

X-ray stacking of Lyman break galaxies in the 4 Ms CDF-S

X-ray luminosities and star formation rates across cosmic time

P.-C. Zinn^{1,2}, S. Blex¹, N. Seymour², and D. J. Bomans¹

¹ Astronomisches Institut Ruhr-Universität Bochum, Universitätsstr. 150, 44801 Bochum, Germany
e-mail: zinn@astro.rub.de

² CSIRO Astronomy & Space Science, PO Box 76, Epping, NSW, 1710, Australia

Received 25 June 2012 / Accepted 3 October 2012

ABSTRACT

Context. Lyman break galaxies (LBGs) are widely thought to be prototypical young galaxies in the early universe, particularly representative of those undergoing massive events of star formation. Therefore, LBGs should produce significant amounts of X-ray emission.

Aims. We aim to trace the X-ray luminosity of LBGs across cosmic time and from that derive constraints on their star formation history.

Methods. We utilize the newly released 4 Ms mosaic obtained with the *Chandra* X-ray Observatory, the deepest X-ray image to date, alongside with the superb spectroscopic data sets available in the CDF-S survey region to construct large but nearly uncontaminated samples of LBGs across a wide range of redshift ($0.5 < z < 4.5$) which can be used as input samples for stacking experiments. This approach allows us to trace the X-ray emission of LBGs to even lower, previously unreachable, flux density limits ($\sim 10^{-18}$ mW m⁻²) and therefore to larger redshifts.

Results. We reliably detect soft-band X-ray emission from all our input redshift bins except for the highest redshift ($z \sim 4$) one. From that we derive rest-frame 2–10 keV luminosities and infer star formation rates and stellar masses. We find that star formation in LBGs peaks at a redshift of $z_{\text{peak}} \approx 3.5$ and then decreases quickly. We also see a characteristic peak in the specific star formation rate ($sSFR = SFR/M_*$) at this redshift. Furthermore, we calculate the contribution of LBGs to the total cosmic star formation rate density (SFRD) and find that the contribution of LBGs is negligible. Therefore, we conclude that most of the star formation in the early universe takes place in lower luminosity galaxies as suggested by hierarchical structure formation models.

Key words. methods: data analysis – galaxies: evolution – galaxies: high-redshift – X-rays: galaxies

1. Introduction

Lyman break galaxies (LBGs) are largely considered as the bright end of the distribution of normal star-forming galaxies across a wide range of redshift. They exhibit a significant scatter in terms of mass (Mannucci et al. 2009), with both a pronounced low-mass (Weatherley & Warren 2003) and high-mass (Barmby et al. 2004) fraction among the entire LBG population. Since they are – thanks to their exceptionally strong star formation activity – relatively easy to select at various redshift ranges from photometric observations only, they have been playing a key role in galaxy evolution studies all over the last two decades. This is foremost due to the very effective “drop out” selection technique established in the early 1990s (see e.g. Steidel & Hamilton 1993; Steidel et al. 1996, 1999). This technique utilizes the strong absorption of all light emitted blueward the rest-frame wavelength of the Lyman limit at 912 Å. Because this produces a very pronounced step in the typical spectrum of an LBG, one can select such objects by searching deep imaging data for sources detected only in longer wavelength filters but not in short wavelength filters. This behavior of a source “dropping out” from being detected below a certain wavelength is nowadays the major tool for selecting candidate high-redshift sources. Since its first application in the early 1990s, the technique underwent a massive development and heavy usage throughout the community (review by Giavalisco 2002). Therefore, large samples of LBGs

across the entire redshift range from $z = 1$ (Burgarella et al. 2007; Habertzettl et al. 2009; Basu-Zych et al. 2011) to $z > 7$ (Bouwens et al. 2010, 2011b; Stark et al. 2011) are publicly available, typically comprising hundreds of objects in a comprehensive form. Even at the highest redshifts, considerable work is being done utilizing forefront equipment such as the new Wide-Field Camera 3 (WFC3) aboard the *Hubble* Space Telescope (HST). Therefore, LBG candidates with redshifts as large as $z = 10$ are being discovered (Bouwens et al. 2011a), but because of their extreme faintness ($H_{\text{AB}} = 28.9$) lacking spectroscopic confirmation which has only been done for LBGs up to $z \sim 7$ (Vanzella et al. 2011). There is also considerable work characterizing LBGs in terms of their environment (Tasker & Bryan 2006; Cooke et al. 2010).

Since LBGs are traditionally selected in the optical wavelength regime, recently extending to the near-infrared (NIR) as the WFC3 aboard HST became operational, ancillary observations are necessary to characterize these objects over the entire electromagnetic spectrum. This has been done by several groups, successfully detecting individual LBGs at moderate redshifts ($z < 3$) also in the mid- and far-infrared (MIR and FIR) regimes (Rigopoulou et al. 2006; Magdis et al. 2008; Burgarella et al. 2011) thanks to other space-based facilities such as the *Spitzer* Space Telescope and, recently, the *Herschel* Space Observatory. At even longer wavelength, ground-based observations in the sub-millimeter and radio regimes have been

conducted, yielding only a few detections of individual objects at $850\ \mu\text{m}$ (e.g. Chapman & Casey 2009) whereas there is no direct detection of a significantly redshifted LBG to date except for strongly lensed systems such as the “8 o’clock arc”, for instance (Volino et al. 2010).

To access even more redshifted LBGs over the entire spectrum, stacking techniques have been successfully applied by many groups. For instance, the radio properties of LBGs have been discussed with respect to their star formation activity by Carilli et al. (2008), utilizing the deep Very Large Array (VLA) 1.4 GHz (or 20 cm) observations of the COSMOS field. Similar stacking investigations of the star formation history of a sample of very high redshift ($z > 7$) LBGs were done by Labbé et al. (2010) in the NIR regime, utilizing ultra-deep *Spitzer* data in the *Chandra* Deep Field South (CDF-S). This particular survey field is also extremely valuable to study LBGs at the shortest wavelength since it comprises the deepest X-ray observations obtained so far. However, with less deep data significant work on stacking LBGs in the X-ray regime has already been done. A first attempt has been done already Brandt et al. (2001) in the *Chandra* Deep Field North (CDF-N) with a 1 Ms exposure. Stacking a sample of only 24 LBGs with redshifts between $2 \leq z \leq 4$, they found a soft-band signal at a significance level of 99.9%. From that, they calculated an average X-ray luminosity of $3.2 \times 10^{41}\ \text{erg s}^{-1}$ in the rest-frame 2–8 keV band, comparable to the most luminous local starburst galaxies such as NGC 3256 (Moran et al. 1999). A similar result was obtained by Nandra et al. (2002) who extended the small sample of Brandt et al. to a statistically more robust number of 148 LBGs. With this larger sample, they were able to exclude LBGs containing active galactic nuclei (AGN) and hence giving an estimate for the star formation rate (SFR) of an average LBG at $z \sim 3$ of about $60 M_{\odot}\ \text{yr}^{-1}$. A next step was done by Lehmer et al. (2005) utilizing *Chandra* data from both the Great Observatories Origins Deep Survey (GOODS) north and south fields with exposure times of 2 Ms and 1 Ms, respectively. With a large sample of LBGs comprising a few thousand galaxies, they were able to do their analysis in different redshift bins in the range $3 \leq z \leq 6$, detecting a significant signal up to $z = 4$. They found the average SFR at this redshift decreases to $10\text{--}30 M_{\odot}\ \text{yr}^{-1}$, and even further to higher redshifts since they were not able to make statistically significant detections at any redshift greater than $z = 4$. A complementary result was obtained by Laird et al. (2006) who identified a sample of direct X-ray detections of LBGs at intermediate redshifts ($z \sim 3$), supporting the notion that LBGs are the “tip of the iceberg” in terms of star formation in their respective epoch. The statistical X-ray detection of high-redshift LBGs ($z > 5$) was attempted for the first time by Cowie et al. (2012) utilizing photometrically selected samples of LBGs comprising several thousand sources in the CDF-S. Although with such a large number of stacked sources, they were not able to detect X-rays from the high- z LBG population, placing an upper limit of $4 \times 10^{41}\ \text{erg s}^{-1}$ on the X-ray luminosity in the rest-frame 4–15 keV band at $z = 6.5$.

With this paper, we attempt to widen the redshift range of X-ray studies of LBGs down to $z = 1$ since space-based facilities such as GALEX and *Swift* now provide reliable LBG samples selected in the ultraviolet (UV). Furthermore, we want to trace the star formation activity and stellar mass build-up of the average LBG from $z = 1$ onwards.

Throughout this paper, we adopt a standard flat Λ CDM cosmology with $H_0 = 70\ \text{km s}^{-1}\ \text{Mpc}^{-1}$ and $\Omega_{\Lambda} = 0.73$ (Komatsu et al. 2011).

2. X-ray data and LBG selection

In this paper, we use the 4 Ms *Chandra* Deep Field South data¹ as basis for the X-ray stacking procedure. These data comprise mosaics of an area of $464.5\ \text{arcmin}^2$ in both the soft (0.5–2 keV) and the hard (2–8 keV) X-ray bands. With an effective exposure time of 3822 ks, this is the the deepest X-ray data set available to date, reaching an on-axis flux limit of $9.1 \times 10^{-18}\ \text{erg s}^{-1}\ \text{cm}^{-2}$ and $5.5 \times 10^{-17}\ \text{erg s}^{-1}\ \text{cm}^{-2}$ in the soft and hard band (Xue et al. 2011).

To create a parent list of LBGs to be stacked, we use the extreme wealth of spectroscopic data available in the CDF-S. To use only spectroscopically confirmed LBGs assures both a clean list with only marginal contamination by mis-classified objects, e.g. lower redshift interlopers, as well as accurate redshifts for the construction of several redshift bins. In particular, we used the VLT/VIMOS spectroscopic surveys (Popesso et al. 2009) and Balestra et al. (2010) as well as the VLT/FORS survey by Vanzella et al. (2005, 2006, 2008). These catalogs provide a detailed spectroscopic classification and are therefore ideal for reliably selecting LBGs. The main photometric selection criterion was based on a U - and B -band drop-out search utilizing a color cut $U - B \geq 1.2$ or $B - V \geq 1.2$, respectively, where the red part of the spectrum must stay flat (e.g. $V - z \leq 1.2$). The main spectral feature used in these surveys for deeming a source an LBG is the characteristic break in their spectrum blueward of the Lyman Limit at rest-frame $912\ \text{\AA}$ which was assessed by comparison to various template spectra of LBGs and other, possibly contaminating, source types. The spectral classifications provided by these surveys are very detailed, allowing e.g. for a split in Ly α emitters and absorbers. However, in order not to bias our samples in a certain way, we chose to not use this criterion but solely select LBGs (whether they are Ly α emitters or absorbers) in different redshift bins. Furthermore, we used only such objects with off-axis angles smaller than $10''$ with respect to the *Chandra*/ACIS pointings in order to not suffer from the strongly varying PSF at greater off-axis angles. For a complete overview of our input list see Table 1.

To extend our LBG sample to redshifts lower than $z = 2.5$, the approximate limit for the spectroscopic classification of LBGs using optical spectra, we had to take a closer look into the UV data of the CDF-S because at these low redshifts, the Lyman break occurs in the near UV range. Using the well-known dropout technique to identify LBG candidates, Basu-Zych et al. (2011) compiled a list of candidates with data from the *Swift* Gamma-ray observatory’s UV/optical telescope (*Swift*-UVOT). The *Swift* satellite, although dedicated to high energy astrophysics, is equipped with a small (30 cm diameter) telescope sensitive to wavelength between $170\ \mu\text{m}$ and $650\ \mu\text{m}$ which observed an area of $266\ \text{arcmin}^2$ in the CDF-S for approximately 60 ks, reaching a limiting magnitude in the U -band of 24.5 AB-mag. These observations lead to a list of candidate LBGs with 43 objects, spanning the desired redshift range. This sample, dubbed BZ-UV, is also listed in Table 1. Because their selection mechanism is based only on photometric data, we point out that the reliability of these objects is not comparable to the spectroscopic lists we use for higher redshift objects. Therefore, and because there are much less LBG candidates identified via UV observations, we expect our stacking results in this redshift range to be significantly less robust than in the other redshift bins.

¹ <http://www2.astro.psu.edu/users/niel/cdfs/cdfs-chandra.html>

Table 1. LBG selection parent tables and redshift bins.

ID	Parent list	Limit ^a	Redshift range	Median z	Age ^b [Gyr]	# Selected objects	# Stacked objects
ref-lowZ	Wuyts et al. (2008)	$R < 24$	$0.5 \leq z \leq 1.0$	0.68	7.52	143	141
BZ-UV	Basu-Zych et al. (2011)	$u < 24.5$	$1.0 \leq z \leq 2.0$	1.33	4.82	43	35
Spec-U1	VIMOS and FORS ^c	$R < 24.5$	$2.5 \leq z \leq 3.0$	2.67	2.49	201	201
Spec-U2	VIMOS and FORS ^c	$R < 24.5$	$3.0 \leq z \leq 3.5$	3.33	1.95	143	92
Spec-B1	VIMOS and FORS ^c	$R < 24.5$	$3.5 \leq z \leq 4.0$	3.70	1.73	63	19
Spec-B2	VIMOS and FORS ^c	$R < 24.5$	$4.0 \leq z \leq 4.5$	4.16	1.50	39	34

Notes. ^(a) Typical magnitude limit of parent spectroscopic survey in AB system. Since Wuyts et al. (2008) collect a variety of spectroscopic redshifts from a large number of surveys targeting the CDF-S, the quoted limiting magnitude is just indicative. The redshifts for the Basu-Zych et al. (2011) sample are photometric estimates, therefore the limit of their survey is given. The VIMOS and FORS campaign targets were chosen not only based on a magnitude limit but also according to the color criteria explained further in the text. ^(b) Time since the Big Bang according to the cosmology defined in Sect. 1 evaluated at median redshift of the bin. ^(c) Popesso et al. (2009), Balestra et al. (2010), Vanzella et al. (2008).

We point out that, due to the plentitude of spectroscopic campaigns targeting the CDF-S region, it is not feasible to establish a single magnitude limit down to which all surveys are complete and hence down to which our samples are complete. The available redshifts are hence covering a heterogeneous family of objects from which we aimed to select LBGs only with no pretense of being complete.

We have taken special care in order to avoid any contamination by X-ray emission from a faint AGN that may be hosted by our stacked sources. The rejection of possible AGN hosts was done in four stages: (i) the optical spectra of all sources in the input lists were examined both by eye and according to the classical BPT diagnostic diagram for AGN activity (Baldwin et al. 1981). Any source with a marginal sign for AGN activity was discarded from the list; (ii) we used the deep *Spitzer* photometry in the CDF-S (GOODS project, see <http://irsa.ipac.caltech.edu/data/GOODS/>) to examine our input LBGs in the mid-infrared color-color diagram according to the criteria from Stern et al. (2005) and excluded all sources lying in the AGN regime of this diagnostic diagram; (iii) we cross-correlated our input source lists to the VLA 20 cm observations by Miller et al. (2008) and plotted them against the *Spitzer*/MIPS 24 μ m flux densities to obtain a measure for the radio–infrared correlation amongst our sources. Such sources that deviate from this correlation as defined in Mao et al. (2011) by showing a more than a 5-fold excess in radio emission most likely originating from an AGN were discarded. This rigorous AGN rejection is crucial for the subsequent stacking since contamination from X-rays produced by a hidden AGN would bias our analysis towards higher SFR. An alternative approach for accounting for possible AGN contamination, e.g. extrapolating the X-ray luminosity function of AGN (see e.g. Ebrero et al. 2009) to obtain an estimated contaminating fraction for the investigated sensitivity range was not used because we probe much higher sensitivities (thanks to the 4 Ms exposure time and the stacking of several hundred objects) then most studies before, therefore the extrapolation would be highly inaccurate.

In order to compare our findings for SFR which will be calculated from the stacked X-ray luminosities, we use data of our selected objects from other survey projects in the CDF-S. Primarily, the COMBO-17 survey (Wolf et al. 2003) is used for its broad range of data available over a large area. A deeper but more narrow view (in terms of covered area) is introduced by the use of data from the GEMS catalog (Rix et al. 2004) with HST observations in two filters and the FIREWORKS (Wuyts et al. 2008) catalog, containing deep data across a wide range of the electromagnetic spectrum (UV to MIR) but for only a

fraction of our selected LBGs. In particular, we used the optical and NIR ancillary data to infer parameters such as stellar mass for all objects in our stacks. Furthermore, FIREWORKS was used to construct a low redshift control sample of galaxies. This sample should cover the very low redshift end in our analysis and be consistent of galaxies of all kinds to obtain a completely unbiased view of X-ray emission from typical galaxies to be later compared to our LBG results. To compile this reference sample (dubbed ref-lowZ in Table 1), we only applied a redshift cut of $0.5 \leq z \leq 1.0$ to all objects classified as galaxy in the COMBO-17 catalog that also have FIREWORKS counterparts with reliable spectroscopic redshifts. This left us with a fairly large sample of more than 200 objects, therefore we expect the stacking to deliver an unambiguous X-ray detection.

3. The stacking procedure

Before starting the stacking process for the various source samples defined above, special care was taken to exclude individually detected sources from the samples. Therefore, in a first step, we cross-matched all our sources as summarized in Table 1 to the *Chandra* 4 Ms source catalog by Xue et al. (2011) and excluded the 13 sources that have X-ray counterparts in this catalog which are mostly lower- z objects. For all remaining sources, $10'' \times 10''$ (resp. 20×20 pixels) cutout images from the 4 Ms soft-band mosaic centered on the (optical or UV derived) source position were created and inspected by eye whether there is a clear but uncatalogued X-ray counterpart. This was only true for two sources which were then also discarded from the further process.

This left us with six source lists (according to our six redshift bins from Table 1) containing only objects without unambiguously detected X-ray counterparts to start the actual stacking algorithm with.

3.1. The stacking algorithm

There has already been some effort in doing X-ray stacking of faint sources as e.g. presented in Nandra et al. (2002), and also stacking procedures in other wavelength ranges, e.g. radio as described by Carilli et al. (2008), are known. A recent attempt also using the new 4 Ms *Chandra* CDF-S mosaic is presented in Cowie et al. (2012). They attempt to trace the X-ray emission of galaxies out to $z = 8$ by utilizing high- z galaxy samples compiled from new HST Wide Field Camera 3 (WFC 3) observations of the *Hubble* Ultra-Deep Field (HUDF) which is part of the CDF-S and applying a weighted mean stacking algorithm based on various quantities (off-axis angle, aperture radii for flux extraction,

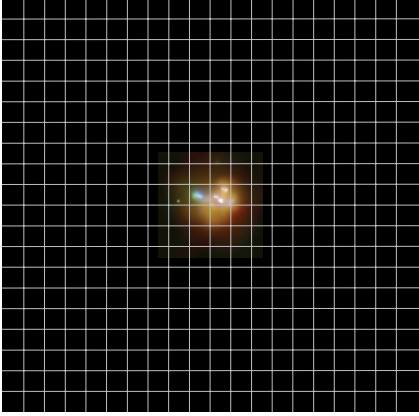


Fig. 1. Local LBG analog VV 114 as seen by *Chandra* (color composite of soft- and hard-band images from Grimes et al. 2006). The overlaid grid indicates the pixel size of the ACIS detector if VV 114 would be at $z = 3$ and our 20×20 pixels stack-image layout.

exact model of noise in aperture) not related to the sources themselves but to the technical layout of the X-ray observations and the reduction process. This technique may enhance the resultant signal-to-noise ratio (S/N) of the final stacked images, it bears the risk of introducing biases or unconstrained statistical effects just because of its complexity. This is discussed in more detail e.g. in Lehmer et al. (2005) and Hickox & Markevitch (2007). Therefore, we chose to apply a much simpler and straightforward stacking method that, on the one hand, guarantees to not introduce any biases but for the eventual cost of S/N in the final stacked images. Our stacking routine works as follows.

We first consider the typical angular size of a galaxy at various redshifts to determine a reasonable size for the images to be stacked. Since within our adopted cosmology the linear scale at every redshift $z > 1$ is nearly constant with $7 \text{ kpc}''$ (Wright 2006), the largest galaxies (determined from their optical morphologies as e.g. by Trujillo et al. 2006) at these redshifts with a diameter of $\sim 10 \text{ kpc}$ would stretch over about $1.5''$ or three pixels given the pixel scale of *Chandra's* ACIS detector. Therefore, we chose to extract fluxes from a 5 pixel diameter aperture to also account for astrometric inconsistencies between the parent galaxy catalogs and the 4 Ms mosaic. Note that this astrophysical motivation of an aperture size is very much different from other stacking studies such as e.g. Nandra et al. (2002); Lehmer et al. (2005). They defined their aperture sizes empirically by testing various diameters and then adopting the one yielding the best S/N. Despite the fact that those two methods are fundamentally different, the final aperture diameters both our astrophysical and the empirical approach yield are in very good agreement. To finally get a measure of the background around every source, we chose the stacking images to be 20 by 20 pixels, four times the aperture diameter. The layout of our images that go into the actual stacking procedure is shown in Fig. 1.

Before actually stacking these 20×20 pixels images, our algorithm sorts out all images that (i) show an integrated flux within the pre-defined 5 pixel aperture (centered on the image center) that exceeds three times the median absolute deviation (MAD) of the noise in the remaining part of the image and (ii) all images with a total flux across the entire image that is zero. The first step in our algorithm is to reject all marginally detected sources to avoid any contamination of such sources that are neither in the source catalog by Xue et al. (2011) nor visible by eye. This can be due to e.g. AGN which are faint at optical and infrared wavelength and therefore are not visible the FORS

or VIMOS spectra (see e.g. Zinn et al. 2011). Since at a redshift $z = 3$ the X-ray luminosity necessary to produce an actual detection in the 4 Ms mosaic is equivalent to a SFR of about $200 M_{\odot} \text{ yr}^{-1}$, an actual X-ray detection is a strong hint for additional AGN activity because LBGs are in general not harboring such extreme starbursts. The second criterion is applied because the 4 Ms mosaic is smaller than the region for which spectroscopic observations have been carried out. Hence, there may be sources in our list that do not have X-ray coverage which should of course be sorted out in order not to contaminate the final stack.

To also exclude X-ray detected sources that are just by chance within our stacking region, we define a second aperture of 10 pixels diameter again centered on the image center. The remaining part of the image is used to determine the background noise level. If the total flux measured within this 10 pixel aperture indicates a detection with a significance of more than 90% with respect to the measured local background, it is rejected for the actual stacking process. This second rejection step is necessary to remove X-ray sources that are close to the stacking location but are not associated with the LBG that should go into the stack. Because those unrelated sources could add flux to the final stack which is not physically related to the LBGs we want to examine, this second rejection step is crucial in terms of a clean stacking sample. Table 1 summarizes the number of sources selected for the six input coordinate lists and the number of sources actually stacked after this rejection step.

The actual stacking then consists of a simple average stacking without any weighting in order to not introduce potential biases due to more complex statistical treatments. We point out that, as common in most other wavelength ranges where the underlying noise distribution is Gaussian, we cannot use a median stacking because most of the pixel values stacked are zero, hence the median would always be zero. This is particularly unfortunate because the median is known to be much more robust against outliers (e.g. one unusually bright pixel that immediately increases the average would just be ignored by the median). This is the main reason for applying the rejection as outlined above.

The flux in the six final stacked images is then extracted by integrating over the pre-defined 5 pixel diameter aperture. To maximize the flux covered by the aperture, it is actively centered on the brightest pixel in the image and not fixed at the image center. To correct for noise within the aperture, the background noise is calculated from the remaining part of the image (every pixel not covered by the flux extraction aperture) and then subtracted from the raw flux measured within the aperture. Because not every pixel in the 4 Ms mosaic has exactly the same exposure time, we used the exposure maps distributed with the actual science frames to determine a precise total exposure time for every stack. By dividing the noise-corrected number of counts within our aperture by this total exposure time, we get the final X-ray flux in the respective band in counts per second. Since our stacking algorithm does not account for the strongly varying PSF of *Chandra* to this stage, we have to correct the counts extracted from the final stacks for the different off-axis angle distributions of the stacked samples. Since the PSF becomes increasingly larger for larger off-axis angles, our pre-defined 2.5 arcsec radius aperture for flux extraction will lose flux for such sources that have large off-axis angles. To compensate for this, we computed the off-axis angle distribution of each stacked sample and its median value which was then used to construct a similarly distributed sample of faint, but detected sources with well known off-axis angles from the Xue et al. (2011) source catalog. By doing so, we found that all our samples show a fairly similar distribution of off-axis angles with median values between $2.8'$ and

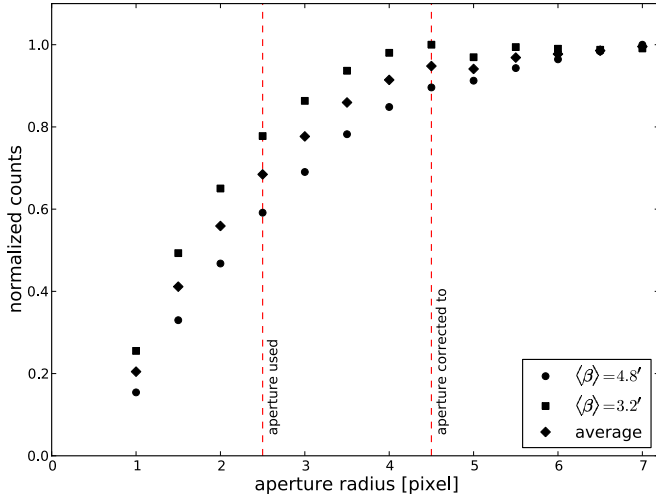


Fig. 2. Distribution of extracted counts with varying aperture radius for the soft *Chandra* band. Our pre-defined 2.5 arcsec aperture was corrected to the 4.5 arcsec aperture marked in red using the conversion factor of $\alpha_{\text{soft}} = 1.385$ as derived from this plot.

3.8 arcmin. This is not surprising since the LBGs in our stacked samples are chosen completely randomly with respect to their off-axis angles in a certain *Chandra* pointing. Therefore, we finally chose to adopt two control samples with off-axis angle distributions showing a median of $2.8''$ and $3.8''$ arcmin, respectively. Those two control samples were then stacked in the soft band according to the procedure outlined above (but of course without rejecting detected sources) and counts were extracted using different aperture sizes (Fig. 2). From this flux extraction experiment, we calculated a correction factor of $\alpha_{\text{soft}} = 1.385$ by comparing the counts extracted in our pre-defined 2.5 arcsec aperture to an aperture where 95% of the maximum number of counts achieved in any aperture were contained. This is the case for the 4.5 arcsec aperture. This factor is then applied to all subsequently extracted count values of our stacked samples. We did not use a 4.5 arcsec aperture directly in order to keep the effects of background noise as small as possible. This procedure ensures a maximum S/N result with realistic flux density estimates as it becomes obvious by comparing the flux density extracted from our two control samples to the flux density that one obtains by averaging over the flux densities as given in Xue et al. (2011). The deviation is just 6% for the $2.8''$ sample and 8% for the $3.8''$ sample. For the hard band, we performed a similar correction, yielding $\alpha_{\text{hard}} = 1.771$. Since we have only one hard band detection (see Sect. 4) and thus our results are entirely based on the soft band, the much larger deviation of 18% does not affect our results. The conversion to physical units (here $\text{erg s}^{-1} \text{cm}^{-2} = \text{mW m}^{-2}$) is then done by extrapolating a conversion factor between counts per second and flux based on the 4 Ms source catalog by Xue et al. (2011). To quantify the significance of detection in our final stacked images, the total number of counts within our extraction aperture, dubbed x , is compared to the noise level of the stacked image, dubbed λ . Assuming a Poissonian noise distribution, this noise level should be the expectation value of the distribution. Hence, we calculate a cumulated probability P which is the probability of measuring x or less counts where we expect λ counts. If this probability exceeds 95%, we assume the stacked detection to be real. We point out that, in most cases, our detections are well above the 99% confidence level. To give an error on our measured flux densities, we compute the number of counts necessary for a 68%

(corresponding to 1σ confidence level in a Gaussian distribution) signal and attribute this quantity to be the error of our measured flux.

3.2. Testing the algorithm

Considering the (mathematical) simplicity of our algorithm, we expect to avoid any biases due to statistical effects. Nevertheless, we ran the following tests on artificial data to ensure the performance of our stacking procedure in terms of reliability of stacked detections. Note that the main assumption of stacking – the noise goes down with the square-root of the number of stacked objects – is also valid for Poissonian noise distributions since they also exhibit (on the scales of the stacked sources) uncorrelated noise.

Therefore, we artificially created images with 500 by 500 pixels and a Poissonian noise distribution (as it is the case for X-ray images). Every image contains 27 point sources according to the PSF of *Chandra* at random off-axis angles and S/N of 3, 1 and 0.1 to obtain nine barely detectable sources, nine not detected sources and nine sources well beyond the detection limit. Because for the *Chandra* 4 Ms mosaic, special care was laid on astrometric calibration (see Xue et al. 2011, Sect. 4.2) resulting in a median positional offset between X-ray and radio sources from Miller et al. (2008) of $0.24''$, so clearly below the pixel size of *Chandra* which is $0.5''$, we ignored effects caused by positional uncertainties and worked with the exact positions of the artificial sources.

We then stacked the sources according to their attributed S/N using our algorithm as previously described. As usually assumed in stacking procedures, the noise level should decrease with the square root of the number of stacked objects. Therefore we expect the original S/N to be increased by a factor of three, making the barely detected ones well detected, the not detected ones barely detected and the sources well beyond detection should remain undetected. These expectations were met in all three cases: the stack of the nine $S/N = 3$ sources contains one nicely centered source for which we measured a new S/N of 8.8, the nine $S/N = 1$ sources were stacked into a single source with $S/N = 2.9$ and the nine $S/N = 0.1$ sources were not visible in the stacked image at all. The fact that the S/N reached in this test are always a bit below the theoretical expectation of the \sqrt{N} -law is mainly just due to small number statistics.

To further verify the algorithm and exclude random effects, e.g. from the fact that there may be a pixel with a high pixel value due to the long tail of the Poissonian noise distribution right at the site of an artificial source and hence unexpectedly increase their S/N, we injected the same 21 sources to the same image again but with a constant offset of five pixels. The sources were then stacked again, resulting into stacked images nearly identical in S/N to the first ones without offsets. For the final test, we stacked the sources with random offsets of up to 5 pixels. As expected, no source could be detected in these stacks (except of course for the nine stacked sources with an initial $S/N = 3$).

These three tests were repeated on 100 images to assure statistical significance. A summary of the tests is presented in Table 2, giving the median S/N of the sources in the final stacked images (in case there is a source detected).

To get a feeling of the stacking as an entire procedure, we also show a panel illustrating our routine (Fig. 3): the left image shows an artificial source with $S/N = 5$ sitting in a Poissonian noise distribution. The image in the middle is of the same size and noise level but with a source of $S/N = 1$ in the image center and the right image shows the resultant stack of nine such

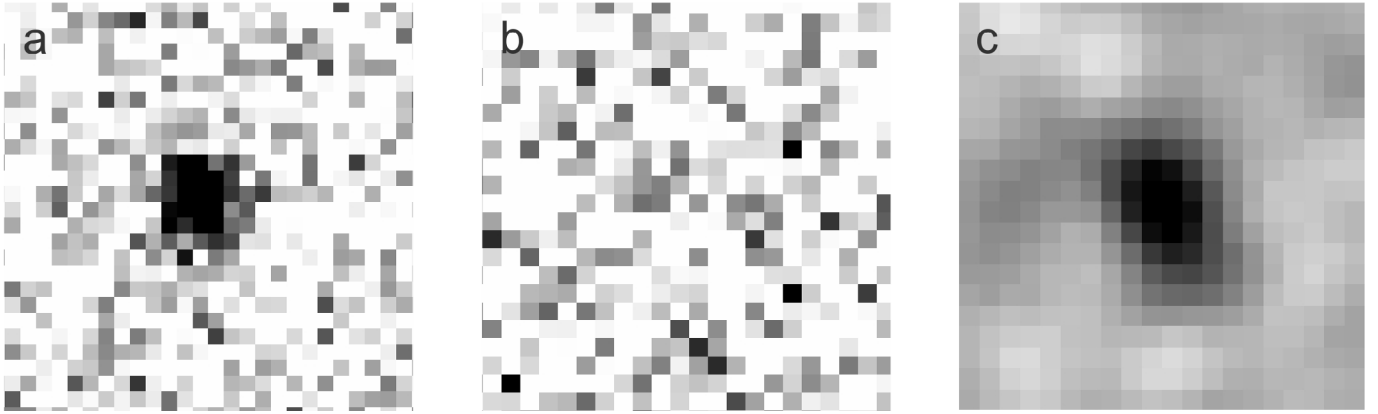


Fig. 3. Plot illustrating the different steps of our stacking procedure: **a)** shows an artificially created source of $S/N = 5$ at the image center within a Poissonian noise distribution, **b)** is the same as **a)** but with a $S/N = 1$ source and **c)** shows the final stacked image of nine such $S/N = 1$ sources (smoothed with a Gaussian of three pixel full-width-half-maximum) using our stacking algorithm.

Table 2. Summary of tests of the stacking algorithm. The values given in the table are the median S/N of the final stacked sources over all 100 test runs.

		Stacking offset		
		none	5 pixels	random
Initial S/N	0.1	nd	nd	nd
	1.0	2.91	2.88	nd
	3.0	8.88	8.89	nd

Notes. “nd” indicates that in the particular stack no source is detected (in the case for a stack of $S/N = 3$ sources no source with a significantly higher S/N).

$S/N = 1$ sources (smoothed with a Gaussian of three pixel full-width-half-maximum) using our stacking algorithm. We measured $S/N = 2.95$ for this stack within a circular aperture of 5 pixels diameter. Compare to Fig. 4 showing our stacks with real data.

4. Stacking results

Our final stacked soft-band images for all six input lists (resp. redshift bins) are shown in Fig. 4, a table summarizing the extracted fluxes and detection probabilities is presented in Table 3.

The flux extraction process was performed in two steps: first, all stacks were visually inspected for whether there is a detection or not. To quantify this, we then extracted the number of counts both within our pre-defined 5 pixel aperture (where the aperture was centered on the pixel with the highest count rate for the images where the visual inspection suggested a detection, for the others the aperture was centered on the image center) as well as for the remaining part of the stacked image. From the distribution of counts outside the aperture, we constructed the noise of the stacked image by fitting a Poissonian distribution to the histogram. This was done to verify the Poissonian behavior of the noise in general and to obtain the expectation value, λ , for every image. From this expectation value, we estimated the noise count rate that would be comprised by our 5 pixel aperture and subtracted it from the actual count rate measured within this aperture (correcting for pixels of which only a fraction lies within the aperture), yielding a count rate of the stacked sources only. Using our fit to the noise distribution, we then calculated the probability that such a count rate occurs just

by chance in the respective image. Its inverse probability (so $1 - \text{“by-chance-probability”}$) is then attributed the confidence of detection. Adopting a common confidence limit of 95%, we consider all stacks with a confidence greater than 0.95 as detections. The source count rates (the ones cleaned from noise counts) and the confidence values are given in Table 3.

We then converted the source count rates to physical flux densities assuming a power-law X-ray spectrum with a photon index typical for star-forming galaxies of $\Gamma = 2.0$ as empirically found by Ptak et al. (1999), see also the discussion on X-ray spectra of LBGs in Laird et al. (2006). We also corrected for Galactic extinction in the CDF-S ($N_{\text{HI}} = 6.8 \times 10^{19} \text{ cm}^{-2}$) following Kalberla et al. (2005). To test our flux calibration, we compared it to the original 4 Ms source catalog by Xue et al. (2011) finding them in very good agreement with their fluxes. Errors were assigned to all measured fluxes by taking into account the image noise (more precisely, the 68% uncertainty of the extracted flux, dubbed δf) and a possible 10% error due to the aperture size (dubbed σ). This error accounts for the fact that we chose our aperture to have a fixed size in contrast to other work where the aperture is empirically determined to maximize the S/N for each stack, see e.g. Lehmer et al. (2005); Laird et al. (2006):

$$\Delta f = \sqrt{(\delta f)^2 + (\sigma \cdot f)^2}, \quad (1)$$

where f is the measured flux of the detected source.

To further validate our results regarding the X-ray luminosity computed assuming a power-law with photon index $\Gamma = 2.0$, we calculated the hardness ratios (or upper limits, respectively) for each stack and check whether they are consistent with X-ray emission following such a power-law distribution. The hardness ratio is accordingly defined as $HR = (H - S)/(H + S)$, where H and S are the counts in the hard- and soft-band, respectively. Since obscured AGN at the considered redshifts would show rather flat X-ray spectra with photon indices $\Gamma \sim 1.1$, corresponding to $HR > 0.0$ Park et al. (2008), the computed hardness ratios from Table 3 reveal that also the final stacks are completely consistent with X-ray emission originating from star formation rather than from AGN activity. For completeness, we note that also the soft-band detections and hard-band non-detections for the two low-redshift control samples are consistent with the X-ray luminosity derived from the soft-band fluxes and a power-law SED with $\Gamma = 2$.

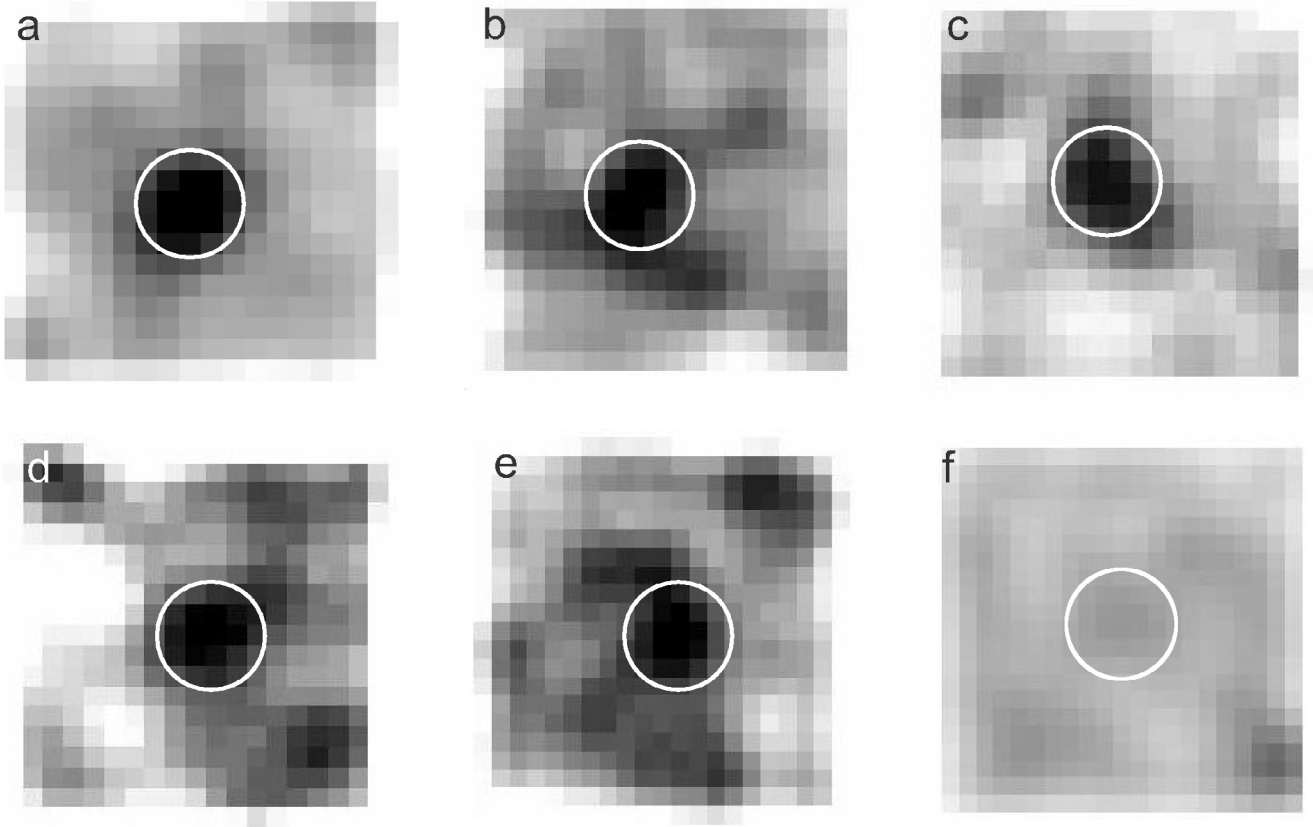


Fig. 4. Final stacked soft-band images (all smoothed with a Gaussian of $FWHM = 3$ pixels for illustration purposes) used for further analysis: **a)** ref-lowZ ($P = 0.999$), **b)** BZ-UV ($P = 0.993$), **c)** Spec-U1 ($P = 0.999$), **d)** Spec-U2 ($P = 0.958$), **e)** Spec-B1 ($P = 0.974$), **f)** Spec-B2 ($P = 0.560$). The white circles indicate the 5 pixels diameter apertures used for flux extraction.

Table 3. LBG stacking results.

ID	Median z	Soft counts	Conf ^a	Hard counts	Conf ^a	HR	Soft flux $10^{-18} \text{ mW m}^{-2}$	Soft flux err ^b $10^{-18} \text{ mW m}^{-2}$	Hard flux $10^{-18} \text{ mW m}^{-2}$	Hard flux err ^b $10^{-18} \text{ mW m}^{-2}$
ref-lowZ	0.68	153.83	0.999	60.01	0.907	< -0.44	2.98	0.51	7.82	u.l.
BZ-UV	1.33	36.36	0.993	25.12	0.850	< -0.18	2.84	0.33	7.44	u.l.
Spec-U1	2.67	112.26	0.999	100.74	0.971	-0.06	1.52	0.32	9.21	3.34
Spec-U2	3.33	39.91	0.958	14.54	0.645	< -0.47	1.18	0.26	1.64	u.l.
Spec-B1	3.70	21.52	0.974	16.83	0.722	< -0.12	3.09	0.96	4.19	u.l.
Spec-B2	4.16	24.14	0.660	7.09	0.626	...	1.40	u.l.	2.16	u.l.

Notes. For the convenience of the reader, we note that $1 \text{ mW m}^{-2} = 1 \text{ erg s}^{-1} \text{ cm}^{-2}$. All counts are given as measured in the pre-defined 2.5 pixel radius aperture, the correction factor α (see Sect. 3.1) was applied subsequently. ^(a) Probability that this detection is real, assuming a Poissonian noise distribution. See text for further explanation. ^(b) u.l. indicates upper limit since no actual detection could be made (confidence < 0.95).

5. Astrophysical interpretation

Table 4 summarizes our derived parameters that are discussed in the next sections.

To obtain the rest-frame 2–10 keV luminosity, we used our soft-band stacked fluxes and the corresponding median redshift of each bin. The median redshifts were calculated from the original spectroscopic redshifts as given in the input catalogs by Popesso et al. (2009), Balestra et al. (2010), and Vanzella et al. (2008) for our spectroscopic samples or, respectively, from the photometric redshifts given in Basu-Zych et al. (2011) for the BZ-UV sample or from the spectroscopic redshifts provided by Wuyts et al. (2008) for the low- z reference sample comprised from the FIREWORKS catalog. To obtain a

2–10 keV luminosity, we used the Portable, Interactive Multi-Mission Simulator (PIMMS) version 4.4² to get a 2–10 keV flux based on an extrapolation of our stacked soft-band fluxes and a photon index $\Gamma = 2.0$ typical for LBGs as e.g. found by Laird et al. (2006). This factor is adopted throughout the stacking literature, see e.g. Nandra et al. (2002) or Lehmer et al. (2005). Note that the stacked fluxes from Table 3 are already corrected for Galactic extinction following Kalberla et al. (2005). These fluxes could then be directly converted to luminosities with respect to the above specified cosmology. We did not account for any errors due to redshift uncertainties because the flux errors are by far dominating the luminosity errors.

² <http://asc.harvard.edu/toolkit/pimms.jsp>

Table 4. Derived parameters for our LBG samples.

ID	Median z	Rest 2–10 keV ^a log(W)	Rest 2–10 keV err ^b log(W)	SFR $M_{\odot} \text{ yr}^{-1}$	SFR err ^b $M_{\odot} \text{ yr}^{-1}$	M_{*} $10^{10} M_{\odot}$	M_{*} err $10^{10} M_{\odot}$	sSFR Gyr^{-1}	sSFR err ^b Gyr^{-1}
ref-lowZ	0.68	33.20	32.44	3.19	0.55	6.73	1.61	0.04	0.01
BZ-UV	1.33	33.79	32.86	11.62	1.45	3.35	0.67	0.37	0.11
Spec-U1	2.67	34.11	33.45	25.86	5.40	1.16	0.29	2.23	0.21
Spec-U2	3.33	34.17	33.52	29.68	6.63	1.09	0.28	2.71	1.11
Spec-B1	3.70	34.67	34.16	93.86	29.04	1.09	0.21	8.66	2.06
Spec-B2	4.16	34.56	u.l.	72.86	u.l.	1.15	0.31	6.32	u.l.

Notes. For the convenience of the reader, we note that $1 \text{ W} = 10^7 \text{ erg s}^{-1}$. ^(a) Assuming a power-law X-ray spectrum with a photon index $\Gamma = 2.0$ typical for LBGs (Laird et al. 2006). This factor is adopted throughout the stacking literature, see e.g. Nandra et al. (2002) or Lehmer et al. (2005). ^(b) u.l. indicates upper limit since no actual detection could be made (confidence < 0.95).

Star formation rates were calculated using the calibration by Ranalli et al. (2003):

$$SFR [M_{\odot} \text{ yr}^{-1}] = 2.0 \times 10^{-33} L_{2-10 \text{ keV}} \quad (2)$$

where $L_{2-10 \text{ keV}}$ is in units of Watts. The corresponding errors take into account both the uncertainties in luminosity and the intrinsic scatter of the calibration. Since Ranalli et al. (2003) did not quantify this scatter, we computed it using their data, finding an rms scatter of roughly 20%. In addition to the intrinsic scatter in that relation we point out that it is calibrated using a local galaxy sample only and that recent studies (e.g. Dijkstra et al. 2012) consider the possibility that the conversion factor between SFR and $L_{2-10 \text{ keV}}$ may increase with redshift, introducing a potential underestimation of SFR. The same conclusion is drawn by Symeonidis et al. (2011) who quantify their deviation to the Ranalli et al. (2003) calibration to be a factor of five. This would directly imply that, adopting the Symeonidis et al. (2011) conversion factor, our SFR would increase by a factor of five. However, for further analysis we chose to stick to the Ranalli et al. (2003) calibration just because it is the most widely used calibration throughout the literature.

Since our final goal is to also quantify the specific star formation rate (sSFR), defined as the SFR divided by the stellar mass of the galaxy ($sSFR = SFR/M_{*}$), we need to get a handle on the stellar mass. To do so, we computed the mean K -band magnitude of the LBGs in every stack using data from FIREWORKS (Wuyts et al. 2008). This mean K -band magnitude was then converted to a stellar mass following the calibration by Daddi et al. (2004):

$$\log(M_{*}/10^{11} M_{\odot}) = -0.4(K - K_{11}) \quad (3)$$

where K is the total K -band magnitude of the respective galaxy and $K_{11} = 20.15$ (Vega system) the typical K -magnitude of a $10^{11} M_{\odot}$ galaxy. A plot illustrating the results for our LBG sample is shown in Fig. 5. Since Daddi et al. (2004) found this relationship for galaxies in the redshift range $1 < z < 3$, we note that our low- z control sample may be affected by a systematic effects caused by its lower redshift. The scatter of this relationship is quantified by Daddi et al. (2004) to be ~ 0.1 dex, therefore the error on M_{*} takes into account both this scatter as well as the 1σ uncertainty of the mean K -band magnitude. The errors of the sSFR are then calculated by simple Gaussian error propagation. Note that, since the derived stellar masses for our LBG sample are just above the completeness limit, the higher-redshift bins might be affected by incompleteness issues leading to an overestimation of stellar mass.

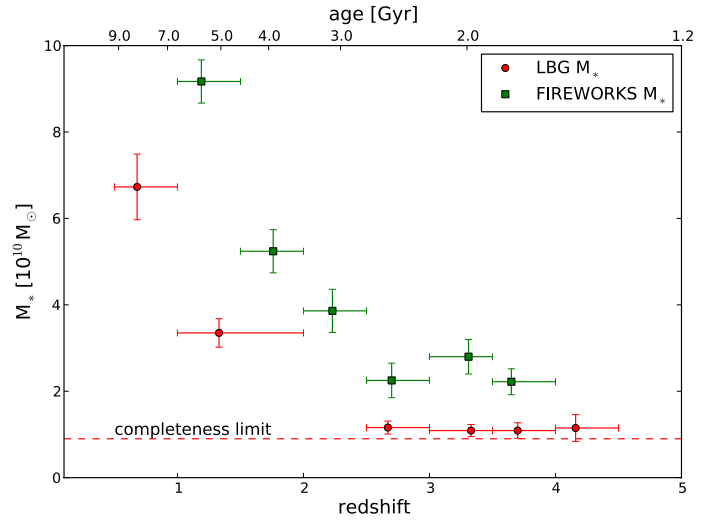


Fig. 5. Stellar masses of our LBG sample with redshift. For comparison, the stellar masses derived from the entire FIREWORKS sample with spectroscopic redshifts are shown, too. The plotted completeness limit corresponds to a 5σ limiting magnitude in the K_s -band of 22.85 Vega-mag.

5.1. The X-ray luminosity of LBGs

Figure 6 shows the measured X-ray fluxes versus redshift. In general, our results agree well with other X-ray stacking studies such as, for example, Lehmer et al. (2005) who found a rest-frame 2–8 keV luminosity for a sample of 449 U -band dropouts ($z \sim 3$) of $(1.5 \pm 0.3) \times 10^{34} \text{ W}$ and $(1.4 \pm 0.6) \times 10^{34} \text{ W}$ for a sample of 395 bright B dropouts ($z \sim 4$), also seeing a nearly constant luminosity with increasing redshift. A similar result was also obtained by Nandra et al. (2002) who give a 2–10 keV rest-frame luminosity of $(3.4 \pm 0.7) \times 10^{34} \text{ W}$ for their sample of 144 U -band dropouts in the *Hubble* Deep Field North. They also stacked a sample of 95 “Balmer break” galaxies which are located at $z \sim 1$, comparable to our BZ-UV sample with a median redshift of 1.33. For their $z \sim 1$ sample, Nandra et al. (2002) report a 2–10 keV rest-frame luminosity of $(0.33 \pm 0.05) \times 10^{34} \text{ W}$ which also agrees well with our value from Table 4. Furthermore, Nandra et al. (2002) only find significant detections in the soft-band stacks just as in this work. The only hard-band stacking detection is made for the Spec-U2 bin (presumably because it is the bin containing the most stacked sources) with a hard-band flux of $(5.20 \pm 1.88) \times 10^{-18} \text{ mW m}^{-2}$. Together with the

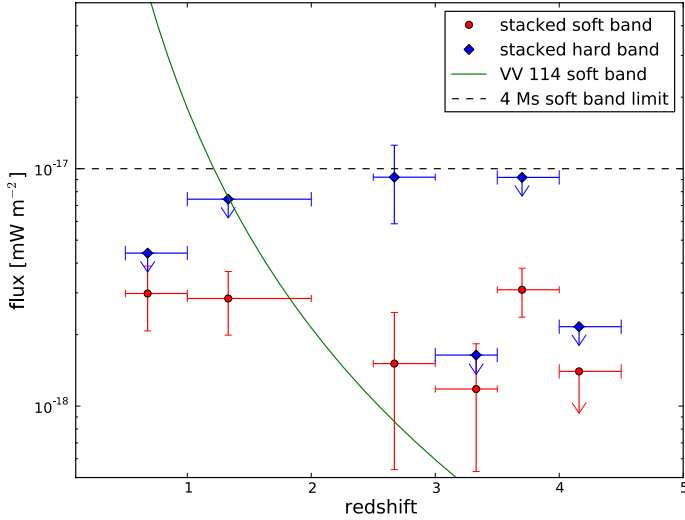


Fig. 6. Stacked fluxes of all samples in the observed-frame soft- and hard-band vs. redshift. Horizontal error bars indicate the width of the redshift bins, the solid line indicates the observed soft-band flux of the low- z LBG analog VV 114 from Grimes et al. (2006). VV 114 has an IR-derived SFR of $48 M_{\odot} \text{ yr}^{-1}$ (Soifer et al. 1989). The horizontal dashed line indicates the soft-band on-axis detection limit of the *Chandra* 4 Ms mosaic as reported by Xue et al. (2011).

corresponding soft-band flux (see Table 3), this gives a photon index $\Gamma \approx 1.1$, significantly lower than the value of 2.0 typical for star-formation activity. Since such low photon indices are more typical for AGN, we assume that this one hard X-ray detection may be due to a small remaining contamination of an AGN which still survived the many steps of AGN rejection previously applied to all stacked sources. An alternative explanation may be the contamination by a strong emission line at 7.47 keV as discussed e.g. in Fiore et al. (2012). However, since the effective area of *Chandra* becomes very small for such high energies, the contribution of photons with energies exceeding 7 keV only makes up a small fraction of the total counts in the 2–8 keV band.

However, we highlight that recent studies by Lehmer et al. (2012) hint a steep rise of number counts of normal star-forming galaxies just below the soft-band flux limit of the *Chandra* 4 Ms mosaic. They report that at flux levels of the order of $10^{-17} \text{ mW m}^{-2}$ normal SFGs make up nearly half the X-ray number counts and that at even lower flux levels those galaxies will completely dominate the population. Therefore, since our stacking analysis probes the faint X-ray population down to $10^{-18} \text{ mW m}^{-2}$, our findings are also consistent with extrapolations from individually detected sources.

5.2. Star formation in LBGs

From the rest 2–10 keV luminosities of our various redshift bins, we estimated the SFR for an average LBG using the method by Ranalli et al. (2003). The resulting SFRs are shown in Fig. 7 (left panel) alongside with several other SFR measurements for LBGs calculated from data taken at different wavelength ranges. Our X-ray derived SFRs peak at a redshift of about $z_{\text{peak}} = 3.5$ with the last measurement at $z \sim 4$ being an upper limit slightly below the peak value.

Compared to SFR estimated from other wavelength ranges, we find at least a good agreement in the trend (increase to $z_{\text{peak}} = 3.5$, then decrease) whilst the absolute numbers differ by

a factor 2 to 5, depending on the particular data set for comparison. Especially the Karim et al. (2011) values show much larger values which could be easily explained by the very different selection criteria they used for their stacking input samples: while we focus on LBGs, they selected their sample based on $3.6 \mu\text{m}$ flux density to be able to split not only in redshift bins but also in stellar mass bins. However, the trend of increasing SFR to at least $z = 3$ is seen in both our LBG as well as their $3.6 \mu\text{m}$ samples. To compare with other LBG samples, we plotted the values obtained by Carilli et al. (2008) for lower redshifts and Smit et al. (2012) for higher redshifts. Carilli et al. (2008) used LBGs in the COSMOS field to create U -, B - and V -dropout samples as input for a 1.4 GHz stacking analysis. They detected only the U -dropouts ($z \sim 3$) in the radio stacks and find $SFR = 31 \pm 7 M_{\odot} \text{ yr}^{-1}$ (adopting the $SFR-L_{1.4 \text{ GHz}}$ calibration by Yun et al. 2001) while the average UV-derived SFR for this sample is $17 M_{\odot} \text{ yr}^{-1}$. These two values agree well with our X-ray based estimate of $29.68 \pm 6.63 M_{\odot} \text{ yr}^{-1}$ at this redshift. The decrease with redshift is confirmed by the results of Smit et al. (2012) who investigated the star formation in higher-redshift LBGs based on their rest-frame UV continuum slopes. Accounting for dust extinction following Meurer et al. (1999), they find SFRs of $35 M_{\odot} \text{ yr}^{-1}$ and $24 M_{\odot} \text{ yr}^{-1}$ at redshifts of 3.8 and 5.0, respectively. Comparing their results to other data, they see a peak in SFR between $3.0 < z_{\text{peak}} < 3.5$ and a peak value of $50\text{--}60 M_{\odot} \text{ yr}^{-1}$, very comparable to our X-ray based findings.

All together, we can confirm a peak in the star formation activity of LBGs at a redshift around $z_{\text{peak}} = 3.5$ which is seen by various authors using various SFR estimation methods. However, the absolute values for SFR differ by quite a large margin which is either due to the different selection criteria for the galaxy samples investigated or inconsistencies in the calibration of the different SFR estimation methods in different wavelength ranges, see e.g. Kurczynski et al. (2012). This again shows that a uniform and consistent cross-calibration of SFR indicators is desperately needed, particularly in advance of upcoming large survey projects observing across the entire electromagnetic spectrum.

Regarding the sSFR, a similar peaking trend with redshift is seen (Fig. 7, right panel). Although our galaxies are in general less massive than in the Karim et al. (2011) selection, the two samples are in better agreement while looking at the sSFR since it normalizes for the mass. Therefore it looks like the LBG sample is smoothly continuing the $3.6 \mu\text{m}$ selection at higher redshifts. This suggests that at redshifts of $z \sim 4$, LBGs are typical for the galaxy population whereas at lower redshifts more massive galaxies are abundant. Since the upper limit obtained for $z = 4.16$ is significantly lower than the actual measurement at $z = 3.70$, a clear decrease for higher redshifts is also visible in sSFR. This peaking behavior of the sSFR supports the widely adopted picture of stellar mass growth having a peak somewhere between redshift 2 and 4, but contradicts findings by other authors who observed (Feulner et al. 2005) or modeled (Khochfar & Silk 2011) the sSFR for a wide range of redshifts and find a more or less constant sSFR (at least for lower mass galaxies) from $z = 2$ onwards to higher redshifts. An explanation for this difference could again be the different sample selection criteria. While we looked at LBGs only, Feulner et al. (2005) did a more complex near-IR selection to again split in stellar mass bins. A comparable LBG sample in the Subaru Deep Field was investigated by Yoshida et al. (2006) who also find a peaking sSFR with a peak value of about 0.1 Gyr^{-1} at $z \approx 4$. This is more than an order of magnitude lower than our peak value, but since they focus on more massive LBGs this difference is not

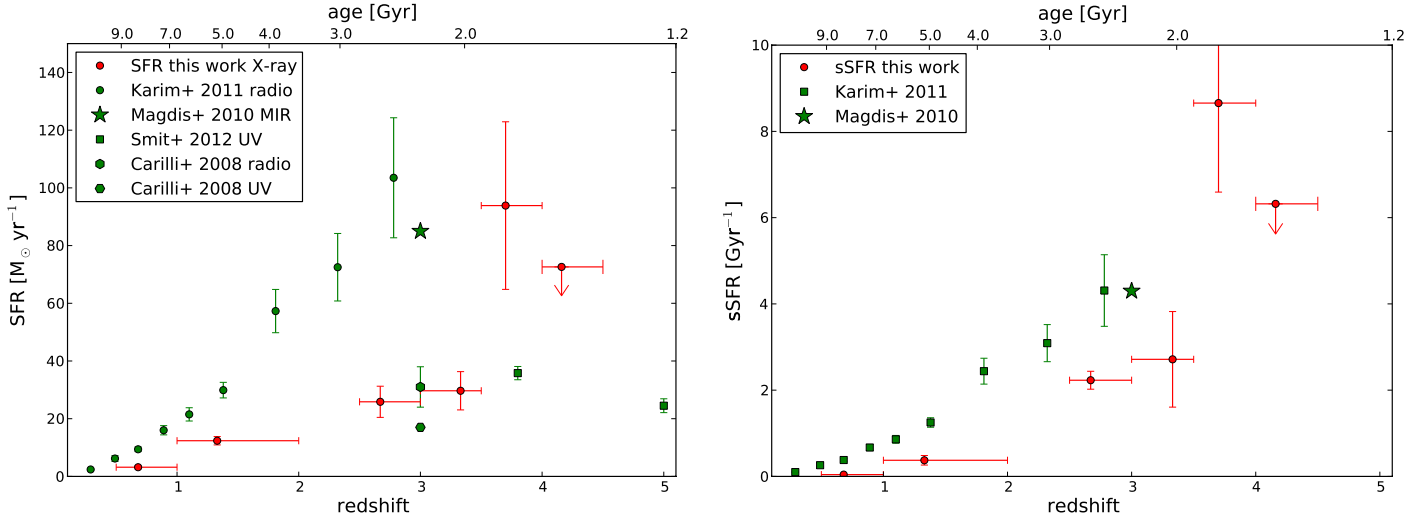


Fig. 7. SFR (*left*) and the sSFR (*right*) for different LBG samples vs. redshift. We compare our findings to the ones by [Karim et al. \(2011\)](#), [Magdis et al. \(2010\)](#), [Carilli et al. \(2008\)](#) and [Smit et al. \(2012\)](#) using various other SFR indicators such as radio or UV luminosities.

a surprise. They argue (according to the analytical galaxy evolution model by [Hernquist & Springel 2003](#)) that the trend of a peaking SFR/sSFR at redshifts between 3 and 4 can be explained by different parameters dominating the star formation process at various redshifts: At lower redshifts, the star formation activity is mostly governed by the cooling rate of the (molecular) gas residing in a dark matter halo whereas at higher redshifts the conversion of cold gas into stars is the dominating parameter regulating star formation. The transition between these two modes is marked by the peak of star formation activity. This hypothesis is recently supported by the work of [Reddy et al. \(2012\)](#) who also see a (mild) peak in the sSFR of several Gyr^{-1} at $z \approx 3$.

5.3. The contribution of LBGs to the cosmic star formation rate density

Figure 8 shows the co-moving star formation rate density (SFRD) as derived from our SFR (Table 4) and a cosmology according to Sect. 1. As one can easily see, the peaking trend also continues in this plot, although the peak is much less pronounced. We compare our estimates to values from [Yoshida et al. \(2006\)](#) and references therein (green diamonds in Fig. 8) to find a similar peaking trend albeit at absolute SFRD values being about one order of magnitude higher than ours. These higher values can be explained by the difference in LBG sample selection between [Yoshida et al. \(2006\)](#) and our work (all our LBGs are spectroscopically confirmed whereas they employ a mixed spectroscopic/photometric selection process).

However, despite the (relatively) small difference between the two LBG samples, there is a much larger difference when comparing SFRD derived from LBG samples to estimates of the total cosmic SFRD as e.g. compiled by [Hopkins \(2004\)](#). The difference between total and LBG SFRD is about four orders of magnitude, hence deeming LBGs to be entirely negligible when investigating the bulk of star formation activity in the universe's history, even when considering potential incompleteness of our LBG sample. This finding agrees very well with previous investigations of the cosmic star formation history (e.g. [Sawicki & Thompson 2006](#); [Bouwens et al. 2007, 2009](#)). From constraints on the UV luminosity functions at various redshifts,

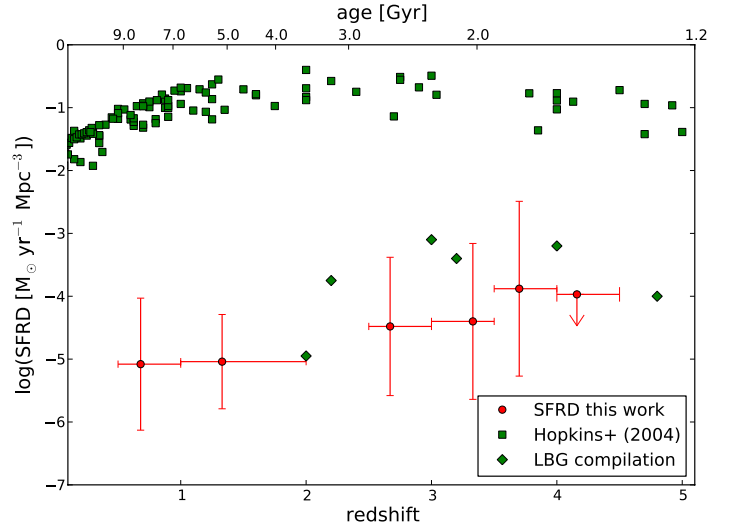


Fig. 8. Contribution of LBGs to the total cosmic SFRD. SFRD measures (green squares) were taken from [Hopkins \(2004\)](#). LBG compilation (green diamonds) refers to measurements compiled by [Yoshida et al. \(2006\)](#) which include values from [Wyder et al. \(2005\)](#) and [Arnouts et al. \(2005\)](#).

these authors drew conclusions for the question on whether the main fraction of star formation activity takes place at the luminous or faint end of the high-redshift galaxy population. Finding extremely steep faint-end slopes at all redshifts $z > 2$, they argue that most of the UV emission (and hence star formation) takes place in low-luminosity galaxies with only a few percent contribution of the bright population. More quantitatively, [Sawicki & Thompson \(2006\)](#) derived characteristic luminosities \mathcal{L}_* (and the corresponding absolute magnitude \mathcal{M}_*) for LBGs by fitting [Schechter \(1976\)](#) luminosity functions (LFs) at redshifts 2.2, 3.0 and 4.0. They find a nearly constant value of $\mathcal{M}_* = -21.0$ at a rest-frame wavelength of 1700 \AA , corresponding to a SFR of about $15 M_\odot \text{ yr}^{-1}$ (uncorrected for potential dust extinction). Because of the steep faint-end slopes of the fitted LFs, they argue that the total UV luminosity density (and hence SFRD) in

this redshift range not dominated by \mathcal{L}_* or even brighter galaxies but by faint galaxies, mostly around luminosities of $0.1 \mathcal{L}_*$. The faint end-slopes of the rest-frame UV LFs becomes even steeper at higher redshifts (Bouwens et al. 2007, 2009). Therefore, the marginal contribution of LBGs to the total cosmic SFRD is not surprising since the LBG samples used in this work are based on a spectroscopic selection. Hence the sources must be bright enough for spectroscopy, which in this case (e.g. for the VIMOS spectroscopic campaign in the CDF-S by Balestra et al. 2010) means that they have to have apparent magnitudes brighter than 24.5 in the B - and R -bands, respectively. At $z = 2.0$, this corresponds to an absolute B -band (so rest-frame $\sim 1500 \text{ \AA}$) magnitude of -21.5 , half a magnitude brighter than \mathcal{M}_* derived by Sawicki & Thompson (2006). At $z = 3.0$, the situation becomes even worse since the spectroscopic sample in this case is limited to objects 1.5 mag brighter than \mathcal{M}_* . We therefore point out that our results are in very good agreement with the previous statements that the bulk of star formation activity takes place not in bright but in faint galaxies. An independent study recently conducted by Tanvir et al. (2012) utilizing gamma-ray bursts (GRBs) as SFR tracers, so following an entirely different premise, comes to the same result. With their data for even higher redshifts ($z \sim 5$), they argue that the bulk of star formation activity is not even accessible in currently available ultra-deep data sets such as the *Hubble* Ultra-Deep Fields which implies mean SFR per “typical” galaxy of less than $0.2 M_\odot \text{ yr}^{-1}$.

6. Summary and conclusions

Utilizing stacking techniques together with the newly acquired *Chandra* 4 Ms mosaic, we have investigated the X-ray luminosity and star formation activity of LBGs across cosmic time. Our stacking input sample spans a redshift range $0.5 < z < 4.5$, making use of the superb spectroscopic data in the *Chandra* Deep Field South region. Spectroscopic selection of LBGs guarantees a highly homogeneous sample with minimal contamination by interloping objects, in particular AGN contaminating the measured X-ray fluxes. Our newly developed stacking algorithm, optimized for highest sensitivities in exchange for the loss of morphological information allows us to probe the LBG population down to formerly unprecedented flux density levels of $10^{-18} \text{ mW m}^{-2}$ in the soft (0.5–2 keV) band. Our findings are summarized below:

1. We reliably (Poissonian confidence level $>95\%$) detect X-ray emission from our LBG input sample out to redshifts of about 4 with a robust upper limit for $z = 4.5$.
2. From the observed soft-band fluxes, we derive rest-frame 2–10 keV luminosities using a photon index $\Gamma = 2.0$ and correcting for Galactic hydrogen absorption. They show a nearly constant value with redshift, underlining the robustness of our stacking procedure.
3. From the rest 2–10 keV luminosities we calculated mean SFR for each redshift bin. Our results show a distinct peak at $z_{\text{peak}} \sim 3.5$, in good agreement with the general trend of various other estimates of SFR for LBGs and other types of galaxies. However, we point out that the various SFR indicators across the electromagnetic spectrum deliver significantly different SFR values, therefore a thorough comparison and cross-calibration is highly needed also in advance of upcoming all-sky survey projects.
4. With ancillary K -band infrared data, we calculated stellar masses and hence sSFR. The peaking behavior of our

LBG sample is also seen in sSFR, underlining the widely adopted picture of stellar mass growth having a peak somewhere between redshift 2 and 4.

5. Considering the contribution of LBGs to the total cosmic SFRD, we find that LBGs only make up a tiny fraction of the total star formation activity at all investigated redshifts, supporting the emergent notion that the bulk of star formation in the universe takes place in very low-mass, low-SFR galaxies currently escaping detection with all available facilities.

All together, LBGs seem to be good examples of “typical” galaxies at their respective redshifts but are not the place in the universe where most the star formation activity takes place. Despite their high typical SFR of a few to several $10 M_\odot \text{ yr}^{-1}$ at all redshifts, LBGs are just not abundant enough to make up a significant fraction of the total star formation in the universe. This highly supports the proposition that the bulk of high-redshift star formation is going on in faint, yet undetected galaxies well below \mathcal{L}_* at all redshifts $z > 2$ (see e.g. Tanvir et al. 2012, and references therein).

Acknowledgements. We thank our anonymous referee for his/her constructive comments, in particular regarding the many technical finesses of X-ray observations.

References

- Arnouts, S., Schiminovich, D., Ilbert, O., et al. 2005, *ApJ*, 619, L43
 Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, *PASP*, 93, 5
 Balestra, I., Mainieri, V., Popesso, P., et al. 2010, *A&A*, 512, A12
 Barmby, P., Huang, J.-S., Fazio, G. G., et al. 2004, *ApJS*, 154, 97
 Basu-Zych, A. R., Hornschemeier, A. E., Hoversten, E. A., Lehmer, B., & Gronwall, C. 2011, *ApJ*, 739, 98
 Bouwens, R. J., Illingworth, G. D., Franx, M., & Ford, H. 2007, *ApJ*, 670, 928
 Bouwens, R. J., Illingworth, G. D., Franx, M., et al. 2009, *ApJ*, 705, 936
 Bouwens, R. J., Illingworth, G. D., González, V., et al. 2010, *ApJ*, 725, 1587
 Bouwens, R. J., Illingworth, G. D., Labbe, I., et al. 2011a, *Nature*, 469, 504
 Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2011b, *ApJ*, 737, 90
 Brandt, W. N., Hornschemeier, A. E., Schneider, D. P., et al. 2001, *ApJ*, 558, L5
 Burgarella, D., Le Floch, E., Takeuchi, T. T., et al. 2007, *MNRAS*, 380, 986
 Burgarella, D., Heinis, S., Magdis, G., et al. 2011, *ApJ*, 734, L12
 Carilli, C. L., Lee, N., Capak, P., et al. 2008, *ApJ*, 689, 883
 Chapman, S. C., & Casey, C. M. 2009, *MNRAS*, 398, 1615
 Cooke, J., Berrier, J. C., Barton, E. J., Bullock, J. S., & Wolfe, A. M. 2010, *MNRAS*, 403, 1020
 Cowie, L. L., Barger, A. J., & Hasinger, G. 2012, *ApJ*, 748, 50
 Daddi, E., Cimatti, A., Renzini, A., et al. 2004, *ApJ*, 617, 746
 Dijkstra, M., Gilfanov, M., Loeb, A., & Sunyaev, R. 2012, *MNRAS*, 421, 213
 Ebrero, J., Carrera, F. J., Page, M. J., et al. 2009, *A&A*, 493, 55
 Feulner, G., Gabasch, A., Salvato, M., et al. 2005, *ApJ*, 633, L9
 Fiore, F., Puccetti, S., Grazian, A., et al. 2012, *A&A*, 537, A16
 Giavalisco, M. 2002, *ARA&A*, 40, 579
 Grimes, J. P., Heckman, T., Hoopes, C., et al. 2006, *ApJ*, 648, 310
 Habertzell, L., Williger, G. M., Lauroesch, J. T., et al. 2009, *ApJ*, 702, 506
 Hernquist, L., & Springel, V. 2003, *MNRAS*, 341, 1253
 Hickox, R. C., & Markevitch, M. 2007, *ApJ*, 661, L117
 Hopkins, A. M. 2004, *ApJ*, 615, 209
 Kalberla, P. M. W., Burton, W. B., Hartmann, D., et al. 2005, *A&A*, 440, 775
 Karim, A., Schinnerer, E., Martínez-Sansigre, A., et al. 2011, *ApJ*, 730, 61
 Khochfar, S., & Silk, J. 2011, *MNRAS*, 410, L42
 Komatsu, E., Smith, K. M., Dunkley, J., et al. 2011, *ApJS*, 192, 18
 Kurczynski, P., Gawiser, E., Huynh, M., et al. 2012, *ApJ*, in press [arXiv:1010.0290]
 Labbé, I., González, V., Bouwens, R. J., et al. 2010, *ApJ*, 708, L26
 Laird, E. S., Nandra, K., Hobbs, A., & Steidel, C. C. 2006, *MNRAS*, 373, 217
 Lehmer, B. D., Brandt, W. N., Alexander, D. M., et al. 2005, *AJ*, 129, 1
 Lehmer, B. D., Xue, Y. Q., Brandt, W. N., et al. 2012, *ApJ*, 752, 46
 Magdis, G. E., Rigopoulou, D., Huang, J.-S., et al. 2008, *MNRAS*, 386, 11
 Magdis, G. E., Elbaz, D., Daddi, E., et al. 2010, *ApJ*, 714, 1740
 Mannucci, F., Cresci, G., Maiolino, R., et al. 2009, *MNRAS*, 398, 1915
 Mao, M. Y., Huynh, M. T., Norris, R. P., et al. 2011, *ApJ*, 731, 79
 Meurer, G. R., Heckman, T. M., & Calzetti, D. 1999, *ApJ*, 521, 64

- Miller, N. A., Fomalont, E. B., Kellermann, K. I., et al. 2008, *ApJS*, 179, 114
- Moran, E. C., Lehnert, M. D., & Helfand, D. J. 1999, *ApJ*, 526, 649
- Nandra, K., Mushotzky, R. F., Arnaud, K., et al. 2002, *ApJ*, 576, 625
- Park, S. Q., Barmby, P., Fazio, G. G., et al. 2008, *ApJ*, 678, 744
- Popesso, P., Dickinson, M., Nonino, M., et al. 2009, *A&A*, 494, 443
- Ptak, A., Serlemitsos, P., Yaqoob, T., & Mushotzky, R. 1999, *ApJS*, 120, 179
- Ranalli, P., Comastri, A., & Setti, G. 2003, *A&A*, 399, 39
- Reddy, N. A., Pettini, M., Steidel, C. C., et al. 2012, *ApJ*, 754, 25
- Rigopoulou, D., Huang, J.-S., Papovich, C., et al. 2006, *ApJ*, 648, 81
- Rix, H.-W., Barden, M., Beckwith, S. V. W., et al. 2004, *ApJS*, 152, 163
- Sawicki, M., & Thompson, D. 2006, *ApJ*, 648, 299
- Schechter, P. 1976, *ApJ*, 203, 297
- Smit, R., Bouwens, R. J., Franx, M., et al. 2012, *ApJ*, 756, 14
- Soifer, B. T., Boehmer, L., Neugebauer, G., & Sanders, D. B. 1989, *AJ*, 98, 766
- Stark, D. P., Ellis, R. S., & Ouchi, M. 2011, *ApJ*, 728, L2
- Steidel, C. C., & Hamilton, D. 1993, *AJ*, 105, 2017
- Steidel, C. C., Giavalisco, M., Dickinson, M., & Adelberger, K. L. 1996, *AJ*, 112, 352
- Steidel, C. C., Adelberger, K. L., Giavalisco, M., Dickinson, M., & Pettini, M. 1999, *ApJ*, 519, 1
- Stern, D., Eisenhardt, P., Gorjian, V., et al. 2005, *ApJ*, 631, 163
- Symeonidis, M., Georgakakis, A., Seymour, N., et al. 2011, *MNRAS*, 417, 2239
- Tanvir, N. R., Levan, A. J., Fruchter, A. S., et al. 2012, *ApJ*, 754, 46
- Tasker, E. J., & Bryan, G. L. 2006, *ApJ*, 642, L5
- Trujillo, I., Förster Schreiber, N. M., Rudnick, G., et al. 2006, *ApJ*, 650, 18
- Vanzella, E., Cristiani, S., Dickinson, M., et al. 2005, *A&A*, 434, 53
- Vanzella, E., Cristiani, S., Dickinson, M., et al. 2006, *A&A*, 454, 423
- Vanzella, E., Cristiani, S., Dickinson, M., et al. 2008, *A&A*, 478, 83
- Vanzella, E., Pentericci, L., Fontana, A., et al. 2011, *ApJ*, 730, L35
- Volino, F., Wucknitz, O., McKean, J. P., & Garrett, M. A. 2010, *A&A*, 524, A79
- Weatherley, S. J., & Warren, S. J. 2003, *MNRAS*, 345, L29
- Wolf, C., Meisenheimer, K., Rix, H.-W., et al. 2003, *A&A*, 401, 73
- Wright, E. L. 2006, *PASP*, 118, 1711
- Wuyts, S., Labbé, I., Schreiber, N. M. F., et al. 2008, *ApJ*, 689, 653
- Wyder, T. K., Treyer, M. A., Milliard, B., et al. 2005, *ApJ*, 619, L15
- Xue, Y. Q., Luo, B., Brandt, W. N., et al. 2011, *ApJS*, 195, 10
- Yoshida, M., Shimasaku, K., Kashikawa, N., et al. 2006, *ApJ*, 653, 988
- Yun, M. S., Reddy, N. A., & Condon, J. J. 2001, *ApJ*, 554, 803
- Zinn, P.-C., Middelberg, E., & Ibar, E. 2011, *A&A*, 531, A14