The Astrophysical Journal, 786:108 (19pp), 2014 May 10 doi:10.1088/0004-637X/786/2/108

THE MOST LUMINOUS $z \sim 9$–10 GALAXY CANDIDATES YET FOUND: THE LUMINOSITY FUNCTION, COSMIC STAR-FORMATION RATE, AND THE FIRST MASS DENSITY ESTIMATE AT 500 Myr*

P. A. Oesch[1,2,11], R. J. Bouwens[3], G. D. Illingworth[1], I. Labbé[3], R. Smit[3], M. Franx[3], P. G. van Dokkum[2], I. Momcheva[2], M. L. N. Ashby[4], G. G. Fazio[4], J.-S. Huang[4], S. P. Willner[4], V. Gonzalez[5], D. Magee[1], M. Trenti[6], G. B. Brammer[7], R. E. Skelton[8], and L. R. Spitler[9,10]
[1] UCO/Lick Observatory, University of California, Santa Cruz, 1156 High St, Santa Cruz, CA 95064, USA; pascal.oesch@yale.edu
[2] Yale Center for Astronomy and Astrophysics, Yale University, New Haven, CT 06520, USA
[3] Leiden Observatory, Leiden University, NL-2300 RA Leiden, Netherlands
[4] Harvard-Smithsonian Center for Astrophysics, Cambridge, MA, USA
[5] University of California, Riverside, 900 University Ave, Riverside, CA 92507, USA
[6] Institute of Astronomy and Kavli Institute for Cosmology, University of Cambridge, Cambridge CB3 0HA, UK
[7] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[8] South African Astronomical Observatory, P.O. Box 9, Observatory 7935, South Africa
[9] Department of Physics and Astronomy, Faculty of Sciences, Macquarie University, Sydney, NSW 2109, Australia
[10] Australian Astronomical Observatory, P.O. Box 915, North Ryde, NSW 1670, Australia
Received 2013 September 11; accepted 2014 January 30; published 2014 April 24

ABSTRACT

We present the discovery of four surprisingly bright ($H_{160} \sim 26$–27 mag AB) galaxy candidates at $z \sim 9$–10 in the complete *HST* CANDELS WFC3/IR GOODS-N imaging data, doubling the number of $z \sim 10$ galaxy candidates that are known, just ∼500 Myr after the big bang. Two similarly bright sources are also detected in a reanalysis of the GOODS-S data set. Three of the four galaxies in GOODS-N are significantly detected at 4.5σ–6.2σ in the very deep *Spitzer*/IRAC 4.5 μm data, as is one of the GOODS-S candidates. Furthermore, the brightest of our candidates (at $z = 10.2 \pm 0.4$) is robustly detected also at 3.6 μm (6.9σ), revealing a flat UV spectral energy distribution with a slope $\beta = -2.0 \pm 0.2$, consistent with demonstrated trends with luminosity at high redshift. Thorough testing and use of grism data excludes known low-redshift contamination at high significance, including single emission-line sources, but as-yet unknown low redshift sources could provide an alternative solution given the surprising luminosity of these candidates. Finding such bright galaxies at $z \sim 9$–10 suggests that the luminosity function for luminous galaxies might evolve in a complex way at $z > 8$. The cosmic star formation rate density still shows, however, an order-of-magnitude increase from $z \sim 10$ to $z \sim 8$ since the dominant contribution comes from low-luminosity sources. Based on the IRAC detections, we derive galaxy stellar masses at $z \sim 10$, finding that these luminous objects are typically 10^9 $M_\odot$. This allows for a first estimate of the cosmic stellar mass density at $z \sim 10$ resulting in $\log_{10} \rho_* = 4.7^{+0.5}_{-0.8}$ $M_\odot$ Mpc^{-3} for galaxies brighter than $M_{\rm UV} \sim -18$. The remarkable brightness, and hence luminosity, of these $z \sim 9$–10 candidates will enable deep spectroscopy to determine their redshift and nature, and highlights the opportunity for the *James Webb Space Telescope* to map the buildup of galaxies at redshifts much earlier than $z \sim 10$.

Key words: galaxies: evolution – galaxies: high-redshift – galaxies: luminosity function, mass function

Online-only material: color figures

1. INTRODUCTION

The very sensitive near-infrared imaging with the Wide Field Camera 3 (WFC3/IR) on board the *Hubble Space Telescope* (*HST*) has enabled the extension of the observational frontier for galaxies to beyond $z \sim 9$, only 500 Myr after the big bang. However, detecting galaxies at such redshifts is clearly approaching the limit of what is possible with the *HST*. Despite extremely deep imaging over the Hubble Ultra-Deep Field (HUDF; Beckwith et al. 2006; Illingworth et al. 2013; Ellis et al. 2013), only one reliable $z \sim 10$ galaxy candidate could be identified over this small field and over additional wider area blank field data in the Chandra Deep Field South region (Oesch et al. 2013a). The first reported $z \sim 10$ galaxy, UDFj-39546284 (Bouwens et al. 2011a), now has an uncertain redshift based on newer data (Ellis et al. 2013; Brammer et al. 2013; Bouwens et al. 2013a). Two more $z \sim 10$ sources were detected in the Cluster Lensing and Supernova survey with Hubble (CLASH; Postman et al. 2012), making use of lensing magnification of massive foreground clusters (Zheng et al. 2012; Coe et al. 2013).

As exciting as these detections are, the small number of $z > 9$ galaxy candidates makes it quite difficult to reliably determine the cosmic star-formation rate density (SFRD) at these early times. In particular, does the cosmic SFRD increase slowly with time at $z > 8$, as seen at $z < 8$, or does it change more rapidly and dramatically as some models suggest? While some authors (e.g., Coe et al. 2013; Ellis et al. 2013) obtained results consistent with the SFRD continuing the same steady decline from $z \sim 8$ to $z \sim 10$ as observed at lower redshifts, the most extensive $z \sim 10$ galaxy search to date (Oesch et al. 2013a) found a significant drop in the SFRD by about an order of magnitude from $z \sim 8$ to $z \sim 10$ when combining all published measurements at $z > 8$. Clearly, enlarging the sample of $z > 8$ galaxies would help to establish the rate at which the cosmic SFRD increased.

With the completion of the Cosmic Assembly Near-Infrared Deep Extragalactic Legacy Survey (CANDELS; Grogin et al. 2011; Koekemoer et al. 2011) in 2013 August, it is now possible

* Based on data obtained with the *Hubble Space Telescope* operated by AURA, Inc. for NASA under contract NAS5-26555. Based on observations with the *Spitzer Space Telescope*, which is operated by the Jet Propulsion Laboratory, California Institute of Technology under NASA contract 1407.
[11] Hubble Fellow, YCAA Fellow.

Table 1
The 5σ Depths of the Observational Data Used in This Paper

Field	Area (arcmin^2)	B_{435}	V_{606}	i_{775}	I_{814}	z_{850}	Y_{105}	J_{125}	JH_{140}	H_{160}	K-band[a]	IRAC 3.6	IRAC 4.5
GOODSN-Deep	64.5	28.0	28.2	27.6	28.9	27.7	27.6	28.1	26.8	27.8	25.0–26.6	27.0	26.7
GOODSN-Wide	69.4	28.0	28.2	27.6	28.2	27.7	27.2	27.2	26.8	27.1	25.0–26.6	27.0	26.7

Notes. Depths are measured in circular apertures and are corrected to total fluxes using the flux growth curves of stars. The aperture diameters were 0.″35 for the ACS and WFC3 data, 0.″6 for the *K*-band, and 2.″0 for the IRAC data. These were chosen to be consistent with the actual aperture sizes used for photometry of our sources of interest.

[a] The depth of the MOIRCS *K*-band data varies significantly across the field due to non-uniform exposure times.

to extend the search area for $z \gtrsim 8$ galaxies. In particular, we will focus on the GOODS-N data set in this paper, where the CANDELS survey acquired F105W (Y_{105}) imaging data and where much more extensive multi-wavelength optical data are available from the Great Observatories Origins Deep Survey (GOODS; Giavalisco et al. 2004) than for the other CANDELS-Wide fields. As we have demonstrated in previous papers, the limits placed on flux measurements at shorter wavelengths play a crucial role in enhancing the reliability of high-redshift galaxy searches by removing probable low-redshift contaminants. The non-detections at wavelengths below the Lyman break (i.e., for $z \sim 9$–10 candidates the optical data together with the Y_{105} imaging) greatly lessen the problem of contamination by lower redshift sources.

With the detection of galaxies at $z > 8$, a new challenge has become to characterize the physical properties of galaxies at 450–650 Myr. A key parameter for any characterization is the stellar mass. This is possible with the use of the *Spitzer* Infrared Array Camera (IRAC), which samples rest-frame optical light even for $z \sim 10$ galaxies with its 4.5 μm channel. Unfortunately, the extremely faint *HST* $z \sim 9$–10 galaxy candidates that were identified down to $H_{160,AB} \sim 30$ mag in the HUDF field are out of reach of IRAC. The two lensed candidates of Coe et al. (2013) and Zheng et al. (2012) are sufficiently bright so that they show weak ($\sim 2\sigma$–3σ) IRAC detections, but the uncertainties are so large that the stellar mass estimates are still not very reliable. In the present paper, we search for bright $z \sim 9$–10 galaxy candidates in the GOODS-N field where very deep *Spitzer*/IRAC data are available. Such deep *Spitzer*/IRAC data would allow for a first estimate of the galaxy stellar mass density at $z \sim 10$.

This paper is an extension of our previous analyses of the GOODS-S data set and our $z \sim 10$ Lyman break galaxy (LBG) search (Bouwens et al. 2011a; Oesch et al. 2012a, 2013a). It is organized as follows. We describe the data used for this analysis in Section 2 and present our high-redshift candidate selection in Section 3. These candidates are subsequently used in Section 4 to derive new constraints on the evolution of the UV bright galaxy population out to $z \sim 10$. Section 5 provides an analysis of the stellar mass density at $z \sim 10$ based on robust IRAC detections. Section 6 summarizes our results and briefly discusses the implications of our findings for planning future surveys such as with the *James Webb Space Telescope* (*JWST*). In the Appendix we note the outcome of a search for bright $z \sim 9$–10 candidates in GOODS-S that was motivated by the discovery of such sources in GOODS-N. Furthermore, we demonstrate the technique for subtracting neighbors in the IRAC images in Appendix B.

Throughout this paper, we will refer to the *HST* filters F435W, F606W, F775W, F814W, F850LP, F105W, F125W, F140W, F160W as B_{435}, V_{606}, i_{775}, I_{814}, z_{850}, Y_{105}, J_{125}, JH_{140}, H_{160}, respectively. Magnitudes are given in the AB system (Oke & Gunn 1983), and we adopt a standard cosmology with $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$, $H_0 = 70$ km s^{-1} Mpc^{-1}, i.e., $h = 0.7$, consistent with the most recent measurements from Planck (Planck Collaboration et al. 2013).

2. DATA

The data set analyzed in this paper consists of deep, high-resolution *HST* imaging covering 0.4–1.6 μm, in addition to ground-based *K*-band data, as well as *Spitzer* IRAC imaging at 3.6 and 4.5 μm. These data sets are discussed in the next section and a summary of their depths is listed in Table 1.

2.1. HST Data in GOODS-North

We base this paper on the entire WFC3/IR and I_{814} Advanced Camera for Surveys (ACS) data over the GOODS-N field from the completed CANDELS survey. The last data were taken on 2013 August 10. For details on the survey layout we refer to the CANDELS team papers (Grogin et al. 2011; Koekemoer et al. 2011). Briefly, the CANDELS GOODS-N field is part of the CANDELS-Deep survey, for which additional Y_{105} imaging was obtained. The Y_{105} imaging is not available in the general CANDELS-Wide component.

The central $\sim$65 arcmin^2 of the GOODS-N field was covered by $\sim$5 orbits of WFC3/IR imaging data in J_{125} and H_{160} reaching to 27.8 mag (5σ) and by $\sim$3 orbits of Y_{105} imaging reaching 27.6 mag. Furthermore, two flanking fields totaling $\sim$70 arcmin^2 of the CANDELS-Wide program completed the WFC3/IR coverage of the GOODS-N field with roughly one orbit in each of the three filters, Y_{105}, J_{125}, and H_{160}, which results in a depth of 27.0–27.2 mag in all three filters. We also include all the JH_{140} imaging data available over GOODS-N. These are very shallow exposures, mostly being used as pre-imaging for the GOODS-N grism program (GO:11600, PI: Wiener) with an exposure time of 800s, in addition to a few supernova follow-up observations from the CANDELS survey. Nevertheless, these can be useful for spectral energy distribution (SED) analyses of brighter sources.

We downloaded all the individual WFC3/IR data from the Barbara A. Mikulski Archive for Space Telescopes (MAST) and the Canadian Astronomy Data Centre (CADC) and reduced the data using standard procedures, as described in detail in Illingworth et al. (2013). We used the persistence masks provided by the Space Telescope Science Institute (STScI) to mask all pixels significantly affected by persistence from previous exposures. We registered all WFC3/IR frames to the official v2 GOODS-N ACS z_{850}-band data before drizzling to a mosaic with final pixel size of 0.″06 and a tangent-plane projection aligned to the GOODS-N ACS data. The RMS maps produced by multidrizzle were rescaled to match the actual fluctuations present in the data as measured through circular apertures of 0.″35 diameter randomly placed on empty sky positions in the images.

The ACS data used here are a somewhat deeper reduction than the publicly released v2 GOODS-N imaging. We included the additional data from supernova follow-up programs (see Bouwens et al. 2007). Furthermore, we reduced the new I_{814}-filter data obtained as part of parallel imaging from the CANDELS program. The reduction includes corrections for charge transfer inefficiency and was performed analogously to the eXtreme Deep Field (XDF) data reduction (Illingworth et al. 2013). The depth of these I_{814} images surpasses all other GOODS ACS images, reaching a 5σ limit of $I_{814} = 28.2$–28.9 mag.

The angular width of the point-spread function (PSF) of the data used here is $\sim$0.″09 and $\sim$0.″16 for the ACS and WFC3/IR imaging, respectively, as measured from unsaturated stars in the field.

2.2. The IRAC Data Set

From previous $z > 8$ analyses it became clear that longer-wavelength constraints from *Spitzer*/IRAC are essential in order to remove contamination from dusty, intermediate redshift sources (e.g., Oesch et al. 2012a). Furthermore, IRAC samples the rest-frame visible at $z > 4$, which is crucial for stellar mass constraints. We therefore include all the IRAC data that are available over the GOODS-N region as part of several programs. In particular, we analyzed the 3.6 μm and 4.5 μm channel IRAC reductions from the Spitzer Extended Deep Survey (SEDS) and S-CANDELS team (PI: Fazio; see also Ashby et al. 2013).

The Spitzer S-CANDELS program (P.I. G. Fazio) is a Cycle 8 Spitzer Exploration Science project to map $\sim$0.2 deg^2 in the five CANDELS fields to 50 hr depth with IRAC, thereby reaching magnitude $\sim$26.8 at 3.6 and 4.5 μm. Data in GOODS-N were obtained in two epochs in 2012 and combined with pre-existing IRAC data from SEDS (Ashby et al. 2013) and the original GOODS program (Dickinson et al. 2003). The corrected, basic calibrated data (cBCD) frames from the *Spitzer* archive were combined into calibrated mosaics following the same procedures as for SEDS (Ashby et al. 2013). The achieved depth at the positions of the $z \sim 9$ candidates (Section 3) is $\sim$50 hr in both IRAC bands except that the 4.5 μm observation of GN-z10-1 has $\sim$72 hr. In blank sky areas the average IRAC 5σ depths are 27.0 and 26.7 mag (within 1″ radius apertures) in the two IRAC channels. Further analysis of these mosaics, including catalog creation and completeness estimates, is ongoing and will be reported elsewhere (M. L. N. Ashby et al., in preparation).

A significant challenge of the IRAC data is the poor angular resolution, FWHM = 1.″7 (Fazio et al. 2004), leading to source confusion. This can be largely overcome, however, through neighbor subtraction based on the high-resolution *HST* data. This is particularly important for the faint sources that we study here. We used an approach outlined in several previous papers (see, e.g., Labbé et al. 2006; Grazian et al. 2006; Laidler et al. 2007) that models a region around a source of interest using its *HST* H_{160} image convolved to IRAC resolution and subtracts all neighbors to give a "cleaned" source. Subsequently we measured fluxes in 1″ radius circular apertures and multiplied by a factor 2.4–2.6 to correct for light outside the aperture. The variation in the aperture correction is due to variations in the position dependent IRAC PSF as measured from nearby stars in the field.

While subtraction of the flux from the neighboring sources does not always work if a candidate is too close to a very bright foreground source, it is very effective in the majority of cases. The modeling and subtraction approach has been refined extensively and now allows us to perform clean IRAC photometry for $\sim$75%–80% of sources in the field, a factor $\sim$2$\times$ larger than without neighbor subtraction.

2.3. MOIRCS K-band Data over GOODS-N

In order to bridge the gap in wavelength range between the 1.6 μm probed by *HST* and the 3.6 μm covered by IRAC, we also made use of a deep *K*-band stack over GOODS-N from the multi-object infrared camera and spectrograph for Subaru (MOIRCS Kajisawa et al. 2006; Bouwens et al. 2008). These data were reduced using the standard MOIRCS pipeline procedures. The average seeing measured from the final stack of the data is 0.″55. Over the region of interest, this stack varies in depth between 25.0 and 26.6 mag AB, as measured in small circular apertures of 0.″6 diameter. Total magnitudes were derived using aperture corrections computed from the profiles of a few bright, non-saturated stars in the field.

2.4. Supporting HST Fields: HUDF09/12/XDF and GOODS-South

In the last section of this paper, we will derive new constraints on the UV luminosity function (LF) and the cosmic SFRD at $z \sim 10$ based on the largest possible data set. We will therefore combine the GOODS-N CANDELS data set with an identical analysis over previous deep *HST* imaging fields. In particular, we directly include the $z \sim 10$ search and analysis from the HUDF09/12/XDF and GOODS-S field from our previous paper (Oesch et al. 2013a). However, motivated by the bright galaxies in GOODS-N, we re-analyzed the GOODS-S data set with detection criteria that are better matched to those used in GOODS-N. The result of this is presented in the Appendix. The combined WFC3/IR+ACS data set spans an area of $\sim$300 arcmin^2 and ranges in depth from $H_{160} = 27.5$–30.0 mag AB (5σ).

3. THE GOODS-NORTH $z > 9$ GALAXY SAMPLE

3.1. Sample Selection

The identification of LBGs in the epoch of reionization makes use of the almost complete absorption of UV photons shortward of the redshifted Lyα line due to a high neutral hydrogen fraction in the inter-galactic medium. At $z > 9$ the Lyα absorption shifts into the J_{125} band, which renders star-forming galaxies red in their $J_{125} - H_{160}$ colors. Our initial selection criterion is therefore to search for galaxies with $J_{125} - H_{160} > 0.5$ and non-detections in the shorter wavelength data. In the second part of this paper, we will restrict the sample to a more conservative criterion with $J_{125} - H_{160} > 1.2$, which includes only galaxies at $z \gtrsim 9.5$. This selection $J_{125} - H_{160} > 1.2$ also matches our previous GOODS-S analysis and allows us to use a larger sample for the subsequent analysis.

Source catalogs were obtained with SExtractor (Bertin & Arnouts 1996), which was run in dual image mode with the H_{160}-band as the detection image. All images were convolved to the H_{160} PSF when performing photometry, and colors were measured in small Kron apertures (Kron factor 1.2), typically 0.″2 radius. Total magnitudes were derived from larger elliptical apertures using the standard Kron factor of 2.5, typically 0.″4 radius, with an additional correction to total fluxes based on the encircled flux measurements of stars in the H_{160} band. This last correction was typically $\sim$0.2 mag but depended on the actual Kron aperture size of individual galaxies.

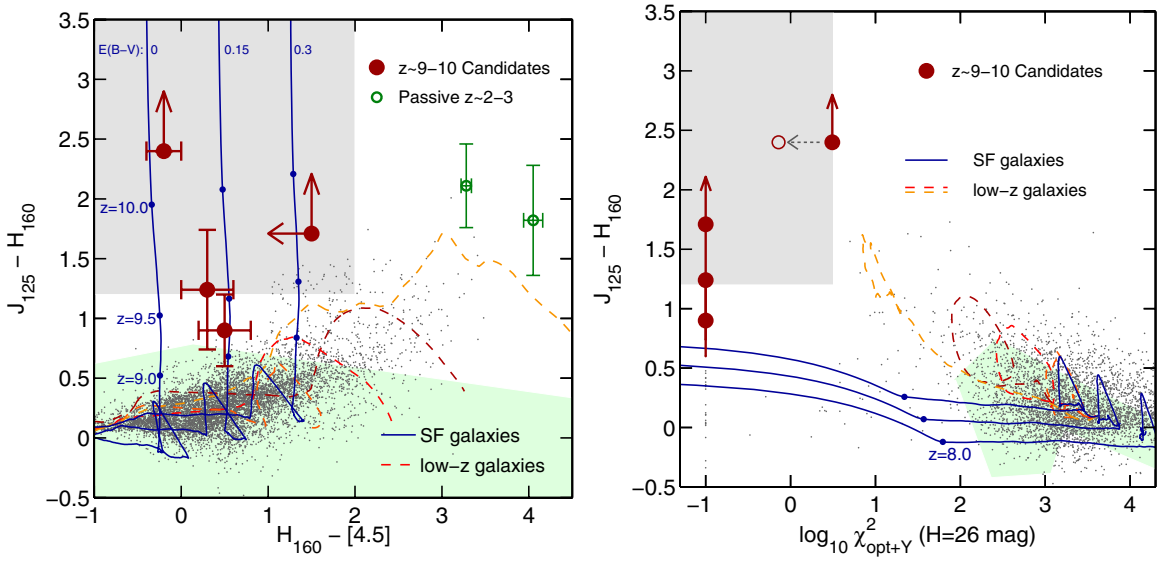

Figure 1. Left: the $J_{125} - H_{160}$ vs. $H_{160} - [4.5]$ color diagram showing the location of the GOODS-N $z > 9$ galaxy candidates relative to other, lower redshift galaxy SED tracks. The four $z \sim 9$–10 candidates are shown with dark red circles with 1σ error bars on their colors. For non-detections, 2σ color limits are shown. The only other sources identified as having no optical detections are shown as green circles. These two sources are very bright in IRAC, consistent with being dusty or passive sources at $z \sim 2$–3. They are thus not included in our subsequent analysis. Also shown as small points are the locations of all sources with reliable IRAC 4.5 μm and H_{160} flux measurements ($>10\sigma$) from the CANDELS GOODS-S catalog of Guo et al. (2013). These sources nicely follow the redshift tracks of evolved $z < 5$ SEDs (dashed yellow to red lines). The gray shaded area indicates the region in color–color space expected for $z \gtrsim 9.5$ star-forming galaxies with $J_{125} - H_{160} > 1.2$. Note that the initial selection of the four GOODS-N candidates was $J_{125} - H_{160} > 0.5$, which selects sources at $z \gtrsim 9$. The blue lines indicate the tracks of star-forming galaxies with different amounts of dust extinction ($E(B-V) = 0, 0.15, 0.3$). Visual inspection and SED fits of the few sources from the CANDELS catalog that lie in the gray area show that they are all relatively compact, intermediate redshift, passive sources. Unlike the real high-redshift candidate sources, these are significantly detected in the optical data and so they can be clearly rejected as contaminants (see the figure to the right). The green shaded area that is well-separated from the gray selection zone indicates the area where Galactic stars are expected, including very low mass M, L, T, and Y dwarfs (see also Coe et al. 2013). Right: plot of $\chi^2_{\rm opt+Y}$ (see Section 3.1) against the $J_{125} - H_{160}$ color, representing our second selection criterion for $z > 9$ galaxies (which minimizes low redshift contaminants). The four $z \sim 9$–10 candidates are again shown with dark red circles, while small gray dots indicate galaxies with 10σ H_{160} detections in our GOODS-N CANDELS-Wide catalog. Sources with $\log_{10}\chi^2_{\rm opt+Y} < -1$ are limited at that value. The shaded areas and lines represent the same as in the left panel. The $\chi^2_{\rm opt+Y}$ values for the SED tracks were normalized to $H_{160} = 26$ mag and were computed for the CANDELS-Wide field depth. The gray arrow and the open circle shows how the value of $\chi^2_{\rm opt+Y}$ changes for GN-z10-1 if the z_{850}-band is excluded in which a background fluctuation causes a positive 1.5σ flux measurement.

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

Based on these catalogs, the following *HST* selection criteria were applied:

$$(J_{125} - H_{160}) > 0.5 \quad (1)$$

$$\mathrm{S/N}(B_{435}\ \mathrm{to}\ Y_{105}) < 2 \quad \wedge \quad \chi^2_{\rm opt+Y} < 3.2$$

in addition to at least 5σ detections in H_{160} (see Figure 1). The $\chi^2_{\rm opt+Y}$ for each candidate source was computed as $\chi^2_{\rm opt+Y} = \Sigma_i \mathrm{SGN}(f_i)(f_i/\sigma_i)^2$ (Bouwens et al. 2011b) where f_i is the flux in band i in a consistent aperture, σ_i is the uncertainty in this flux, and SGN(f_i) is equal to 1 if $f_i > 0$ and -1 if $f_i < 0$, and the summation is over the B_{435}, V_{606}, i_{775}, I_{814}, z_{850}, and Y_{105} bands. The limit of $\chi^2_{\rm opt+Y} = 3.2$ was chosen to result only in a small reduction in the selection volume of real $z > 9$ sources (20% based on Gaussian statistics), while efficiently excluding lower redshift contamination (see the right panel of Figure 1). This reduction in the selection volume is accounted for in our calculations of the UV LF in Section 4.

These *HST* selection criteria resulted in a total of six potential candidates in the full GOODS-N WFC3/IR data set. However, two of these sources are extremely bright in the *Spitzer*/IRAC bands with $H_{160} - [4.5] > 3.2$. As shown in the left panel of Figure 1, this is much redder than expected for a real high-redshift galaxy. However, it is consistent with similarly red sources with photometric redshifts $z \sim 2$–4 that we identified as contaminants in our GOODS-S high-redshift search (Oesch et al. 2012a, 2013a). Such sources are very interesting for $z \sim 2$–4 studies (see, e.g., Huang et al. 2011; Caputi et al. 2012), but will not be discussed further here. We exclude these two sources from our analysis and proceed with only four potential $z > 9$ galaxy candidates.

While three out of the four remaining sources show negative values of $\chi^2_{\rm opt+Y}$, the brightest candidate (GN-z10-1) lies very close to the selection limit (see right panel of Figure 1). This is mostly driven by a 1.5σ positive flux measurement in the z_{850}-band. From a visual inspection of that image, however, it appears that this is due to a feature in the background and is not associated with real flux from the source. If we remove this band, the $\chi^2_{\rm opt+Y}$ drops to 0.7. Nevertheless, the $\chi^2_{\rm opt+Y}$ near the selection limit indicates that this source could be a potential contaminant. Based on SED fitting, however, we will show later in Section 3 that no low redshift galaxy SED or stellar SED that we know of can reproduce the very red $J_{125} - H_{160}$ color break of this source together with its flat continuum longward of 1.6 μm. Taken together, these results suggest that the most likely interpretation is that GN-z10-1 is at high redshift.

Stamps of the four viable high-redshift candidates are presented in Figure 2, and their positions and photometry are listed in Tables 2 and 3. As is evident from Figure 2, the four sources are all detected at $\geqslant 7\sigma$ in the H_{160} band. The brightest source is 15σ. Furthermore, all sources are seen in observations at other wavelengths, albeit at lower significance. With the exception of GN-z10-1, all show weak detections in J_{125}, and two are even seen weakly in the very shallow JH_{140} data. Furthermore, the

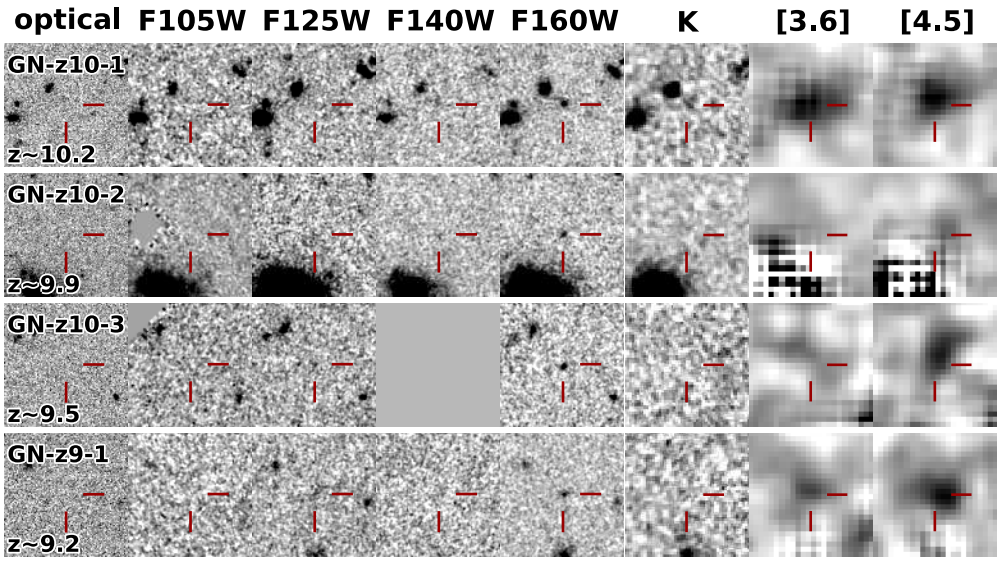

Figure 2. $6'' \times 6''$ images of the four $z \geqslant 9$ galaxy candidates identified in the CANDELS GOODS-N data. From left to right, the images show a stack of all optical bands, Y_{105}, J_{125}, JH_{140}, H_{160}, MOIRCS K, and neighbor-subtracted IRAC 3.6 μm and 4.5 μm images. The stamps are sorted from high to lower photometric redshift from SED fits (indicated in the lower left; see also Table 2). The IRAC neighbor-subtraction works well for all sources except for GN-z10-2, where the nearby foreground source is too bright, and clear residuals are visible at the location of the candidate. Only IRAC upper limits are therefore included for this source in the following analysis. Clearly, all other sources show significant (>4.5σ) detections in the 4.5 μm channel. The brightest source (GN-z10-1) is also detected at 6.9σ in the 3.6 μm channel. With the exception of the brightest candidate, which is weakly detected in the *K*-band (at 2σ), the MOIRCS *K*-band data provide only upper limits.

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

Table 2
Coordinates and Basic Photometry of $z > 9$ LBG Candidates in the GOODS-N Field

Name	ID	R.A.	Decl.	H_{160}	$J_{125} - H_{160}$	$H_{160} - [4.5]$	z_{phot}[a]
GN-z10-1	GNDJ-625464314	12:36:25.46	+62:14:31.4	25.95 ± 0.07	>2.4	-0.2 ± 0.2	10.2 ± 0.4
GN-z10-2	GNDJ-722744224	12:37:22.74	+62:14:22.4	26.81 ± 0.14	>1.7	(<1.5)[b]	9.9 ± 0.3
GN-z10-3	GNWJ-604094296	12:36:04.09	+62:14:29.6	26.76 ± 0.15	1.2 ± 0.5	0.3 ± 0.3	9.5 ± 0.4
GN-z9-1	GNDJ-652258424[c]	12:36:52.25	+62:18:42.4	26.62 ± 0.14	0.9 ± 0.3	0.5 ± 0.3	9.2 ± 0.3

Notes. Color limits are 2σ. The numbers in the source IDs are a combination of the last 5 digits of the R.A. and the last 4 digits of the declination, which results in a unique name for all sources in the GOODS-N field.
[a] Photometric redshifts listed here are derived with ZEBRA. The EAZY code and template set returns consistent redshifts within $\Delta z = 0.1$.
[b] 3σ upper limit due to uncertainties in the neighbor flux subtraction.
[c] The source GN-z9-1 does not satisfy the criterion $J_{125} - H_{160} > 1.2$ and is therefore not included in our analysis of the UV LF and SFRD evolution at $z > 8$ in Section 4, for which we combine data from several previous analyses which used that stricter criterion. Nonetheless, this is considered to be a robust detection of a $z \sim 9$ candidate galaxy. It is excluded from the analysis only because of our intent to use consistent selection criteria for the overall sample analysis.

brightest source is detected at 2σ in the ground-based *K*-band data.

Neighbor-subtraction was applied to the IRAC data of all four $z \gtrsim 9$ galaxy candidates. The resulting cleaned IRAC images are shown in the two right-hand columns of Figure 2. As can be seen, three of these sources are detected in at least one IRAC band. For source GN-z10-2, the residuals of the bright foreground neighbor are still visible, and its IRAC flux measurements are therefore highly uncertain. In order to provide some photometric constraints for this source from IRAC, we use conservative upper limits based on the RMS fluctuations in the residual image at the position of the bright foreground source. All flux measurements for these sources, together with the uncertainties are listed in Table 3.

3.2. Photometric Redshift Analysis

Figure 3 shows the SED fits to the fluxes of the four high-redshift galaxy candidates. These are derived with the photometric redshift code ZEBRA (Feldmann et al. 2006; Oesch et al. 2010b) using a large library of stellar population synthesis template models based on the library of Bruzual & Charlot (2003). Additionally, we added nebular line and continuum emission to these template SEDs in a self-consistent manner, i.e., by converting ionizing photons to H and He recombination lines (see also, e.g., Schaerer & de Barros 2009). Emission lines of other elements were added based on line ratios relative to Hβ tabulated by Anders & Fritze-v. Alvensleben (2003). The template library adopted for the SED analysis is based on both constant and exponentially declining star-formation histories of varying star-formation timescales ($\tau = 10^8$ to 10^{10} yr). All models assume a Chabrier initial mass function and a metallicity of 0.5 $Z_\odot$, and the ages range from $t = 10$ Myr to 13 Gyr. However, only SEDs with ages less than the age of the universe at a given redshift are allowed in the fit. Dust extinction is modeled following Calzetti et al. (2000).

As is evident in Figure 3, all candidates have a best-fit photometric redshift at $z \geqslant 9$ with uncertainties of

Table 3
Flux Densities of $z > 9$ LBG Candidates in the GOODS-N Field

Filter	GN-z10-1	GN-z10-2	GN-z10-3	GN-z9-1
B_{435}	7 ± 9	3 ± 6	-2 ± 6	-11 ± 11
V_{606}	2 ± 7	-3 ± 5	-5 ± 5	2 ± 8
i_{775}	5 ± 10	6 ± 7	6 ± 7	-9 ± 11
I_{814}	3 ± 7	1 ± 5	-2 ± 8	0 ± 9
z_{850}	17 ± 11	-7 ± 6	-7 ± 8	-14 ± 13
Y_{105}	-7 ± 9	-7 ± 7	-2 ± 10	-18 ± 8
J_{125}	11 ± 8	12 ± 7	23 ± 8	36 ± 9
JH_{140}	102 ± 47	85 ± 34	...	86 ± 54
H_{160}	152 ± 10	68 ± 9	73 ± 8	82 ± 11
K	137 ± 67	-45 ± 51	85 ± 261	76 ± 55
IRAC 3.6 μm	139 ± 20	(<81)[a]	39 ± 21	65 ± 18
IRAC 4.5 μm	122 ± 21	(<119)[a]	93 ± 21	125 ± 20

Notes. Measurements are given in nJy with 1σ uncertainties.
[a] 3σ upper limit due to uncertainties in the neighbor flux subtraction.

$\sigma_z = 0.3$–0.4 (1σ). These uncertainties could be reduced with deeper JH_{140} imaging data in the future. However, the high redshift nature of these sources is quite secure given the current photometry. The right panels of Figure 3 show the redshift likelihood function. The integrated probability for $z < 5$ solutions is strikingly small for each of these sources. GN-z10-3 has the highest low-redshift likelihood with only 0.2%. In this case, the best-fit low redshift SED is a combination of high dust extinction and extreme emission lines, which line up to boost the fluxes in the H_{160} and IRAC 4.5 μm bands. It is unclear how likely the occurrence of such an SED really is. However, deeper Y_{105} data or spectroscopic observations could rule out such an SED. Some first spectroscopic constraints are already available from shallow WFC3 grism observations (see Section 3.3.2).

As a cross-check, we also tested and confirmed the high-redshift solutions with the photo-z code EAZY (Brammer et al. 2008). In particular, we fit photometric redshifts with templates that include emission lines as included in the v1.1 distribution of the code.[12] The best-fit EAZY redshifts are all within 0.1 of the ZEBRA values listed in Table 2.

3.3. Possible Sample Contamination

As we will show in Section 4.1, the detection of such bright $z \sim 9$–10 galaxy candidates in the GOODS-N data set is surprising given previous constraints on UV LFs at $z > 8$. A detailed analysis of possible contamination is therefore particularly important. We discuss several possible sources of contamination in the next sections.

3.3.1. Emission Line Galaxies

Strong emission line galaxies have long been known to potentially contaminate very high-redshift sample selections. These are a particular concern in data sets which do not have very deep optical data to establish a strong spectral break through non-detections (see, e.g., Atek et al. 2011; van der Wel et al. 2011; Hayes et al. 2012). Sources with extreme rest-frame optical line emission may also contaminate $z \gtrsim 9$ samples if the $z \sim 10$ candidate UDFj-39546284 (Bouwens et al. 2011a; Oesch et al. 2012a) is any guide. In that case, the extremely deep supporting data did not result in any detection shortward of the H_{160} band, but other evidence (tentative detection of an emission line at 1.6 μm and the high luminosity of UDFj-39546284) indicates that an extreme emission line galaxy at $z \sim 2.2$ is a more likely interpretation of the current data (see Bouwens et al. 2013a; Ellis et al. 2013; Brammer et al. 2013; Capak et al. 2013).

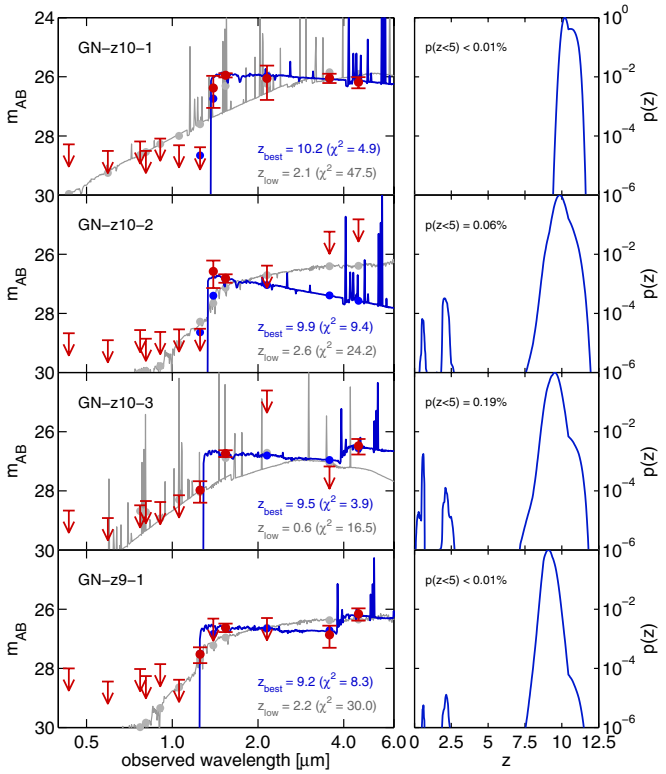

Figure 3. Spectral energy distribution fits to the *HST* and *Spitzer*/IRAC photometry of the four GOODS-N $z \sim 9$–10 galaxy candidates (left) together with the redshift likelihood functions (right). The measurements and their upper limits (2σ) are shown in dark red. Best-fit SEDs are shown as blue solid lines, in addition to the best low redshift solutions in gray. The corresponding SED magnitudes are shown as filled circles. For all sources, the $z \geqslant 9$ solution fits the observed fluxes significantly better than any of the possible low-redshift SEDs. The integrated likelihoods for $z_{\rm phot} < 5$ are all <0.2% as shown by the labels in the right panels.

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

In our SED analysis in Section 3.2, we specifically included line emission in order to test for contamination from strong emission line sources. Indeed, for two of the candidates, the best-fit low-redshift photometric redshift solutions are obtained from a combination of extreme emission lines and high dust extinction. However, all candidates are detected (although sometimes faintly) in several non-overlapping filters. For example, with the exception of GN-z10-1, all sources show some flux in the J_{125} filter, as well as a clear detection in H_{160}. It is therefore unlikely that the detected *HST* flux originates from emission lines alone. Furthermore, three of the four candidates show robust detections in the IRAC bands, which further limits the likelihood of contamination by pure line emitters. For example, GN-z10-1 (the brightest source), shows evidence for a flat continuum from the *HST* H_{160} to the IRAC 3.6 μm and 4.5 μm bands. As can be seen from Figure 3, while this can be mimicked with the combination of [O III]/Hβ contamination in the H_{160} band and continuum emission in the IRAC channels, the shorter wavelength flux limits rule out such a lower redshift solution.

Taken together, the likelihood that the sources here are lower-redshift emission line galaxies is low. The emission line constraints from the IRAC filters are discussed further in Section 5 where we present galaxy stellar mass estimates.

[12] Available at http://code.google.com/p/eazy-photoz/.

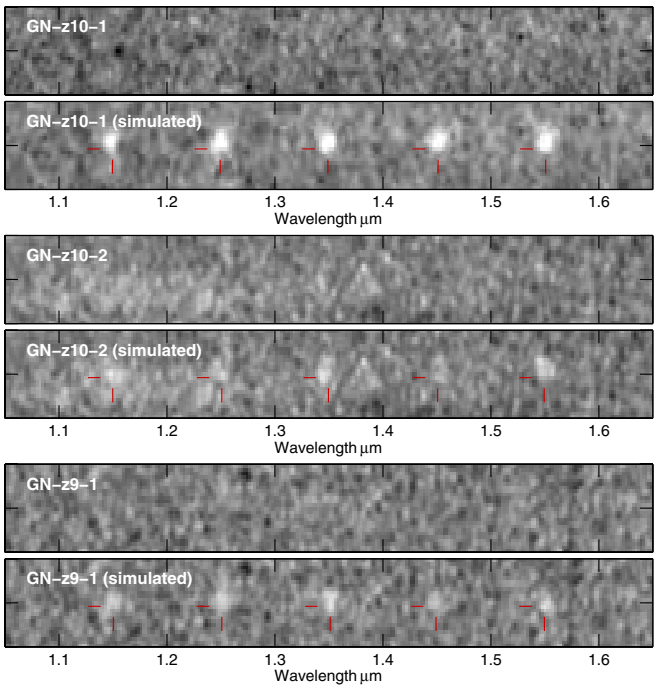

Figure 4. 2D WFC3/IR grism G141 spectra for the three sources for which data are available. These are GN-z10-1 (top two panels), GN-z10-2 (middle two panels), and GN-z9-1 (bottom panels) as labeled in the plots. The spectra of these sources are expected to run along the center of each panel in the horizontal direction. The spectra were slightly smoothed with a Gaussian. No significant line emission is detected for any of the three sources. Below the original data, we show a panel with a simulation of pure emission line sources at five different wavelengths, as indicated by red tick marks, with a line flux corresponding to the H_{160} photometry (5.5×10^{-17} erg s^{-1} cm^{-2} for the brightest source, and 2.5×10^{-17} erg s^{-1} cm^{-2} for the fainter two). Despite some residual contamination from a foreground source in the spectrum of GN-z10-2, such strong emission lines would have been significantly detected at $>4\sigma$. The grism data rule out pure emission line source contamination for these three sources.

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

3.3.2. Constraints from HST Grism Data

Quantitative constraints on pure emission line sources can be obtained from the WFC3/G141 grism observations over GOODS-N from HST program 11600 (PI: Wiener). These spectra cover $\sim$1.05–1.70 μm at low resolution, reaching a 5σ emission line flux limit for compact sources of $\sim$2–5 $\times$ 10^{-17} erg s^{-1} cm^{-2} (see Brammer et al. 2012). If the H_{160}-band flux originated from a single emission line, the observed magnitudes of our sources (H_{160} = 26.0–26.8 mag) would correspond to line fluxes of 2.5–5.5 $\times$ 10^{-17} erg s^{-1} cm^{-2}. These lines should thus be detectable as $\sim 5\sigma$ features. We have therefore analyzed the grism spectra using reductions developed by the 3D-HST team. Three sources (excluding GN-z10-3) were covered with such data.

The grism spectra were reduced using a newly developed pipeline by the 3D-HST team (G. B. Brammer et al., in preparation). The spectrum of GN-z10-1 was heavily affected by a varying sky background during the exposures, and it was necessary to manually exclude some of the affected readouts, which effectively reduced the total exposure time by less than 25%.

The final 2D spectra of the three covered sources are shown in Figure 4. For sources GN-z10-1 and GN-z9-1, the spectra are blank, showing no significant features. For source GN-z10-2, the spectrum is partially contaminated by the trace of a nearby bright source. Two extremely faint features may be visible at low significance ($\sim 2\sigma$) at $\sim$1.38 μm and 1.50 μm. However, given the foreground contamination, the data are currently inconclusive regarding these faint features.

The spectra of these three sources provide very useful constraints on contamination by pure emission line sources. Figure 4 also shows simulated spectra of sources for which a single line could explain the whole flux in the H_{160} filter, with lines at five different wavelengths. The simulated lines were computed based on the actually observed H_{160}-band profile of these sources in the spatial direction and assuming a Gaussian emission line with FWHM = 100 Å in the dispersion direction. As can be seen, such a line would be significantly detectable even in the contaminated spectrum of GN-z10-2.

The grism data therefore rule out the contamination of a pure, single emission line source for these three candidates. However, based on the current data, we cannot rule out contamination by lower redshift sources with less extreme emission lines. Deeper data would be required to do so. As we shall see, while any one constraint is not definitive, the grism emission line limits, the constraints on low-redshift contamination based on SED fitting, and those from photometric scatter discussed in Sections 3.2 and 3.3.4 indicate that the sources identified here are highly likely to be at high redshift.

3.3.3. Stellar Contamination?

Stellar contamination can be a problem for high-redshift galaxy selections due to strong absorption features in dwarf stars, and the unusual brightness of our GOODS-N sources led us to give particular attention to this aspect. We checked the surface brightness profiles of the sources, their colors, and their SEDs.

For the brighter two sources GN-z10-1 and GN-z9-1, we measure half-light radii of 0.″17 for both. After a simple correction for the stellar PSF, these result in intrinsic half-light radii of only 0.″11. At $z \sim 10$, this would correspond to a physical size of only 0.5 kpc, which, while small, is consistent with the expectations from extrapolating the $z \sim$ 4–8 size trends to $z \sim 10$ (e.g., Oesch et al. 2010a; Ono et al. 2012). Furthermore, both sources have SExtractor FWHM measurements more than 1.8 times wider than for non-saturated stars in the H_{160}-band imaging. The data suggest that these sources are resolved and so they are unlikely to be stellar contaminants. The fainter two sources GN-z10-2 and GN-z10-3, while bright, are quite compact and are not detected with high enough signal-to-noise ratio (S/N) to rule out stellar contamination purely based on their surface brightness profile.

As Figure 1 shows, contamination by stars in our sample is only expected from significant photometric scatter. Stars show colors of $J_{125} - H_{160} \lesssim 0.5$, even for very low mass dwarfs (see also Oesch et al. 2012a; Coe et al. 2013), which are typically the most important contaminants in high-redshift samples. These colors are sufficiently blue that our primary selection with $J_{125} - H_{160} > 1.2$ is not expected to show significant stellar contamination. Furthermore, stars are excluded from our selection based on the $\chi^2_{\rm opt+Y}$ measure. Even a cool dwarf star with $H_{160} = 26$ mag shows $\chi^2_{\rm opt+Y} \gtrsim 100$ (see right panel of Figure 1).

The low probability of stellar contamination is confirmed through SED fitting. In addition to galaxy templates, we fit all four candidates with stellar templates including observed dwarf spectra from Burgasser et al. (2004). None of these fit any of our galaxy candidates (likelihood for stellar contamination $<10^{-4}$),

based purely on the *HST* photometry. We therefore conclude that it is very unlikely that any of our candidates is a Galactic star.

3.3.4. Photometric Scatter Simulations

As we previously demonstrated (Oesch et al. 2013a), photometric scatter of lower-redshift sources into the selection regions can be the most important source of contamination for high-redshift LBG samples. This can be tested with photometric scatter simulations based on real galaxies in the photometric catalogs of fields where much deeper data are available in the same filters as for the CANDELS data. In particular, we make use of the XDF and the HUDF09-2 data sets, which have *HST* optical (and NIR) data in the same filters but have limits that are up to 2.5 mag fainter.

From the deeper fields, we selected sources in the magnitude range of the GOODS-N $z \sim 10$ galaxy candidates, i.e., $H_{160} =$ 26–27.5 mag, and we applied photometric scatter as measured for real sources at those magnitudes in the CANDELS data. In detail, we computed the average flux uncertainty of real sources in our CANDELS catalogs and used these to add a Gaussian perturbation to the flux measurements from the deep data.

The contamination fraction can then be estimated based on applying the LBG selection criteria to this simulated catalogs, correcting for the GOODS-N CANDELS survey area, and repeating this many times. The resulting average number of contaminants per realization is only 0.19. Not unexpectedly, these contaminants are all found closer to the magnitude limit at $H_{160} > 27.0$ mag rather than at $H_{160} < 27.0$ where the GOODS-N $z \sim 9$–10 candidates lie. This indicates that photometric scatter is not significantly contaminating our sample at the observed brightness of these candidates.

An independent estimate of the contamination fraction in the sample can also be obtained from the best low-redshift SED fits by estimating the probability with which such galaxy SEDs would be selected in the CANDELS data. We therefore applied photometric scatter to the expected magnitudes of the low-redshift SEDs shown in Figure 3 and then applied our LBG selection criteria. Repeating this 10^6 times for all four candidates results in an average contamination fraction of only 0.15 source per realization, consistent with the 0.19 estimated above from the scatter simulations based on the XDF and HUDF09-2 fields.

To summarize, the high-quality multiwavelength *HST* data and our use of all information in the optical data based on the $\chi^2_{\rm opt+Y}$ measure allows us to conclude that photometric scatter of lower redshift galaxies with known SEDs is unlikely (consistent with our analysis of the photometric redshift likelihood functions). However, we cannot rule out contamination by unusual, extremely rare sources not represented in the simulation databases. The simulations we carried out assumed sources that have SEDs or photometric characteristics consistent with documented results. While it is quite unlikely that we have identified sources with very unusual SEDs, the possibility remains, though finding four such undocumented sources seems a remote possibility. Deep spectroscopy of these candidates is the only way to ultimately establish their true nature.

3.4. AGN Contribution?

As pointed out earlier, the GOODS-N $z \sim 10$ galaxy candidates are very compact and very bright. This raises the possibility that some of these sources host an active galactic nucleus (AGN), which contributes or even dominates the observed fluxes in the H_{160}-band. While it would be surprising (though very interesting) to see significant AGN activity just a few hundred million years after the formation of the first stars, without spectroscopic observations, it is of course nearly impossible to reliably assess such a contribution. However, our imaging data can provide some first constraints.

At least the two brighter sources in our sample appear to be resolved. However, as is usually the case for compact high redshift sources, the current data do not exclude a point source in the center of a more extended star forming region. For such objects, probably the most stringent constraint on an AGN contribution can be obtained from a variability analysis. AGN flux variations are seen over essentially all timescales and would be a clear indicator for nuclear accretion activity (for a review see, e.g., Ulrich et al. 1997).

Given that the GOODS-N CANDELS data were acquired over almost a two-year timescale, we can directly test for variability. We have therefore split the data into five separate subsplits, sorted by acquisition time, and have reduced these frames separately. This was done for the H_{160}-band images of each of our candidates. We then remeasured the magnitudes of these sources via the dual image mode of SExtractor with the full stack as the detection image. The resulting flux variation is shown in Figure 5.

For two of our sources, we do not detect a significant signal ($>2\sigma$) at a couple of epochs. This results from the low exposure times in those stacks when the source falls on masked regions of the WFC3/IR detector. Thus the lack of data in those epochs is not indicative of real variation.

Evaluation of the measurements for each source shows that none of the sources displays a statistically significant variation. This provides some indicative evidence against AGN contributions in these candidates. Due to the limited depth of the data, however, smaller amplitude flux variability (of the order of a couple of tenths of a magnitude) cannot be ruled out.

3.5. Possible Lensing Magnification

Given the brightness of our candidates, it is interesting to ask whether any of these could be significantly magnified by a foreground source. Even though none of the candidates appears to be highly magnified (none have the significant elongation which would be a clear sign for very high magnification), smaller values of magnification are possible. Wyithe et al. (2011) estimated that very high-redshift luminosity functions could be significantly distorted due to a magnification bias once the characteristic magnitude lies ~2 mag below the survey limit. This could indeed be the case with our GOODS-N data.

We therefore examined the neighbors of all four candidates. The two lowest-redshift ones do not show a very bright source nearby and are therefore unlikely to be affected by lensing. However, the two highest-redshift sources do show neighbors within 2.″9 and 1.″2 (see the top two rows in Figure 2). To see if magnification was contributing to their unusual brightness, we estimated their possible magnification bias based on the simplified assumption of a Singular Isothermal Sphere (SIS) lens (see, e.g., Schneider et al. 2006).

1. *GN-z10-1.* Our highest-redshift candidate shows a neighbor with $H_{160} = 24.7$ mag and half-light radius $r_{1/2} =$0.″2 at a distance of 1.″2. Our photometric redshift analysis of this source indicates that it likely lies at $z_{\rm phot} = 1.8$, and has a stellar mass of $\log_{10} M = 9.1\ M_\odot$. While we do not have any information on the velocity dispersion σ_v of this galaxy, we can obtain a rough estimate based on the virial

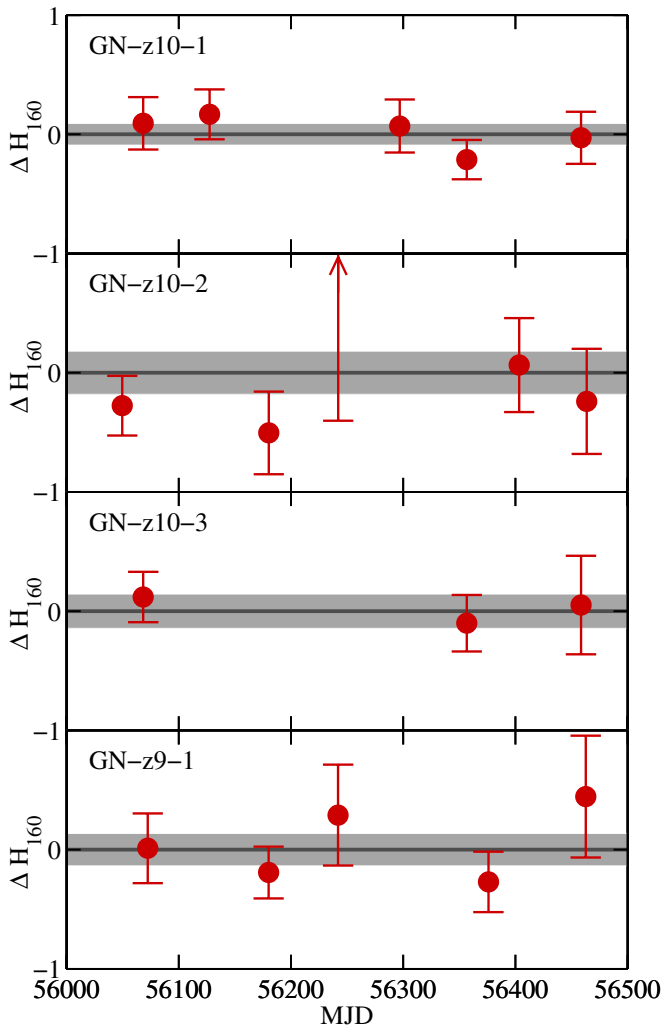

Figure 5. Variability of our four GOODS-N $z \sim 9$–10 candidates in the H_{160}-band. The CANDELS GOODS-N H_{160} exposures in the relevant regions were sorted by their acquisition time and binned in ∼100 day wide bins in order to obtain five independent H_{160} images at different times. The bins were chosen as a compromise between having a significant number of exposures and having adequate S/N per bin. Given the data acquisition schedule, the times vary significantly from region to region. The figure shows the magnitude differences with 1σ error bars for each source at the median acquisition time of each stack around that particular source as a function of the Modified Julian Date (MJD). Positive values correspond to fainter magnitude measurements. The gray region corresponds to the final magnitude uncertainty for each source. For almost all subsets a $>2\sigma$ detection is seen at the expected position. The exceptions occurred for the middle two sources. Source GN-z10-3 is in the CANDELS-Wide area and is thus covered by fewer exposures than the other sources. In two of our splits, the source landed on a masked part of the WFC3/IR detector and therefore could not be detected in those images. For GN-z10-2, we find no significant detection in the middle stack, where we show a 2σ upper limit on the source flux. Overall, the testing showed that there was no evidence for variability on timescales of 100 days at levels over a couple of tenths of a magnitude for any of these candidates.

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

theorem. Assuming that all the mass is contained within $2r_{1/2}$, we estimate $\sigma_v = 27$ km s^{-1}. This seems low and so we took a more conservative assumption of a SIS with $\sigma_v = 50$ km s^{-1}. Even for this dispersion this source has an Einstein radius for lensing a $z \sim 10$ source of $<0\farcs04$, resulting in a magnification of $<4\%$ at the separation of the candidate.

2. *GN-z10-2.* This candidate lies $2\farcs9$ from a bright galaxy at a spectroscopic redshift $z_{\rm spec} = 1.02$ (Barger et al. 2008). The foreground source has $\log_{10} M = 10.8\,M_\odot$ and half-light radius of $r_{1/2} = 0\farcs5$, resulting in an estimate of $\sigma_v \sim 125$ km s^{-1}. A SIS model with these parameters has an Einstein radius of $0\farcs3$ for lensing a $z \sim 10$ galaxy resulting in a possible magnification of 11% for this source.

From the above considerations, we conclude that lensing magnification is most likely not significant for our sample, amounting to at most 0.1 mag. Given the small magnifications and the uncertainties in the above estimates, we do not correct for any possible magnification.

4. IMPLICATIONS FOR THE GALAXY POPULATION AT $z > 8$

The detection of four very bright $z > 9$ galaxy candidates in GOODS-N is quite surprising given the dearth of candidates in the very similar GOODS-S data as well as in the much deeper data in the three HUDF09 fields. In this section, we present the implications of these detections for the interpretation of galaxy evolution at $z > 8$. In particular, we will combine the GOODS-N data with our previous $z \sim 10$ search over the GOODS-S and the HUDF09/XDF fields. In order to do so, we restrict our analysis to the three sources in the GOODS-N sample that satisfy a more stringent $J_{125} - H_{160} > 1.2$ criterion. This excludes only the $z \sim 9$ candidate GN-z9-1, while the other three GOODS-N candidates which have $z_{\rm phot} = 9.5$–10.2 are included.

Motivated by the discovery of the bright sources in GOODS-N, we also re-analyzed the CANDELS GOODS-S data with search criteria that are better matched to those used in GOODS-N. As discussed in detail in the Appendix, we indeed identified two potential $z > 9$ galaxy candidates which are also relatively bright ($H_{160} = 26.6$ and 26.9). One of them (GS-z10-1) shows $J_{125} - H_{160} > 1.2$ and has a photometric redshift of $z_{\rm phot} = 9.9 \pm 0.5$. We will therefore also include this additional, new source in our subsequent analysis.

In Section 4.5, we will determine the evolution of the cosmic SFRD to $z \sim 10$, for which we include the two additional $z \sim 10$ candidates from the CLASH survey (Zheng et al. 2012; Coe et al. 2013). Given the uncertainties in the lensing magnification we will not include these sources in our constraints on the UV LF, however.

4.1. The Expected Abundance of $z \sim 10$ Galaxies in GOODS-N

In order to compute the expected number of sources for a given UV LF, we use extensive simulations of artificial galaxies inserted in the real data to estimate the selection volume. The artificial sources were detected and re-selected in the same manner as the original sources, from which we estimate the completeness $C(m)$ and selection probabilities $S(z, m)$ as a function of H_{160} magnitude m and redshift z. These simulations allow us to statistically correct for the fact that our catalogs are missing a fraction of real high-redshift sources due to blending with foreground galaxies which amounts to a typical incompleteness of ∼20% across these HST fields. For more information on the simulation setup, see Oesch et al. (2013a).

Given the selection function and the completeness, we can compute the number of expected sources in bins of magnitude for a given LF $\phi(M)$:

$$N^{\rm exp}(m) = \int_{\Delta m} dm \int dz \frac{dV}{dz} S(m,z) C(m) \phi(M[m,z]). \quad (2)$$

Figure 6 shows the results of this calculation for several assumed LFs. Most importantly, for the best-fit $z \sim 10$ LF from

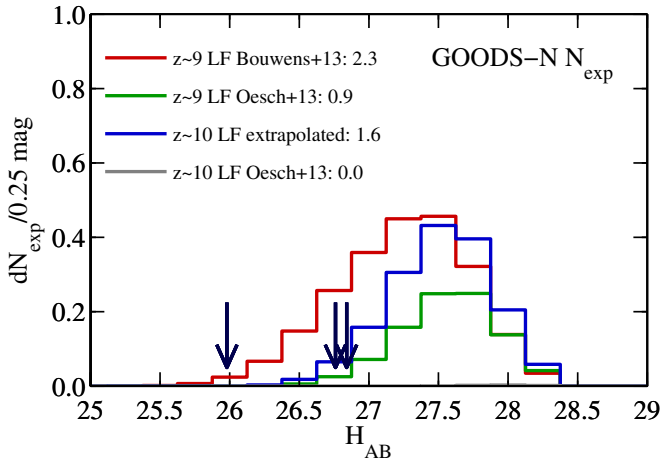

Figure 6. Expected number of $z \sim 10$ galaxy candidates per bin of 0.25 mag in the GOODS-N field for different assumptions about the UV LFs. These include previous estimates at $z \sim 9$, $z \sim 10$, and an extrapolation of the lower redshift UV LF trends to $z \sim 10$. The total expected numbers for each assumption are indicated in the legend. The magnitudes of the three observed candidate $z \sim 10$ galaxies in GOODS-N are indicated by black arrows. All the tested LFs produce a distribution which peaks at significantly fainter magnitudes than observed in the sample. Furthermore, all assumed LFs result in a lower number of total expected sources at the magnitudes of the three that are observed. In particular, not a single source would be expected to be seen in GOODS-N if the estimated UV LF from the HUDF09/12 and GOODS-S fields (Oesch et al. 2013a) was correct. The detection of three bright $z \sim 10$ candidates has significant consequences on the best-fit UV LF at $z \sim 10$ as we show in Section 4.4.

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

the GOODS-S+HUDF09/12 (Oesch et al. 2013a) not a single $z \sim 10$ galaxy candidate was expected to be seen in GOODS-N.

Additionally, we tested the expected number of sources based on a simple extrapolation of lower redshift LF trends to $z \sim 10$. This extrapolation is based on the parameterization of Bouwens et al. (2011b), who measured the UV LF evolution across $z \sim 4$ to $z \sim 8$ and found: $\phi^* = 1.14 \times 10^{-3}$ Mpc^{-3} mag^{-1} = const, $\alpha = -1.73$ = const and $M*(z) = -20.29+0.33\times(z-6)$. This extrapolation results in an assumed $M*(z = 10) = -18.97$, and predicts 1.6 sources overall in GOODS-N, but only 0.14 at $H_{160} < 27$ mag.

Very similar numbers are predicted by the theoretical $z \sim 10$ UV LF model of Tacchella et al. (2013), from which one would have expected to see 1.9 sources overall in GOODS-N, of which only 0.5 were expected at $H_{160} < 27$ mag. Even for other test LFs, such as the $z \sim 9$ UV LF estimates from Oesch et al. (2013a) and Bouwens et al. (2012a), only ~2 sources were expected to be seen in GOODS-N. However, they are expected to be fainter with $H_{160} > 27$ mag. The detection of three bright sources at $H_{160} < 27$ mag is therefore quite unexpected, given our previous constraints on the UV LFs from multiple surveys. The three GOODS-N $z \sim 10$ candidates should therefore have interesting implications for the LF at early times and also for the cosmic SFRD evolution at $z > 8$.

It is instructive to see what must be done with current UV LFs to get as many as three bright $z \sim 9$–10 galaxies. For example, 3.4 bright ($H_{160} < 26.8$ mag) $z \sim 9$–10 candidates are predicted in the GOODS-N data only if the UV LF at $z \sim 9$–10 was the same as at $z \sim 8$ (using, e.g., the results of McLure et al. 2013). However, using the same, unchanged $z \sim 8$, LF predicts a total of 11 $z \sim 10$ candidates in the GOODS-N data including fainter sources, and it predicts 42 sources in the whole combined search field (see Section 4.3). This clearly is not the case and is securely ruled out.

4.2. Discussion of Cosmic Variance

As we outlined in previous sections, we find no reason that the GOODS-N sample is heavily contaminated by lower redshift sources. However, the detection of three bright candidates at $H_{160} < 27$ mag was quite unexpected based on previous analyses of the GOODS-S data set. The question arises how this finding could be affected by cosmic variance.

We used the publicly available cosmic variance calculator[13] of Trenti & Stiavelli (2008) to estimate the likely impact of this on $z \sim 10$ galaxy searches (see also Robertson 2010). Based on a simple halo abundance matching, one expects a cosmic variance of 40%–45% per 4.7 arcmin^2 WFC3/IR pointing, depending on the assumptions about the halo occupation fraction. For the field layout of the ~150 arcmin^2 GOODS-N or GOODS-S WFC3/IR data, the expected cosmic variance ranges between 15 to 20%.

Given the very low number of expected sources in each survey, the variance is completely dominated by Poissonian statistics. For instance, the chance of finding three or more $z \sim 10$ galaxy candidates in the GOODS-N field when 1.6 sources are expected is 22%, independent of whether one assumes a 20% cosmic variance or not, on top of Poissonian statistics.

What the analysis of the GOODS-N data shows, is that larger data sets have to be analyzed for reliable measurements of the UV LFs at very high redshifts in order to overcome the limitations of Poissonian statistics. It will therefore be interesting to explore the upcoming HST Frontier Fields, which will add another 8 to 12 deep field pointings in which one would expect ~0.5–1 sources each for a (hopefully) much more reliable sampling of the UV LF, particularly at intermediate magnitudes.

4.3. The Combined $z \sim 10$ Galaxy Sample from the GOODS-N+S and HUDF09/XDF

In order to constrain the cosmic SFRD evolution at $z > 8$, we combine our analysis of GOODS-N with previous $z \sim 10$ galaxy searches. In particular, we directly use the results from Oesch et al. (2012b) and Oesch et al. (2013a), who analyzed all three ultra-deep HUDF09 fields (including the new HUDF12 data in the XDF reduction) as well as the complete CANDELS GOODS-S data. In these fields, only one viable $z \sim 10$ candidate was previously identified satisfying $J_{125} - H_{160} > 1.2$. This source (XDFj-38126243) is extremely faint with $H_{160} = 29.8$. It was found in the deepest WFC3/IR imaging available in the HUDF/XDF field.

We have re-analyzed the CANDELS GOODS-S data and found one additional bright $z \sim 10$ galaxy candidate. This source, GS-z10-1, together with a slightly lower redshift candidate, is discussed in detail in the Appendix.

The total $z \sim 10$ galaxy sample with $J_{125} - H_{160} > 1.2$ that is used in the reminder of this paper thus comprises three bright sources in GOODS-N, one bright candidate from GOODS-S, and one faint source from the XDF data. The most striking feature of this sample is that no sources are found at intermediate magnitudes at $H_{160} = 27$–29 (see, e.g., Figure 7). This, along with the very small sample size, will make it very challenging to derive a reliable LF (see the next section).

The selection functions and completeness curves for the GOODS-S and HUDF09 fields have previously been computed by Oesch et al. (2012a, 2013a). The use of an identical approach

[13] http://casa.colorado.edu/~trenti/CosmicVariance.html

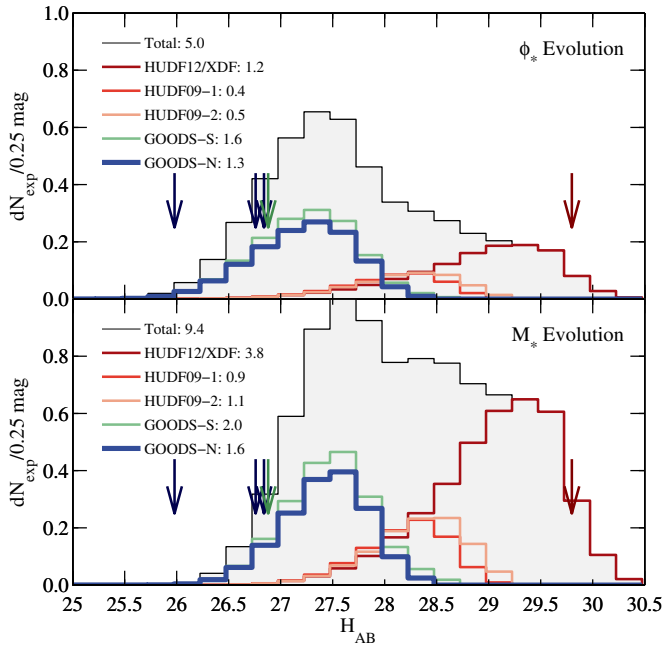

Figure 7. Expected number of $z \sim 10$ candidates per bin of 0.25 mag in the different fields used in our analysis for the best-fit UV LFs under the different assumptions of ϕ_*-only (top) and M_*-only evolution (bottom). The different lines correspond to different survey fields as shown in the legend, while the shaded gray region corresponds to the total in our analysis. The legends list the breakdown of expected sources in each field for the two assumed LF evolutions. The downward pointing arrows indicate the magnitudes of the three GOODS-N $z \sim 10$ candidates (dark blue), the one GOODS-S source (green), and the XDF candidate (dark red). Interestingly, the detected candidates only cover the tails of the expected magnitude distribution, with no sources being detected around any of the peaks. The two assumptions result in quite different magnitude distributions and expected number of candidates in fields of different depths, which has important consequences for planning future surveys for such high-redshift sources.

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

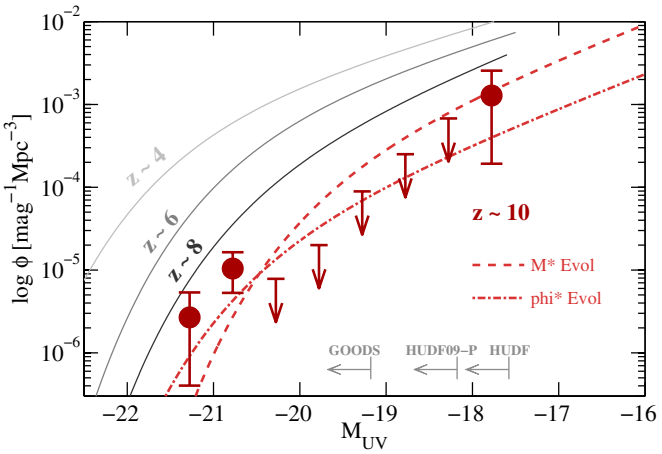

Figure 8. Improved constraints on the $z \sim 10$ UV LF from the combined $z \sim 10$ search using the blank-field GOODS-N, GOODS-S, and HUDF09/12/XDF WFC3/IR data sets. The additional data from fields other than the GOODS-N are taken directly from Oesch et al. (2013a). The dark red circles indicate the step-wise UV LF estimates in bins of 0.5 mag using the four GOODS-N+S and the one HUDF12/XDF $z \sim 10$ galaxy candidates, satisfying $J_{125} - H_{160} > 1.2$. Upper limits are 1σ. The dashed line represents the best-fit M_*-only evolution relative to the $z \sim 8$ UV LF, while the dot-dashed line shows the same for ϕ_*-only evolution. Lower redshift LFs are shown as gray solid lines for illustration of the LF evolution trends (Bouwens et al. 2007, 2012b; McLure et al. 2013). Evolution in ϕ_* appears to better match the full data set including the new results from GOODS-N, but non-Schechter LFs may provide a better fit. We have not included the CLASH survey candidates (Zheng et al. 2012; Coe et al. 2013) given the uncertainties in the lensing magnification.

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

to the GOODS-N analysis in this paper allows us to directly combine all search fields for a total measurement of the galaxy number density at $z \sim 10$.

4.4. Improved Constraints on the UV Luminosity Function at $z \sim 10$

The dearth of $z \sim 10$ candidate sources in the intermediate magnitude range $H_{160} = 27$–29 mag (see Figure 7), makes it challenging to provide a meaningful Schechter LF fit (Schechter 1976) to the observed sources. A simple power-law might provide a better description of the UV LF at such high redshifts. However, the widespread use of Schechter LF fits at lower redshifts $z \sim 4$–8 and in previous papers at $z \sim 9$–10 suggests that use of the same formalism at $z \sim 10$ is useful for comparative purposes. Furthermore, theoretical models and simulations still point toward a Schechter-like function (e.g., Trenti et al. 2010; Lacey et al. 2011; Tacchella et al. 2013). We thus update our previous estimates of the Schechter function parameters based on the combined data set of GOODS-N/S+HUDF09/HUDF12/XDF.

In our previous analysis we assumed the characteristic luminosity, M_*, to be the main parameter of the Schechter function to evolve to higher redshift. This was motivated by previous $z \sim 4$–8 measurements of the UV LF. However, this assumption is called into question with the three detections in GOODS-N and the one bright source in GOODS-S, because, as we shall see below, an M_*-only evolution results in a substantial over-prediction of the total number of candidates in our search fields.

In order to show this, we first determine our baseline lower redshift UV LF model, relative to which we will measure the evolutionary trends. Over the last few years, several $z \sim 8$ UV LF determinations have been published by several teams based on WFC3/IR data sets (e.g., Bouwens et al. 2010b, 2011b; Yan et al. 2011; Bradley et al. 2012; Oesch et al. 2012b; Lorenzoni et al. 2013; Schenker et al. 2013; McLure et al. 2013). The most recent determinations among these that use several search fields are all in good agreement with each other, and returned consistent estimates of the $z \sim 8$ UV LF Schechter function parameters. As a baseline model we adopt the values from McLure et al. (2013) which represents the widest area study to date and its UV LF parameters represent a good average of recent results from several teams (see, e.g., Table 6 of Schenker et al. 2013).

Hence, for the $z \sim 8$ baseline, we adopt $\log_{10}\phi_*(z = 8) = -3.35$ Mpc^{-3} mag^{-1}, $M_*(z = 8) = -20.12$ mag, and $\alpha(z = 8) = -2.02$ (McLure et al. 2013). We then estimate the $z \sim 10$ UV LF parameters relative to this baseline model by varying one parameter at a time. In particular, we test for M_*- and ϕ_*-evolution.

The best-fit parameters were determined by minimizing the Poissonian likelihood of observing $N_{\rm obs}$ sources in a given magnitude bin when $N_{\rm exp}$ are expected from the LF: $\mathcal{L} = \prod_j \prod_i P(N^{\rm obs}_{j,i}, N^{\rm exp}_{j,i})$, where j runs over all fields, i runs over the magnitude bins of width 0.5 mag, and P is the Poissonian probability.

Doing so for M_*-only evolution relative to the baseline model results in a best-fit estimate of $M_*(z = 10) = -19.36 \pm 0.15$. The expected magnitude distribution of $z \sim 10$ candidates for this LF is shown in the lower panel of Figure 7, and the LF itself is shown as the dashed line in Figure 8. This determination lies

Table 4
Summary of $z \sim 10$ UV LF and SFRD Estimates

	$\log_{10}\phi_*$ ($\mathrm{Mpc^{-3}\ mag^{-1}}$)	$M_{\rm UV}$[a] (mag)	α	$N_{\rm exp}^{\rm tot}$ [a]	$\log_{10}\dot{\rho}_*$ [b] ($M_\odot\ \mathrm{yr^{-1}\ Mpc^{-3}}$)
This Work (from candidates)	...	...	...	...	-3.25 ± 0.35
This Work (ϕ_* evolution)	-4.27 ± 0.21	-20.12 (fixed)	-2.02 (fixed)	$5.0^{+3.4}_{-2.2}$	-3.22 ± 0.21
This Work (M_* evolution)	-3.35 (fixed)	-19.36 ± 0.15	-2.02 (fixed)	$9.4^{+4.2}_{-3.0}$	-2.80 ± 0.11
Oesch et al. (2013a)	-2.94 (fixed)	-17.7 ± 0.7	-1.73 (fixed)	$0.6^{+2.5}_{-0.5}$	$-3.7^{+0.7}_{-0.9}$

Notes.
[a] Total number of $z \sim 10$ candidates with $J_{125} - H_{160} > 1.2$ expected to be seen in all the search fields of this paper. These include GOODS-North, GOODS-South, the HUDF09 parallel fields as well as the HUDF12/XDF field, in which we identified a total of five candidate galaxies.
[b] The SFRD measurement is limited at $M_{\rm UV} < -17.7$, the luminosity limit of the HUDF12/XDF data.

significantly above the upper limits at intermediate magnitudes. With $9.4^{+4.2}_{-3.0}$, the total expected number of $z \sim 10$ galaxy candidates from this LF is also larger than the five observed sources, though the difference is not very significant. A larger disagreement arises since these nine sources would all be expected at magnitudes fainter than $H_{160} \sim 27$ mag in the different search fields (Figure 7). Yet these are not seen.

A somewhat more consistent result is achieved from a fit with ϕ_* evolution. Assuming again the baseline $z \sim 8$ LF parameters and varying only ϕ_*, we find a best-fit $\log_{10}\phi_* = -4.27 \pm 0.21$ $\mathrm{Mpc^{-3}\ mag^{-1}}$, almost an order of magnitude lower than the $z \sim 8$ normalization. This LF is shown in Figure 8. As can be seen, it represents a better compromise between the detections at the bright and faint end and the upper limits at intermediate luminosities. The total expected number of sources for this model is $5.0^{+3.4}_{-2.2}$, consistent with the observed number of sources, but again with a magnitude distribution which peaks at $H_{160} = 27$–29 mag, where we do not detect any candidates (see Figure 7). These best-fit UV LF parameters and the corresponding total number of expected sources are summarized in Table 4, and the stepwise $z \sim 10$ UV LF constraints are tabulated in Table 5.

While previous results in the literature suggested the main evolving parameter of the UV LF at $z \sim 4$–8 to be predominantly the characteristic luminosity L^*, evolution that is dominantly ϕ^* is consistent within the uncertainties. In fact, evolution in the normalization of the UV LF may be more easily accommodated by current theoretical models. As discussed in Section 5.5 of Bouwens et al. (2008), one challenge with L^* evolution being the dominant form of evolution in the LF is that it requires some physical mechanism to impose a cut-off at a specific luminosity (and likely mass) in the *UV* LF and for that luminosity to depend on redshift. Since it is not clear what physical process would cause the cut-off to depend on redshift, simulators often find very little evolution in L^* (e.g., Jaacks et al. 2012a).

Our concern, however, is that even the best-fit UV LFs we find here do not really reproduce well the high number of bright galaxy candidates we found in the data. Even for the best-fit LF from ϕ_*-evolution, we only expect to see 0.4 sources brighter than $H_{160} = 27$ mag in GOODS-N, and 0.9 in all fields combined. Detecting three such bright sources in the GOODS-N alone is quite unlikely, with a probability that is only 0.8% from Poissonian statistics (i.e., it is a 2.4σ event).

Given our somewhat-improbable mix of $z \sim 10$ detections and limits, i.e., with our $z \sim 10$ candidate galaxies only being found at the extrema of the luminosity range probed (see Figure 8), one other possibility we have to consider is a non-Schechter-like form for the LF at $z > 6$, as has already been speculated elsewhere (e.g., Bouwens et al. 2011b; Bowler et al. 2012). It may be that this could be caused by very bursty and highly biased star-formation events at very high redshift with low duty cycle (see, e.g., Jaacks et al. 2012b; Wyithe et al. 2013).

Table 5
Stepwise $z \sim 10$ UV LF Based on the Full Data Set

$M_{\rm UV}$ (mag)	ϕ_* ($10^{-3}\ \mathrm{Mpc^{-3}\ mag^{-1}}$)
-21.28	$0.0027^{+0.0027}_{-0.0023}$
-20.78	$0.010^{+0.006}_{-0.005}$
-20.28	<0.0078
-19.78	<0.020
-19.28	<0.089
-18.78	<0.25
-18.28	<0.68
-17.78	$1.3^{+1.3}_{-1.1}$

Note. Limits are 1σ for a non-detection.

4.5. The Evolution of the Cosmic SFRD at $z > 8$

The most recent WFC3/IR data sets from several independent surveys have enabled determinations of the cosmic SFRD at redshifts $z \sim 9$ and $z \sim 10$ that were thought to be largely inaccessible for quantitative constraints on LFs or the SFRD. Estimates of the SFRD have been performed independently from a small sample of sources from the CLASH program (Bouwens et al. 2012a; Zheng et al. 2012; Coe et al. 2013) and from the HUDF09 and HUDF12 surveys (Bouwens et al. 2011b; Oesch et al. 2012a, 2013a; Ellis et al. 2013; McLure et al. 2013). While the conclusions from these separate analyses disagree on the SFRD evolutionary trends, Oesch et al. (2013a) have shown that combining all the results from the literature, the SFRD appears to increase by an order of magnitude in just 170 Myr from $z \sim 10$ to $z \sim 8$, down to the current detection limits of the HUDF09/12/XDF data set.

Here we extend our previous analysis from the CANDELS GOODS-S and HUDF09/12/XDF for an updated estimate of the cosmic SFRD at $z \sim 10$ with the inclusion of the new sources found in the GOODS-N field, and in the GOODS-S. Figure 9 shows the new results plus other measurements at $z > 8$ from the literature and the cosmic SFRD evolution constrained by these data.

The SFRD was computed directly from the observed UV luminosity density (LD) of the three GOODS-N, the single bright GOODS-S, and the one XDF $z \sim 10$ galaxy candidates. The LDs were converted to a SFRD using the conversion factor of Madau et al. (1998), assuming a Salpeter initial mass

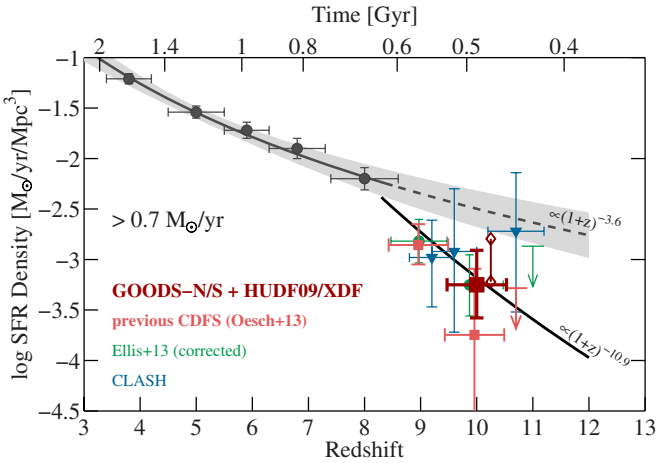

Figure 9. Redshift evolution of the cosmic star-formation rate density (SFRD) $\dot{\rho}_*$ above a star-formation limit $>0.7\,M_\odot$ yr^{-1} including the new GOODS $z\sim10$ galaxy candidates. The lower redshift SFRD estimates are based on LBG UV LFs from Bouwens et al. (2007, 2012b) including dust corrections. The gray band represents their 1σ uncertainty. The new measurement from the five detected candidates in the combined CANDELS GOODS-N/S and the HUDF09/12/XDF data set is shown as the dark red square. The individual SFRD with error bars were computed from the UV LD of the individually detected sources. Open diamonds connected with a vertical line represent the SFRDs as estimated based on integrating the best-fit UV LFs down to the corresponding luminosity limit of $M_{\rm UV}=-17.7$ (see Table 4). The upper diamond represents M_* evolution, and the lower diamond is derived from ϕ_* evolution. These estimates are offset to $z=10.25$ for clarity. Previous measurements of the SFRD at $z>8$ are shown from a combination of HUDF09/12+GOODS-S (pale red; Oesch et al. 2013a) as well as from CLASH cluster detections (blue triangles; Bouwens et al. 2012a; Coe et al. 2013; Zheng et al. 2012). Additionally, we also show the results of the HUDF12 field only from Ellis et al. (2013; green circles). We corrected down their $z\sim10$ point by a factor $2\times$ to account for our removal of a source that was shown to be a diffraction spike (see Oesch et al. 2013a). When combining all the measurements of the SFRD at $z\geqslant8$ from different fields we find $\log\dot{\rho}_*\propto(1+z)^{-10.9\pm2.5}$ (black solid line), significantly steeper than the lower redshift trends which only fall off as $(1+z)^{-3.6}$ (gray line). The current data at $z>8$ show that the cosmic SFRD is very likely to increase dramatically, by roughly an order of magnitude, in the 170 Myr from $z\sim10$ to $z\sim8$.

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

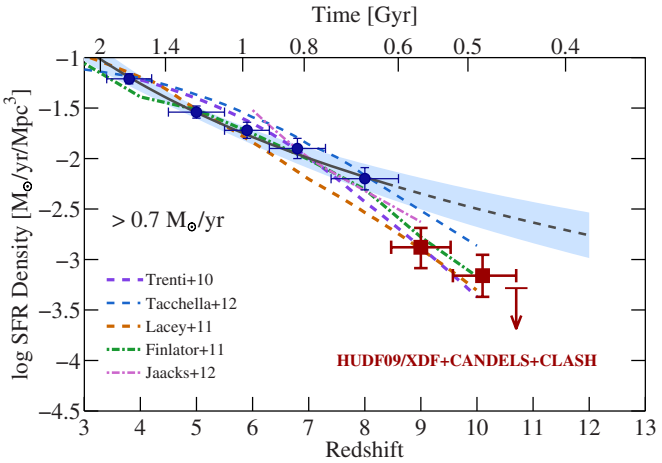

Figure 10. A comparison of the observed SFRD evolution (limited at $>0.7\,M_\odot$/yr) with model predictions. Here we combined all the measurements from our analysis with the CLASH results (Bouwens et al. 2012a; Coe et al. 2013; Zheng et al. 2012) at $z\sim9$ and $z\sim10$ (dark red squares). The lower redshift points (dark blue) are the same as in the previous figure. The model curves are based on semi-analytical/empirical modeling (Trenti et al. 2010; Lacey et al. 2011; Tacchella et al. 2013) and on SPH simulations (Finlator et al. 2011; Jaacks et al. 2013). The simulation curves were converted from UV luminosity densities and corrected for dust extinction by the same amount as the observational measurements (where necessary). All of these models predict a significant downturn in the observed SFRD at $z\sim8$, consistent with the observations.

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

function.[14] Since the HUDF12/XDF data reaches down to $M_{\rm UV}=-17.7$ mag, the derived SFRD is limited at SFR $>0.7\,M_\odot$ yr^{-1}. For the $z>8$ points, we did not perform any dust correction, because it is expected to be negligible based on the evolution of the UV continuum slope distribution at lower redshift (e.g., Bouwens et al. 2012c, 2013b; Dunlop et al. 2013; Finkelstein et al. 2012; Wilkins et al. 2011).

The direct SFRD from the five observed candidates is $\log_{10}\dot{\rho}_*=-3.25\pm0.35\ M_\odot$ yr^{-1} Mpc^{-3}. As can be seen from the summary in Table 4, this is a factor 0.45 dex higher compared to our previous estimate using only the one HUDF12/XDF candidate. However, it remains quite consistent with our previous estimate that a very large change occurs in the SFRD in the 170 Myr from $z\sim10$ to $z\sim8$. With the new data, the build-up remains strong at 1.05 ± 0.38 dex from $z\sim10$ to $z\sim8$, i.e., by an order of magnitude.

Together with the direct SFRD as measured from the five detected sources in the XDF and GOODS-N+S, Figure 9 also shows the SFR densities of the two best-fit UV LFs we derived in the previous section (for a summary see also Table 4). In particular, the best-fit M_* evolution results in a significantly higher SFRD, essentially equal to the current $z\sim9$ estimates.

[14] SFR($M_\odot$ yr^{-1}) $=1.4\times10^{-28}\,L_{1500}$ (erg s^{-1} Hz^{-1}) (Kennicutt 1998).

However, we stress again that M_* evolution should have resulted in nine detected $z\sim10$ candidates in our search fields and should therefore be considered an upper limit.

Combining our updated SFRD estimate with previous analyses from different data sets in the literature at $z>9$ (Bouwens et al. 2012a; Coe et al. 2013; Zheng et al. 2012), the new best-fit evolution of the cosmic SFRD at $z\geqslant8$ is $\log\dot{\rho}_*\propto(1+z)^{10.9\pm2.5}$, which is almost unchanged from our previous determination without the new luminous sources in GOODS-N and GOODS-S ($(1+z)^{11.4\pm3.1}$; Oesch et al. 2013b). The small change is mostly due to the fact that our new, combined SFRD measurement from all the CANDELS-Deep and HUDF09/XDF data almost exactly falls on the previously estimated trend and that the LF is so steep that the integrated flux is dominated still by the lower luminosity sources.

As we show in Figure 10, the rapid decline at $z>8$ is not completely unexpected. It seems to be a very generic prediction of a wide range of theoretical models, which reproduce the UV LF evolution across $z\sim4$–7. These models include semi-empirical estimates of the SFRD evolution (Trenti et al. 2010; Tacchella et al. 2013), semi-analytical models (Lacey et al. 2011), as well as hydrodynamic simulations (Finlator et al. 2011; Salvaterra et al. 2011; Dayal et al. 2013; Jaacks et al. 2013). Given that these models all use very different prescriptions and techniques, it is likely that this rapid decline is mostly driven by the underlying evolution of the dark matter halo mass function, which is also evolving very rapidly at $z>8$.

5. ROBUST REST-FRAME OPTICAL DETECTIONS OF $z\sim10$ GALAXIES: NEBULAR EMISSION LINES AND STELLAR MASSES

The most important result from our IRAC analysis in the previous sections is that for all three sources in GOODS-N for which the neighbor subtraction was successful, we detect

a significant signal in at least one IRAC band. The detection significances in the 4.5 μm channel are 5.8σ (GN-z10-1), 4.5σ (GN-z10-3), and 6.2σ (GN-z9-1). Furthermore, the brightest of our candidates (GN-z10-1) is detected at 6.9σ also in the 3.6 μm band of the IRAC data (see also Figure 2). For this source, we therefore have two secure IRAC flux measurements in addition to its H_{160}-detection. These measurements allow us to place significant constraints on the SED of GN-z10-1 and derive relatively robust physical properties based on SED fitting. Additionally, the lower redshift source in GOODS-S (GS-z9-1) is detected in both IRAC channels, while only a faint 4.5 μm detection is found for the other GOODS-S $z \sim 10$ candidate GS-z10-1 (see Figure 12 and Table 8 in the Appendix).

Note that the *Spitzer* 4.5 μm channel only probes up to rest-frame *B*-band at the very high redshift of our targets. The ability to identify an underlying old population of stars is therefore limited with these data. However, this is unlikely to be a significant issue for these galaxies given the young age of the universe at $z \sim 9$–10.

The most striking feature of the SED of GN-z10-1 is that it does not show any Balmer break, having an IRAC color of [3.6]–[4.5] $= -0.14 \pm 0.29$. Overall, the SED of this source is extremely flat with few features. Using a constant star-formation history and a Bruzual & Charlot (2003) model with 0.5 $Z_{\odot}$, the SED is best fit with the minimum allowed age of only 10 Myr and a stellar mass of $10^{8.7\pm0.3}$ $M_{\odot}$. Interestingly, the SED-based SFR for this source is 55 $M_{\odot}$ yr^{-1}, which is significantly higher than the observed SFR based on its uncorrected UV luminosity and standard conversion factor (this gives 14 $M_{\odot}$ yr^{-1}). The high SED-derived SFR can be attributed largely to the non-negligible dust extinction required in the SED fit of $A_{\rm UV} = 1.1$ mag, together with the young SED age.

Given that GN-z10-1 lies at $z_{\rm phot} \sim 10.2$, we can use its $H_{160} - [3.6]$ color to estimate its UV continuum slope β.[15] This results in $\beta = -2.0 \pm 0.2$, which is consistent with the UV continuum slope of its best-fit SED ($\beta = -2.09$). This may appear relatively red for such a young, high-redshift source. However, it is consistent with the UV continuum slope evolutionary trends with luminosity found by Bouwens et al. (2013b), given the high luminosity of this source.

The SED fitting of the remaining two sources with clean IRAC photometry gives stellar masses of $10^{9.2\pm0.3}$ $M_{\odot}$ for GN-z10-3 and $10^{9.4\pm0.3}$ $M_{\odot}$ for GN-z9-1 with SED ages of ~300 Myr. Similarly, the GOODS-S sources show stellar masses of $10^{8.5\pm0.4}$ $M_{\odot}$ and $10^{9.5\pm0.3}$ $M_{\odot}$. The typical stellar mass for these bright $z \sim 9$–10 sources therefore appears to be $10^{9.0}$ $M_{\odot}$. The mass estimates, together with estimates of the specific SFR (SSFR) of these sources, are listed in Table 6. The stellar mass estimate of source GN-z10-2 is highly uncertain, because of the lack of clean IRAC photometry.

As can be seen from Table 6, the SSFRs of these sources are not very well constrained, given the current photometry and redshift uncertainties. However, they all lie higher than 2 Gyr^{-1}, where a possible flattening of the SSFR had been suggested at $z \sim 3$–6 (Stark et al. 2009; González et al. 2010), contrary to expectations from models. These galaxies therefore provide some tentative evidence for a continued increase in the SSFRs at higher redshifts (see also Smit et al. 2013; Stark et al. 2013; Gonzalez et al. 2012; de Barros et al. 2012; Schaerer & de Barros 2010).

[15] Defined as $f_{\lambda} \propto \lambda^{\beta}$.

Table 6
Physical Properties of GOODS-North and South $z \sim 9$–10 Candidates

Source	$M_{\rm UV}$	log M_* ($M_{\odot}$)	log SSFR (yr^{-1})
GN-z10-1	-21.6 ± 0.1	8.7 ± 0.3	-7.0 ± 0.6
GN-z10-2	-20.7 ± 0.1	(7.9)[a]	(−7.1)[a]
GN-z10-3	-20.6 ± 0.2	9.2 ± 0.3	-8.5 ± 0.8
GN-z9-1	-20.7 ± 0.1	9.4 ± 0.3	-8.5 ± 0.9
GS-z10-1	-20.6 ± 0.2	8.5 ± 0.4	-7.5 ± 0.6
GS-z9-1	-20.8 ± 0.2	9.5 ± 0.3	-8.3 ± 0.5

Note. [a] The stellar population parameters for source GN-z10-2 are highly uncertain due to uncertainties in the IRAC neighbor subtraction.

Emission lines, which can significantly contaminate the IRAC photometry, are implicitly accounted for in our SED library as discussed in Section 3.2. At the redshifts of our candidates ($z = 9.2$–10.2), the only significant line emission in the IRAC bands would be expected from [O II] or weaker Balmer lines. While these lines are likely not as strong as [O III]$_{4959,5007}$ or Hα, which have been identified as significant contributors to IRAC fluxes of high-redshift galaxies at $z \sim 4$–8 (e.g., Schaerer & de Barros 2009; Shim et al. 2011; Labbé et al. 2013; de Barros et al. 2012; Stark et al. 2013), their combined equivalent width can still be substantial at very young ages 10–50 Myr. In the most extreme cases it could be enough to boost the IRAC photometry by up to 0.5 magnitude, depending on the exact implementation of the emission lines in the models and on poorly constrained quantities such as metallicity, ionization parameter, reddening of the nebular regions, and escape fraction.

While some uncertainties remain, these mass estimates allow us to derive a first estimate of the stellar mass density at $z \sim 10$ using the four bright $z > 9.5$ sources from GOODS-N and GOODS-S. Using the volume density for these sources which we derived above, we estimate a mass density for such galaxies of $\log_{10} \rho_{*,z\sim10}(M_{\rm UV} < -20.5) = 4.0^{+0.5}_{-0.6}$ $M_{\odot}$ Mpc^{-3}.

To compare this with recent measurements in the literature at lower redshifts, which include contributions to the stellar mass from sources with $M_{\rm UV} < -18$ (e.g., González et al. 2011; Stark et al. 2013), we have to fold in the contribution of fainter sources at $z \sim 10$. This is quite uncertain for several reasons: (1) we only have one faint candidate in our combined data set, and (2) this XDF source is too faint to be detected by IRAC, even in the existing ultra-deep IUDF10 data set over the XDF (see Labbé et al. 2013; Oesch et al. 2013b).

In order to get a first, simple estimate of the $z \sim 10$ stellar mass density to $M_{\rm UV} < -18$, we thus assume that the XDF candidate has the same $M_*/L_{\rm UV}$ relation as the average bright GOODS-N+S galaxy, which are one order of magnitude more UV-luminous. With this assumption, we derive $\log_{10} \rho_{*,z\sim10}(M_{\rm UV} < -18) = 4.7^{+0.5}_{-0.8}$ $M_{\odot}$ Mpc^{-3}.

Figure 11 shows that the estimated value of this stellar mass density lies almost an order of magnitude below the extrapolated empirical trends from recent determinations (González et al. 2011; Labbé et al. 2013; Stark et al. 2013). The large drop is mostly a consequence of the drop in the UV LD from $z \sim 8$ to $z \sim 10$, given that the stellar mass densities are all derived down to a fixed UV luminosity. The magnitude of the drop is still very uncertain given the current uncertainties on the stellar mass density measurements at $z \geqslant 7$. However, it is in overall very good agreement with theoretical model predictions as shown in the figure (Finlator et al. 2011; Dayal et al. 2013; Tacchella et al.

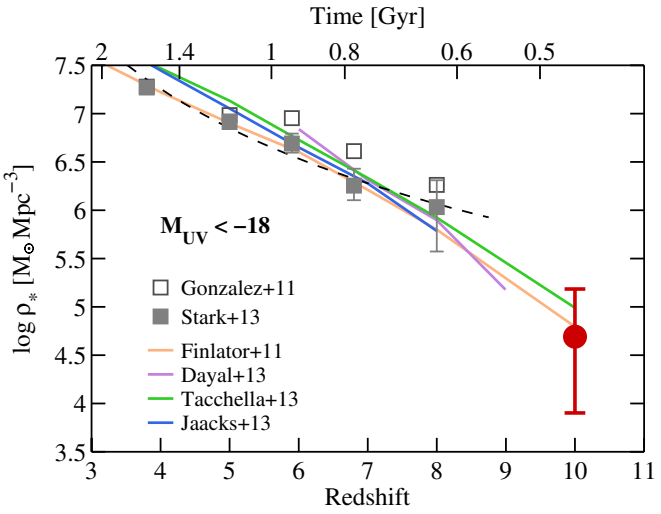

Figure 11. Redshift evolution of the cosmic stellar mass density in LBGs at $M_{\rm UV} \lesssim -18$ to $z \sim 10$. Gray squares represent the most recent determination of Stark et al. (2013), who used the measurements of González et al. (2011) and empirically corrected these for the likely biases due to strong emission lines in the IRAC photometry for $z \sim 5$–8 sources. Their empirical trends are indicated as a dashed black line. The dark red point is our first estimate of the $z \sim 10$ stellar mass density based on the three GOODS-N $z = 9.5$–10.2 galaxies, the one source in GOODS-S, and the XDF candidate. For the latter, we have assumed the same $M_*/L_{\rm UV}$ ratio as for the average of the GOODS-N+S sources, for which these measurements can be done directly given the IRAC detections. The stellar mass density is found to be more than an order of magnitude lower at $z \sim 10$ than at $z \sim 8$, consistent, as would be expected, with the drop in the cosmic SFRD down to $M_{\rm UV} < -18$. This decrease to higher redshift is also consistent with the latest simulations and models, which are overplotted as colored lines (Finlator et al. 2011; Dayal et al. 2013; Tacchella et al. 2013; Jaacks et al. 2013). The current uncertainties on the stellar mass density measurements at $z \geqslant 7$ are still significant, however. Deeper IRAC data over larger high-redshift samples would be ideal for reducing the uncertainties in the near future.

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

2013). Larger $z > 8$ galaxy samples with direct IRAC detections are required to derive more robust estimates of the stellar mass densities in the future in order to be able to discriminate between different models.

Our analysis demonstrates the power of *Spitzer*/IRAC to probe galaxy masses as early as 500 Myr after the big bang and indicates the potential for further progress in the near future with a dedicated IRAC survey.

6. DISCUSSION AND CONCLUSIONS

We discuss the discovery of four very bright $z \sim 9$–10 galaxy candidates identified in the complete CANDELS GOODS-N data set. These sources have magnitudes in the range $H_{160} = 26$–27 mag, comparable to the two highly magnified $z \sim 10$ galaxy candidates found in the CLASH cluster data set (Zheng et al. 2012; Coe et al. 2013). However, these four GOODS-N candidates do not show signs of significant magnification ($\lesssim$0.1 mag), and they are thus by far the most luminous $z \sim 10$ galaxy candidates detected with *HST* to date.

Furthermore, these sources result in the first $>5\sigma$ *Spitzer*/IRAC detections of $z > 9$ galaxy candidates. The brightest candidate is securely detected at 6.9σ in the 3.6 μm band and 5.8σ in the 4.5 μm band of the 50 hr *Spitzer*/IRAC data from the combined GOODS, SEDS, and S-CANDELS data. For two of the three remaining candidates, those for which bright neighbor subtraction was successful in the IRAC bands, we additionally detect significant signal in the 4.5 μm band at 4.5σ and 6.2σ. *Spitzer*/IRAC has thus been able to clearly detect rest-frame optical light out to ~500 Myr after the big bang.

Motivated by the bright galaxies in GOODS-N, we re-analyzed the GOODS-S data set for bright sources using similar selection criteria as in GOODS-N ($J_{125} - H_{160} > 0.5$) and updated SExtractor detection parameters. This analysis resulted in the detection of two similarly luminous sources in GOODS-S, both of which were also detected in *Spitzer*/IRAC, though one had marginal 2σ detections in both IRAC bands. These sources are discussed in the Appendix.

The *Spitzer*/IRAC flux measurements, together with the extensive deep *HST* data, allow us to constrain any lower-redshift contamination. Extensive tests were done (see Section 3) to evaluate the likelihood that these sources could be other than high-redshift $z \sim 9$–10 galaxies. The results suggest that the most plausible outcome is that these galaxy candidates are really at $z \sim 9$–10. Yet we cannot rule out that they constitute very unusual objects at lower redshift. Fortunately these objects are so bright that the opportunity exists for these sources to be measured by current near-IR spectrographs on 8–10m class telescopes to establish their redshifts.

The *Spitzer*/IRAC flux measurements further enable the first derivation of stellar mass estimates for $z \sim 10$ galaxy candidates. The four sources with clean IRAC photometry all show masses in the range $\log_{10} M = 8.7$–$9.5\ M_\odot$ (see Table 6). The brightest source, which is securely detected in both IRAC bands, shows effectively no Balmer break with its [3.6]–[4.5] color and overall reveals a featureless SED, indicative of very recent onset of star-formation. Its UV continuum slope is measured to be $\beta = -2.0 \pm 0.2$, which requires a non-negligible amount of dust to be present already at a cosmic age of 500 Myr. While somewhat redder than the UV slopes reported for faint galaxies at $z \sim 7$ (e.g., Bouwens et al. 2012c, 2013b; Finkelstein et al. 2012), this observed β fits well with previously reported β-luminosity trends (Bouwens et al. 2013b) and is not surprising given the galaxy's luminosity.

The $z > 8$ Luminosity Function: Luminosity or Density Evolution or More Complex? The three highest-redshift GOODS-N sources and one of the GOODS-S sources satisfy the previously adopted $J_{125} - H_{160} > 1.2$ criterion for $z \sim 10$ galaxy selections in the GOODS-S and HUDF09/12 fields. We use these four new sources to update our previous estimates of the $z \sim 10$ UV LF and of the cosmic SFRD evolution across $z \sim 8$–10, now using the complete CANDELS-Deep data set of GOODS-N and South, together with the three ultra-deep HUDF09/12 fields. The UV LF determination at $z > 9.5$ thus consists of five sources: the three bright GOODS-N candidates, one bright GOODS-S candidate, and the very faint candidate XDFj-38126243 found in the XDF.

Based on previous estimates of the UV LF at $z > 8$, the detection of four such luminous $z \sim 10$ galaxy candidates in the GOODS-N+S fields is unexpected.

The full $z \sim 10$ sample studied here populates only the bright and the faint tails of the expected magnitude distribution of an assumed Schechter function LF, with a dearth of intermediate luminosity galaxy candidates (see, e.g., Figure 7).

The updated $z \sim 10$ UV LF estimates are presented in Section 4.4. As expected, the detection of the four new bright GOODS-N+S sources significantly tilts the UV LF parameters to higher number densities at bright luminosities compared to previous work (Oesch et al. 2013a). The UV LF parameters are derived relative to a baseline $z \sim 8$ LF and are summarized

in Table 4. In particular, we test two scenarios for the UV LF evolution to $z \sim 10$: ϕ_*-only or M_*-only. While the ϕ_*-only evolutionary scenario results in a somewhat better fit to the current $z \sim 10$ search results, the high number densities at the bright end are in clear tension with the upper limits at intermediate magnitudes.

Even with these new Schechter function parameters only one bright source is expected at $H_{160} < 27$ mag in the full search area. This highlights the need for spectroscopy to determine the true nature of these bright candidates and the need to search larger areas for $z \sim 9$–10 sources to determine the distribution and abundance of star-forming galaxies in the very early universe.

If ϕ_* evolution is correct and if it is not offset by a steepening in the faint-end slope of the UV LF (i.e., $\alpha < -2$), this would result in a significant drop in the total ionizing flux density of galaxies and may thus have important consequences for cosmic reionization.

The current lack of intermediate luminosity $z \sim 9$–10 galaxy candidates may also point to a non-Schecter-like form of the LF at $z > 6$ (e.g., Bouwens et al. 2011b; Bowler et al. 2012). This could potentially be caused, e.g., by bursty and highly biased star-formation in the very early universe (see, e.g., Jaacks et al. 2012b; Wyithe et al. 2013). Larger samples of $z \sim 9$–10 galaxy candidates are needed, however, e.g., from the HST Frontier Fields program to address this in the near future.

The Star Formation Rate Density since $z \sim 10$. We re-evaluated the cosmic SFRD given these new $z \sim 9$–10 sources. We included these new sources together with all prior $z > 8$ candidates from the literature to obtain an updated estimate of the cosmic SFRD at $z \sim 10$ and found it to be a factor 1.05 ± 0.38 dex below the $z \sim 8$ value. The addition of the four luminous sources found here did not significantly change the rapid evolution in the SFRD that we reported in Oesch et al. (2013a). This is due to the steepness of the LF and the fact that the luminosity density at $z \sim 10$ (as at $z \sim 6$–8) is dominated by the faint, lower luminosity sources. The cosmic SFRD evolves rapidly in just 170 Myr from $z \sim 10$ to $z \sim 8$ (down to the current limit of $>0.7\,M_\odot$ yr^{-1}) as shown in Figure 9.

The Mass Density at $z \sim 10$. Based on the individual stellar mass measurements of the GOODS-N $z \sim 10$ candidates, we attempt a first estimate of the cosmic stellar mass density at $z \sim 10$. After correcting from the bright, IRAC detected GOODS sources to LBGs down to $M_{\rm UV} < -18$, we find $\log_{10}\rho_* = 4.7^{+0.5}_{-0.8}$ $M_\odot$ Mpc^{-3}. This is more than an order of magnitude lower than the current measurements at $z \sim 8$. However, this decline is in good agreement with model predictions (see Figure 11). Larger galaxy samples with direct IRAC detections will be necessary, to verify this first stellar mass density estimate at $z > 8$ in the future. As our analysis has shown, IRAC is capable of detecting faint galaxies at very high redshifts. With a larger survey it would thus be possible to significantly reduce the uncertainties in current stellar mass density measurements at $z \geqslant 7$.

Future Opportunities. The discovery of such luminous candidate galaxies at $z \sim 10$—together with the existence of similarly luminous galaxies in $z \sim 6$–8 probes (Bunker et al. 2003; Bouwens et al. 2010a, 2014; Oesch et al. 2012b; Trenti et al. 2011; Yan et al. 2012; Bowler et al. 2012)—gives us hope that comparably luminous sources, i.e., −22 to −21 mag, may be found in searches at even higher redshifts, i.e., $z > 10$, albeit with substantially reduced volume densities.

If the current $z \sim 9$–10 candidates are indeed at such high redshift, as we fully expect given the low likelihood of lower redshift contamination, it would argue for the use of deep but moderately wide-area searches to maximize the number of $z > 8$ galaxies that are found. There are still sufficiently large uncertainties in the evolution of the UV LF at $z > 6$ at present, however, that the optimal survey strategy is far from clear. Better constraints on the UV LFs at $z \sim 9$ and $z \sim 10$ with present and upcoming *HST* data will therefore be key to inform the optimal survey strategies, e.g., for *JWST*.

The unusual brightness of these $z \sim 9$–10 candidates makes them obvious targets for spectroscopy, both from the ground and from space. The brightest source is within 0.1 mag of the highly magnified CLASH $z > 10$ source of Coe et al. (2013). Deep spectroscopy could reach $z > 9$ Lyα emission lines with an equivalent width as small as 10 Å and will rule out contamination by lower level emission line sources with significant dust extinction. Spectroscopic redshift measurements could show if these surprisingly luminous candidates are really at high redshift as all the photometric tests suggest. They could significantly reduce the uncertainties on the physical parameter estimates of these sources and provide the basis for a more detailed modeling of their SEDs and star-formation histories.

The bright $z \sim 9$–10 candidates highlight the importance of probing a large volume and several independent fields for accurate cosmic average measurements at high redshift. Although somewhat reduced by lensing magnification, the upcoming HST Frontier Fields program will provide additional search volume for faint $z \sim 10$ galaxy candidates and, together with its *Spitzer*/IRAC component, will provide new opportunities for exploring the cosmic frontier before the advent of *JWST*.

We thank the anonymous referee for very helpful feedback that improved the paper. We also thank Jason Jaacks, Sandro Tacchella, and Kristian Finlator for providing us with the predictions of their models. Support for this work was provided by NASA through Hubble Fellowship grant HF-51278.01 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS 5-26555. Additionally, this work was supported by NASA grant NAG5-7697, NASA grant HST-GO-11563, NASA grant HST-GO-12177, NASA grant JPL-1416188/1438944, ERC grant HIGHZ 227749, and NWO vrij competitie grant 600.065.140.11N211. This work was further supported in part by the National Science Foundation under grant PHY-1066293 and the Aspen Center for Physics. Some of the data presented in this paper were obtained from the Mikulski Archive for Space Telescopes (MAST). This research used the facilities of the Canadian Astronomy Data Centre operated by the National Research Council of Canada with the support of the Canadian Space Agency.

Facilities: HST (ACS/WFC3), Spitzer (IRAC)

APPENDIX A

TWO BRIGHT $z > 9$ CANDIDATE GALAXIES IN GOODS-S

Motivated by the discovery of the four bright $z > 9$ galaxy candidates in the GOODS-N field, we re-analyzed the GOODS-S data set. In particular, we ran new SExtractor catalogs with a higher deblending efficiency to split neighboring sources, and we searched for additional sources with a bluer

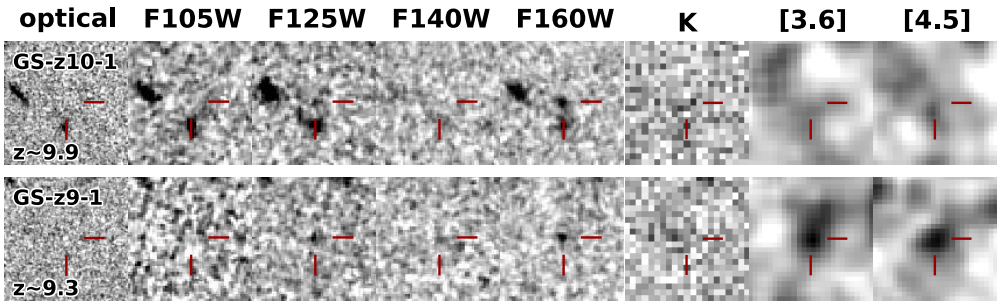

Figure 12. $6'' \times 6''$ negative images of the two new $z \geqslant 9$ galaxy candidates identified in our reanalysis of the CANDELS GOODS-S data. From left to right, the images show a stack of all optical bands, Y_{105}, J_{125}, H_{160}, HAWKI K, and neighbor-subtracted IRAC 3.6 μm and 4.5 μm images. The K-band image is a very deep stack (26.5 mag, 5σ) of ESO/VLT HAWK-I data from the HUGS survey (PI: Fontana). Both sources are only weakly detected in these data.
(A color version of this figure is available in the online journal.)

Table 7
Coordinates and Basic Photometry of Two New $z > 9$ LBG Candidates in the GOODS-S Field

Name	ID	R.A.	Decl.	H_{160}	$J_{125} - H_{160}$	$H_{160} - [4.5]$	$z_{\rm phot}$
GS-z10-1	GSDJ-2269746283	03:32:26.97	−27:46:28.3	26.88 ± 0.15	1.7 ± 0.6	-0.4 ± 0.6	9.9 ± 0.5
GS-z9-1[a]	GSDJ-2320550417	03:32:32.05	−27:50:41.7	26.61 ± 0.18	1.1 ± 0.5	0.5 ± 0.3	9.3 ± 0.5

Note. [a] The source GS-z9-1 does not satisfy the criterion $J_{125} - H_{160} > 1.2$ and is not included in the UV LF analysis.

color cut of $J_{125} - H_{160} > 0.5$, as we did in GOODS-N, rather than the more conservative cut of $J_{125} - H_{160} > 1.2$ as adopted in our previous work (e.g., Oesch et al. 2013a).

These new catalogs revealed two possible, bright $z > 9$ galaxy candidates in the CANDELS GOODS-S data set, GS-z9-1 and GS-z10-1. They have magnitudes of $H_{160} = 26.6 \pm 0.2$ and $H_{160} = 26.9 \pm 0.2$, respectively. The latter candidate also shows a color of $J_{125} - H_{160} > 1.2$ (namely 1.7 ± 0.6), while the first is only slightly too blue to satisfy this criterion ($J_{125} - H_{160} = 1.1 \pm 0.5$).

Given its red color, GS-z10-1 could already have been in the previous catalog of Oesch et al. (2013a) who analyzed the same CANDELS GOODS-South data set. The reason this source was not previously selected is due to a very faint neighbor that was included in the Kron aperture in the earlier SExtractor catalog. This caused the candidate to be rejected due to apparent optical flux in the aperture. With careful visual inspection we assessed that the optical flux in the previous aperture was due to a faint neighboring galaxy and is not likely associated with the high-z candidate. With the new deblending parameters for our SExtractor run, this source is now confirmed to be a legitimate $z > 9$ galaxy candidate. Its photometric redshift is found to be $z_{\rm phot} = 9.9 \pm 0.5$. We thus include this candidate in the full analysis of the main body of this paper. We have verified that the GOODS-N data returns the same candidates when using these updated deblending parameters.

The inclusion of this $z \sim 10$ candidate does not significantly change the results. For instance, including this candidate only causes a change of 0.1 dex in $\phi*$ when assuming density evolution or a change of only 0.1 in $M*$ for luminosity evolution. The total cosmic SFRD changes by only 0.02 dex, because this is dominated by the luminosity from lower luminosity sources as indicated by the faintest candidate in the XDF (and by the steep slopes found at slightly later times at $z \sim 7$–8).

The other source, GS-z9-1, was already in the previous SExtractor catalogs. However, it was not included in the analysis due to its bluer color of $J_{125} - H_{160} < 1.2$. For completeness, we present this source here as well, particularly since it is so close to our $z \sim 10$ color cutoff. Interestingly, it also shows significant IRAC detections in both 3.6 and 4.5 μm bands with fluxes consistent with a significant Balmer break at $z \sim 9$, giving added weight to our identification of this source as a probable $z \sim 9$ candidate. From SED fitting we find a photometric redshift of $z_{\rm phot} = 9.3 \pm 0.5$ for this source.

Images of both new GOODS-S candidates are shown in Figure 12, and their SED fits and photometric redshift likelihood functions are shown in Figure 13. Table 7 lists the basic information of these sources, and Table 8 list all their flux measurements.

Table 8
Flux Densities of Two New $z > 9$ LBG Candidates in the GOODS-S Field

Filter	GS-z10-1	GS-z9-1
B_{435}	-1 ± 9	7 ± 10
V_{606}	1 ± 6	0 ± 8
i_{775}	-6 ± 9	-5 ± 12
I_{814}	5 ± 6	-3 ± 9
z_{850}	-4 ± 9	-5 ± 16
Y_{105}	0 ± 6	-14 ± 9
J_{125}	13 ± 7	29 ± 11
JH_{140}	12 ± 23	55 ± 33
H_{160}	66 ± 9	85 ± 14
K−HAWKI	33 ± 19	54 ± 18
IRAC 3.6 μm	32 ± 17	58 ± 24
IRAC 4.5 μm	44 ± 22	131 ± 23

Note. Measurements are given in nJy with 1σ uncertainties.

APPENDIX B

IRAC NEIGHBOR SUBTRACTION

The point-spread function of *Spitzer*/IRAC is ∼10× broader than for WFC3/IR. A crucial aspect of using the *Spitzer*/IRAC data to constrain the rest-frame optical fluxes of faint galaxies at high redshift is therefore to reliably subtract neighboring sources to deal with source confusion. Several teams have developed techniques to perform efficient neighbor subtraction based on modeling the IRAC fluxes from the high-resolution

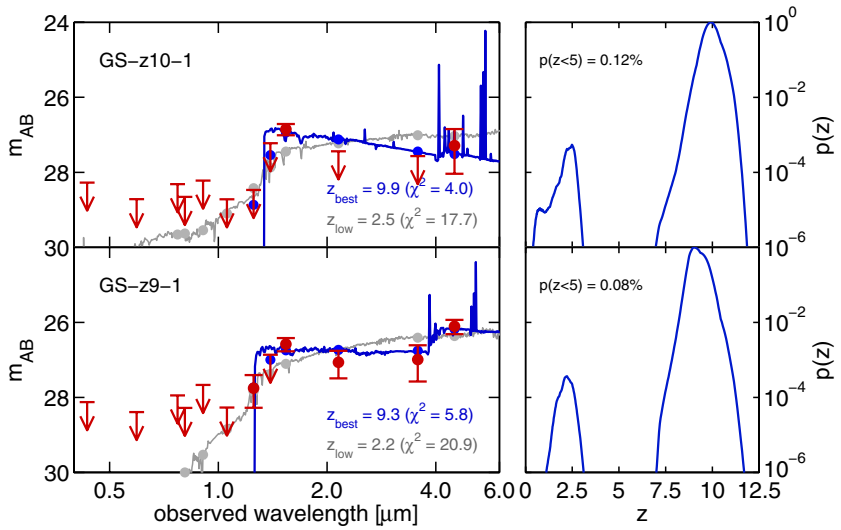

Figure 13. Spectral energy distribution fits (left) and redshift likelihood functions (right) for the two bright $z \sim 9$–10 sources in GOODS-S. In the left panel, photometry and 2σ upper limits are shown in dark red. Best-fit SEDs are shown as blue solid lines, in addition to the best low redshift solutions in gray. The legend lists their respective redshift and χ^2 values. The redshift likelihood distributions from SED fitting in the right panels are shown on logarithmic axes. For both sources, the likelihood of a low redshift solution is less than 1% as indicated by the labels.

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

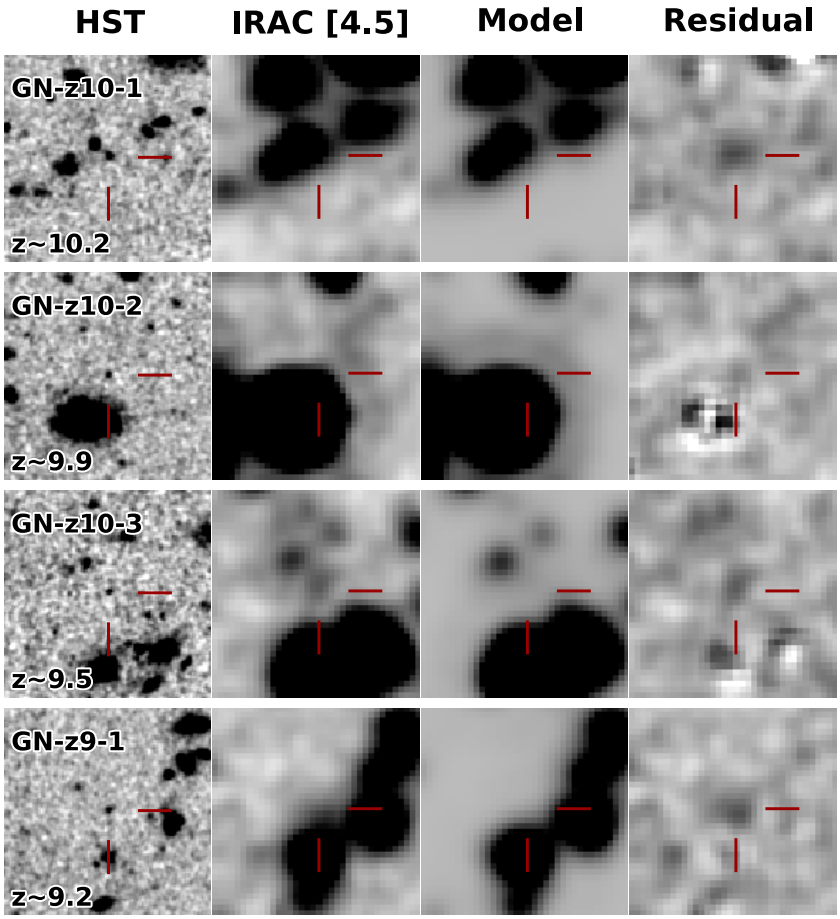

Figure 14. Demonstration of the IRAC neighbor subtraction routine. The left column shows the $13'' \times 13''$ stamps of the original H_{160} image of the four GOODS-N high-redshift candidates. The second column is the original IRAC image at 4.5 μm. We use a kernel to convolve the H_{160} image to the PSF of this IRAC image and fit the combined light profile of each source in the H_{160}-band image. The third column shows the result of this fit for all the neighbors of the target source. This model is subsequently subtracted from the original IRAC image which results in a residual image (right column), on which we perform aperture photometry of the source of interest.

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

HST imaging and convolving these to the IRAC PSF (see, e.g., Labbé et al. 2006; Grazian et al. 2006; Laidler et al. 2007).

A key uncertainty in this modeling is the derivation of a reliable convolution kernel from the HST to the IRAC images. In particular, given the significant asymmetry of the IRAC PSF, this kernel can be highly dependent on the source position. Our modeling technique (see I. Labbé et al., in preparation) therefore first maps out the kernel as a function of position based on stars in the field. This additional step led to a very significant improvement of the neighbor subtraction compared to other procedures used in the literature. Our tests show that we can map the kernel to high precision and perform excellent subtractions (see, e.g., Labbé et al. 2010a, 2010b, 2013; González et al. 2010, 2012).

The efficiency of our subtraction routine on the IRAC dataset used here is illustrated in Figure 14, where we show the process from the original H_{160} image to the residual image on which we perform aperture photometry of the sources, as described in Section 5. While, e.g., the source GN-z10-1 is clearly confused with a foreground source, the neighbor subtraction works extremely well. The uncertainties we compute on the IRAC fluxes include the additional contribution from the neighboring flux, and we have verified that our measurements are robust against small uncertainties in the image registration.

REFERENCES

Anders, P., & Fritze-v. Alvensleben, U. 2003, A&A, 401, 1063
Ashby, M. L. N., Willner, S. P., Fazio, G. G., et al. 2013, ApJ, 769, 80
Atek, H., Siana, B., Scarlata, C., et al. 2011, ApJ, 743, 121
Barger, A. J., Cowie, L. L., & Wang, W.-H. 2008, ApJ, 689, 687
Beckwith, S. V. W., Stiavelli, M., Koekemoer, A. M., et al. 2006, AJ, 132, 1729
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Bouwens, R., Bradley, L., Zitrin, A., et al. 2012a, arXiv:1211.2230
Bouwens, R. J., Illingworth, G. D., Franx, M., & Ford, H. 2007, ApJ, 670, 928
Bouwens, R. J., Illingworth, G. D., Franx, M., & Ford, H. 2008, ApJ, 686, 230
Bouwens, R. J., Illingworth, G. D., González, V., et al. 2010a, ApJ, 725, 1587
Bouwens, R. J., Illingworth, G. D., Labbe, I., et al. 2011a, Natur, 469, 504
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2010b, ApJL, 709, L133
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2011b, ApJ, 737, 90
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2012b, ApJL, 752, L5
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2012c, ApJ, 754, 83
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2013b, arXiv:1306.2950
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2014, arXiv:1403.4295
Bouwens, R. J., Oesch, P. A., Illingworth, G. D., et al. 2013a, ApJL, 765, L16
Bowler, R. A. A., Dunlop, J. S., McLure, R. J., et al. 2012, MNRAS, 426, 2772
Bradley, L. D., Trenti, M., Oesch, P. A., et al. 2012, ApJ, 760, 108
Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503
Brammer, G. B., van Dokkum, P. G., Franx, M., et al. 2012, ApJS, 200, 13
Brammer, G. B., van Dokkum, P. G., Illingworth, G. D., et al. 2013, ApJL, 765, L2
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
Burgasser, A. J., McElwain, M. W., Kirkpatrick, J. D., et al. 2004, AJ, 127, 2856
Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682
Capak, P., Faisst, A., Vieira, J. D., et al. 2013, ApJL, 773, L14
Caputi, K. I., Dunlop, J. S., McLure, R. J., et al. 2012, ApJL, 750, L20
Coe, D., Zitrin, A., Carrasco, M., et al. 2013, ApJ, 762, 32
Dayal, P., Dunlop, J. S., Maio, U., & Ciardi, B. 2013, MNRAS, 434, 1486
de Barros, S., Schaerer, D., & Stark, D. P. 2012, A&A, in press (arXiv:1207.3663)
Dickinson, M., Giavalisco, M., & GOODS Team., 2003, in The Mass of Galaxies at Low and High Redshift, ed. R. Bender & A. Renzini (Dordrecht: Springer), 324
Dunlop, J. S., Rogers, A. B., McLure, R. J., et al. 2013, MNRAS, 432, 3520
Ellis, R. S., McLure, R. J., Dunlop, J. S., et al. 2013, ApJL, 763, L7
Fazio, G. G., Hora, J. L., Allen, L. E., et al. 2004, ApJS, 154, 10
Feldmann, R., Carollo, C. M., Porciani, C., et al. 2006, MNRAS, 372, 565
Finkelstein, S. L., Papovich, C., Salmon, B., et al. 2012, ApJ, 756, 164
Finlator, K., Oppenheimer, B. D., & Davé, R. 2011, MNRAS, 410, 1703
Giavalisco, M., Dickinson, M., Ferguson, H. C., et al. 2004, ApJL, 600, L103
González, V., Bouwens, R., Ilingworth, G., et al. 2012, arXiv:1208.4362
González, V., Bouwens, R. J., Labbé, I., et al. 2012, ApJ, 755, 148
González, V., Labbé, I., Bouwens, R. J., et al. 2010, ApJ, 713, 115
González, V., Labbé, I., Bouwens, R. J., et al. 2011, ApJL, 735, L34
Grazian, A., Fontana, A., de Santis, C., et al. 2006, A&A, 449, 951
Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, ApJS, 197, 35
Guo, Y., Ferguson, H. C., Giavalisco, M., et al. 2013, ApJS, 207, 24
Hayes, M., Laporte, N., Pelló, R., Schaerer, D., & Le Borgne, J.-F. 2012, MNRAS, 425, L19
Huang, J.-S., Zheng, X. Z., Rigopoulou, D., et al. 2011, ApJL, 742, L13
Illingworth, G. D., et al. 2013, ApJS, 209, 6
Jaacks, J., Choi, J.-H., Nagamine, K., Thompson, R., & Varghese, S. 2012a, MNRAS, 420, 1606
Jaacks, J., Nagamine, K., & Choi, J. H. 2012b, MNRAS, 427, 403
Jaacks, J., Thompson, R., & Nagamine, K. 2013, ApJ, 766, 94
Kajisawa, M., Konishi, M., Suzuki, R., et al. 2006, PASJ, 58, 951
Kennicutt, R. C., Jr. 1998, ARA&A, 36, 189
Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, ApJS, 197, 36
Labbé, I., Bouwens, R., Illingworth, G. D., & Franx, M. 2006, ApJL, 649, L67
Labbé, I., González, V., Bouwens, R. J., et al. 2010a, ApJL, 708, L26
Labbé, I., González, V., Bouwens, R. J., et al. 2010b, ApJL, 716, L103
Labbé, I., Oesch, P. A., Bouwens, R. J., et al. 2013, ApJL, 777, L19
Lacey, C. G., Baugh, C. M., Frenk, C. S., & Benson, A. J. 2011, MNRAS, 412, 1828
Laidler, V. G., Papovich, C., Grogin, N. A., et al. 2007, PASP, 119, 1325
Lorenzoni, S., Bunker, A. J., Wilkins, S. M., et al. 2013, MNRAS, 429, 150
Madau, P., Pozzetti, L., & Dickinson, M. 1998, ApJ, 498, 106
McLure, R. J., Dunlop, J. S., Bowler, R. A. A., et al. 2013, MNRAS, 432, 2696
Oesch, P. A., Bouwens, R. J., Carollo, C. M., et al. 2010a, ApJL, 709, L21
Oesch, P. A., Bouwens, R. J., Illingworth, G. D., et al. 2012a, ApJ, 745, 110
Oesch, P. A., Bouwens, R. J., Illingworth, G. D., et al. 2012b, ApJ, 759, 135
Oesch, P. A., Bouwens, R. J., Illingworth, G. D., et al. 2013a, ApJ, 773, 75
Oesch, P. A., Carollo, C. M., Feldmann, R., et al. 2010b, ApJL, 714, L47
Oesch, P. A., Labbé, I., Bouwens, R. J., et al. 2013b, ApJ, 772, 136
Oke, J. B., & Gunn, J. E. 1983, ApJ, 266, 713
Ono, Y., Ouchi, M., Curtis-Lake, E., et al. 2013, ApJ, 777, 155
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2013, arXiv:1303.5076
Postman, M., Coe, D., Benítez, N., et al. 2012, ApJS, 199, 25
Robertson, B. E. 2010, ApJ, 713, 1266
Salvaterra, R., Ferrara, A., & Dayal, P. 2011, MNRAS, 414, 847
Schaerer, D., & de Barros, S. 2009, A&A, 502, 423
Schaerer, D., & de Barros, S. 2010, A&A, 515, A73
Schechter, P. 1976, ApJ, 203, 297
Schenker, M. A., Robertson, B. E., Ellis, R. S., et al. 2013, ApJ, 768, 196
Schneider, P., Kochanek, C. S., & Wambsganss, J. 2006, in Saas-Fee Advanced Course 33: Gravitational Lensing: Strong, Weak and Micro, ed. G. Meylan, P. Jetzer, & P. North (Berlin: Springer)
Shim, H., Chary, R.-R., Dickinson, M., et al. 2011, ApJ, 738, 69
Smit, R., et al. 2013, ApJ, submitted (arXiv:1307.5847)
Stark, D. P., Ellis, R. S., Bunker, A., et al. 2009, ApJ, 697, 1493
Stark, D. P., Schenker, M. A., Ellis, R., et al. 2013, ApJ, 763, 129
Tacchella, S., Trenti, M., & Carollo, C. M. 2013, ApJL, 768, L37
Trenti, M., Bradley, L. D., Stiavelli, M., et al. 2011, ApJL, 727, L39
Trenti, M., & Stiavelli, M. 2008, ApJ, 676, 767
Trenti, M., Stiavelli, M., Bouwens, R. J., et al. 2010, ApJL, 714, L202
Ulrich, M.-H., Maraschi, L., & Urry, C. M. 1997, ARA&A, 35, 445
van der Wel, A., Straughn, A. N., Rix, H.-W., et al. 2011, ApJ, 742, 111
Wilkins, S. M., Bunker, A. J., Stanway, E., Lorenzoni, S., & Caruana, J. 2011, MNRAS, 417, 717
Wyithe, J. S. B., Yan, H., Windhorst, R. A., & Mao, S. 2011, Natur, 469, 181
Wyithe, S., Loeb, A., & Oesch, P. 2013, MNRAS, in press (arXiv:1308.2030)
Yan, H., Finkelstein, S. L., Huang, K.-H., et al. 2012, ApJ, 761, 177
Yan, H., Yan, L., Zamojski, M. A., et al. 2011, ApJL, 728, L22
Zheng, W., Postman, M., Zitrin, A., et al. 2012, Natur, 489, 406