Searching for massive galaxies at $z \geq 3.5$ in GOODS-North[★]

C. Mancini^{1,2,4}, I. Matute², A. Cimatti³, E. Daddi⁴, M. Dickinson⁵, G. Rodighiero⁶, M. Bolzonella⁷, and L. Pozzetti⁷

¹ Dipartimento di Astronomia e Scienza dello Spazio, Università degli Studi di Firenze, Largo E. Fermi 3 50125 Firenze, Italy
 e-mail: mancini@arcetri.astro.it

² Osservatorio Astrofisico di Arcetri (OAF), INAF-Firenze, Largo E. Fermi 5, 50125 Firenze, Italy

³ Dipartimento di Astronomia, Università di Bologna, via Ranzani 1, 40127 Bologna, Italy

⁴ CEA-Saclay, DSM/DAPNIA/Service d'Astrophysique, 91191 Gif-Sur Yvette Cedex, France

⁵ NOAO, 950 N. Cherry Ave. PO 26732, Tucson, AZ 85726-6732, USA

⁶ Dipartimento di Astronomia, Università di Padova, Vicolo Osservatorio 2, 35122 Padova, Italy

⁷ INAF-Bologna, via Ranzani, 40127 Bologna, Italy

Received 17 July 2008 / Accepted 21 January 2009

ABSTRACT

Aims. We constrain the space density and properties of massive galaxy candidates at redshifts of $z \geq 3.5$ in the Great Observatories Origin Deep Survey North (GOODS-N) field. By selecting sources in the *Spitzer* + IRAC bands, a sample highly complete in stellar-mass is assembled, including massive galaxies that are very faint in the optical/near-IR bands and would be missed by samples selected at shorter wavelengths.

Methods. The $z \geq 3.5$ sample was selected to $m_{AB} = 23$ mag at $4.5 \mu\text{m}$ using photometric redshifts obtained by fitting the galaxies spectral energy distribution at optical, near-IR bands, and IRAC bands. We also require that the brightest band (in AB scale) in which candidates are detected is the IRAC $8.0 \mu\text{m}$ band to ensure that the near-IR $1.6 \mu\text{m}$ (rest-frame) peak is falling in or beyond this band.

Results. We found 53 $z \geq 3.5$ candidates, of masses in the range $M_{\star} \sim 10^{10}$ – $10^{11} M_{\odot}$. At least $\sim 81\%$ of these galaxies are missed by traditional Lyman Break selection methods based on ultraviolet light. *Spitzer* + MIPS emission is detected for 60% of the sample of $z \geq 3.5$ galaxy candidates. Although in some cases this might suggest a residual contamination from lower redshift star-forming galaxies or Active Galactic Nuclei, 37% of these objects are also detected in the sub-mm/mm bands in SCUBA, AzTEC, and MAMBO surveys, and have properties fully consistent with vigorous starburst galaxies at $z \geq 3.5$. The comoving number density of galaxies with stellar masses of above $5 \times 10^{10} M_{\odot}$ (a reasonable stellar-mass completeness limit for our sample) is $2.6 \times 10^{-5} \text{ Mpc}^{-3}$ (using the volume within $3.5 < z < 5$), and the corresponding stellar mass density is $\sim (2.9 \pm 1.5) \times 10^6 M_{\odot} \text{ Mpc}^{-3}$, or about 3% of the local density above the same stellar mass limit. For the subsample of MIPS-undetected galaxies, we measure a number density of $\sim 0.97 \times 10^{-5} \text{ Mpc}^{-3}$ and a stellar mass density of $\sim (1.15 \pm 0.7) \times 10^6 M_{\odot} \text{ Mpc}^{-3}$. Even in the unlikely case that these are all truly quiescent galaxies, this would imply an increase in the space density of passive galaxies by a factor of ~ 15 between $z \sim 4$ and $z = 2$, and by ~ 100 to $z = 0$.

Key words. cosmology: observations – galaxies: formation – Galaxy: evolution – infrared: galaxies – galaxies: high-redshift – galaxies: photometry

1. Introduction

The question of how galaxy mass is assembled at early cosmological epochs remains open. In a Λ CDM Universe, the standard formation scenario predicts hierarchical growth of structures. Local galaxies formed by repeated merging events. The first semi-analytic galaxy-formation models (Kauffmann & Charlot 1998) predicted that the assembly of the most massive systems was completed only during most recent epochs ($z < 1$). Nevertheless a significant population of massive galaxies has been found beyond $z \sim 1$. Up to $z \approx 2.5$, the most massive objects spectroscopically confirmed are similar to local early-type galaxies (ETGs): old (age of ~ 1 Gyr), passively evolving, and of mass $M_{\star} > 10^{11} M_{\odot}$ (Cimatti et al. 2004; McCarthy et al. 2004; Daddi et al. 2005; Saracco et al. 2005; Kriek et al. 2006).

At higher redshift the availability of multiwavelength data from deep surveys enabled searches for massive galaxy candidates to $z \approx 5$ – 6.5 . The dropout-technique pioneered by Steidel et al. (1996) to identify Lyman-break galaxies (LBGs) at $z \approx 3$, was also successful in identifying blue star-forming galaxies at higher redshifts ($z \sim 4$ – 6 ; Steidel et al. 1999; Dickinson 1998; Giavalisco et al. 2004b; McLure et al. 2006; Yan et al. 2006).

A substantial amount of spectroscopic confirmation is available for LBGs at $z \sim 5$ – 6 (Bunker et al. 2003; Stanway et al. 2004; Dickinson et al. 2004; Eyles et al. 2007; Stark et al. 2007; Vanzella et al. 2008). This allowed early estimates of the stellar mass density at those redshifts. For instance, Yan et al. (2006) measured the comoving stellar mass density at $z \sim 6$ in the total GOODS field and found good agreement with the simulations of Night et al. (2006). The stellar mass density found in the GOODS-S field by Eyles et al. (2007) for a sample of *i*-dropout LBGs in the same redshift range was consistent with the Yan et al. (2006) results, despite being higher by a factor of 4

[★] Figures 13 to 15 are only available in electronic form at <http://www.aanda.org>

(~ 2 , if the different IMFs are taken into account) than the Bower et al. (2006) semi-analytical model predictions. The general picture remains controversial and further studies are necessary to better constrain galaxy formation models. The Lyman-break technique can only be used to select actively star-forming galaxies with sufficiently luminous UV emission to allow their detection and color selection. Galaxies, and especially massive ones, could be faint or undetected in the optical and near-infrared bands. These sources could be either old systems with small quantities of dust and low levels of star-formation rate, or dust-reddened starburst galaxies. If these $z \geq 3.5$ objects really exist, they would be too faint to be studied spectroscopically with the instrumentation currently available, and the only possible approach that can be used to constrain their number density and their properties remains a photometric one, based on the Spectral Energy Distribution (SED) fitting analysis.

The most popular criterion used so far to select high- z galaxies with red rest-frame colors is based on the red-shifted 4000 Å/Balmer break, i.e., the typical features of galaxies with evolved stellar populations. It was pioneered by Franx et al. (2003) to detect “distant red galaxies” (DRGs) at $z > 2$ ($J_s - K_s > 2.3$, AB system), and extended by Brammer & van Dokkum (2007) to identify DRGs to $z \sim 3.7$ ($H - K > 0.9$, AB system). By comparing the properties and space densities of DRGs at $z \sim 3.7$ and $z \sim 2.4$ in the same field, they found that the stellar mass density was about a factor of 5 lower in the higher redshift bin.

The availability of *Spitzer* data has enabled us to search for galaxies with red rest-frame optical colors at high redshift, and to constrain their stellar masses (Fontana et al. 2006; Dunlop et al. 2007; Rodighiero et al. 2007; Wiklind et al. 2008; Eyles et al. 2005; Pérez-González et al. 2008). Some of the objects identified may be among the most massive stellar systems found at $z \sim 4-5$. Rodighiero et al. (2007) extracted a *Spitzer*+IRAC selected sample in the GOODS-South field, limited to galaxies undetected in the optical and close to the detection limit in the K -band. Their criterion complemented those adopted by previous studies (i.e., the $K < 23.5$ selection used by Dunlop et al. 2007) and was designed to identify high- z massive galaxies missed by conventional selection techniques based on optical and near-IR observations. They detected a potential population of optically obscured massive galaxies at $z \geq 4$. In the same field, Wiklind et al. (2008) found 11 K -selected $z \geq 5$ massive and evolved galaxy candidates. They selected these objects by applying color criteria based on the identification of the 4000 Å/Balmer break between the K and $3.6 \mu\text{m}$ passbands, in combination with spectral energy distribution (SED) fitting.

We note that in both of these works approximately half of the high- z massive candidates are detected at $24 \mu\text{m}$. The $24 \mu\text{m}$ emission might originate in obscured AGN activity at high redshift, and/or Polycyclic Aromatic Hydrocarbon (PAH) emission in dusty star forming galaxies at lower redshift. One must consider the degeneracy between the effects of reddening and variations in redshift when selecting high redshift sources using photometric data. Studies of sub-mm galaxies (SMGs) illustrated that moderate $24 \mu\text{m}$ emission can arise in massive starburst galaxies at high redshift ($z \geq 3.5$; Daddi et al. 2009; Greve et al. 2008; Perera et al. 2008). Hence a primary difficulty in selecting bona fide high- z galaxies by means of photometric redshifts is to remove in an efficient way contamination by dust-reddened galaxies at low redshift. An example of the difficulties involved was shown by the case of the galaxy HUDF-JD2 in the Hubble Ultra Deep Field. Mobasher et al. (2005) proposed that

this galaxy is a massive ($M \simeq 6 \times 10^{11} M_\odot$), old galaxy with small amounts of dust reddening (with an age of $\sim 10^8$ yr and formation redshift of $z_{\text{form}} > 9$) at $z \simeq 6.5$. However, other studies (e.g., Dunlop et al. 2007; Rodighiero et al. 2007; Chary et al. 2007) suggested that HUDF-JD2 is instead a galaxy at lower redshift, with significant amounts of dust reddening ($z \simeq 2.2$ and $A_V \simeq 3.8$, Dunlop et al. 2007).

The principal aim of this paper is to study the properties and space density of massive galaxies at $z \geq 3.5$ in the Great Observatory Origin Deep Survey-North (GOODS-N). This survey is especially suited to our purposes because of the quality and depth of the available multiwavelength data sets. In particular, it is of wider areal coverage at sub-millimeter wavelengths than GOODS-S. While similar searches have been completed in the GOODS-South field, no study has yet been performed in the northern hemisphere field of the GOODS Survey. We selected the sample from the IRAC channel 2 ($4.5 \mu\text{m}$) data for two reasons. First, we wished to identify the most massive systems, and at $z \geq 3.5$ the $4.5 \mu\text{m}$ band is sensitive to rest-frame near-IR emission, which is correlated well with the galaxy stellar mass (Kauffmann & Charlot 1998). Second, the $4.5 \mu\text{m}$ band selection allows us to construct a more complete sample in terms of stellar mass than an optical or near-IR selection. This is an important difference from previous work: we attempted to construct a sample that was as complete as possible, including objects undetected in optical and near-IR bands, and hence missed by the optical or K -band selection criteria used so far in the literature (Drory et al. 2004; Fontana et al. 2006; Dunlop et al. 2007; Wiklind et al. 2008; Brammer & van Dokkum 2007). As already mentioned, Rodighiero et al. (2007) also used an IRAC band ($3.6 \mu\text{m}$) selection, but their additional conditions about the lack of optical emission and the faint near-IR detection ($K > 23$, AB system) ensured that their sample was incomplete by construction.

The structure of the paper is as follows. In Sect. 2, we present the optical, near-IR and IR data available for the GOODS-N field. In Sect. 3 we describe the IRAC selection and the photometric data of the selected sample. Section 4 describes the results of the SED fitting analysis, including the measurement of photometric redshift and stellar mass. In Sect. 5, we discuss the selection of $z \geq 3.5$ galaxy candidates. In the same section, we discuss the MIPS- $24 \mu\text{m}$ and sub-mm/mm detection of part of the sample. In Sect. 6, we study the effects of incompleteness and the bias against low mass galaxies in our magnitude-limited sample. We provide estimates of the comoving number density and stellar mass density for both the full sample and for the subsample of MIPS-undetected galaxies. Section 7 presents the summary and conclusions.

Throughout this paper, we adopt the following cosmological parameters $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_\Lambda = 0.7$, and $\Omega_m = 0.3$, and all magnitudes are given in the AB system.

2. The GOODS-N data

This work is based on the analysis of the Great Observatory Origins Deep Survey – North (GOODS-N) data. The GOODS-N is a multiwavelength deep survey centered on the Hubble Deep Field North (HDFN $12^{\text{h}}36^{\text{m}}55^{\text{s}}$, $+62^\circ14'15''$, Giavalisco et al. 2004b) of area $\sim 10' \times 16'$. The field is part of the GOODS *Spitzer* Legacy Program (PI: Dickinson) and the *Hubble Space Telescope* (HST) Treasury Program (PI: Giavalisco), divided into two fields (GOODS-N and GOODS-S), from the northern and southern hemispheres (see Dickinson et al. 2003). The GOODS-N field observations are for wavelengths between the

X-rays and sub-mm. In particular, for the analysis presented in this paper, we used the following data sets:

- *U*-band imaging, taken at the Kitt Peak National Observatory (KPNO) 4 m telescope, from the Hawaii Hubble Deep Field North (H-HDFN) survey data sets (Capak et al. 2004).
- Optical imaging from the *Hubble Space Telescope* (HST) with the Advanced Camera for Survey (ACS) in the *F435W* (*B*), *F606W* (*V*), *F775W* (*i*), *F850LP* (*z*) passbands (Giavalisco et al. 2004b).
- Near-infrared imaging, in the *J*, *H*, and *Ks* bands, from the KPNO-4 m telescope (Dickinson private communication).
- Infrared observations from the *Spitzer* Space Telescope (Werner et al. 2004) + Infra-Red Array Camera (IRAC; Fazio et al. 2004).
- 24 μm observations from the *Spitzer* Space Telescope and Multi-band Imaging Photometer for *Spitzer* (MIPS; Rieke et al. 2004).

We describe specific properties of the data below.

2.1. KPNO 4 m *U*-band imaging

The H-HDFN Survey is a deep multi-color imaging survey of 0.2 degrees², centered on the Hubble Deep Field North. The data consist of deep images in the *U*, *B*, *V*, *R*, *I*, and *z'* bands. Of these, we used only the *U*-band images, due to the availability of HST/ACS imaging from *B* to *z* band. The *U*-band observations were taken in March 2002, using the KPNO-4 m telescope with the MOSAIC prime focus camera (Muller et al. 1998), covering a field of view of $36' \times 36'$ with a seeing of $1''.26$. We analyzed the data that overlap the GOODS-N field (an area of about 165 arcmin² around the HDF-N). The *U* image reaches a 5σ sensitivity limit of 27.1 mag.

2.2. ACS+HST optical imaging

The GOODS-N optical data from the Advanced Camera for Surveys (ACS), onboard HST, were acquired for the *F435W* (*B*), *F606W* (*V*), *F775W* (*i*), and *F850LP* (*z*) passbands (v1.0 data products). The 5σ point-source sensitivity limits are 28.5, 28.8, 28.1, and 27.6 for each passband respectively, scaled from values reported by Giavalisco et al. (2004b) to the full exposure time of ACS v1.0 images. The HST/ACS high-resolution data is organized into 17 images, called “sections”. Each section is an image 8192×8192 pixels in size, drizzled from the native ACS pixel scale ($0''.05/\text{pixel}$) to an image subsampled at $0''.03/\text{pixel}$. The enormous size ($\sim 40\,000 \times 40\,000$ pixels) of the combined HST images in the field meant that source extraction was an extremely intensive computing task. We therefore rebinned the final image using an 8×8 pixel kernel and insisting on flux conservation. The new rebinned mosaic obtained in this way has a size of 5120×5120 pixels, and a pixel scale of $0''.24/\text{pixel}$.

2.3. KPNO 4 m FLAMINGOS near-infrared imaging

We complemented the public data with near-IR imaging taken at the KPNO-4m telescope with the Florida Multi-object Imaging Near-IR Grism Observational Spectrometer (FLAMINGOS; observations completed by Dickinson et al.). The FLAMINGOS images in the *J*, *H*, and *Ks* passbands, are mosaics of data taken at different pointings and orientations to cover the entire

GOODS-N field. The average exposure times per band in the main part of the GOODS-N field area were 20 250 s, 15 770 s, and 35 440 s for *J*-, *H*- and *Ks*-band, respectively. The three images achieved 5σ point-source sensitivities of 24.03, 23.77, and 23.81 mag, respectively. The pixel scale was $0''.316/\text{pixel}$, and seeing $1''.27$, $1''.2$, and $1''.2$, for the *J*, *H*, and *Ks* passbands, respectively.

We performed some photometric checks before measuring the source fluxes. We verified the zero-point reliability of FLAMINGOS images by comparing the bright and unsaturated star magnitudes with those of stars from the 2MASS catalogue.

2.4. Spitzer imaging at 3.6 to 24 μm

The infrared imaging was acquired by *Spitzer* Space Telescope + IRAC between 3.6 and 8.0 μm , and *Spitzer* + MIPS at 24 μm .

The four IRAC channels, centered on 3.6 μm , 4.5 μm , 5.8 μm , and 8.0 μm , have formal 5σ limits for isolated point sources of 26.1, 25.5, 23.5, and 23.4, respectively. The mean FWHM of a point source in the IRAC band images was $1''.66$, $1''.72$, $1''.88$, and $1''.98$, for channels 1–4 respectively. The initial image pixel-scale of $1''.22$ was reduced to $0''.6/\text{pixel}$, after the dithering and drizzling process.

The MIPS public data set includes calibrated maps and a catalogue of 24 μm sources with flux densities of $S_{24} > 80 \mu\text{Jy}$. The PSF was generated from isolated sources in the image, and renormalized based on the aperture correction published in the MIPS data Handbook (v2.1, Sect. 3.7.5, Table 3.12). An independent PSF algorithm was then executed to extend the 24 μm sample to fainter magnitudes (see Rodighiero et al. 2006). In this procedure, the detection threshold was extended to $S_{24} > 35 \mu\text{Jy}$. Since the 24 μm information can constrain the SFR more accurately (Chary & Elbaz 2001) and the nature of high-redshift sources, we also used a deeper 24 μm catalogue, reaching a 5σ limit of $20 \mu\text{Jy}$ (Chary et al. 2008, in preparation; Daddi et al. 2007).

3. Sample selection

To obtain an accurate object-flux measurement in all bands, we performed aperture photometry in each band using the *SExtractor* software (Bertin & Arnouts 1996). The initial catalogue was constructed by selecting sources in the 4.5 μm public image obtained with *Spitzer* Space Telescope + IRAC, to a limiting magnitude of $m_{4.5} = 23.0$ (2.3 μJy). Simulations demonstrated that a sample selected in this way was complete at the 80% level to $m_{4.5} \sim 23.0$ (Dickinson, private communication), where 20% of the sources were lost due to blending.

The infrared selection allowed us to search for massive galaxies to high redshifts. For galaxies at $3.5 < z < 7$, the rest-frame light emitted at $\lambda > 0.6 \mu\text{m}$ is red-shifted into the IRAC 4.5 μm band, and the 4.5 μm selection is therefore directly sensitive to stellar mass, rather than ongoing star-formation activity. From a technical point of view, the choice of selecting the sample at 4.5 μm represents the optimal compromise amongst the IRAC bands in terms of image quality, blending problems, and sensitivity.

To detect galaxies in the 4.5 μm IRAC band and perform the photometry, we used *SExtractor* with the following set of parameters. The detection limit was set to be $\sim 1\sigma$ above the sky background. A Gaussian filter was used to improve the detection of faint, extended objects. Because of the significant crowding of IRAC images, and the large PSF, the minimum contrast pa-

parameter was set to be a small value of 5×10^{-9} , to improve the source deblending.

We measured fluxes in $4''.0$ diameter apertures, since this size allowed us to minimize the uncertainties in the photometry for the *Spitzer* +IRAC bands (as also suggested by the SWIRE team¹). In the $4.5 \mu\text{m}$ IRAC filter, we applied an aperture correction of 0.25 mag, which we computed independently from point-source objects, by measuring the total flux in $8''.0$ diameter apertures.

Our final IRAC selected catalogue contained 4142 objects ($m_{4.5} < 23.0$).

3.1. Multi-band photometry

We searched for the counterparts of the 4142 IRAC-selected objects in the other data sets. We detected the sources and measured the photometry independently in each data set and associated counterparts with the $4.5 \mu\text{m}$ IRAC detected galaxies using a search radius of $1''.0$.

For the *U*-band, we measured the fluxes inside $4''.0$ (diameter) apertures. We applied an aperture-correction of 0.1 mag, computed by considering the total flux in an $8''.0$ diameter aperture. For undetected sources, we used the 3σ flux upper-limit of 26.7 mag. This value was computed by measuring the standard deviation in the sky signal at random positions across the field.

For the HST/ACS bands, circular apertures of diameters $2''.0$ were used. The following aperture corrections for point sources were found by measuring the total flux in a $5''.0$ diameter-aperture, for the *B*, *V*, *i*, and *z* bands: 0.050, 0.040, 0.030, and 0.045 mag, respectively. For undetected objects, we obtained 3σ upper limits as described above for the *U*-band. We measured the upper limits to be 26.3, 26.9, 26.3, and 25.5, for the *B*, *V*, *i*, and *z* bands, respectively.

The near-IR magnitudes were measured in $4''.0$ diameter apertures, and corrected to total magnitudes (within $8''.0$ diameter) by subtracting 0.1, 0.14, and 0.15 mag for the *J*-, *H*-, and *Ks*-band, respectively. The 3σ detection limits were 23.3, 23.0, and 23.1, for the *J*-, *H*- and *Ks*-band, respectively.

In the other IRAC channels, we used the same aperture diameters as in IRAC- $4.5 \mu\text{m}$ ($4''.0$), and the same approach in computing aperture corrections. The final IRAC magnitudes were obtained by applying aperture corrections of 0.23, 0.25, 0.40, and 0.44 mag for 3.6, 4.5, 5.8, and $8.0 \mu\text{m}$ bands, respectively.

To check the reliability of our photometry, we compared the colors of stars with those expected, based on the models of Lejeune et al. (1997) in color-color diagrams involving one band from each one of the three data sets studied (ACS/HST optical, Near-Infrared/FLAMINGOS and infrared/IRAC-*Spitzer*). The diagram $(b - J)$ versus $(J - m_{3.6})$ is shown in Fig. 1. The star sequence is clearly visible on the left side of the panel. As it is clear from the figure, we found good agreement between the observed (green circles) and expected colors (red asterisks from the “corrected” templates of Lejeune et al. 1997) for spectroscopically identified stars (taken from the public database of the TKRS survey Wirth et al. 2004). Therefore, because of the good agreement between the optical, near-IR, and IR data sets, we were not required to apply any further photometric correction to our data magnitudes.

The photometric errors given by *SExtractor* software are generally underestimated, because the software does not take account of any possible correlation between neighboring pixels

Fig. 1. The color-color diagram $(b - J)$ versus $(J - m_{3.6})$. The star sequence is clearly visible on the left side of the diagram. We found a very good agreement between the observed colors (green circles) and the expected ones (red asterisks from the “empirically corrected” templates of Lejeune et al. 1997) for spectroscopically identified stars (public database of the TKRS survey Wirth et al. 2004).

caused by the image processing. Hence, we applied corrections to the photometric errors separately in each data set. For the four IRAC channels, we added 0.1 mag in quadrature to each of the *SExtractor* errors. This was a conservative choice to account for uncertainties in both the photometric zeropoints and the aperture corrections, as suggested by Maraston et al. (2006). In the other filters, we measured the signal at various random sky positions in the field, within the same aperture used to perform galaxy photometry in each band. By comparing the standard deviation in these measurements with the true errors in the photometry of faint galaxies as inferred by *SExtractor*, we derived scaling factors, which we applied to the formal errors reported in the *SExtractor* catalogues.

4. Spectral energy distribution (SED) fitting analysis

4.1. Photometric redshifts estimate

To estimate the photometric redshifts of our selected galaxies, we used the *Hyperz* code (Bolzonella et al. 2000). This code compares the multiwavelength photometry of each source with a database of theoretical and/or empirical templates at different redshifts. The best-fit solution, for a given set of templates, is derived by a standard χ^2 minimization.

To determine the photometric redshifts, we excluded the $24 \mu\text{m}$ band, because the stellar-population templates used did not account for dust emission. For a similar reason, we also decided to omit the $8 \mu\text{m}$ band, which for a given survey includes thermal dust and PAH (i.e., non-stellar) emission for low redshift galaxies and can be contaminated by star formation or AGN at high redshift (Daddi et al. 2007). We used the template library of the observed SEDs constructed by Coleman et al. (1980, hereafter CWW), extended towards UV

¹ http://data.spitzer.caltech.edu/popular/swire/20050603_enhanced_v1/

and IR wavelengths by means of [Bruzual & Charlot \(2003\)](#) synthetic spectra, as explained in the *Hyperz* User's manual². This set of templates includes four types of spectra with characteristics related to the main local types of galaxies: elliptical/lenticular (E/S0), spiral (Sbc and Scd), and irregular (Im). The CWW library was complemented with a spectrum of a young starburst galaxy with a constant star-formation rate (SFR) and age of 0.1 Gyr, from the [Bruzual & Charlot \(2003\)](#) models.

We took into account the dust extinction by applying the [Calzetti et al. \(2000\)](#) law. The extinction parameter (A_V) was allowed to vary from 0 to 0.8 in steps of $d(A_V) = 0.1$. Only a small range of A_V was allowed, since the CWW templates already included intrinsically reddened SEDs. The permitted redshift range spanned the wide interval of $0 < z < 10$, in steps of $d(z) = 0.07$.

We used the prescription of [Madau \(1995\)](#) to represent the average Lyman- α forest opacity as a function of wavelength and redshift.

4.2. Photometric redshift reliability

We checked the reliability of our photometric redshifts for 836 objects in the GOODS-N field with high quality spectroscopic redshifts. Unobscured AGN were excluded from the list by means of the X-ray information from the publicly available X-ray catalogue of the Chandra Deep Field-North Survey ([Alexander et al. 2003](#)). The largest number of spectroscopic redshifts (721) were taken from the public database of the Team Keck Treasury Redshift Survey (TKRS; [Wirth et al. 2004](#)), and have an accuracy of 100 km s^{-1} . The median redshift of this survey is rather low ($z = 0.65$). To increase the number of available objects with the highest spectroscopic redshifts, we complemented the TKRS database with galaxies found by [Reddy et al. \(2006\)](#) in the redshift interval $2.0 \lesssim z \lesssim 3.5$ (108). We also used data of five objects out to $z \simeq 5$ from the NASA/IPAC Extragalactic Database³ (NED), and three spectroscopic redshifts from [Daddi et al. \(2009\)](#).

The comparison between photometric (z_{phot}) and spectroscopic (z_{spec}) redshifts is shown in Fig. 2. After removing 5σ outliers of the best-fit solution, we found a combined mean offset of $(z_{\text{spec}} - z_{\text{phot}})/(1 + z_{\text{spec}}) = 0.004$ and a rms scatter of $\sigma [(z_{\text{spec}} - z_{\text{phot}})/(1 + z_{\text{spec}})] = 0.09$. This result was obtained by means of an iterative optimization procedure applied to the *Hyperz* results, which consisted repeating the SED fitting analysis for the same objects, fixing the redshifts to their known spectroscopic values. For each filter, the median offset between empirical and “theoretical” magnitudes (the latter being derived from the best-fit solution SEDs) was obtained. We found the largest photometric offsets for the U -, z -, and K -bands (+0.2, +0.09, +0.14, respectively) and $|\Delta \text{mag}| \sim 0.05$ for the others. By adding these corrections to the observed magnitudes, the rms of the $\Delta z/(1 + z_{\text{spec}})$ quantity was slightly reduced. The same offsets were finally applied to the entire sample when estimating the photometric redshifts.

Since the current study is focuses on galaxies at $z \geq 3.5$, we carefully examined three outliers (#6955, #8734 and #8903) in the photometric-spectroscopic redshift comparison with $z_{\text{spec}} \geq 3.5$. These are indicated by red diamonds in Fig. 2, and information about them is presented in Table 1. In Fig. 3, we show the photometric SEDs of these three objects, and in Fig. 4 we show cutout images from U -band to $24 \mu\text{m}$. For the first

Fig. 2. Comparison between photometric z_{phot} and spectroscopic z_{spec} redshifts for 836 objects in our sample. The $z_{\text{spec}} - z_{\text{phot}}$ relation has a combined mean offset of $(z_{\text{spec}} - z_{\text{phot}})/(1 + z_{\text{spec}}) = 0.004$ with an rms scatter of $\sigma [(z_{\text{spec}} - z_{\text{phot}})/(1 + z_{\text{spec}})] = 0.09$. For the flagged objects (red open diamonds), appearing as catastrophic outliers, we suggest the lower redshift solution, as detailed in the text.

two objects (#6955 and #8734, [Dawson et al. \(2001\)](#) reported Keck spectra containing isolated emission lines with no continuum detection (their quality class 4; see Tables 2 and 3 of that paper), which they identified as weak Ly- α emission. However, object #6955 is detected with a significance level of $>3\sigma$ limit in the U -band, which excludes the proposed $z = 3.826$. Keck/DEIMOS spectra acquired between 2005 and 2006 (Dickinson and Stern, private communication) detected an emission line at $9273 \text{ \AA}$ that is likely to be [OII] $3727 \text{ \AA}$ at $z = 1.488$, which is consistent with our photometric redshift estimate for this object. Object #8734 corresponds to a red point-source in the HST/ACS images, which has a faint, bluer, extended emission to the southeast. This is likely to represent a superposition of a cool Galactic star and a faint galaxy. [Cowie et al. \(2004\)](#) reported that this is a star based on DEIMOS spectroscopy. The faint, extended component is certainly a galaxy, although it is clearly detected in the ACS B -band image, and is therefore unlikely to be at $z = 4.887$ as reported by [Dawson et al. \(2001\)](#). Overall, the photometry at all wavelengths redward of the V -band is dominated by the point-source component, and the photometric SED for this object (Fig. 3) clearly resembles that of a cool, late-type star (probably an M-dwarf). Finally, for object #8903, [Dawson et al. \(2001\)](#) reported $z = 4.762$. However, the ACS $V - i$ and $i - z$ colors of this galaxy imply that it is unlikely to be a V -dropout Lyman break galaxy, since they coincide with the locus of colors for ordinary foreground galaxies. The ACS color images display a faint, red, early-type galaxy, and our photometric redshift estimate is $z_{\text{phot}} = 0.98$, close to that proposed by [Capak et al. \(2004, \$z_{\text{phot}} = 0.89\$ \)](#). Although [Cowie et al. \(2004\)](#) did not provide details about the spectral features that they observed, it is possible that they mistook faint [OII] emission at $z = 0.88$ for faint Ly α at $z = 4.76$.

4.3. Galaxy stellar-mass estimate

We used the *Hyperzmass* code to estimate the mass of each source. The *Hyperzmass* code operates in a similar way to *Hyperz* (see [Pozzetti et al. 2007](#)). In this task, the redshifts were fixed to either the derived photometric values, or spectroscopic ones, when available. The observed data points were fitted using

² <http://webast.ast.obs-mip.fr/hyperz/>

³ <http://nedwww.ipac.caltech.edu/>

Table 1. Outliers of our $z_{\text{phot}} - z_{\text{spec}}$ comparison (Fig. 2).

id	RA ^a	Dec ^a	z_{phot}	z_{spec}^a	Dist ^b
#6955	12:36:40.510	+62:13:34.90	1.48	3.826	0.27
#8734	12:36:37.630	+62:14:53.70	0.68	4.886	0.16
#8903	12:37:21.048	+62:15:01.73	0.98	4.762	0.41

^a The coordinates and the spectroscopic redshifts are from NED database; ^b distance (arcsec) between our coordinates (from *Spitzer* data set) and the NED position.

both the synthetic spectral models of [Bruzual & Charlot \(2003, hereafter BC03\)](#) and [Maraston \(2005b, hereafter MA05\)](#). Both template sets consist of stellar-populations models with star-formation rates (SFRs) that decrease exponentially with time ($\psi(t) = \tau^{-1} \exp^{-t/\tau}$). In our analysis, we considered models with different values of the timescale τ (in unit of Gyr): 0.1, 0.3, 1, 2, 3, 5, 10, 15, 30, and ∞ . The last model corresponds to a stellar population with a constant SFR.

When adopting the BC03 models, we used the [Chabrier \(2003\)](#) initial mass function (IMF), and for the MA05 models, we used the [Kroupa \(2001\)](#) IMF, which is similar to the Chabrier IMF. To compare the different results, we computed a systematic correction to convert the masses obtained with a Kroupa IMF to values corresponding to a Chabrier IMF ($\log M_{\text{Chabrier}} = \log M_{\text{Kroupa}} - 0.04$).

In the SED fitting analysis, the extinction parameter, A_V , was allowed to span the range of values $0 \leq A_V \leq 6$, in steps of 0.1. This wide reddening range was required because the synthetic stellar population models had not been intrinsically reddened in a similar way to the real galaxies. We used $A_V = 6$ as upper limit to the extinction parameter, since values of A_V higher than ~ 5 – 6 mag have not been observed so far in high-redshift galaxies. The galaxy age was allowed to vary as a free parameter, spanning ages between 5×10^6 yr to 11.7×10^9 yr (and always requiring ages less than the age of the Universe at each redshift). To minimize the number of free parameters, we limited the fits to those for solar metallicity.

5. Massive galaxies at $z \geq 3.5$

Evidence for massive galaxies at $z > 3$ has been found by several studies. However, results have not been conclusive because the number of objects is low, and frequently biased against objects that are faint or undetected in optical and near-IR bands. For most of these objects, no spectroscopic redshift information is available. Possible contamination by lower redshift interlopers is also a source of uncertainty. To shed some light on both the nature and number density of the high redshift population of massive galaxies, we selected a final sample of $z_{\text{phot}} \geq 3.5$ candidates, by applying the following selection criteria (summarized in Table 2). Among the initial 4142 IRAC-4.5 μm selected galaxies, we found 190 galaxies with photometric redshifts of $z_{\text{phot}} \geq 3.5$.

All of these objects were carefully inspected for possible blending issues. One important problem in using *Spitzer* data of crowded fields is the blending between nearby sources. Neighboring objects that are well separated in optical and near-IR images, can still be blended in the IRAC bands, due to the limited spatial resolution of the *Spitzer* telescope. In constructing reliably a bona-fide sample of high- z galaxies, we had therefore to take into account that a fraction of the IRAC-selected objects could be affected by blending. We decided to inspect

Fig. 3. SEDs of the outliers in the $z_{\text{spec}} - z_{\text{phot}}$ space (see objects flagged in Fig. 2). Our best fits (black curves) produce photometric redshifts of ~ 1.48 , 0.68 and 0.98 , respectively. We compare these results with the best fits (red curves) obtained by fixing the redshifts to the spectroscopic values taken from the literature ($z = 3.826$, $z = 4.886$, and $z = 4.762$; NED database). Obviously, the black curves fit the observed data better than the red ones. See also Table 1.

visually all the $z \geq 3.5$ objects and excluded from the sample any sources that the *SExtractor* software was unable to deblend. To identify them, we visually examined the *SExtractor* “aperture image” for each filter, which was a “check-image” provided as output by the software, in which the circular apertures were shown for all the objects detected above the threshold level. We classified an object as blended if the IRAC position did not correspond to a single source in the higher resolution optical and/or near-IR images (e.g., objects that are well separated in the optical or near-IR images may be blended in the IRAC images due to their poorer angular resolution). About 25% of the pre-

Fig. 4. Multiwavelength identification of the three outlier objects in the $z_{\text{spec}} - z_{\text{phot}}$ diagram. The cut-outs have a size of $10'' \times 10''$. North is up, East is at left.

Table 2. Sample selection criteria.

Selection criteria ^a	Number of objects ^b
$m_{4.5} \leq 23$	4142
$z_{\text{phot}} \geq 3.5$	140
No hard X-ray emission	125
IRAC/channel 4 peak	53

^a Subsequent steps to select the final sample; ^b number of objects satisfying each criterion.

selected $z_{\text{phot}} \geq 3.5$ objects were flagged as blended and therefore discarded.

By matching our catalogue with the publicly available X-ray catalogue of the Chandra Deep Field-North Survey³, we found 14 objects that were detected in the hard X-ray band, and one that had a reliably measured soft X-ray flux. We excluded these 15 sources from the final sample since X-ray emission is indicative of AGNs in those galaxies, which can alter the observed flux and induce an overestimate in the stellar masses.

Finally, we introduced an additional criterion, based on IRAC colors. We considered genuine $z \geq 3.5$ candidates to be only those sources exhibiting an SED peak in the $8.0 \mu\text{m}$ observed frame ($1-2 \mu\text{m}$ rest-frame for $z \geq 3.5$ galaxies) in the AB scale. The reason is that the near-infrared emission of all but the extremely young galaxies show a distinct peak at $1.6 \mu\text{m}$ in the rest-frame, which is red-shifted into the $8.0 \mu\text{m}$ IRAC band for $z \geq 3.5$ sources. This peak corresponds to the combination of the Planck spectral peak of cool stars and the effects of a minimum in the H^- opacity of the stellar atmospheres. It was used as a photometric-redshift indicator in several previous studies (e.g., Sawicki 2002; Berta et al. 2007). This criterion was applied because at this step we expected that our high- z sample could be contaminated by a substantial fraction of low redshift galaxies, because of the intrinsic uncertainty in the photometric-redshift estimate. The degeneracy between reddening and redshift could introduce lower-redshift contaminants in the final high- z sample (i.e., dust-reddened galaxies at lower redshift can be mistaken for higher- z old stellar populations). From a statistical point of view, we could expect a larger population of lower redshift contaminating sources that entered our $z \geq 3.5$ redshift selection than the number of genuine $z \geq 3.5$ galaxies that were placed at $z < 3.5$ because of the uncertainties in their photometric redshifts. This is true because galaxies at $z \geq 3.5$ represent the minority ($\sim 4\%$) of galaxies selected at $4.5 \mu\text{m}$, while most objects are located at $z < 2$.

As already pointed out, the $8 \mu\text{m}$ data point was excluded from the SED fitting analysis because, for objects at $z < 1.8$, it could sample the PAH emission lines for dusty galaxies, and the stellar population models that we used did not account for dust emission. By excluding the $8 \mu\text{m}$ band from the fitting procedure, we may, however, be unable to distinguish between galaxies at $z \sim 2-3$, where the $1.6 \mu\text{m}$ peak lies in the $5.8 \mu\text{m}$ band, and galaxies at $z \geq 3.5$. By requiring the SED to increase between 5.8 and $8 \mu\text{m}$, we may be able to exclude lower redshift contaminants more reliably.

The parameterization of our adopted color criterion is given by the following:

$$(m_{8.0} \leq m_{5.8} + 0.1) \cap (m_{8.0} \leq m_{4.5} + 0.1) \cap (m_{8.0} \leq m_{3.6} + 0.1) \quad (1)$$

where the 0.1 mag terms were introduced to account for the median uncertainty in the IRAC photometry of ~ 0.1 .

In the two panels of Fig. 5, we show the expected $m_{5.8} - m_{8.0}$ color as a function of redshift for different stel-

Fig. 5. The two panels show the $m_{5.8} - m_{8.0}$ color in function of redshift for the Maraston stellar population models (Maraston 2005a). *Top panel:* Single-burst Simple Stellar Populations (SSPs) with three different ages (0.1, 0.3 and 1 Gyr). *Bottom panel:* 0.1 Gyr stellar populations with constant star formation rate (SFR) and four different reddening values ($E_{B-V} = 0.0, 0.3, 0.5, 0.8$).

lar population models. The top panel shows Simple Stellar Populations (SSPs) with three different ages from 0.1 to 1 Gyr, and the bottom panel shows a constant star formation rate model with a fixed age of 0.1 Gyr and reddening between $E_{B-V} = 0$ and 0.8. The templates shown are from the Maraston stellar population models (Maraston 2005a). These models, however, are not supposed to be used for $\lambda > 2.5 \mu\text{m}$, where they exhibit a significant and unphysical. This discontinuity affects the predicted IRAC colors for $0 < z < 2.7$. Hence, we also tested the reliability of this color criterion with BC03 templates, which are not affected by this problem. We verified that the expected IR colors for BC03 stellar population models at $z < 2.7$ are equally bluer than the color cuts defined by Eq. (1).

From the upper panel of Fig. 5, it is clear that the $1.6 \mu\text{m}$ peak criterion may exclude some blue and unreddened ($E_{B-V} = 0$) galaxy at $z \geq 3.5$ of age < 0.1 Gyr. This is because galaxies dominated by young stellar populations do not show the $1.6 \mu\text{m}$ peak in their SEDs (see Sawicki 2002). Because of our adoption of this criterion, we excluded from the final sample a spectroscopic confirmed LBG (object #5455, from Spinrad et al. 1998, with $z_{\text{spec}} = 5.34$ and $z_{\text{phot}} = 5.18$), for which the best fit model solution yielded a predicted age of ~ 7 Myr, and mass of $\sim 10^{10} M_{\odot}$. Nonetheless, we decided to apply this criterion as a conservative way obtaining a reliable sample of the most

massive galaxies at $z \geq 3.5$. The objects that we missed in this way are few in number (see next section), and are also the least massive ones. Their inclusion would not strongly affect our estimate of the comoving stellar mass density (for more details see Sects. 5.2, 6.1, and 6.2).

This criterion is efficient in excluding SSP galaxies with redshifts lower than ~ 2.5 . On the other hand, we note that it cannot completely remove any contamination by interlopers at $2.5 < z < 3.5$, since SSP models of ages > 1 Gyr and $3 < z < 3.5$, and dusty star-forming galaxies at $2.5 < z < 3.5$ can still satisfy this criterion (see the bottom panel of the same figure).

Finally, we note that the adopted color criterion would also select the so-called “IRAC power-law sources”. These objects are in general selected by requiring that the SED increases with the increasing wavelength of the four IRAC bands, and they are believed to be $z < 2.5$ AGNs with a hot dust component that swamps the stellar emission (Alonso-Herrero et al. 2006; Donley et al. 2007). Since Donley et al. (2007) selected a sample of (62) IRAC power-law galaxies in the Chandra Deep Field North (CDFN), 27 of which are in the GOODS-N field, we investigated our overlap with their sample. We found 10 sources in our $4.5 \mu\text{m}$ IRAC-selected ($m_{4.5} \leq 23$) sample (of 4142 objects) in common with their “IRAC power-law” sample. All of these sources but one have an infrared part of the SED that peaks at $8.0 \mu\text{m}$, but most of them were excluded from our $z \geq 3.5$ final sample because of their lower photometric redshifts. Only one object of the Donley et al. (2007) sample fulfilled all of our selection criteria. This is object #14793 in our sample (see Table 4 and Fig. 14), corresponding to the object CDFN:[DRP2007]19080 in Donley et al. (2007). This galaxy has no X-ray emission ($< 4.41 \times 10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$) and a rather high $24 \mu\text{m}$ flux ($111 \pm 3.5 \mu\text{Jy}$). The high redshift solution is still consistent with these properties.

We obtained a final sample of 53 galaxies ($\sim 43\%$ of the pre-selected $z \geq 3.5$ sample) that we considered to be robust candidates at $z \geq 3.5$. About 12% of the 125 pre-selected high- z candidates were discarded due to the low signal-to-noise ratio (S/N) of the $8.0 \mu\text{m}$ -band data, but we cannot exclude the possibility that some could be at high redshift. The other 45% probably consist of either younger systems (age < 0.1 Gyr) at the same redshift, or lower redshift $z < 3.5$ contaminants. We show in the following that the former are only a minority (see Sect. 5.2). In Table 6, the multi-band photometry of the final sample is shown.

As a further test of the sample quality we checked how many candidates satisfy the empirical blending criterion tested on extensive simulations by the GOODS team and used by Daddi et al. (2007) and Dickinson et al. (2009, in prep.). This is based on the measurements of the angular separation ($\Delta\theta$) between positions of IRAC sources and their K -band counterparts (we also used optical coordinates for K -undetected sources). According to this criterion, if a galaxy has $\Delta\theta > 0''.5$, it is probably contaminated by blending in the IRAC bands. However, we note that this criterion proved to be successful for objects detected at the $5\text{--}10\sigma$ in both the K and IRAC bands. For sources with fainter optical and near-IR detections, the shift in the centroid position throughout the different bands could be larger due to higher noise. Hence, a larger $\Delta\theta$ could be caused by coordinate fluctuations rather than blending contamination. Because of the extremely red colors of our candidate high- z galaxies, this criterion is not applicable to our full sample, which consists of one half non-detections in both the optical and near-IR bands. Of the sources with at least a reliable optical or near-IR detection ($> 3\text{--}5\sigma$), we verified that the majority (18/25) passed this criterion. On the other hand, we found 7 galaxies (#1098, #2172, #2796, #13857, #15268,

#15541, #15761) with $\Delta\theta > 0''.5$. However, for these objects the lack of visual evidence of blending led us to conclude that the quite large $\Delta\theta$ separation is caused by the S/N . For two of these galaxies, the confirmed spectroscopic redshifts (#13857, #15541, see Sect. 5.4) also prove that the measured IRAC photometry leads to reliable SED fitting results.

5.1. Comparison between the BC03, MA05 models

The SED fitting analysis with BC03 and MA05 templates provided discordant results for both galaxy stellar masses and ages. The major cause of this discrepancy is the different treatments of the Thermally Pulsing Asymptotic Giant Branch (TP-AGB) phase of stellar evolution in the population models, as debated by several authors (e.g., Maraston et al. 2006; Kannappan & Gawiser 2007; Bruzual A 2007).

In the MA05 models, stars in the TP-AGB phase are a dominant source of bolometric and near-infrared light for stellar populations in the age range of 0.2 to 2 Gyr, more dominant than found in the BC03 models, for which the TP-AGB phase was calibrated using local galaxy stellar populations. Another difference is in the treatment of convective overshooting. Maraston et al. (2006) tested both the BC03 and MA05 models on a sample of high- z galaxies with reliable spectroscopic redshifts ($1.4 \leq z \leq 2.5$) in the Hubble Ultra Deep Field (HUDF). They found that the results of MA05 models imply younger ages by factors of up to 6 and lower stellar masses, by 60% on average, with respect to BC03 (and in general to other population synthesis models). The overestimate of galaxy stellar masses by BC03 models, due to the lack of the TP-AGB contribution to the integrated luminosity, become considerable for evolved stellar populations of ages between 0.5 and 2 Gyr.

In Fig. 6, we compare the stellar masses estimated for the $z \geq 3.5$ galaxies with the two different template libraries (BC03 versus MA05). It appears that the BC03 models tend to overestimate the galaxy masses of our sample with respect to the MA05 models, by 0.15 dex on average (see the smaller panel on the top of Fig. 6), in good agreement with previous studies (e.g., Maraston et al. 2006; Wuyts et al. 2007; Cimatti et al. 2008; Werner et al. 2004). In the rest of this paper, we adopt the MA05 stellar templates in computing the galaxy stellar masses and the comoving stellar mass density.

5.2. Comparison with Lyman-break selection techniques

One of the most popular method used to identify high- z galaxies is the “dropout” technique, a color selection based on the redshifted 912 Å Lyman-break caused by neutral hydrogen in the galaxy SEDs. This technique was introduced by Guhathakurta et al. (1990) and Steidel & Hamilton (1992) to select Lyman Break Galaxies (LBG) at $z \approx 3$ (u -band dropout). It has also been largely used to identify galaxies in the redshift range $z \sim 4$ –6, as B -, V - or i -dropouts (Steidel et al. 1999; Dickinson 1998; ?; Giavalisco et al. 2004b). However, this color criterion is strongly biased against selecting red, massive galaxies at high redshift, which are faint or totally undetected in optical and near-IR bands.

To obtain a quantitative estimate of the fraction of red objects missed by combining optical magnitude-limited selection and dropout techniques, we correlated our final sample of IRAC-selected high-redshift galaxies (53 objects) with a sample of B - and V -dropout LBGs in the GOODS-N field from Giavalisco et al. (2004a). We found that only 7 of our massive high- z

Fig. 6. Comparison between estimated masses by means of different template libraries for our final sample of 53 objects: $\log M(\text{BC03})$ versus $\log M(\text{MA05})$. It is visible that BC03 models tend to overestimate galaxy masses with respect to the MA05 ones. The inserted panel shows the distribution of the $\log M(\text{BC03}) - \log M(\text{MA05})$ relation as a function of the galaxy mass.

candidates are included in the B -dropout sample, and none in the V -dropout sample. This means that the remaining IR bright ($m_{4.5} < 23$) $z \geq 3.5$ candidates have been missed by those selection criteria, due to their faint emission at optical and near-IR wavelengths (most of them having $z_{850} > 27$).

While the vast majority of our candidates are missed by the dropout selection technique, a significant population of massive galaxies with $m_{4.5} < 23$ identifiable by the dropout technique may be missed by our selection criteria. We tested whether these galaxies are present using the sample of Giavalisco et al. (2004a).

Giavalisco et al. (2004a) listed 684 LBGs candidates (B -dropout + V -dropout). Only 7 of these are also part of our sample of $z \geq 3.5$ candidates. Only 32 of the Giavalisco et al. (2004a) galaxies have $m_{4.5} < 23$ (15 B -dropout and 17 V -dropout). Of these, 20 were excluded from our sample because of their photometric redshifts $z < 3.5$ (8 B -dropout and 12 V -dropout). One more object (V -dropout) was excluded because of the presence of X-ray emission and 4 objects (V -dropouts) were excluded because of the lack of an IRAC-8.0 μm peak. One of these 4 V -dropout galaxies has a confirmed spectroscopic redshift of $z = 5.2$ (see also Sect. 5). It is reasonable to suppose that all the 4 V -dropouts galaxies that we excluded are genuine $z \geq 3.5$ galaxies with young ages, lacking a distinct peak in the IRAC-8.0 μm band. Even taking them into account would increase our sample of $z \geq 3.5$ galaxy candidates by only 7% and would reduce to 81% the fraction of $z \geq 3.5$ massive galaxies lost by the UV dropout selection techniques.

It is also interesting to compare with the Lyman-break galaxies samples selected by Eyles et al. (2007) and Stark et al. (2007) in the GOODS-South field. The former work reports 6 galaxies at $z \sim 6$ with masses in the range of 1 – $2.4 \times 10^{10} M_{\odot}$ and 4.5 μm magnitude $m_{4.5} > 23$. Since the authors used Salpeter IMF and BC03 models to derive galaxy SEDs, we have to consider that the galaxy masses of these 6 objects are overestimated by a factor of ~ 2.4 ($2.4 = 1.4 \times 1.7$, where 1.4 and 1.7 are the conversion factors from BC03 to MA05 and from Salpeter to Chabrier IMF, respectively; Pozzetti et al. 2007; Cimatti et al. 2008, see also Sects. 5.1 and 6.2). When rescaled, masses span the range 0.4 – $1 \times 10^{10} M_{\odot}$, only one galaxy having $M \sim 10^{10} M_{\odot}$. Lyman Break Galaxies similar to the ones

Table 3. Best fitting parameters for the MIPS-u sample.

ID	z_{phot}	χ^2_v	$P_{\chi^2_v}$	A_V	$z_{\text{inf-90\%}}$	$z_{\text{sup-90\%}}$	$z_{2-\text{phot}}$	$P_{\chi^2_{v,2}}$	$\log M$
#522	3.928	0.410	94.28	0.80	3.060	5.230	10.060	43.52	9.960
#1098	4.369	1.465	14.55	0.20	4.250	4.460	4.922	0.04	10.143
#2574	3.620	0.307	97.98	0.40	2.570	4.110	6.098	0.47	10.581
#3081	3.543	0.466	91.26	0.70	3.340	3.760	0.456	0.00	9.816
#3302	4.292	0.306	98.01	0.30	3.620	4.600	10.060	0.00	10.565
#4008	9.024	0.201	99.63	0.40	2.150	10.060	3.753	99.57	10.069 ^a
#5073	4.481	0.738	68.90	0.20	4.250	4.670	3.417	20.62	11.073
#5460	4.236	0.854	57.65	0.00	3.480	4.390	3.683	50.87	10.402
#6876	3.620	0.683	74.09	0.20	2.220	4.040	3.172	72.67	10.485
#7042	3.774	0.646	77.54	0.50	3.340	3.900	3.529	75.10	9.723
#7286	3.564	5.582	0.00	0.10	2.850	3.620	3.193	0.00	10.610
#7309	3.613	1.267	24.30	0.00	3.130	3.830	3.417	22.90	9.723
#7785	3.970	1.964	3.28	0.30	3.900	4.040	0.540	0.00	10.391
#8403	4.264	2.018	2.76	0.00	4.110	4.390	3.347	0.57	10.710
#9537	4.103	0.150	99.89	0.40	3.200	4.390	6.651	0.00	10.801
#9561	4.026	0.838	59.14	0.00	2.080	4.320	3.396	51.28	10.412
#11682	4.320	0.367	96.11	0.00	4.040	4.530	2.941	64.89	10.186
#12327	4.971	0.202	99.41	0.70	2.920	7.750	3.256	96.17	11.185
#14130	4.432	0.315	97.06	0.10	1.590	5.090	3.067	96.30	10.235
#14260	4.152	1.243	25.71	0.50	4.040	4.250	0.652	0.01	9.935
#15268	4.775	2.534	0.47	0.80	4.390	5.090	7.260	0.07	10.373

Column description: ID: identification number; z_{phot} : photometric redshift best-fit solution; χ^2_v : reduced χ^2 ; $P_{\chi^2_v}$: probability relative to z_{phot} ; A_V : dust extinction parameter; $z_{\text{inf-90\%}}$ and $z_{\text{sup-90\%}}$: upper and lower bound for the 90% confidence interval; $z_{2-\text{phot}}$: secondary solution for photometric redshift; $P_{\chi^2_{v,2}}$: probability relative to $z_{2-\text{phot}}$; $\log M$: stellar mass in logarithmic scale, computed by using MA05 stellar population models.

We note that for galaxies with $z_{\text{phot}} \sim 8-10$, the best fit redshifts are uncertain, and the range of equally probable solutions ($z_{\text{inf}} - z_{\text{sup}}$) within the 90% of confidence is large. Also for some of these objects – marked by the footnote^a – the *Hypermass* procedure did not find any physical solution for the best-fit SEDs by fixing the redshifts to these extremely high values.

^a This value for the galaxy mass was found for the secondary solution of photometric redshift, $z = 3.753$, because the *Hypermass* software did not find a plausible solution at $z = 9.024$.

found by Eyles et al. (2007) would not be included in our sample, since these are fainter than our magnitude cut of $m_{4.5} = 23$. This agrees with the lower masses found for these LBGs compared to our high- z sample.

The latter work also reports 6 galaxies, spectroscopically confirmed at $z \sim 4.4-5.6$. We used the same conversion factor to rescale galaxy masses from BC03 and Salpeter IMF and we found that only three galaxies of the Stark et al. (2007) sample have masses $> 10^{10} M_{\odot}$. Only two of these three sources have $m_{4.5} \leq 23$.

Hence, only 4 galaxies of the 12 LBGs at $4.4 \leq z \leq 6$ from both works, are within our mass range ($M \sim 10^{10}-10^{11} M_{\odot}$), and only two would pass our magnitude selection ($m_{4.5} \leq 23$). The two IRAC brightest and most massive sources from Stark et al. (2007) have ages of ~ 150 Myr, if rescaled by a factor of 6 due to the overestimation by BC03 models (Maraston et al. 2006), and $E_{B-V} \sim 0.0$. We can suppose that they are similar to the youngest, unreddened population excluded from our final sample by the $8 \mu\text{m}$ peak criterion (see Sect. 5 and Fig. 5, top panel), such as the 4 V-band dropout galaxies of the sample of Giavalisco et al. (2004b). The other two galaxies with masses $\sim 10^{10} M_{\odot}$ and fainter than our IRAC magnitude selection would be even younger (ages < 100 Myr).

In conclusion, it seems that we do not miss a significant population of sources with our selection criteria. A few blue star-forming galaxies with young ages are indeed excluded but these objects have generally low stellar masses that would not meet our stellar-mass completeness-limits ($M_{\star} \geq 5 \times 10^{10} M_{\odot}$; see Sects. 6.1 and 6.2).

5.3. $24 \mu\text{m}$ emission

We used observations at $24 \mu\text{m}$ with *Spitzer*+MIPS to provide some constraints on the “activity” of our high-redshift galaxy candidates. Detection of $24 \mu\text{m}$ flux for $z \geq 3.5$ galaxy candidates could be explained in terms of radiation from star-formation that has been reprocessed and re-emitted by the dust (as emission lines from PAH molecules), or the presence of highly obscured AGN in relatively quiescent galaxies. We divided our sample of high redshift candidates into two subsamples: MIPS-detected (MIPS-d) and MIPS-undetected (MIPS-u) samples and discuss them separately in the following.

5.3.1. The MIPS-u sample

Twenty-one objects in our sample are undetected in MIPS to the detection limit of $20 \mu\text{Jy}$. The best-fit model spectral energy distributions of the MIPS-u subsample are presented in the left panels of Fig. 13. The right panels of the same figure show the χ^2_v distribution as a function of redshift for each of the sources. We also report the most probable photometric redshift solution (z_{phot}). The best-fit model parameters and the estimated galaxy stellar masses are presented in Table 3. Most of the MIPS-u redshifts are dispersed about a median value of $z \sim 4$, in the redshift range $z = 3.5-5$. We found that the best-fit model SEDs corresponded to $\chi^2_v < 2$ for 20 of 21 objects and there was no degeneracy in the redshift solution for the majority of the sources. In contrast, we also identified possible lower redshift solutions within the 90% confidence intervals of the χ^2_v distribution for 5 objects (#4008, #9561, #12327 and #14130), caused

Fig. 7. Comparison between the observed $z - m_{4.5}$ colors of the MIPS-u galaxies (black open diamonds) and the ones predicted at the same redshift for SSPs models (MA05) with different ages (black curves in the figure, with age of 0.1, 0.3 and 0.5 Gyr, respectively, from the lower part upward). Most MIPS-u objects have $z - m_{4.5}$ colors consistent with SSPs with ages in the range of 0.3–0.5 Gyr. On the contrary, a smaller part of the sub-sample (7 objects) shows bluer colors, more compatible with younger SSPs (0.1–0.3 Gyr), probably exhibiting star formation activity.

mainly by the redshift-reddening degeneracy (see Table 3 and Fig. 13).

To constrain the nature of the stellar populations of MIPS-u galaxies further, we compared their $z - m_{4.5}$ colors with the predictions of the MA05 SSP models for different ages (0.1, 0.3 and 0.5 Gyr). The results are shown in Fig. 7. Although the MIPS-u objects span a wide range of $z - m_{4.5}$ colors, most have colors that are consistent with older starlight (0.3–0.5 Gyr). In contrast, a smaller part of the subsample (7 objects) exhibits bluer colors, which are more consistent with younger SSPs (0.1–0.3 Gyr), implying that ongoing star-formation activity is present.

The results of the SED fitting analysis and the lack of mid-IR $24 \mu\text{m}$ emission could suggest that many of the MIPS-u galaxies might be “quiescent”, non-dusty candidates at $z \geq 3.5$. However, we note that the MIPS observations in the field are of insufficient depth to constrain unequivocally the lack of star formation in these galaxies. The $24 \mu\text{m}$ emission is used generally as a SFR indicator in high-redshift galaxies ($z \sim 2$ to 3; Daddi et al. 2005; Papovich et al. 2007; Yan et al. 2007; Rigby et al. 2008). Following Chary & Elbaz (2001), we can estimate that a source at $z \sim 4$ with a $24 \mu\text{m}$ flux of $20 \mu\text{Jy}$ will have a total IR luminosity of $\sim 1.5 \times 10^{13} L_{\odot}$, corresponding to a SFR of $\sim 1500 M_{\odot} \text{ yr}^{-1}$ for a Chabrier IMF (Kennicutt 1998). Therefore, the $24 \mu\text{m}$ non-detection imposes only a loose upper limit to the SFR. However, the $24 \mu\text{m}$ non-detection is a necessary prerequisite for being a quiescent system (given the MIPS flux limit). In the following analysis, we use the MIPS-u galaxy sample to derive upper limits to both the space density and stellar mass density of “quiescent” candidates at $z \sim 4$ (Sect. 6.3).

5.3.2. The MIPS-d sample

Among the 53 galaxies in our high- z sample, more than half (32) are detected at $24 \mu\text{m}$. The best-fit SED model results for the MIPS-d objects and their relative χ^2_{ν} distributions within the 90% confidence intervals are shown in Fig. 14. In Table 4, we report the best-fit model parameters for the MIPS-d sample. The

quality of the fit for the MIPS-d SEDs is high, with χ^2_{ν} values being below 1.5 for almost all objects (except for #5511, which has $\chi^2_{\nu} = 2.46$). However, most galaxies in the subsample have degenerate redshift solutions, caused by the different models used and the amount of reddening present (#702, #2560, #2796, #3393, #3850, #5157, #5367, #6099, #9301, #9306, #10176, #11304, #13129, #14105).

It is difficult to understand the nature of these objects, given that high and low photometric redshift solutions are often equally probable. In these cases, the $24 \mu\text{m}$ emission could be interpreted in different ways, depending on both the object redshift and the reddening parameter. By assuming that $z \geq 3.5$, and taking account of the lack of hard-X ray emission, the MIPS flux could originate in highly obscured AGNs hosted by relatively quiescent galaxies. However, for this high redshift solution, the $24 \mu\text{m}$ emission could also be due to star formation of extremely high rates. In the next section, we examine this possibility using sub-mm/mm photometry from the SCUBA “supermap” (Borys et al. 2003; Pope et al. 2005), and the MAMBO (Greve et al. 2008) and AzTEC (Perera et al. 2008) imaging surveys of GOODS-N.

On the other hand, if the low redshift solution ($z \sim 2$ –3) is considered, the $24 \mu\text{m}$ emission could be explained in terms of PAH emission from dust heated by star-formation activity, and those objects could be dusty starburst galaxies at $z \sim 2$ –3.

5.4. Sub-mm detections: a large population of massive starburst galaxies at $z \sim 4$?

Sub-mm/mm selected galaxies (hereafter SMGs) are the brightest star-forming galaxies known, being far more luminous than the local ultra-luminous IR galaxies – ULIRGs – with $L_{\text{IR}} \gg 10^{12} L_{\odot}$. These galaxies are massive, young objects being observed during their formation epochs, that have high SFRs (Lilly et al. 1999; Scott et al. 2002), about one order of magnitude higher than that of more typical systems of similar masses (Daddi et al. 2007). They are also fairly rare objects, probably due to the short duration of their bright phase ($\ll 100 \text{ Myr}$; Greve et al. 2005). These objects could represent the common early phase in the formation of massive elliptical galaxies, and hence they might be the crucial link in understanding the massive-galaxy formation process.

The numerous sub-mm data sets available in the GOODS-N field can be used to improve the constraints on the nature of our $z \geq 3.5$ candidates.

We completed direct comparisons with three sub-mm maps of the field:

- the Sub-millimeter Common User Bolometer Array (SCUBA, Holland et al. 1999) “supermap” of 35 SMGs selected to have data of $S/N \geq 4$ at $850 \mu\text{m}$ (Borys et al. 2003; Pope et al. 2005, 2006);
- the Max Planck Millimeter Bolometer Array (MAMBO, IRAM 30-m telescope, Bertoldi et al. 2000) $1200 \mu\text{m}$ map from Greve et al. (2008), composed by 27 sources detected with $3.5 \leq S/N < 4$ and 20 with $S/N \geq 4$;
- the AzTEC (Wilson et al. 2008, 15 m James Clerk Maxwell Telescope – JCMT) $1100 \mu\text{m}$ map of 28 sources detected with $S/N \leq 3.75$ (Perera et al. 2008).

We used a simple approach, looking for overlap between our sample and the positions of the submm/mm selected galaxies in the above works. We have used radial separation limits of $7''$, $6''$, and $9''$, to search for matches with galaxies in the SCUBA, MAMBO and AzTEC, respectively.

Table 4. Best fitting parameters for the MIPS-d sample.

ID	z_{phot}	χ^2_v	$P_{\chi^2_v}$	A_V	$z_{\text{inf-90\%}}$	$z_{\text{sup-90\%}}$	$z_{2-\text{phot}}$	$P_{\chi^2_{v,2}}$	$\log M$
#588	4.285	0.785	64.36	0.60	3.690	4.740	9.388	0.05	9.973
#702	8.814	0.067	99.99	0.40	1.450	10.060	3.340	99.99	10.563
#1720	3.522	0.526	87.33	0.60	2.850	4.250	3.921	84.30	10.399
#2172	3.711	0.536	86.56	0.70	3.550	4.180	4.117	70.42	11.097
#2560	3.830	0.044	100.00	0.40	2.780	8.520	7.274	100.00	10.762
#2796	10.060	0.279	98.59	0.80	3.360	10.060	3.536	95.36	10.870 ^a
#2894	3.893	0.628	79.14	0.70	2.990	4.460	10.060	0.37	10.441
#3393	5.503	0.477	90.59	0.80	3.830	5.860	5.167	84.73	10.647
#3850	4.166	0.063	100.00	0.70	3.410	9.290	8.009	99.90	10.940
#5157	4.740	0.344	96.90	0.00	4.110	7.400	6.504	92.59	11.433
#5367	4.355	0.494	89.51	0.80	3.900	4.880	10.060	0.03	10.806
#5511	4.152	2.521	0.50	0.50	4.040	4.250	1.310	0.00	10.551
#6099	5.251	0.367	96.10	0.10	3.200	8.450	4.999	96.02	11.178
#6245	5.062	0.108	99.98	0.00	4.460	5.370	2.934	81.31	10.662
#6463	9.437	0.193	99.68	0.10	3.690	10.060	4.215	99.64	9.973 ^a
#8520	3.683	0.573	83.71	0.00	3.340	3.830	3.137	44.99	10.730
#9007	3.669	1.117	34.42	0.60	2.780	4.110	2.969	34.41	10.824
#9301	8.226	0.059	100.00	0.20	2.640	10.060	8.037	100.00	11.035
#9306	8.660	0.127	99.95	0.80	2.360	10.060	8.219	99.92	11.011
#9782	3.928	1.472	14.28	0.20	3.480	4.320	8.723	0.02	11.116
#10176	5.818	0.002	100.00	0.20	3.550	9.570	5.517	100.00	11.280
#11304	4.145	0.378	95.67	0.80	3.550	8.100	5.461	89.78	10.759
#12699	3.543	1.760	6.21	0.30	3.200	3.690	3.354	5.62	10.558
#13129	7.960	0.030	100.00	0.20	3.410	10.060	8.296	100.00	11.056
#13857	3.746	0.863	56.74	0.10	3.270	4.040	3.403	43.23	10.707
#14505	3.564	0.014	100.00	0.80	2.920	10.060	7.757	100.00	10.025
#14668	4.600	0.209	99.56	0.10	1.870	5.230	4.789	99.53	10.177
#14722	4.145	1.816	5.23	0.40	4.110	4.180	0.617	0.36	10.180
#14793	4.474	1.225	27.41	0.00	4.040	5.160	9.185	0.01	11.411
#15541	3.949	1.216	27.45	0.40	3.690	4.110	0.680	0.00	11.146
#15761	4.292	0.177	99.78	0.10	3.900	4.530	6.567	0.00	11.126
#15771	3.564	0.360	96.35	0.20	3.060	3.830	6.854	0.00	11.283

See Table 3 for the column description. For the objects with extremely high redshift best fit solutions, the same caveat we gave in Table 3 should be taken into account.

^a These values for the galaxy masses were found for the secondary solutions of photometric redshift, because the *hyperzmass* software did not find a acceptable solution at $z \sim 9-10$.

We found eight of our high- z candidates in the error box of the SCUBA “supermap” sources (#5157, #6099, #6463, #8520, #10176, #11682, #13857, #15541), but only seven (except #11682) correspond to the IRAC counterparts identified by Pope et al. (2006). For two of these galaxies, GN20 and GN20.2, (objects #15541, #13857 in our sample) spectroscopic redshifts around $z \sim 4$ were determined by Daddi et al. (2009, hereafter D09). Those objects are among the most distant spectroscopically-SMGs known at the present. Our photometric redshifts for these galaxies are consistent within the margins of error (~ 0.2) with the spectroscopic redshifts, which are 3.75 and 3.95 for GN20 and GN20.2 respectively. It is interesting to remark that D09 discuss the evidence of a galaxy proto-cluster at $z \sim 4$ in the GN20 and GN20.2 area, due to the finding of an over-density of B -dropout galaxies. Four objects in total from our $z \geq 3.5$ sample (#13857, #14722, #15268, #15541) are also within a $25''$ radius centered on the coordinates of GN20. This represents an overdensity of a factor of 18 with respect to the expected number density of IRAC-selected galaxies in the field. Massive galaxies seem to be tracing this proto-cluster structure at $z = 4.05$.

We found that five sources in our high- z sample match the MAMBO galaxies within the correlation distance. Of them only #15541 (GN20) is also SCUBA-detected. In addition, 8 objects from the AzTEC survey correspond to galaxies in our

$z \geq 3.5$ sample, within the adopted search radius. Three of them are also SCUBA-detected, and three are found to be in common with the MAMBO survey from G08.

Candidate $z \geq 3.5$ galaxies in our sample that are consistent with being counterparts of SCUBA, MAMBO and/or AzTEC sources are listed in Table 5.

It should be noted that the bulk of the sub-mm counterparts, except the objects #522 (GN1200.5/AzGN06) and #7785 (GN1200.46), are $24 \mu\text{m}$ -detected (MIPS-d sample). The latter is also a less secure MAMBO detection, being among the sources with a S/N below the 4.5σ limit, and is not part of the “robust deboosted catalogue”.

We can check the reliability of the high- z solution for the sub-mm detections by comparing their S_{850}/S_{24} and $S_{24}/S_{8.0}$ colors with the values measured by D09 for the two confirmed SMGs galaxies at redshift ~ 4 (GN20 and GN20.2). We show this comparison in Fig. 8, where GN20 and GN20.2a are represented as black, bold diamonds. As suggested by D09, starburst galaxies at $z \sim 4$ would have relatively blue Spitzer MIPS-IRAC colors, in a similar way to GN20 and GN20.2a ($\log(S_{24}/S_{8.0}) < 0.7$, $\log(S_{24}/S_{4.5}) < 1.11$). Because of the negative sub-mm k -correction, the S_{850}/S_{24} ratio should also increase with redshift. Hence, we expect that starbursts at $z \sim 4$ should occupy the upper-left part of the diagram in Fig. 8 with $\log(S_{850}/S_{24}) > 2$ (such as GN20 and GN20.2a).

Table 5. SCUBA, MAMBO and AzTEC counterparts found in our $z \geq 3.5$ sample.

ID (this paper)	SCUBA name	MAMBO/AzTEC name	z_{phot} (this paper)	$z_{\text{min}90\%}$	$z_{\text{max}90\%}$	z_{phot} (P06)	$z_{\text{radio-IR}}$ (D09)	S_{850} [mJy]	$S_{1200/1100}$ [mJy]	d [arcsec]
#588	...	.../AzGN12	4.28	3.69	4.74	...	...	...	3.07	5.2
#3850	...	.../AzGN20N	4.17	3.41	9.29	...	...	...	2.79	2.3
#5367	...	.../AzGN27N	4.35	3.90	4.88	...	...	...	2.31	6.5
#14505	...	GN1200.9/AzGN28	3.56	2.92	10.06	...	...	...	2.7/2.31	4.1/4.2
#522	...	GN1200.5/AzGN06	3.93	3.06	5.23	...	...	...	3.8/4.13	2.9/1.1
#7785	...	GN1200.46/...	3.97	3.90	4.04	...	...	...	2.1 ^a	4.5
#11304	...	GN1200.13/...	4.14	3.55	8.10	...	...	...	2.2	0.8
#6099	GN03	...	5.25	2.95	10.0	2.0	1.9	1.8	...	2.5
#10176	GN09	.../AzGN31	5.818	3.55	10.0	2.4	13.42	8.9	2.13	3.9/5.3
#5157	GN11	...	4.74	4.11	7.40	2.3	2.45	7.0	...	2.4
#8520	GN12	.../AzGN08	3.68	3.34	3.83	3.1	3.20	8.6	3.83	5.16/8.7
#6463	GN18	...	9.437/4.215	3.7	10.0	2.2	2.50	3.2	...	3.2
#15541	GN20	GN1200.1/AzGN01	3.949	3.69	4.11	2.95	4.055	20.3	9.3/10.69	1.3/1.9/1.4
#13857	GN20.2a	...	3.746	3.27	4.04	2.83	4.051	9.9	...	6.2

In the columns we report: the ID number of the objects (this paper); the names from the sub-mm surveys; photometric redshift estimated in this paper and their 90% confidence intervals; photometric redshifts by [Pope et al. \(2006, P06\)](#); radio-IR redshifts by [Daddi et al. \(2009, D09\)](#) – with the two spectroscopic values in boldface –; the SCUBA, MAMBO and AzTEC fluxes. The last column is the distances between the proposed IRAC-4.5 μm counterparts and the sub-mm positions.

^a Not deboosted flux.

Fig. 8. S_{850}/S_{24} versus $S_{24}/S_{8.0}$ colors for the SCUBA (black diamonds), MAMBO (red triangles) and AzTEC (green crosses) detections of our high- z sample. The comparison with the GN20 and GN20.2 ([Daddi et al. 2009](#)) seems to support the high- z hypothesis for most of these sources (left part of the diagram; see Sect. 5.4). The solid black lines represent the color limits suggested by ([Daddi et al. 2009](#)) for sub-mm galaxies at $z \geq 4$ ($\log(S_{24}/S_{8.0}) < 0.7$). The small violet squares are the SCUBA “supermap” sources that do not overlap with our high- z sample, most of them spectroscopically confirmed at $z \sim 1$ –2.5.

Of the SCUBA sources (black diamonds), two (#10176/GN09 and #8520/GN12) have mid-IR colors comparable with those of GN20 and GN20.2a, and $\log(S_{850}/S_{24}) > 2$. This finding can be seen as further confirmation of the $z \geq 3.5$ redshifts estimated for these objects. It also agrees, within the 99% confidence interval, with the radio-IR redshifts estimated by D09 (see their Table 3). In the same figure, the “SCUBA-blank” MAMBO (red triangles) and AzTEC (green crosses) detections are also shown. To represent them in this diagram, we scaled the S_{1200} and S_{1100} fluxes to those of S_{850} , by multiplying them by a factor of 2, i.e., the typical S_{850}/S_{1200} ratio found for GN20 and GN20.2 (D09). This is also consistent with the S_{850}/S_{1100} flux ratio derived by [Perera et al. \(2008, \$2.08 \pm 0.18\$ \)](#). For the MIPS-u #522 and #7785 sources,

we show the limits in the two-color diagram, computed with $S_{24} = 20 \mu\text{Jy}$ (the 5σ 24 μm -MIPS flux lower limit). The sub-mm-IR colors of these sources are also consistent, within the error bars, with the expected colors for high- z SMGs. For comparison, we also show the positions occupied in Fig. 8 by the other sub-mm galaxies from the [Pope et al. \(2006\)](#) sample of secure IRAC counterparts that do not overlap with our sample. Almost all these sources (violet small squares) are spectroscopically confirmed at $z \sim 1$ –2.5 (see [Pope et al. 2006](#)). The evidence that most occupy the lower-right part of Fig. 8 strongly support the above statements. The other two galaxies from the SCUBA “supermap” have colors similar to GN20 and GN20.2, i.e., GN10, and GN22. The first one was claimed by several authors (D09, [Dannerbauer et al. 2008](#); [Wang et al. 2007, 2009](#)) to be a $z > 3.7$ starburst, although [Pope et al. \(2006\)](#) report a $z = 2.2$ photometric redshift. The second one is a spectroscopically confirmed galaxy at $z = 2.509$ ([Chapman et al. 2005](#); [Pope et al. 2006](#)). It represents the only case in the SCUBA sample of a lower-redshift sub-mm/mm emitter with both S_{850}/S_{24} ratio and IRAC colors similar to spectroscopically confirmed $z \geq 3.5$ SMGs. Nevertheless, from a statistical point of view, the diagram shown in Fig. 8 seems to be a valid diagnostic for identifying high- z starbursts in sub-mm/mm sources.

Hence, we can conclude that, for $\sim 78\%$ of the galaxies in our sample that are likely sub-mm/mm detected, these results support the hypothesis of extreme SFR activity at high- z . Instead the sub-mm (SCUBA) sources lying in the right part of Fig. 8 (objects #5157, #6099 and #6463, i.e., $\sim 22\%$ of the sub-mm detections) are probably at lower redshifts, as proposed by [Pope et al. \(2006\)](#).

5.5. Mid-IR versus IR colors

The properties and overall spectral energy distributions of the MIPS-d galaxies are comparable to that of the most debated object in the literature, HUDF-JD2. This object, identified as a $z \sim 3.4$ galaxy by [Yan et al. \(2004\)](#) and [Chen & Marzke \(2004\)](#), was suggested by [Mobasher et al. \(2005\)](#) to be a massive post-starburst galaxy at $z \approx 6.5$ hosting an obscured AGN. However,

Table 6. Multi-band photometry for the final sample of 44 candidates at $z \geq 3.5$. The coordinates reported here are from the *Spitzer* data-set.

ID	U	B	V	I	z	J	H	Ks	m _{3.6}	m _{4.5}	m _{5.8}	m _{8.0}	m ₂₄	RA	Dec
# 522	>26.7	>26.3	26.485	26.980	>25.5	>23.3	>23.0	>23.1	23.288	22.895	22.517	21.583	>20.6	12:36:27.2	62:06:06.1
# 588	>26.7	27.774	27.187	26.210	>25.5	>23.3	>23.0	>23.1	23.379	22.960	22.274	21.836	20.218	12:36:32.7	62:06:21.8
# 702	>26.7	--	>26.9	>26.3	>25.5	>23.3	>23.0	>23.1	--	22.964	22.857	22.254	20.081	12:36:15.7	62:06:43.4
# 1098	>26.7	26.968	25.998	24.907	24.317	>23.3	23.242	23.127	22.422	22.465	22.065	21.807	>20.6	12:36:36.8	62:07:14.0
# 1720	>26.7	26.397	25.778	26.033	26.033	>23.3	>23.0	23.021	22.954	22.611	22.354	22.218	20.161	12:36:51.0	62:08:29.8
# 2172	>26.7	27.262	25.325	24.536	24.243	>23.3	23.187	22.482	21.360	21.118	20.898	20.928	19.156	12:36:13.3	62:09:02.0
# 2560	>26.7	--	>26.9	>26.3	>25.5	>23.3	>23.0	23.222	21.999	21.587	21.227	21.137	18.527	12:35:53.0	62:09:30.2
# 2574	>26.7	>26.3	26.722	26.810	26.070	>23.3	>23.0	23.778	22.958	22.559	22.510	22.183	>20.6	12:36:04.7	62:09:25.5
# 2796	>26.7	>26.3	>26.9	>26.3	>25.5	>23.3	>23.0	22.404	21.869	21.528	21.167	21.227	18.029	12:36:03.1	62:09:47.5
# 2894	>26.7	>26.3	26.926	26.256	27.044	>23.3	>23.0	>23.1	23.399	22.997	22.608	22.521	20.496	12:37:10.0	62:09:56.4
# 3081	>26.7	27.283	25.263	24.600	24.272	24.240	>23.0	>23.1	22.859	22.638	--	--	>20.6	12:36:44.4	62:10:03.9
# 3302	>26.7	>26.3	27.155	26.106	25.940	>23.3	>23.0	>23.1	22.834	22.638	22.146	22.169	>20.6	12:35:59.9	62:10:08.4
# 3393	>26.7	>26.3	>26.9	26.751	>25.5	>23.3	23.329	>23.1	22.423	22.064	21.615	21.376	19.581	12:37:19.5	62:10:21.4
# 3850	>26.7	>26.3	>26.9	>26.3	>25.5	>23.3	>23.0	23.202	22.073	21.579	21.125	21.066	20.037	12:37:12.5	62:10:36.1
# 4008	>26.7	>26.3	>26.9	>26.3	>25.5	>23.3	>23.0	23.330	--	22.069	21.817	21.405	>20.6	12:36:53.7	62:11:12.9
# 5073	>26.7	>26.3	26.382	25.038	25.135	>23.3	23.058	23.489	22.107	21.831	21.602	21.165	>20.6	12:36:36.1	62:11:54.7
# 5157	>26.7	27.265	>26.9	27.550	>25.5	>23.3	>23.0	23.846	21.479	21.102	20.764	20.856	18.519	12:36:37.5	62:11:57.2
# 5367	>26.7	27.241	27.154	26.063	25.954	>23.3	>23.0	>23.1	22.729	22.373	22.781	22.567	>20.6	12:36:31.3	62:12:18.3
# 5460	>26.7	>26.3	25.754	24.742	24.932	>23.3	24.431	24.036	23.019	22.691	21.219	20.221	18.084	12:36:37.1	62:12:31.1
# 5511	>26.7	27.209	25.492	24.543	23.806	24.163	22.865	22.320	21.444	21.678	21.269	21.209	18.305	12:36:08.6	62:12:51.2
# 6099	>26.7	27.438	>26.9	>26.3	>25.5	>23.3	>23.0	>23.1	22.129	21.678	22.433	22.222	19.809	12:37:20.6	62:13:00.2
# 6245	>26.7	>26.3	>26.9	25.766	25.270	>23.3	24.283	24.065	22.791	22.688	22.433	22.222	19.177	12:37:02.5	62:13:02.6
# 6463	>26.7	>26.3	>26.9	>26.3	>25.5	>23.3	>23.0	23.287	22.875	22.426	21.948	21.903	>20.6	12:36:30.4	62:13:33.0
# 6876	>26.7	>26.3	26.322	26.938	>25.5	>23.3	>23.0	>23.1	23.042	22.801	22.486	22.386	>20.6	12:37:25.5	62:13:39.8
# 7042	>26.7	>26.3	25.418	24.665	24.685	>23.3	>23.0	23.463	23.224	22.992	23.337	23.059	>20.6	12:37:09.5	62:13:50.3
# 7286	>26.7	27.443	25.467	26.915	24.953	24.288	24.464	22.810	22.519	22.265	21.941	22.033	>20.6	12:36:41.6	62:14:12.5
# 7309	>26.7	27.718	25.498	24.823	24.915	>23.3	24.377	23.068	23.031	22.875	23.306	22.830	>20.6	12:36:55.9	62:14:51.6
# 7785	>26.7	27.099	24.698	23.850	23.618	25.240	23.265	23.026	22.584	22.616	22.367	22.473	>20.6	12:37:03.8	62:14:48.8
# 8403	>26.7	>26.3	25.844	24.729	24.927	23.792	22.877	22.841	22.198	21.971	22.106	21.684	>20.6	12:37:34.7	62:15:15.8
# 8520	>26.7	26.790	25.090	24.743	24.539	24.971	23.162	22.617	21.867	21.739	21.504	21.488	19.433	12:38:03.6	62:15:30.5
# 9007	>26.7	>26.3	25.750	25.303	>25.5	>23.3	>23.0	23.721	22.399	22.033	21.749	21.528	20.255	12:37:33.8	62:15:32.1
# 9301	>26.7	>26.3	>26.9	>26.3	>25.5	>23.3	23.802	>23.1	22.924	22.484	22.332	22.270	19.661	12:36:39.4	62:15:42.8
# 9306	>26.7	>26.3	>26.9	>26.3	>25.5	>23.3	>23.0	22.973	22.360	21.967	21.735	21.555	>20.6	12:37:52.7	62:15:49.3
# 9537	>26.7	>26.3	26.557	25.606	25.600	>23.3	>23.0	22.824	22.133	21.873	21.650	21.485	>20.6	12:37:39.5	62:15:59.0
# 9561	>26.7	>26.3	26.310	25.492	26.253	>23.3	>23.0	23.056	--	22.686	22.597	22.463	>20.6	12:36:22.0	62:16:16.2
# 9782	>26.7	26.988	26.705	>26.3	26.493	>23.3	>23.0	23.298	22.203	21.487	21.199	21.046	18.889	12:37:39.5	62:17:14.8
# 10176	>26.7	>26.3	>26.9	>26.3	>25.5	>23.3	>23.0	>23.1	22.428	21.882	21.439	21.237	20.034	12:36:39.2	62:17:36.8
# 11304	>26.7	27.318	27.906	>26.3	>25.5	>23.3	>23.0	>23.1	21.918	21.373	20.939	20.617	>20.6	12:38:05.4	62:18:16.4
# 11682	>26.7	>26.3	26.464	25.416	25.536	>23.3	>23.0	24.296	22.702	22.189	22.091	21.367	>20.6	12:37:24.0	62:18:33.7
# 12327	>26.7	--	>26.9	26.574	>25.5	>23.3	>23.0	24.705	22.645	22.682	21.996	21.895	19.942	12:37:28.1	62:19:20.5
# 12699	>26.7	26.925	25.271	24.616	24.460	23.892	24.640	22.082	22.107	21.959	22.029	21.515	20.186	12:37:08.8	62:22:02.2
# 13129	>26.7	>26.3	>26.9	>26.3	>25.5	>23.3	24.368	>23.1	22.963	22.364	22.029	21.515	20.637	12:37:36.3	62:21:25.3
# 13857	>26.7	>26.3	25.150	24.638	24.404	>23.3	24.107	>23.1	22.605	22.465	21.913	21.578	>20.6	12:37:13.0	62:21:11.4
# 14130	>26.7	--	>26.9	25.899	26.511	>23.3	>23.0	23.387	22.048	22.900	22.846	22.596	>20.6	12:36:44.0	62:19:38.8
# 14260	>26.7	>26.3	25.352	24.183	23.941	23.682	24.668	22.678	22.552	22.680	22.814	22.647	20.211	12:37:12.3	62:23:03.4
# 14505	>26.7	>26.3	>26.9	>26.3	>25.5	>23.3	>23.0	>23.1	22.924	22.388	22.140	21.573	20.610	12:37:11.5	62:21:55.8
# 14668	>26.7	>26.3	>26.9	25.723	25.684	>23.3	>23.0	23.497	22.936	22.749	22.956	22.385	19.934	12:36:58.9	62:22:15.3
# 14722	>26.7	>26.3	25.010	23.929	23.658	23.133	23.167	23.125	22.352	22.714	--	--	20.6	12:37:09.6	62:22:02.5
# 14793	>26.7	--	26.436	25.824	25.501	>23.3	>23.0	23.757	21.952	21.566	20.716	20.492	>20.6	12:37:11.9	62:22:12.4
# 15268	>26.7	>26.3	>26.9	25.710	>25.5	>23.3	>23.0	>23.1	22.790	21.887	21.316	21.071	19.359	12:36:55.3	62:21:07.9
# 15541	>26.7	>26.3	25.185	24.487	24.155	>23.3	>23.0	22.753	21.874	21.573	21.010	20.595	19.140	12:37:01.5	62:20:25.3
# 15761	>26.7	>26.3	26.783	25.670	25.788	24.071	>23.0	22.707	21.688	21.299	20.985	20.931	19.659	12:37:01.5	62:20:25.3
# 15771	>26.7	>26.3	26.170	26.087	25.386	>23.3	23.687	22.170	21.418	21.117	20.795	20.743	19.659	12:37:01.5	62:20:25.3

Fig. 9. Infrared colors $m_{8,0} - m_{24}$ versus $m_{3,6} - m_{8,0}$ of our high- z sample of galaxies (black diamonds) compared with the objects found by Rodighiero et al. (2007) (red open triangles). The bold red triangle represents HUDF-JD2 (Mobasher et al. 2005). Bold, black, open diamonds are the SCUBA detected galaxies of our sample. The two objects with confirmed $z \sim 4$, GN20 and GN20.2, are highlighted (filled black diamonds). The MIPS-undetected objects are shown as upper limits in the $m_{8,0} - m_{24}$ color: black and red arrows for our sample and Rodighiero et al. (2007) sample, respectively. Two evolutionary color tracks of an ULIRG (violet dot-dashed curve; Chary & Elbaz 2001), and of a galaxy hosting an obscured AGN (dashed green curve; Seyfert 2 galaxy NGC 1068), with increasing redshift values, are also shown. The black horizontal line represents the upper limit suggested by D09 and also shown in Fig. 8 for colors of starburst galaxies at $z \geq 3.5$.

subsequent studies (Dunlop et al. 2007; Rodighiero et al. 2007; Chary et al. 2007) claimed that HUDF-JD2 is more likely to be a lower redshift ($z \sim 1.5$ – 2.2) star-forming galaxy with significant dust reddening. Our galaxies are also similar to candidate $z > 4$ galaxies found by Rodighiero et al. (2007) and Wiklind et al. (2008). In Fig. 9, the infrared colors ($m_{8,0} - m_{24}$ versus $m_{3,6} - m_{8,0}$) of our high- z sample (open black diamonds) are compared with the objects found by Rodighiero et al. (2007, open red triangles). The bold red triangle represents HUDF-JD2 (Mobasher et al. 2005). Bold, black, open diamonds are the sub-mm detected galaxies of our high- z sample. The two objects confirmed to have redshift of $z \sim 4$, GN20 and GN20.2, are represented by filled black diamonds. The MIPS-undetected objects are shown as upper limits to the $m_{8,0} - m_{24}$ color (black and red arrows for our sample and the sample of Rodighiero et al. 2007, respectively). The evolutionary color tracks of an ULIRG (Chary & Elbaz 2001) and of a galaxy hosting an obscured AGN (the Seyfert 2 galaxy NGC 1068) with increasing redshift values are also shown. It is clear, however, that such locally calibrated templates are not useful in interpreting the colors of distant galaxies. This can be argued by considering that the observed $m_{8,0} - m_{24}$ colors of GN20 and GN20.2 are ~ 1 mag redder with respect to the ULIRG track at $z \sim 4$. As discussed in D09, this is probably due to a lower stellar mass-to-light ratio in the distant galaxies.

The black horizontal line represents the upper limit suggested by D09 and also shown in Fig. 8 for colors of starburst galaxies at $z \geq 3.5$ ($\log(S_{24}/S_{8,0}) < 0.7$, in AB magnitude: $m_{8,0} - m_{24} < 1.7$; see also Sect. 5.4). If the high- z solution is correct, objects bluer than this limit should be strongly dominated by star formation, such as GN20 and GN20.2. Hence, the sub-mm detected galaxies in our sample that are located down to the black line in the figure should have properties similar to the

two starbursts spectroscopically confirmed at $z \sim 4$. The MIPS-d sources with similar colors but lacking a sub-mm counterpart could also be high- z starbursts. Alternatively, the lack of sub-mm detection should be justified either by the shallow sensitivity of the currently available sub-mm instruments (SCUBA, MAMBO, AzTEC), or by the particularly “warm” SEDs of this objects.

On the other hand, the $z \geq 3.5$ candidates with $m_{8,0} - m_{24} > 1.7$ in Fig. 9 could represent a different population. Their redder colors, more similar to the bulk of the Rodighiero et al. (2007) sample, could be due to heavily obscured AGN, as also suggested by Rodighiero et al. (2007) for most of their sources, or they might be $z < 3.5$ contaminants. The position of object HUDF-JD2, which overlaps the solid black line in the middle part of the diagram, does not allow us to place tighter constraints on its nature.

6. The comoving stellar-mass density at $3.5 < z < 5$

Several studies (Drory et al. 2005; Fontana et al. 2006; Stark et al. 2007; Verma et al. 2007; Eyles et al. 2007; Yan et al. 2006) have extended measurements of the total stellar mass density to $z \sim 4$ – 5 , using deep multicolor data from optical and near-IR selected samples. One of the goals of this paper was to determine how our IRAC-4.5 selected sample contributes to the comoving stellar mass density at $z \geq 3.5$. As already discussed, observations of the high-redshift Universe are generally biased against red galaxies with large mass-to-light ratios, potentially missing galaxies that could represent a significant part of the total mass density (see Fontana et al. 2006). Here, we emphasize that we can only determine a lower limit to the comoving stellar mass density in the redshift interval $3.5 \leq z \leq 5$, for several reasons. First, the combination of our IRAC 4.5 μm magnitude limit and the 1.6 μm peak SED criterion leads to miss objects with lower masses at these redshifts, such as the fainter, bluer Lyman-break galaxies, that we discussed in Sect. 5. More generally, a magnitude-limited sample (selected at any wavelength) is not the same as a mass-limited sample, as we discuss further in the next section.

6.1. Incompleteness effects and mass selection criteria

In estimating the comoving stellar-mass density, we consider that our sample could be affected by incompleteness. As several other authors (Fontana et al. 2004; Drory et al. 2004; Labbé et al. 2005) have noted, IR-selected samples are not equivalent to mass-selected samples. At any redshift, galaxies detected above the sample magnitude limit can have a fairly wide range of M_*/L ratios, depending on their spectral properties, ages, dust extinction, and metallicities. At higher redshifts, magnitude-limited samples therefore become biased against objects with lower masses and high M_*/L ratios, such as galaxies that are not currently forming stars, or which are highly extinguished.

We attempted to evaluate the consequences and minimize the effects of incompleteness by determining the threshold mass limit as a function of redshift for a variety of galaxy models spanning a range of ages and degrees of dust extinction. We used the MA05 template library with the Chabrier IMF, as used in preceding sections to derive the galaxy stellar masses. We considered both the simplest case of SSP models with different ages (0.1, 0.5 and 1 Gyr), and a set of models with a constant star formation rate, an age of 0.1 Gyr, and different amounts of dust extinction ($E_{B-V} = 0.0, 0.3, 0.5$, applied using the Calzetti attenuation law; Calzetti et al. 2000). The observed 4.5 μm flux as a function of redshift was derived from these models, and from

Fig. 10. The redshift dependence of the stellar mass limit for our IRAC magnitude-limited sample ($m_{4.5} \leq 23$), derived from MA05 stellar population synthesis models. The red curves represent SSP models with different ages (0.1, 0.5 and 1 Gyr). The green curves are models with constant SFR, an age of 0.1 Gyr, and dust extinction, with three different values of reddening E_{B-V} (0.0, 0.3 and 0.5). The open diamonds are the objects of our $z \geq 3.5$ sample, with error bars indicating the 68% confidence ranges on their stellar masses.

this we translated our $4.5 \mu\text{m}$ magnitude limit ($m_{4.5} \leq 23$) into a limiting stellar mass at each redshift. The results are shown in Fig. 10. For passively evolving galaxies (i.e., the SSPs) with ages less than 0.5 Gyr, and for star-forming galaxies with $E_{B-V} < 0.3$, our sample should be fairly complete at $M \geq 5 \times 10^{10} M_\odot$ to $z \leq 5$. This is another advantage of our $4.5 \mu\text{m}$ sample selection from the deep GOODS IRAC data, which is complete to significantly lower mass thresholds at these high redshifts than K -selected samples (see Fontana et al. 2006).

Assuming that the redshift estimates of these objects in our high- z sample are correct, we derive their contribution to the comoving number density and the total stellar mass density, both for the total sample (MIPS-u+MIPS-d) and for the MIPS-u sample only, in the redshift interval $3.5 \leq z \leq 5$ as described in the following sections. The comoving volume in this redshift range over the total GOODS-N solid angle ($\sim 165 \text{ arcmin}^2$, coincident with the ACS z -band area) is $V_{\text{com}} \sim 0.725 \times 10^6 \text{ Mpc}^3$.

6.2. Contribution to the stellar mass density from our total sample in the redshift bin $z = 3.5\text{--}5$

Based on the preceding analysis, we assumed a limiting stellar mass $M \sim 5 \times 10^{10} M_\odot$ as the completeness threshold for our magnitude-limited sample. For quiescent galaxies with intermediate ages (~ 0.5 Gyr), we should be complete above this mass threshold to $z \sim 5$. However, for $z > 4$, we may be incomplete close to this mass limit for the oldest possible passive stellar populations, or for heavily dust-obscured galaxies.

Given this potential incompleteness, and the fact that we are undoubtedly missing a large percentage of bluer, lower-mass LBGs in the same redshift range (as discussed in Sect. 5.2), we must consider our estimate for the stellar mass density at $3.5 \leq z \leq 5$ to be a lower limit. Figure 10 shows the derived stellar masses versus redshift for the objects in our IRAC-selected sample. We count objects and sum their stellar masses for galaxies above our mass threshold ($M \sim 5 \times 10^{10} M_\odot$). The resulting comoving number density and the stellar mass density for this sample are $2.6 \times 10^{-5} \text{ Mpc}^{-3}$ and $(2.9 \pm 1.5) \times 10^6 M_\odot \text{ Mpc}^{-3}$, respectively.

Fig. 11. Comparison between our estimate of the lower limit of the stellar mass density (ρ_*) at $3.5 \leq z \leq 5$ (red triangle) and the stellar mass density estimated by Fontana et al. (2006, F06) at $3 \leq z \leq 4$, for galaxies in the mass range $M_* \sim 3 \times 10^{10} - 3 \times 10^{11} M_\odot$. The curve shows the evolution of stellar mass density with redshift from the Millennium Simulation models, computed for galaxies above our threshold limit of $\geq 5 \times 10^{10} M_\odot$.

In Fig. 11, we show our lower limit to the comoving stellar mass density (ρ_*) at $3.5 \leq z \leq 5$, and compare it to another estimate (Fontana et al. 2006, F06) at $3 \leq z \leq 4$ for galaxies in a similar mass range ($M_* = 3 \times 10^{10} - 3 \times 10^{11} M_\odot$). We considered the F06 value from the compilation published by Wilkins et al. (2008), and corrected it for the M/L differences between the Salpeter and Chabrier IMFs ($\log M(\text{Chabrier}) = \log M(\text{Salpeter}) - 0.23$), as well as by a factor of 1.4 to account for the difference between the BC03 models used by F06 and the MA05 models used in our study (Sect. 5.1). We also show a curve derived from the Millennium Simulation models for the redshift evolution of the stellar-mass density, computed for galaxy masses above the threshold limit of $\geq 5 \times 10^{10} M_\odot$. Assuming that the photometric redshifts of all the selected high- z candidates are correct, our results are in good agreement with the Millennium Simulation predictions.

6.3. Stellar-mass density for the MIPS-u candidates in the redshift bin $z = 3.5\text{--}5$

As described above, our sample should be largely complete for “quiescent”, non-star-forming galaxies with ages ≤ 0.5 Gyr and masses $M \geq 5 \times 10^{10} M_\odot$. However, at these high redshifts, the flux limit at $24 \mu\text{m}$, even for deep data from GOODS, corresponds to quite high star-formation rates ($\sim 1000\text{--}1500 M_\odot/\text{yr}$). Therefore, we cannot affirm that the MIPS-undetected (MIPS-u) objects in our sample are truly “quiescent”, and thus we can only derive an upper limit for the number and stellar mass densities of quiescent galaxies. For the MIPS-u subsample with stellar masses $\geq 5 \times 10^{10} M_\odot$ at $z = 3.5\text{--}5$, we find a comoving number density of $0.97 \times 10^{-5} \text{ Mpc}^{-3}$, and a stellar mass density of $(1.15 \pm 0.7) \times 10^6 M_\odot \text{ Mpc}^{-3}$ (MA05 templates and Chabrier IMF).

In Fig. 12, we compare our results for the MIPS-u sample with estimates of the stellar mass density for quiescent galaxies at other redshifts from the literature. To facilitate this comparison, we scaled all values to correspond to the Chabrier IMF and MA05 models by means of the M/L conversion relations given above (see Sect. 6.2). We also computed the comoving stellar

Fig. 12. Comparison between the upper limit to stellar mass density (ρ_*) found in this paper for “quiescent” galaxies in the redshift bin $z = 3.5\text{--}5$ and results from the literature at other redshifts. All points have been scaled to Chabrier IMF and MA05 masses (see Sect. 6.2). Our estimate is represented as a red bold triangle. The horizontal red dashed line represent the stellar mass density we derive by considering all MIPS-u galaxies to be SSPs with age of 0.5 Gyr. The local mass density in passive galaxies is shown by the black asterisk (Baldry et al. 2004). The black open square is for Distant Red Galaxies (DRGs) at $z = 2\text{--}3$ in the Hubble Deep Field South (Labbé et al. 2005). The black filled rhombus is from Daddi et al. (2005) for passive-BzK galaxies in the Hubble Ultra Deep Field (HUDF) at $z = 1.39\text{--}2$, and the open rhombus is for Balmer-break MIPS undetected galaxies at $z = 5\text{--}7$ in GOODS-South from Wiklind et al. (2008).

mass density of the MIPS-u galaxies by considering the unlikely case that they were all truly passive and consistent with being single-burst stellar populations (SSPs) with no dust. The masses were derived using the ratio of the observed IRAC-4.5 μm flux for each galaxy to the expected flux for an MA05 SSP model at the same redshift, assuming an age of ~ 0.5 Gyr and a mass normalized to $1 M_\odot$. In this case, the number density of MIPS-u candidates above the mass threshold would be higher ($1.52 \times 10^{-5} \text{ Mpc}^{-3}$). The corresponding stellar-mass density would be $(1.46 \pm 0.7) \times 10^6 M_\odot \text{ Mpc}^{-3}$. This value is also represented in Fig. 12 by the red dashed line. From this exercise, we concluded that even if all the MIPS-u galaxies were considered as old SSPs, the number density of “quiescent” galaxies at $z = 3.5\text{--}5$ would remain low in comparison with that found at $z \leq 2\text{--}2.5$ in previous studies.

The stellar mass density that we have found for the MIPS-u galaxies in the redshift range $z = 3.5\text{--}5$ is ~ 2 orders of magnitude lower than the local value (black asterisk; Baldry et al. 2004), and more than a factor of 10 lower than the values found at $z \sim 1.7$ by Daddi et al. (2005, black filled rhombus) for the passive-BzK galaxies in the Hubble Ultra Deep Field (HUDF) and at $z \sim 2.5$ by Labbé et al. (2005, black open square) for Distant Red Galaxies (DRGs) in the Hubble Deep Field South. Our result is also in good agreement with the mass density estimated by Wiklind et al. (2008) for $z \sim 5.5$ Balmer-break-selected galaxies without MIPS detections in the GOODS South Field at redshift $z \sim 5.5$ (open rhombus in the figure). Considering that our estimate is an upper limit to the comoving stellar-mass density of massive and quiescent galaxies, we conclude that there is evidence of strong evolution beyond $z \sim 2$ (see also Daddi et al. 2005). The decreasing number density of quiescent objects at high redshifts could mean that most massive

galaxies at early epochs experienced significant star-formation activity.

7. Summary and conclusions

A substantial number of massive galaxies have been found at redshifts as high as $z \sim 5$. Although the range of galaxy stellar populations and observed colors at all redshifts has been heterogeneous, most surveys at $z \geq 3.5$ have selected primarily star-forming blue galaxies, and have been biased against red massive objects that are faint or completely undetected by optical and near-IR surveys.

We have studied galaxies selected at IRAC-4.5 μm in the GOODS-N field, in a complementary way to optical or near-IR selection. This has allowed us to select massive, evolved galaxies at high redshift, for which the optical rest-frame emission (which is most closely related to stellar mass) is sampled by the red-shifted, rest-frame near-infrared portion of the spectrum. On the other hand, this selection may also be sensitive to “active” obscured galaxies at high redshift, for which the IR emission could represent reprocessed radiation from absorbing material surrounding an AGN, or from dust heated by star formation, as seen in dusty starburst galaxies such as local ULIRGs.

Other studies based on near-IR selection have detected red, massive galaxy candidates at $z_{\text{phot}} \geq 4$, but have not obtained complete samples. In most of these studies, the samples have been selected in the K -band, with the possibility of missing K -undetected objects visible in the longer-wavelength IRAC filters (e.g., Mobasher et al. 2005; Wiklind et al. 2008). Rodighiero et al. (2007) used IRAC-3.6 μm selection ($m_{3.6} < 23.26$) to identify a sample of high- z galaxies missed by optical and K -band selection. However, they assumed as additional conditions the non-detection of their candidates in optical bands, and a detection close to the sky threshold limit in the K band ($K > 23$, AB system), which ensured that their sample was a priori incomplete.

Here, we have attempted to select a sample of massive high-redshift red galaxies that is as complete as possible, by means of IRAC selection, to allow us to recover massive objects missed by previous studies. The most important results of this paper are summarized below:

- By selecting at a wavelength of IRAC-4.5 μm ($m_{4.5} \leq 23$), we extracted a sample of 4142 objects in the $\sim 165 \text{ arcmin}^2$ of the GOODS-N field. We performed an SED fitting analysis to estimate photometric redshifts for those objects, and found fifty-three candidates at $z \geq 3.5$, by also requiring that the peak of the f_ν stellar spectral energy distribution (at $1.6 \mu\text{m}$ in the rest-frame) was within the IRAC-8.0 μm band. We excluded unobscured AGN from our final sample, which were identified by their hard and/or soft-X ray emission. Almost 81% of our candidates were completely missed by the B - and V -band Lyman break dropout techniques designed to identify UV-bright, star-forming galaxies at similar redshifts. For each object, we evaluated the reliability and the confidence limits of our photometric redshift estimates with a χ^2 test. Two objects have spectroscopically confirmed redshifts (GN20 and GN20.2; Daddi et al. 2009) in good agreement with our estimates.
- We divided our final sample in two subsamples of (32) MIPS-detected and (21) MIPS-undetected objects (the MIPS-d and MIPS-u samples, respectively).

MIPS-d sample: in the subsample detected at $24 \mu\text{m}$, we found 18 galaxies with unambiguous high photometric-redshift solutions and 14 with degenerate solutions. If the

high-redshift solution is correct, the $24\ \mu\text{m}$ emission could indicate the presence of a heavily obscured AGN (perhaps Compton thick due to the absence of X ray emission), or perhaps strong star-formation activity in a dusty, hyperluminous starburst. We cannot firmly exclude the possibility of these galaxies being dusty starburst galaxies at lower redshifts, for which the $24\ \mu\text{m}$ flux could be explained with PAH emission at $z \sim 2-3$.

MIPS-u sample: twenty-one of our $z \geq 3.5$ candidates are undetected at $24\ \mu\text{m}$. The lack of $24\ \mu\text{m}$ emission is a necessary but insufficient prerequisite for indicating that these are quiescent galaxies. The upper limit to the SFR set by the MIPS detection limit ($f_{24} < 20\ \mu\text{Jy}$) corresponds to a star-formation rate $< 1500\ M_{\odot}\ \text{yr}^{-1}$, assuming bolometric corrections derived from spectral templates for local ultraluminous infrared galaxies. For this reason, we can use this subsample only to provide an upper limit to the total stellar mass of “quiescent” galaxies at $z \geq 3.5$.

- Fourteen galaxies from our $z \geq 3.5$ sample are detected at submillimeter and millimeter wavelengths with SCUBA (Pope et al. 2006), MAMBO (Greve et al. 2008) and AzTEC (Perera et al. 2008). All but two are also detected at $24\ \mu\text{m}$. Two objects, GN20 and GN20.2, have spectroscopically confirmed redshifts at $z \sim 4$ (Daddi et al. (2009)). By comparing their colors with other SMGs galaxies in a S_{850}/S_{24} versus S_{24}/S_8 diagram, we found that $\sim 78\%$ of the sub-mm detected objects among our candidates have properties that are fully consistent with those of starburst galaxies at $z \geq 3.5$. The remaining $\sim 22\%$ are more likely to be at lower redshift.
- We have computed the contribution to the number density and the stellar mass density by both the total (MIPS-d+MIPS-u) sample and the subsample undetected at $24\ \mu\text{m}$ (MIPS-u) for galaxies with stellar masses above a limit of $5 \times 10^{10}\ M_{\odot}$, chosen to minimize incompleteness at $z \leq 5$. For the redshift range $3.5 \leq z \leq 5$, we inferred a comoving number density of $2.6 \times 10^{-5}\ \text{Mpc}^{-3}$ for average stellar masses of $\sim 10^{11}\ M_{\odot}$. The corresponding stellar-mass density is $(2.9 \pm 1.5) \times 10^6\ M_{\odot}\ \text{Mpc}^{-3}$. Our results are in good agreement with previous determinations in the literature, and with the predictions of the Millennium simulations. For the MIPS-u sample, we determined a number density of $0.97 \times 10^{-5}\ \text{Mpc}^{-3}$ and a corresponding stellar-mass density of $(1.15 \pm 0.7) \times 10^6\ M_{\odot}\ \text{Mpc}^{-3}$. We also considered the extreme case that all of the MIPS-undetected galaxies were purely quiescent, with ages ~ 0.5 Gyr. This results in higher upper limits to their number and stellar-mass densities, of $1.52 \times 10^{-5}\ \text{Mpc}^{-3}$ and $(1.46 \pm 0.7) \times 10^6\ M_{\odot}\ \text{Mpc}^{-3}$, respectively. Our MIPS-u sample of quiescent candidates accounts for $\sim 5-6\%$ of the mass density in place at $z \sim 2$ and for only $\sim 1\%$ of local stellar mass density (Baldry et al. 2004). This provides evidence of significant growth in the number of massive galaxies between redshifts $z = 4-5$ and $z = 2$.

Acknowledgements. We thank Kyoung-Soo Lee for her reductions of the KPNO-4 m FLAMINGOS imaging data of the GOODS-North field, and for information about the details of this data set. C.M. and E.D. acknowledge support from the French ANR grant numbers ANR-07-BLAN-0228 and ANR-08-JCJC-0008.

References

Alexander, D. M., Bauer, F. E., Brandt, W. N., et al. 2003, *AJ*, 126, 539
 Alonso-Herrero, A., Pérez-González, P. G., Alexander, D. M., et al. 2006, *ApJ*, 640, 167

- Baldry, I. K., Glazebrook, K., Brinkmann, J., et al. 2004, *ApJ*, 600, 681
 Berta, S., Lonsdale, C. J., Siana, B., et al. 2007, *A&A*, 467, 565
 Bertin, E., & Arnouts, S. 1996, *A&AS*, 117, 393
 Bertoldi, F., Menten, K. M., Kreysa, E., Carilli, C. L., & Owen, F. 2000 [arXiv:astro-ph/0010553]
 Bolzonella, M., Miralles, J.-M., & Pelló, R. 2000, *A&A*, 363, 476
 Borys, C., Chapman, S., Halpern, M., & Scott, D. 2003, *MNRAS*, 344, 385
 Bower, R. G., Benson, A. J., Malbon, R., et al. 2006, *MNRAS*, 370, 645
 Brammer, G. B., & van Dokkum, P. G. 2007, *ApJ*, 654, L107
 Bruzual A., G. 2007 [arXiv:astro-ph/0703052]
 Bruzual, G., & Charlot, S. 2003, *MNRAS*, 344, 1000
 Bunker, A. J., Stanway, E. R., Ellis, R. S., McMahon, R. G., & McCarthy, P. J. 2003, *MNRAS*, 342, L47
 Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, *ApJ*, 533, 682
 Capak, P., Cowie, L. L., Hu, E. M., et al. 2004, *AJ*, 127, 180
 Chabrier, G. 2003, *PASP*, 115, 763
 Chapman, S. C., Blain, A. W., Smail, I., & Ivison, R. J. 2005, *ApJ*, 622, 772
 Chary, R., & Elbaz, D. 2001, *ApJ*, 556, 562
 Chary, R.-R., Teplitz, H. I., Dickinson, M. E., et al. 2007, *ApJ*, 665, 257
 Chen, H.-W., & Marzke, R. O. 2004, *ApJ*, 615, 603
 Cimatti, A., Daddi, E., Renzini, A., et al. 2004, *Nature*, 430, 184
 Cimatti, A., Cassata, P., Pozzetti, L., et al. 2008, *A&A*, 482, 21
 Coleman, G. D., Wu, C.-C., & Weedman, D. W. 1980, *ApJS*, 43, 393
 Cowie, L. L., Barger, A. J., Hu, E. M., Capak, P., & Songaila, A. 2004, *AJ*, 127, 3137
 Daddi, E., Renzini, A., Pirzkal, N., et al. 2005, *ApJ*, 626, 680
 Daddi, E., Dickinson, M., Morrison, G., et al. 2007, *ApJ*, 670, 156
 Daddi, E., Dannerbauer, H., Stern, D., et al. 2009, *ApJ*, 694, 1517
 Dannerbauer, H., Walter, F., & Morrison, G. 2008, *ApJ*, 673, L127
 Dawson, S., Stern, D., Bunker, A. J., Spinrad, H., & Dey, A. 2001, *AJ*, 122, 598
 Dickinson, M. 1998, in *The Hubble Deep Field*, ed. M. Livio, S. M. Fall, & P. Madau, 219
 Dickinson, M., Giavalisco, M., & The Goods Team 2003, in *The Mass of Galaxies at Low and High Redshift*, ed. R. Bender, & A. Renzini, 324
 Dickinson, M., Stern, D., Giavalisco, M., et al. 2004, *ApJ*, 600, L99
 Donley, J. L., Rieke, G. H., Pérez-González, P. G., Rigby, J. R., & Alonso-Herrero, A. 2007, *ApJ*, 660, 167
 Drory, N., Bender, R., Feulner, G., et al. 2004, *ApJ*, 608, 742
 Drory, N., Salvato, M., Gabasch, A., et al. 2005, *ApJ*, 619, L131
 Dunlop, J. S., Cirasuolo, M., & McLure, R. J. 2007, *MNRAS*, 376, 1054
 Eyles, L. P., Bunker, A. J., Stanway, E. R., et al. 2005, *MNRAS*, 364, 443
 Eyles, L. P., Bunker, A. J., Ellis, R. S., et al. 2007, *MNRAS*, 374, 910
 Fazio, G. G., Hora, J. L., Allen, L. E., et al. 2004, *ApJS*, 154, 10
 Fontana, A., Pozzetti, L., Donnarumma, I., et al. 2004, *A&A*, 424, 23
 Fontana, A., Salimbeni, S., Grazian, A., et al. 2006, *A&A*, 459, 745
 Franx, M., Labbé, I., Rudnick, G., et al. 2003, *ApJ*, 587, L79
 Giavalisco, M., Dickinson, M., Ferguson, H. C., et al. 2004a, *ApJ*, 600, L103
 Giavalisco, M., Ferguson, H. C., Koekemoer, A. M., et al. 2004b, *ApJ*, 600, L93
 Greve, T. R., Bertoldi, F., Smail, I., et al. 2005, *MNRAS*, 359, 1165
 Greve, T. R., Pope, A., Scott, D., et al. 2008, *MNRAS*, 389, 1489
 Guhathakurta, P., Tyson, J. A., & Majewski, S. R. 1990, *ApJ*, 357, L9
 Holland, W. S., Robson, E. I., Gear, W. K., et al. 1999, *MNRAS*, 303, 659
 Kannappan, S. J., & Gawiser, E. 2007, *ApJ*, 657, L5
 Kauffmann, G., & Charlot, S. 1998, *MNRAS*, 297, L23
 Kennicutt, Jr., R. C. 1998, *ARA&A*, 36, 189
 Kriek, M., van Dokkum, P. G., Franx, M., et al. 2006, *ApJ*, 649, L71
 Kroupa, P. 2001, *MNRAS*, 322, 231
 Labbé, I., Huang, J., Franx, M., et al. 2005, *ApJ*, 624, L81
 Lejeune, T., Cuisinier, F., & Buser, R. 1997, *A&AS*, 125, 229
 Lilly, S. J., Eales, S. A., Gear, W. K. P., et al. 1999, *ApJ*, 518, 641
 Madau, P. 1995, *ApJ*, 441, 18
 Maraston, C. 2005a, *MNRAS*, 362, 799
 Maraston, C. 2005b, in *Multiwavelength Mapping of Galaxy Formation and Evolution*, ed. A. Renzini, & R. Bender, 290
 Maraston, C., Daddi, E., Renzini, A., et al. 2006, *ApJ*, 652, 85
 McCarthy, P. J., Le Borgne, D., Crampton, D., et al. 2004, *ApJ*, 614, L9
 McLure, R. J., Cirasuolo, M., Dunlop, J. S., et al. 2006, *MNRAS*, 372, 357
 Mobasher, B., Dickinson, M., Ferguson, H. C., et al. 2005, *ApJ*, 635, 832
 Muller, G. P., Reed, R., Armandroff, T., Boroson, T. A., & Jacoby, G. H. 1998, in *Optical Astronomical Instrumentation*, ed. S. D’Odorico, *Proc. SPIE*, 3355, 577
 Night, C., Nagamine, K., Springel, V., & Hernquist, L. 2006, *MNRAS*, 366, 705
 Papovich, C., Rudnick, G., Le Floch, E., et al. 2007, *ApJ*, 668, 45
 Perera, T. A., Chapin, E. L., Austermann, J. E., et al. 2008, *MNRAS*, 391, 1227
 Pérez-González, P. G., Rieke, G. H., Villar, V., et al. 2008, *ApJ*, 675, 234
 Pope, A., Borys, C., Scott, D., et al. 2005, *MNRAS*, 358, 149
 Pope, A., Scott, D., Dickinson, M., et al. 2006, *MNRAS*, 370, 1185
 Pozzetti, L., Bolzonella, M., Lamareille, F., et al. 2007, *A&A*, 474, 443

- Reddy, N. A., Steidel, C. C., Erb, D. K., Shapley, A. E., & Pettini, M. 2006, *ApJ*, 653, 1004
- Rieke, G. H., Young, E. T., Engelbracht, C. W., et al. 2004, *ApJS*, 154, 25
- Rigby, J. R., Marcellac, D., Egami, E., et al. 2008, *ApJ*, 675, 262
- Rodighiero, G., Lari, C., Pozzi, F., et al. 2006, *MNRAS*, 371, 1891
- Rodighiero, G., Cimatti, A., Franceschini, A., et al. 2007, *A&A*, 470, 21
- Saracco, P., Longhetti, M., Severgnini, P., et al. 2005, *MNRAS*, 357, L40
- Sawicki, M. 2002, *AJ*, 124, 3050
- Scott, S. E., Fox, M. J., Dunlop, J. S., et al. 2002, *MNRAS*, 331, 817
- Spinrad, H., Stern, D., Bunker, A., et al. 1998, *AJ*, 116, 2617
- Stanway, E. R., Bunker, A. J., McMahon, R. G., et al. 2004, *ApJ*, 607, 704
- Stark, D. P., Bunker, A. J., Ellis, R. S., Eyles, L. P., & Lacy, M. 2007, *ApJ*, 659, 84
- Steidel, C. C., & Hamilton, D. 1992, *AJ*, 104, 941
- Steidel, C. C., Giavalisco, M., Pettini, M., Dickinson, M., & Adelberger, K. L. 1996, *ApJ*, 462, L17
- Steidel, C. C., Adelberger, K. L., Giavalisco, M., Dickinson, M., & Pettini, M. 1999, *ApJ*, 519, 1
- Steidel, C. C., Adelberger, K. L., Shapley, A. E., et al. 2003, *ApJ*, 592, 728
- Vanzella, E., Cristiani, S., Dickinson, M., et al. 2008, *A&A*, 478, 83
- Verma, A., Lehnert, M. D., Förster Schreiber, N. M., Bremer, M. N., & Douglas, L. 2007, *MNRAS*, 377, 1024
- Wang, W.-H., Cowie, L. L., van Saders, J., Barger, A. J., & Williams, J. P. 2007, *ApJ*, 670, L89
- Wang, W.-H., Barger, A. J., & Cowie, L. L. 2009, *ApJ*, 690, 319
- Werner, M. W., Roellig, T. L., Low, F. J., et al. 2004, *ApJS*, 154, 1
- Wiklind, T., Dickinson, M., Ferguson, H. C., et al. 2008, *ApJ*, 676, 781
- Wilkins, S. M., Trentham, N., & Hopkins, A. M. 2008, *MNRAS*, 385, 687
- Wilson, G. W., Austermann, J. E., Perera, T. A., et al. 2008, *MNRAS*, 386, 807
- Wirth, G. D., Willmer, C. N. A., Amico, P., et al. 2004, *AJ*, 127, 3121
- Wuyts, S., Labbé, I., Franx, M., et al. 2007, *ApJ*, 655, 51
- Yan, H., Dickinson, M., Eisenhardt, P. R. M., et al. 2004, *ApJ*, 616, 63
- Yan, H., Dickinson, M., Giavalisco, M., et al. 2006, *ApJ*, 651, 24
- Yan, H., Dickinson, M., Giavalisco, M., et al. 2007, in *ASP Conf. Ser.*, ed. J. Afonso, H. C. Ferguson, B. Mobasher, & R. Norris, 380, 35

Fig. 13. Spectral Energy Distribution best-fit models for the MIPS-u sample in units of AB magnitude. The red diamonds represent the observed magnitudes from the photometric data. Non-detections are represented by upper limits. See also Table 3.

Fig. 13. continued.

Fig. 13. continued.

Fig. 13. continued.

Fig. 14. Spectral Energy Distributions of MIPS-d galaxies of our final high- z sample (on the left panels), and the correspondent χ^2_{ν} distributions as a function of redshift (right panels) in 90% confidence intervals. For each candidate the most probable redshifts (first and second redshift solutions: z_{phot} and $z_{\text{phot}2}$) and the χ^2_{ν} values with the respective probability are also labeled. See also Table 4.

Fig. 14. continued.

Fig. 14. continued.

Fig. 14. continued.

Fig. 14. continued.

Fig. 14. continued.

Fig. 15. Multiwavelength images of the fifty-three galaxies of the final sample (MIPS-d+MIPS-u). The cut-outs have a size of $10'' \times 10''$. North is up, East is at left.

Fig. 15. continued.

Fig. 15. continued.

Fig. 15. continued.

Fig. 15. continued.

Fig. 15. continued.