THE REST-FRAME UV-TO-OPTICAL COLORS AND SPECTRAL ENERGY DISTRIBUTIONS OF $z \sim 4$–7 GALAXIES

VALENTINO GONZÁLEZ1, RYCHARD J. BOUWENS2, IVO LABBÉ2, GARTH ILLINGWORTH1, PASCAL OESCH1,3, MARIN FRANX2, AND DAN MAGEE1

1 Astronomy Department, University of California, Santa Cruz, CA 95064, USA
2 Leiden Observatory, Leiden University, NL-2300 RA Leiden, The Netherlands

Received 2011 October 28; accepted 2012 June 18; published 2012 August 6

ABSTRACT

We use the ultra-deep HUDF09 and the deep Early Release Science data from the HST WFC3/IR camera, along with the wide-area Spitzer/IRAC data from GOODS-S to derive spectral energy distributions (SEDs) of star-forming galaxies from the rest-frame UV to the optical over a wide luminosity range ($M_{1500} \sim -21$ to $M_{1500} \sim -18$) from $z \sim 7$ to $z \sim 4$. The sample contains $\sim 400$ $z \sim 4$, $\sim 120$ $z \sim 5$, $\sim 60$ $z \sim 6$, and $36$ $z \sim 7$ galaxies. Median stacking enables the first comprehensive SED study of very faint high-$z$ galaxies at multiple redshifts (e.g., $[3.6] = 27.4 \pm 0.1$ AB mag for the $M_{1500} \sim -18$ sources at $z \sim 4$). At $z \sim 4$, we study the stacked SEDs over a range of $5$ mag reaching down to $\sim 0.06L_\ast_{z=4}$. We use all available SEDs and template fits to derive rest-frame UV-to-optical colors ($U-V$) at all redshifts and luminosities. We find that this color does not vary significantly with redshift at a fixed luminosity. The UV-to-optical color does show a weak trend with luminosity, becoming redder at higher luminosities. This is most likely due to dust. At $z \gtrsim 5$, we find blue $[3.6]-[4.5]$ colors $\sim -0.3$ mag that are most likely due to rest-frame optical emission lines contributing to the flux in the IRAC filter bandpasses. Such contributions would lower both ages and masses by $\sim 2\times$. The scatter in our derived SEDs remains large, but the results are most consistent with a lack of any evolution in the SEDs with redshift at a given luminosity. The uniformity of the SEDs suggests a self-similar mode of evolution over a timespan from $0.7$ Gyr to $1.5$ Gyr after the big bang that encompasses very substantial growth in the stellar mass density in the universe (from $\sim 4 \times 10^6$ to $\sim 2 \times 10^7 M_\odot$ Mpc$^{-3}$).

Key words: galaxies: evolution – galaxies: high-redshift

Online-only material: color figures

1. INTRODUCTION

As a result of the installation of the WFC3/IR camera on the Hubble Space Telescope (HST), it is now possible to obtain extremely deep, high resolution observations in the near-infrared. These observations match the exquisite optical data provided by the Advanced Camera for Surveys (ACS) over deep and ultra-deep fields like the GOODS and the HUDF (Giavalisco et al. 2004; Bouwens et al. 2007). The upgrade has enabled several teams to extend their searches for Lyman break galaxies (LBGs) to very high redshifts (e.g., Bouwens et al. 2010; Oesch et al. 2010, 2012; McClure et al. 2010; Yan et al. 2010; Bunker et al. 2010; Finkelstein et al. 2010). In combination, these two cameras have allowed for high signal-to-noise (S/N) measurements of the rest-frame UV light of high-redshift galaxies over a wide wavelength baseline.

As a result of these new observations, the UV luminosity functions at $z \gtrsim 4$ are known to high accuracy. Another direct result from these observations is the determination of the UV slope of the spectral energy distributions (SEDs), generally characterized by the parameter $\beta$ ($f_\lambda \propto \lambda^\beta$, e.g., Meurer et al. 1999; Bouwens et al. 2009). This slope has been shown to become steeper (bluer) both with decreasing luminosity and with increasing redshift (Bouwens et al. 2009, 2011a; Wilkins et al. 2011; but see Dunlop et al. 2012; Finkelstein et al. 2011). For an intermediate aged (100 Myr) stellar population, this slope generally depends most strongly on the dust reddening so that this has been interpreted as an increase in the dust content of galaxies with cosmic time and as galaxies increase in luminosity (Bouwens et al. 2011a; Finkelstein et al. 2011).

Since the rest-frame UV light primarily probes young short-lived stellar populations (mostly O and B stars), it is interesting to extend these observations to longer wavelengths and study the relative importance of more evolved stellar populations. This is very challenging though and is now only possible thanks to the incredible sensitivity of the IRAC camera on the Spitzer Space Telescope. Some of the first observations of rest-frame optical light from $z \gtrsim 4$ galaxies showed quite sizable UV-to-optical breaks (Eyles et al. 2005; Yan et al. 2005). Under the usual assumption of smooth, exponentially declining star formation histories (SFHs), these colors were interpreted as indicative of an evolved stellar population (ages $\gtrsim 100$ Myr) in combination with the younger populations probed by the UV light. Follow-up work on $z \sim 4$–7 galaxies continued to reveal systems with moderate size UV-to-optical breaks (e.g., Eyles et al. 2007; Yan et al. 2006; Labbé et al. 2006; Stark et al. 2009; González et al. 2010), even possibly to $z \sim 8$ (Labbé et al. 2010a). However interesting, these studies are only possible for the brightest $z \gtrsim 4$ galaxies due to the extreme depths required.

Some alternative explanations have been proposed for the red UV-to-optical colors observed in $z > 6$ galaxies. For example, for young star-forming galaxies, one might expect strong line emission. At specific redshifts, these lines could contaminate the 3.6 and 4.5 $\mu$m IRAC fluxes, resulting in red UV-to-optical colors similar to those produced by the Balmer break (Schaerer & de Barros 2009, 2010). For emission lines to be strong enough to dominate the rest-frame optical fluxes, extremely young ages ($\sim 6$ Myr) are required. Such models have ages shorter than the dynamical times of these high-$z$ sources and are also somewhat uncertain. It would be surprising if the majority of these galaxies, which are selected over a redshift range that spans $\gtrsim 200$ Myr,
had such young ages. Even if rest-frame optical emission nebular emission lines do not dominate the measured IRAC fluxes, it is expected that they will contribute some of the flux measured in the mid-IR, but the exact contribution is difficult to establish at these high redshifts.

It has been routine, then, to compare the colors of these high-redshift galaxies with synthetic stellar population (SSP) models of relatively evolved galaxies with smooth SFHs to derive physical properties from the observations. By studying LBG samples at different redshifts in the aforementioned way, one intriguing result has emerged. At a given intrinsic luminosity, the specific star formation rate (SFR) of galaxies seems to remain remarkably constant with redshift (Stark et al. 2009; González et al. 2010, 2011; McLure et al. 2011) though accounting for the latest dust corrections by Bouwens et al. (2011a) would modify this somewhat. This result has called the attention of theorists and has proved very hard to reproduce in simulations (e.g., Weinmann et al. 2011; Khochfar & Silk 2011). The assumptions used to derive these results vary somewhat from group to group. In this work, we present the basic observations that lead to such a result, namely the observed SEDs of galaxies at $z \sim 4$–7.

In the present work, we study the observational properties of the typical high-redshift galaxy. We do this by splitting our large sample in both redshift and UV-luminosity bins and constructing mean SEDs. The large number of galaxies allows us to study the typical colors at high S/N but also extend our measurements to faint limits that are currently inaccessible on an individual basis with the current data. Although the scatter about the typical properties is also highly interesting, we defer their study to a future work which relies on even deeper Spitzer observations. In Section 2, we present the data and in Section 3 the sample selection and photometry. Section 4 describes the stacking procedure and we discuss the stacked SEDs and colors in Section 5. A summary is presented in Section 6.

Throughout, we use a ($H_0$, $\Omega_M$, $\Omega_{\Lambda}$) = (70 km s$^{-1}$, 0.3, 0.7) cosmology when necessary and we quote all magnitudes in the AB system (Oke & Gunn 1983).

2. DATA

In this paper, we make use of the large number of $z \gtrsim 4$ galaxies found in the ultra-deep HUDF data and the wide-area Early Release Science (ERS) data to construct median SEDs for star-forming galaxies at $z \sim 4$–7. The deep optical, near-IR, and deep IRAC coverage over these fields allows us to create SEDs extending from the rest-frame UV to the rest-frame optical down to very faint fluxes.

2.1. HST ACS and WFC3/IR

In constructing our median SEDs, we make use of the deep optical and near-IR HST data from ACS (GOODS program, Giavalisco et al. 2004) and WFC3/IR (e.g., Bouwens et al. 2011a). Over the ERS field (Windhorst et al. 2011), both the ACS optical $(B_{435}, V_{606}, i_{775})$ and the WFC3/IR $(Y_{998}, J_{125}, H_{160})$ images from HST reach depths of $\sim$28 mag ($5\sigma$ measured on 0.35 diameter apertures). In the HUDF field, the ACS optical data $(B'Viz)$ are 1.5–2 mag deeper, and the WFC3/IR data $(Y_{105}, J_{125}, H_{160})$ are 1.5 times deeper than the ERS. The HUDF was observed with the $Y_{105}$ filter instead of the $Y_{998}$ filter used in the ERS.

2.2. Spitzer/IRAC

Because these fields are located inside the GOODS-S field, they both have deep Spitzer/IRAC coverage that amounts to $\sim$23.3 hr over the ERS, and twice that for the HUDF which was observed to the same depths but in two different epochs with the IRAC camera rotated by 180° (Giavalisco et al. 2004). For IRAC channel 1 images ($\lambda_{\nu} = 3.6 \mu m$), each single epoch reaches depths of 27.8 mag, measured over 2.5 diameter apertures (1$\sigma$). For channel 2 ($\lambda_{\nu} = 4.5 \mu m$), this limit corresponds to 27.2 mag.

3. SAMPLE SELECTION AND BASIC PHOTOMETRY

The criteria used to identify high-redshift star-forming galaxies have already been presented in numerous previous works. For details on the selection procedure and discussion of sample contamination please see Bouwens et al. (2007, 2011b). In short, the selection criteria used to find the sources presented here are as follows:

- $z \sim 4$ B-dropouts:

  $$(B_{435} - V_{606} > 1.1) \land [B_{435} - V_{606} > (V_{606} - z_{850}) + 1.1] \land (V_{606} - z_{850} < 1.6)$$

- $z \sim 5$ V-dropouts:

  $$[V_{606} - i_{775} > 0.9(i_{775} - z_{850}) + 1.9] \lor (V_{606} - i_{775} > 2) \land (V_{606} - i_{775} > 1.2) \land (i_{775} - z_{850} < 1.3)$$

- $z \sim 6$ i-dropouts:

  $$i_{775} - z_{850} > 1.3 \land (z_{850} - J_{125} < 0.8)$$

The combined samples contain a total of 729 sources at $z \sim 4$, 178 at $z \sim 5$, and 83 at $z \sim 6$ (see Table 1).

We complement the $z \sim 4$, 5, and 6 samples with the sample of 36 $z \sim 7$ sources presented by Labbé et al. (2010a). We have adopted the $z \sim 7$ stacks presented in that work.

At these depths, the Spitzer/IRAC images are fairly crowded due to the size and shape of its point spread function (PSF). As a consequence, standard aperture photometry of the sources generally results in fluxes that are contaminated by flux from nearby sources. We are able to de-blend the fluxes from the different sources by making use of the higher resolution information available through the HST images. Our method has been explained in previous works by our group (e.g., Labbé et al. 2006, 2010b; González et al. 2010, 2011) but we give a short description here. We start by registering the IRAC images to the higher resolution HST image that will be used as a light profile template. In this case, we use the WFC3/IR $H_{160}$-band image as our template because it is the closest in wavelength and has high S/N. We use the SExtractor software (Bertin & Arnouts 1996) to create segmentation maps. We isolate all the sources in a small area around each galaxy in our sample using

| $z$ | ERS | HUDF | Total |
|-----|-----|------|-------|
| 4   | 270/524 | 77/123 | 407/729 |
| 5   | 137/205 | 41/55 | 178/360 |
| 6   | 118/178 | 58/83 | 276/260 |

Notes. Summary of the number of sources in the sample. IRAC photometry of these sources requires fitting and subtraction of the flux from surrounding foreground neighbors. This is not possible in all the cases. The table shows the number of sources with clean IRAC photometry/total number of sources found in the field.

* We adopt the sample presented in Labbé et al. (2010a).
these segmentation maps. We then convolve them with a kernel derived from the PSFs of the HST and the IRAC images. We fit for the normalization fluxes of all sources (including the source we are trying to clean) simultaneously. Finally, we remove the flux from the unrelated sources to obtain a clean image of the galaxies in our sample. Later, we use these cleaned images to produce median stacks.

The procedure described above is not always successful in subtracting flux from neighboring sources at a satisfactory level such that reliable flux measurements are possible. The reasons why it can fail vary but the most common is the presence of a neighboring source that is too bright, extended, and close to our region of interest. In these cases, the amount of flux subtracted off our source of interest is large and more uncertain and in most cases results in large residuals. We have carefully inspected the residual images (all sources subtracted, including the one in our sample) and the cleaned images (only unrelated neighbors subtracted) of each of the sources in the original sample and come up with a sub-sample of sources with clean IRAC photometry. This “by eye” inspection criterion generally results in good agreement with a criterion that selects only the IRAC photometry. This selection does not introduce biases relevant to the determination of the mean rest-frame optical colors of these sources.

4. STACKED PHOTOMETRY AND MEDIAN SEDs

It has been shown that in star-forming galaxies, the total SFR appears to correlate with stellar mass, the so-called main sequence of star-forming galaxies (Noeske et al. 2007). This sequence has been observed at high redshifts ($z \gtrsim 4$) but contrary to what has been observed at $z < 2$, the normalization of the relation at fixed stellar mass does not seem to evolve over $2 < z < 7$ (Stark et al. 2009; González et al. 2011). The lack of deeper rest-frame optical from Spitzer/IRAC for these sources is one of the big limitations to study this sequence to fainter magnitudes. In terms of the bulk of the population, however, progress can be made by stacking large numbers of sources to study their mean properties. In this section, we seek to construct the median SEDs of galaxies at $z \gtrsim 4$.

Due to the nature of the dropout search and the filters available, our full sample is naturally divided into redshift bins centered at $\langle z \rangle = 3.8, 5.0, 5.9$ plus the $\langle z \rangle = 6.9$ sample from Labbè et al. (2010a). In this section, we further split each redshift sub-sample based on their observed rest-frame UV luminosity. The filters used as reference UV luminosity are the $i_{775}$ filter for the $z \sim 4$ sources; the $z_{850}$ filter for the $z \sim 5$ sources; the $y_{098}$ filter for the $z \sim 6$ sources in the ERS and the $y_{105}$ filter for the $z \sim 6$ sources in the HUDF; and finally, the $j_{125}$ filter for the $z \sim 7$ sources. These filters were chosen to be close to rest frame 1500 Å at the different redshifts and make the splitting criterion fairly uniform across redshifts. These bins are presented in Figure 1.

In the case of the HST ACS and WFC3 imaging, the fluxes utilized correspond to the MAG_AUTO values determined using the SExtractor code plus an additional correction to account for the flux outside the MAG_AUTO aperture. These corrections are estimated assuming a point-source profile and are generally in the range 0.1–0.3 mag. The corrections are larger for the fainter sources (because they generally have smaller apertures). In each UV-luminosity bin, the median flux corresponds to the median of the individual measurements. Only sources with reliable IRAC fluxes were included in the median. The uncertainty in the median was estimated through bootstrap resampling simulations performed as follows. These simulations involved creating a large number of realizations of the sample, by drawing, with replacement, $N_{\text{bin}}$ fluxes (from the total $N_{\text{bin}}$ values available). The fluxes are perturbed within their individual uncertainties before estimating the median of the realization. The uncertainties on individual sources are generally
on the stacked images using a circular aperture of 2
sizeable residuals that we may have overlooked in inspecting the cleaning procedure. The pixel by pixel median also removes any sources within a bin that have clean residuals from our IRAC median combine (pixel by pixel) the cleaned stamps of the integration images. Instead of using individual fluxes, we first apertures of 2′′. The magnitudes included in the stamps were measured using circular apertures of 2′′ in diameter and corrected to total assuming stellar profiles which amount to ∼−0.6 mag in both the [3.6] and the [4.5] channels. The background was estimated from the stacks and subtracted off the images. To estimate the uncertainties on the fluxes measured, we created 200 random realizations of the stack. In each realization, sources were drawn with replacement and then median stacked pixel by pixel.

Figure 2. Median-stacked images of the z ∼ 4 sources. The sample has been split according to their i_{775,AB} magnitude (approximate rest frame 1500 Å) in bins of 1 mag. The stamps are 10′′ on a side. The number of sources stacked along with the total equivalent integration time has been included to the left of the stamps. The magnitudes included in the stamps were measured using circular apertures of 2′′ in diameter and corrected to total assuming stellar profiles which amount to ∼−0.6 mag in both the [3.6] and the [4.5] channels. The background was estimated from the stacks, in an annulus between 6′′ and the ∼−4 sources. The sample has been quite small, given the very high S/N of the images. There are, however, systematic uncertainties in the absolute calibration of the different bands and instruments across the SED. These uncertainties are typically no smaller than 5%, so we adopt this value as our minimum allowed uncertainty for any given flux.

In the case of the IRAC imaging, most sources are too faint to be individually detected even in these deep ∼ 23.3 hr integration images. Instead of using individual fluxes, we first median combine (pixel by pixel) the cleaned stamps of the sources within a bin that have clean residuals from our IRAC cleaning procedure. The pixel by pixel median also removes any sizeable residuals that we may have overlooked in inspecting the individual fits. We perform standard aperture photometry on the stacked images using a circular aperture of 2′′ diameter and correct the fluxes to the total assuming PSF profiles. This amounts to a factor of 1.8 correction to the fluxes in both the [3.6] and the [4.5] IRAC filters. The local background is estimated from the stacks, in an annulus between 6′′ and 10′′ diameters around the stamp center, where the source is expected to be, and then subtracted off the image.

It is important to note that this procedure is very different from the way the HST stacking is done. We have checked that this does not impose significant biases in our measurement by applying the same procedure as described for the IRAC stacks to the H_{160}—band stacks. We first masked out the neighboring sources and then convolved the H_{160} image with a kernel to degrade it to the resolution of the IRAC images. We have median combined these images and performed simple aperture photometry in the same way as we did for the IRAC bands, using 2′′ diameter apertures and correcting to the total fluxes assuming stellar profiles (factor 1.8). We find that the H_{160}—IRAC colors measured with the two methods are in excellent agreement (agreeing to within 0.05 mag), indicating that we are not introducing any significant bias in the estimates of the UV-to-optical colors.

To estimate the uncertainty in the IRAC stacks, we also create bootstrap realizations. We treat the two epochs of IRAC data over the HUDF as independent stamps in the stack, thereby effectively weighting with exposure time, i.e., N_{bin} = N_{ERS} + 2N_{HUDF}. In each bootstrap draw then, we draw with replacement N_{ERS} + 2N_{HUDF} such stamps and estimate fluxes as described above. The image noise in each stack is dominated by the background noise (shot noise from the sources will contribute less than 2% even for the brightest sources in the sample). This image noise is expected to decrease with the total equivalent exposure time approximately as 1/\sqrt{t_{exp,eq}} (t_{exp,eq} = (N_{ERS} + 2N_{HUDF}) × 23.3 hr). The actual photometric uncertainties that we derive are always greater than this because they also include the intrinsic scatter in the fluxes of the sources being stacked together (see the Appendix). Stamps for all the IRAC median stacks along with the derived magnitudes and uncertainties are presented in Figures 2–4. Table 2 is a compilation of all the stacked photometry including the Labbé et al. (2010a) stacks and information on the number of sources that go in each stack. It can be seen in the figures that even for sources as faint as 27 mag, the stacks show good detections in both IRAC bands.

5. REST-FRAME UV-TO-OPTICAL COLORS OF HIGH-z STAR-FORMING GALAXIES

The observed colors of galaxies can give us important information about the stellar populations that compose them.
| Ref. Mag | $N_{\text{ERS}}$ | $N_{\text{HUDF}}$ | $B_{455}$ | $V_{606}$ | $i_{775}$ | $z_{850}$ | $Y_{098}$ | $Y_{105}$ | $J_{125}$ | $H_{160}$ | $[3.6]$ | $[4.5]$ |
|----------|----------------|----------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| $i_{775}$ | $z \sim 4$ B-dropouts |
| 25       | 35/72          | 0/4            | 33.0 ± 6.1 | 226.3 ± 15.3 | 356.2 ± 17.6 | 375.9 ± 21.9 | 359.9 ± 21.0 | 389.3 ± 30.2 | 467.9 ± 50.4 | 687.5 ± 50.2 | 635.4 ± 63.6 |
| 26       | 93/176         | 12/18          | 13.5 ± 2.3 | 106.6 ± 3.5 | 154.3 ± 4.2 | 153.2 ± 5.3 | 146.4 ± 6.3 | 142.4 ± 9.0 | 149.5 ± 5.5 | 162.0 ± 6.9 | 236.9 ± 13.8 | 246.4 ± 16.2 |
| 27       | 113/218        | 34/51          | 8.3 ± 1.3 | 46.8 ± 1.4 | 59.6 ± 1.7 | 53.6 ± 2.6 | 57.0 ± 3.1 | 54.0 ± 3.3 | 57.1 ± 2.3 | 62.0 ± 2.9 | 72.1 ± 5.5 | 61.8 ± 6.8 |
| 28       | 24/44          | 38/62          | 3.3 ± 1.2 | 25.0 ± 1.5 | 26.3 ± 1.4 | 27.3 ± 1.9 | 32.1 ± 4.3 | 21.3 ± 1.4 | 26.5 ± 1.8 | 29.3 ± 2.6 | 40.5 ± 5.0 | 35.7 ± 7.3 |
| $z_{850}$ | $z \sim 5$ V-dropouts |
| 25       | 4/9            | 0/0            | −10.1 ± 11.7 | 47.2 ± 8.6 | 249.5 ± 49.4 | 305.1 ± 25.1 | 309.2 ± 25.2 | 315.6 ± 20.9 | 288.7 ± 21.4 | 681.9 ± 67.1 | 498.8 ± 50.2 |
| 26       | 26/40          | 3/4            | −3.6 ± 3.6 | 17.0 ± 3.6 | 114.5 ± 11.1 | 130.5 ± 7.7 | 129.3 ± 11.3 | 126.7 ± 15.3 | 132.2 ± 7.5 | 140.2 ± 11.0 | 268.8 ± 26.0 | 202.3 ± 26.2 |
| 27       | 41/62          | 9/10           | −1.1 ± 2.1 | 8.5 ± 1.4 | 58.7 ± 1.4 | 65.5 ± 3.5 | 69.1 ± 5.6 | 56.5 ± 9.4 | 67.4 ± 5.0 | 62.4 ± 5.2 | 95.7 ± 12.9 | 69.2 ± 10.4 |
| 28       | 6/12           | 9/14           | −2.3 ± 2.1 | 2.6 ± 1.5 | 25.9 ± 2.6 | 25.2 ± 2.9 | 21.0 ± 10.4 | 22.0 ± 4.2 | 24.9 ± 3.3 | 27.7 ± 4.9 | 41.3 ± 15.9 | 34.2 ± 16.0 |
| $Y$      | $z \sim 6i-$dropouts |
| 26       | 4/8            | 3/3            | 0.0 ± 3.9 | −3.3 ± 3.6 | 10.4 ± 7.0 | 97.8 ± 15.5 | 116.9 ± 12.8 | 157.3 ± 24.1 | 124.9 ± 16.7 | 148.8 ± 26.3 | 255.4 ± 61.2 | 172.5 ± 51.6 |
| 27       | 12/17          | 3/6            | −0.6 ± 3.5 | 0.8 ± 2.6 | 11.7 ± 6.6 | 71.2 ± 7.0 | 64.7 ± 7.4 | 49.9 ± 5.4 | 71.2 ± 6.4 | 72.3 ± 9.1 | 134.8 ± 24.6 | 84.2 ± 27.4 |
| 28       | 2/3            | 9/14           | −2.7 ± 2.5 | 1.8 ± 2.0 | 2.6 ± 2.7 | 28.2 ± 4.7 | 29.7 ± 14.0 | 21.6 ± 2.9 | 27.5 ± 4.1 | 21.2 ± 3.9 | 27.1 ± 10.7 | 44.6 ± 17.3 |
| $H_{160}$ | $z \sim 7$ z-dropouts |
| 26       | ...            | ...            | −3.2 ± 2.3 | 2.8 ± 2.3 | −7.4 ± 4.7 | 23.5 ± 5.4 | 63.7 ± 7.7 | 129.7 ± 16.0 | 128.2 ± 17.0 | 262.0 ± 28.0 | 181.0 ± 37.0 |
| 27       | ...            | ...            | −2.4 ± 0.8 | −0.8 ± 0.6 | 2.2 ± 1.8 | 14.6 ± 2.3 | 49.1 ± 4.3 | 64.8 ± 4.1 | 57.0 ± 3.4 | 107.0 ± 16.0 | 83.8 ± 25.0 |
| 28       | ...            | ...            | 1.2 ± 0.8 | 0.4 ± 0.5 | −0.1 ± 0.7 | 5.6 ± 1.1 | 27.2 ± 1.8 | 29.6 ± 2.0 | 24.5 ± 1.6 | 45.1 ± 9.5 | 39.3 ± 17.0 |

Notes. The median SEDs of high-$z$ galaxies in bins of observed UV luminosity. The fluxes are in units of nJy. $N_{\text{HUDF}}$ and $N_{\text{ERS}}$ are the numbers of sources with reliable IRAC photometry over the total available in the given bin. The effective number of stacked images in a bin corresponds to $N_{\text{ERS}} + 2N_{\text{HUDF}}$. Only those galaxies with reliable photometry were considered in the stack. The $z \sim 7$ stacks were taken directly from Labbé et al. (2010a).
was split according to their $Y$ Salpeter (1955) IMF, constant SFH, and 0 the Bruzual & Charlot (2003, BC03) models and we assume a on luminosity and redshift. When models are invoked, we use

and what can be learned from them and from their dependence work, we chiefly explore the observed UV-to-optical colors

those of Synthetic Stellar Population (SSP) models. In this

Figure 4. Same as Figure 2, but for the $z \sim 6$ sample. In this case, the sample was split according to their $Y_{98}$ magnitude in the case of the ERS sources and $Y_{105}$ for the HUDF sources (approximately rest frame 1500 Å). As in the faintest bin in Figure 2, the faintest bin at this redshift also shows a rather weak detection in IRAC.

Usually, this is accomplished by comparing their colors with those of Synthetic Stellar Population (SSP) models. In this work, we chiefly explore the observed UV-to-optical colors and what can be learned from them and from their dependence on luminosity and redshift. When models are invoked, we use the Bruzual & Charlot (2003, BC03) models and we assume a Salpeter (1955) IMF, constant SFH, and 0.2 $Z_{\odot}$ metallicity, as these quantities cannot be constrained from our data. To derive best fits, we use the SED fitting code FAST (Kriek et al. 2009).

The rest-frame UV-to-optical color is of particular interest at high redshifts. Several studies have detected significant rest-frame UV-to-optical colors ($\gtrsim 0.5$ mag) in $z \gtrsim 4$ galaxies (Eyles et al. 2005, 2007; Yan et al. 2005; Stark et al. 2009; González et al. 2010, 2011; Labbé et al. 2006, 2010a; McLure et al. 2011). The usual interpretation is that these colors derive from large Balmer breaks which are associated with evolved stellar populations with ages $\gtrsim 200$ Myr. This is surprising given how young the universe is at the time these galaxies are being observed. The main limitation for field studies is the depth of the Spitzer/IRAC imaging available, which constrains these studies to the brightest LBGs. Searches behind clusters provide an alternative to probe the fainter populations (see, e.g., Ellis et al. 2001; Bradley et al. 2008; Richard et al. 2008; Laporte et al. 2011, and the recent tentative determination of a 1.5 mag Balmer break in a $z \sim 6$ galaxy by Richard et al. 2011), but the number of sources amenable to such studies remains low. We work around this limitation by stacking large numbers of sources which allows us to estimate the typical rest-frame UV-to-optical colors of faint $z \gtrsim 4$ sources.

We start by presenting the results found at $z \sim 4$, where the similar and complementary work of Lee et al. (2011) allows us to study these colors over an unprecedentedly large range of UV luminosities.

5.1. The UV-to-optical Colors at $z \sim 4$

Our sample of $B$-dropout has a mean redshift of $(z) = 3.8 \pm 0.3$. We divide this sample based on their $i_{775}$ luminosity (which approximately corresponds to rest frame 1500 Å) in four bins of $\Delta i_{775} = 1$ mag. These are the same bins as shown in the histogram in Figure 1. The bins at $i_{775} > 28$ mag are too faint to be detected in Spitzer/IRAC and will not be included in the analysis. Assuming the mean redshift of the sample, these bins correspond approximately to $M_{1500} = -21, -20, -19$, and $-18$, which covers the moderately bright to faint end of the population ($M_{1500}^{*} = -20.24$; Bouwens et al. 2007).

In a recent study, Lee et al. (2011) focus on the median SEDs of the brightest $z \sim 4$ galaxies found in the ground-based NOAO Deep Wide-Field Survey. Their sample has a mean redshift $(z) = 3.7 \pm 0.4$, and their SEDs also sample the rest-frame UV and optical using a combination of optical filters: $B_{n}, R, I$ (Mosaic Camera), near-IR filters: $J, H, K_{s}$ (NEWFIRM Camera), and the Spitzer/IRAC mid-IR channels 1 through 4. Similar to the present work, they divide their sample according to the $I$-band luminosity, although with uneven bins. Their sources are intrinsically brighter than in this work, corresponding to $L > L_{*}^{\odot}$ sources roughly covering the range $-23 \lesssim M_{1500}^{*} \lesssim -21$.

7 The combination of the Lee et al. (2011) sample with our data spans the UV-luminosity range $-23 \lesssim M_{1500}^{*} \lesssim -17.5$. This represents an unprecedented data set to study the SEDs, and especially the UV-to-optical colors, of $z \sim 4$ galaxies over a large range of luminosities. The SEDs from both works are presented in Figure 5. In this figure, the $H_{160}$-band flux measurements are shown with open symbols since it is likely biased and should be ignored as explained in detail later. An overall trend to bluer colors (both in the rest-frame UV and the rest-frame UV-to-optical colors) toward fainter luminosities is already apparent in this figure.

We quantify the UV-to-optical colors by estimating the interpolated rest-frame $(U_n - V_n)$ colors for the median-stacked SEDs in both sets. Figure 6 shows a systematic trend of redder $(U_n - V_n)^{\text{rest}}$ colors as a function of increasing UV luminosity.

The interpolated $(U_n - V_n)$ colors were determined from the best fits to the full observed SED, where the fits correspond to a linear combination of a set of template SEDs. These SEDs correspond to the default template set from the photometric code EAZY (Brammer et al. 2008). This template set was constructed and calibrated empirically to determine accurate photometric redshifts of galaxies of various types from very blue starbursts to red ellipticals and has been shown to work well up to $z \sim 5$–6 (Brammer et al. 2008). For the $U_n$ and $V_n$ filters, we have assumed narrow ideal filters that correspond to step functions of width of 100 Å, centered at 3500 Å and 5500 Å for the $U_n$ and $V_n$ filters, respectively. The colors were calculated from the best-fit combination of templates using the filters described above. To estimate the uncertainty in the $(U_n - V_n)^{\text{rest}}$ estimates, we perturbed the photometry and the redshift of each median-stacked SED within the allowed uncertainty, and a new fit was obtained. A new set of colors was measured from this best fit and the procedure was repeated a large number of times to derive the confidence intervals. A similar procedure was used to derive rest-frame $M_{1500}$ magnitudes for an ideal filter centered at 1500 Å and with a width of 100 Å.

4 Although relatively deep IRAC channels 3 and 4 data (sampling 5.8 $\mu$m and 8.0 $\mu$m respectively) are available for our sample in the GOODS-S, these data are still $\lesssim 0.5$ mag shallower than the channels 1 and 2. It is, therefore, much harder to obtain reliable stacked photometry in those bands, especially considering that our sample is much smaller and intrinsically fainter than the Lee et al. (2011) sample. We have not included these bands in the median-stacked SEDs presented in this work.
In recent works, Bouwens et al. (2009, 2011a) accurately determine the UV slope of LBGs at $z > 3$ as a function of UV luminosity and redshift. For a star-forming stellar population, the UV slope, parameterized by $\beta$ ($f_{\lambda} \propto \lambda^{\beta}$), is most sensitive to the dust content of the galaxy. In consequence, the authors argue that the dependence of $\beta$ on UV luminosity can be fully explained by an increase in the amount of dust that these galaxies contain as they become brighter. If this is the case, then a trend of redder ($U_n - V_n$)$_{rest}$ colors for UV-righter sources like the one observed here is also expected, since this color is also affected by dust extinction. The bottom panel of Figure 6 shows the ($U_n - V_n$)$_{rest}$ colors for UV-brighter sources corrected for dust extinction. These were derived as in Bouwens et al. (2009) from the mean $\beta$ versus UV-luminosity trend and the classic Meurer et al. (1999) relation...
that relates the dust content with the UV-slope \( \beta \) according to
\[
A_{1600} = 4.43 + 1.99 \times \beta,
\]
where \( A_{1600} \) is the extinction in magnitudes at 1600 Å. A Calzetti et al. (2000) extinction curve is used to estimate the extinction at other wavelengths.

As can be seen from the figure, the dust-corrected \( (U_n - V_n) \) color is essentially independent of the UV luminosity. This again suggests that the main origin for the correlation of \( \beta \) with the UV luminosity is an increasing dust content at brighter magnitudes. Otherwise, we would expect some residual trend in the dust-corrected color. This would be the case, for example, if a strong age trend existed with UV luminosity. As will be shown later, a similar flattening of the corrected colors is observed at all redshifts in the \(-21 < M_{1500} < -18\) range.

### 5.2. The UV-to-optical Colors at Higher Redshift

We now extend our determination of the SED of high-\( z \) galaxies to \( z \sim 5 \) and \( z \sim 6 \), and combine our determinations with the stacked SEDs of sources at \( z \sim 7 \) recently presented in Labbé et al. (2010a). Figure 7 shows the SEDs at \( 4 \lesssim z \lesssim 7 \). The expected mean redshifts of these samples correspond to \((z) = 3.8, (z) = 5.0, (z) = 5.9, \) and \((z) = 6.9, \) with typical spread in redshift \( \Delta z \) of \( \pm 0.3 \). The samples at each redshift have been split into bins of 1 mag, using as reference the observed magnitude in the band closest to 1500 Å (Table 2 contains the SEDs and information about the samples in each bin). Best-fit BC03 models with constant star formation (CSF) and 0.2 Z⊙ metallicity are overlaid for reference. As explained in the previous section, we have excluded from the fitting the bands that are potentially biased (marked by the open symbols). As can be seen in this figure, all these SEDs are remarkably similar and show sizable UV-to-optical colors.

The similarity of the SEDs can also be appreciated in Figure 8, where the SEDs from the different redshift samples have been grouped according to their approximate \( M_{1500} \) luminosities. The SEDs in the three \( M_{1500} \) bins have been re-normalized according to their \( M_{1500} \) luminosity and combined to produce oversampled SEDs, which take advantage of the fact that the different filters probe different rest-frame wavelengths at the different redshifts. The combined SEDs show the typical flat UV slopes, but also rather flat colors in the rest-frame optical, albeit with a relatively large scatter. This scatter can in part be a consequence of the normalization which is done in the far UV, which effectively can also be contributing to this scatter.

#### 5.2.1. Emission Lines in the Rest-frame Optical

We first consider the implications if the sizable red \( J_{125} - [3.6] \) colors that we observe are assumed to be exclusively measuring the stellar continuum in these high-redshift galaxies. The red color would arise from large Balmer breaks and would be indicative of evolved stellar populations. As can be seen in Figure 9(a), the \( J_{125} - [3.6] \) colors characteristic of these sources would then require fairly large ages of \(~1 \) Gyr if no dust is allowed in the models (the model tracks in the figure are CSF models). Allowing some reddening by dust (using in this case a Calzetti et al. 2000 law) reduces their ages considerably (Figure 9(c)). The amount of dust reddening is constrained however. Based on the UV continuum slopes typically observed in these galaxies (Bouwens et al. 2009), and using the \( z = 0 \) Meurer et al. (1999) relation between this slope and the dust content, it has been shown that the reddening at \( z \gtrsim 4 \) is typically \( A_V \lesssim 0.7 \) mag. The reddening is likely to be even smaller at higher redshifts (Bouwens et al. 2009, 2011a). With this added constraint, the ages derived from SED fitting are generally estimated to be 300–400 Myr (e.g., Yan et al. 2005; Eyles et al. 2005; Stark et al. 2009; González et al. 2010; Labbé et al. 2010a).

In our SEDs at \( z \gtrsim 5 \), we also find that the \( ([3.6] - [4.5]) \) color is consistently blue, with typical values of \(~-0.3 \) mag (the only exception is the faintest \( z \sim 6 \) bin but this color is very uncertain). This color is very hard to reproduce with a model with any SFH that is based only on the stellar continuum. The addition of dust only makes the comparison worse (Figures 9(c) and (d)). It should be noted, however, that the uncertainties in this color are considerable and that in most cases the observed color is formally consistent with the stellar continuum only models.

It has been claimed (Schaefer & de Barros 2009, 2010) that at some of these redshifts, particularly \( z \sim 6 \) and 7, the SEDs could be better fitted by very young stellar population models \((\lesssim 10 \) Myr old) with moderate dust content and very strong nebular emission lines. These lines would dominate the rest-frame optical fluxes. As the galaxies are forming stars, it is likely that they have nebular emission lines, and so we have assessed whether emission lines could be influencing our fluxes and fits. The observed position of the most prominent optical nebular emission lines at each redshift is marked in Figure 7 by the dashed vertical lines (assuming the mean redshift of each sample). It can be seen that at the different redshifts involved, some of the rest-frame optical emission lines would lie within the IRAC \([3.6]\) and \([4.5]\) filters.

Atek et al. (2011) showed recently that at \( z < 2.8 \), there are sources that show emission lines with large enough equivalent widths to dominate the broadband fluxes in the optical. However, even for their sample of high equivalent width selected sources, such objects are not very common. For the more typical cases, the lines that they find in their sample would have moderate contributions to the IRAC broadbands, usually making them \(0.2-0.3 \) mag brighter. At higher redshifts, \( 3.8 \lesssim z \lesssim 5 \), the work of Shim et al. (2011), based on spectroscopic redshifts and IRAC broadband photometry, finds similar results. They find that the sources in this redshift range show an excess of flux in \([3.6]\) over \([4.5]\), which the authors think can be explained by Hα emission. The majority of the sources in their sample have \( ([3.6] - [4.5]) \) colors that are consistent with an Hα contribution of \(0.2-0.3 \) mag to the \([3.6]\) filter, indicating moderate contributions from emission lines.

We consider the effects of a similar (moderate) contribution of emission lines to the colors of a 280 Myr old CSF model. For the purpose of this exercise, we assume an Hα rest-frame equivalent width \( EW_{\text{rest}} = 300 \) Å. At \( z \sim 6 \), this would increase the luminosity in the \([4.5]\) filter by \(0.23 \) mag. We only consider the strongest lines in the optical and assume the strength ratios presented by Anders & Fritze-v. Alvensleben (2003) for a 0.2 Z⊙ metallicity system. This corresponds to \( EW_{\text{rest}}(\text{H}α, \text{H}β, \text{O}iii) = (189, 105, 670) \) Å, respectively. The open squares in Figures 9(a) and (b) correspond to the median colors for a model like the one described (the median takes into account the expected redshift distribution of each sample).

As can be seen in Figure 9(a), the main effect of the contribution from the emission lines is that it allows for significantly younger ages to reproduce the \( J_{125} - [3.6] \) colors (by a factor...
Figure 7. Stacked SEDs of galaxies in units of observed magnitudes (see also Table 2). The x-axis shows wavelength and the approximate filter that it corresponds to in our filter set for reference (note that this filter set is different to the one in the x-axis of Figure 5). In the case of the optical bands, the errors were derived by bootstrap re-sampling the measured fluxes. The individual uncertainties were set to a minimum of 5% to account for systematic uncertainties in the filter to filter absolute calibrations. The errors in the IRAC bands were derived by bootstrap re-sampling the individual images and repeating the pixel by pixel median-stack process. In all cases, these errors on the median should include both the image noise and the uncertainty arising from variations within the population included in the stack. Simple best-fit BC03 models with CSF are also included for reference. In the fitting process, the redshifts were fixed to the median redshift of each sample, i.e., \(\langle z \rangle = 3.8\), \(\langle z \rangle = 5.0\), \(\langle z \rangle = 5.8\), and \(\langle z \rangle = 6.9\), respectively. Due to the intrinsic redshift distribution of each sample, the median fluxes measured for bands near Ly\(\alpha\) or the Balmer break are likely biased and in consequence they were excluded from the fitting process (excluded bands are marked by the open symbols). The position of the most prominent possible emission lines in the rest-frame optical region of the SED are marked by the vertical dashed lines (assuming the mean redshift of each sample). At the different redshifts sampled here, different lines could be contaminating the continuum fluxes measured by the Spitzer/IRAC filters. The overall shape and UV-to-optical colors of the SEDs, however, remain remarkably constant with redshift. The small relative excess of [3.6] over [4.5] fluxes observed in most SEDs at \(z \gtrsim 5\) suggests that the nebular emission lines indeed contribute to the observed flux in the rest-frame optical (the exception is the faintest bin at \(z \sim 6\) which has very uncertain UV-to-optical colors due to the weak IRAC detections—see Figure 4).

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

of two or more). Since this implies lower stellar continuum fluxes than previously assumed, this would also result in lower best-fit stellar masses. Typical masses (log(M/M\(_{\odot}\)) \sim 8.7–10; González et al. 2011) will decrease by a factor \(\gtrsim 2\). Simultaneously, given the line ratios and the redshift distributions expected for these samples, a model like this can also help in reproducing the blue [3.6]–[4.5] colors of the SEDs at \(z \gtrsim 5\) (Figure 9(b)). While formally models with just stellar continua can fit the data, given the current large uncertainties, such colors are hard to obtain with stellar continuum only models. It is more likely that some degree of contribution of emission lines provides a better solution.

The assessment presented here is necessarily very simplistic, given the current uncertainties, and is only meant to show the possible effects of emission lines on the colors of these galaxies. Better estimates of the contribution of these lines and its effects on the properties derived from SED fitting will depend on the actual strengths and ratios of the lines. Moderate improvement can be achieved if more precise redshifts are known for these sources; such data do not exist at this time for such faint high-redshift galaxies and so this remains an open question until improved spectroscopic capability becomes available.

5.2.2. UV-to-optical Color versus Luminosity

Regarding the trend of redder UV-to-optical colors for brighter UV luminosities presented in Figure 6 for the \(z \sim 4\) sample, it can also be observed for the other samples. In fact, within the uncertainties in the color determination, the \(J_{125} - [3.6]\) versus \(M_{1500}\) trend is roughly the same at \(z \sim 5\) and 6 (Figure 10, top). This is particularly remarkable in view of the fact that this color is an observed color, not a rest-frame color, and in consequence is probing different wavelengths for the
Intrinsic magnitude of the SEDs that are being combined. At different redshifts, our set of HST UV-to-optical colors of the SEDs depend weakly on redshift. Here, we perform a simultaneous fit to the rest of the sample excluding the $z \sim 4$ points results in a slightly flatter but consistent relation $(U_n - V_n) = -0.024(\pm 0.039) \times M_{1500} - 0.039(\pm 0.783)$. The slope derived for the full sample is in very good agreement with the $-0.04$ slope found for the $z \sim 4$ data only, also plotted for reference as the black dashed line in Figure 10. The latter trend included the bright sample of Lee et al. (2011) and is consistent within the uncertainties with the fit to all the points.

In the bottom panel of Figure 10, we present an estimate of the reddening corrected $(U - V)$ colors. The reddening corrections are derived assuming a Calzetti et al. (2000) extinction curve and using the local Meurer et al. (1999) relation between dust reddening and the UV-slope $\beta$ (Equation (1)). The values of $\beta$ are extracted from the best-fit trends measured by Bouwens et al. (2011a) at the corresponding luminosities. These corrected colors show large scatter around $(U - V)_{\text{rest,corrected}} \sim 0.3$, but there is no indication of a residual trend with UV luminosity except for the $z \sim 7$ sample (black symbols). This may indicate that again at $4 \lesssim z \lesssim 6$, the evolution can be fully explained by a change in dust content and there is no strong evolution in the ages of these sources as a function of UV luminosity. This is consistent with the report by Stark et al. (2009) based on stellar population modeling of slightly brighter LBGs in this redshift range. It is also consistent with the predictions from the hydrodynamical simulations of Finlator et al. (2011), which find no trend of age with UV luminosity.

At $z \sim 7$, there seems to be a residual trend in $(U_n - V_n)_{\text{rest,corrected}} \sim 0.4$ versus UV luminosity in the sense that brighter sources have systematically bluer intrinsic $(U_n - V_n)$ colors. In principle, this could indicate that the brighter sources are actually younger than the fainter sources, which would be difficult to imagine. Alternatively, it is possible that simple dust corrections (calibrated at $z = 0$) are not adequate at these high redshifts, due to, for example, changes in the IMF, metallicity, or escape fraction, all of which could alter the UV-slope $\beta$ for a given dust content. It should also be noted that the $\beta$ versus UV luminosity at $z \sim 7$ is slightly steeper than at the other redshifts (Bouwens et al. 2011a) and so the corrections to the $(U_n - V_n)_{\text{rest}}$ versus UV-luminosity trend are more extreme. Nevertheless, the uncertainties are sufficiently large that the $z \sim 7$ trend is also fully consistent with being constant.

6. SUMMARY

Determining the rest-frame UV-to-optical colors of individual UV-faint galaxies at $z \gtrsim 4$ is challenging. At such redshifts, the rest-frame optical lies at $\lambda > 2.5 \mu m$ in the mid-IR and so it is hard to measure at the extremely faint magnitudes typical of galaxies in the first 1.5 Gyr. To gain insight into the typical spectral properties of $z \sim 4$–6 galaxies, we have taken advantage of the large samples of $z \sim 4$–6 sources found in the deep HST ACS and WFC3/IR images of the ERS and HUDF fields. The HST data give us rest-frame UV fluxes. We then use Spitzer/IRAC [3.6] and [4.5] $\mu m$ data from the deep GOODS survey of these fields to measure rest-frame optical fluxes by stacking the IRAC images and determining median flux values.

\begin{table}[h]
\centering
\caption{Rest-frame Magnitude and Colors} 
\begin{tabular}{|c|c|c|c|}
\hline
$M_{1500}$ & $J_{125} - [3.6]$ & $(U - V)_{\text{rest}}$ & $(U - V)_{\text{corrected}}$
\hline
$-20.84 \pm 0.22$ & $0.62 \pm 0.12$ & $0.48 \pm 0.08$ & $0.29$
\hline
$-19.91 \pm 0.17$ & $0.50 \pm 0.07$ & $0.38 \pm 0.05$ & $0.25$
\hline
$-18.84 \pm 0.16$ & $0.25 \pm 0.09$ & $>0.32$ & $>0.24$
\hline
$-17.97 \pm 0.15$ & $0.46 \pm 0.15$ & $0.36 \pm 0.05$ & $0.33$
\hline

$z \lesssim 5$ & & & 
\hline
$-21.14 \pm 0.13$ & $0.84 \pm 0.13$ & $0.48 \pm 0.06$ & $0.29$
\hline
$-20.21 \pm 0.14$ & $0.77 \pm 0.12$ & $0.46 \pm 0.06$ & $0.33$
\hline
$-19.46 \pm 0.13$ & $0.38 \pm 0.17$ & $0.31 \pm 0.03$ & $0.22$
\hline
$-18.39 \pm 0.16$ & $0.55 \pm 0.44$ & $0.42 \pm 0.12$ & $0.39$
\hline

$z \gtrsim 6$ & & & 
\hline
$-20.42 \pm 0.19$ & $0.78 \pm 0.30$ & $0.43 \pm 0.12$ & $0.37$
\hline
$-19.68 \pm 0.13$ & $0.69 \pm 0.22$ & $0.46 \pm 0.10$ & $0.43$
\hline
$-18.59 \pm 0.14$ & $-0.02 \pm 0.46$ & $0.40 \pm 0.12$ & $0.42$
\hline

$z \lesssim 7$ & & & 
\hline
$-20.71 \pm 0.20$ & $0.76 \pm 0.18$ & $0.48 \pm 0.11$ & $0.35$
\hline
$-19.89 \pm 0.12$ & $0.54 \pm 0.18$ & $0.45 \pm 0.12$ & $0.40$
\hline
$-19.00 \pm 0.12$ & $0.46 \pm 0.24$ & $0.48 \pm 0.16$ & $0.53$
\hline
\end{tabular}
\end{table}

Notes. Rest-frame $M_{1500}$ estimated from the best-fit power law to the rest-frame UV photometry. Possibly contaminated bands were ignored in the fit (see open points in Figure 7).

Different samples. This shows again how similar and flat these SEDs are. If all the SEDs are considered simultaneously, then a trend of $J_{125} - [3.6] = -0.17(\pm 0.07) \times M_{1500, AB} - 2.80(\pm 2.43)$ is found. For simple CSF models, this trend implies that the $M/L$ ratio depends on the UV luminosity. González et al. (2011) find that at $z \sim 4$, the $M/L$ ratio changes by a factor $\sim 5$ between $M_{1500} = -18$ and $M_{1500} = -21$. So, despite the steep UV-luminosity functions characteristic at $z \gtrsim 4$, the contribution of the faintest sources to the total stellar mass density is more modest than their contribution to the SFR density.

Similar to the analysis performed at $z \sim 4$, we have derived rest-frame $(U_n - V_n)$ colors for the median SEDs at all redshifts. The computed colors are summarized in Table 3, and shown in Figure 8. The SEDs of the combined samples in three bins of intrinsic UV luminosity: $M_{1500} \sim -21$, $-20$, and $-19$. As seen in Figure 10, the UV-to-optical colors of the SEDs depend weakly on redshift. Here, we perform the exercise of combining all the photometry from SEDs at different redshifts into a single SED. At different redshifts, our set of HST/ACS+WFC3/IR+Spitzer/IRAC filters probe different rest-frame wavelengths. The combined SEDs are plotted in terms of their magnitudes relative to $M_{1500}$, which is measured from the best power-law fit to the UV continuum of each median SED. An offset magnitude of $-21$, $-20$, and $-19$ is added to reflect the approximate intrinsic magnitude of the SEDs that are being combined.

(A color version of this figure is available in the online journal.)
This allows us to determine the SEDs of these galaxies covering both the rest-frame UV and the optical. These SEDs represent a first comprehensive study of the rest-frame UV-to-optical properties of $z \gtrsim 4$ galaxies as a function of both redshift and UV luminosity down to very faint limits.

At $z \gtrsim 4$, we have also combined our faint SEDs with the $L>L^*$ stacks presented in Lee et al. (2011). This allows us to examine the colors of $z \gtrsim 4$ sources over an unprecedentedly large range of luminosities. Our main findings are as follows.

1. At $z \gtrsim 4$, the $(U_n - V_n)$ rest-frame color (interpolated from the data using a fitted SED from a set of templates) is $(U_n - V_n)\sim 0.4$ mag. There is a shallow but systematic trend of redder colors for brighter UV luminosities over the range $-23 \lesssim M_{1500,AB} \lesssim -17.5$ (Figure 6, Section 5.1). A linear fit to this trend is
   $$(U_n - V_n)_{\text{rest}} = -0.04 \times M_{1500} - 0.38$$

2. The SEDs of star-forming galaxies at $4 \lesssim z \lesssim 7$ are remarkably similar at all luminosities from $-23 \lesssim M_{1500,AB} \lesssim -17.5$, showing fairly flat rest-frame UV and rest-frame optical colors (Section 5.2 and Figures 7, 8 and 10). At a given redshift, there are weak indications of a subtle trend for redder colors for brighter sources. A simple fit to the data at all redshifts results in a $(U_n - V_n)_{\text{rest}} = -0.03 \times M_{1500} - 0.16$ relation (Figure 10, center, Section 5.2).

3. The UV-to-optical color, as measured by the observed $J_{125} - [3.6]$ color, remains fairly constant with redshift, despite the Spitzer/IRAC bands probing different wavelength regions of the SED (Figure 10, top). This suggests that the optical colors are fairly flat at all redshifts $z \sim 4-7$, although there is significant scatter, which can be caused by optical emission-line contamination to the IRAC filters.

4. The $z \gtrsim 5$ SEDs show consistently blue $[3.6] - [4.5] \sim -0.3$ colors. This is hard to reproduce with models that only include stellar continuum, although it is still formally possible given the current uncertainties. Nonetheless, it can be more naturally explained with moderate flux contributions from optical nebular emission lines in the two IRAC bands (of the order of 0.2–0.3 mag). Including such a contribution from emission lines leads to lower best-fit ages and lower stellar masses than previously estimated for galaxies at these redshifts by about a factor of two for both ages and masses. We caution though that detailed assessments are not yet possible with the current data.
The dashed line is a best fit to all the points (all redshifts simultaneously): UV-to-optical colors with increasing luminosity is observed at all redshifts. At $z \sim 7$, in particular, no age dependence on luminosity seems to be required. vs. UV-luminosity relation), the dust-corrected colors of all SEDs seem to show following a simple prescription based on the UV slopes and the local Meurer colors (Section 5.1). The best-fit trend to all the data is shown by the brown dashed line (slope $\beta \equiv -0.170 \times M_{1500}$). This is different from Figure 6 in Figure 6, this trend is in very good agreement with the trend exhibited by the full sample. Bottom: the $(U_n - V_n)_\text{rest}$ colors after being corrected by dust extinction following a simple prescription based on the UV slopes and the local Meurer et al. (1999) relation. Except for the $z \sim 7$ SEDs (which have a much steeper $\beta$ vs. UV-luminosity relation), the dust-corrected colors of all SEDs seem to show a flat trend (although the scatter is large). This indicates that a change in dust in the 4 $\lesssim z \lesssim 7$ redshift range. The stacked SEDs allow us to study the UV-to-optical colors of high-z star-forming galaxies down to very low luminosities. We find a mild trend to bluer color at fainter luminosities. Interestingly, these stacked SEDs also show a remarkable similarity across redshift in a period of considerable stellar mass growth from about 0.8 Gyr to 1.5 Gyr, suggesting a smooth self-similar mode of evolution from $z \sim 7$ to $z \sim 4$.

We acknowledge support from NASA grants HST-GO-11563 and HST-GO-11144. We also acknowledge funding from ERC grant HIGHZ No. 227749. This work is based on observations made with the Hubble Space Telescope and with the Spitzer Space Telescope. Support for this work was provided by NASA through an award issued by JPL/Caltech. P.O. acknowledges support 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.

APPENDIX

DEPTH OF THE STACKS

To quantify the rest-frame optical properties of very faint high-redshift galaxies, we stack the IRAC images for a large number of faint galaxies, after removing the flux contribution from nearby neighbors (Section 4). While we do not expect large systematics in the photometry of individual sources from our de-blending procedure to remove neighboring sources, it is possible that modest systematics could arise when stacking a large number of sources. The role of this appendix is to determine how large such systematics might be. Fortunately, we are able to demonstrate that the fluxes that we measure in our de-blending procedure to remove neighboring sources, it is possible that modest systematics could arise when stacking a large number of sources. The role of this appendix is to determine how large such systematics might be. Fortunately, we are able to demonstrate that the fluxes that we measure in our de-blending procedure to remove neighboring sources, it is possible that modest systematics could arise when stacking a large number of sources.

We started by selecting 200 empty areas in the HUDF field. We select them based on the segmentation maps derived from the high-resolution HST images. We processed each of the selected empty areas as if they were the position of one of the real sources in our catalog, i.e., we ran them through the de-blending code to remove the flux from nearby unassociated sources. To determine the rms for a given $N_{\text{stack}}$ number of stacked stamps, we randomly draw (with replacement) $N_{\text{stack}}$ stamps from the 200 empty areas, median combine them, and perform aperture photometry on the median stack in the same fashion as for our...
real stacks. We repeat the drawing 300 times and measure the rms of these 300 trials. This determination of the noise only considers the sky background noise. This is in fact the dominant source of noise in the image; the contribution of shot noise from the faint sources considered here is less than 2%. As can be seen in the upper panel of Figure 11, the limiting flux decreases as $\propto 1/\sqrt{N_{\text{stacked}}}$, as expected for background-limited noise. In the case of the stacks that contain real sources, the rms is larger than the image noise because it is also affected by the intrinsic variation in the distribution of fluxes that are being stacked.

As can be seen in the lower panel, the flux of the median stacks is very close to the expected zero flux.

REFERENCES

Anders, P., & Fritz–v. Alvensleben, U. 2003, A&A, 401, 1063
Atek, H., Siana, B., Scarlata, C., et al. 2011, ApJ, 743, 121
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Bouwens, R. J., Illingworth, G. D., Franx, M., & Ford, H. 2007, ApJ, 670, 928
Bouwens, R. J., Illingworth, G. D., Franx, M., et al. 2009, ApJ, 705, 936
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2010, ApJ, 709, L133
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2011a, arXiv:1109.0994
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2011b, ApJ, 737, 90
Bradley, L. D., Bouwens, R. J., Ford, H. C., et al. 2008, ApJ, 678, 647
Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Bunker, A. J., Wilkins, S., Ellis, R. S., et al. 2010, MNRAS, 409, 855
Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682
Dunlop, J. S., McLure, R. J., Robertson, B. E., et al. 2012, MNRAS, 420, 901
Ellis, R., Santos, M. R., Kneib, J.-P., & Kuijken, K. 2001, ApJ, 560, L119
Eyles, L. P., Bunker, A. J., Ellis, R. S., et al. 2007, MNRAS, 374, 910
Eyles, L. P., Bunker, A. J., Stanway, E. R., et al. 2005, MNRAS, 364, 443
Finkelstein, S. L., Papovich, C., Giavalisco, M., et al. 2010, ApJ, 719, 1250
Finkelstein, S. L., Papovich, C., Salmon, B., et al. 2011, arXiv:1110.3785
Finlator, K., Oppenheimer, B. D., & Davé, R. 2011, MNRAS, 410, 1703
Giavalisco, M., Ferguson, H. C., Koekemoer, A. M., et al. 2004, ApJ, 600, L93
González, V., Labbé, I., Bouwens, R. J., et al. 2010, ApJ, 713, 115
González, V., Labbé, I., Bouwens, R. J., et al. 2011, ApJ, 735, L34
Khochfar, S., & Silk, J. 2011, MNRAS, 410, L42
Kriek, M., van Dokkum, P. G., Labbé, I., et al. 2009, ApJ, 700, 221
Labbé, I., Bouwens, R. J., Illingworth, G. D., & Franx, M. 2006, ApJ, 649, L67
Labbé, I., González, V., Bouwens, R. J., et al. 2010a, ApJ, 716, L103
Labbé, I., González, V., Bouwens, R. J., et al. 2010b, ApJ, 708, L26
Laporte, N., Pelló, R., Schaerer, D., et al. 2011, A&A, 531, A74
Lee, K.-S., Dey, A., Reddy, N., et al. 2011, ApJ, 733, 99
McLure, R. J., Dunlop, J. S., Cirasuolo, M., et al. 2010, MNRAS, 403, 960
McLure, R. J., Dunlop, J. S., de Ravel, L., et al. 2011, MNRAS, 418, 2074
Meurer, G. R., Heckman, T. M., & Calzetti, D. 1999, ApJ, 521, 64
Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007, ApJ, 660, L43
Oesch, P. A., Bouwens, R. J., Illingworth, G. D., et al. 2010, ApJ, 709, L16
Oesch, P. A., Bouwens, R. J., Illingworth, G. D., et al. 2012, ApJ, 745, 110
Oke, J. B., & Gunn, J. E. 1983, ApJ, 266, 713
Richard, J., Kneib, J.-P., Ebeling, H., et al. 2011, MNRAS, 414, L31
Richard, J., Stark, D. P., Ellis, R. S., et al. 2008, ApJ, 685, 705
Rudnick, G., Rix, H.-W., Franx, M., et al. 2003, ApJ, 599, 847
Salpeter, E. E. 1955, ApJ, 121, 161
Schaerer, D., & de Barros, S. 2009, A&A, 502, 423
Schaerer, D., & de Barros, S. 2010, A&A, 515, A73
Shim, H., Chary, R.-R., Dickinson, M., et al. 2011, ApJ, 738, 69
Stark, D. P., Ellis, R. S., Bunker, A., et al. 2009, ApJ, 697, 1493
Taylor, E. N., Frantz, M., van Dokkum, P. G., et al. 2009, ApJS, 183, 295
Weinmann, S. M., Neistein, E., & Dekel, A. 2011, MNRAS, 417, 2737
Wilkins, S. M., Bunker, A. J., Stanway, E., Lorenzoni, S., & Caruana, J. 2011, MNRAS, 417, 717
Windhorst, R. A., Cohen, S. H., Hathi, N. P., et al. 2011, ApJS, 193, 27
Yan, H., Dickinson, M., Giavalisco, M., et al. 2006, ApJ, 651, 24
Yan, H., Dickinson, M., Stern, D., et al. 2005, ApJ, 634, 109
Yan, H., Windhorst, R. A., Hathi, N. P., et al. 2010, Res. Astron. Astrophys., 10, 867