3.5, respectively. We stress, however, that the estimated evolution from $z > 4$ is very uncertain and affected both by large uncertainties in the SMF of galaxies at $z > 4$ (especially at the high-mass end) and poor constraints on the low-mass end slope of the SMF of galaxies at $3.0 \leq z < 4.0$.

5. MASSIVE, OLD, AND DUSTY GALAXIES AT $2 < z < 3$?

In this section, we consider the case of adding an additional template to the EAZY template set used to derive photometric redshifts.

Previous searches for old and massive galaxies at $z > 4$ highlighted the difficulty in unambiguously identifying old and massive objects at extreme redshifts on the basis of spectral fitting. In particular, Dunlop et al. (2007) have shown that equally acceptable solutions could be obtained at $z \sim 5$ with high stellar masses ($M_{\text{star}} \sim 6 \times 10^{11} M_{\odot}$) and low extinction ($A_V \sim 0.4$ mag) and at $z \sim 2$ with moderate stellar masses ($M_{\text{star}} \sim 7 \times 10^{10} M_{\odot}$) and high extinction ($A_V \sim 3.8$ mag).

To robustly estimate photometric redshifts, the template set needs to be large enough that it spans the broad range of multi-band galaxy colors and small enough that the color and redshift degeneracies are kept to a minimum (e.g., Benitez 2000). The default template set used in this work was carefully constructed and tested in Brammer et al. (2008). It has been shown to satisfy the requirements for a satisfactory template set, providing significantly reduced systematic effects and smaller scatter in z_{phot} versus z_{spec} at all redshifts. This template set already includes a dusty starburst model (50 Myr old and $A_V = 2.75$ mag). Here, we include an additional template representative of an old (1 Gyr; $\tau = 100$ Myr) and very dusty ($A_V = 3$ mag) galaxy, similar to the reddest template used in Blanton & Roweis (2007), and we repeat the whole analysis. We note that in the local universe old stellar populations are usually not (very) dust-obscured, and that it remains to be seen whether this template is physically plausible for the high-redshift universe.

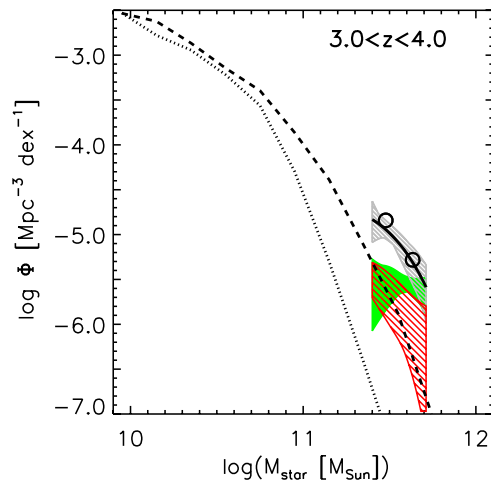


Figure 10. SMFs of galaxies at $3.0 \leq z < 4.0$ measured by adopting the additional old and dusty template in the estimate of the photometric redshifts. The gray hatched region, the black solid line, and empty circles represent the SMF measured adopting the default EAZY template set. The colored regions represent the SMFs derived by adopting the additional old and dusty template in the estimate of the photometric redshifts. The green region is obtained by using the Bruzual & Charlot (2003) stellar population synthesis models, while the hatched red region is obtained by using the Maraston (2005) models. The dotted and dashed curves represent the predicted SMFs from the semi-analytic model of Somerville et al. (2008).

(A color version of this figure is available in the online journal.)

For eight galaxies (C1-4890, C1-6110, C1-7340, C1-15367, C1-18825, A2-6835, A2-15753, and A2-18070), the resulting photometric redshift estimates formally lie at $z < 3$. The three objects in the AEGIS field now have $z \sim 2.9$, hence just below our redshift selection window, consistent with their redshift probability functions and with A2-15753 having rest-frame UV color typical of LBGs. The five objects in the COSMOS field are instead shifted to much lower redshifts, i.e., $z \sim 2.4$, significantly lower than allowed for by their redshift probability functions plotted in Figure 1.

We used FAST to refit stellar population synthesis models to the eight galaxies that moved to $z < 3$. Six of the eight are best fitted by an old stellar population (age ≈ 2 Gyr, i.e., as old as the age of the universe at their redshifts), large stellar masses ($M_{\text{star}} \approx 2 \times 10^{11} M_{\odot}$), and large values of extinction ($A_V \approx 2.1$ mag). The remaining two objects are instead best fitted by very dusty ($A_V \approx 3.2$ mag), young (age ≈ 50 Myr) starbursts.

The SMFs of galaxies at $3.0 \leq z < 4.0$ measured by adopting the additional old and dusty template in the estimate of the photometric redshifts are shown in Figure 10. As shown in this figure, the systematic effect on the derived SMF of galaxies at $3.0 \leq z < 4.0$ due to the inclusion of this additional (although generally not included) template is significant, larger than the systematic effect caused by the stellar population synthesis model assumption alone, and brings the observed SMF of galaxies at $3.0 \leq z < 4.0$ in better agreement with the SMF predicted from semi-analytic models.

We note, however, that the EAZY and FAST best-fit models using the additional old and dusty template do not provide statistically better modeling of the observed SEDs with respect to the default template set without the old and dusty template. This further shows the ambiguity and difficulty in characterizing the population of massive galaxies at $z > 3$, even with our adopted dataset, which is characterized by exquisite wavelength coverage from the ultraviolet to the mid-infrared.

In summary, our sample of massive galaxies at $3 \leq z < 4$ could potentially be contaminated (up to $\sim 50\%$) by a previously unrecognized population of massive, old, and very dusty galaxies at $z \sim 2.6$. We note that the existence of such a population would be an important and puzzling result in itself.

6. SUMMARY AND CONCLUSIONS

In this paper, we have used the far-ultraviolet to mid-infrared coverage of the NMBS to derive the observed and rest-frame properties of a complete sample of galaxies at $3.0 \leq z < 4.0$ with $M_{\text{star}} > 2.5 \times 10^{11} M_{\odot}$, and to provide more accurate measurements of the high-mass end of the SMF of galaxies at $3.0 \leq z < 4.0$. With the addition of five medium-bandwidth NIR filters, NMBS delivers accurate photometric redshift, $\sigma_z/(1+z) \approx 0.02$, for K -selected sources at $z > 1.5$, and provides well-sampled SEDs in the critical wavelength regime around the Balmer/4000 Å breaks, allowing, for the first time, the accurate detection of the Lyman and the Balmer/4000 Å breaks simultaneously. Combined with its large surveyed area, $\sim 0.5 \text{ deg}^2$, it allowed us to construct a statistically significant and representative sample of 14 very massive ($M_{\text{star}} > 2.5 \times 10^{11} M_{\odot}$) galaxies over the redshift range $3.0 \leq z < 4.0$.

The typical very massive galaxy at $3.0 \leq z < 4.0$ is red and faint in the observer's optical, with a median r -band magnitude of $\langle r_{\text{tot}} \rangle = 26.1$. The median $H - K$ color is 1.2, with only two objects having $H - K < 0.9$ at a significant level, highlighting the efficiency of the $H - K$ color technique in selecting galaxies at $z > 3$ with prominent breaks in the rest-frame optical (Brammer & van Dokkum 2007). The median rest-frame $U - V$ color $\langle U - V \rangle = 1.6$ is similar to local Sb spiral galaxies, although we find a range in $U - V$ colors, from the typical color of nearby irregular to those of local elliptical galaxies. The median $U - V$ slope is $\langle \beta \rangle = -0.36$, indicating a relatively flat spectrum in F_{λ} . Intriguingly enough, the distribution of UV slopes of the mass-selected sample at $3.0 \leq z < 4.0$ is very different from the distributions of UV slopes of UV-selected galaxies at $z > 2.5$ as well as the $H - K$ galaxies at $z \sim 3.7$ so far discovered, which show distributions of β peaked at $\beta < -1.6$. This difference is most likely due to the very different ranges in stellar mass probed by the different samples with the $H - K > 0.9$ galaxies in Brammer & van Dokkum (2007) being on average a factor of ~ 15 less massive than our sample, and the typical UV-selected galaxies having masses in the range $10^9 - 10^{11} M_{\odot}$ (e.g., Shapley et al. 2001; Magdis et al. 2010).

By constructing a mass-limited sample from a K -selected catalog with accurate photometric redshifts rather than the typical color-selection techniques, we were able to find a population of galaxies mostly complementary to the typical population of dropout galaxies at $z \sim 3-4$. Specifically, we have shown that only 57% of the mass-selected sample has observed optical colors that satisfy either the U - or the B -dropout color criteria. However, $\sim 50\%$ of these galaxies are too faint in the observed optical to be included in typical spectroscopic samples of LBGs.

From the SED modeling, our complete sample of massive galaxies at $3.0 \leq z < 4.0$ seems to show a range in stellar population properties. About 40%–60% of the sample is characterized by ages consistent with the age of the universe at the targeted redshifts, suggesting that the bulk of the stellar mass in these systems was formed at very early times. Dust seems to be quite ubiquitous in massive galaxies at $3.0 \leq z < 4.0$, with a median extinction of $\langle A_V \rangle = 1.0$ mag. About 30% of

the sample has SFR estimates from SED modeling consistent with no star formation activity, while the rest of the sample is characterized by significant star formation activity, as high as several hundreds of solar masses per year. Of particular interest is the $z = 3.54$ galaxy C1-22857, which has an estimated stellar mass of $\sim 3 \times 10^{11} M_{\odot}$, a maximally old age, and completely suppressed star formation. This galaxy is also not detected in the *Spitzer*-MIPS $24 \mu\text{m}$ data, further supporting its quiescent nature. Recent ultra-deep NIR spectroscopic observations have confirmed a massive galaxy at $z = 2.18$ with strongly suppressed star formation, pushing back its formation redshift to $z \gtrsim 4-7$ (Kriek et al. 2009). Spectroscopic confirmation of the quiescent nature of C1-22857, as well as other objects in our mass-selected sample, is of paramount importance, as it would provide even stronger evidence that massive galaxies formed their stars extremely efficiently very early in time.

Surprisingly, most (>80%) of the massive galaxies $3.0 \leq z < 4.0$ are detected in the MIPS $24 \mu\text{m}$ data. The total IR luminosities estimated from the observed $24 \mu\text{m}$ fluxes range from 5×10^{12} to $4.0 \times 10^{13} L_{\odot}$, typical of ULIRGs and HLIRGs, implying extreme dust-enshrouded SFRs ($\sim 600-4300 M_{\odot} \text{yr}^{-1}$, tens to several hundreds of times larger than the SFRs estimated from SED modeling), or very common heavily obscured AGNs, or both in the most massive galaxies at $z = 3.5$. Whereas it is not possible to discriminate between the AGN or starburst as the dominant source responsible in heating the dust, we favor the AGN as a significant, if not a dominant, contributor. Specifically, the reasons for this are trifold. First, the extreme MIPS-derived SFRs cannot be sustained for more than $\sim 10^8$ yr without overprediction of the high-mass end of the SMFs of galaxies at $z < 3$. This seems in contradiction with the very large fraction of MIPS detections, which implies a long duty cycle of the star formation. Second, 80% of the MIPS-detected sources are HLIRGs. In the local universe, the AGN is thought to be the dominant source of radiation responsible for the far- and mid-IR SEDs of HLIRGs. Third, for the targeted redshift range, the $24 \mu\text{m}$ band probes the rest-frame wavelengths from ~ 4.8 to $\sim 7.1 \mu\text{m}$, where hot dust dominates the MIR emission, and the contribution from an AGN as the source of the radiation field heating the dust becomes increasingly more likely. If the MIPS $24 \mu\text{m}$ emission is dominated by AGN-heated dust, the large fraction of very massive galaxies at $3.0 \leq z < 4.0$ with MIPS detection suggests that AGNs are very common, providing further supporting evidence for the co-evolution of massive galaxies and AGNs. We note that three galaxies are detected in the X-ray, with 2–7 keV luminosities and hardness ratios typical of obscured, high-luminosity AGNs. Observations at longer wavelengths (e.g., in the far-IR), as well as other signatures of AGNs (e.g., detection of narrow-emission lines), are necessary to constrain the occurrence of AGNs in this sample and to discriminate between dust-enshrouded star formation and heavily obscured AGNs.

We have significantly improved the measurements of the high-mass end of the SMF of galaxies at $3.0 \leq z < 4.0$ (the complete analysis of the evolution of the SMF of galaxies over the redshift range $0.5 \lesssim z < 4.0$ from NMBS will be presented in D. Marchesini et al., 2010, in preparation). The accurate photometric redshifts and the large surveyed area allowed us to significantly reduce the contributions of photometric redshift errors and cosmic variance to the total error budget. The measured high-mass end is in very good agreement with previous measurements, providing further supporting evidence for the existence of a significant number of very massive galaxies

at $z = 3.5$. Combined with the results from Marchesini et al. (2009), the number density of the most massive galaxies appears to have evolved little from $z = 3.5$ to 1.6, with a larger subsequent evolution down to $z \sim 0.1$. These results are broadly consistent with the growth by a factor of ~ 2 in the stellar mass of massive galaxies from $z = 2$ to 0.1 recently estimated in van Dokkum et al. (2010), once the different sample selection and systematic uncertainties are taken into account. The contribution of $M_{\text{star}} > 2.5 \times 10^{11} M_{\odot}$ galaxies to the total stellar mass budget at $3.0 \leq z < 4.0$ in galaxies with $M_{\text{star}} > 10^{10} M_{\odot}$ is $\sim 8_{-3}^{+13}\%$, although this estimate strongly depends on the value of the low-mass end slope of the SMF, which is still very poorly constrained. The stellar mass density in galaxies more massive than $2 \times 10^9 M_{\odot}$ seems to evolve by a factor of 5 ± 3 and 13 ± 10 from $z \sim 5$ and ~ 6 , respectively, to $z = 3.5$.

Our measurement of the high-mass end of $3.0 \leq z < 4.0$ seems to exacerbate the disagreement between the observed number densities of massive galaxies and those predicted by the latest generation of galaxy formation models (e.g., Somerville et al. 2008). The disagreement between the observed and the model-predicted high-mass end of the SMF at $3.0 \leq z < 4.0$ is significant at the $\sim 3\sigma$ level if only random errors are considered. However, systematic errors now dominate the total error budget at $3.0 \leq z < 4.0$, leading to uncertainties of a factor of ~ 8 in the densities at the high-mass end. When systematic uncertainties due to different SED-modeling assumptions are included, the disagreement found between observed and model-predicted SMFs is only marginally significant. We finally note that additional systematic uncertainties on the high-mass end of the $3.0 \leq z < 4.0$ SMF could be potentially introduced by either (1) the intense star formation activity and/or the very common AGN activity as inferred from the MIPS $24 \mu\text{m}$ detections, and/or (2) contamination by a significant population of massive, old, and dusty galaxies at $z \sim 2.6$ previously unrecognized. This might indicate that the high-mass end of the SMF cannot be properly constrained without further spectroscopic data.

The NMBS has allowed us to study with unprecedented accuracy the population of very massive galaxies at $3.0 \leq z < 4.0$, thanks to its wide surveyed area, the accurate photometric redshifts, and the well-sampled SEDs in the rest-frame optical. To further improve the characterization of the galaxy population at the high-mass end at $3.0 \leq z < 4.0$, it is necessary to significantly increase the sample size, which is currently comprised of only 14 sources. NMBS-II, a shallow-wide accepted NOAO Survey Program specifically designed to further constrain the population of very massive galaxies at $z > 2$, will image an area of sky a factor of ~ 10 larger than NMBS, resulting in a significant increase in the number of very massive galaxies out to $z \sim 3.5$ with accurate photometric redshifts and well-sampled SEDs.

Follow-up multi-object spectroscopic observations in both the optical and NIR are of vital importance to confirm the redshifts and to better characterize the properties of the most massive galaxies at $z = 3.5$, including AGNs and Lyman- α emitter fractions, AGN and/or starburst contamination of the optical-to-MIR SEDs, superwind outflows, SFRs, and mass-to-light ratios. However, to probe the rest-frame wavelength regime redward of $\sim 5000 \text{ \AA}$, and hence to robustly constrain the SFHs and SFRs, measure metallicities, absorption lines and velocity dispersions from rest-frame optical features will require NIR-SPEC on the *James Webb Space Telescope*. The estimated total IR luminosities typical of HLIRGs make the very massive galaxies at $3 \leq z < 4.0$ in our sample ideal candidates for

follow-up observations with the Atacama Large Millimeter Array (ALMA). ALMA will be crucial in constraining the amount of dust and gas in these systems, as well as discriminating between dust-enshrouded star-formation and obscured AGN activity. Moreover, it will also allow for measurements of the kinematics in these systems, providing an independent estimate of the dynamical masses of the most massive galaxies at $3.0 \leq z < 4.0$.

Finally, to fully characterize the population of galaxies at $3 < z < 4$, the analysis performed in this work has to be extended to lower stellar masses. This will necessarily require very deep imaging with NIR medium-bandwidth filters to provide very accurate photometric redshifts and well-sampled SEDs down to faint *K*-band magnitudes.

We are grateful to the anonymous referee whose comments and suggestions helped to significantly improve this paper. Ron Probst and the NEWFIRM team are thanked for their work on the instrument and help during the observations. This paper is partly based on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/DAPNIA, at the Canada–France–Hawaii Telescope (CFHT), which is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences de l'Univers of the Centre National de la Recherche Scientifique (CNRS) of France, and the University of Hawaii. This work is based in part on data products produced at TERAPIX and the Canadian Astronomy Data Centre as part of the CFHT Legacy Survey, a collaborative project of NRC and CNRS. Support from NSF grants AST-0449678 and AST-0807974, and NASA LTSA NNG04GE12G is acknowledged.

REFERENCES

- Adelberger, K. L., & Steidel, C. C. 2000, *ApJ*, **544**, 218
- Avni, Y., & Bahcall, J. N. 1980, *ApJ*, **235**, 694
- Barmby, P., Huang, J.-S., Ashby, M. L. N., Eisenhardt, P. R. M., Fazio, G. G., Willner, S. P., & Wright, E. L. 2006, *ApJ*, **642**, 126
- Baugh, C. M. 2006, *Rep. Prog. Phys.*, **69**, 3101
- Benitez, N. 2000, *ApJ*, **536**, 571
- Bertin, E., & Arnouts, S. 1996, *A&AS*, **117**, 393
- Blanton, M. R., & Roweis, S. 2007, *AJ*, **133**, 734
- Bouwens, R. J., et al. 2009, *ApJ*, **705**, 936
- Brammer, G. B., & van Dokkum, P. G. 2007, *ApJ*, **654**, L107
- Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, *ApJ*, **686**, 1503
- Brammer, G. B., et al. 2009, *ApJ*, **706**, L173
- Brusa, M., et al. 2009, *ApJ*, **716**, 348
- Bruzual, G., & Charlot, S. 2003, *MNRAS*, **344**, 1000
- Calzetti, D., Armus, L., Bohlin, R. C., Kinney, A. L., Koornneef, J., & Storchi-Bergmann, T. 2000, *ApJ*, **533**, 682
- Calzetti, D., Kinney, A. L., & Storchi-Bergmann, T. 1994, *ApJ*, **429**, 582
- Capak, P., et al. 2007, *ApJS*, **172**, 99
- Cole, S., et al. 2001, *MNRAS*, **326**, 255
- Conselice, C. J., Blackburne, J. A., & Papovich, C. 2005, *ApJ*, **620**, 564
- Dale, D. A., & Helou, G. 2002, *ApJ*, **576**, 159
- Davis, M., et al. 2003, *Proc. SPIE*, **4834**, 161
- Davis, M., et al. 2007, *ApJ*, **660**, L1
- Dickinson, M., Papovich, C., Ferguson, H. C., & Budavári, T. 2003, *ApJ*, **587**, 25
- Draine, B. T., & Li, A. 2007, *ApJ*, **657**, 810
- Drory, N., Salvato, M., Gabasch, A., Bender, R., Hopp, U., Feulner, G., & Pannella, M. 2005, *ApJ*, **619**, L111
- Dunlop, J. S., Cirasuolo, M., & McLure, R. J. 2007, *MNRAS*, **376**, 1054
- Efstathiou, G., Ellis, R. S., & Peterson, B. A. 1988, *MNRAS*, **232**, 431
- Elsner, F., Feulner, G., & Hopp, U. 2008, *A&A*, **477**, 503
- Elvis, M., et al. 2009, *ApJS*, **184**, 158
- Erben, T., et al. 2009, *A&A*, **493**, 1197
- Eyles, L. P., Bunker, A. J., Ellis, R. S., Lacy, M., Stanway, E. R., Stark, D. P., & Chiu, K. 2007, *MNRAS*, **374**, 910
- Fontana, A., et al. 2006, *A&A*, **459**, 745
- Fontanot, F., De Lucia, G., Monaco, P., Somerville, R., & Santini, P. 2009, *MNRAS*, **397**, 1776
- Förster Schreiber, N. M., et al. 2006, *AJ*, **131**, 1891
- Franx, M., van Dokkum, P. G., Förster Schreiber, N. M., Wuyts, S., Labbé, I., & Toft, S. 2008, *ApJ*, **688**, 770
- Fukugita, M., Shimasaku, K., & Ichikawa, T. 1995, *PASP*, **107**, 945
- Gehrels, N. 1986, *ApJ*, **303**, 336
- Genzel, R., et al. 1998, *ApJ*, **498**, 579
- Giavalisco, M., et al. 2004, *ApJ*, **600**, 103
- Hathi, N. P., Malhotra, S., & Rhoads, J. E. 2008, *ApJ*, **673**, 686
- Hildebrandt, H., Pielorz, J., Erben, T., van Waerbeke, L., Simon, P., & Capak, P. 2009, *A&A*, **498**, 725
- Ilbert, O., et al. 2009, *ApJ*, **690**, 1236
- Kajisawa, M., et al. 2009, *ApJ*, **702**, 1393
- Kendall, M. G., & Stuart, A. 1961, *The Advanced Theory of Statistics*, Vol. 2 (London: Griffin & Griffin)
- Kennicutt, R. C. 1998, *ApJ*, **498**, 541
- Kriek, M., van Dokkum, P. G., Labbé, I., Franx, M., Illingworth, G. D., Marchesini, D., & Quadri, R. F. 2009, *ApJ*, **700**, 221
- Kriek, M., et al. 2007, *ApJ*, **669**, 776
- Kriek, M., et al. 2008, *ApJ*, **677**, 219
- Kroupa, P. 2001, *MNRAS*, **322**, 231
- Labbé, I., et al. 2003, *AJ*, **125**, 1107
- Laird, E. S., Nandra, K., Hobbs, A., & Steidel, C. C. 2006, *MNRAS*, **373**, 217
- Laird, E. S., et al. 2009, *ApJS*, **180**, 102
- Lee, S.-K., Idzi, R., Ferguson, H. C., Somerville, R. S., Wiklind, T., & Giavalisco, M. 2009, *ApJS*, **184**, 100
- Lilly, S. J., et al. 2007, *ApJS*, **172**, 70
- Lutz, D., Spoon, H. W. W., Rigopoulou, D., Moorwood, A. F. M., & Genzel, R. 1998, *ApJ*, **505**, L103
- Magdis, G. E., Rigopoulou, D., Huang, J.-S., & Fazio, G. G. 2010, *MNRAS*, **401**, 1521
- Maíz Apellániz, J. 2006, *AJ*, **131**, 1184
- Mancini, C., Matute, I., Cimatti, A., Daddi, E., Dickinson, M., Rodighiero, G., Bolzonella, M., & Pozzetti, L. 2009, *A&A*, **500**, 705
- Maraston, C. 2005, *MNRAS*, **362**, 799
- Maraston, C., Pforr, J., Renzini, A., Daddi, E., Dickinson, M., Cimatti, A., & Tonini, C. 2010, *MNRAS*, **407**, 830
- Marchesini, D., van Dokkum, P. G., Förster Schreiber, N. M., Franx, M., Labbé, I., & Wuyts, S. 2009, *ApJ*, **701**, 1765
- Marchesini, D., et al. 2007, *ApJ*, **656**, 42
- McLure, R. J., Cirasuolo, M., Dunlop, J. S., Foucaud, S., & Almaini, O. 2009, *MNRAS*, **395**, 2196
- Mobasher, B., et al. 2005, *ApJ*, **635**, 832
- Moster, B. P., Somerville, R. S., Newman, J. A., & Rix, H.-W. 2010, *ApJ*, submitted (arXiv:1001.1737)
- Murphy, E. J., Chary, R.-R., Alexander, D. M., Dickinson, M., Magnelli, B., Morrison, G., Pope, A., & Teplitz, H. I. 2009, *ApJ*, **698**, 1380
- Muzzin, A., Marchesini, D., van Dokkum, P. G., Labbé, I., Kriek, M., & Franx, M. 2009, *ApJ*, **701**, 1839
- Muzzin, A., van Dokkum, P. G., Kriek, M., Labbé, I., Cury, I., Marchesini, D., & Franx, M. 2010, *ApJ*, in press (arXiv:1003.3479)
- Nandra, K., & Pounds, K. A. 1994, *MNRAS*, **268**, 405
- Ouchi, M., et al. 2004, *ApJ*, **611**, 660
- Papovich, C., et al. 2006, *ApJ*, **640**, 92
- Papovich, C., et al. 2007, *ApJ*, **668**, 45
- Pérez-González, P. G., et al. 2008, *ApJ*, **675**, 261
- Quadri, R. F., et al. 2007, *AJ*, **134**, 1103
- Reddy, N. A., & Steidel, C. C. 2009, *ApJ*, **692**, 778
- Reddy, N. A., Steidel, C. C., Erb, D. K., Shapley, A. E., & Pettini, M. 2006, *ApJ*, **653**, 1004
- Rigby, J. R., et al. 2008, *ApJ*, **675**, 262
- Rodighiero, G., Cimatti, A., Franceschini, A., Brusa, M., Fritz, J., & Bolzonella, M. 2007, *A&A*, **470**, 21
- Salpeter, E. E. 1955, *ApJ*, **121**, 161
- Sandage, A., Tammann, G. A., & Yahil, A. 1979, *ApJ*, **232**, 352
- Sanders, D. B., et al. 2007, *ApJS*, **172**, 86
- Schechter, P. 1976, *ApJ*, **203**, 297
- Scoville, N., et al. 2007, *ApJS*, **172**, 1
- Shapley, A. E., Steidel, C. C., Adelberger, K. L., Dickinson, M., Giavalisco, M., & Pettini, M. 2001, *ApJ*, **562**, 95
- Somerville, R. S., Hopkins, P. F., Cox, T. J., Robertson, B. E., & Hernquist, L. 2008, *MNRAS*, **391**, 481
- Somerville, R. S., & Primack, J. R. 1999, *MNRAS*, **310**, 1087
- Somerville, R. S., Primack, J. R., & Faber, S. M. 2001, *MNRAS*, **320**, 504
- Stark, D. P., Bunker, A., Ellis, R. S., Eyles, L. P., & Lacy, M. 2007, *ApJ*, **659**, 84

- Stark, D. P., Ellis, R. S., Bunker, A., Bundy, K., Targett, T., Benson, A., & Lacy, M. 2009, [ApJ](#), **697**, 1493
- Steidel, C. C., Adelberger, K. L., Shapley, A. E., Pettini, M., Dickinson, M., & Giavalisco, M. 2003, [ApJ](#), **592**, 728
- Steidel, C. C., Hunt, M. P., Shapley, A. E., Adelberger, K. L., Pettini, M., Dickinson, M., & Giavalisco, M. 2002, [ApJ](#), **576**, 653
- Taylor, E. N., et al. 2009, [ApJ](#), **694**, 1171
- Treister, E., et al. 2009, [ApJ](#), **706**, 535
- van Dokkum, P. G., et al. 2006, [ApJ](#), **638**, 59
- van Dokkum, P. G., et al. 2009, [PASP](#), **121**, 2
- van Dokkum, P. G., et al. 2010, [ApJ](#), **709**, 1018
- Whitaker, K. E., et al. 2010a, [ApJ](#), **719**, 1715
- Whitaker, K. E., et al. 2010b, [ApJ](#), submitted
- Wiklind, T., Dickinson, M., Ferguson, H. C., Giavalisco, M., Mobasher, B., Grogin, N. A., & Panagia, N. 2008, [ApJ](#), **676**, 781
- Wolf, C., Meisenheimer, K., Rix, H.-W., Borch, A., Dye, S., & Kleinheinrich, M. 2003, [A&A](#), **401**, 73
- Wuyts, S., Labbé, I., Förster Schreiber, N. M., Franx, M., Rudnick, G., Brammer, G. B., & van Dokkum, P. G. 2008, [ApJ](#), **682**, 985
- Wuyts, S., et al. 2007, [ApJ](#), **655**, 51
- Yan, H., Dickinson, M., Giavalisco, M., Stern, D., Eisenhardt, P. R. M., & Ferguson, H. C. 2006, [ApJ](#), **651**, 24