

GOODS-*HERSCHEL*: STAR FORMATION, DUST ATTENUATION, AND THE FIR–RADIO CORRELATION ON THE MAIN SEQUENCE OF STAR-FORMING GALAXIES UP TO $z \simeq 4^*$

M. PANNELLA^{1,2,25}, D. ELBAZ¹, E. DADDI¹, M. DICKINSON³, H. S. HWANG^{1,4}, C. SCHREIBER¹, V. STRAZZULLO^{1,5}, H. AUSSEL¹, M. BETHERMIN^{1,6}, V. BUAT⁷, V. CHARMANDARIS^{8,9}, A. CIBINEL^{1,10}, S. JUNEAU¹, R. J. IVISON^{6,11}, D. LE BORGNE^{2,12}, E. LE FLOC’H¹, R. LEITON^{1,13}, L. LIN¹⁴, G. MAGDIS^{8,15}, G. E. MORRISON^{16,17}, J. MULLANEY^{1,18}, M. ONODERA¹⁹, A. RENZINI²⁰, S. SALIM²¹, M. T. SARGENT^{1,10}, D. SCOTT²², X. SHU^{1,23}, AND T. WANG^{1,24}

¹ Laboratoire AIM-Paris-Saclay, CEA/DSM/Irfu—CNRS—Université Paris Diderot, CEA-Saclay, F-91191 Gif-sur-Yvette, France

² Institut d’Astrophysique de Paris, UMR 7095, CNRS, 98bis boulevard Arago, F-75005 Paris, France

³ National Optical Astronomy Observatory, 950 North Cherry Avenue, Tucson, AZ 85719, USA

⁴ School of Physics, Korea Institute for Advanced Study, 85 Hoegiro, Dongdaemun-Gu, 130-722 Seoul, Korea

⁵ Ludwig-Maximilians-Universität, Department of Physics, Scheinerstr.1, D-81679 München, Germany

⁶ European Southern Observatory, Karl-Schwarzschild-Strasse 2, D-85748 Garching, Germany

⁷ Aix-Marseille Université, CNRS, LAM (Laboratoire d’Astrophysique de Marseille) UMR7326, F-13388, Marseille, France

⁸ Institute for Astronomy, Astrophysics, Space Applications & Remote Sensing, National Observatory of Athens, 15236, Penteli, Greece

⁹ Department of Physics, University of Crete, 71003, Heraklion, Greece

¹⁰ Astronomy Centre, Dept. of Physics & Astronomy, University of Sussex, Brighton BN1 9QH, UK

¹¹ Institute for Astronomy, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ, UK

¹² Sorbonne Universités, UPMC Univ Paris 06, UMR 7095, Institut d’Astrophysique de Paris, F-75005, Paris, France

¹³ Astronomy Department, Universidad de Concepción, Casilla 160 C, Concepción, Chile

¹⁴ Institute of Astronomy & Astrophysics, Academia Sinica, Taipei 106, Taiwan (R.O.C.)

¹⁵ Department of Physics, University of Oxford, Keble Road, Oxford OX1 3RH, UK

¹⁶ Institute for Astronomy, University of Hawaii, Honolulu, Hawaii, HI-96822, USA

¹⁷ Canada–France–Hawaii Telescope, Kamuela, Hawaii, HI-96743, USA

¹⁸ Department of Physics & Astronomy, University of Sheffield, Sheffield, S3 7RH, UK

¹⁹ Institute for Astronomy, ETH Zürich, Wolfgang-Pauli-strasse 27, 8093 Zürich, Switzerland

²⁰ INAF-Osservatorio Astronomico di Padova, Vicolo dell’Osservatorio 5, I-35122 Padova, Italy

²¹ Indiana University, Department of Astronomy, Swain Hall West 319, Bloomington, IN 47405-7105, USA

²² Department of Physics and Astronomy, University of British Columbia, Vancouver, BC V6T 1Z1, Canada

²³ Department of Physics, Anhui Normal University, Wuhu, Anhui, 241000, China

²⁴ School of Astronomy and Space Sciences, Nanjing University, Nanjing, 210093, China

²⁵ Max-Planck-Institut für Extraterrestrische Physik, Giessenbachstrasse 1, D-85748 Garching, Germany

Received 2014 July 10; accepted 2015 May 12; published 2015 July 9

ABSTRACT

We use deep panchromatic data sets in the GOODS-N field, from *GALEX* to the deepest *Herschel* far-infrared (FIR) and VLA radio continuum imaging, to explore the evolution of star-formation activity and dust attenuation properties of star-forming galaxies to $z \simeq 4$, using mass-complete samples. Our main results can be summarized as follows: (i) the slope of the star-formation rate– M_* correlation is consistent with being constant $\simeq 0.8$ up to $z \simeq 1.5$, while its normalization keeps increasing with redshift; (ii) for the first time we are able to explore the FIR–radio correlation for a mass-selected sample of star-forming galaxies: the correlation does not evolve up to $z \simeq 4$; (iii) we confirm that galaxy stellar mass is a robust proxy for UV dust attenuation in star-forming galaxies, with more massive galaxies being more dust attenuated. Strikingly, we find that this attenuation relation evolves very weakly with redshift, with the amount of dust attenuation increasing by less than 0.3 mag over the redshift range $[0.5–4]$ for a fixed stellar mass; (iv) the correlation between dust attenuation and the UV spectral slope evolves with redshift, with the median UV slope becoming bluer with redshift. By $z \simeq 3$, typical UV slopes are inconsistent, given the measured dust attenuations, with the predictions of commonly used empirical laws. (v) Finally, building on existing results, we show that gas reddening is marginally larger (by a factor of around 1.3) than the stellar reddening at all redshifts probed. Our results support a scenario where the ISM conditions of typical star-forming galaxies evolve with redshift, such that at $z \geq 1.5$ Main Sequence galaxies have ISM conditions moving closer to those of local starbursts.

Key words: dust, extinction – galaxies: evolution – galaxies: formation – galaxies: ISM – surveys – radio continuum: galaxies

1. INTRODUCTION

Our understanding of galaxy formation and evolution has made substantial progress in recent years. Having reached a robust measurement, at least up to $z \simeq 4$, of the stellar mass growth and star-formation rate (SFR) density evolution over cosmic time (e.g., Dickinson et al. 2003; Drory et al. 2005; Fontana et al. 2006; Hopkins & Beacom 2006; Pannella et al. 2006, 2009b; Pérez-González et al. 2008; Bouwens et al. 2009; Marchesini et al. 2009; Ilbert et al. 2010, 2013; Karim

* Based on observations collected at the *Herschel*, *Spitzer*, Keck, NRAO-VLA, Subaru, KPNO, and CFHT observatories. *Herschel* is an European Space Agency Cornerstone Mission with science instruments provided by European-led Principal Investigator consortia and with significant participation by NASA. The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc.

et al. 2011; Burgarella et al. 2013; Muzzin et al. 2013; Madau & Dickinson 2014), we must now try to understand which main processes drive the star-formation histories (SFHs) in galaxies and, more specifically, the timescales of the mass growth in star-forming galaxies and the main reasons for the downsizing pattern observed in the passive galaxy population (e.g., Thomas et al. 2005; Cimatti et al. 2006; Pannella et al. 2009a; Renzini 2009; Peng et al. 2010; Elbaz et al. 2011).

A preferred tool for investigating this topic is to study the tight correlation between the galaxy SFR and stellar mass (M_*) content, which is present at all explored redshifts, and also known as the “main sequence” (MS) of star-forming galaxies (e.g., Brinchmann et al. 2004; Daddi et al. 2007; Elbaz et al. 2007, 2011; Noeske et al. 2007; Salim et al. 2007; Pannella et al. 2009a; Magdis et al. 2010; Salimi et al. 2012; Whitaker et al. 2012). The slope and scatter of this correlation, the evolution of its normalization with cosmic time, and also the detailed dissection of its demographics—i.e., the exact percentages of objects that live above it (starburst galaxies, SB) or below it (quiescent galaxies)—contain crucial, and still poorly known, pieces of the galaxy evolution puzzle (e.g., Karim et al. 2011; Rodighiero et al. 2011; Wuyts et al. 2011; Sargent et al. 2012).

While photometric redshift and spectral energy distribution (SED)-fitting techniques have become common and are robust ways to estimate galaxy distances and stellar masses, we are still heading down an unpaved way toward obtaining accurate SFR measurements for cosmologically relevant galaxy samples. Substantial progress has been made thanks to tracers of star formation such as the mid-infrared *Spitzer*, radio continuum (VLA), and, more recently, far-infrared (FIR) *Herschel* surveys, which are not subject to dust attenuation corrections. Due to sensitivity limits, however, the bulk of the star-forming galaxy population is not detected in these kinds of data at redshift greater than one (i.e., when most of the stellar mass growth in the universe took place). The only way to populate and study the $SFR-M_*$ plane with individual detections at $z > 1$ is to use other, dust attenuated tracers of star-formation activity. The most common and easily accessible tracer in most multi-wavelength databases is the UV continuum light emitted by young massive stars. A viable, although more expensive alternative to the rest-frame UV light, is to use line emissions, such as [O II] or H α , which are also good tracers of SFR (e.g., Kennicutt 1998). These can be probed through narrow-band imaging (e.g., Garn et al. 2010; Sobral et al. 2012) or by spectroscopic surveys (e.g., Gilbank et al. 2010; Kashino et al. 2013).

The main issue with estimating star formation from the emerging UV light, as well as from [O II] or H α line emission, is the need to correct, at least in a statistical sense, for the intervening dust attenuation. A common approach is to use the correlation between the slope of the UV spectrum of galaxies and dust attenuation (e.g., Meurer et al. 1999; Calzetti et al. 2000; Daddi et al. 2004; Overzier et al. 2011). The effectiveness of such correlations has been questioned in the local universe (Kong et al. 2004) and is still debated in the literature (e.g., Buat et al. 2005; Reddy et al. 2010; Oteo et al. 2013). Dust attenuation affecting line measurements can in principle be derived by measuring the Balmer decrement (i.e., H α /H β) in galaxy spectra (e.g., Brinchmann et al. 2004; Garn et al. 2010; Zahid et al. 2013b), but this information is rarely available in the distant universe due to the scarcity of

near-infrared (NIR) spectroscopy and, it cannot be determined using narrow-band imaging data. For this reason, line emission studies at high redshift often rely on an indirect way to correct for dust attenuation, such as the comparison with a robust SFR tracer like the galaxy FIR or radio continuum emission (see, e.g., Garn et al. 2010) or a SFR derived through dust-corrected UV/SED-fitting (e.g., Förster Schreiber et al. 2009; Mancini et al. 2011).

By using a sample of *BzK* selected star-forming galaxies, Pannella et al. (2009a) showed that star-forming galaxies at $z \simeq 1.7$ were growing in mass in a self-similar, exponential way. Consequently, the galaxies cannot have lived on the MS for their entire lives, as this would imply a dramatic overproduction of mass compared to the measured evolution of the stellar mass density over cosmic time (see also, e.g., Heinis et al. 2014 for similar conclusions). Pannella et al. (2009a) also found that the correlation between UV slope and UV dust attenuation allowed them to retrieve accurate—again, in a statistical sense—galaxy SFRs, and that dust attenuation is a strong function of the galaxy stellar mass. Finally they were also able to show that the measured emerging UV light is poorly correlated (or perhaps even anti-correlated) with the actual star formation present in a galaxy, as measured by the 1.4 GHz radio continuum luminosity.

The main aim of this paper is to extend the work presented in Pannella et al. (2009a) by focusing on a mass-complete sample of star-forming galaxies up to $z \simeq 4$. Accurate SFR measurements are obtained from the deepest *Herschel* FIR and radio 1.4 GHz continuum imaging available to date, permitting an unbiased derivation of the MS evolution with cosmic time. For the first time, in this paper, we study the radio–FIR correlation from a pure stellar mass selection perspective. Finally, we are able to thoroughly study the dust attenuation properties of star-forming galaxies over a vast range of cosmic time by comparing the derived SFR to the measured UV light.

This paper is organized as follows: in Section 2 we describe the selection and stellar population properties of the star-forming galaxy sample used in this work. In Section 3 we describe the stacking analysis performed on the GOODS-*Herschel* FIR and VLA radio continuum images. In Section 4 we show our results on the MS of star-forming galaxies, their dust attenuation properties, and the FIR–radio correlation. In Section 5 we discuss our results and show some implications of the extra attenuation suffered by line emission compared to the stellar continuum, and the correlation between gas phase metallicity and dust attenuation. We close in Section 6 by summarizing our main results.

Throughout this paper we use AB magnitudes, a Salpeter (1955) initial mass function (IMF), and adopt a Λ CDM cosmology with $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$ and $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$. As a matter of notation, we refer to the rest-frame *GALEX* far-ultraviolet (FUV) bandpass and to the total integrated IR light in the range 8–1000 μm when using the subscripts “UV” and “IR,” respectively.

2. THE GOODS-N STAR-FORMING GALAXY SAMPLE: PHYSICAL PROPERTIES AND MASS-COMPLETENESS

2.1. Multi-wavelength Data and Catalog Production

The galaxy sample we use in this work is drawn from a K_S -band selected multi-wavelength catalog in the GOODS-North field, spanning 19 passbands from *GALEX* near-

ultraviolet (NUV) to IRAC $8\mu\text{m}$ (namely, *GALEX* NUV, KPNO Mosaic U, Subaru Suprime-Cam B-V-IA624-R-I-z-Y, Canada–France–Hawaii Telescope (CFHT) WIRCam *J-H-Ks*, Subaru MOIRCS *J-H-Ks* and *Spitzer* IRAC 3.6-4.5-5.8 and $8\mu\text{m}$). The *GALEX* data are part of the *GALEX* GR6 data release.²⁶ The optical, NIR, and IRAC data are described in Capak et al. (2004), Ouchi et al. (2009), Wang et al. (2010), Kajisawa et al. (2011), L. Lin et al. (2012, 2015, in preparation), M. Dickinson et al. (2015, in preparation), and M. Onodera et al. (2015, in preparation).

We first register all images to a common pixel grid and then use SExtractor (Bertin & Arnouts 1996) in dual image mode to measure photometry. The K_S -band CFHT WIRCAM image has been adopted as the primary detection image because it represents the best compromise, among all available bands, between the need for a robust tracer of galaxy stellar mass and an angular resolution ($\simeq 0''.8$) matching that of most of the other bands, which simplifies catalog assembly and photometry measurements. The whole catalog contains 56,144 objects over the CFHT WIRCAM K_S image field of approximately 900 arcmin^2 and down to an AB magnitude of 24.5 (i.e., the image 5σ limiting magnitude, see Wang et al. 2010 for details).

The images used here have very different resolutions. Rather than convolving all images to the worse resolution, which would result in a significant loss of information, we account for this in the estimate of aperture colors by applying point-spread function (PSF)-matching corrections based on the growth-curve of point-like sources. To limit uncertainties in such corrections we use $2''$ diameter apertures to sample the galaxy SEDs.

Band-merging with IRAC photometry was done with a cross-match of the NUV–optical–NIR catalog to the K_S –IRAC catalog released by Wang et al. (2010). The latter study used a “real-cleaning” procedure to perform PSF modeling in the IRAC images by assuming as priors the positions of sources detected in the same K_S CFHT WIRCAM image we have used. We refer the reader to Wang et al. (2010) for a more detailed description and assessment of the extracted IRAC photometry. Here we want to stress that because we use the same K_S image that was publicly released by Wang and collaborators, our cross-matching procedure becomes a very robust way of associating IRAC fluxes to our K_S -based multi-wavelength catalog.

The whole catalog contains 56,144 objects over the CFHT WIRCAM K_S image field of approximately 900 arcmin^2 and down to an AB magnitude of 24.5 (i.e., the image 5σ limiting magnitude, see Wang et al. 2010 for details).

We classified 2,072 objects as stars, and as such exclude them from the sample according to their SExtractor stellarity index, at bright ($K_S < 20$) mag, and to their position in the BzK diagram, as in Daddi et al. (2004).

The morphological selection of stars is also effective in removing those Type 1 (i.e., optically unobscured) active galactic nuclei (AGNs) whose emission from the active nucleus dominates over the host galaxy in almost all bands. In any case, these objects are extremely difficult to deal with in terms of both the photometric redshift derivation and stellar mass estimate (e.g., Salvato et al. 2009; Cisternas et al. 2011; Bongiorno et al. 2012), so it is preferable to remove them from our sample. However, we do not attempt to remove from the

star-forming galaxy sample those X-ray sources detected in the 2-Ms *Chandra* data (Alexander et al. 2003; Bauer et al. 2004) whose optical-to-NIR emission is dominated by the host galaxy. A number of recent studies (see, e.g., Mullaney et al. 2012; Santini et al. 2012; Juneau et al. 2013; Rosario et al. 2013) have shown that these AGNs are mostly Type II obscured AGN and LINERS, which have similar SFR to normal (i.e., non-AGN) star-forming and passive galaxies, respectively. Thus their inclusion has no net effect on our results. In Section 3 we discuss our approach to minimize the contribution of the AGNs present in our sample to the IR budget.

2.2. Photometric Redshifts

Photometric redshifts were estimated by running EAZY²⁷ (Brammer et al. 2008) on the multi-wavelength catalog. We used EAZY in its standard setup and modeled the observed galaxy SED with a linear combination of the seven standard EAZY galaxy templates (Whitaker et al. 2011) to maximize the likelihood as a function of the galaxy redshift. Following a common procedure for producing accurate photometric redshifts (see, e.g., Capak et al. 2007; Gabasch et al. 2008), we allowed for a photometric offset in each measured band. For each band in our catalog, we computed a photometric zero-point shift by iteratively running the code on a subsample of high fidelity spectroscopic redshifts. For each iteration and for each band, we computed the median of the ratio between the measured flux density in the catalog and the carefully computed model value (i.e., the integral of the best-fit solution template through the theoretical response curve), and applied these median offsets before starting the next iteration. Computing median offsets, instead of mean ones, is an effective way to filter out catastrophic events, such as badly measured SEDs or wrong spectroscopic redshifts determinations. The procedure described here usually converges to a robust solution within a limited number of iterations.

The photometric offsets between observed and modeled data may be attributed to many different effects, including actual zero-point errors, differences of the actual system response curves with respect to adopted ones, and inaccurate PSF-matching corrections, as well as uncertainties and limitations in the template library adopted.

In the following sections, we use a multi-wavelength catalog that has been corrected for these systematic offsets. In all bands these amount to less than 20% of the measured flux so that, while substantially improving the photometric redshift accuracy, applying these flux corrections does not have any major effect on the derived galaxy properties, nor in general on the main results of this study. We will discuss in the relevant sections when notable differences arise from the use of the corrected versus uncorrected photometry.

When comparing to the spectroscopic sample of Barger et al. (2008) and D. Stern et al. (2015, in preparation), we reach a relative (i.e., $\Delta z = (z_{\text{phot}} - z_{\text{spec}})/(1 + z_{\text{spec}})$) accuracy of 3% (see Figure 1), with less than 3% catastrophic outliers (i.e., objects with $\Delta z > 0.2$).

²⁶ Publicly available at <http://galex.stsci.edu/GR6/>.

²⁷ Publicly available at <http://www.github.com/gbrammer/ezay-photoz>.

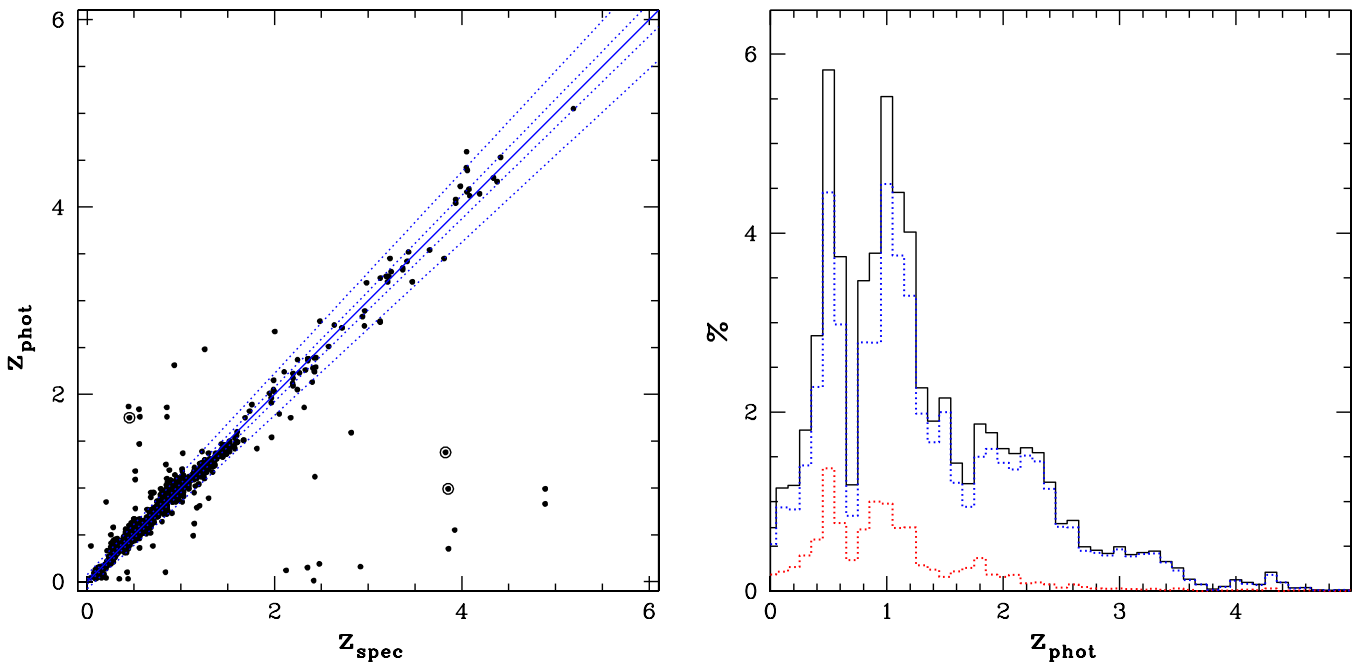


Figure 1. Left: comparison between spectroscopic and photometric redshifts used in this work. The solid line is the bisector, while the dotted lines show the 3% ($\pm 1\sigma$) and 9% ($\pm 3\sigma$) bands. The three circled dots are galaxies for which two different spectroscopic determinations are available. For all three, the second determination agrees with our photometric redshift estimate. Right: redshift distribution of our catalog in the GOODS-N field down to $K_{AB} = 24.5$, the 5σ limiting magnitude. The solid line shows the actual sample of the 18,416 galaxy sample used in this work, (i.e., restricted to the FIR deep imaging area). Dashed blue (red) lines show the subsample of star-forming (quiescent) galaxies. We note that the two main redshift peaks at $z \simeq 0.5$ and $z \simeq 1$, and as well as the underdensity at $z \simeq 0.75$, are real, large-scale structure features because they are also present in the spectroscopic redshift distribution.

2.3. Derivation of Stellar Masses

Stellar masses were determined with *FAST*²⁸ (Kriek et al. 2009) on the U to $4.5\,\mu\text{m}$ PSF-matched aperture photometry, using Bruzual & Charlot (2003) delayed exponentially declining SFHs ($\psi(t) \propto \frac{t}{\tau^2} \exp(-t/\tau)$) with $0.01 < \tau < 10$ Gyr, solar metallicities ($Z_{\odot} = 0.02$), Salpeter IMF, and the Calzetti et al. (2000) reddening law with A_V up to 4 mag.

The choice of a fixed solar metallicity has essentially no impact on the derived stellar masses, because the age-metallicity degeneracy leaves mass-to-light ratios basically unchanged. Moreover, solar metallicities are still a fair assumption for our sample at high redshift given that, due to the increasing mass completeness with redshift, it contains at all cosmic epochs the most massive and most metal-rich systems.

Derived (aperture) masses were extrapolated to “total” masses using the ratio between the total (*FLUX_AUTO*) and aperture flux in the K_s -band detection image. While this approach corrects for the bulk of the flux loss, it is based on one band only and thus neglects any color gradient within the galaxy.

With respect to the choice of the SFH, we note that it has been shown that other forms of SFHs are more appropriate for star-forming galaxies at high redshift (Pannella et al. 2009a; Maraston et al. 2010; Papovich et al. 2011). For what directly concerns this work, rising or constant (possibly truncated) SFHs would change the stellar masses of our star-forming galaxy sample by an amount that is well within the median estimated accuracy of about 0.2 dex (see, e.g., Papovich

et al. 2006, 2011; Buat et al. 2014; de Barros et al. 2014; and the results presented in Appendix C).

As a further accuracy test, we compared stellar mass estimates derived with *FAST* with the SED-fitting code described in Drory et al. (2004, 2009). The results from runs with the different codes and different adopted SFHs show no systematic differences in the stellar mass estimates, except for the highest stellar masses ($M_* \gtrsim 2 \times 10^{11} M_{\odot}$), where the two-component fitting from Drory et al. (2004) predicts masses that are larger by 0.2 dex, with a global scatter of about 0.2 dex. We provide more details on the estimated stellar mass errors and on the comparison between the two codes in Appendix C.

2.4. The Star-forming Galaxy Sample: UVJ Selection, β Slope, and Mass-completeness

The main aim of this study is to characterize the relation between star formation, UV dust attenuation, and the stellar mass content of galaxies over cosmic time. To do so, it is important not to mix galaxies that might still have some residual ongoing star formation (but are substantially quenched), with galaxies that are genuinely star-forming and reddened by dust attenuation. A selection based on direct SFR indicators, such as UV or line emission, inevitably mixes the two populations. Ideally, one would like to remove galaxies with extremely low specific star-formation rates (SSFRs) from the sample, but this is not easily achieved for individual galaxies because it requires accurate dust attenuation corrections or, alternatively, FIR/radio derived star-formation rate estimates. Following established procedures, we excluded quiescent galaxies from the sample by using the $U - V$ versus $V - J$ rest-frame color plot (Wuyts et al. 2007; Williams et al. 2009). This approach builds on two main ingredients: (1) the intrinsic rest-frame $U - V$ color is a good proxy for the

²⁸ Publicly available at <http://astro.berkeley.edu/~mariska/FAST.html>.

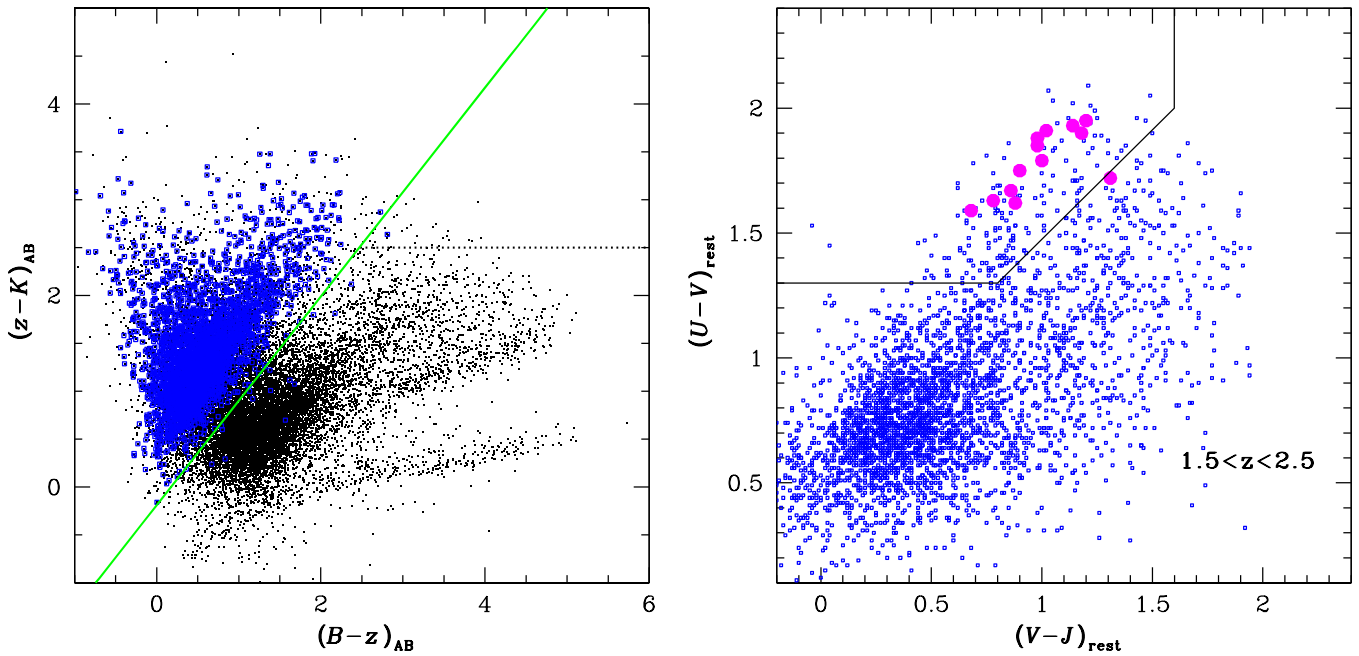


Figure 2. Left: The BzK diagram used for the selection of star-forming galaxies in the redshift range 1.5–2.5 according to Daddi et al. (2004). The solid line divide the regions populated by star-forming galaxies at $z \simeq 2$ ($sBzK$, above the solid line) from the lower redshift objects (below the solid line and below the dotted line), and the passive galaxies at $z \simeq 2$ ($pBzK$, below the solid line and above the dotted line). Black points show all the galaxies in our catalog, whereas open squares represent the UVJ -selected star-forming galaxies in the redshift range 1.5–2.5. Right: UVJ diagram used for the selection of the sample of star-forming galaxies in this work. The solid lines identify the two regions used to divide quiescent (upper-left) from star-forming (lower-right) galaxies. Here we show the galaxies in the redshift range 1.5–2.5, which would have been selected as star-forming $sBzK$ (blue squares) and passive quiescent $pBzK$ (magenta points).

SSFR of a galaxy (e.g., Salim et al. 2005; Pannella et al. 2009b); (2) the color–color plot allows us to break the age–dust degeneracy and hence efficiently split the sample into two classes of galaxy, which are either essentially quiescent or still actively forming stars. We selected star-forming galaxies at all redshifts as

$$(U - V) < 1.3, (V - J) > 1.6, (U - V) < (V - J) \times 0.88 + 0.59.$$

These UVJ selection limits were originally defined by Williams et al. (2009) in order to maximize the difference in SSFRs between regions. However, the rest-frame color distributions might be slightly different than Williams et al. (2009) due to photometric coverage, band selection, and the specific redshift where the analysis is performed (the rest-frame color of quiescent galaxies is becoming redder with decreasing redshift). As a result, different studies have used slightly different dividing lines, sometimes changing with redshift (e.g., Cardamone et al. 2010; Brammer et al. 2011; Whitaker et al. 2011; Muzzin et al. 2013; Strazzullo et al. 2013; Viero et al. 2013). The differences between these studies are always lower than 0.2 mag in color and often comparable with the same accuracy to which the rest-frame colors are known. In order to be conservative and minimize the contamination of quiescent galaxies in our sample, we chose to use the same selection region at all redshifts. We also checked whether slightly changing our assumptions on the selection limits (i.e., applying ± 0.1 mag shifts on the rest-frame colors) would quantitatively change our results.

We checked our selection against the BzK selection of star-forming galaxies at $z \simeq 2$ (Daddi et al. 2004) (see Figure 2),

and the results are in good agreement over the common redshift range. Finally, we tested whether all the results of this paper would remain the same if instead of using the UVJ selection to separate the active from passive/quenched galaxies we used a cut in SSFR, namely selecting as star-forming galaxies all the objects with $\text{SSFR} \geq 10^{-11} \text{ yr}^{-1}$ as derived from the *FAST* SED-fitting output.

Hereafter, we will concentrate on the UVJ -selected sub-sample of 14,483 star-forming galaxies that fall within the *GOODS-Herschel* area, the deepest FIR imaging available in the northern sky (Elbaz et al. 2011).

Following Brammer et al. (2011), rest-frame magnitudes are computed with *EAZY* for all objects in the catalog from the best-fit SED model integrated through the theoretical filters. The β slope of the UV continuum was calculated as

$$\beta = \frac{\log(f_{\lambda_1}/f_{\lambda_2})}{\log(\lambda_1/\lambda_2)} = \frac{0.4(m_2 - m_1)}{\log(\lambda_1/\lambda_2)} - 2, \quad (1)$$

where m_1 and m_2 are the magnitudes at wavelengths λ_1 and λ_2 , respectively (see, e.g., Overzier et al. 2011; Nordon et al. 2013). In this study, we adopt the *GALEX* FUV ($\lambda_c \simeq 1530 \text{ \AA}$) and NUV ($\lambda_c \simeq 2315 \text{ \AA}$) response curves to compute the photometric measure of β , but we also checked that the results obtained would remain largely unchanged by using different rest-frame bands sampling the slope, or by obtaining a direct fit of the form $f_{\lambda} \propto \lambda^{\beta}$ to the best-fit SED in the rest-frame range 1250 and 2600 \AA . In Figure 3 we show the results of this comparison using the best-fit SEDs of *FAST*. The *GALEX*-derived estimate of β provides an accurate description of the UV slope, with a median offset of 0.1 and a scatter of

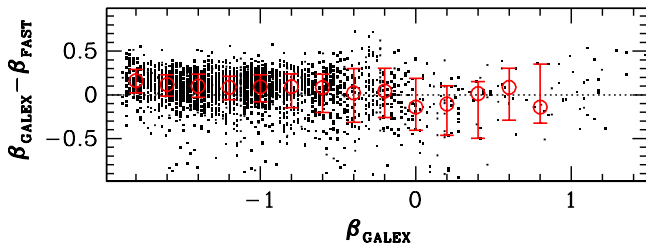


Figure 3. Comparison between the β values derived from the rest-frame *GALEX* bands and the one derived from a fit, in the rest-frame range 1250 Å and 2600 Å, to the best-model SEDs of *FAST*. Red open circles show median values of the difference, and error bars represent 16th and 84th percentiles of the distribution. The β_{GALEX} values are only slightly redder, by around 0.1, than the SED-derived ones with a median uncertainty value of 0.2.

around 0.2. This result is in good agreement with the finding reported in Finkelstein et al. (2012).

An important issue, which is often not given careful consideration in high redshift studies dealing with star-forming galaxies, is the assessment of the sample’s mass completeness. This is because the amount of dust attenuation plays an important role in the selection function: low mass galaxies (and the definition of *low* depends of course on the selection band and detection limit of the specific data set) can only be detected if they suffer from low dust attenuation. In other words, as redshift increases, magnitude-limited samples become more and more biased against objects with lower masses and/or high M_*/L ratios, such as galaxies that are not currently forming stars or are highly dust obscured. Because we are interested in dissecting dust attenuation properties, we want to make sure that this issue has no impact on our results. Following Rodighiero et al. (2010), we estimated our completeness mass by using a stellar population with a 1 Gyr old constant SFH and highly attenuated ($E[B - V] = 0.6$). As a function of redshift, there is a one-to-one correspondence between the observed K_s magnitude and the stellar mass built up by this model. We hence define as the completeness mass at a given redshift the mass corresponding to the completeness magnitude (≈ 23.8 AB mag) of our K_s selected catalog. All galaxies more massive than our completeness mass will have a brighter flux at K_s and hence be detected in our survey; the same is true for galaxies suffering less dust attenuation. Obviously, this definition of mass completeness is dependent on the specific stellar population parameters adopted, and can only apply to galaxies that can be likely represented by such an SFH (i.e., star-forming galaxies). Figure A.1 of Rodighiero et al. (2010) illustrates how the actual mass completeness varies by adopting different ages and attenuations for the stellar population models. However we note that our modeling is very conservative and the quoted mass completenesses (see Table 1) should be regarded as safe and robust, although theoretical, values. As previously stated, our sample, which is composed of a priori selected star-forming galaxies, is also cut in SSFR (at approximately 10^{-11} yr^{-1} , see discussion above), meaning that, even if we are complete down to a certain mass and able to robustly trace the $\text{SFR}-M^*$ made by the bulk of the star-forming galaxy population, we would still be able to detect galaxies with extremely low SSFR if they suffer low dust attenuation. This can be translated in an SFR limit at the completeness mass that we also quote in Table 1. We want to stress here that, as our primary selection is a mass selection, the secondary limits

Table 1
Stellar Mass and Star Formation Rate Completeness vs. Redshift

z	0.7	1	1.3	1.7	2.3	3.3
$\log M_*/M_\odot$	9.0	9.5	10.0	10.3	10.5	10.7
SFR ($M_\odot \text{ yr}^{-1}$)	1	2	10	20	30	90

in SFR have no influence on our results because they apply to galaxies well below the MS locus.

3. THE STACKING ANALYSIS

We study the FIR properties of star-forming galaxies at high redshift using the deepest *Herschel* imaging available to date. Obviously, any analysis based only on FIR detected sources would be an SFR-biased view of the universe at all redshifts. This is best exemplified by the fact that only 1,095 sources are detected at more than 3σ in the deep GOODS-*Herschel* images (see Elbaz et al. 2011), while more than 11,000 star-forming galaxies are present in our sample over the same area. At all redshifts the flux density limit of the *Herschel* images allows us to explore only the highest values of star-formation rates, as opposed to a purely mass-selected sample, where UV-based star-formation rates can reach down to extremely low SFR levels. Up to $z \approx 1$ the FIR data are deep enough to probe the MS down to relatively low stellar masses (i.e., $M_* \geq 10^9 M_\odot$). At higher redshifts, only the most extreme events of star formation and the high mass end of the mass function can be detected in the GOODS-*Herschel* data. In order to reach the FIR flux densities typical of normal MS galaxies, and to gain a comprehensive picture of the physics driving star-forming galaxies at high redshift, it is therefore mandatory to adopt a stacking approach.

The GOODS-*Herschel* images were generated at different pixel scales in order to allow a fair sampling ($\approx \text{FWHM}/5$) of the beam. For each of the sources in our star-forming galaxy sample we produced a cutout—centered on the nearest pixel to the sky position of the source, in the PACS (100, 160 μm) and SPIRE (250, 350, 500 μm) images of 60×60 pixels—which corresponds to an angular scale of about 10 times the image beam. The cutouts were then stacked to create median images in selected bins of redshift and stellar mass. Median stacking is more robust than the mean against the tails of the distribution (mainly due to the relatively few detections, but also to possible photo- z catastrophic outliers), while the rms still goes down by approximately \sqrt{N} and image statistics are well preserved. Total fluxes are retrieved by performing a PSF plus background modeling with the *GALFIT* code (Peng et al. 2002) on the stacked images. In order to calculate realistic errors on the retrieved flux densities we use a bootstrapping approach. In each redshift and mass bin we randomly select (with replacement) $1/N$ of the sample within that bin and re-estimate the median flux density of this subsample. This is performed 1,000 times for each bin and the error on the median value is calculated by taking $\sqrt{1/N}$ times the standard deviation of the results. Running this procedure with $N = 1, 2, 3$, and 4 yields consistent results.

We corrected the estimated *Herschel* flux densities for the bias due to clustering in a statistical way. The clustering is expected to introduce an extra signal in the PSF modeling of the stacking results due to the positive correlation of the faint sources present in the field and the stacked sample. Different

Table 2
Corrections Applied to SPIRE Fluxes

Band	250 μm	350 μm	500 μm
Scaling factor	0.8	0.7	0.6

formalisms have been proposed in the last years in order to estimate the impact of the source clustering on the stacking results (see, e.g., Béthermin et al. 2010, 2012b; Chary & Pope 2010; Kurczynski & Gawiser 2010; Bourne et al. 2011; Viero et al. 2013). A natural expectation is that the bias is directly proportional to the beam size of the image (i.e., the worse is the image angular resolution, the higher the expected bias), while the dependence on stellar mass and redshift is less clear.

The novel technique adopted to estimate the flux boosting due to clustering will be thoroughly detailed in Schreiber et al. (2015). Here for the sake of clarity we will briefly outline the main steps involved in the procedure. A simulated map in each band is created using the sky position, redshift, and stellar mass for all galaxies in our sample. The analytical model described in Sargent et al. (2012) is then used to associate an SFR, and hence an IR luminosity, to every galaxy. Finally *Herschel* flux densities are estimated by adopting a library of FIR templates and injected, with PSFs, in a pure instrumental noise map. The stacking procedure is then run on the simulated map in the same bins of mass and redshift as for the real data. The correction factor is derived from the comparison between the median input flux and output stacked result.

We report in Table 2 the median corrections adopted for the SPIRE bands. These values are in nice agreement with values already reported in the literature for similar experiments in the same bands (Viero et al. 2013) or in similar angular resolution data (Bourne et al. 2011). Within the uncertainties, we are not able to see any dependence of flux boosting on stellar mass and redshift, and hence adopt a fixed scaling in all bins. No flux correction has been applied to the PACS photometry. Here the estimated boosting is less than 10% and is approximately counterbalancing the loss of flux due to the high-pass filtering in the data reduction process (Popesso et al. 2012).

After de-boosting the SPIRE flux densities, we derived the total IR luminosity by fitting the observed IR SED from 100 to 500 μm with the Chary & Elbaz (2001) library models. The derived IR luminosities are deemed more accurate than methods based on bolometric correction extrapolation of single IR band measurements because they use the more extended photometric information, which can also account for the slight evolution in redshift of the median IR SED (Béthermin et al. 2012a; Magdis et al. 2012). In Appendix A we compare our results to IR luminosities estimated by using bolometric correction techniques and test their fidelity and relative accuracy. Finally, we want to remind the reader that we chose to retain Type II AGNs in our sample because they do not prevent the accurate determination of photometric redshifts and stellar masses of their host galaxies. Still, the AGN itself could dominate the mid-IR total budget. In order to minimize the impact of AGN emission to the IR stacked SEDs and on our results, we discarded all the points below the 30 μm rest-frame from the IR SED-fitting. Indeed, as shown by, for example, Mullaney et al. (2011), above this wavelength the contribution of the AGN to the measured SED becomes negligible.

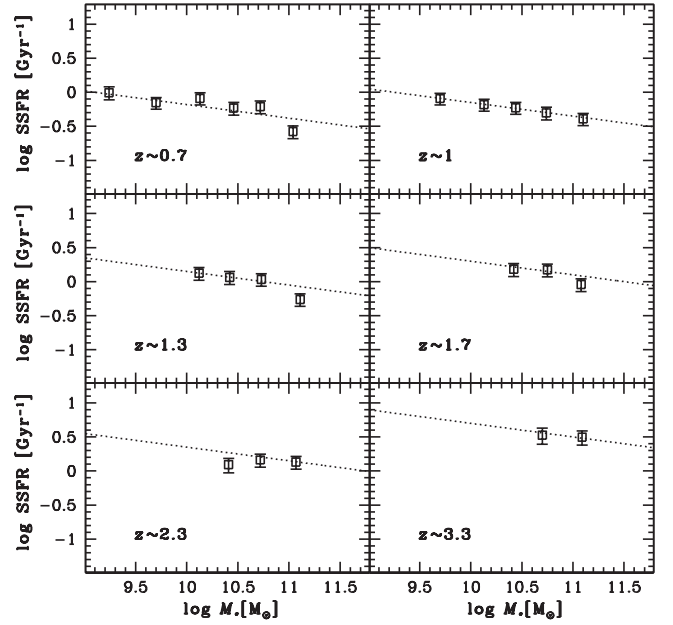


Figure 4. Evolution of the specific star-formation rate (SSFR) vs. stellar mass (M_*) at different redshifts from our FIR stacking analysis. We find a slope of 0.8 up to $z \simeq 1.5$. For higher redshifts, we are simply assuming that the slope is the same and allow for an evolution in normalization.

To compute star-formation rates from IR and UV luminosities, we adopt the conversions of Kennicutt (1998) and Daddi et al. (2004), respectively:

$$\frac{\text{SFR}_{\text{IR}}}{M_{\odot} \text{ yr}^{-1}} = 1.71 \times 10^{-10} \frac{L_{\text{IR}}}{L_{\odot}}; \quad (2)$$

$$\frac{\text{SFR}_{\text{UV}}}{M_{\odot} \text{ yr}^{-1}} = 1.13 \times 10^{-28} \frac{L_{\text{UV}}}{\text{erg s}^{-1} \text{ Hz}^{-1}}. \quad (3)$$

The detection of UV emission from star-forming galaxies clearly indicates that at least part of the UV radiation is not absorbed by dust (and reprocessed into the infrared). Therefore, following, for example, Bell (2003), to derive the total SFRs we use $\text{SFR}_{\text{total}} = \text{SFR}_{\text{IR}} + \text{SFR}_{\text{UV}}$, where SFR_{UV} is computed using the observed UV luminosity in Equation (3).

4. RESULTS: THE PROPERTIES OF STAR-FORMING GALAXIES UP TO $z \simeq 4$

In the next sections we will present our results on the star formation and dust attenuation properties of galaxies in the GOODS-N field up to $z \simeq 4$.

4.1. The SSFR- M_* Correlation: Evolution of Slope and Normalization

The main goal of this work is to define, in a consistent way, the locus of the MS of star-forming galaxies over a significant range of cosmic time. Many studies have been published on this topic in recent years (e.g., Rodighiero et al. 2010; Karim et al. 2011). We use data probing the full FIR SED (using all *Herschel* bands), so as not to rely on the L_{IR} derived from the bolometric correction of a single band. In Figure 4 we show the MS of star-forming galaxies (i.e., the correlation between the galaxy stellar mass and SSFR) in different redshift bins. Only results from mass-complete samples are shown. We fit a linear trend in the first three redshift bins and find a slope consistent

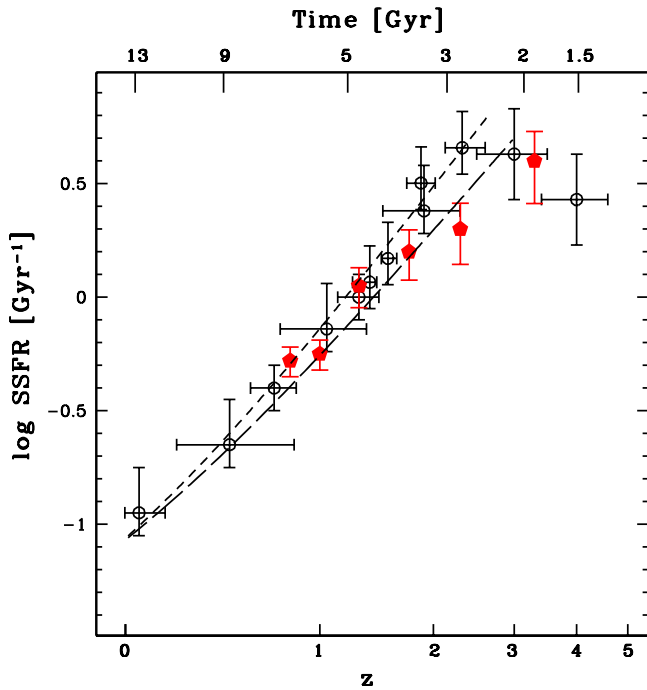


Figure 5. FIR-derived SSFRs from this work (red pentagons) for star-forming galaxies with $M_* \simeq 3 \times 10^{10} M_\odot$ as a function of redshift. A sample of previously published estimates from Brinchmann et al. (2004), Daddi et al. (2007, 2009), Elbaz et al. (2007), Noeske et al. (2007), Pannella et al. (2009a), and Magdis et al. (2010) is plotted as empty circles. The dashed and long-dashed lines show the SSFR as a function of redshift according to the relations published in Pannella et al. (2009a) and Elbaz et al. (2011).

with -0.2 ± 0.08 , corresponding to a slope of 0.8 ± 0.08 for the proper $\log M_*$ – $\log \text{SFR}$ correlation, up to redshift $z \simeq 1.5$. This value of the slope is consistent with a number of previous studies (e.g., Santini et al. 2009; Rodighiero et al. 2011, 2014; Lin et al. 2012; Heinis et al. 2014; Renzini & Peng 2015). We then assume the same slope for the higher redshift bins, and fit only the normalization of the MS up to $z \simeq 4$.

In Figure 5 we show the evolution of the MS normalization for a stellar mass of $\log(M_*/M_\odot) = 10.5$ as obtained from this work, together with some previously published results (Brinchmann et al. 2004; Daddi et al. 2007, 2009; Elbaz et al. 2007; Noeske et al. 2007; Pannella et al. 2009a; Magdis et al. 2010) and the analytical descriptions of the MS evolution from Pannella et al. (2009a) and Elbaz et al. (2011).²⁹

Our median SSFR values are slightly lower than previous studies at $z \geq 1.5$. In principle this could be explained by the fact that we are dealing with a mass-complete sample, while previous studies were biased toward brighter ($K_S < 23$ in Pannella et al. 2009a) or star-formation selected ($24 \mu\text{m}$ detected sources in Elbaz et al. 2011) samples that almost by construction are bound to overshoot the median value of the whole star-forming population. At the same time we cannot exclude the possibility that at least part of the offset could be due to some misclassified galaxies that entered the *UVJ*-selected star-forming sample.

²⁹ Here we plot the original derivation of Elbaz et al. (2011), which is indeed a good match to our stacking results. However, we note that after taking into account a possible stellar mass systematic offset of 0.19 dex due to the different stellar population models used in the Elbaz et al. (2011) study and in the one presented here (see also, e.g., Bell et al. 2012), the relation of Elbaz et al. (2011) would shift upward and almost perfectly overlap the Pannella et al. (2009a) result. We will investigate this issue in more detail in a future work.

Finally, it is worth noticing how the SSFR tends to keep increasing up to the highest redshift without any evidence of the possible flattening at $z \geq 2$ obtained in previous studies (e.g., Daddi et al. 2009; Stark et al. 2009; González et al. 2010), which were based on SFR estimates derived from dust-attenuation-corrected UV luminosities. We will discuss this in more detail in the following sections, but based on our results this might be related to a change in the dust and/or stellar population properties of star-forming galaxies with cosmic time.

4.2. An Unbiased Analysis of the Radio–FIR Correlation: the Contribution of Old Stellar Populations to the FIR

Taking advantage of the deep 1.4 GHz map available in the GOODS-N field (Morrison et al. 2010), we perform a stacking analysis of the radio continuum in the same bins of mass and redshift as done in the FIR *Herschel* bands in order to study the evolution of the radio–FIR correlation for a uniformly selected, mass-complete sample of star-forming galaxies over cosmic time. Sargent et al. (2010) discussed the impact of sample selection, specifically the FIR versus radio selection, in explaining the controversial results obtained in literature studies on the evolution of the FIR–radio correlation. Based on *Spitzer* $24 \mu\text{m}$ -derived SFR estimates, they showed how the correlation was not evolving up to redshift 1 and possibly even at higher redshift, although this was plagued by the uncertain extrapolation from the measured $24 \mu\text{m}$ to total IR luminosity at $z \geq 1.5$. Here we have the advantage of being able to stack mass-complete samples.

Some recent *Herschel*-FIR based studies have claimed a redshift evolution of the correlation (e.g., Ivison et al. 2010; Casey et al. 2012; Magnelli et al. 2012, but see also, e.g., Strazzullo et al. 2010; Bourne et al. 2011; Barger et al. 2012, 2014; Del Moro et al. 2013 for different conclusions), with increasing radio luminosities for the same IR luminosity at higher redshift.

On the other hand, simple theoretical arguments predict the opposite trend (e.g., Condon 1992; Carilli et al. 2008), because of the increased inverse Compton cooling of the relativistic electron population due to the scattering off of the increasing cosmic microwave background at high redshift, $U_{\text{CMB}} \propto (1 + z)^4$.

We derive 1.4 GHz monochromatic luminosities from the measured radio flux densities assuming a synchrotron spectral index, α , of -0.8 and the median redshift of each stacked subsample, as

$$L_{1.4} = 1.19 \times 10^{14} \text{DM}^2 \times S_{1.4} \times (1 + z_{\text{med}})^{-(\alpha+1)} \left(\text{WHz}^{-1} \right), \quad (4)$$

where DM is the distance in Mpc at the median redshift z_{med} of the stacked subsample, and $S_{1.4}$ is the measured flux density in the stack in units of μJy . We notice that our assumption of a nonevolving spectral index is justified by most observational studies dealing with statistical samples of star-forming galaxies (see e.g., Ibar et al. 2009; Bourne et al. 2011). We adopt the Yun et al. (2001) relation to convert such luminosities in

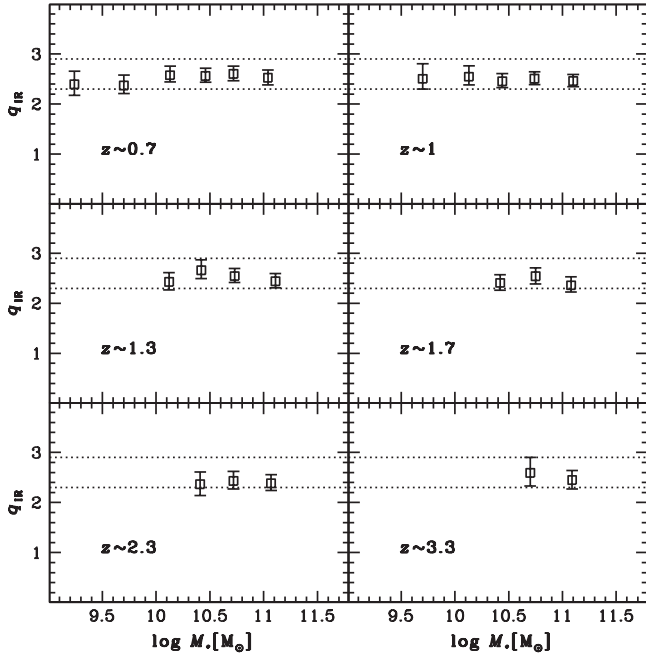


Figure 6. The luminosity ratio $q_{\text{IR}} = \log(L_{\text{IR}}/3.75 \times 10^{12}) - \log(L_{1.4})$ as a function of redshift and for the different mass bins from stacking of mass-complete star-forming samples. The dashed lines encompass the $\pm 1\sigma$ scatter of the local FIR–radio correlation. We do not see any evolution of the median q_{IR} up to the highest redshifts probed.

star-formation rates:

$$\text{SFR}_{1.4} = 5.9 \times 10^{-22} L_{1.4} (M_{\odot} \text{ yr}^{-1}). \quad (5)$$

In order to study the evolution of the correlation, we compute, following Helou et al. (1985), the luminosity ratio:

$$q_{\text{IR}} = \log(L_{\text{IR}}/(3.75 \times 10^{12} \text{ W})) - \log(L_{1.4}/\text{WHz}^{-1}) \quad (6)$$

for all redshift and mass bins where we have performed our stacking analysis. We show the results in Figure 6 together with the $\pm 1\sigma$ confidence locus of the local universe q_{IR} value (e.g., Yun et al. 2001; Bell 2003). We find striking agreement with the local correlation at all redshifts and masses explored.

We stress that our result is the first so far obtained for a mass-complete uniformly selected sample of star-forming galaxies and hence is expected to be significantly more robust against selection bias, as compared to all previous analyses.

This finding has some non-trivial implications. First, it shows that a FIR–radio correlation is well defined up to $z \simeq 4$ with basically the same q_{IR} value measured in the local universe, and thus radio continuum data are an ideal tool for estimating dust-unbiased SFRs at high redshift.

Second, since radio luminosity is only contributed by young stellar populations, our result suggests that the IR emission of star-forming galaxies in the redshift range $0.5 < z < 4$ can only have, in a statistical sense, a marginal contribution from old stellar populations, and likely smaller than the 20% estimated in the local universe (see, e.g., Law et al. 2011).

4.3. UV Dust Attenuation Over Cosmic Time

In order to study the redshift evolution of the UV dust attenuation, we define as an effective measurement of dust

attenuation the quantity:

$$A_{\text{UV}} = 2.5 \times \log(\text{SFR}_{\text{IR}}/\text{SFR}_{\text{UV}}^{\text{obs}} + 1), \quad (7)$$

(i.e., the magnitude correction to the measured UV emission). This means that, by construction, we obtain

$$\text{SFR}_{\text{UV}}^{\text{corr}} = \text{SFR}_{\text{total}} = \text{SFR}_{\text{IR}} + \text{SFR}_{\text{UV}}^{\text{obs}}.$$

For each subsample stacked in each redshift and mass bin, we estimate the median FUV emission and UV spectral slope (see Equation (1)).

Assuming the Calzetti et al. (2000) empirical correlation between UV dust attenuation and slope β :

$$A_{\text{UV}}^{\beta} = 4.85 + 2.31 \times \beta, \quad (8)$$

we then define the quantity

$$\text{SFR}_{\text{UV}}^{\beta} = \text{SFR}_{\text{total}}^{\beta} = \text{SFR}_{\text{UV}}^{\text{obs}} \times 10^{A_{\text{UV}}^{\beta}/2.5} \quad (9)$$

which provides an estimate of the total SFR by correcting the observed UV luminosity assuming the measured β slope and the Calzetti et al. (2000) law, which is a standard prescription adopted for high redshift studies.

In the following we discuss how the measured UV dust attenuation, A_{UV} , depends on galaxy properties, such as the stellar mass and the UV spectral slope over the redshift range 0.5–4.

4.3.1. A_{UV} Versus M_{*}

In the top-left panel of Figure 7 we show the correlation between dust attenuation and galaxy stellar mass. The correlation evolves only marginally, less than 0.3 mag in A_{UV} , over the redshift range explored:

$$A_{\text{UV}} = 1.6 \times \log M_{*} - 13.5 \text{ mag for } z = 1.2\text{--}4. \quad (10)$$

The stellar mass content of a star-forming galaxy turns out to be a robust proxy of the dust attenuation affecting its UV emission (see, e.g., Pannella et al. 2009a, 2013; Schaerer & de Barros 2010; Buat et al. 2012; Kashino et al. 2013; Heinis et al. 2014; Oteo et al. 2014). Previous results focused on a relatively small redshift range, and sometimes on a UV-selected sample, which, by construction, tends to be biased against the most massive and most obscured star-forming galaxies. Here, we test this correlation over a wide redshift range and in a mass-complete way. Testing the real dispersion of a correlation obtained through stacking is not trivial and often impractical. To at least obtain an estimate of the correlation scatter, we used all *Herschel* detected sources in the field (Elbaz et al. 2011) up to $z \simeq 1.3$, where the GOODS-*Herschel* data are deep enough to allow a good statistical description of the parent sample, and estimate a dispersion of $\simeq 1$ mag in A_{UV} .

The mild evolution of this correlation lends support to a number of earlier studies claiming that the same amount of star formation suffered less dust extinction at high redshift compared to the local universe (e.g., Buat et al. 2007; Reddy et al. 2012). Because of the evolution of the MS with redshift, the same amount of star formation is hosted in galaxies that are less and less massive as redshift increases, and hence they will suffer a correspondingly lower UV dust extinction.

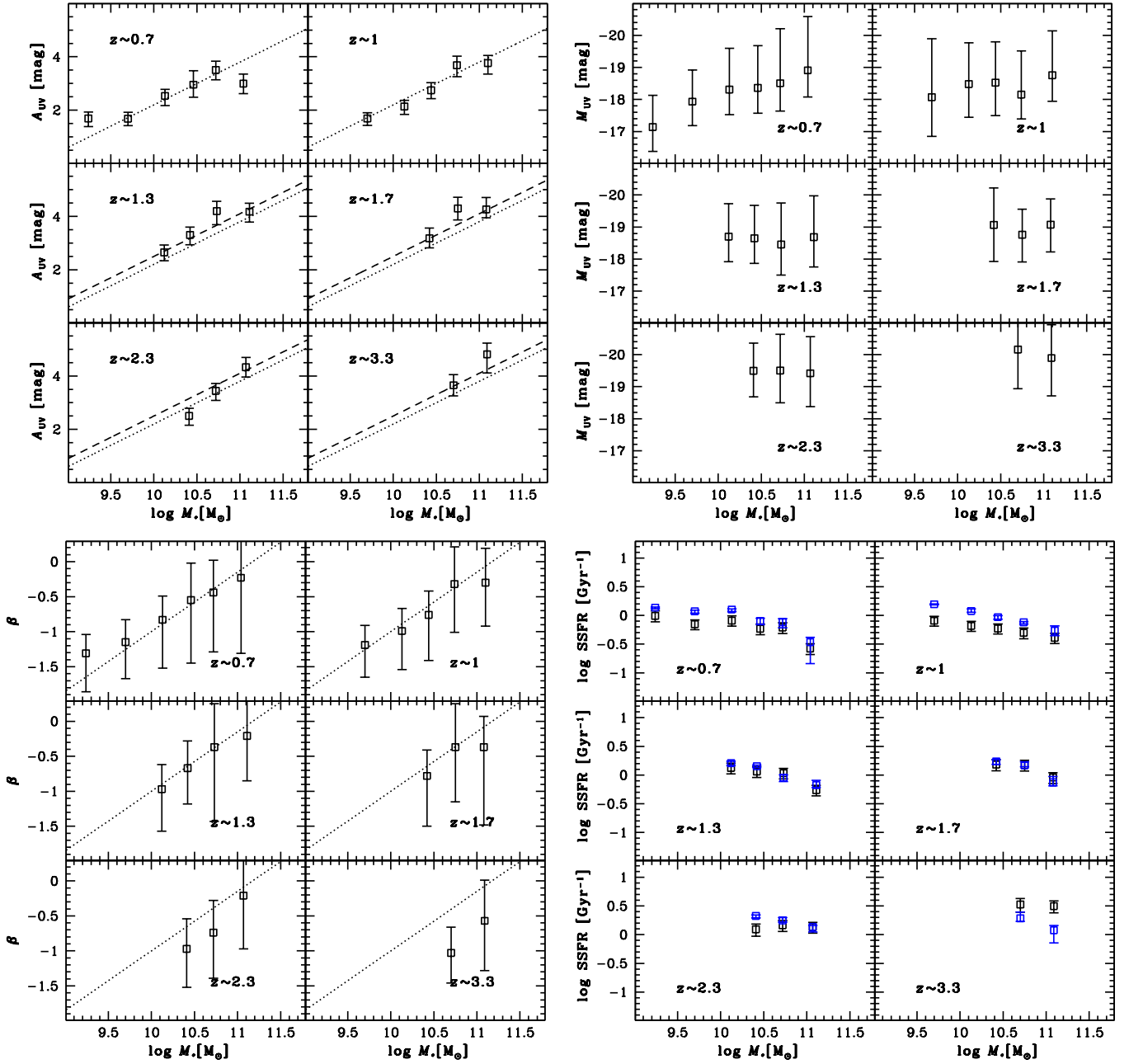


Figure 7. Top-left: measured A_{UV} vs. M_* correlation at different redshifts and for mass-complete bins. The dotted (dashed) lines are the linear fit to the measured values in the redshift range 0.5–1.2 (1.2–4.0). As described in the text, above redshift 1 we see a higher normalization by only $\simeq 0.3$ mag of attenuation at a fixed stellar mass. Using the *Herschel* detections from Elbaz et al. (2011), we estimate a dispersion in the correlation of $\simeq 1$ mag. Top-right: median absolute magnitudes M_{UV} for each bin of stellar mass where stacking has been performed. Galaxies with very different stellar masses, and hence very different star-formation rates, emit about the same dust attenuated UV light in average. Error bars show the rms of UV absolute magnitude distributions, rather than the error on the median values. Bottom-left: median UV spectral slope, β , for each bin of stellar mass where stacking has been performed. The dotted line is the best fit to the lowest redshift bin and is kept constant in all panels in order to show the evolution of the UV slope at a fixed stellar mass: the median β slope becomes bluer with redshift by about 0.5 over the whole redshift range, but most of the evolution seems to happen at $z \geq 2$. Error bars represent the rms of β distributions in the stellar mass bins. Bottom-right: comparison between the IR (black) and UV corrected (blue) derived SSFRs as a function of redshift and stellar mass. At $z \geq 3$ SFRs derived from the UV corrected emission underestimate the real SFR by a factor of around two.

On the other hand, dust attenuation is known to correlate with both the ISM metallicity and the mode and geometry of star formation in the host galaxy (i.e., normal versus starbursting galaxies; e.g., Heckman et al. 1998; Cortese et al. 2006). The ISM metallicity is expected to decrease with increasing redshift at a fixed mass, according to the well-studied evolution of the mass–metallicity relation (e.g., Tremonti et al. 2004; Savaglio et al. 2005; Erb et al. 2006;

Zahid et al. 2013a; Steidel et al. 2014; Wuyts et al. 2014). By assuming that the ISM conditions of star-forming galaxies do not change with redshift, one would see an effective decrease with redshift of UV attenuation at a fixed stellar mass; if anything we measure a weak trend going in the opposite direction. This suggests that the ISM conditions of MS galaxies are evolving, becoming somehow more extreme with redshift, which might be understood as a combination of both

geometrical effect (star-forming galaxies are becoming smaller with redshift) and physical effects (gas fractions and SSFRs are also increasing), leading to an overall enhanced volume density of the UV radiation field, as also suggested by their higher dust temperature (e.g., Hwang et al. 2010; Magdis et al. 2012; Symeonidis et al. 2013). We will come back to this topic in Section 5.

In the top-right panel of Figure 7 we plot the median UV luminosities for the same stacking samples as in the top-left panel. We show that at all redshifts explored, there is basically no correlation between stellar mass—a robust proxy of the ongoing star formation and UV dust attenuation—and the emerging UV photons. The amount of UV radiation that escapes dust absorption rises systematically with redshift, mimicking the rise of the general SFR level. This is consistent with a number of previous studies (e.g., Pannella et al. 2009a; Buat et al. 2012; Heinis et al. 2013). Here we explore a wider redshift range and find basically the same result: the emerging UV light is not (or only marginally in the lowest redshift bin) correlated to the actual SFR present in a galaxy, because the tight correlation between stellar mass (a proxy for the SFR; i.e., the intrinsic UV luminosity) and the dust attenuation conspires so that the emerging, average UV light is basically the same for very different SFR levels.

In the bottom-left panel of Figure 7 we plot the median values of the galaxy UV spectral slope, β , per bin of stellar mass at different redshifts. The two quantities, β and M_* , clearly correlate at all redshifts, but while the attenuation is fairly constant, or slightly increasing with redshift, β values are becoming systematically bluer with redshift. The fact that high mass galaxies at high redshift have a similar dust attenuation compared to similar mass galaxies at lower redshift, but have bluer UV slopes, has important implications for UV-derived SFRs in the high redshift universe. Using standard recipes to correct the observed UV emission by means of the β slope would systematically underpredict galaxy SFRs. This is better quantified in the bottom-right panel of Figure 7 where we compare the MS evolution plot derived from *Herschel* data (black squares, as in Figure 4) to the one derived from UV data that have been corrected by dust attenuation according to Equation (9). The comparison shows three main features mainly dependent on redshift: (1) at redshift lower than $\simeq 1$, the $\text{SFR}_{\text{UV}^\beta}$ tends to overpredict the real SFR, possibly because old stellar populations present in star-forming galaxies contribute to create redder UV slopes, which is in agreement with previous results (see, e.g., Kong et al. 2004 and Overzier et al. 2011); (2) in the redshift range 1.5–2.5 the $\text{SFR}_{\text{UV}^\beta}$ provides a nice match to the real SFR, which is in agreement with the results of Daddi et al. (2007), Pannella et al. (2009a), and more recently Rodighiero et al. (2014); (3) at redshift $z \gtrsim 3$, $\text{SFR}_{\text{UV}^\beta}$ are systematically lower than the real SFR and underpredict the SFR by a factor of around two. Intriguingly enough, this underprediction of SFR conspires to create a plateau of the UV-derived SSFR values at $z \gtrsim 2$, which was indeed observed in previous studies (see, e.g., Daddi et al. 2009; Stark et al. 2009; González et al. 2010). Moreover, the underestimate of SFR has an impact on the estimates of the global star-formation rate density (SFRD) at $z \gtrsim 3$, which were obtained by applying standard recipes as in Equation (9), which should be revised upward by a factor of around two. A similar conclusion was already presented in de Barros et al. (2014) and Castellano et al. (2014) for a sample of low mass

Lyman-break selected galaxies. We extend their result to the high mass end of the star-forming galaxy population at high redshift, putting forward a more compelling case for an upward revision of the SFRD estimates obtained so far at redshift $z \gtrsim 3$. We will further discuss this point in the next section.

4.3.2. A_{UV} Versus β Slope

In this section we explore in more detail the correlation between UV dust attenuation and UV spectral slope by repeating the stacking procedure in bins of β slope. This correlation is possibly the most widely used method in the literature to derive UV dust attenuation and hence SFR from UV emission. Daddi et al. (2004) calibrated it for a sample of star-forming galaxies at $z \simeq 2$ using the observed $B-z$ color, which is a proxy for the UV spectral slope, and this calibration has been shown to behave well in a number of studies (e.g., Daddi et al. 2007; Pannella et al. 2009a; Rodighiero et al. 2014). Previously, Meurer et al. (1999) had shown that local starburst galaxies and $z \simeq 3$ Lyman Break Galaxies (LBGs) were following the same tight correlation between dust attenuation and UV spectral slope. After Meurer et al. (1999), numerous investigations have tried to understand whether the so-called Meurer law would be followed by galaxies at all redshifts. Different studies have reported contradicting results on this topic. One thing that became clear within a few years after the Meurer et al. (1999) study was that local normal star-forming galaxies do not follow the Meurer relation, but instead have “redder” spectra for the amount of dust attenuation they suffer (e.g., Kong et al. 2004; Buat et al. 2005; Overzier et al. 2011; Boquien et al. 2012 and Grasha et al. 2013).

We now investigate how the median UV spectral slope relates to dust attenuation. We plot at all redshifts the predicted correlations between dust attenuation and β slope according to Meurer et al. (1999), Calzetti et al. (2000), Daddi et al. (2004), and Overzier et al. (2011) in the top-left panel of Figure 8. The first three determinations overlap to a large extent at blue UV slopes, whereas they spread out at red slopes. The determination of Overzier stands out, clearly predicting redder slopes at a fixed attenuation. However, we remind the reader that the sample used in Overzier et al. (2011) was strictly based on blue star-forming objects with no data points having a UV slope redder than $\beta = -0.5$. For this reason, the derived relation (R. Overzier 2015, private communication) should not be used for very attenuated galaxies.

When binned in β values, the median dust attenuation of the MS galaxy population follows the prediction of the Calzetti law already in our lowest redshift bin with a possible hint of a dust attenuation overprediction only in the reddest slope bins and up to redshift $z \simeq 1$. This latter feature might be due to old stellar populations of massive galaxies already contributing partially to the measured red slope. The departure from the Meurer relation of normal star-forming galaxies in the local universe has been thoroughly investigated in the last years and it is now well accepted to explain it as an age effect of the stellar populations hosted in local galaxies (see e.g., Kong et al. 2004; Boquien et al. 2012 and Grasha et al. 2013). In the lowest redshift bin we show, for illustrative purpose, the relation obtained by Boquien et al. (2012) for a sample of seven face-on and spatially resolved local star-forming galaxies drawn from the *Herschel* Reference Survey.

Already at $z \simeq 1.3$ and up to $z \lesssim 3$ the correlation matches well the observed data points, agreeing with the results of

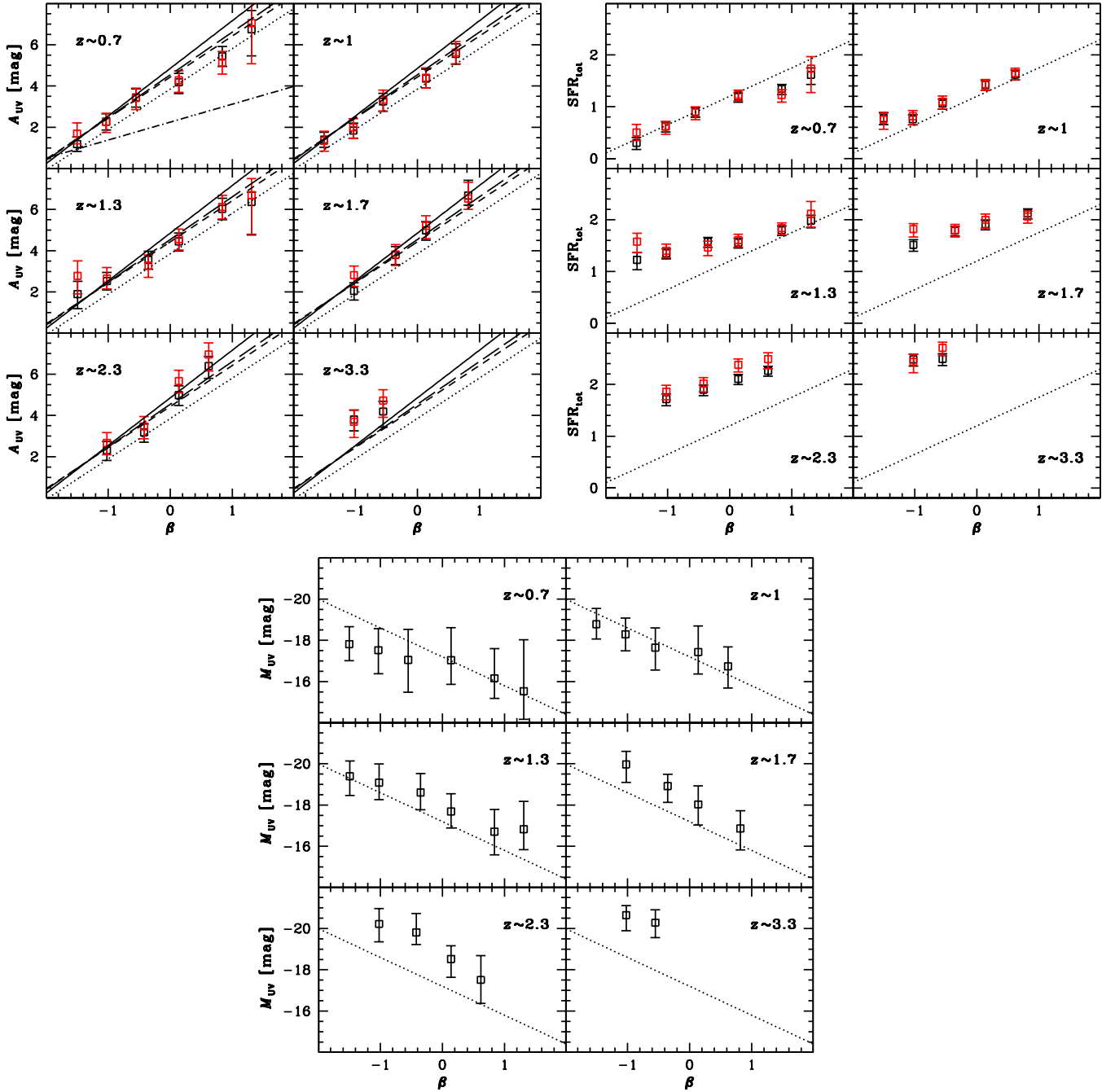


Figure 8. Top-left: the measured A_{UV} vs. β correlation at different redshifts for the whole sample of star-forming galaxies in the GOODS-N field. The A_{UV} measurements obtained from the stacked IR GOODS-H and radio VLA data are plotted with black and red squares. The predictions of different attenuation laws from Calzetti et al. (2000), Meurer et al. (1999), Daddi et al. (2004), and Overzier et al. (2011), are shown with solid, long-dashed, dashed, and dotted lines, respectively. Only in the lowest redshift bin, at $z \sim 0.7$, do we show, with a dotted-dashed line, the relation found for local star-forming galaxies by Boquien et al. (2012), see text for more details. Top-right: total SFR from stacking the IR and radio data in bins of UV slope and redshifts. The dotted line is the fit to the lowest redshift bin stacking result and is kept identical in all panels to show the evolution in the correlation between the two quantities. At a fixed value of β (i.e., at fixed UV dust attenuation; to within a factor of 2 accuracy), the star-formation rate grows by more than an order of magnitude. Bottom: median UV absolute magnitude in the bins of β used to produce the stacking results. A negative correlation is found, which points toward higher attenuation for galaxies with a fainter measured UV. Error bars show the rms values of the UV luminosities in the bins of β .

Daddi et al. (2007), who found the UV slope derived SFR to be in good agreement with those derived from $24\mu m$ and radio continuum data. In the last redshift bin, between redshift three and four, we see that UV dust attenuation tends to overshoot the prediction of the Calzetti law, so that galaxies tend to have bluer UV slope (an offset of about -0.5) compared to that predicted by the Calzetti law at a fixed amount of dust

attenuation. These results are overall in very good agreement with the ones discussed in the previous section.

There are several possible explanations for such a finding. At face value, our result seems to support the suggestion of Maiolino et al. (2004) that high redshift galaxies have a grayer attenuation law than that predicted by the Calzetti law. More recently, Castellano et al. (2012, 2014), Alavi et al. (2014), and

de Barros et al. (2014) have suggested that the bluer slopes can be explained by galaxy stellar populations with young ages and/or low metallicities. On the other hand, Buat et al. (2012) and Kriek & Conroy (2013) have suggested that the presence of a dust bump, the feature at 2175Å, in the spectra of high redshift galaxies may result in underestimated values for the UV slope. Finally, blue spectral slopes at high redshift have been reported previously by a number of authors as the possible outcome of a measurement bias when estimating slopes directly from the observed photometry rather than, as we actually do in this study, using the best-fit SED (see, e.g., Buat et al. 2011; Bouwens et al. 2012; Finkelstein et al. 2012).

We remind the reader that our β estimate is directly proportional to the difference between two rest-frame magnitudes, and as such, one might wonder about the possible impact of the shifts applied to the photometric catalog on our conclusions. The only sensible difference we find by using the uncorrected photometry is in the first redshift bin. The uncorrected photometry would produce redder slopes ($\Delta\beta \sim 0.3$) that would possibly move our dust attenuation further away from the Meurer relation and possibly closer to the correlation found in the local universe (see e.g., Boquien et al. 2012). All the other redshift bins remain unchanged with possibly a weak trend, but well within the estimated β uncertainty, to obtain slightly bluer values compared to the ones we used. If anything, this would reinforce our conclusions.

In the top-right panel of Figure 8 we show the evolution of the correlation between β slope and total star-formation rate at different redshifts. In all panels the dotted line is the best fit to the correlation in the lowest redshift bin, and is kept constant at all redshifts to aid comparison. As redshift increases, the amount of star-formation rate at fixed UV slope increases steadily. This can be interpreted as further evidence that at a fixed star-formation rate there is less and less UV dust attenuation with increasing redshift.

In the bottom panel of Figure 8 we plot the median measured UV absolute magnitudes from stacks in bins of β . When binning in β values, again basically in UV dust attenuation, we find an anti-correlation between the measured UV luminosities and UV slopes: galaxies with redder UV slopes (i.e., that are more dust attenuated) tend to have fainter UV-measured luminosities, and galaxies that are less attenuated have brighter luminosities. Pannella et al. (2009a) found a similar result at $z \simeq 1.7$; here we can extend this at all probed redshifts (see also Heinis et al. 2013, 2014 for similar conclusions).

Overall, the anti-correlation is consistent with an identical slope at all redshifts, but with an increasing normalization, with the UV-observed luminosity increasing by about 2 mag (more than a factor of six in luminosity) over the redshift range probed.

Reddy et al. (2009) found that low luminosity objects are less attenuated, statistically speaking, than more UV luminous ones. This is contrary to what we found here. The difference between our result and the study of Reddy et al. (2009) might easily be explained by the different selection bands used to build the galaxy samples. We use a K band selection, which is very close to a stellar mass selection at these redshifts and is therefore sampling the mass function in a fairly complete way at the high mass end. Massive galaxies are intrinsically UV-bright but also extremely dust attenuated, so it is reasonable to expect these objects to be very faint in the UV and even fainter

than lower mass galaxies. On the other hand, the LBG-like selection of the Reddy et al. (2009) sample, which is mostly an observed frame UV selection, will sample the mass function in a sparse way, giving more weight toward the low-to-intermediate range of the stellar mass function, thereby missing almost completely the high mass end as well as very obscured galaxies, which is a very well-known drawback of UV-selected samples. These UV-selected low mass galaxies preferentially have low star-formation rates, low dust attenuation, and faint UV output, and they outnumber by several orders of magnitude the high mass galaxies.

5. THE EVOLVING ISM OF STAR-FORMING GALAXIES OVER COSMIC TIME

In this section we will discuss our results and investigate how the interstellar medium of star-forming galaxies and the physical conditions of star formation evolved with redshift. The main ingredient we will use in this section is the fact that the correlation between UV dust attenuation and galaxy stellar mass remains fairly constant with redshift. Notwithstanding a 10-fold increase in star-formation rate, indeed over more than 6 Gyrs ($0.5 \leq z \leq 4$), on average a galaxy of a given stellar mass suffers only a slight increase of dust attenuation. We measure at most an increase of about 0.3 mag, which corresponds to a factor of 1.3 in UV flux attenuation. We will complement this nonevolving $A_{UV}-M_*$ relation with other measured scaling relations that describe the conditions of the star-forming galaxy ISM, and finally derive some evolutionary trends and conclusions.

5.1. The Attenuation of Continuum versus Line Emission

$H\alpha$ line emission is a well-known and robust SFR tracer (e.g., Kennicutt 1998). Lying at 6563 Å, $H\alpha$ suffers much less dust extinction compared to the UV rest-frame of a galaxy spectrum, and hence the uncertainty on SFR due to the poorly known dust attenuation can be an order of magnitude less severe as compared to UV-based estimates. On the other hand, $H\alpha$ is mainly produced by extremely massive ($\geq 10 M_\odot$), short lived ($\simeq 10$ Myr) stars that are deeply embedded in the giant molecular cloud (GMC) H II regions, whereas stars producing the UV rest-frame stellar continuum are less massive ($< 10 M_\odot$), shine over timescales 10 times longer, and have time to migrate out of the dense H II regions. The net outcome of this process is that $H\alpha$ emission suffers from an extra attenuation that has, as we just described, an origin driven by the distribution and density of H II regions within the galaxy itself.

The extra amount of attenuation suffered by nebular emission has been quantified in the local universe by Calzetti et al. (2000) to be a factor of about 1.7. This can be derived by accounting for the fact that the stellar continuum suffers roughly half of the reddening suffered by the ionized gas, encoded into $E(B-V)_{\text{star}} = 0.44 E(B-V)_{\text{gas}}$, and for the fact that Calzetti et al. (2000) used two different extinction curves for the nebular (Fitzpatrick 1999) and continuum (Calzetti et al. 2000) emission,³⁰ which have similar shapes but different normalizations (i.e., $R_V = 3.1$ and 4.0, respectively) and thus

³⁰ We refer the reader to the review paper of Calzetti (2001) where all the details about the actual assumptions are extensively discussed.

for example:

$$\frac{A_{V_{\text{gas}}}}{A_{V_{\text{star}}}} = \frac{R_{V_{\text{gas}}}}{R_{V_{\text{star}}}} \frac{E[B - V]_{\text{gas}}}{E[B - V]_{\text{star}}} \sim 1.7. \quad (11)$$

Here we assume that this ratio is constant in wavelength. Although this is obviously not strictly true, because it assumes that the two mentioned extinction curves have exactly the same shape and only differ in their normalizations, it turns out not to be a bad approximation at the wavelengths of interest for this study, namely at the FUV (1500 Å) and H α (6563 Å) wavelengths.

At higher redshift, things are much less constrained. Erb et al. (2006) and Reddy et al. (2010) claimed a vanishing difference (i.e., a factor close to 1) at $z \simeq 2$, but this has not been confirmed by more recent studies (e.g., Förster Schreiber et al. 2009; Mancini et al. 2011; Wuyts et al. 2013).

The main problem for high redshift studies is the lack of a robust SFR indicator to directly compare with the H α and UV stellar continuum measurements. Most of the quoted studies rely on indirect arguments by estimating the SFR from SED-fitting or from the β -corrected UV luminosity. This approach is, by construction, extremely uncertain and model dependent, and the end product of the comparison is closer to a correction factor, which produces in the end a consistent result, rather than a real physical measure.

In this study we are able to look at the redshift evolution of the extra attenuation affecting H II regions in a way that is relatively model-independent. The idea is to compare the $A_{\text{UV}}-M_*$ relation we found in Section 4.3.1 to the one that similarly relates $A_{\text{H}\alpha}$ to M_* as a function of redshift. The latter relation was first established in the local universe by Garn & Best (2010), and more recently confirmed by Zahid et al. (2013b), by using the observed Balmer decrement (i.e., H_α/H_β) measured in the SDSS spectra, as previously done in Brinchmann et al. (2004).

In the last few years, several different studies have looked at the redshift evolution, at least up to $z \simeq 1.5$, of the $A_{\text{H}\alpha}-\log M_*$ relation (e.g., Sobral et al. 2012; Domínguez et al. 2013; Ibar et al. 2013; Kashino et al. 2013; Momcheva et al. 2013; Price et al. 2014). Compared to the SDSS studies mentioned, these are affected by additional uncertainties and systematics mainly due to the fact that the H α line shifts into the observed NIR for $z \geq 0.5$, which this often prevents a robust measurement of the Balmer decrement. Despite the diverse caveats, most of the results obtained so far, up to $z \simeq 1.5$, suggest that there is no redshift evolution of the $A_{\text{H}\alpha}-\log M_*$ correlation (see, e.g., Figure 6 in Price et al. 2014, for a compilation of available results).

In order to investigate the nebular extinction, we start from the $A_{\text{H}\alpha}-\log M_*$ correlation defined in Garn & Best (2010):

$$A_{\text{H}\alpha} = 0.91 + 0.77 \cdot X + 0.11 \cdot X^2 - 0.09 \cdot X^3 \quad (12)$$

where the quantity $X = (\log M_* - 10 + 0.23)$ is the stellar mass term in units of $10^{10} M_\odot$, which has been rescaled from the Chabrier (2003) IMF used in Garn & Best (2010) to the Salpeter (1955) IMF we are using in this study. We then assume a Calzetti et al. (2000) attenuation law to derive the line reddening $E[B - V]_{\text{gas}}$ as

$$E[B - V]_{\text{gas}} = A_{\text{H}\alpha}/3.33 \quad (13)$$

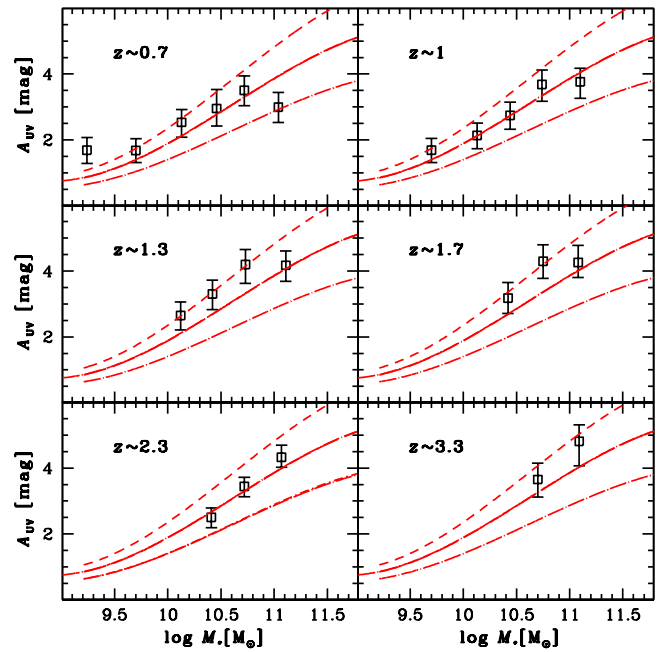


Figure 9. Comparison between the continuum and line attenuation as a function of stellar mass and redshift. Black squares show our stacking measurements of the UV dust attenuation as a function of mass and redshift, while the red lines (dashed, solid, dot-dashed) represent the Garn & Best (2010) relation between H α attenuation and stellar mass, which has been extrapolated to the UV rest-frame assuming the Calzetti et al. (2000) attenuation law and scaled by 1, 1.3, and 1.7, respectively. The last value (1.7) is the factor found by Calzetti (2001) in the local universe.

and the dust attenuation for nebular emission in the rest-frame UV as

$$A_{\text{UV-gas}} = 10.4 E[B - V]_{\text{gas}}. \quad (14)$$

In Figure 9 we overplot the rescaled Garn & Best (2010) relation divided by different scaling factors (1.7, 1.3 and 1) to the measured $A_{\text{UV}}-\log M_*$ relation at different redshifts. The factor 1.7 corresponds to the value found by Calzetti et al. (2000) in the local universe. We find that this factor is clearly smaller at higher redshift, being 1.3 at $z \simeq 1$, and becomes close to unity at higher z (i.e., there is a vanishing difference between the reddening of the H II regions and that affecting the young massive, but slightly older stars responsible for the rest-frame continuum UV emission). This can be understood through the fact that, on the one hand, galaxies are becoming slightly smaller in physical size (e.g., Ferguson et al. 2004; Elbaz et al. 2011; M. Pannella et al. 2015, in preparation), but at the same time their SFR density is dramatically increasing from the local universe out to $z \simeq 4$, by more than a factor of 40 according to the evolution of the normalization of the $\log M_*-\log \text{SFR}$ correlation. For this reason, the volume density of H II regions is becoming so high that it almost entirely fills up the available galaxy surface. This suggests that the ISM of star-forming galaxies was more and more dense and opaque to UV radiation with increasing redshift, eventually resembling, at $z \geq 1.5$, the physical properties of local H II regions, as has already been suggested by Reddy et al. (2010). More recently Wild et al. (2011), by analysing a sample of 15,000 galaxies drawn from the Sloan Digital Sky Survey, have found a correlation between the galaxy SSFR and the ratio between line and continuum reddening, such that the ratio becomes smaller with increasing SSFR (i.e., as galaxies move toward the

starburst regime locus). The correlations found by Wild et al. (2011) nicely explain our findings and also support our conclusions.

On the other hand, Kreckel et al. (2013) have shown, for a sample of eight nearby spatially resolved star-forming galaxies, that the ratio between line and continuum reddening increases with increasing SFR density, which seems to be in contradiction with the Wild et al. (2011) result and hence our conclusions. Very recently, Boquien et al. (2015) obtained results on M33 that are in nice agreement with those presented in our study, and suggest possible causes explaining the controversial results of Kreckel et al. (2013).

5.2. The Relation Between Dust Attenuation and ISM Metallicity

A more physical way to look at ISM conditions in star-forming galaxies and their possible evolution with redshift is to consider the correlation between dust attenuation and ISM metal content for star-forming galaxies. In the local universe, the correlation was first explored by Heckman et al. (1998) for a sample of starburst galaxies, and found to yield a tight correlation for star-forming galaxies, with more dusty galaxies being more metal-rich. The correlation is expected, as metals and more complex molecules are needed to build up the actual dust that absorbs the UV radiation.

In a more recent study, Cortese et al. (2006) looked at the properties of normal star-forming galaxies in the local universe, which is a sample of objects much closer to what in this paper we are calling MS galaxies. They found that these objects were also following a tight correlation between dust attenuation and ISM metallicity, but the correlation was offset with respect to the one of starburst galaxies: at a fixed metallicity, starburst galaxies are more obscured compared to normal star-forming galaxies. The offset in the correlation can be interpreted in terms of the different ISM conditions found in normal star-forming and starburst galaxies. The latter host star-formation rate densities that are much higher than the former—both because star formation happens on smaller scales and because of the higher overall levels of star formation—and they can enrich most of their ISM much earlier and more quickly than normal smoothly star-forming MS galaxies.

Assembling mass-complete samples of galaxies at high redshift with measured metallicities is practically impossible with the presently available facilities. Most of the existing data in the high redshift universe faces the unavoidable bias of being limited by line sensitivity, and hence selected against the more dusty, more massive, but also more metal-rich systems. As already discussed, dust attenuation has an impact on the source detection itself, which becomes more and more severe with increasing redshift. This bias can in principle become extreme when it comes to metallicity measurements; where easiest, and possibly only, the objects for which it is possible to measure metallicity will be the ones with the lowest metallicities. This effect, which seems to nicely fit and explain some recent observations of metallicity and gas masses at high redshift (see, e.g., Ouchi et al. 2013; Tan et al. 2013; Troncoso et al. 2014), should be carefully taken into account when dealing with high redshift sources and deriving more general conclusions on the global galaxy population.

Here we estimate the median ISM metallicity for our samples of star-forming galaxies in bins of stellar mass and redshift, by applying the fundamental metallicity relation,

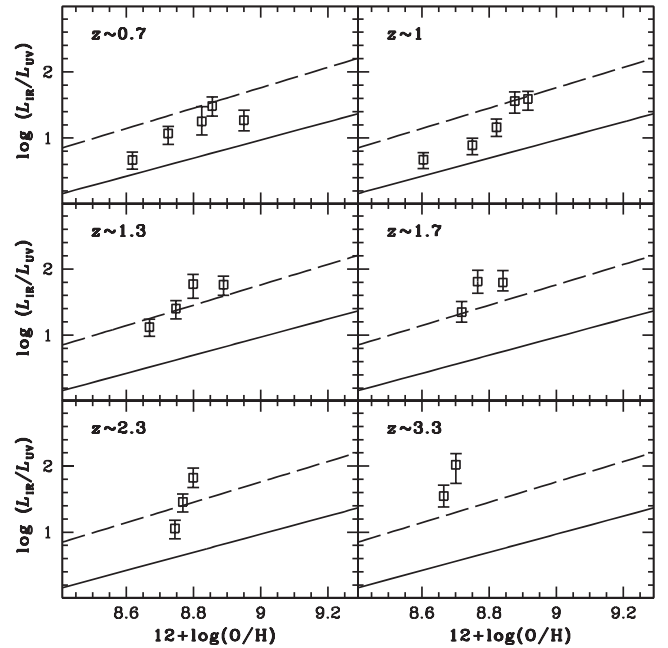


Figure 10. ISM metallicity vs. $\log(L_{\text{IR}}/L_{\text{UV}})$ at different redshifts. Here we estimate the ISM metal content using the fundamental metallicity relation (FMR) defined in Mannucci et al. (2010). We show at all redshifts the correlations found in the local universe for starburst (Heckman et al. 1998) and normal star-forming galaxies (Cortese et al. 2006) with long-dashed and solid lines, respectively. MS galaxies tend to catch up at high redshift with the correlation followed by starburst galaxies in the local universe, suggesting an evolution of the median star-forming galaxy ISM.

FMR, found by Mannucci et al. (2010). The relation is defined as a tri-dimensional surface linking the SFR, metallicity, and stellar mass of star-forming galaxies. It was first defined in a sample of local SDSS galaxies and then found to be essentially identical at higher redshift at least out to $z \simeq 2$. The relation implies that galaxies obey an anti-correlation between SFR and metallicity: at a fixed stellar mass, the drop in metal content of star-forming galaxies at higher redshift (e.g., Tremonti et al. 2004; Savaglio et al. 2005; Erb et al. 2006; Zahid et al. 2013a) is compensated by the increase of SFR.

As already stated, the FMR has been verified to remain unchanged up to $z \simeq 2$ (but see also, e.g., Cullen et al. 2014; Steidel et al. 2014; Wuyts et al. 2014), whereas a sudden change has been claimed recently by the same authors in Troncoso et al. (2014) for a sample of $z \simeq 3$ galaxies. Of course, we have no way to test their claim here because of how the mass completeness of their sample could possibly affect their conclusions according to that discussed in the previous paragraph. Very recently, Maier et al. (2014) have shown that, when using the formalism put forward in Lilly et al. (2013) or equivalently the Equation (2) of Mannucci et al. (2010), as opposed to the Equation (5) of the same paper, to describe the FMR, measured high redshift metallicities are overall well described by the FMR predictions. Here we assume that the FMR is still valid in our last redshift bin at $z \geq 3$.

We show in Figure 10 the evolution of dust attenuation (here shown as the classical quantity $\log(L_{\text{IR}}/L_{\text{UV}})$ instead of A_{UV} as in the previous figures, to be consistent with the studies of Heckman et al. 1998 and Cortese et al. 2006) versus the ISM metallicity at different redshifts. We see a clear evolution in redshift, with MS galaxies moving progressively from the locus occupied locally by normal star-forming galaxies (solid line),

to overshooting the one occupied locally by starburst galaxies (long-dashed line). We stress here that our result does not depend dramatically on the use of the FMR, but that a qualitatively similar outcome would have been obtained by simply using the evolution of the mass–metallicity relation. In other words, the shift in redshift we are seeing is mostly driven by the fact that galaxies of a fixed stellar mass are becoming less and less metal-rich with increasing time while keeping the same level of UV dust attenuation.

Galaxies are becoming overall more compact in sizes at higher redshift and at the same time they are also significantly increasing their SFRs, and this trend makes them more and more similar to local starbursts in terms of ISM conditions. This similarity is also confirmed by the reported evolution of the IR SED shape of MS galaxies up to redshift $z \simeq 2$ by Magdis et al. (2012) and more recently by Magnelli et al. (2014). It is extended at higher redshift by C. Schreiber et al. (2015, in preparation), with the median dust temperature becoming higher with increasing redshift and indeed more similar to the dust temperatures of highly star-forming galaxies seen in the local universe.

6. CONCLUSIONS

We have presented the first results of an ongoing project aimed at better understanding the physics of star formation across cosmic time. We take advantage of one of the deepest panchromatic data sets available at present to select a star-forming galaxy sample in the GOODS-N field and to obtain a complete sampling of galaxy redshifts, SFRs, stellar masses, and UV rest-frame properties. We quantitatively explore, with a mass-complete sample, the evolution of star formation and dust attenuation properties up to $z \simeq 4$ as a function of the host galaxy properties. Our main results can be summarized as follows.

1. We find that the slope of the $SFR-M_*$ correlation is equal to 0.8, and is consistent with being identical at all redshifts up to $z \simeq 1.5$, while the normalization increases continuously up to the highest redshift we are able to probe.
2. We explored the FIR–radio correlation for a mass-selected sample of star-forming galaxies and found that the correlation does not evolve up to $z \simeq 4$.
3. We confirm that galaxy stellar mass is a robust proxy for the UV dust attenuation in star-forming galaxies: more massive galaxies are more dust attenuated than less massive ones.
4. We find that the correlation between UV attenuation and mass evolves very weakly with redshift: the amount of dust attenuation increases by less than 0.3 mag over the redshift range 0.5–4 for fixed stellar mass. This finding explains the already reported evolution of the $SFR-A_{UV}$ relation: the same amount of star formation is less attenuated at higher redshift because it is hosted in less massive, and less metal-rich, galaxies.
5. We explored the correlation between the UV dust attenuation and UV spectral slope, a widely used proxy for dust attenuation. At all redshifts the two quantities correlate, as already shown in many previous studies, but the correlation evolves with redshift with star-forming galaxies having a bluer UV spectral slope for the same amount of dust attenuation, reaching values inconsistent

with the ones predicted by the Calzetti law at $z \geq 3$. At these redshifts the SFRs derived from the UV corrected emission underestimate the real SFR by a factor of around two. Consequently, this should lead, as already suggested in Castellano et al. (2014), to an upward revision, by the same factor, of the SFRD estimates at redshift $z \geq 3$.

6. Combining our findings with previously published results from the $H\alpha$ line emission surveys, we find that at $z \simeq 1$ the line reddening is larger than the continuum reddening by a factor $\simeq 1.3$ and becomes closer to a factor of one by redshift $z \simeq 3$. This is substantially lower than the value of around 1.7, which is found in the local universe, and points toward a more compact and more dense star-formation distribution in high redshift star-forming galaxies.
7. Finally, using the fundamental metallicity relation to estimate ISM metallicities, we find that the amount of dust attenuation at a fixed ISM metallicity increases with redshift and reaches the local value for highly star-forming galaxies.

We speculate that our results point toward an evolution of the ISM conditions of the median MS star-forming galaxy, such that at $z \geq 1.5$ MS galaxies have ISM properties more similar to those found in local starbursts.

We thank the referee for thorough and constructive comments that helped clarify and improve this paper. We are grateful to N. Drory for sharing the SED-fitting code used to estimate galaxy stellar masses, to Masami Ouchi for kindly sharing the Subaru SuprimeCam deep z and ZR imaging of the GOODS-N field, and to Gabe Brammer for kindly answering (too) many questions about EAZY. This research was supported by the French Agence Nationale de la Recherche (ANR) projects ANR-09-BLAN-0224 and ANR-08-JCJC-0008 and by the ERC-StG grant UPGAL 240039. We acknowledge the contribution of the FP7 SPACE project “ASTRODEEP” (Ref. No. 312725), supported by the European Commission. S.J. acknowledges support from the EU through grant ERC-StG-257720. R.J.I. acknowledges support from the European Research Council in the form of the Advanced Investigator Programme, 321302, COSMICISM. This research has made use of the NASA/IPAC Infrared Science Archive, which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. This work is based in part on observations obtained with WIRCAM, a joint project of CFHT, Taiwan, Korea, Canada, France, at the Canada–France–Hawaii Telescope (CFHT), which is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences de l’Univers of the Centre National de la Recherche Scientifique of France, and the University of Hawaii. This work is also based in part on data collected at the Subaru Telescope, which is operated by the National Astronomical Observatory of Japan, and on observations made with the NASA *Galaxy Evolution Explorer*. GALEX is operated for NASA by the California Institute of Technology under NASA contract NAS5-98034. PACS has been developed by a consortium of institutes led by MPE (Germany) and including UVIE (Austria); KU Leuven, CSL, IMEC (Belgium); CEA, LAM (France); MPIA (Germany); INAF/IFSI/OAA/OAP/OAT, LENS, SISSA (Italy); IAC (Spain). This development

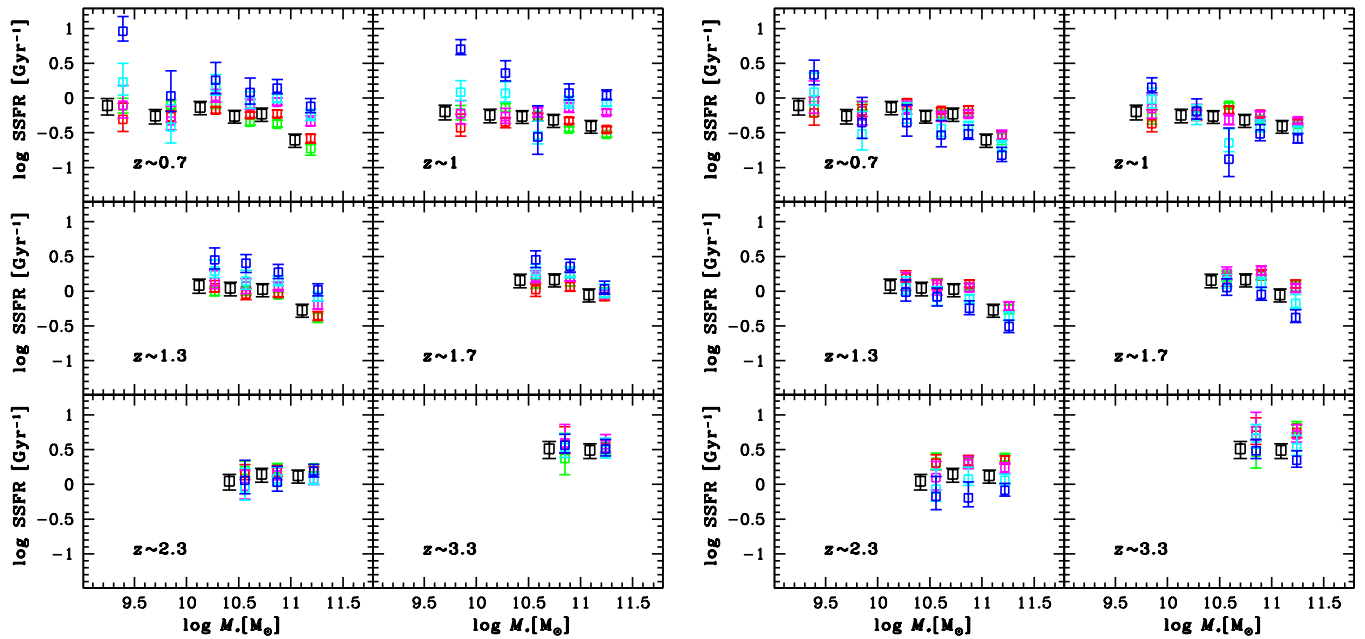


Figure 11. Left: single band derived SSFRs (green-100 μm , red-160 μm , magenta-250 μm , cyan-350 μm , and blue-500 μm) after applying a bolometric correction to each stacking result using the Chary & Elbaz (2001) library. We also plot in black, horizontally offset for the sake of clarity, the SSFR derived from the multi-band IR fitting, again using the same library but allowing for a free normalization of the SEDs. Right: the same information as in the other panel, with the only difference being that this time we have used the MS SED described in Elbaz et al. (2011) to derive bolometric corrections for all the IR band flux densities output from the stacking routine.

has been supported by the funding agencies BMVIT (Austria), ESA-PRODEX (Belgium), CEA/CNES (France), DLR (Germany), ASI/INAF (Italy), and CICYT/MCYT (Spain). SPIRE has been developed by a consortium of institutes led by Cardiff University (UK) and including Univ. Lethbridge (Canada); NAOC (China); CEA, LAM (France); IFSI, Univ. Padua (Italy); IAC (Spain); Stockholm Observatory (Sweden); Imperial College London, RAL, UCL-MSSL, UKATC, Univ. Sussex (UK); and Caltech, JPL, NHSC, Univ. Colorado (USA). This development has been supported by national funding agencies: CSA (Canada); NAOC (China); CEA, CNES, CNRS (France); ASI (Italy); MCINN (Spain); Stockholm Observatory (Sweden); STFC (UK); and NASA (USA).

APPENDIX A ON THE BOLOMETRIC CORRECTIONS FOR STACKING RESULTS

In this appendix we discuss the impact of bolometric corrections in the different *Herschel* bands on the derivation of the total IR luminosity. The GOODS-N field is one of the few fields observed deeply in the IR *Herschel* bands, which makes it special compared to other deep extragalactic fields. Previously only deep *Spitzer* 24 μm data were available to infer the total IR emission, and hence an estimate of the SFR. Here we test this with all the PACS/SPIRE bands of *Herschel* by comparing the outcome of the Chary & Elbaz (2001) library³¹ and the new MS SED defined in Elbaz et al. (2011), as shown in Figure 11. A similar comparison was presented in Elbaz et al. (2010, 2011), but this was limited to *Herschel* detections. Here we extend this to IR stacking results, which could be useful in case only relatively shallow data are available. By comparing to the results from the multi-band IR

SED-fitting, we can draw the following conclusions: (a) the new MS template (Elbaz et al. 2011) significantly improves the bolometric corrections for the SPIRE bands only up to redshift $z \simeq 2$; and (b) the Chary & Elbaz (2001) recipe works best for the bolometric corrections of the PACS bands at all the redshifts explored.

APPENDIX B TESTING RADIO STACKING AT DIFFERENT RESOLUTIONS

In this appendix we discuss the accuracy of the derived radio stacking results. Following Pannella et al. (2009a), in order to derive the radio fluxes of the stacked images, we fit a two-dimensional Gaussian, convolved with the image dirty beam, plus a background pedestal level using the *GALFIT* software (Peng et al. 2002). Allowing for a range of sizes, this kind of fitting naturally takes into account the physical sizes of the galaxies at different stellar masses and redshifts. One possible concern of such an approach could be a systematic bias introduced by the variable sizes, which might, in principle, impact the results of our study, such as the constancy of the radio-FIR correlation. To test for possible biases, we used the VLA image at 6'' resolution (G.E. Morrison 2015, private communication), which has been produced with the same visibility data set, but with a different tapering of the data that gives more weight to short VLA baselines. The 6'' resolution image is less deep than the natural 1''.5 image and therefore not optimal for the science goals of this study. Still, we can fruitfully use the lower resolution image by assuming that flux densities of the stacking results in such image can be derived robustly by a simple PSF fitting (i.e., not allowing for any impact of the source sizes). We show the comparison in the left panel of Figure 12. The results are in excellent agreement and usually within less than 10% difference, that is a value close to

³¹ Publicly available at http://david.elbaz3.free.fr/astro_codes/chary_elbaz.html.

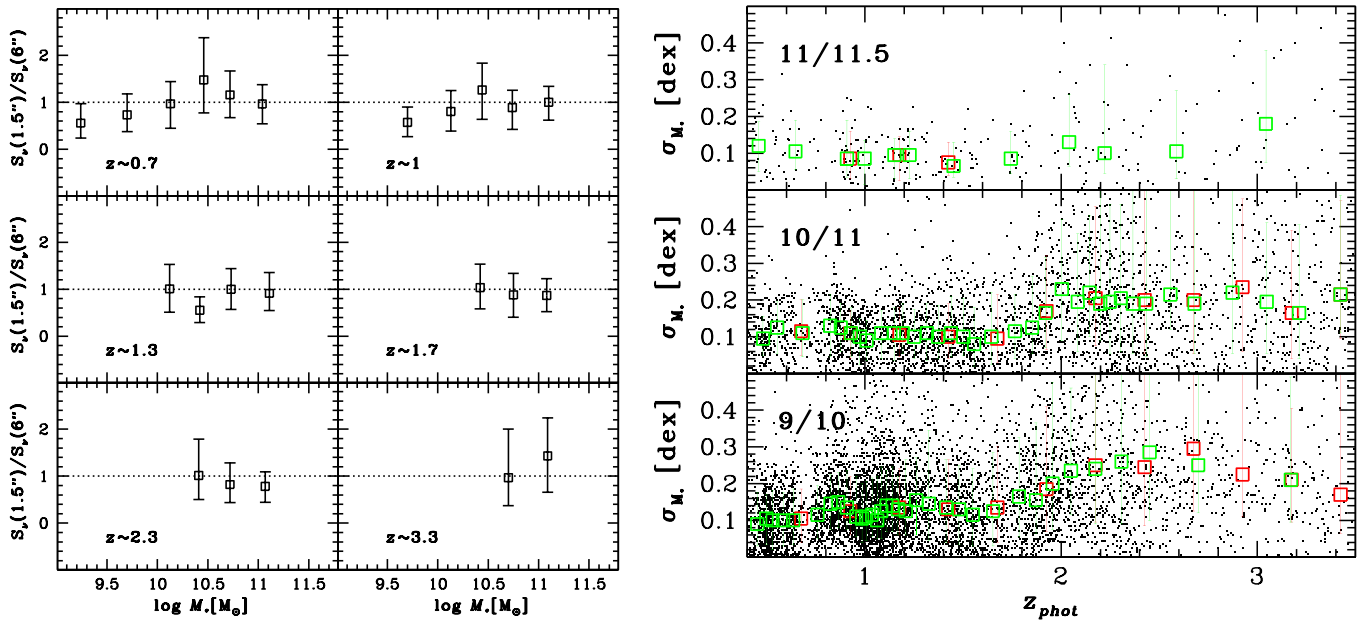


Figure 12. Left: ratios between the total radio flux densities obtained from stacking at two different image resolutions, 1.5 and 6 arcsec. We note here that a Gaussian-convolved model is used to fit the higher resolution image to account for the physical sizes of sources, while a simple PSF fitting has been performed for the 6'' resolution image. Right: the stellar mass error (see text for details) as a function of redshift and in three bins of stellar masses (indicated on the top-left of each panel) are all plotted with black dots. There is an increase of the median error (green and red empty squares) with redshift, passing from ~ 0.1 dex below redshift $z = 1.5$ to ~ 0.2 dex above redshift $z = 2$, and with decreasing stellar mass.

the minimum statistical error derived from our bootstrapping simulations.

APPENDIX C TESTING GALAXY STELLAR MASS ESTIMATES WITH MOCK CATALOGS AND A DIFFERENT SED-FITTING CODE

In this appendix we try to assess the accuracy of our *FAST*-derived stellar mass estimates by using both mock catalogs and the results from the SED-fitting code described in detail in Drory et al. (2004, 2009).

To test mass accuracies on mock catalogs we proceeded as follows. We created a suite of 200 mock catalogs by randomly shifting the photometric points within the measured accuracies. We then ran *FAST* on each mock catalog and derived for each object a distribution of best-fit stellar masses. In the right panel of Figure 12 we plot the stellar mass error as a function of redshift in three stellar mass bins. There is an increase of the median error with redshift, specifically passing from 0.1 dex below redshift $z = 1.5$ to 0.2 above redshift $z = 2$, which likely depends on the fact that when going to higher redshift the photometry gets more noisy in the optical rest-frame. It is also worth noting the increase of errors as stellar masses get smaller, which is again likely driven by the global signal-to-noise ratio decrease. This level of inaccuracy should be considered a good approximation to the real errors involved, but is still a lower limit because it does not factor in the discrepancies between different stellar population models.

Finally, we also tested our *FAST* masses against the masses obtained by the SED-fitting code described in Drory et al. (2004, 2009). This code has been used in a number of published GOODS-*Herschel* studies (Kirkpatrick et al. 2012; Magdis et al. 2012; Mullaney et al. 2012; Penner et al. 2012; Del Moro et al. 2013), making use of an earlier version of the photometric catalog in the GOODS-North field. For this reason,

a comparison between the two codes is useful both in terms of global accuracies, but also to understand the possible impact on the published results. We parameterize the possible SFHs by a two-component model consisting of a main, smooth component, described by an exponentially declining star-formation rate ($\psi(t) \propto \exp(-t/\tau)$) that is linearly combined with a secondary burst of star formation. The main component timescale τ varies in the range 0.1–20 Gyr, and its metallicity is fixed to solar. The age of the main component, t , is allowed to vary between 0.01 Gyr and the age of the universe at the object's redshift. The secondary burst of star formation, which cannot contain more than 10% of the galaxy's total stellar mass, is modeled as a 100 Myr old constant SFR episode of solar metallicity. We adopt a Salpeter (1955) IMF for both components, with lower and upper mass cutoffs of 0.1 and 100 M_{\odot} , respectively. Adopting the Calzetti et al. (2000) attenuation law, both the main component and the burst are allowed to exhibit a variable amount of attenuation by dust with $A_V^m \in [0, 1.5]$ and $A_V^b \in [0, 2]$ for the main component and the burst, respectively.

The results from the two runs with different codes and different SFHs show no systematic differences in stellar mass estimates, except at the highest stellar masses ($\log M_{*} \geq 11.2$), where the two-component fitting predicts larger masses by 0.2 dex, with a scatter of about 0.2 dex.

REFERENCES

- Alavi, A., Siana, B., Richard, J., et al. 2014, *ApJ*, **780**, 143
- Alexander, D. M., Bauer, F. E., Brandt, W. N., et al. 2003, *AJ*, **126**, 539
- Barger, A. J., Cowie, L. L., Chen, C.-C., et al. 2014, *ApJ*, **784**, 9
- Barger, A. J., Cowie, L. L., & Wang, W.-H. 2008, *ApJ*, **689**, 687
- Barger, A. J., Wang, W.-H., Cowie, L. L., et al. 2012, *ApJ*, **761**, 89
- Bauer, F. E., Alexander, D. M., Brandt, W. N., et al. 2004, *AJ*, **128**, 2048
- Bell, E. F. 2003, *ApJ*, **586**, 794
- Bell, E. F., van der Wel, A., Papovich, C., et al. 2012, *ApJ*, **753**, 167
- Bertin, E., & Arnouts, S. 1996, *A&AS*, **117**, 393

- Béthermin, M., Daddi, E., Magdis, G., et al. 2012a, *ApJL*, **757**, L23
- Béthermin, M., Dole, H., Beelen, A., & Aussel, H. 2010, *A&A*, **512**, A78
- Béthermin, M., Le Floc'h, E., Ilbert, O., et al. 2012b, *A&A*, **542**, A58
- Bongiorno, A., Merloni, A., Brusa, M., et al. 2012, *MNRAS*, **427**, 3103
- Boquien, M., Buat, V., Boselli, A., et al. 2012, *A&A*, **539**, A145
- Boquien, M., Calzetti, D., Aalto, S., et al. 2015, *A&A*, in press (arXiv:1502.01347)
- Bourne, N., Dunne, L., Ivison, R. J., et al. 2011, *MNRAS*, **410**, 1155
- 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. 2012, *ApJ*, **754**, 83
- Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, *ApJ*, **686**, 1503
- Brammer, G. B., Whitaker, K. E., van Dokkum, P. G., et al. 2011, *ApJ*, **739**, 24
- Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, *MNRAS*, **351**, 1151
- Bruzual, G., & Charlot, S. 2003, *MNRAS*, **344**, 1000
- Buat, V., Giovannoli, E., Heinis, S., et al. 2011, *A&A*, **533**, A93
- Buat, V., Heinis, S., Boquien, M., et al. 2014, *A&A*, **561**, A39
- Buat, V., Iglesias-Páramo, J., Seibert, M., et al. 2005, *ApJL*, **619**, L51
- Buat, V., Marcellac, D., Burgarella, D., et al. 2007, *A&A*, **469**, 19
- Buat, V., Noll, S., Burgarella, D., et al. 2012, *A&A*, **545**, A141
- Burgarella, D., Buat, V., Gruppioni, C., et al. 2013, *A&A*, **554**, A70
- Calzetti, D. 2001, *PASP*, **113**, 1449
- Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, *ApJ*, **533**, 682
- Capak, P., Aussel, H., Ajiki, M., et al. 2007, *ApJS*, **172**, 99
- Capak, P., Cowie, L. L., Hu, E. M., et al. 2004, *AJ*, **127**, 180
- Cardamone, C. N., Urry, C. M., Schawinski, K., et al. 2010, *ApJL*, **721**, L38
- Carilli, C. L., Lee, N., Capak, P., et al. 2008, *ApJ*, **689**, 883
- Casey, C. M., Berta, S., Béthermin, M., et al. 2012, *ApJ*, **761**, 140
- Castellano, M., Fontana, A., Grazian, A., et al. 2012, *A&A*, **540**, A39
- Castellano, M., Sommariva, V., Fontana, A., et al. 2014, *A&A*, **566**, A19
- Chabrier, G. 2003, *PASP*, **115**, 763
- Chary, R., & Elbaz, D. 2001, *ApJ*, **556**, 562
- Chary, R.-R., & Pope, A. 2010, *ApJ*, submitted (arXiv:1003.1731)
- Cimatti, A., Daddi, E., & Renzini, A. 2006, *A&A*, **453**, L29
- Cisternas, M., Jahnke, K., Bongiorno, A., et al. 2011, *ApJL*, **741**, L11
- Condon, J. J. 1992, *ARA&A*, **30**, 575
- Cortese, L., Boselli, A., Buat, V., et al. 2006, *ApJ*, **637**, 242
- Cullen, F., Cirasuolo, M., McLure, R. J., Dunlop, J. S., & Bowler, R. A. 2014, *MNRAS*, **440**, 2300
- Daddi, E., Cimatti, A., Renzini, A., et al. 2004, *ApJ*, **617**, 746
- Daddi, E., Dannerbauer, H., Stern, D., et al. 2009, *ApJ*, **694**, 1517
- Daddi, E., Dickinson, M., Morrison, G., et al. 2007, *ApJ*, **670**, 156
- de Barros, S., Schaerer, D., & Stark, D. P. 2014, *A&A*, **563**, A81
- Del Moro, A., Alexander, D. M., Mullaney, J. R., et al. 2013, *A&A*, **549**, A59
- Dickinson, M., Papovich, C., Ferguson, H. C., & Budavári, T. 2003, *ApJ*, **587**, 25
- Domínguez, A., Siana, B., Henry, A. L., et al. 2013, *ApJ*, **763**, 145
- Drory, N., Bender, R., & Hopp, U. 2004, *ApJL*, **616**, L103
- Drory, N., Bundy, K., Leauthaud, A., et al. 2009, *ApJ*, **707**, 1595
- Drory, N., Salvato, M., Gabasch, A., et al. 2005, *ApJL*, **619**, L131
- Elbaz, D., Daddi, E., Le Borgne, D., et al. 2007, *A&A*, **468**, 33
- Elbaz, D., Dickinson, M., Hwang, H. S., et al. 2011, *A&A*, **533**, A119
- Elbaz, D., Hwang, H. S., Magnelli, B., et al. 2010, *A&A*, **518**, L29
- Erb, D. K., Steidel, C. C., Shapley, A. E., et al. 2006, *ApJ*, **647**, 128
- Ferguson, H. C., Dickinson, M., Giavalisco, M., et al. 2004, *ApJL*, **600**, L107
- Finkelstein, S. L., Papovich, C., Salmon, B., et al. 2012, *ApJ*, **756**, 164
- Fitzpatrick, E. L. 1999, *PASP*, **111**, 63
- Fontana, A., Salimbeni, S., Grazian, A., et al. 2006, *A&A*, **459**, 745
- Förster Schreiber, N. M., Genzel, R., Bouché, N., et al. 2009, *ApJ*, **706**, 1364
- Gabasch, A., Goranova, Y., Hopp, U., Noll, S., & Pannella, M. 2008, *MNRAS*, **383**, 1319
- Garn, T., & Best, P. N. 2010, *MNRAS*, **409**, 421
- Garn, T., Sobral, D., Best, P. N., et al. 2010, *MNRAS*, **402**, 2017
- Gilbank, D. G., Balogh, M. L., Glazebrook, K., et al. 2010, *MNRAS*, **405**, 2419
- González, V., Labbé, I., Bouwens, R. J., et al. 2010, *ApJ*, **713**, 115
- Grasha, K., Calzetti, D., Andrews, J. E., Lee, J. C., & Dale, D. A. 2013, *ApJ*, **773**, 174
- Heckman, T. M., Robert, C., Leitherer, C., Garnett, D. R., & van der Rydt, F. 1998, *ApJ*, **503**, 646
- Heinis, S., Buat, V., Béthermin, M., et al. 2013, *MNRAS*, **429**, 1113
- Heinis, S., Buat, V., Béthermin, M., et al. 2014, *MNRAS*, **437**, 1268
- Helou, G., Soifer, B. T., & Rowan-Robinson, M. 1985, *ApJL*, **298**, L7
- Hopkins, A. M., & Beacom, J. F. 2006, *ApJ*, **651**, 142
- Hwang, H. S., Elbaz, D., Magdis, G., et al. 2010, *MNRAS*, **409**, 75
- Ibar, E., Ivison, R. J., Biggs, A. D., et al. 2009, *MNRAS*, **397**, 281
- Ibar, E., Sobral, D., Best, P. N., et al. 2013, *MNRAS*, **434**, 3218
- Ilbert, O., McCracken, H. J., Le Fèvre, O., et al. 2013, *A&A*, **556**, A55
- Ilbert, O., Salvato, M., Le Floc'h, E., et al. 2010, *ApJ*, **709**, 644
- Ivison, R. J., Magnelli, B., Ibar, E., et al. 2010, *A&A*, **518**, L31
- Juneau, S., Dickinson, M., Bournaud, F., et al. 2013, *ApJ*, **764**, 176
- Kajisawa, M., Ichikawa, T., Tanaka, I., et al. 2011, *PASJ*, **63**, 379
- Karim, A., Schinnerer, E., Martínez-Sansigre, A., et al. 2011, *ApJ*, **730**, 61
- Kashino, D., Silverman, J. D., Rodighiero, G., et al. 2013, *ApJL*, **777**, L8
- Kennicutt, R. C., Jr. 1998, *ARA&A*, **36**, 189
- Kirkpatrick, A., Pope, A., Alexander, D. M., et al. 2012, *ApJ*, **759**, 139
- Kong, X., Charlot, S., Brinchmann, J., & Fall, S. M. 2004, *MNRAS*, **349**, 769
- Kreckel, K., Groves, B., Schinnerer, E., et al. 2013, *ApJ*, **771**, 62
- Kriek, M., & Conroy, C. 2013, *ApJL*, **775**, L16
- Kriek, M., van Dokkum, P. G., Labbé, I., et al. 2009, *ApJ*, **700**, 221
- Kurczynski, P., & Gawiser, E. 2010, *AJ*, **139**, 1592
- Law, K.-H., Gordon, K. D., & Misselt, K. A. 2011, *ApJ*, **738**, 124
- Lilly, S. J., Carollo, C. M., Pipino, A., Renzini, A., & Peng, Y. 2013, *ApJ*, **772**, 119
- Lin, L., Dickinson, M., Jian, H.-Y., et al. 2012, *ApJ*, **756**, 71
- Madau, P., & Dickinson, M. 2014, arXiv:1403.0007
- Magdis, G. E., Daddi, E., Béthermin, M., et al. 2012, *ApJ*, **760**, 6
- Magdis, G. E., Elbaz, D., Daddi, E., et al. 2010, *ApJ*, **714**, 1740
- Magnelli, B., Lutz, D., Saintonge, A., et al. 2014, *A&A*, **561**, A86
- Magnelli, B., Lutz, D., Santini, P., et al. 2012, *A&A*, **539**, A155
- Maier, C., Lilly, S. J., Ziegler, B., et al. 2014, *ApJ*, in press (arXiv:1406.6069)
- Maiolino, R., Schneider, R., Oliva, E., et al. 2004, *Natur*, **431**, 533
- Mancini, C., Förster Schreiber, N. M., Renzini, A., et al. 2011, *ApJ*, **743**, 86
- Mannucci, F., Cresci, G., Maiolino, R., Marconi, A., & Gnerucci, A. 2010, *MNRAS*, **408**, 2115
- Maraston, C., Pforr, J., Renzini, A., et al. 2010, *MNRAS*, **407**, 830
- Marchesini, D., van Dokkum, P. G., Förster Schreiber, N. M., et al. 2009, *ApJ*, **701**, 1765
- Meurer, G. R., Heckman, T. M., & Calzetti, D. 1999, *ApJ*, **521**, 64
- Momcheva, I. G., Lee, J. C., Ly, C., et al. 2013, *AJ*, **145**, 47
- Morrison, G. E., Owen, F. N., Dickinson, M., Ivison, R. J., & Ibar, E. 2010, *ApJS*, **188**, 178
- Mullaney, J. R., Alexander, D. M., Goulding, A. D., & Hickox, R. C. 2011, *MNRAS*, **414**, 1082
- Mullaney, J. R., Pannella, M., Daddi, E., et al. 2012, *MNRAS*, **419**, 95
- Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013, *ApJ*, **777**, 18
- Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007, *ApJL*, **660**, L43
- Nordon, R., Lutz, D., Saintonge, A., et al. 2013, *ApJ*, **762**, 125
- Oteo, I., Bongiovanni, Á., Magdis, G., et al. 2014, *MNRAS*, **439**, 1337
- Oteo, I., Magdis, G., Bongiovanni, Á., et al. 2013, *MNRAS*, **435**, 158
- Ouchi, M., Ellis, R., Ono, Y., et al. 2013, *ApJ*, **778**, 102
- Ouchi, M., Mobasher, B., Shimasaku, K., et al. 2009, *ApJ*, **706**, 1136
- Overzier, R. A., Heckman, T. M., Wang, J., et al. 2011, *ApJL*, **726**, L7
- Pannella, M., Carilli, C. L., Daddi, E., et al. 2009a, *ApJL*, **698**, L116
- Pannella, M., Elbaz, D., & Daddi, E. 2013, in IAU Symp. 292, GOODS-Herschel: Dust Attenuation up to $z \sim 4$, ed. T. Wong & J. Ott (Cambridge: Cambridge Univ. Press), **289**
- Pannella, M., Gabasch, A., Goranova, Y., et al. 2009b, *ApJ*, **701**, 787
- Pannella, M., Hopp, U., Saglia, R. P., et al. 2006, *ApJL*, **639**, L1
- Papovich, C., Finkelstein, S. L., Ferguson, H. C., Lotz, J. M., & Giavalisco, M. 2011, *MNRAS*, **412**, 1123
- Papovich, C., Moustakas, L. A., Dickinson, M., et al. 2006, *ApJ*, **640**, 92
- Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2002, *AJ*, **124**, 266
- Peng, Y.-J., Lilly, S. J., Kovač, K., et al. 2010, *ApJ*, **721**, 193
- Penner, K., Dickinson, M., Pope, A., et al. 2012, *ApJ*, **759**, 28
- Pérez-González, P. G., Rieke, G. H., Villar, V., et al. 2008, *ApJ*, **675**, 234
- Popesso, P., Magnelli, B., Buttiglion, S., et al. 2012, *A&A*, submitted (arXiv:1211.4257)
- Price, S. H., Kriek, M., Brammer, G. B., et al. 2014, *ApJ*, **788**, 86
- Reddy, N., Dickinson, M., Elbaz, D., et al. 2012, *ApJ*, **744**, 154
- Reddy, N. A., Erb, D. K., Pettini, M., Steidel, C. C., & Shapley, A. E. 2010, *ApJ*, **712**, 1070
- Reddy, N. A., & Steidel, C. C. 2009, *ApJ*, **692**, 778
- Renzini, A. 2009, *MNRAS*, **398**, L58
- Renzini, A., & Peng, Y.-J. 2015, *ApJL*, **801**, L29
- Rodighiero, G., Cimatti, A., Gruppioni, C., et al. 2010, *A&A*, **518**, L25
- Rodighiero, G., Daddi, E., Baronchelli, I., et al. 2011, *ApJL*, **739**, L40
- Rodighiero, G., Renzini, A., Daddi, E., et al. 2014, *MNRAS*, **443**, 19
- Rosario, D. J., Mozena, M., Wuyts, S., et al. 2013, *ApJ*, **763**, 59
- Salim, S., Charlot, S., Rich, R. M., et al. 2005, *ApJL*, **619**, L39
- Salim, S., Rich, R. M., Charlot, S., et al. 2007, *ApJS*, **173**, 267
- Salmi, F., Daddi, E., Elbaz, D., et al. 2012, *ApJL*, **754**, L14
- Salpeter, E. E. 1955, *ApJ*, **121**, 161
- Salvato, M., Hasinger, G., Ilbert, O., et al. 2009, *ApJ*, **690**, 1250

- Santini, P., Fontana, A., Grazian, A., et al. 2009, [A&A](#), **504**, 751
- Santini, P., Rosario, D. J., Shao, L., et al. 2012, [A&A](#), **540**, A109
- Sargent, M. T., Béthermin, M., Daddi, E., & Elbaz, D. 2012, [ApJL](#), **747**, L31
- Sargent, M. T., Schinnerer, E., Murphy, E., et al. 2010, [ApJS](#), **186**, 341
- Savaglio, S., Glazebrook, K., Le Borgne, D., et al. 2005, [ApJ](#), **635**, 260
- Schaerer, D., & de Barros, S. 2010, [A&A](#), **515**, A73
- Schreiber, C., Pannella, M., Elbaz, D., et al. 2015, [A&A](#), **575**, A74
- Sobral, D., Best, P. N., Matsuda, Y., et al. 2012, [MNRAS](#), **420**, 1926
- Stark, D. P., Ellis, R. S., Bunker, A., et al. 2009, [ApJ](#), **697**, 1493
- Steidel, C. C., Rudie, G. C., Strom, A. L., et al. 2014, [ApJ](#), **795**, 165
- Strazzullo, V., Gobat, R., Daddi, E., et al. 2013, [ApJ](#), **772**, 118
- Strazzullo, V., Pannella, M., Owen, F. N., et al. 2010, [ApJ](#), **714**, 1305
- Symeonidis, M., Vaccari, M., Berta, S., et al. 2013, [MNRAS](#), **431**, 2317
- Tan, Q., Daddi, E., Sargent, M., et al. 2013, [ApJL](#), **776**, L24
- Thomas, D., Maraston, C., Bender, R., & Mendes de Oliveira, C. 2005, [ApJ](#), **621**, 673
- Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, [ApJ](#), **613**, 898
- Troncoso, P., Maiolino, R., Sommariva, V., et al. 2014, [A&A](#), **563**, A58
- Viero, M. P., Monceli, L., Quadri, R. F., et al. 2013, [ApJ](#), **779**, 32
- Wang, W.-H., Cowie, L. L., Barger, A. J., Keenan, R. C., & Ting, H.-C. 2010, [ApJS](#), **187**, 251
- Whitaker, K. E., Labbé, I., van Dokkum, P. G., et al. 2011, [ApJ](#), **735**, 86
- Whitaker, K. E., van Dokkum, P. G., Brammer, G., & Franx, M. 2012, [ApJL](#), **754**, L29
- Wild, V., Charlot, S., Brinchmann, J., et al. 2011, [MNRAS](#), **417**, 1760
- Williams, R. J., Quadri, R. F., Franx, M., van Dokkum, P., & Labbé, I. 2009, [ApJ](#), **691**, 1879
- Wuyts, E., Kurk, J., Förster Schreiber, N. M., et al. 2014, [AARA](#), **52**, 415
- Wuyts, S., Förster Schreiber, N. M., Nelson, E. J., et al. 2013, [ApJ](#), **779**, 135
- Wuyts, S., Förster Schreiber, N. M., van der Wel, A., et al. 2011, [ApJ](#), **742**, 96
- Wuyts, S., Labbé, I., Franx, M., et al. 2007, [ApJ](#), **655**, 51
- Yun, M. S., Reddy, N. A., & Condon, J. J. 2001, [ApJ](#), **554**, 803
- Zahid, H. J., Geller, M. J., Kewley, L. J., et al. 2013a, [ApJL](#), **771**, L19
- Zahid, H. J., Yates, R. M., Kewley, L. J., & Kudritzki, R. P. 2013b, [ApJ](#), **763**, 92