UNIVERSITÀ DEGLI STUDI DI PADOVA



DIPARTIMENTO DI FISICA e ASTRONOMIA "G. Galilei"

SCUOLA DI DOTTORATO DI RICERCA IN ASTRONOMIA
CICLO XXVI

Searching for High-redshift Galaxies in Hubble Space Telescope Deep Data

Direttore della Scuola: Ch.mo Prof. Giampaolo Piotto
Supervisori: Ch.mo Prof. Alessandro Pizzella
Dott. Massimo Stiavelli (STScI)

Dottoranda: Valentina Calvi

2014

In memory of my grandfather, Orso & Riky

If you can dream it, you can do it.
-Walt Disney-

With increasing distance our knowledge fades and fades rapidly. Eventually we reach the dim boundary, the outmost limits of our telescope. The search will continue. Not until the empirical resources are exhausted need we pass on to the dreamy realm of speculation.
-Edwin Hubble-

Contents

1 Introduction **1**
- 1.1 A Brief History of the Universe . . . 1
- 1.2 Cosmic Reionization . . . 3
  - 1.2.1 The Reionization Process . . . 6
  - 1.2.2 Possible Sources for Reionization . . . 6
- 1.3 The Lyman-Break Technique . . . 8
  - 1.3.1 Lower-redshift Contaminants . . . 11
- 1.4 Deep Imaging Surveys . . . 11
  - 1.4.1 A Brief History of HST Deep Surveys . . . 11
  - 1.4.2 The Great Observatories Origins Deep Survey . . . 13
  - 1.4.3 The Hubble Ultra Deep Field . . . 13
  - 1.4.4 The UDF05 . . . 13
  - 1.4.5 The eXtreme Deep Field . . . 14
- 1.5 Radiation Background . . . 14
- 1.6 The Luminosity Function . . . 16
- 1.7 The Power Spectrum . . . 17
- 1.8 Cosmological Surface Brightness Dimming . . . 19
- 1.9 Aim and Scheme of the Thesis . . . 21

2 Constraining the Luminosity Function of Faint Undetected i-dropout Galaxies **23**
- 2.1 The Data . . . 23
- 2.2 Data Analysis . . . 24
  - 2.2.1 Creation of the Mask . . . 25
  - 2.2.2 The Power Spectrum Technique . . . 27
  - 2.2.3 Different Contributions . . . 31
  - 2.2.4 Possible Sources of Contamination . . . 33
- 2.3 Monte Carlo Simulations . . . 40
  - 2.3.1 Elements of the Simulations . . . 40
  - 2.3.2 Results of the Simulations . . . 48
  - 2.3.3 Detected i_{775} -dropouts . . . 49
  - 2.3.4 The z ~ 6 Faint Galaxies Contribution to Cosmic Reionization . . . 51
- 2.4 Conclusions . . . 51

3 Looking for Faint Galaxies at $z \sim$ 7-8: Analysis of WFC3 Data from the HUDF09 Project **55**
- 3.1 The Data . . . 55

3.2 Data Reduction . . . 59
3.2.1 Persistence . . . 60
3.2.2 Cosmetic of the WFC3/IR Detector . . . 62
3.2.3 Dithering . . . 63
3.2.4 MultiDrizzle . . . 63
3.3 A Customized Version of the HUDF09 Dataset . . . 66
3.3.1 Median Subtraction . . . 67
3.4 Data Analysis: Looking for Faint Galaxies at $z \sim 8$ . . . 67
3.4.1 Identification of the Sources and Creation of the Mask . . . 68
3.4.2 Power Spectra . . . 69
3.4.3 Effect of the Source Masking . . . 73
3.4.4 Simulations . . . 73
3.5 Data Analysis: Looking for Faint Galaxies at $z \sim 7$ . . . 75
3.6 Summary and Conclusions . . . 76

4 HUDF Main Field Data Reduction **81**
4.1 Analysis of the XDF data . . . 81
4.2 Description and Aim of the Reprocessing Procedure . . . 85
4.3 Raw Data . . . 85
4.4 Calibration Reference Files . . . 86
4.4.1 Bias . . . 87
4.4.2 Dark . . . 88
4.4.3 Flat Field . . . 89
4.5 Satellite Trails . . . 90
4.6 Alignment . . . 90
4.6.1 WCS Coordinates Update . . . 92
4.6.2 Cosmic Ray Rejection and SExtractor Catalogs . . . 93
4.6.3 Tweakreg . . . 94
4.6.4 Drizzling and Second Alignment . . . 95
4.6.5 Tweakback . . . 95
4.6.6 Drizzling, Alignment, and Final Combination . . . 95
4.7 Correcting for the Bias Jump . . . 96
4.8 Electronic Crosstalk Ghosts . . . 97
4.9 Preliminary Results on the XDF v2.0 Dataset . . . 98
4.9.1 Photometry Quality Check . . . 98
4.9.2 Power Spectra . . . 99
4.10 Summary and Conclusions . . . 102

5 Surface Brightness Dimming Effect **105**
5.1 The Data . . . 106
5.1.1 The Hubble Ultra Deep Field . . . 106
5.1.2 The Great Observatories Origins Deep Survey . . . 106
5.2 Rescaling the RMS Maps . . . 109
5.3 Source Selection . . . 110
5.4 Results . . . 112
5.5 Simulations . . . 117
5.6 Conclusions . . . 120

6 A Preliminary Study on the Lower-redshift Contaminant Population Entering the Dropout Samples at $z \sim$ 4-5-6 **123**
6.1 CANDELS GOODS-S Multi-wavelength Catalog . . . 124
6.1.1 Dropouts Selection . . . 124
6.1.2 Preliminary Results . . . 126
6.2 3D-HST Catalogs . . . 126
6.2.1 Preliminary Results . . . 129
6.3 Next Steps . . . 134

7 Summary and Conclusions **135**
7.1 The Contribution to Reionization from Faint Galaxies . . . 136
7.2 The Effect of Cosmological Dimming on Deep Survey . . . 137
7.3 Lower-redshift Contaminants Entering the Dropout Selection . . . 137
7.4 Conclusions . . . 138
7.5 Future Perspective . . . 139

A Gas-phase Metallicity of 27 Galaxies at Intermediate Redshift **141**
A.1 Introduction . . . 141
A.2 Observations and Data Reduction . . . 142
A.3 Sample Selection . . . 142
A.4 Data Analysis and Gas-phase Metallicity . . . 145
A.5 Conclusions . . . 147

B Stellar Population Properties for a Sample of Hard X-ray AGNs **151**
B.1 Introduction . . . 151
B.2 Sample Selection and Data Properties . . . 153
B.2.1 The Sample . . . 153
B.2.2 The data . . . 153
B.3 Measurements of the Stellar Populations . . . 154
B.4 Results and Discussion . . . 155
B.5 Summary and Conclusions . . . 162

Bibliography **163**

Acknowledgments **181**

List of Figures

1.1 The evolution of the Universe. 2
1.2 Expected cumulative SB density of reionization sources as a function of the apparent magnitude 5
1.3 Color selection criteria for LBGs. 9
1.4 Probing the Universe with HST 12
1.5 XDF size and pointings 14
1.6 Schechter luminosity function 17
1.7 Visualization of the power spectrum. 18

2.1 Pointings of the UDF05 program 24
2.2 Power spectrum of M32 (Tonry and Schneider, 1988) 28
2.3 i_{775} and z_{850} band power spectra 29
2.4 Ratio between the z_{850} and i_{775} power spectra 30
2.5 Power spectra of pairs of point sources 32
2.6 Ratios between power spectra using different masks 34
2.7 Effect of `MultiDrizzle` on power spectra (noise) 35
2.8 Effect of `MultiDrizzle` on power spectra (noise + flat field residual) 35
2.9 Comparison between the v1.0 and v2.0 power spectra 37
2.10 Ratio between the power spectra obtained from the v1.0 and v2.0 data . . . 37
2.11 Power spectra obtained applying a rotation 38
2.12 Ratio between the power spectra obtained applying a rotation 38
2.13 Power spectra obtained from unbinned and binned frames 39
2.14 Ratio between the power spectra obtained from unbinned and binned frames 39
2.15 Ratios between the power spectra obtained from simulations including only noise and large scale residuals. 41
2.16 Distribution of the total magnitude and effective radius 43
2.17 Distribution of the ellipticity and PA for the ACS data 43
2.18 Comparison of the ratio between the simulated power spectra derived assuming different faint-end slopes. 45
2.19 Comparison of the ratio between the simulated power spectra derived assuming different faint-end slopes and pairs of galaxies. 45
2.20 Example of distribution of the effective radius of the faint i-dropout galaxies in arcseconds. 46
2.21 Effect of the effective radius distribution on the power spectrum 47
2.22 Ratio between the powerspectra considering different percentages of E and S galaxies 48

2.23 Comparison of the observed ratio between power spectra obtained simulating mock galaxies in different magnitude ranges . . . 49
2.24 Distribution of the residuals . . . 50
2.25 Comparison between data and simultions . . . 51
2.26 Distribution of the total magnitude and the effective radius of the candidate i_{775} -dropouts . . . 52
2.27 Distribution of the ellipticity and PA of the i_{775} -dropouts candidates . . . 53

3.1 Lyman-break galaxy at $z \sim 7$ . . . 57
3.2 Pointings of the HUDF09 program . . . 58
3.3 Surface density of z_{850} -dropouts in each of the HUDF09 program field . . . 59
3.4 Surface density of Y_{105} -dropouts in each of the HUDF09 program field . . . 59
3.5 Persistence and cosmetic of the IR detector . . . 61
3.6 Effect of mock persistence images on single band power spectra . . . 62
3.7 Effect of mock persistence images on the ration between power spectra . . . 62
3.8 Scheme of how `MultiDrizzle` works . . . 65
3.9 Effect of the data reduction on the ratio between the J_{125} and Y_{105} power spectra for the HUDF09 main field . . . 68
3.10 Distribution of the total magnitude and effective radius in WFC3/IR data . . . 70
3.11 Distribution of the ellipticity and position angle in the WFC3/IR data . . . 70
3.12 Y_{105} and J_{125} -band power spectra for the HUDF09-01 field . . . 71
3.13 Y_{105} and J_{125} -band power spectra for the HUDF09 main field . . . 71
3.14 Ratio between the Y_{105} and J_{125} power spectra for the HUDF09-01 field . . . 72
3.15 Ratio between the Y_{105} and J_{125} power spectra for the HUDF09 main field . 72
3.16 Power spectra from simulations: noise . . . 74
3.17 Power spectra from simulations: noise and flat field . . . 74
3.18 Power spectra from simulations: noise and mock galaxies . . . 75
3.19 Effect of different masks on the Y_{105} and z_{850} -band power spectra . . . 77
3.20 Ratio between the z_{850} and Y_{105} power spectra for the HUDF09-1 field (square kernel) . . . 78
3.21 Ratio between the z_{850} and Y_{105} power spectra for the HUDF09-1 field (point kernel) . . . 78

4.1 XDF i_{775} and z_{850} -band power spectra . . . 83
4.2 Ratio between the z_{850} and i_{775} -band power spectra for the XDF . . . 83
4.3 Ratio between the Y_{105} and J_{125} -band power spectra for the XDF . . . 84
4.4 Satellite trail . . . 91
4.5 ACS quadrants readout directions and electronic ghosts . . . 98
4.6 Our new version of the XDF i_{775} and z_{850} -band image . . . 99
4.7 Comparison of the total magnitude for detected galaxies in HUDF, XDF, and our images . . . 100
4.8 Comparison of the S/N for detected galaxies in HUDF, XDF, and our images 101
4.9 Ratio between the powerspectra obtained from the v1.0 and v2.0 of the XDF dataset . . . 102

5.1 GOODS sections layout . . . 107
5.2 B_{435} and V_{606} total magnitude (HUDF-XDF) . . . 112

5.3 i_{775} and z_{850} total magnitude (HUDF-XDF) . . . 112
5.4 B_{435} and V_{606} total magnitude (GOODS v1.0-v2.0) . . . 113
5.5 i_{775} and z_{850} total magnitude (GOODS v1.0-v2.0) . . . 113
5.6 Comparison of the total magnitude for 21 B_{435} -dropouts matched in the HUDF and GOODS v2.0 . . . 115
5.7 GOODS v2.0-XDF comparison of the total magnitude . . . 116
5.8 Distribution of the normalized residuals for each simulation . . . 119

6.1 Location of candidate V_{606} -dropouts and contaminants with respect to the color-color selection box . . . 125
6.2 Location of candidate i_{775} -dropouts and contaminants with respect to the color selection box . . . 125
6.3 Number counts of contaminants and dropout galaxies (1) . . . 127
6.4 Number counts of contaminants and dropout galaxies (2) . . . 128
6.5 B_{435} -dropouts and contaminants in the color-color plot . . . 130
6.6 V_{606} -dropouts and contaminants in the color-color plot . . . 130
6.7 Number counts of B_{435} -dropouts and contaminants . . . 131
6.8 Number counts of V_{606} -dropouts and contaminants . . . 132
6.9 Distribution of photometric redshifts for B_{435} -dropouts and contaminants . 133
6.10 Distribution of photo-z for V_{606} -dropouts and contaminants . . . 133

A.1 Redshift distribution . . . 143
A.2 Bandpasses adopted for measuring equivalent widths . . . 145
A.3 $L - Z$ relation . . . 149

B.1 Histograms of the main properties of the galaxies in our sample . . . 154
B.2 Galaxy spectrum compared to the fitted linear combination of template stellar spectra and emission lines . . . 156
B.3 Age and metallicity obtained from the spectral fitting (1) . . . 157
B.4 Age and metallicity obtained from the spectral fitting (2) . . . 158
B.5 Age and metallicity obtained from the spectral fitting (3) . . . 159
B.6 Mass-weighted ages and metallicities for the dominant stellar component . 160
B.7 Distribution of luminosity-weighted ages and metallicities . . . 161
B.8 Correlation between morphological type and stellar population properties . 162

List of Tables

2.1 Properties of the NICP12 data . . . 25
2.2 Characteristics of the mock galaxies . . . 42
2.3 χ^2 statistics . . . 47

3.1 Properties of the WFC3/IR fields . . . 56
3.2 Properties of the HUDF09 data . . . 56
3.3 z_{850} and Y_{105} dropouts detected using WFC3/IR . . . 58

4.1 Data used for the reprocessing of the HUDF main field . . . 86
4.2 Images affected by satellite trails . . . 92

5.1 Properties of the HUDF dataset . . . 106
5.2 HDF-N and CDF-S coordinates in GOODS . . . 107
5.3 Characteristics of the GOODS dataset . . . 109
5.4 Final rescaling factors for the rms maps . . . 110
5.5 Best linear least-square fit . . . 114
5.6 Best linear least-square fit . . . 114
5.7 Best linear least-square fit parameters . . . 117
5.8 Rms values from the GOODS images and simulations . . . 117
5.9 Linear least-square fit from Monte Carlo simulations . . . 120
5.10 Results from the Kolmogorov-Smirnov statistic . . . 120

6.1 Number of dropouts . . . 126
6.2 Number of candidate dropouts and contaminants for the B_{435} and V_{606} -dropout samples . . . 129
6.3 Number of candidate dropouts and contaminants for the B_{435} and V_{606} -dropout samples with enlarged color cuts . . . 129

A.1 Measured properties of the sample galaxies. Col. (1): SDSS name of the galaxy. Col.(2-5): Equivalent widths of the $[O_{II}]\lambda 3727$, H_β, $[O_{III}]\lambda 4959$, and $[O_{III}]\lambda 5007$ emission lines, respectively. Col.(6): Measured redshift. . . . 144
A.2 EW intervals definition . . . 146
A.3 Properties of the 27 galaxies in our sample . . . 148

B.1 Properties of the galaxies in the sample . . . 164
B.2 Mass-weighted age and metallicity measured for the galaxies in our sample. 165
B.3 Mass-weighted age and metallicity measured for the young and old stellar populations considered separately. . . . 165

Abstract

The history of our Universe spans 13.7 billions of years and could be divided into several stages from the Big Bang up to now.
Around 370 000 years after the Big Bang ($z \sim 1100$) the temperature of the Universe lowered enough for the first simple atoms to form. Matter and radiation decoupled, and the Universe became transparent to radiation. Cosmic microwave background (CMB) photons that we detect nowadays were last scattered at $z \sim 1100$ and, since then, have been traveling in straight line. This is the reason why CMB is usually defined as the picture of the Universe at that redshift. Right after the CMB was emitted, the Universe entered the so called Dark Ages when no sources of light exist. Studying the early Universe, one of the most important phases is the subsequent phase-transition named reionization, i.e. the process that reionized the matter in the Universe after the formation of the first sources of light, namely the first stars and galaxies. Consequently, the detection and study of these objects are the key to unveil the early stages of the history of the Universe.

Imaging plays a more important role than spectroscopy in searching for high-redshift galaxies because it permits to observe more objects at the same time, better managing the telescope time. Deep surveys, obtained by observing the same sky area for several days, are the answer to the need for detections of high-redshift galaxies.

This thesis is focused on the study of the galaxy population existing when the Universe was less than 1.5 Gyr old. When studying the early Universe, the detection of high-redshift sources depends strongly on the detection limit of the survey and the surface brightness of the objects themselves. Taking this into account, we made use of the deepest datasets currently available obtained with the Hubble Space Telescope (HST) in both the optical and near-infrared (NIR) domain to carefully study how these two issues affect the identification and photometry of high-redshift galaxies.

The important role played by high-redshift galaxies in cosmic reionization is no longer debated and, lately, most of studies agreed on the key relevance of galaxies that are below the current detection limit. While we are waiting for the James Webb Space Telescope (JWST) to directly observe these faint galaxies, finding an alternative way to estimate their overall light contribution is mandatory. To this aim we developed a technique based on the power spectrum to analyze background fluctuations. Relying on a Lyman break-like approach we compared the power spectra of background signal derived from observations obtained in two adjacent bands to identify the light contribution from a population of galaxies lying within a specific redshift range. Then, Monte Carlo simulations permitted us to disentangle the information embedded in the light excess identified via power spectra, in particular deriving a constraint on the faint-end slope of the luminosity function.

The UDF05 dataset, follow-up of the original Hubble Ultra Deep Field (HUDF), consists in observations in the optical bands obtained with the Advanced Camera for Survey (ACS). It permitted us to constrain the slope of the luminosity function at $z \sim 6$ (0.95 Gyr after the Big Bang), which turned out the be steep enough to allow bright and faint galaxies at that redshift to account for the ionizing photon budget required for cosmic reionization.

The subsequent analysis aimed at deriving similar constraints on the faint-end slope of the luminosity function at $z \sim 7-8$ (between 0.64 and 0.77 Gyr after the Big Bang) using deep observations in the near-infrared obtained with the infrared channel of the Wide Field Camera 3 (WFC3-IR) during the HUDF09 program.

Regarding $z \sim 8$, the quality of the NIR dataset did not permit to disentangle any light produced by the faint galaxy population from the background noise and spurious signals. On the basis of the drop in the star formation rate density from $z \sim 6$ to $z \sim 7$ and beyond, there should be a more relevant contribution in terms of photoionizing photons at $z \sim 7$ than at $z \sim 8$ and we expected to be able to detect it. Unfortunately, the analysis at $z \sim 7$ implied dealing with different detectors that are characterized by systematics that can not be erased by simply considering the ratio of the power spectra. Up to now the understanding of all WFC3/IR related problems is not as good as for ACS and a further analysis is needed before being able to use the IR dataset for the analysis of surface brightness fluctuations.

Since a perfect reduction procedure of the images turned out to be an essential requirement to study any background signal, we performed an advanced data reduction to get an improved version of the deepest image of the Universe currently available, the so called eXtreme Deep Field (XDF). The goal was to create an image that allows to verify our findings on the faint-end slope of the luminosity function at $z \sim 6$ since the XDF did not permit us to get any constraint on background fluctuations. We started from raw frames obtained from several proposals over 10 years and created hyperbiases and hyperdarks taking into account all the issues affecting ACS data, including the minor ones such as the herringbone effect. Then, we masked the satellite trails, aligned all the frames, and corrected for the chip-to-chip jump. We are still working on the dataset, in particular we are focused on modelling and correcting for the electronic ghost. Anyway, the preliminary check on photometry suggests a promising, even though small, achievement in term of signal to noise of the sources.

The effect of surface brightness on the detection of primordial galaxies in deep surveys is directly depending on the cosmological surface brightness dimming that can be express in the form $(1 + z)^{-4}$ and that affects all the sources. The strong dependence of surface brightness dimming with increasing redshift suggests the presence of a selection bias when searching for high-redshift galaxies, i.e. we tend to detect only those galaxies with a high surface brightness. Unresolved knots of emission are not affected by surface brightness dimming, thus allowing, in principle, to test clumpiness within high-redshift galaxies. We followed an empirical approach based on HST legacy datasets characterized by different depth to study the surface brightness dimming of galaxies. We selected a sample of Lyman-break galaxies at $z \sim 4$ (1.5 Gyr after the Big Bang) detected in the XDF, HUDF, and the Great Observatories Origins Deep Survey (GOODS) datasets and found no significant trend when comparing the total magnitudes measured from images with different depth. Then, we compared our results to the prediction for mock sources derived from Monte Carlo simulations. In particular, considering different surface brightness profiles for the mock galaxies we were able to rule out all the extended profiles as fit for our data, getting a confirmation on the clumpy distribution of the light in high-redshift galaxies.

The study of cosmological surface brightness dimming is also important since it could affect our prediction of what the upcoming JWST can observe at higher redshifts, where younger galaxies may exhibit a larger fraction of clumpiness. Our direct comparison showing that galaxies detected in GOODS do not become significantly brighter in the HUDF suggests that most of their light is compact and hints to the fact that JWST will likely not find diffuse star forming components.

Finally, to complete the study on high-redshift galaxies we also focused on lower-redshift galaxies that could enter the high-redshift sample due to photometric scatter. In general

interlopers are galaxies at $z \sim 1-2$ showing colors similar to those of real dropout galaxies due to the 4000 Å break. Even though their colors are likely to include them in the dropout sample, contaminants have a non negligible detection in the bands blueward of the Lyman break.

The preliminary study we performed using the multi-wavelength catalog obtained from CANDELS GOODS-South shows that the number counts of contaminants are significantly different from those of dropout galaxies at $z \sim 5-6$ suggesting a clear difference in the luminosity functions of the two populations and little or no evolution in the population of interlopers entering the sample at different redshifts. Finally, we used the 3D-HST catalogs for the GOODS-South field that provided us with photometric data in ground based, HST, and Spitzer/IRAC bands as well as with photometric redshfits. This catalog allowed a study on the interlopers at $z \sim 4-5$.

Riassunto

I 13.7 miliardi di anni di storia dell'Universo possono essere suddivisi in diverse fasi a partire dal *Big Bang* fino ad arrivare al giorno d'oggi. Dopo l'emissione della radiazione cosmica di fondo (CMB) avvenuta 370 000 anni dopo il *Big Bang* ($z \sim 1100$), in seguito al disaccoppiamento fra materia e radiazione l'Universo è divenuto trasparente a quest'ultima dando inizio alla fase chiamata età oscura (*"Dark Ages"*), durante cui non era presente alcuna sorgente di luce. Nello studio dell'Universo primordiale riveste un ruolo chiave lo studio della transizione di fase, nota come reionizzazione, avvenuta in seguito alla nascita delle prime stelle e galassie. Lo studio delle prime sorgenti di luce che hanno popolato l'Universo è, quindi, la chiave per scoprire l'Universo primordiale e capirne l'evoluzione.

La ricerca di galassie ad alto *redshift* ha ottenuto una spinta fondamentale grazie alla fotometria e alle tecniche basate sull'acquisizione di immagini che, a differenza della spettroscopia, permettono lo studio simultaneo di più oggetti, ottimizzando il tempo di osservazione con i telescopi. In particolare le *survey* profonde, ottenute osservando la medesima regione di cielo per più giorni, sono la risposta alla necessità di identificare il maggior numero possibile di oggetti.

Questa tesi è focalizzata sullo studio delle prime galassie, già formate a meno di 1.5 miliardi di anni dal *Big Bang*. Lo studio dell'Universo primordiale dipende fortemente sia dal limite in magnitudine delle *survey* che dalla brillanza superficiale delle galassie che vogliamo osservare. Per caratterizzare l'effetto di entrambi questi fattori, in questa tesi abbiamo analizzato dati ottenuti con il telescopio spaziale Hubble (HST) sia nelle bande ottiche, che in quelle del vicino infrarosso (IR). L'obiettivo del nostro studio è stato capire come questi effetti influiscano e limitino l'identificazione e la caratterizzazione fotometrica delle galassie ad alto *redshift*.

Il ruolo chiave giocato dalle prime galassie nell'ambito del processo di reionizzazione è ormai assodato, ma studi recenti hanno mostrato come le galassie meno brillanti, e dunque al di sotto dell'attuale limite di osservabilità, possano aver avuto un'importanza maggiore rispetto alle galassie brillanti che sono state identificate fino ad ora. In attesa che il telescopio spaziale James Webb (JWST) possa osservare direttamente queste galassie poco brillanti, è fondamentale trovare un modo alternativo per stimare quale sia il contributo di luce totale proveniente da questi oggetti. Proprio a tale scopo abbiamo sviluppato una tecnica, basata sull'utilizzo dello spettro di potenza, per analizzare le fluttuazioni del segnale di *background*. Basandoci su un approccio simile al tecnica del *Lyman break*, utilizzata comunemente per l'identificazione di galassie ad alto *redshift*, abbiamo confrontato lo spettro di potenza del segnale di fondo in due bande adiacenti per isolare la luce prodotta da una popolazione di galassie deboli entro un ristretto intervallo di *redshift*. Grazie ad una serie di simulazioni di tipo Monte Carlo siamo riusciti a ricavare dall'eccesso di segnale nella banda più rossa un vincolo sulla pendenza α della funzione di luminosità alle magnitudini più deboli.

In particolare, utilizzando i dati del progetto UDF05, continuazione del programma Hubble Ultra Deep Field (HUDF), siamo riusciti ad ottenere un valore della pendenza α della funzione di luminosità a magnitudini deboli a $z \sim 6$ (corrispondente a 0.95 miliardi di anni dopo il *Big Bang*). I valori di $\alpha = -1.8$ e $\alpha = -1.9$ che abbiamo ottenuto tenendo o meno conto del *clustering* sono tali da permetterci di dire che la quantità di fotoni ionizzanti prodotti da tutte le galassie a $z \sim 6$ è sufficiente a giustificare il processo di reioniz-

zazione dell'idrogeno nell'Universo.

L'analisi che abbiamo condotto successivamente è indirizzata ad ottenere un valore di α a *redshift* più alto, nello specifico $z \sim 7-8$ (corrispondenti, rispettivamente, a 0.64 e 0.77 miliardi di anni dopo il *Big Bang*). Con questo obiettivo abbiamo fatto uso delle osservazioni profonde ottenute nell'ambito del programma HUDF09 nel vicino infrarosso.

Per quanto riguarda $z \sim 8$, la qualità delle immagini infrarosse non ci ha permesso di isolare alcun segnale prodotto dalle galassie deboli dal rumore di fondo e da segnali spuri. Sulla base del crollo della densità del tasso di formazione stellare andando da $z \sim 6$ a *redshift* più alti il contributo, in termini di fotoni ionizzanti, delle galassie a $z \sim 7$ dovrebbe essere maggiore rispetto a quello della popolazione a $z \sim 8$ e quindi ci aspettavamo che isolare un segnale prodotto da quelle galassie fosse più facile che a $z \sim 8$. In realtà, però, l'analisi a $z \sim 7$ prevede il confronto di dati ottenuti con strumenti diversi, caratterizzati da problemi diversi ed effetti sistematici che non vengono eliminati semplicemente considerando il rapporto fra gli spettri di potenza, come accade, invece, nel caso di immagini ottenute con la stessa camera. Inoltre, al momento, la nostra conoscenza dei problemi della camera infrarossa Wide Field Camera 3 (WFC3/IR) di HST non è ancora allo stesso livello di quella relativa alla camera Advanced Camera for Survey (ACS) ed è, quindi, necessario un ulteriore e più approfondito studio sugli effetti legati allo strumento prima di poter usare i dati infrarossi per lo studio delle fluttuazioni di *background*.

Avendo constatato che una perfetta procedura di riduzione dei dati costituisce un requisito essenziale per poter studiare le fluttuazioni di *background* e, quindi, il segnale proveniente dalle galassie meno brillanti, il passo successivo è stato ottenere una versione migliorata dell'immagine più profonda dell'Universo attualmente disponibile ottenuta nell'ambito del progetto eXtreme Deep Field (XDF). Il nostro obiettivo era di ottenere un'immagine tale da permetterci di verificare i nostri risultati sulla pendenza della funzione di luminosità a $z \sim 6$ dato che l'XDF non ci ha permesso di individuare alcuna fluttuazione nel segnale di fondo.

Il nostro lavoro è iniziato acquisendo dall'archivio le immagini non ridotte ottenute in diversi progetti durante un arco temporale di 10 anni e creando i nostri *hyperbias* e *hyperdark*. In questo abbiamo potuto tenere in considerazione tutte le problematiche legate alla camera ACS, anche quelle minori, come l'*herringbone effect*, che spesso non vengono corrette dalla normale *pipeline* di riduzione dati. Successivamente abbiamo mascherato le tracce dei satelliti, allineato tutte le immagini e corretto per la differenze di livello di fondo esistenti fra un *chip* e l'altro.

Al momento stiamo ancora lavorando su queste immagini, nello specifico con l'idea di modellare i *ghost* elettronici e correggere le immagini per questo problema. I test preliminari sulla fotometria delle nostre immagini sono promettenti e suggeriscono che ci sia un leggero guadagno in termini di rapporto segnale-rumore rispetto alla versione originale dell'XDF, ma solo un ulteriore lavoro ci permetterà di ottenere le immagini finali.

La brillanza superficiale incide direttamente sull'identificazione di galassie primordiali in immagini profonde sulla base dall'effetto noto come *dimming* cosmologico. Questo consiste in una diminuzione della brillanza superficiale di tutti gli oggetti astronomici estesi che scala con il *redshift* stesso di un fattore $(1+z)^4$. La forte dipendenza dell'effetto dal *redshift* suggerisce un effetto di selezione per il quale si individuerebbero più facilmente le galassie primordiali con un'alta brillanza superficiale rispetto a quelle più deboli. Va notato che, siccome l'effetto del *dimming* si ha soltanto su oggetti estesi, essendo legato alla brillanza superficiale, la ricerca di tracce di questo effetto può essere utilizzata per

testare quale sia la distribuzione di luce nelle sorgenti ad alto *redshift*. Nello specifico il *dimming* può aiutare a capire se l'emissione di luce sia concentrata in strutture compatte o meno.

Nel nostro studio abbiamo adottato un approccio empirico, confrontando dati provenienti da *survey* con profondità diversa, ma ottenute tutte con lo stesso strumento, in particolare HST/ACS. Abbiamo concentrato la nostra attenzione su un campione di galassie a $z \sim 4$ (corrispondente a 1.5 miliardi di anni dopo il *Big Bang*) identificate nelle immagini XDF, HUDF e GOODS e, confrontando le magnitudini totali derivate dalle diverse *survey*, non è emerso alcun andamento nei dati che fosse imputabile al *dimming* cosmologico. Per completare il lavoro abbiamo, poi, effettuato delle simulazioni Monte Carlo per ricavare quale sarebbe il riscontro sui dati se il *dimming* fosse in atto a secondo del tipo di distribuzione di luce nelle galassie ad alto *redshift*. Confrontando i risultati delle simulazioni con quelli ricavati dai dati è stato possibile escludere i profili di brillanza superficiale caratteristici delle sorgenti estese. I nostri dati sono, quindi, in accordo con una distribuzione della formazione stellare ad alto *redshift* in strutture compatte, come sostenuto anche da altri gruppi di ricerca.

In generale lo studio degli effetti del *dimming* cosmologico riveste un ruolo molto importante nella determinazione del tipo di oggetti che JWST potrà osservare. Essendo in grado di spingere il nostro orizzonte verso fasi della storia dell'Universo ancora più vicine al *Big Bang*, JWST permetterà di verificare in maniera più accurata sia la distribuzione della luce nelle galassie primordiali sia il contributo proveniente dalle sorgenti che al momento non riusciamo a vedere singolamente.

Infine, per completare il nostro studio sulle galassie ad alto *redshift* abbiamo preso in considerazione gli oggetti a $z \sim 1-2$ che possono contaminare i cataloghi di galassie primordiali. Lo spettro di questi oggetti mostra un *break* a 4000 Å che può portarli ad avere dei colori molto simili a quelli delle vere galassie ad alto *redshift*. Ciò che distingue i contaminanti dalle vere galassie primordiali è l'essere chiaramente identificabili nelle bande più blu del *Lyman break*.

Lo studio preliminare che è stato condotto sulla base di due cataloghi pubblici multibanda di tutte le sorgenti presenti nel campo GOODS-South mostra che la distribuzione del numero di contaminanti e delle vere galassie primordiali è diversa nell'intervallo di *redshift* che abbiamo analizzato ($z \sim 4-5-6$). Questo suggerisce che la popolazione contaminante abbia una funzione di luminosità differente rispetto alle galassie ad alto *redshift*. Inoltre, si ricava dai dati che la distribuzione dei conteggi di contaminanti a basso *redshift* mostra solo una leggera, se non assente, evoluzione. La popolazione contaminante è, quindi, sempre all'incirca la stessa, e questo è confermato dal leggero spostamento del *break* a 4000 Å richiesto perché questi oggetti soddisfino i criteri di selezione in colore delle galassie a $z \geq 4$. Anche i *redshift* fotometrici sono in accordo con l'ipotesi che i contaminanti siano oggetti a *redshift* più basso.
Una più completa caratterizzazione della popolazione delle galassie che hanno colori simili a quelle primordiali sarà possibile quando nuovi cataloghi saranno resi disponibili.

List of Acronyms and Symbols

General

AGN	Active Galactic Nuclei
EBL	Extragalactic Background Light
FWHM	Full Width Half Maximum
HI	neutral Hydrogen
HII	ionized Hydrogen
He	neutral Helium
IR	infrared
IGM	Inter-Galactic Medium
IMF	Initial Mass Function
IR	infrared
ISM	Inter-Stellar Medium
LBG	Lyman Break Galaxy
LF	Luminosity Function
Lyα	Lyman α line
PSF	Point-Spread Function
SB	Surface Brightness
SED	Spectral Energy Distribution
S/N	signal-to-noise
SFR	Star Formation Rate
QSO	Quasi-Stellar Object
UV	ultraviolet
z	redshift

Telescopes and Instruments

ACS	Advanced Camera for Survey
CTE	Charge Transfer Efficiency
HST	Hubble Space Telescope
JWST	James Webb Space Telescope
NICMOS	Near Infrared Camera and Multi-Object Spectrometer
VLT	Very Large Telescope
WFC3-IR	infra-red channel of the Wide Field Camera 3
WFPC2	Wide Field and Planetary Camera 2

Surveys, Archives, and Fields

CANDELS	The Cosmic Assembly Near-IR Deep Extragalactic Legacy Survey
GOODS	Great Observatories Origins Deep Survey
HDF-N/S	Hubble Deep Field North/South
HUDF	Hubble Ultra Deep Field
HUDF09	WFC3/IR follow-up survey of the HUDF in the IR
HUDF12	Hubble Ultra Deep Field 2012, follow-up of the HUDF
MAST	Mikulski Archive for Space Telescopes Archive
NICP12	parallel field observed in the UDF05 survey
NICP34	second parallel field in the UDF05 survey
UDF05	ACS follow-up survey of the HUDF parallel fields
XDF	eXtreme Deep Field

Chapter 1

Introduction

A cosmological model is a mathematical description of the Universe which aims to explain its current behavior and evolution over time.

Recently, a model of expanding Universe, based on a small number of accurately constrained parameters, was developed. It is the so-called Λ Cold Dark Matter (ΛCDM) cosmological model. In this cosmological scenario, only $\sim 4\%$ of the total energy density of the Universe is due to the well studied baryonic matter, the remaining is due to dark matter (hereafter DM, $\sim 24\%$) and, mainly, dark energy ($\sim 71\%$). The ΛCDM model is based on six parameters: baryon density, dark matter density, dark energy density, scalar spectral index, curvature fluctuation amplitude, and reionization optical depth.

As shown by theoretical, semi-analytic or simulation studies, such as Ciardi et al. (2000); Gnedin (2000), in the ΛCDM hierarchical scenario star-forming galaxies are able to ionize the Universe in the redshift range $z \sim 6 - 15$ with reasonable parameter choices.

1.1 A Brief History of the Universe

The history of the Universe can be divided into a series of subsequent phases, as shown in Figure 1.1. Following the Big Bang scenario, the Universe formed 13.7 billion of years ago. The main stages after the initial explosion were:

- Inflation, started 10^{-37}s after the Big Bang. During this phase the Universe exponentially expanded and the volume increased, at least, by a factor 10^{78} (Guth, 1981).
- Emission of the cosmic microwave background radiation (CMB). At $z \sim 1100$ the first atoms formed, matter and radiation decoupled, and the Universe became transparent to radiation.
- "Dark Ages", when the Universe was characterized by no sources of light.
- Era of galaxies. The birth of the first stars, galaxies, black holes and supernovae. These first sources of light marked the end of the "Dark Ages".
- Mergers of protogalaxies, which led to the formation of the actual, currently existing, galaxies.

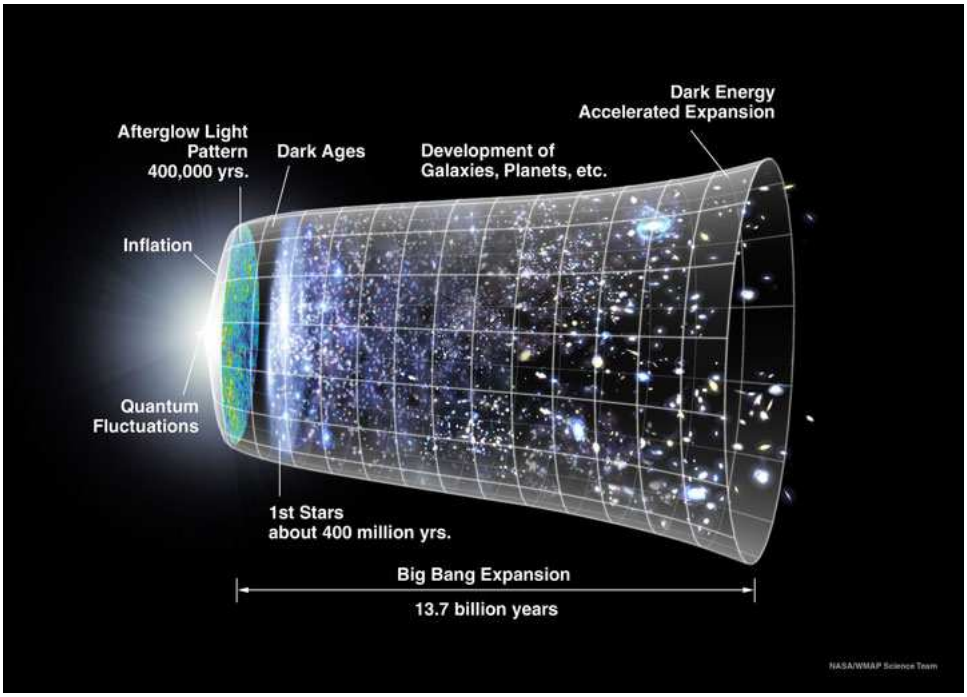


Figure 1.1: The evolution of the Universe over 13.7 billion years, from the Big Bang (left side) to the current epoch (right side). The main events, such as inflation, Dark Ages, the birth of the first stars, are highlighted.

Despite the recent progresses of observational cosmology the history of the Universe is still characterized by a period, spanning between 370,000 and 1 billion years after the Big Bang and called "Dark Ages", which is completely unknown. In fact, the CMB can not give us information on it because the baryonic matter and radiation already decoupled during this epoch. Moreover, the existing telescopes are only able to detect high-z galaxies and quasars at the very end of the "Dark Ages". Probing this epoch is, indeed, at the forefront of modern astrophysics and cosmology.

The phase transition at the end of the Dark Ages was connected to the formation of the first sources of light according to the following steps:

- the growth and virialization of DM halos
- the collapse of baryonic structures within the DM potential wells
- HII cooling which led to star formation in low mass halos.

After that the Universe entered the phase termed epoch of reionization when the very first luminous sources ionized the neutral intergalactic medium (hereafter IGM). The cosmic reionization marked a significant shift in the evolution of the Universe, from the relatively simple evolution of the first phases to more complicated and interconnected processes. This period will be probed most thoroughly by the next generation of telescopes, in the optical, infrared, and radio range.

1.2 Cosmic Reionization

Cosmic reionization is one of the most important topics in observational cosmology today. As we said in the previous section, the deep observations obtained in the last years permitted us to improve the study of the history of Universe, especially from $z \sim 6$ up to the current day, but we are still trying to investigate both the "Dark Ages" and the reionization epoch. There are multiple reasons for studying, in particular, the reionization (Stiavelli, 2009). First of all, reionization of hydrogen represents a phase transition which affected the range of viable masses for galaxies. In fact, before the reionization, small galaxies formed easily since they were shielded by neutral hydrogen, but, after this phase, the formation of such systems was hampered (Efstathiou, 1992; Quinn et al., 1996; Dijkstra et al., 2004). Moreover, studying the reionization epoch it is possible to probe the power spectrum of those density fluctuations which arose from the recombination phase and which are at smaller scales than the ones accessible by the experiments regarding the cosmic microwave background (CMB). Finally, in the hierarchical scenario the first objects producing photons can be considered the seeds for the formation of the following larger objects, so the reionization epoch could have constrained the formation of current giant galaxies and could explain the connection between the growth of black holes and the evolution of their hosting galaxies.

The most relevant issue in studying reionization is the identification and subsequent analysis of the signal from high-z sources, in particular the disentanglement from Galactic and extragalactic foreground contaminants, as well as from noise.

One of the most promising ways to study the transition from Dark Ages to the epoch of reionization relies on the redshifted 21cm emission line from neutral hydrogen at high redshifts (Zaroubi, 2013). For this reason, a number of dedicated telescopes were recently built up, such as the Low Frequency Array (LOFAR), the Murchison Widefield Array (MWA), the Precision Array to Probe Epoch of Reionization (PAPER) and the Giant Metrewave Radio Telescope (GMRT).

The main questions regarding cosmic reionization are:

- When did reionization occur?
- Which sources caused it?

The answers are likely to be different for hydrogen (H) and helium (He).

The answer to the first question can be inferred from observational evidences. The absorption spectra of quasars at $z \sim 6$ in the Sloan Digital Sky Survey (SDSS) indicate a relevant increase in the neutral fraction of hydrogen at $z > 6$ (Becker et al., 2001; Fan et al., 2002), while, from the UV spectrum of quasars, we can derive that helium is fully ionized only at $z < 3$ (Kriss et al., 2001; Smette et al., 2002; McQuinn, 2012). Regarding quasars, their photoionizing contribution increases as we look back in time from $z = 0$ to $z \sim 2$, as the peak of the quasar luminosity function is approached (Fanidakis et al., 2011). Beyond $z \sim 2$ their contribution significantly decreases and, probably, the stellar emission is responsible for the photoionizing budget required (Siana et al., 2008; Faucher-Giguère et al., 2009; Haardt and Madau, 2012). Star-forming galaxies at $z > 3$ are, therefore, the main candidates to provide the remaining ionizing photons.

The full reionization of helium at $z \sim 3$ was the last major phase change of the IGM. Helium reionization began at redshifts $z \sim 3.5 - 4$ (Becker et al., 2011), maybe even at

higher redshift as stated by Bolton et al. (2012).

It is probable that the same sources that reionized hydrogen caused also the single reionization of helium, from HeI to HeII. Neutral helium can be ionized by photons with an energy of 24.6 eV or higher, and its recombination rate is roughly equal to the hydrogen one. On the other hand, HeII requires an ionizing energy of 54.4 eV, and it recombines at least 5 times faster than H. This means that, even considering quasars or galaxies, HeII should reionize later than HI, even though the number of helium atoms is 13 times smaller than the number hydrogen ones. On the basis of above, since HeII reionize at lower redshift, this process is more accessible to observations and can be considered as an observational preview of hydrogen reionization (Barkana and Loeb, 2001).

Regarding the role of galaxies in cosmic reionization an important unknown elements is the so called escape fraction f_c, i.e. the fraction of ionizing radiation that escapes the galaxy into the IGM. These ionizing photons are the relevant one in order to reionize the Universe. Determining the escape fraction of ionizing radiation on the basis of observations is very difficult because it is hard to get information at high redshift. Nevertheless, Inoue et al. (2005); Shapley et al. (2006); Iwata et al. (2009) found the value of this parameter to be in the range 0.1-0.5. Even theoretical studies (Ciardi et al., 2002) considering different sources of reionization agree with the value to be in that range.

In Stiavelli et al. (2004b) a general model, which can be applied to a variety of different reionizing sources, is described. They considered galaxies with either Population III or Population II stars, modeled as black bodies with $T = 10^5$ K and $T = 5 \cdot 10^4$ K, respectively. In Figure 1.2 the expected cumulative surface density of reionizing sources is shown as a function of the apparent AB magnitude in the non-ionizing UV continuum at $\lambda = 1400$ Å rest frame. Assuming a Schechter luminosity function (for more details on this see Sec. 1.6), characterized by the knee luminosity M_* and the faint-end slope α, and considering the reionization process ended at $z = 6$, the 3 rows of plots differ in the values of M_* and α, as specified in the bottom label in each stamp plot. In particular, the middle panels have a brighter knee and the bottom ones a steeper α slope. The left panels refer to predictions for Population III stars, the right ones for Population II. The key idea is that the mean surface brightness (SB) of reionizing sources depends on the fraction of photons in the Lyman continuum escaping from them (f_c) and, at the same time, on the clumping factor characteristic of the IGM (C). For this reason the different dotted lines plotted in each stamp are labeled with the corresponding values for f_c and C. The lower limit or the minimum mean surface brightness required to justify the reionization lies on (f_c, C)=(1, 1), i.e. all the ionizing photons are escaping from the galaxy and the recombination is the lowest possible. The upper limit, instead, is constrained by the amount of metals produced. Assuming $Z < 0.01 Z_\odot$ at $z \sim 6$, the shaded regions in Figure 1.2 are excluded, while the clear ones are allowed.

The predictions from Stiavelli et al. (2004b) are based on the assumption that reionization occurred over a $\Delta z \sim 1$ interval just above $z \sim 6$. The observed integrated surface brightness is consistent with the one required to reionize the Universe assuming reasonable values for both the escape fraction and the clumping factor, i.e. $0.05 < f_c < 0.5$ and $1 < C < 30$, and considering sources which are composed either of Population III or Population II stars with a top-heavy IMF. Reionization by sources with a Salpeter IMF is more difficult, and requires, at the same time, a high value for f_c and a very low metallicity. Even in the demanding case of reionization limited to a short redshift interval of $\Delta z \sim 1$, the observed objects are sufficient to reionize the IGM. Of course, if the reionization process

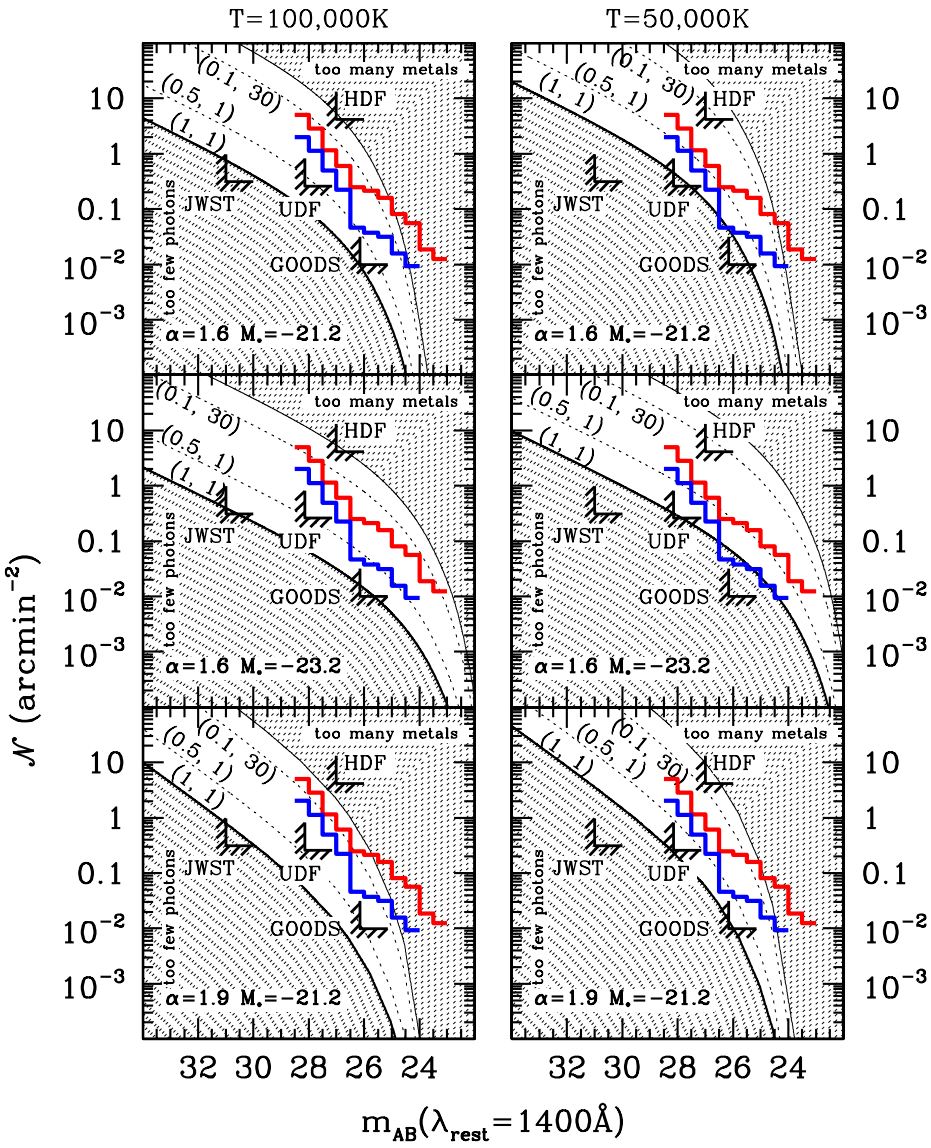


Figure 1.2: Expected cumulative surface brightness density of reionization sources as a function of the apparent magnitude (AB mag). Comparison between model predictions for Lyman Break Galaxies (LBGs) and data from UDF and GOODS survey. The left and right panels differs in the stellar population consider in the models, in particular a very metal poor top heavy one (Pop. III) on the left, and a top heavy one with the typical Pop. II metallicity on the right. The different stamps plot refer to different values of the parameters α and M_*, as specified in the label. The white areas are the only allowed. The dotted lines show the surface brightness required to complete the reionization and depends on f_c and C, as labeled in parenthesis (Stiavelli, 2009).

began earlier the requirement would be less strict.

1.2.1 The Reionization Process

The reionization of hydrogen itself can be split in sub-epochs basically on the basis of the percentage of ionized IGM in the Universe.

Gnedin (2000) proposed to divide the reionization process into three different phases. During the first one, called the "pre-overlap" phase, the ionized regions of individual sources propagated into the neutral intergalactic medium. In the second one, the "overlap" phase, the ionized regions started overlapping and, then, ionized most of the IGM, except for high-density peaks. The Universe became transparent to the UV radiation and, consequently, the mean free path of photons increased dramatically. Finally, during the third and last phase, the ever-continuing "post-overlap" phase, the ionization fronts propagated into the neutral high-density regions.
On the basis of the previous facts, the product of star formation efficiency and ionizing photon escape fraction from galaxies at high-z is well constrained and this determines the evolution of the reionization process.

Unlike the scheme followed by Gnedin (2000), Cen (2003) separated the entire reionization process into four stages. From $z \sim 30$ to $z \sim 15-16$ the first stars (Population III stars) gradually heated up and reionized the IGM up to reach the first reionization. Then, during a period of the order of $\Delta z \sim 1$, the IGM remained completely ionized thanks to the ionizing photons produced by forming Population III stars. The third stage started with the transition to Population II stars at $z \sim 13$. This caused a decrease in the amount of ionizing photons of a factor of ~ 10, consequently the hydrogen rapidly recombined and the Universe became opaque to Lyα and Lyman continuum photons, leading to the second recombination phase. Finally, the last stage started at $z \sim 6$ when the Universe reionized for the second time.

Apart from the details, we can summarize how reionization unfolded. At the very beginning the Universe was neutral, then the first sources appeared. These objects, producing ionizing photons, were rare and only able to reionize small bubbles of IGM surrounding them. Then, as the Universe expanded, the density of the IGM decreased and, consequently, the number and size of the ionized bubbles increased. As the bubbles started to overlap the process evolved quickly leading to a fully ionized Universe.

Recent studies on the cosmic microwave background it constrains reionization to be roughly 50% complete between $z = 9$ and $z = 11.8$, although the results depend on the shape of the assumed reionization history (Mitra et al., 2012; Hinshaw et al., 2013).

1.2.2 Possible Sources for Reionization

The sources of UV photons expected to contribute to the reionization process are mainly two, star forming galaxies and quasars. The former probably have been fundamental for the hydrogen reionization while the latter dominated the helium reionization (Madau et al., 1999; Shapiro and Giroux, 1987). However, the role of QSOs in the cosmic reionization is still debated. According to Fan et al. (2001); Meiksin (2005); Shankar and Mathur (2007), $z \sim 6$ quasars are not able to provide enough ionizing photons to contribute significantly to reionization. On the contrary, Fontanot et al. (2012) found that at $z \sim 6$ we cannot

neglect the QSO contribution at the observational limit $M_{UV} \sim -20.6$, even if it is still insufficient to provide the required rate of ionizing photons. At higher redshift they found a rapid decrease in the contribution of AGNs to the ionizing background, down to the order of a few percent at $z \sim 8$, and a totally negligible one thereafter. So, in order to achieve the reionization of hydrogen at $z \sim 7$, the AGN population should present extreme properties (i.e. steep faint end of the LF and contribution of very faint objects), but this would imply the reionization of HeII to take place above redshift $z \sim 5$, which is in contrast with the observational results by Fechner et al. (2006) and Zheng et al. (2008). Summarizing AGN alone cannot be responsible for the reionization of hydrogen.

Nevertheless, we should mention that the possibility of other sources contributing to cosmic reionization can not be completely ruled out, at least from an observational point of view. For this reason, in the last years various other possible sources have been studied too, the early ones being the supernova- driven winds (Tegmark et al., 1993), hard photons from structure formation (Sasaki and Takahara, 1994) early formed massive black holes (Sasaki and Umemura, 1996) and more exotic sources, for example decaying dark matter particles (Sciama, 1988; Dodelson and Jubas, 1992, 1994; Sciama, 1994, 1995; Sethi and Nath, 1997). Accreting black holes in the redshift range between $z = 6$ and $z = 8$ are rare and possibly obscured by significant amounts of gas and dust, so their UV emission does not contribute considerably to the reionization process (Treister et al., 2011).

Population III stars have a chemical composition very different from Population I and II stars since they formed in the primordial Universe and, thus, contain only hydrogen and helium. The pre-stellar clouds are poor radiators until they reach high temperatures. For this reason they are very massive and, consequently, very efficient and abundant sources of UV photons, even if their lifetime is short. Nevertheless, Population III stars are unlikely to dominate the ionizing UV production. In the standard scenario, in which only a single metal free star per halo is formed, the contribution to reionization given by Population III stars is mostly indirect. In fact, while the cumulative number of ionizing photons produced by them is not enough, they enrich the intergalactic medium of metals and allow the more efficient formation of following generations of Population II stars (Trenti and Stiavelli, 2009). It is important to notice that, according to Loeb and Barkana (2001), these first stars are likely located in faint high-redshift galaxies.

On the basis of above, star-forming galaxies represent the most obvious source of ionizing photons. However, since, even in the Hubble Ultra Deep Field, Lyman break galaxies at $z \sim 6$ (Stiavelli et al., 2004b) and $z \sim 7$ (Bouwens et al., 2004b; Bolton and Haehnelt, 2007; Oesch et al., 2009) are not bright enough to account for the ionizing budget required and given the low volume densities of high-redshift galaxies at high luminosities, a fundamental point is whether reionization can be accomplished through a significant population of very faint galaxies as stated by Yan and Windhorst (2004), Richard et al. (2008), and Trenti et al. (2010). According to Fontanot et al. (2012) in order to achieve reionization at high redshift considering only the ionizing photons coming from LBGs, a substantial contribution to the ionizing background is required from sources fainter than the actual observational limit or it is necessary an increasing f_c with respect to their bright counterparts. Theoretical models (Trenti et al., 2010; Salvaterra et al., 2011) predict the existence of this faint galaxies population, which can be proved through the measurement of the faint-end slope of the luminosity function and extrapolation to fainter limits, as well as through future extraordinarily deep observations with the James Webb Space Telescope.

However, we should remind that a relevant contribution to reionization from the faint

galaxies would imply an extended reionization epoch, and this result is disfavored by recent constraints on kinetic Sunyaev-Zeldovich effect (Zahn et al., 2012; Kuhlen and Faucher-Giguère, 2012). Therefore, if the faint dwarf galaxies would not provide the required amount of ionizing photons, it would be necessary to test, for example, a strong luminosity/redshift evolution of the escape fraction f_c or additional sources of ionizing photons.

In the last few years the reionization of neutral hydrogen has been investigated by using a variety of techniques. Observationally, we have constraints on reionization from the Gunn-Peterson troughs in the spectra of $z \sim 6$ QSOs (Fan et al., 2002, 2006), the Compton optical thickness measured by WMAP (Jarosik et al., 2011), and the luminosity function and clustering properties of Lyα emitters (Ouchi et al., 2010). According to Komatsu et al. (2011) reionization should have begun at least as soon as $z \sim 11$, while Fan (2007) and Ouchi et al. (2010) constrained the end of the process to be happened no later than $z \sim 6$.

While waiting for the new generation of telescopes that will permit us to investigate faint objects in the early phases of cosmic history, some research groups are exploiting strong lensing to study faint galaxies at $z \sim 6-8$ (Bradley et al., 2013). Vanzella et al. (2013) report the first two spectroscopically confirmed faint galaxies at $z = 6.4$ selected from strong lensed images but the sample is still too small to be able to address the questions regarding the role of this galaxies as potential drivers of the reionization process.

Thus, in order to make further progress in this field, it is necessary to find ways to study the faint end of the galaxy luminosity function at $z \geq 6$ using the integrated light contribution of individually undetected galaxies. This idea is not new and it has been tried in the context of high-redshift studies using Spitzer IRAC and ACS data (Kashlinsky et al., 2007) and is, in fact, also analogous to the study of surface brightness fluctuations of nearby galaxies (Tonry and Schneider, 1988).

Hereafter we will call undetected galaxies those galaxies below the instrument detection limit. The aim of this thesis is to describe a new technique to analyze the contribution of faint, individually, undetected galaxies and to use it to constraint the faint-end slope of the luminosity function (hereafter LF) at $z \sim 6$, $z \sim 7$ and $z \sim 8$. On the basis of these extended LFs the contribution to cosmic reionization will be calculated.

1.3 The Lyman-Break Technique

High-redshift galaxies are very distant objects and, since light propagates through space at a speed $c \sim 300000 \text{ km s}^{-1}$, they appear to us as they were in the past, when the light departed them. For this reason observations of very high-z objects play a central role in cosmology, providing an insight into the primeval Universe and permitting to unveil the mechanisms leading galaxy formation.

Being extremely far away, galaxies at high redshift appear as faint objects, irrespective of their absolute magnitude, and that is the reason why the search for such objects is not trivial and relies on different techniques with respect to nearby objects.

Nearby star-forming galaxies show spectra which are dominated, in the range between 3500 and 7500 Å, by strong, narrow emission lines of hydrogen, oxygen and nitrogen produced in the nebulae associated with young massive stars. On the contrary, once an high-redshift galaxy is observed, the spectrum obtained using an optical telescope is the rest-frame ultraviolet one, between 1000 and 2000 Å, where the only strong nebular emission line is the Lyα one, characterized by a rest wavelength of 1215.67 Å. Unfortunately, the first

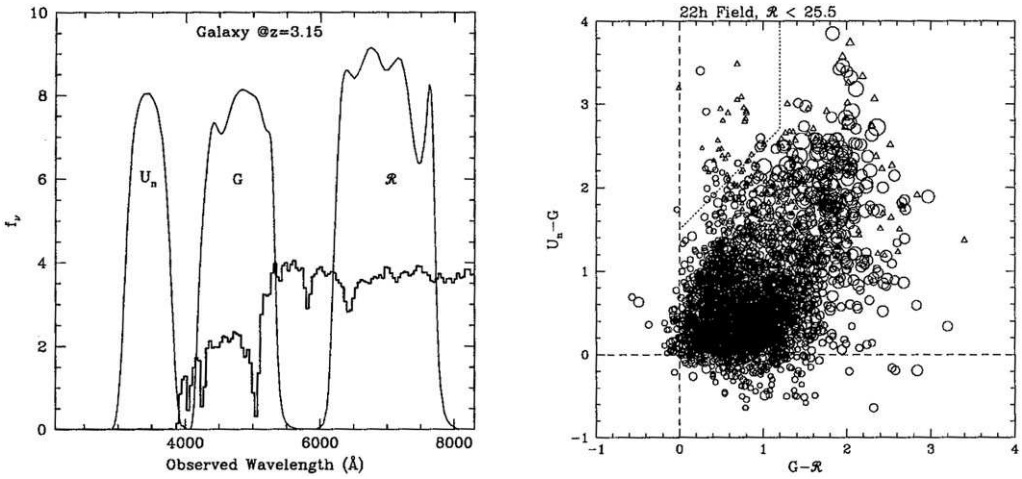


Figure 1.3: Left panel: example of photometry in the region across the 912 Å Lyman-continuum discontinuity. The U, G, and R filters are located on both sides of the break on the UV SED of a $z \sim 3.5$ star-forming galaxy (Giavalisco, 1998). Right panel: example of the color selection criteria used to identify LBGs. In the color-color plot the dashed lines indicate the selection window for LBGs at $z \sim 3$ derived from deep UGR imaging in the SSA22 field taken with the COSMIC camera at the 200-inch Palomar telescope (Giavalisco, 1998).

spectroscopic observations searching for Lyα emission did not give good results since it is difficult for Lyα photons to escape the galaxy where they are produced, given the large cross-section for absorption and subsequent scattering by HI atoms.

Since the spectroscopic identification of high-redshift galaxies was not successful, a new method was developed by Steidel and Hamilton (Steidel and Hamilton, 1992, 1993). It relies on the existence of a discontinuity in the far-UV spectrum of any star forming galaxy which is due to the limit of the Lyman series near 912 Å, the wavelength of photons with sufficient energy to ionize hydrogen. In the absence of dust extinction, an actively star-forming galaxy should have a blue continuum at rest frame ultraviolet wavelengths, nearly flat in f_ν units. Blueward of the 912 Å Lyman limit, however, photoelectric absorption by sources of HI will sharply truncate the spectrum (Dickinson, 1998; Pettini et al., 1998).

This Lyman-break is characterized by a threefold origin: first of all the drop in the spectra of hot O and B stars which dominate the integrated spectrum in the ultraviolet (Cassinelli et al., 1995), then the absorption due to neutral interstellar medium within a star-forming galaxy (Leitherer et al., 1995), and, finally, the opacity of the intergalactic medium.

Basically, there are two methods which can be used to select high-redshift galaxies using broad-band images obtained in the optical and/or IR domain: photometric redshifts and the Lyman-break technique.

The technique used to derive photometric redshifts, introduced for the first time by Baum (1962), is based on the detection of strong spectral features, like the 4000 Å break, Balmer-break, Lyman-break or strong emission lines, and relies on multiple band observations. It has the advantage of sensitivity which enables to get photometric redshifts of large samples of faint and high-redshift galaxies. On the other hand, photometric red-

shifts, derived from broadband fluxes, tend to be affected by strong degeneracy between low-z galaxies, showing a Balmer break, and the high-z ones, which are characterized by the Lyman-break feature.

The Lyman-break technique, based on a color-color selection, is a simple form of photometric redshift selection. The aim of this technique is not an accurate estimate of the redshift of individual objects, but the use of color criteria to select galaxies within a particular redshift interval and, at the same time, to exclude foreground and background spurious objects. The selection function of the method depends strongly on the adopted color criteria, on the intrinsic dispersion in the UV spectral properties of star forming galaxies at high-z, on cosmic variance in the intergalactic transmission along different lines of sight, and on the distribution of photometric measurement errors (Dickinson, 1998).

The Lyman-break technique permits to identify star-forming galaxies at high redshift using color criteria derived from a multi-band photometry in the region across the 912 Å Lyman-continuum discontinuity. The left panel of Figure 1.3 shows the position of the U, G, and R filters (200-inch Palomar telescope) on both sides of the break for a galaxy at $z \sim 3.5$ (Giavalisco, 1998). Comparing the flux within this set of filters it is possible to detect $z \sim 3$ objects. In fact, they are really faint in the U band, i.e. basically no flux is collected by the bluer filter, but are regularly detected in the redder bands. This is the reason why $z \sim 3$ galaxies are usually defined U-dropouts.

Using a different set of filters, shifted toward the red part of the spectrum, the Lyman-break technique permits to look for galaxies at higher redshift. So galaxies at $z \sim 4$, which are not detected in the B-band are called B-dropouts, those at $z \sim 5$ are not detected in the V band and are called V-dropouts, those at $z \sim 6$ are i-dropouts, and so on.

As can be seen from the right panel of Figure 1.3, high-redshift galaxies fall in a well-defined region of the color-color plot, irrespective of their spectroscopic type. So high-redshift galaxies, observed in three bands, can be identified on the basis of the Lyman-break and color selection criteria.

Partridge and Peebles (1967) were the first who considered the idea of exploiting the Lyman limit as a spectral feature able to identify star-forming galaxies which are far away, but to implement it observationally it was necessary to wait for sensitive detectors, especially in the U band. Guhathakurta et al. (1990) and Songaila et al. (1990) obtained the first observations of high-z galaxies using the multi-band imaging and constrained the surface density of $z > 3$ galaxies, but Steidel and Hamilton (1993) were the very first to use the Lyman-break technique to select high-z objects. However, their data were not really deep and covered such a limited sky area to not present a significant sample of robust candidates. Even using successive, deeper, observations taken at the New Technology Telescope, the William Herschel Telescope and the 200-inch Hale Telescope, the very high-redshift galaxy population was not detected. It was the Hubble Deep Survey which permitted to push even further the Lyman-break technique, looking for galaxies at $z > 3$.

Since the Hubble Deep Field (HDF) data were obtained through several filters, it was possible the use color selection techniques to isolate and study populations of galaxies at different redshifts. Moreover, because the HDF data were deeper than those previously used, the colors were measured with an higher precision, leading to a more precise identification of LBGs.

A comprehensive study on the Lyman-break technique and possible contaminants can be found in Stanway et al. (2008). One of the main drawbacks of the technique they highlight is being biased towards objects with a bright rest-frame UV continuum, i.e. starbursts.

Galaxies that are over their starburst phase and are passively evolving are not likely to be identified by this technique, as well as those that show a very young sturburst.

1.3.1 Lower-redshift Contaminants

Each sample of dropout galaxies identified through the Lyman-break technique is affected by contaminants. There are two classes of interlopers that can enter any color-based high-z selection: cool galactic stars and galaxies at lower redshift.

As described in detail in Stanway et al. (2008), M stars satisfy the color criteria for the V and R-dropouts. L and T stars show, instead, the same colors of i-dropouts. These stars, anyway, mainly do not affect HST catalogs of high-z galaxies since they are rejected on the basis of their full-width half-maximum (FWHM).
In HST-based selection the major source of contamination is, instead, from galaxies at lower redshift that enter the color-color selection box due to photometric scatter.

Among the strong spectral features redward of the Ly$-\alpha$ line, the 4000 Å Balmer break is the main responsible for the red colors that lead to a wrong identification. As shown in Hayes et al. (2012), claims of detection at very high-z based on color-color selection can be easily refuted by a spectroscopic follow-up if contamination from lower-z galaxies is not properly taken into account.

No study has been done yet to study the luminosity function of the contaminant population, even though a better understanding of the interlopers issue would be very helpful for both existing high-z studies and the upcoming JWST surveys.

Up to know the only estimates on the percentage of interlopers in dropout samples come from simulations, as shown, for example, by Su et al. (2011) and Pirzkal et al. (2013). These studies agree on an estimate of about 20% of contaminants.

1.4 Deep Imaging Surveys

As we said before, high-redshift galaxies are possible sources of ionizing photons. When searching for these objects the imaging technique plays a more important role than spectroscopy because it permits to observe more objects at the same time, better managing the telescope time.

Deep surveys are obtained observing for several days the same sky area. This permit to reach faint magnitude limits but has the drawback to investigate only a tiny stamp-size region in the sky.
To select the suitable field for a deep survey it is necessary to satisfy a few general criteria, as listed in Stiavelli (2009). First of all the area should be characterized by low galactic dust extinction and by lack of cirrus. Moreover, the absence of bright sources of all kinds is required.

1.4.1 A Brief History of HST Deep Surveys

To date, the deepest imaging data are part of surveys carried out with the Hubble Space Telescope.

In 1995 the Hubble Deep Field North (hereafter HDFN, Williams et al. 1996; Ferguson et al. 2000) was the first public imaging survey, followed in 1998 by the Hubble Deep Field

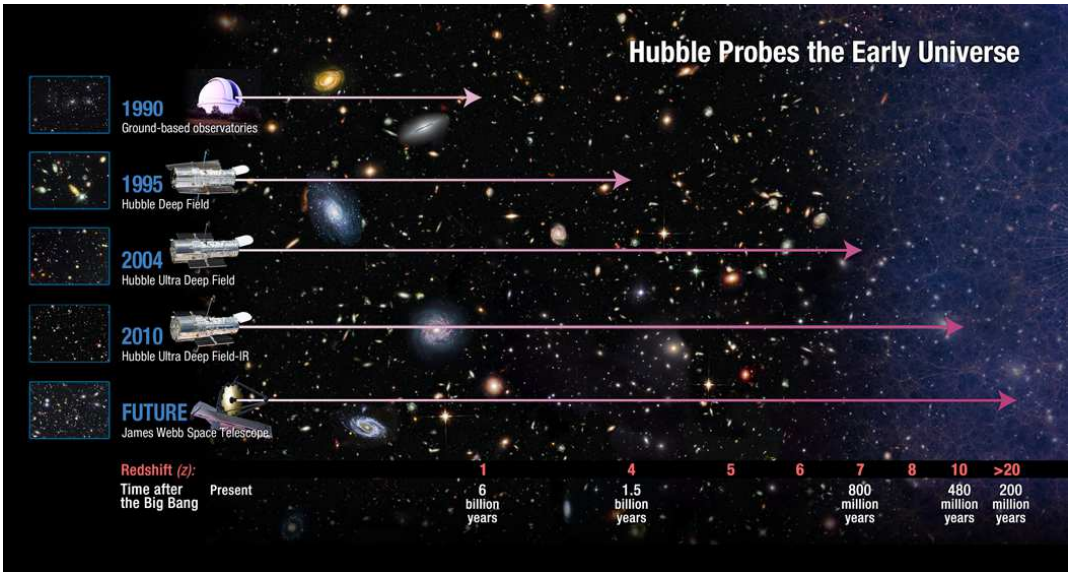


Figure 1.4: This figure shows how Hubble Space Telescope improved our capabilities to look back the history of the Universe compared to ground-based observations. HST permitted us to probe the early Universe with the HDF, HUDF, and HUDF09 surveys. On the bottom is represented the expected contribution from the James Webb Space Telescope which will overcome even the best results achieved by HST.

South (hereafter HDFS, Williams et al. 2000). Both these surveys were obtained with the Wide Field and Planetary Camera (WFPC2) in four optical bands, namely using the F300W, F450W, F606W, and F814W filters.

After the Advanced Camera for Surveys (hereafter ACS) was installed during HST servicing mission 3B in March 2002, the HST field of view was doubled, allowing better survey capabilities. In 2002 the first deep survey carried out with ACS was the treasury program led by Giavalisco aimed at imaging two sky areas as part of the Great Observatories Origins Deep Survey (GOODS). Then, between September 2003 and January 2004 a deeper survey was carried out, the Hubble Ultra Deep Field (hereafter HUDF).

As follow up of the original HUDF, in 2006 the UDF05 project was constructed to search for galaxies at $z > 6.5$ in the HUDF (GO 10632; P.I.: M. Stiavelli). After the installation of the Wide Field Camera 3 (WFC3) during servicing mission 4 in May 2009, the HUDF09 program (GO 11563; P.I.: G. Illingworth) made it possible to identify samples of galaxies at $z \sim$ 7-8 (Oesch et al., 2010c; Bouwens et al., 2010; McLure et al., 2010; Bunker et al., 2010; Yan et al., 2010; Finkelstein et al., 2010) and even up to $z \sim 8.5$. Later on, The HUDF12 program (GO12498, P.I: Ellis) was designed to focus on the contribution of early galaxies to cosmic reionization (Ellis et al., 2013; Koekemoer et al., 2013) on the basis of WFC3 and ACS observations.

CANDELS is another imaging survey carried out with ACS and WFC3 focusing on the high-z Universe. It is the largest project ever, with 902 orbits allocated over three years (2010-2013).

Recently Illingworth et al. (2013) presented the eXtreme Deep Field (XDF), obtained by combining data taken in ten years with ACS and WFC3/IR.

Even though the sky area covered by the XDF (left panel in Figure 1.4.5), and by deep surveys in general, is really small, the combination of such a huge amount of data permits to investigate high-redshift galaxies down to very faint magnitudes ($i_{775} \sim 30$ mag, $z_{850} \sim 29.5$ mag) and to get a background signal characterized by an high S/N.

In the following we will list and briefly describe the HST deep surveys that were used for the projects highlighted in this thesis. We will provide comprehensive description of datasets and issues related to each of these surveys in the following chapters.

1.4.2 The Great Observatories Origins Deep Survey

The Great Observatories Origins Deep Survey (GOODS) combines some of the deepest observations from space and ground-based telescopes covering a total of ~ 320 square arcminutes in two fields centered on the Hubble Deep Field North (HDF-N) and the Chandra Deep Field South (CDF-S). The GOODS HST/ACS program (GO: 9425 and 9583, P.I.: M. Giavalisco) consists of a full, multi-epoch stacked mosaics of the GOODS data obtained with ACS in 4 bands (F435W, F606W, F775W, and F850LP) for both the HDF-N and CDF-S fields (Giavalisco et al., 2004).

1.4.3 The Hubble Ultra Deep Field

The HUDF consists in 10^6 s exposure of an 11 arcmin^2 region in the southern sky obtained using ACS. The exposure time, for a total of 400 orbits, was divided among the following four filters: F435W (B_{435}), F606W (V_{606}), F775W (i_{775}), and F850LP (z_{850}). The limiting magnitude reached is ~ 29 $\mathrm{m_{AB}}$ for point sources. The field contains at least 10000 objects, mostly galaxies (Beckwith et al., 2006). Using the Lyman-break technique, three different redshift ranges were studied, identifying the objects on the basis of the missing flux in the shortest wavelength bands. Objects which dropped out of the B_{435}, V_{606}, or i_{775} filters, lie, approximately, within the redshift ranges $3.5 < z < 4.7$, $4.6 < z < 5.5$, and $5.7 < z < 7$, respectively. Two parallel fields (hereafter NICP12 and NICP34) were acquired with NICMOS while the HUDF was imaged with the ACS. These HUDF-NICMOS parallel fields are characterized by a total exposure time of $\sim 2 \cdot 10^5$ s in both the F110W and F160W filters, providing the deepest near-IR data available at the time.

1.4.4 The UDF05

The UDF05 program is a 204-orbit HST Large Program of ultra deep ACS and NICMOS observations of multiple fields, which was originally constructed to observe with the Wide Field Camera of ACS the HUDF-NICMOS parallel fields, but unfortunately, the NICMOS observations of the main ACS field were carried on only for one-third of the exposure time originally planned. The aim of this survey was to obtain new ultra deep data to study the evolution of the faint-end of the luminosity function within the redshift range $4 \leq z \leq 8$ (Oesch et al., 2009), reaching in the visible bands a depth comparable to the NICMOS data one.

The ACS observations for the NICP12 field spanned a total of 101 HST orbits in the V_{606}, i_{775}, and z_{850} filters (Oesch et al., 2007).

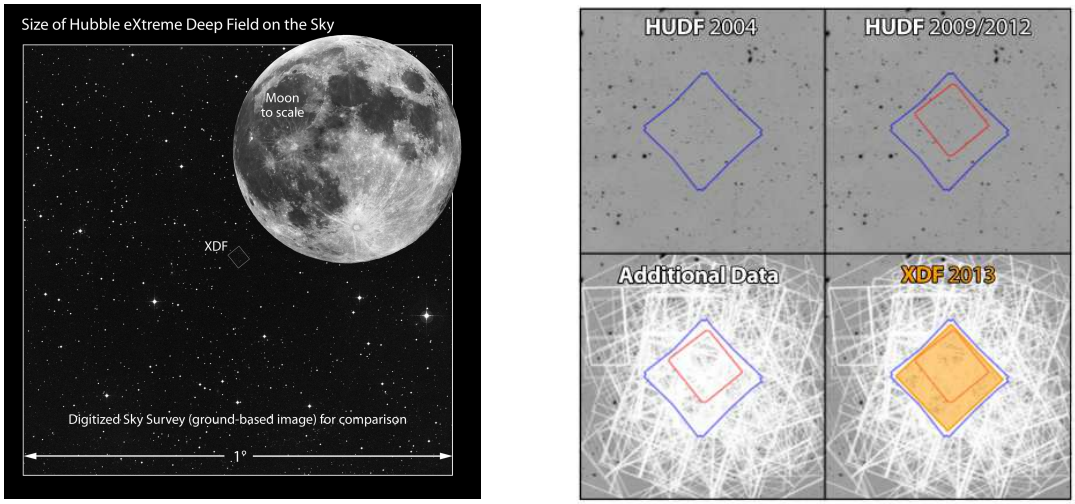


Figure 1.5: Left panel: Comparison between the angular size of the XDF field and the angular size of the full moon, not related to the relative position in the sky. (Illustration Credit: NASA; ESA; and Z. Levay, STScI; Moon Image Credit: T. Rector; I. Dell'Antonio/NOAO/AURA/NSF). Right panel: Picture showing the position and orientations of the data combined to obtain the final XDF images (Illingworth et al., 2013).

1.4.5 The eXtreme Deep Field

Recently Illingworth et al. (2013) released to the archive the entire dataset obtained combining 10 years (from 2002 to 2012) of optical and infrared data obtained with both ACS and WFC3/IR[1]. Images from 19 different HST programs were combined to get the deepest images of the Universe in both the optical (B_{435} , V_{606} , i_{775} , and z_{850} -bands) and near-IR (Y_{105} , J_{125} , and H_{160} -bands) wavelenght range for a 10.8 and 4.7 arcmin^2 area, respectively. In total, the XDF consists of a total of 2963 HST exposures, 1972 in the optical bands and 991 in the near-IR.

1.5 Radiation Background

Going back in time towards the Big Bang the identification of sources like the first stars, AGNs and galaxies starts to be challenging and the study of the background radiation becomes important at all wavelengths. In particular, we are interested in infrared and optical backgrounds, representing the integrated light contribution from populations of faint galaxies, which cannot be individually resolved, over cosmic history. Due to the faintness of the sources themselves, high-quality measurements of the background flux can be hard to obtain and several groups struggled to find a way to study it (Kashlinsky et al., 2005, 2007; Cooray et al., 2007, 2009). Basically, the background radiation is generated by light emitted from luminous objects during the entire history of the Universe, including those epochs currently not accessible to telescopes.

The contribution of young galaxies to the background radiation in the visible and infrared was studied for the first time by Partridge and Peebles (1967). They concluded that,

[1]http://archive.stsci.edu/prepds/xdf/

to be detected, the first galaxies should be brighter than they are now and that any detection of background radiation in the IR would be helpful to constrain an initial bright phase of the Universe. The study of correlation properties of the unresolved optical background was pioneered by Shectman (1974). Using Schmidt plates to detect optical extragalactic background light (EBL) fluctuations on arcminute scales, he found the EBL power spectrum to be consistent with clustering of galaxies fainter than $R = 18$ mag, as predicted by Gunn (1966) and Shectman (1973). Later on, Stecker et al. (1977) found the expected brightness spectrum of the background light in the far-IR to have two peaks, like the spectrum of regular dusty sources, while Martin and Bowyer (1989) derived an evidence of UV background signal from galaxies on the basis of fluctuations detected on arcminute-scales.

In the past, star formation was characterized by more dust, and this is the reason why the contribution of high-z galaxies to the background radiation is expected to decrease from redshifting even though intrinsically the peak in the spectrum is shifted towards redder wavelengths.

In general, the EBL can be used as indicator of the total luminosity of the Universe and could provide us with constraints on the evolution history of faint objects in addition to constraints on the role faint galaxies played in the cosmic reionization scenario. Moreover, since redshifted photons produced by the first stellar objects, protogalaxies, AGNs, and galaxies are recorded in the EBL, information on the evolution of cosmic structures at all epochs is embedded in it.

In the last decade, thanks to space telescopes such as the Diffuse Infrared Background Experiment (DIRBE; Hauser and Dwek 2001) and the Infrared Telescope in Space (IRTS; Matsumoto et al. 2005), it was possible to estimated the absolute IR background.

Kashlinsky et al. (2005), in particular, focused on the cosmic infrared background (CIB) in deep Spitzer data and followed a procedure similar to the one we are presenting in the following chapters. After subtracting detected galaxies (down to $m_{AB} \sim 22-25$ mag) from the data, they obtained the angular power spectrum of the background fluctuations. These detected fluctuations are due to extragalactic undetected sources in addition to the sum of the following contributions: systematic and random effects instrument-related, source artifacts, solar system and Galactic foregrounds (Kashlinsky et al., 2007). The angular power spectrum of the anisotropies they found can be reasonably attributed to the diffuse light produced by Population III stars, since it differs from the one we expect from Solar System and Galactic sources as well as from residual faint galaxies. As we will describe in detail later on, the measure of background signal is not trivial because often the fluctuations are small and hidden by the contribution of other sources and systematic effects.

The contribution from AGN nuclear emission to background radiation was studied and found to be just about 35%, or even less, of the total background in most infrared bands while the bulk of radiation comes from starburst and normal galaxies (Silva et al., 2004). The contribution increases up to 10-20% of the bolometric luminosity of typical submillimeter galaxies when correcting for the significant amount of X-ray absorption present, but still remains far from being dominant (Brandt and Hasinger, 2005).

The papers by Helgason et al. (2012) and Helgason et al. (2014) focused on the near-IR and X-ray backgrounds claiming that the emitters producing the observed background fluctuations on arcminute scales cannot be identified with known galaxies. Their results are consistent with a population of faint galaxies, highly clustered, lying below the current detection limits.

The contribution from high-z star-forming galaxies to the background light is still not

well understood, but, of course, as long as new telescopes and deeper observations come out the number of detected galaxies increases and the unresolved background contribution decreases.

1.6 The Luminosity Function

A quantitative study on the ultraviolet luminosity function (LF) at different redshifts permits to characterize the galaxy population during cosmic time. In particular the LF is related to the star formation rate (SFR) occurring at a certain epoch and it allows us to determine the contribution to the reionization process from high-redshift galaxies.

The Schechter parametrization is usually introduced to represent the LF. It can be written as a function of the magnitude $\Phi(M)$ or of the luminosity $\Phi(L)$:

$$\Phi(M) = 0.4 \cdot ln(10) \cdot \Phi^* \cdot 10^{0.4(M-M^*)^{1+\alpha}} \cdot e^{-10^{0.4(M-M^*)}} \tag{1.1}$$

$$\Phi(L) = \left(\frac{\Phi^*}{L^*}\right)\left(\frac{L}{L^*}\right)^{\alpha} exp\left(\frac{-L}{L^*}\right) \tag{1.2}$$

This parametrization is characterized by an exponential cut-off at high luminosities and a power-law shape at the fainter ones (Figure 1.6).

The Schechter function is characterized by three parameters: L^*, α, and n^*. L^* is the luminosity that separates the low luminosity range from the bright one. At $L < L^*$, i.e. at low luminosities, the function is described by a power-law with slope α, i.e. $\Phi(L) \propto L^{\alpha}$. This means that galaxies with lower luminosities are more common. At $L > L^*$, i.e. at high luminosities, the function is represented by an exponential cutoff in the form $\Phi(L) \propto e^{-L}$, which means that very luminous galaxies are rare. Finally, n^* is a normalization, set at L*.

Working very well from low to high redshifts, the Schechter parametrization has been considered almost universal in characterizing the shape of the galaxy LF for a long time. Despite that, following the recent studies by Castellano et al. (2010) and Capak et al. (2011) at $z \sim 7$ and the work by Yan et al. (2012) on $z \sim 8$ galaxies, at very high redshift the decline of the bright-end of the LF seems to be not as steep as the Schechter parametrization predicts. Capak et al. (2011) suggested that this trend at bright luminosities can be explained considering a less efficient AGN feedback in the early Universe, which is related to the time necessary to supermassive black holes to built up their masses.

According to Stark et al. (2009) and González et al. (2010) the UV LF is correlated with the stellar mass, while Lee et al. (2006), Overzier et al. (2006), and McLure et al. (2009) found a correlation with the mass of the halo. These correlations are useful to constrain the build-up of galaxies in the early phases of the Universe. The comparison between the observed LF and the one derived from theoretical models constrains the physical processes occurring in the galaxies across cosmic time. For example, Bouwens et al. (2007) found that the faint-end slopes of LFs up to $z \sim 6$ are flatter with respect to the dark matter halo function, and this can be explained considering a feedback from supernovae winds (Cole, 1991) or the effect of photoevaporation and heating during cosmic reionization (Barkana and Loeb, 2001). Moreover, the bright-end of the LF shows a sharp decline, different from the one typical of the halo mass function and this difference can be ascribed to AGN feedback or to inefficiency in the gas cooling in high-mass halos (Benson et al., 2003). Even at

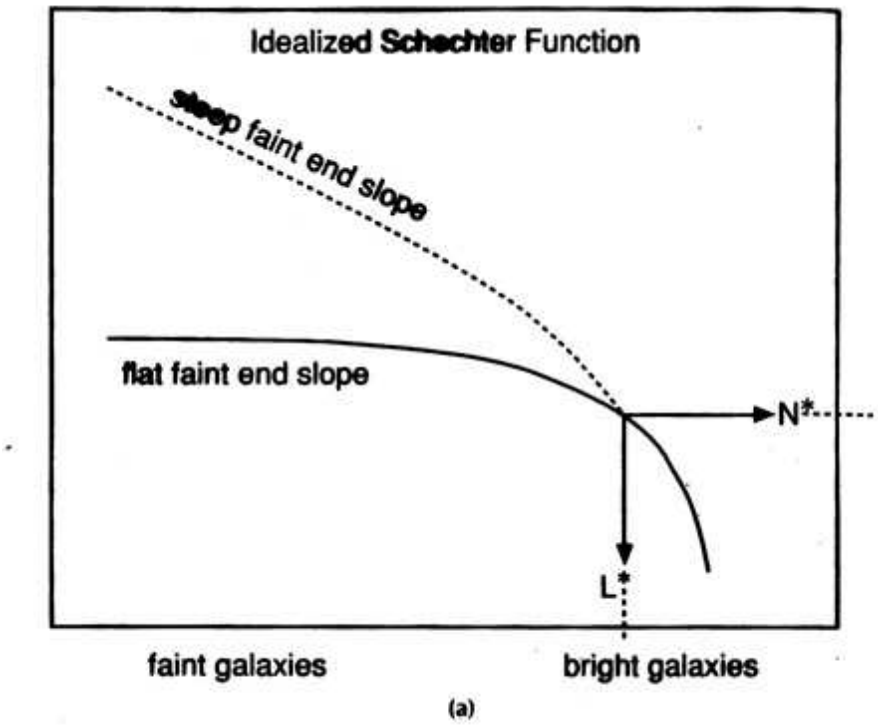


Figure 1.6: Scheme of the Schechter luminosity function. On the left part are shown two possible faint end slopes, a steep one (dotted line) and a flat one (solid line). The knee of the function determines the values of both L^* and n^*.

high redshift, it was found that feedback processes govern the shape of the bright-end of the LF, in particular the exponential decline (Willott et al., 2013).

A comprehensive study on UV LFs at high-z was done by Bouwens et al. (2014) combining the wide-area CANDELS data (Grogin et al., 2011; Koekemoer et al., 2011) with deeper data from the XDF. In particular, they found that, overall, the shape of the LF from $z \sim 4$ to $z \sim 8$ is well in agreement with the Schechter function. Focusing on the LF from $z \sim 4$ up to $z \sim 10$ they derived significant evidences of evolution for the faint-end slope α and the volume density Φ.

1.7 The Power Spectrum

Fourier series provides an alternative way of representing data. In fact, instead of studying the signal amplitude as a function of space or time, it permits to represent the signal by how much information is contained at different spacial or time frequencies.

Fourier's theory is based on the idea that any periodic function, no matter how complex it is along the period, can be represented, without loosing information, as a weighted sum of simple sinusoids. This can be easily applied to images. Irrespective of how irregular may be the analyzed image, it can be decomposed into a set of sinusoidal components

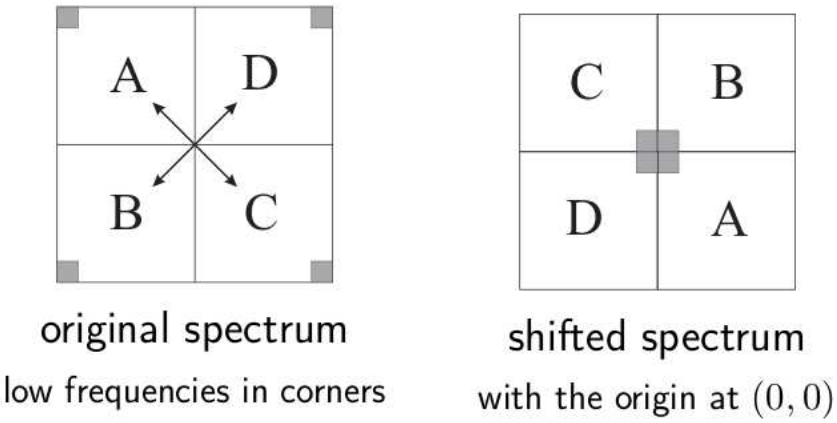


Figure 1.7: The visualization of the power spectrum with the corners representing positions of low frequencies (left panel) is usually hard is usually replaced by the one, obtained swapping diagonally the quadrants, in which the low frequencies are in the middle of the image (right panel).

having each a well-defined frequency. The sine and cosine functions used to decompose are called the basis functions of the decomposition. The following weighted sum of the basis functions is called a Fourier series:

$$f(x) = \sum_{n=0}^{\infty} a_n \cdot cos\left(\frac{2\pi nx}{L}\right) + b_n \cdot sin\left(\frac{2\pi nx}{L}\right) \tag{1.3}$$

In equation 1.3 the weighting factors for each sine and cosine function, i.e. a_n and b_n, respectively, are called Fourier coefficients.

The Fourier series decomposition can be equally done even for 2D images. In this case the basis consists of 2D sine and cosine functions. A Fourier series representation of a 2D function, f(x,y), having a period L in both the x and y directions is:

$$f(x,y) = \sum_{u=0}^{\infty}\sum_{v=0}^{\infty} a_{u,v} \cdot cos\left[\frac{2\pi ux}{L_x} + \frac{2\pi vy}{L_y}\right] + b_{u,v} \cdot sin\left[\frac{2\pi ux}{L_x} + \frac{2\pi vy}{L_y}\right] \tag{1.4}$$

The 2D Fourier transform decomposes the image function f(x, y) into a linear combination of harmonic (sines and cosines, more generally orthogonal) functions.

The Fourier transform is one of the most commonly used techniques in linear signal processing. It provides a one to one transform of the signals from the spatial domain representation to the frequency domain representation. The Fourier transform, F_m, is a complex function.

The best way to compute the Fourier transform of discrete data is the so called Fast Fourier Transform (FFT). The FFT algorithm allows to compute the Fourier coefficients faster.

The power spectrum, or power spectral density, is the square of the norm of the Fourier transform:

$$P_m = |F_m|^2 = (\mathrm{Re}F_m)^2 + (\mathrm{Im}F_m)^2 \tag{1.5}$$

P_m is a real valued and non-negative function, and, for real valued F_m, it is also even. According to the Wiener-Khinchine Theorem, the power spectrum is the Fourier Transform of the autocorrelation function. The power spectrum shows power as the mean squared amplitude at each frequency line, but does not include any phase information.

In the frequency domain, the features characterized by the smallest scale in the image affect the largest scale in the power spectrum. The sources of noise, which affect the large scale range of spatial frequency in the power spectrum, determine the so called white power spectrum, which is basically flat. On the contrary, the large scale features in the image, like, for example, the residuals remaining after the flat field correction, affect mostly the low frequency region of the power spectrum. On the basis of above, it is easy to identify and distinguish these two categories of features.

Usually it is useful to visualize a centered spectrum with the origin of the coordinate system (0,0) in the middle of the spectrum. Referring to the top panel of Figure 1.7, if we assume that the original spectrum is divided into four quadrants, the small gray-filled squares in the corners represent positions of low frequencies. Due to the symmetries of the spectrum the quadrant positions can be swapped diagonally so as the low frequencies are located in the middle of the image (Figure 1.7, bottom panel). In this way subsequently it is possible to derive the power spectrum profile and plot it.

1.8 Cosmological Surface Brightness Dimming

According to Robertson and Walker, that discovered independently the same metric in the mid-1930s, there are only three possible space-time metrics for an homogeneous and isotropic Universe and these metrics can be written as

$$ds^2 = -c^2dt^2 + a^2(t)[dr^2 + S_k^2(r)d\gamma^2],$$

where r is the comoving coordinate and $a(t)$ is a scale factor that is function of time only. By fixing the distances between all points, the metric also defines the geometry of space-time. Since homogeneity and isotropy are stated, there are only three possible geometries of space and the following convention is assumed:

- Positive Curvature ($k = +1$). This geometry is characterized by having the sum of the three angles of a triangle greater than 180° and can be assimilated to the surface of a sphere. Space is, indeed, finite and the model of Universe is called closed.

- Flat space ($k = 0$). This is the Euclidean geometry we are used to. Because it is balanced between the other two models, this is sometimes called a critical Universe.

- Negative Curvature ($k = -1$). It is characterized by having the sum of the three angles of a triangle less than 180°. It can be compared to the surface of a saddle and the correspondent model of Universe is named open.

These 3 possibilities determine the possible forms for S_k:

$$\begin{aligned} S_k(r) &= R_0\, sin(\tfrac{r}{R_0}) && \text{if } k=+1 \quad \text{(positively curved space)} \\ &= r && \text{if } k=0 \quad \text{(flat Universe)} \\ &= R_0\, sinh(\tfrac{r}{R_0}) && \text{if } k=-1 \quad \text{(negatively curved space i.e. saddle)} \end{aligned}$$

Considering the comoving distance r we can define the angular diameter distance of a source as:

$$D_A = \frac{S_k(r)}{1+z}.$$

Towards higher redshift D_A decreases first, and then, at $z \gtrsim 1$ it increases, so as more distant objects appear larger than the closer ones.

Photons that were emitted at a time dt_e are received over an interval $dt_0 = (1+z)dt_e$ by the observers, and they are shifted downward in energy by $(1+z)$. The bolometric flux F is, therefore:

$$F = \frac{L}{4\pi S_k^2(r)(1+z)^2} = \frac{L}{4D_L^2}$$

where L is the source bolometric luminosity and $D_L = (1+z)^2 D_A$ (Hogg, 1999).
The solid angle subtended by a source of projected area A is

$$\Omega = \frac{A}{D_A^2}$$

and the surface brightness is

$$I_0 = \frac{F}{\Omega} = \frac{L D_A^2}{4\pi D_L^2 A} = \frac{L}{4\pi A(1+z)^4} = \frac{I_e}{(1+z)^4}.$$

The last formula shows the surface brightness dimming that affects all cosmological sources (Tolman, 1930, 1934). The surface brightness of every extended source is no longer redshift-independent as in an Euclidean Universe, but the $(1+z)^4$ term makes high-z galaxies difficult to detect when they are faint. A factor $(1+z)^2$ comes from the expansion of space, another factor $(1+z)^2$ from the redshifting of light. In particular, this effect is relevant for sources at $z \geq 4$ that dim by a factor greater than 600. The observational effect of cosmological dimming is a bias in the detection of very high-z galaxies, i.e. we tend to detect only those galaxies with an high surface brightness.

The first studies on cosmological dimming and the consequent observational bias towards more compact objects at higher redshifts were conducted by Phillipps et al. (1990) but only up to redshift about 0.3. Later on Pascarelle et al. (1998) and Lanzetta et al. (2002) stated that a correction is needed in order to sample high and low-z galaxies at a same SB threshold and that at high-z only the bright regions of galaxies are observable since the faint ones cannot be detected against the background. A comprehensive discussion on Tolman dimming can be found in Disney and Lang (2012) who studied the surface brightness selection in the framework of the search for progenitors of low-z galaxies. On the basis of Tolman dimming effect, they claim that, looking back in time, there should be different kind of objects at different epochs, in particular more compact objects at higher redshifts. Moreover, the high-z galaxy population, mostly hidden because of dimming effects, could be the source of the ionizing flux required to reionize the Universe (for a comprehensive study on the role of faint undetected galaxies in cosmic reionization see Calvi et al. 2013).

In the last 20 years the Hubble Space Telescope (HST) has enabled the astronomical community to study the morphology of Lyman Break Galaxies (LBGs) at different redshifts,

from $z \sim 1.5$ up to $z \sim 7-8$, and all those objects were found to be compact (Giavalisco et al., 2004; Lotz et al., 2006; Oesch et al., 2010a; Williams et al., 2014). Bouwens et al. (2004a) found that the principal effect of depth in galaxy surveys is to add galaxies at fainter magnitudes, but there is approximately a $(1+z)^{-1}$ trend on the sizes. In particular, since galaxies at high-z are mostly compact objects (Bouwens et al., 2004a; Oesch et al., 2010a) the effect of cosmological dimming is generally not expected to substantially affect their measured fluxes when comparing results from surveys with different magnitude limits.

1.9 Aim and Scheme of the Thesis

Part of this thesis is already published.

The aim of this thesis is a study on high-redshift galaxies, in particular at $z \geq 4$, using data from different deep surveys carried out with the Hubble Space Telescope. The questions we would like to answer range from the origin of the ionizing radiation that caused cosmic reionization and, in particular, the role played by high-redshift star forming galaxies, to the effect of cosmological dimming on the detectability of galaxies in deep imaging surveys to what kind of galaxies enter the high-z samples as contaminants.

In Chapter 2 we will present a new technique to quantify the light contribution coming from those high-redshift galaxies that are below the detection threshold of imaging data. The technique will be applied to HST/ACS images in the F775W and F850LP filters of the Ultra Deep Field parallel field NICP12. The aim of this analysis is to extend by a few magnitudes the faint-end of the luminosity function at $z \sim 6$ and to constrain the α slope.

Chapter 3 is devoted to the application of the power spectrum technique, introduced in Chapter 2, to the IR dataset obtained with WFC3/IR as part of the HUDF09 program. The goal is to push the studies on the LF at $z \sim 7-8$ beyond the magnitude limit of that deep survey, trying to constrain the faint-end slope of the LF, as done in Chapter 3.

In Chapter 4 we will describe the analysis we performed on the background signal of the XDF images and the consequent data-reduction procedure we performed to get a new version of the XDF in the optical bands, in particular the i_{775} and z_{850} ones. This work is mandatory to get an independent constraint on α at $z \sim 6$ applying the power spectrum technique to the sky area covered by the HUDF main field. In order to get the best data reduction possible we will fully describe all the issues affecting ACS data and those the user faces when dealing with a huge dataset to get a final combined image.

The goal of Chapter 5 is to quantify the effects of cosmological surface brightness dimming in the selection of galaxies at high-redshift. Our empirical strategy relies on the comparison of the total flux detected for the same source in surveys characterized by different depth and makes use of Monte Carlo simulations to derive the expected trend when assuming different surface brightness profiles. To this aim we will use of datasets taken with HST/ACS that are characterized by different depths. By using independent datasets, our results should be more robust against statistical errors or systematics.

In Chapter 6 we present the preliminary results of the study we conducted to constrain the luminosity function of lower-redshift galaxies that enter the samples of high-z objects selected applying the Lyman-break technique.

Finally, in Chapter 7 we will summarize all the work done on the search for high-z galaxies in deep HST surveys and recap our results.

Most of this work can be used to foreseen the capabilities of the upcoming James Webb Space Telescope. All future observations carried with this new telescope will be able to verify our findings, pushing our limits towards the higher redshift Universe.

In addition to the work based on HST deep data, Appendix A and B are devoted to two side projects carried out in parallel. In particular, Appendix A describes a study on the gas-phase metallicity of a sample of galaxies at intermediate redshift. Appendix B presents the analysis of the stellar population properties for a sample of galaxies hosting a hard X-ray emitting Active Galactic Nucleus.

Throughout this thesis we will use the AB magnitude system (Oke and Gunn, 1983).

Chapter 2

Constraining the Luminosity Function of Faint Undetected i-dropout Galaxies

Based on **V. Calvi**, *A. Pizzella, M. Stiavelli, L. Morelli, E. M. Corsini, E. dalla Bontà, L. Bradley, and A. M. Koekemoer, 2013, MNRAS: 432, 3474-3484*

In this Chapter we will present a new technique to quantify the light contribution coming from the faint high-redshift galaxies that lie below the detection threshold of imaging data, set conventionally at $S/N = 4.5$. We illustrate the technique with an application to Hubble Space Telescope Advanced Camera for Surveys images in the F775W and F850LP filters of the Ultra Deep Field parallel field NICP12. The aim of this analysis is to extend by a few magnitudes the faint-end of the luminosity function at $z \sim 6$. After masking all the detected sources in the field we apply the Fast Fourier Transform to obtain the spatial power spectrum of the background signal. The power spectrum permits us to separate the background noise signal, the residuals due to the data reduction of the wide field, and the overall signal produced by faint galaxies. The ratio of the signal in the i_{775} and z_{850} -bands is used to estimate the contribution of the faint i_{775} -dropout galaxies. We rely on extensive Monte Carlo simulations to characterize various sources of uncertainty and quantify the number of faint i_{775} -dropout galaxies in the field. The analysis allows us to put constraints on the luminosity function at $z \sim 6$ down to $z_{850} = 30$ mag, 2.5 mag fainter than with standard techniques on the same data. The data are consistent with a faint-end slope of the luminosity function of $\alpha = -1.9$. Assuming a specific set of values for the clumping factor, escape fraction, and spectral energy distribution, we find that the $z \sim 6$ undetected galaxies down to $z_{850} = 30$ mag could have driven cosmic reionization.

2.1 The Data

Nowadays, the deepest data available are from the Hubble Ultra Deep Field (HUDF) project (Beckwith et al., 2006), the UDF05 follow up of the HUDF (Oesch et al., 2007), and the HUDF09 program (Bouwens et al., 2012a). These ultra deep observations of multiple fields have been obtained with the Advance Camera for Survey (ACS), the Near Infrared Camera

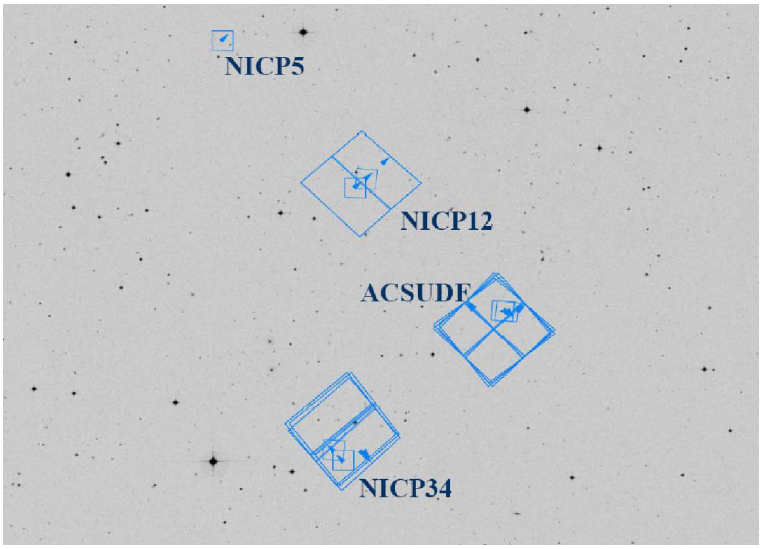


Figure 2.1: UDF05 fields relative to the original HUDF ACS and NICMOS-parallel pointings. The larger and smaller squares correspond to the ACS and NICMOS pointings, respectively (Oesch et al., 2007).

and Multi-Object Spectrometer (NICMOS) and the Wide Field Camera 3 (WFC3).

For the first application of our technique we selected the optical images of the parallel field NICP12 rather than the main UDF field. This field, as well as the NICP34 field, was imaged by ACS using the F606W (hereafter V_{606}), F775W (hereafter i_{775}), and F850LP (hereafter z_{850}) filters during 2005 and 2006 in parallel with deep HST NICMOS NIC3 observations of the original UDF as part of program GO-10632 (P.I. M. Stiavelli). The position of the parallel fields NICP12 and NICP34 observed during the UDF05 program is shown in Figure 2.1, as well as the main field pointing.
The reason for this choice is that the version 2.0 (hereafter v2.0) NICP12 images were produced with a more advanced data reduction than that used in the main field (Oesch et al., 2007). In particular, an herringbone effect introduced inadvertently by the subtraction of compressed bias frame was eliminated as well as electronic ghost images of the bright sources. Thus, the NICP12 images are, nowadays, the cleanest available among the deepest.
The coordinates of the center of the NICP12 field are:
R.A.(J2000.0)=$03^{\rm h}33^{\rm m}03.60^{\rm s}$
Dec.(J2000.0)=$-27^\circ 41' 01.80''$
The characteristics of the observations, i.e. the number of HST orbits, the exposure time in units of seconds, and the zero point in unit of AB magnitudes, for the observations in the V_{606}, i_{775}, and z_{850}-bands are listed in Table 2.1.

2.2 Data Analysis

Bouwens et al. (2007) and Oesch et al. (2007) worked on the UDF and UDF05 data looking for galaxies at $z \sim 6$, i.e. i-dropouts galaxies, constraining the luminosity function at that

Filter	HST orbits	Exposure Time$^{(a)}$ [s]	Zero Point$^{(b)}$ [AB mag]
F606W	9	21,600	26.486
F775W	23	54,000	25.654
F850W	69	168,000	24.862

Table 2.1: Properties of the NICP12 data. Col.(1): filter name. Col.(2): number of HST orbits. Col.(3): exposure time in seconds. Col(4): zero point in AB magnitudes. (a) Data from Oesch et al. (2007), (b) Data from de Marchi et al. (2004).

redshift and the contribution of those galaxies to the reionization process. Their studies are limited to the bright range of magnitudes because of the selection in S/N. On the contrary, our technique allowed us to push the analysis, and consequently the luminosity function, towards fainter magnitudes.

One of the key point of our method, which permits to distinguish it from the previous attempts at high redshift, is our special attention to systematic effects. Thus, we focus on the i-dropout, $z \sim 6$ component that can be studied using images obtained with the same camera and characterized by, essentially, the same systematic effects. The population we seek is detected by an excess in the power spectrum in the z_{850} -band over that of the i_{775} -band, i.e. it is a Lyman break population.

A first step to reveal the contribution of undetected galaxies was to mask all the light coming from stars and detected galaxies, including their bright halos. Once the appropriate mask had been created we applied the power spectrum technique to analyze the different contributions existing in the background.

2.2.1 Creation of the Mask

The UDF05 observations were obtained using a dithering strategy and, then, combined using the `MultiDrizzle` task (Koekemoer et al., 2006). Since when running `MultiDrizzle` on dithered observations of the same field only a few pointings in the outer regions of the field can be combined, and we wanted to work on the deepest data possible without any difference in the depth within the field, we used the weight maps associated to the v2.0 data to select the central region for the following analysis.

The data we focused on were binned to have a scale of 90 mas pixel^{-1} and a size of 3500 $\times$ 3500 pixel^2. We used the weight maps associated to the data to exclude the regions where only few exposures have been combined with `MultiDrizzle` (Koekemoer et al., 2003). This rebinning is intended to limit the computational volume, but it does not affect the final result, as discussed in Section 2.2.4.

First of all, to study the signal coming from faint galaxies it is necessary to remove all the defects and objects that could interfere with our analysis, such as the detected galaxies, their bright halos, residual cosmic rays, and bad pixels. Galaxies were detected independently in the i_{775} and z_{850} -band images using the `SExtractor` photometry package by Bertin and Arnouts (1996), version 2.5.0. `SExtractor` parameters were optimized to maximize the number of detected galaxies while minimizing the number of spurious sources. The detection threshold (DETECT_MINAREA) was set to be a minimum of 5 connected pixels with an intensity of 0.55σ (DETECT_THRESH) above the background (Oesch et al., 2007).

Oesch et al. (2009) considered reliable only those galaxies detected with $S/N > 4.5$. According to Stiavelli (2009) a good estimate for the value of S/N for the detected sources is the ratio of the ISO_FLUX over the ISO_FLUX_ERROR. We, indeed, used this ratio to derive the S/N and, then, we applied the cut from Oesch et al. (2009).

Since we were interested in galaxies which are too faint to be individually studied, but that produce a relevant overall light contribution, we had to mask all the reliable single galaxies and focus on the total contribution of the fainter ones. We noted that, by doing this, we were sensitive to both galaxies just below our detection threshold and much fainter ones. This was intentional, as we were trying to constrain all contributions below those of reliably detected galaxies.

`SExtractor` allows the user to specify the name and type of the output catalog by using the CATALOG_NAME and CATALOG_TYPE parameters in the configuration file **sex*. PARAMETERS_NAME permits to specify the name of the file containing the parameters to be listed in the in the output catalog. Moreover, `SExtractor` can give as output some of the maps used during the intermediate steps. These output fits files are specified in the CHECKIMAGE_TYPE parameter and their names using CHECKIMAGE_NAME. The available options are:

- APERTURES: shows ellipses of different sizes, shapes, and orientations overlapped to the detected objects. It is a good diagnostic on whether or not the threshold used is the correct one.
- SEGMENTATION: permits to check if the objects are broken up too much or not. In the segmentation map all the pixels belonging to a detected object have a value which correspond to the identification number associated to that object in the catalog. For this reason the segmentation map is useful for figuring out what is what from the catalogs.
- BACKGROUND: shows the interpolated background with no sources.
- -BACKGROUND: is the difference between the image and the interpolated background.
- BACKGROUND_RMS: is an interpolated background noise map characterized by full resolution.
- MINIBACKGROUND: is a background map with low resolution.
- MINIBACK_RMS: is a background noise map with low resolution.
- OBJECTS: shows all the objects detected above the threshold.
- -OBJECTS: is the background subtracted image with all the detected sources blanked.
- FILTERED: is the background subtracted filtered image.

Starting from the segmentation maps we created two masks (one for each band) to reject all the detected sources existing in the field of view. All the rejected pixels were set to 0.0, as done by Kashlinsky et al. (2005) and Arendt et al. (2010). Since most of the objects are surrounded by a bright halo, we decided to enlarge the masked area to avoid contaminating signal from bright galaxies. We reversed each mask to be 1 where there are sources and 0 elsewhere, we convolved two times the mask with a Gaussian filter with

FWHM $= 30$ pixel =2".7 enlarging it each time by adding all the pixels with a value > 0.5. Finally we reversed again the mask. We obtained the final mask by merging the masks in the two passbands and convolving again. The final masked area corresponds to 16.8% of the field.

2.2.2 The Power Spectrum Technique

A description of random fluctuations is fundamental to extract information hidden in the background of images and to provide a quantitative measure for the comparison with simulations.

The spatial power spectrum is the key of our method, allowing us to determine the amplitude of surface brightness fluctuations and to distinguish the sources of noise from the fluctuations due to the unresolved galaxies.

It should be noticed that the power spectrum we are focusing on is not the primordial one considered first by Harrison (1970) and Zeldovich (1972), and then by White (1994) to study, e.g., the cosmic microwave background anisotropies. We are, instead, interested in the spatial distribution of the faint $z \sim 6$ galaxies, so as to be able to extend the faint-end of the luminosity function. This power spectrum is related to the primordial one, but departs from it due to non-linear evolutions, gas physics , and conversion of gas to stars. For this reason we did not rely on theoretical expectation, but we used the Fourier Transform to derive the spatial power spectrum of the luminosity fluctuation due to these galaxies. For point sources only their spatial distribution contributes to the power spectrum, but, for extended objects, this depends on both their size and their distribution. On the basis of above we are able to derive the typical angular scale of the faint galaxies (see, however, comments at the end of Section 2.2.3).

A technique similar to ours was used for the first time by Tonry and Schneider (1988). They discovered that the distance of a galaxy is inversely proportional to the amplitude of the luminosity fluctuations due to unresolved red giant stars and they used the spatial power spectrum to directly measure these fluctuations. After computing the Fourier transform and the power spectrum of the data, they fitted the power spectra with the sum of a constant value and the power spectrum of the point spread function (PSF):

$$P_{image} = P_0 + P_1 \cdot P_{PSF} \tag{2.1}$$

The resulting fit is shown in Figure 2.2 using the data obtained for M32.
The main difference between their work and ours is the density of sources per pixel. They were studying nearby galaxies whose regions have an average projected density of stars of one hundred per pixel so the fluctuations between adjacent pixels have rms variations equal to 10% of the mean signal. On the contrary, we are dealing with images where the projected density of faint galaxies is much less then one per pixel, therefore we have smaller fluctuations and our results are more sensitive to the presence of spurious signals.

Our approach consists in applying the `IDL`[1] Fast Fourier Transform (FFT) routine to the masked images to compute the two-dimensional Fourier transform and, then, to derive the spatial power spectrum of the signal. Therefore, the power spectrum image obtained is two-dimensional, but to plot it and to do all the calculations in the following it is convenient to derive the radial trend.

[1] Interactive Data Language is distributed by ITT Visual Information Solutions.

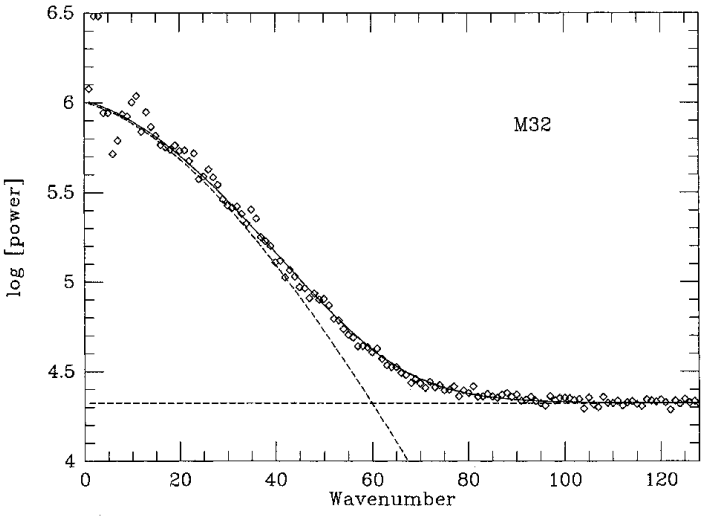


Figure 2.2: Power spectrum obtained from the data of M32 (Tonry and Schneider, 1988). The observations, plotted as diamonds, are well reproduced by the sum of a constant value and the power spectrum of the PSF.

The power spectra for the two bands can be compared by calculating their difference or ratio. The difference is the most natural choice, but relies on an extremely accurate color calibration of the power spectra in order to properly subtract the contribution of lower redshift objects from the z_{850} -band power spectrum. The ratio of the power spectra could introduce a loss of sensitivity of the method, because it does not allow us to detect faint galaxies with the same power spectrum shape as the bright ones. On the other hand, the ratio allows us to avoid the need of a very accurate and challenging color calibration of the power spectra. Thus, from here on, we focus our analysis on the ratio of power spectra.

First of all, we obtained the power spectra for both the i_{775} and z_{850} -band images (Figure 2.3).

The units of measurement for each single power spectrum can be derived as follow. First of all we should consider the definition of magnitude and AB magnitude:

$$mag = -2.5 \cdot log_{10}(signal[DN/s]) + Z_{pt} \tag{2.2}$$

$$AB_{mag} = -2.5 \cdot log_{10}(flux[erg/s/cm^2/Hz]) - 48.6 \tag{2.3}$$

Equalizing equation 2.2 and 2.3 we obtain

$$-2.5 \cdot log_{10}(signal[DN/s]) + Z_{pt} = -2.5 \cdot log_{10}(flux[erg/s/cm^2/Hz]) - 48.6 \tag{2.4}$$

$$log_{10}\left(\frac{flux}{signal}\right) = -29.7016 \tag{2.5}$$

Since 1erg/s$= 10^2$ nW and 1cm$^2 = 0.0001$m^2, we can derive the flux in units of nW/m^2/Hz

$$flux = signal \cdot 10^{-29.7016} \cdot 10^2 \cdot 10^4 \tag{2.6}$$

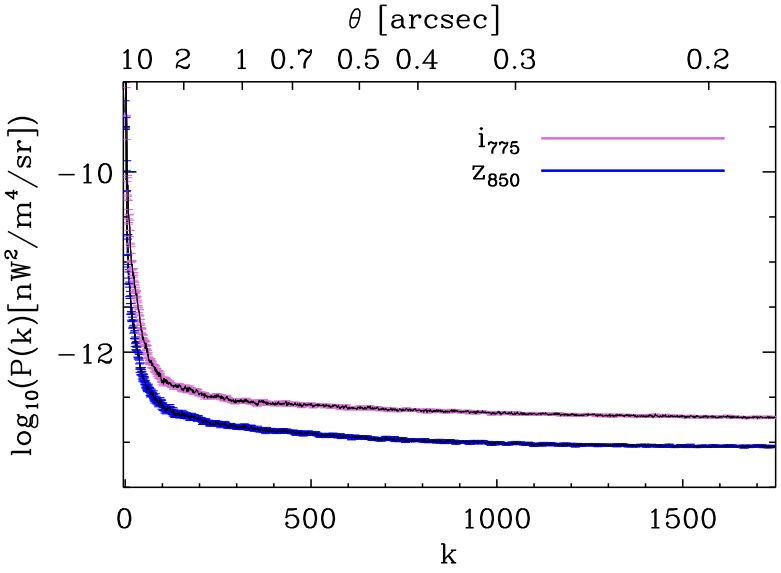


Figure 2.3: Power spectra with error bars of the background signal of the i_{775} (purple top line) and z_{850} images (blue bottom line).

Considering the band width of the filter, in this case $\Delta\nu = 7.11 \cdot 10^{13}$ Hz for the i_{775} one, we obtain the flux in units of nW/m^2

$$flux = signal \cdot 7.11 \cdot 10^{-29.7016} \cdot 10^{6} + 10^{13} = signal \cdot 7.11 \cdot 10^{-10.7016} \tag{2.7}$$

We should also consider the pixel scale, which permits to convert the flux from nW/m^2 to nW/m^2/sr

$$\left(\frac{0.03}{206264.8}\right)^2 = 2.11 \cdot 10^{-14} \tag{2.8}$$

$$\left(\frac{0.09}{206264.8}\right)^2 = 1.9 \cdot 10^{-13} \tag{2.9}$$

Finally, to derive the calibrated image in physical units, i.e. nW/m^2/sr, it is simply necessary to multiply the non calibrated one by a conversion factor

$$flux = signal \cdot CF \tag{2.10}$$

The final conversion factor CF depends on the pixel scale and the zero point. The value for CF is 6698.65 for the i_{775} -band image with a pixel scale of 30 mas/pix, 743.9 for the i_{775} -band image with a pixel scale of 90 mas/pix, 9578.86 for the z_{850} -band image with a pixel scale of 30 mas/pix, and 1063.76 for the z_{850} -band image with a pixel scale of 90 mas/pix.

When `IDL` is computing the Fast Fourier Transform, a normalizing factor $1/\sqrt{N}$ is introduced, being N the total number of pixels in the image. To derive the real FFT from the `IDL` output it is, indeed, necessary to multiply the output by $\sqrt{N}$.

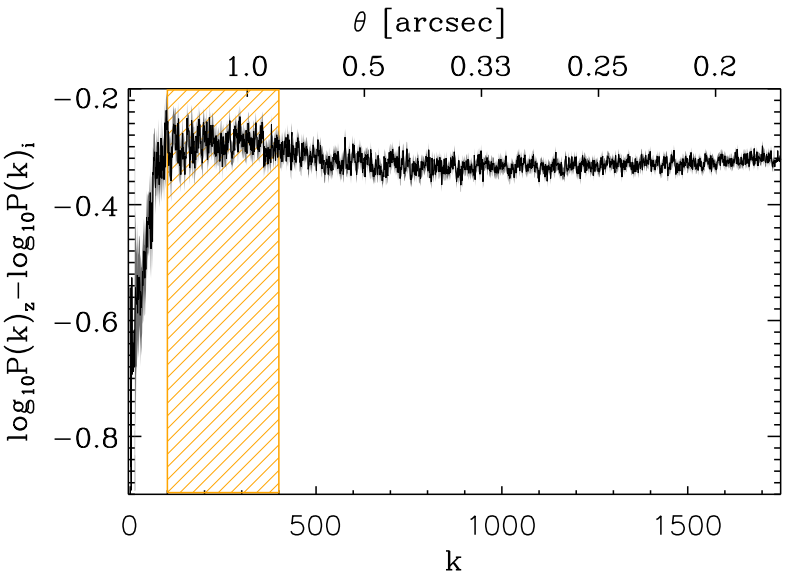


Figure 2.4: Ratio between the z_{850} and i_{775} power spectra. The gray shaded area represents the errors associated to the ratio. The region between $\mathrm{k} = 100$ and $\mathrm{k} = 400$, where the signal from the faint galaxies was measured, respectively, is highlighted in orange.

Considering the Fourier Transform of the calibrated image and the power spectrum derived from it, calling them FFT and PS, respectively

$$FFT \cdot \sqrt{N} \cdot \frac{\sqrt{scale}}{206264.8} \tag{2.11}$$

$$PS \cdot N \cdot \frac{scale}{206264.8^2} \tag{2.12}$$

Finally we have the power spectrum in units of $\mathrm{nW}^2/\mathrm{m}^4/\mathrm{sr}$, which is consistent with the finding of Kashlinsky et al. (2004, 2005, 2007, 2012).

After obtaining the power spectrum for each band, to highlight the light contribution coming from the undetected galaxies, which are bona fide $z \sim 6$ candidates, we plotted the ratio between the z_{850} and i_{775} -band power spectra (Figure 2.4). The undetected galaxies are responsible for the peak visible in Figure 2.4. We tested the reliability of the peak using a χ^2 statistics. Comparing the values of the ratio between the power spectra in the range of the peak, i.e. $100 \lesssim k \lesssim 400$, and in the range $1400 \lesssim k \lesssim 1700$ dominated only by the noise, we obtained $\chi^2 = 3.08$. This value made us confident that the peak is a real feature and it is can not be ascribed to random noise since its confidence level is basically equal to zero.

The angular scale θ, plotted on the top x-axis of the figures, is derived from k as follows:

$$\theta\ [\mathrm{arcsec}] = \frac{\mathrm{D}_{\mathrm{frame}}\ [\mathrm{pixel}] \cdot \mathrm{scale}\ [\mathrm{arcsec\ pixel}^{-1}]}{k} \tag{2.13}$$

where $\mathrm{D_{frame}}$ is the dimension of the frame in pixels and scale is the pixel scale of the image.

Following our technique, we can derive the number of faint galaxies from the height of the signal excess they produce in the ratio between the power spectra (Figure 2.4), comparing it with the one obtained from our detailed simulations (see Section 2.3).

It must be noticed that our result can be affected by a fraction of interlopers at lower redshift (for a general overview on the topic see Stanway et al. 2008 and Chapter 6). This is an issue affecting all studies based on the dropout selection and it is due to the aliasing between the Lyman break and the 4000 Å break, as discussed by Dahlen et al. (2010). Su et al. (2011), on the basis of accurate simulations, estimated that the fraction of $z \sim 1.2$ galaxies that can be selected as $z \sim 6$ population could be as high as 24%. This value is a pessimistic one since Malhotra et al. (2005) found half that number of contaminant galaxies. More recently Pirzkal et al. (2013) estimated the fraction of low redshift interlopers to be on average 21%.
A preliminary study on interlopers affecting high-z galaxy samples at $z \sim 4-5-6$ is described in Chapter 6.

2.2.3 Different Contributions

The use of the power spectrum technique requires a preparatory study to understand how the different components of a scientific image affect the final power spectrum.

The ratio between the spatial power spectra in Figure 2.4 can be divided into three different parts, the first one at low wave numbers ($\mathrm{k} \lesssim 100$), the second one in the range of high wave numbers ($\mathrm{k} \gtrsim 400$), and the remaining one at intermediate wave numbers, which corresponds to the contribution from the undetected galaxies.

The power spectrum of random noise is termed white power spectrum and has approximately a constant value over all the range of wave numbers. At high wave numbers, meaning at the pixel scale, the white noise contribution is dominant. At low wave numbers the principal contribution is due to residuals of the flat field correction process (for more details on this see Section 2.3.1).

The highlighted region of the plot in Figure 2.4 is the most interesting one, where the signal is dominated by the contribution of the galaxies, therefore it permits to infer the existence of galaxies below the usual detection limit. In particular, the amplitude of the peak visible between $\mathrm{k} = 100$ and $\mathrm{k} = 400$ is proportional to the number of undetected galaxies in the field.

The resulting sizes are, indeed, somewhat large for galaxies at $z \sim 6$. Even though it is not possible to deeply investigate this issue, one possible interpretation is that these sizes can be ascribed to the effect of multiple halo occupancy with individually more compact sources. The sizes we inferred would imply halo masses of $10^9 M_\odot$, or even larger. We considered this different from ordinary correlation contribution since multiple halo occupancy is usually not considered in correlation calculations and could have much smaller scales. The regular correlation length for galaxies at $z \sim 6$ is not known but, in order to correspond to the peak we observed (equivalent to 100 kpc comoving or less), it would have to decrease by more than an order of magnitude compared to the measured values at low redshift.

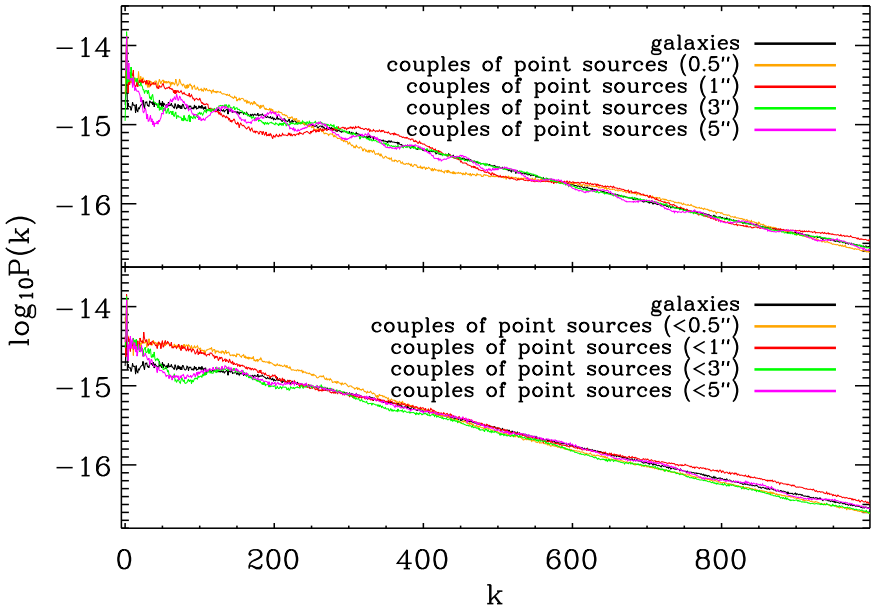


Figure 2.5: The power spectrum obtained from a frame with 120 mock galaxies (no noise, mask, nor flat field residuals) is plotted in black and compared to those derived for 60 pairs of point sources with the same total luminosity of the extended galaxies. As reported in the legend, the color code refers to different distances, in arcseconds, between the members of the pairs.

Hints on the Effect of Grouped Sources

It is known that galaxies at high redshift are not necessarily randomly distributed, but can be grouped. Even though the clustering properties of high-z objects are not well known, nor even investigated, we built up a toy model and decided to perform a series of simulations to constrain the possible effect of grouped galaxies on the power spectrum. In particular, our goal was to figure out if, grouping the sources, the contribution from galaxies characterized by smaller sizes, up to the point-like ones, can resemble the one from single extended galaxies randomly distributed, i.e. with no correlation between the relative positions on the frame. To maximize the effect of groups of galaxies and disentangle it from the one related to the intrinsic size of extended galaxies, we considered point sources, with the same total luminosity of extended mock galaxies, and grouped them in pairs or small groups with three or four elements. Moreover, we did not introduce either the background noise, nor the flat field residuals, to not bias the results and study only the pure effect of groups of galaxies.

The power spectrum in the case of grouped sources is characterized by a sinusoidal trend, with the frequency depending on the separation between the components of the pair or group. We found that for really close by point sources, i.e. those with distances lower than $1''$, the frequency is lower, otherwise the frequency is higher (see top panel of Figure 2.5).
Since it is not realistic to have all the members of the groups at the same distance from

each other, we introduced different distances. In this way the sinusoidal trend is soften, especially at high wave numbers (see bottom panel of Figure 2.5).
Then, we analyzed the effect of changing the maximum value for the relative distance between the objects in a pair or small group. We found that, with a separation lower than 1 arcsec, grouped point sources can reproduce the trend of extended galaxies with no grouping.

Later on, for a more precise analysis, we considered the state of the art regarding clustering of high-redshift dropout galaxies in literature. Overzier et al. (2006) studied the clustering properties of galaxies at $z \sim 6$ finding an angular correlation for galaxies brighter than $z_{850} = 28.5$ mag on scales larger than $10''$ and no correlation on the same scales going down to $z_{850} = 29.0$ mag due to low S/N. Studying a sample of galaxies at $z \sim 5$, Lee et al. (2006) found a further correlation on smaller angular scales, which was interpreted as due to the so-called one-halo term. The clustering properties, due to both the one-halo and the two-halo term, are still not well explored for i_{775} -dropouts with magnitudes in the range between $z_{850} = 28$ mag and $z_{850} = 30$ mag. Moreover, our field is too small to be affected by the large angular scale clustering which is better constrained (Overzier et al., 2006) than the small scale one. On the basis of this, we explored the effect of galaxy grouping on the power spectrum building up a toy model which introduces pairs of sources instead of randomly distributing them according to a non-zero angular correlation function only on scales smaller than $10''$.
We assumed the one-halo term in the form $\omega(\theta) = A_\omega \theta^{-\beta}$ found by Lee et al. (2006) for V_{606} -dropout galaxies to be valid also for i_{775} -dropouts and placed the mock galaxies in the simulated frame accordingly. We found that clustering could amplify the signal from faint galaxies. In particular, simulations with no clustering lead to over-estimating the number of faint galaxies (see Section 2.3 for more details). In any case, since the clustering/grouping properties of i_{775} -dropouts on the smaller angular scales are not known at the present epoch, our findings need to be revisited as soon as a proper angular correlation function is found at $z \sim 6$.

2.2.4 Possible Sources of Contamination

There are many sources of uncertainties that can affect our results and we will shortly list and characterize them in the following:

- *Mask:* The primary source of uncertainty in our results is the size of the mask. Our aim is to mask all the galaxies which can be reliably and individually identified and to study the overall signal coming from those which are too faint to be detected one by one. If we are masking to a signal level which is too low, we will not see any signal coming from faint galaxies. On the other hand, if the mask brightness limit is too high, the faint $z \sim 6$ undetected galaxies will be overwhelmed by the light of the foreground galaxies and their bright halos. The ratio between the power spectra obtained masking all the sources detected by `SExtractor` with $S/N > 4.5$ and $S/N > 3$ are compared in Figure 2.6. In the region between $\mathrm{k} = 100$, which corresponds to objects with a diameter $\mathrm{D_{galaxies}} = 3.''15$, and $\mathrm{k} = 400$, which refers to objects with $\mathrm{D_{galaxies}} = 0.''79$, we notice that the peak due to the signal of faint i-dropouts is dropped and almost invisible. As the percentages of masked area, i.e. 16.8% and 20% for $S/N > 4.5$ and $S/N > 3$, respectively, are quite similar, we can in-

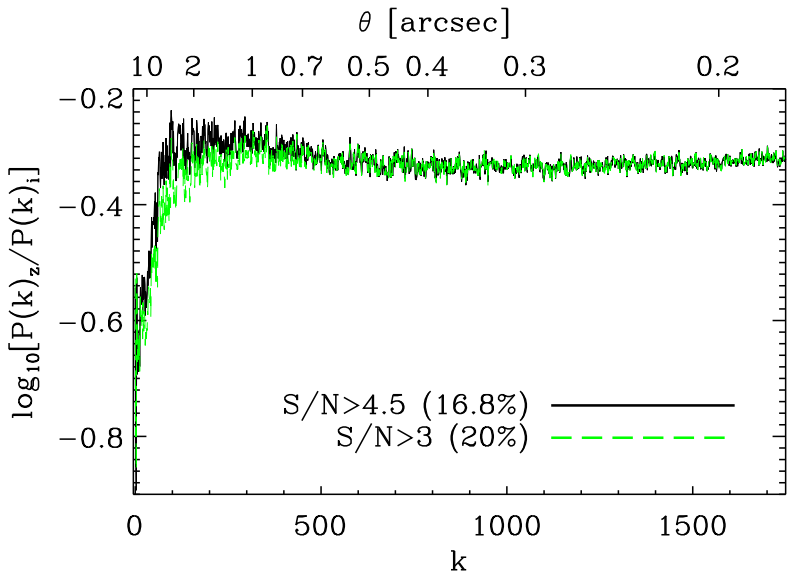


Figure 2.6: Ratios between the z_{850} and i_{775} -band power spectra obtained masking all the sources detected by `SExtractor` with $S/N > 4.5$ (black line) and $S/N > 3$ (green dashed line). The percentages within brackets indicate the masked area.

fer that we are detecting a signal from galaxies barely below our detection threshold rather than from a multitude of very faint objects.

- *Calibration files and dither pattern.* The use of `MultiDrizzle` could introduce spurious signals in the final scientific frame obtained after the combination of all the **_flt.fits* files. To test this, we downloaded all the **_flt.fits* data available for the NICP12 field in the i_{775} and z_{850} bands from the *Hubble Space Telescope* (HST) archive and replaced the scientific frame with random noise. The resulting mock frames have no sources and maintain the same coordinates of the real ones. Then, for each filter we processed these noise-only frames with `MultiDrizzle`, and we applied the power spectrum technique to the final scientific images. As can be noticed from Figure 2.7 no features are introduced by `MultiDrizzle` neither in the single power spectra (top panel) nor in the ratio (bottom panel). Moreover, we performed a similar test combining mock frames with noise multiplied by a random residual of flat field (with a maximum of 0.3 and 1 per cent) and, then, applying the power spectrum technique. It should be noted that the flat field residual in this case introduces a pixel-to-pixel uncertainty, not a large-scale one. Once again, we concluded that `MultiDrizzle` is not creating spurious signals (see Fig. 2.8) and, consequently, it is not affecting the final result of our technique.

- *Cosmic Ray Residuals:* The presence of cosmic ray residuals in the frame determines a gentle increasing slope in the range of high wave numbers, which depends on the intensity and number of pixels affected by the cosmic rays.

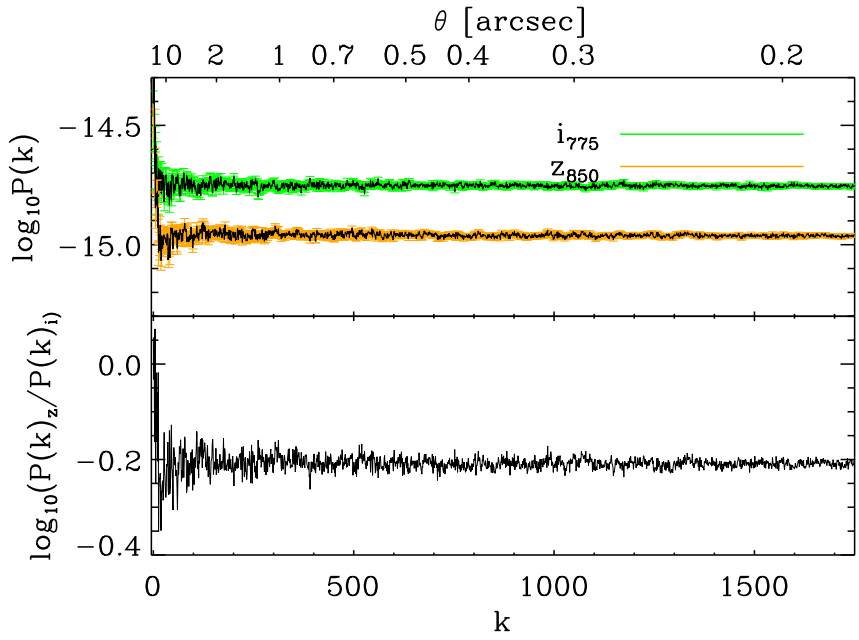


Figure 2.7: Power spectra derived by combining the noise-only frames with `MultiDrizzle` (top panel) and their ratio (bottom panel). The flatness of the profile indicates that no features were introduced when combining the images with `MultiDrizzle`.

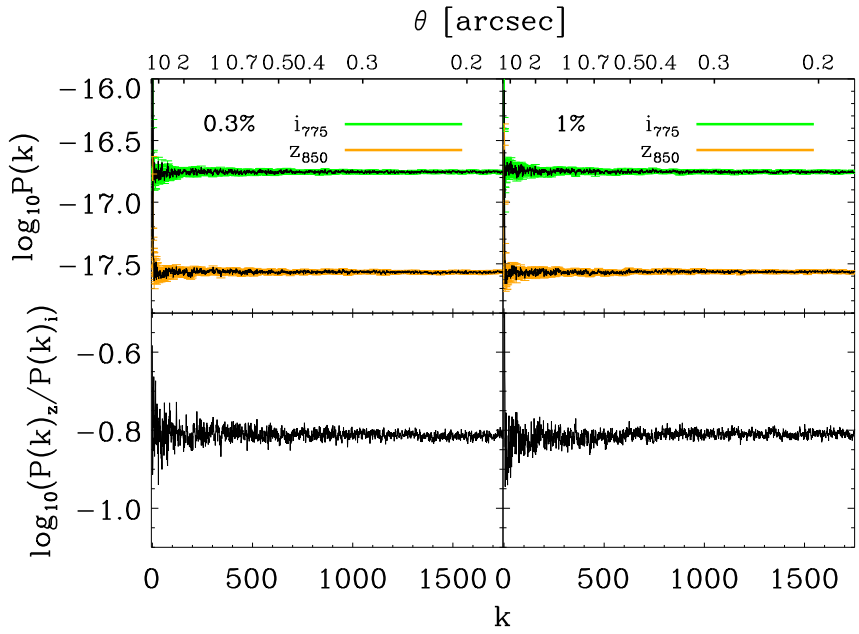


Figure 2.8: Power spectra derived by combining with `MultiDrizzle` noise-only frames after being multiplied by a random residual of flat field with a maximum of 0.3% (top-left panel) and 1% (top-right panel) and their ratio (bottom panel). As in Figure 2.7, no spurious feature is introduced in the final ratio when combining the images with `MultiDrizzle`.

- *Zodiacal Light:* The zodiacal light is the dominant source of sky brightness in the near-IR (Reach, 1997). We were only interested in the fluctuations due to the zodiacal light cloud changing in the sky during the time. According to Kashlinsky et al. (2005), considering two observations of the same field taken six months apart and computing their difference, it is possible to isolate the zodiacal light fluctuations from the contribution coming from galactic and extragalactic objects. With this method Kashlinsky et al. (2005) were able to derive an upper limit for the zodiacal light contribution and to affirm that it did not affect their results on the cosmic infrared background (CIB) fluctuations.
  Unfortunately, a few issues prevented us from applying the same analysis to our HST data. First of all, observations taken six months apart imply different orientations of the camera which could introduce detector-related systematics, leaving a signature in the power spectrum. Moreover, ACS is affected by the contribution of scattered sunlight whenever the angle is close to (or less than) 90° and this could determine a wrong estimate of the fluctuation ascribable to zodiacal light. Finally, at the moment we do not have observations of our fields taken with such a long temporal delay, so we cannot study relevant zodiacal fluctuations. On the basis of above, we decided to assume a negligible zodiacal contribution, following Kashlinsky et al. (2005), and to better handle the zodiacal light fluctuation as soon as all the data will be available, analyzing observations of the same field obtained with a long temporal separation, so as to highlight relevant differences in the zodiacal light.

- *Flat Field Correction and Bias Subtraction Residuals:* If the flat field correction or bias subtraction process were not correctly done, the data could be affected by a diffuse signal not coming from any real source. It can easily overcome the small amount of photons produced by the undetected galaxies we are looking for. These residuals modify the trend of each single power spectrum, increasing its slope in the low and medium wave numbers ranges and making difficult to see any peak in the power spectra ratio.

- *Cleaning Level:* The cleanliness of the data is a crucial issue to take into account. In fact the worse is the the data reduction, the greater is the amount of residuals which could give rise to a spurious signal and alter the results. To test this we applied the same steps described in Section 2.2.2 to the first public release of the UDF05 data (v1.0) of our field and we compared the power spectra in each band (Figure 2.9). The v1.0 data are affected by spurious signals (Oesch et al., 2007) which are revealed by a series of wiggles extending all the way to the white noise region. This is even more evident if we analyze the ratio between the power spectra for the v1.0 and v2.0 data plotted in Figure 2.10. In the v1.0 the light coming from the faint galaxies is completely overcome by the presence of a herringbone artifact and electronic ghosts.

- *Rotation:* The images we used, oriented with North up and East left, were obtained from `MultiDrizzle` without introducing any additional correlation between adjacent pixels because of the use of the point kernel (Beckwith et al., 2006). We did not introduce any further rotation in our data, nevertheless, we decided to investigate this effect for sake of completeness since this were doing a technical analysis aimed at describing all the aspects of our new technique. If we were dealing with rotated frames it would be important to consider the following effect. The rotation of the images be-

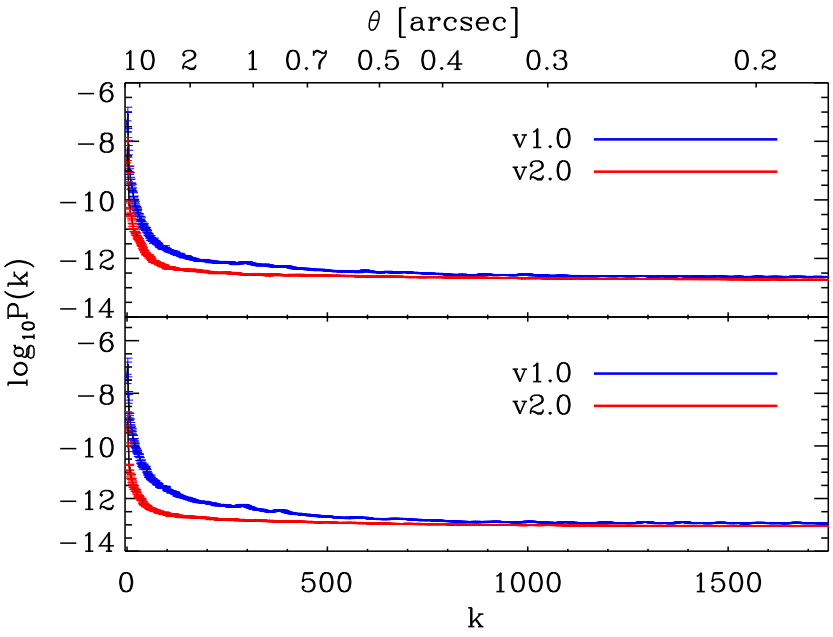


Figure 2.9: Power spectra obtained from the v1.0 (top line with blue error bars) and v2.0 (bottom line with red error bars) versions of the data for the i_{775} (top panel) and z_{850} -band images (lower panel).

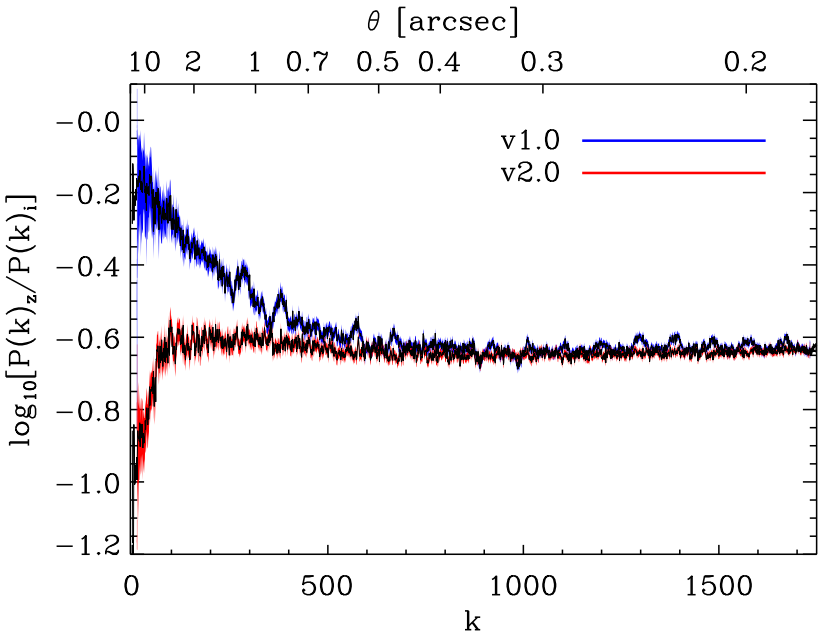


Figure 2.10: Comparison of the ratios between the z_{850} and i_{775} -band power spectra for the v1.0 (top line with blue error bars) and the v2.0 (bottom line with red error bars) data.

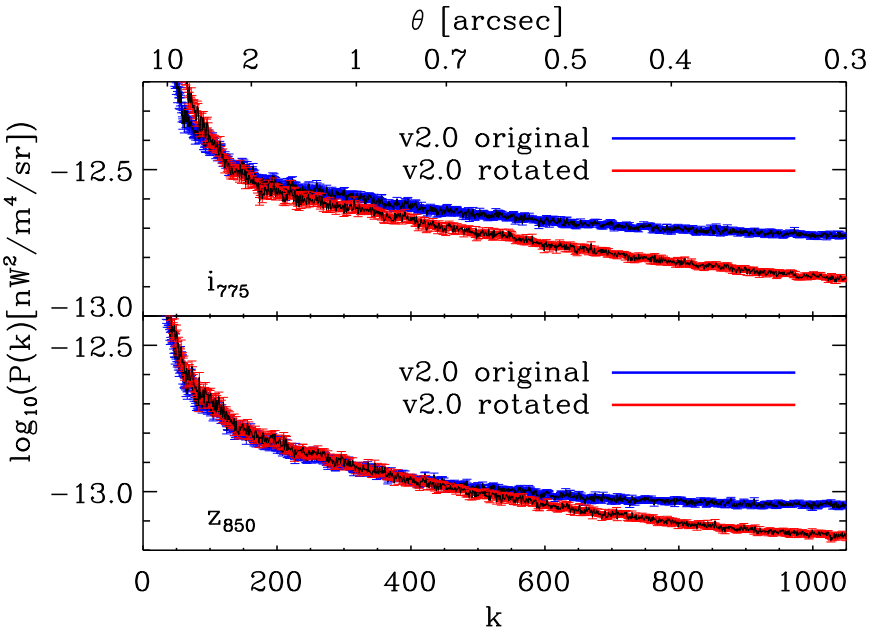


Figure 2.11: Power spectra obtained from the i_{775} (top panel) and z_{850}-band (bottom panel) images with (red bottom line) and without (blue top line) applying a 45° clockwise rotation. In each panel the power spectrum referring to the rotation was vertically shifted to match the other one at $\mathrm{k} = 125$ ($\mathrm{D}_{\mathrm{galaxies}} = 2.''52$) to better show the effect of the rotation.

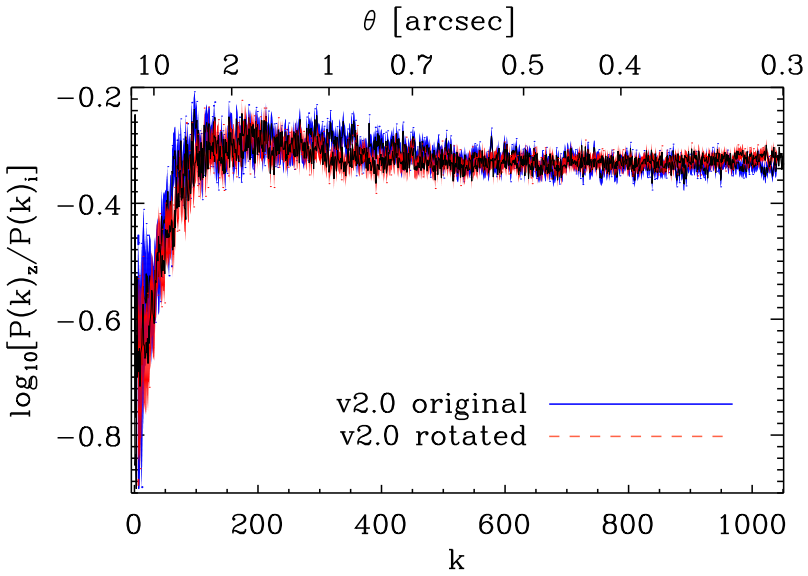


Figure 2.12: Comparison of the ratios between the z_{850} and i_{775}-band power spectra for the v2.0 data with (red dashed line) and without (blue solid line) applying a 45° clockwise rotation.

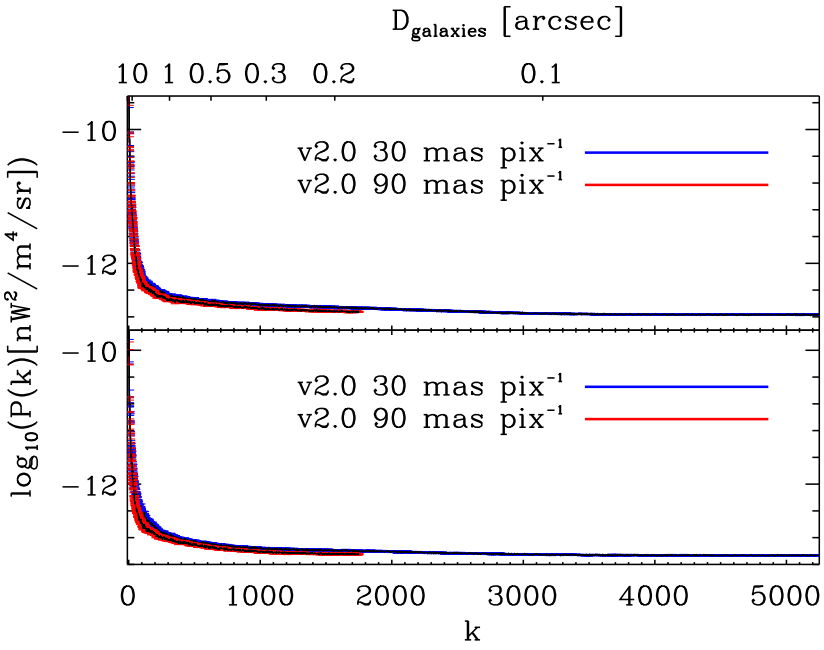


Figure 2.13: Power spectra obtained from the v2.0 i_{775} (top panel) and the z_{850} images (lower panel) with a pixel scale of 30 (blue line) and 90 mas pixel^{-1} (red line).

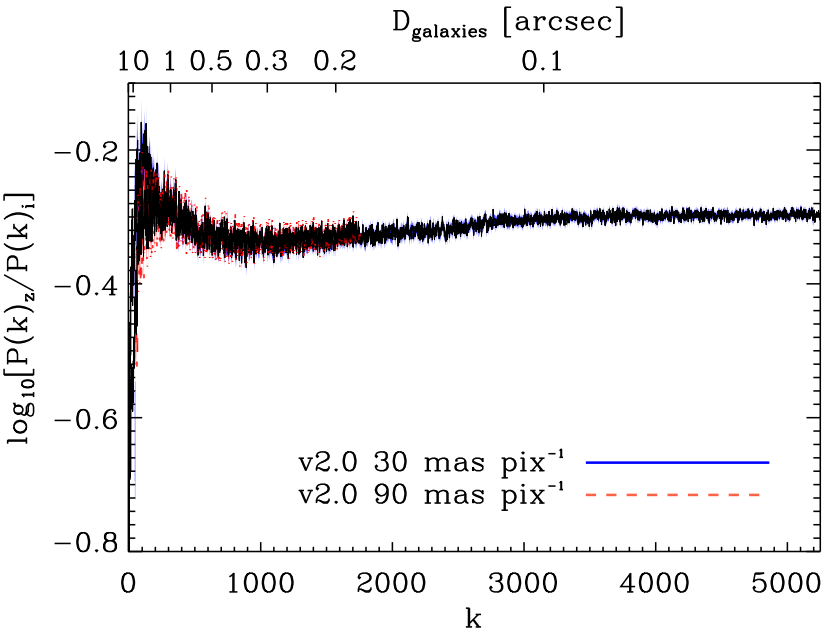


Figure 2.14: Ratios between the z_{850} and i_{775} power spectra obtained for the v2.0 images with a pixel scale of 30 (blue solid line) and 90 mas pixel^{-1} (red dashed line).

fore extracting the power spectrum determines a correlation between adjacent pixels, which is detected as a clear slope at high wave numbers of both the i_{775} and z_{850} power spectra (Figure 2.11). The magnified panel shows that the power spectra of the original and rotated data present different slopes, but, as shown in Figure 2.12, the ratio of the power spectra is not affected by any effect concerning the rotation since the slope introduced is exactly the same in both the i_{775} and z_{850} -band power spectra. The ratio of the rotated data fits perfectly with that of the original ones.

- *Binning:* To test the effect of data binning we compared the results obtained from the images with a pixel scale of 90 mas pixel^{-1} and a dimension of 3500 pixel each side, and the original images, characterized by a pixel scale of 30 mas pixel^{-1} and a 10500 pixel side (Figure 2.13). By binning the data, we lost details and so the power spectrum covers a more limited range of wave numbers. However, the ratio between the power spectra plotted in Figure 2.14 shows that the white power spectrum is well reproduced by both the 30 and the 90 mas pixel^{-1} cases.

2.3 Monte Carlo Simulations

The power spectrum permits us to see the existence of an overall signal coming from faint galaxies, but we are not able to directly obtain neither the number of those galaxies nor their magnitudes. To derive the number and properties of the undetected galaxies responsible for the excess of signal detected in the FFT we performed a series of Monte Carlo simulations.
To create the mock i_{775} and z_{850} -band images of the NICP12 field we started by deriving the values of the magnitudes, below the detection limit, and the morphology to assign to the simulated galaxies. Then, we focused on the way to resemble as well as possible the noise and the large scale effects characteristic of the data. Finally, we compared the power spectra obtained from the ACS data with the results coming from the simulations. We iterated the simulation 100 times to estimate the variability of the power spectra depending on the elements we used.

2.3.1 Elements of the Simulations

To reproduce as well as possible the ACS data, it was necessary to introduce in the simulations all the different contributions with the best characterization possible. In the following all the elements used in the simulations will be fully discussed.

- *Random Noise:* We used the `IDL RANDOMN` routine to create a frame with the same size of our images and randomly distributed values with a mean of zero and a standard deviation of one. Subsequently, to reproduce the typical noise of the ACS data, we multiplied this frame by the characteristic standard deviation of the i_{775} and z_{850} images we measured with the `IRAF`[2] task `IMSTATISTICS`. The resulting two random noise frames have a white power spectrum and reproduce exactly the trend observed at high wave numbers in Figure 2.3.

[2]Imaging Reduction and Analysis Facilities (IRAF) is distributed by the National Optical Astronomy Observatories which are operated by the Association of Universities for Research in Astronomy (AURA) under cooperative agreement with the National Science Foundation.

- *Large Scale Effects:* The residuals of the flat field correction produce the peak at low wave numbers in both the i_{775} and z_{850} -band power spectra (Figure 2.3). Since we were not interested in studying the trend in the low wave number range, we simply searched a way to recover the shape of the power spectrum in this range. We created a model of flat field residuals to reproduce the large scale effects and, then, we added it to the simulated frame (additive flat field residuals). To be sure our model was not affecting the results of the simulations, we performed tests using mock frames with only noise and large scale residuals, assuming different models for the latter. In particular, we considered a Gaussian model, a uniform model (i.e. a constant flat), and one with a linear slope. As can be seen in Figure 2.15 the differences in the results are confined in the very low wave number range ($k < 100$), up to the dashed vertical line and it is not possible to reproduce the signal excess due to faint galaxies with only noise and flat field residuals. On the basis of these results we can affirm that the adopted model is not affecting the range we are interested in, and the signal in $100 < k < 400$ is not affected by our choice regarding the modeling of the residuals of the flat field correction.

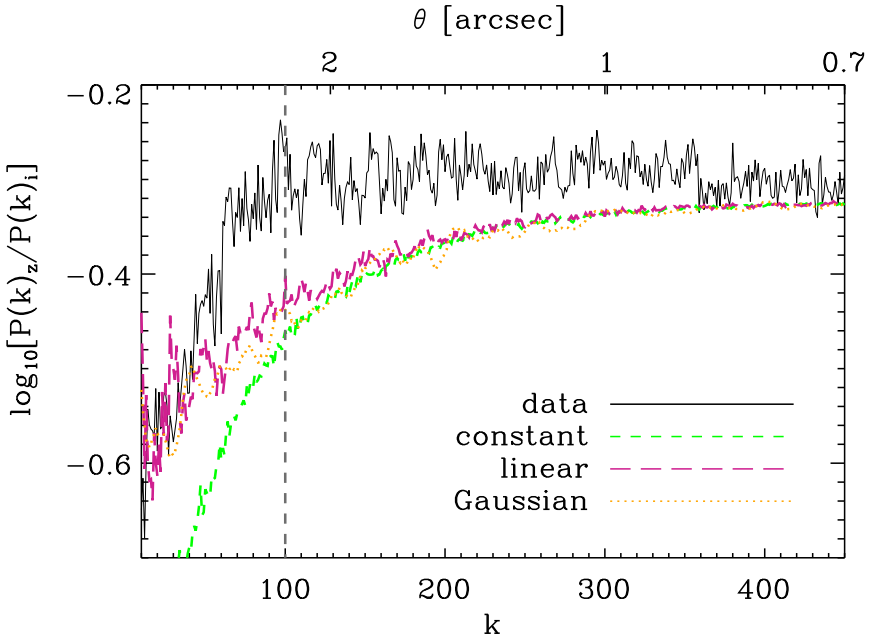


Figure 2.15: Comparison of the ratios between the z_{850} and i_{775} -band power spectra obtained from simulations including only noise and large scale residuals. Different models of large scale residuals were tested: a constant (green short dashed line), a linear one (magenta long dashed line), and a Gaussian one (orange dotted line). In black is plotted the ratio between the power spectra derived from the data. The dashed vertical line indicates the limit of the wave numbers range affected by the type of model considered for the large scale residuals.

- *Faint Galaxies:* We simulated both the i_{775} and z_{850} -band images of the NICP12 field. To determine the magnitudes of the simulated faint galaxies, we considered all the detected galaxies of the field. The histogram of the distribution of the values of the to-

tal magnitude, which were obtained from the MAG_AUTO parameter in SExtractor, for the i_{775} -band image is shown in the top left panel of Figure 2.16. The value of 28.5 mag, where the number of galaxies decreases by a factor larger than 50%, is assumed as the brighter limit for the range of magnitudes used in the simulation of the i_{775} -band image. The fainter limit is $i_{775} = 30.5$ mag, as we will see later in this Section.

The number of faint galaxies with $z \leq 5$ to be inserted in the mock i_{775} -band image was derived extrapolating the trend of the number counts beyond the cut off assuming a slope $\alpha = -1.3$ according to Metcalfe et al. (1995). The number of galaxies for each bin of magnitude used in the simulation is specified in Table 2.2.
To reproduce the z_{850} -band image, we considered both the $z \sim 6$ faint galaxies and those at $z \leq 5$. The latter include the same mock galaxies used to simulate the i_{775} -band image, whose magnitudes were corrected with a color term randomly drawn from a distribution reproducing the $i_{775} - z_{850}$ colors observed for the detected galaxies. The former population consists of $z \sim 6$ galaxies not detectable in the i_{775} image. The brighter limit for the magnitude range of the mock $z \sim 6$ galaxies is assumed to be the value of 28 mag where the histogram of the total magnitudes shows a drop larger than 50% (Figure 2.16, bottom left panel). The fainter limit of the range considered is $z_{850} = 30$ mag. We extrapolated to fainter magnitudes the luminosity function at $z \sim 6$ to obtain the number of galaxies to be created.

All the simulations hereafter include a flat field residual model. After a series of tests, we found a good agreement in reproducing this effect by combining 20 Gaussian functions with a standard deviation of about one third of the image dimension: 10 of them have a peak of 10^{-4} to match the typical fluctuations of the flat field, and the other 10 Gaussians have a peak of -10^{-4} to create a smooth background with soft large scale fluctuations. The chi-square statistics results obtained comparing the data to the simulations including the three different flat field models plus mock galaxies and noise in the range $100 < k < 400$ were very similar (see Columns 3 and 4 in Table 2.3 and panel (b) of Figure 2.24).

We tested different values for the slope α, in particular $\alpha = -1.3$, $\alpha = -1.9$, and $\alpha = -2.4$. On the basis of a χ^2 statistics (see columns 7 and 8 of Table 2.3 and (b) panel of Figure 2.24) we assumed our extrapolation to have the slope $\alpha = -1.9$ (Figure 2.18) and we anchored it to the LF value at a magnitude of 28 mag given by Bouwens et al.

$z \leq 5$, $\alpha = -1.3$		$z = 6$, $\alpha = -1.9$	
Magnitude [mag]	Number	Magnitude [mag]	Number
28.5	714	28.0	32
29.0	1605	28.5	87
29.5	3008	29.0	233
30.0	5503	29.5	622
30.5	10028	30.0	1665

Table 2.2: Number and magnitudes of the simulated galaxies used to reproduce the population of galaxies at $z \leq 5$ (left) and at $z \sim 6$ (right). The slope used for the extrapolation is specified on the top.

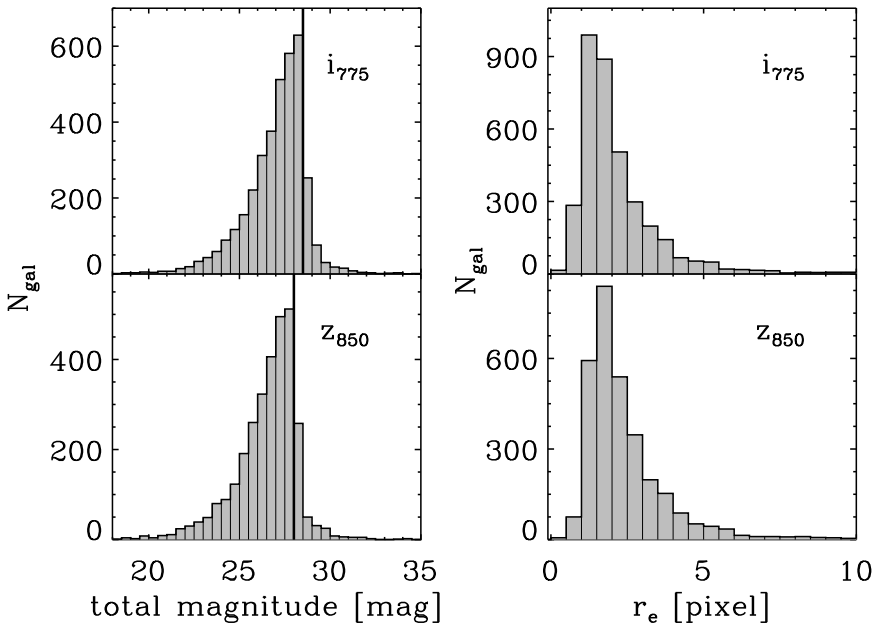


Figure 2.16: Distribution of the values of the total magnitude (left panels) and the effective radius (right panels) of the detected galaxies in the i_{775} (top panel) and z_{850} -band (bottom panel) images. The vertical lines plotted on the left histograms correspond to the magnitude where we observe a drop in the galaxy number and which was assumed as the brightest magnitude used in our simulations.

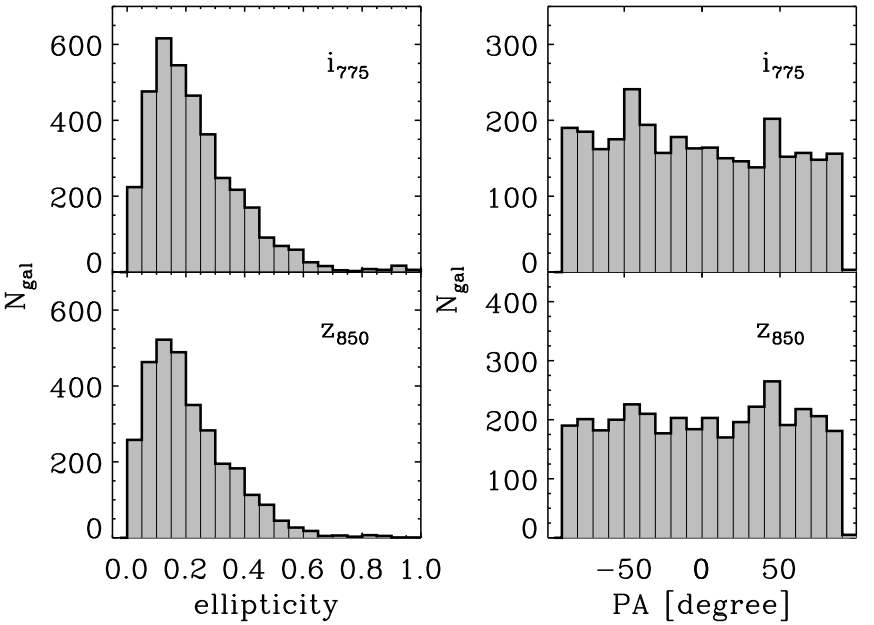


Figure 2.17: Distribution of the values of the ellipticity (left panels) and position angle (right panels) of the galaxies detected in the i_{775} (top panel) and z_{850} (bottom panel) images.

(2007). This value is in agreement with the expectation of Su et al. (2011) who found the faint-end slope of the luminosity function at $z \sim 6$ to be in the range $-1.90 < \alpha < -1.55$.

Testing the effect of groups of galaxies as described in Section 2.2.4, we considered different slopes α, i.e. $\alpha = -1.7$, $\alpha = -1.8$, and $\alpha = -1.9$ introducing pairs of mock galaxies in the simulation (Figure 2.19). The χ^2 statistics shows that $\alpha = -1.8$ fits best the data ($\chi^2_\nu = 1.06$, see Columns 9 and 10 of Table 2.3).

Finally, we also considered the effect of contaminants at lower redshift changing the low-luminosity slope of the distribution of galaxies with $z \leq 5$. To this aim we considered higher and lower values for α_i and we noted that this change produces a similar one, with opposite sign, in the derived α value at $z \sim 6$.

In building the set of galaxies, we considered a range of geometrical parameters such as the position and orientation on the frame, as well as the structural ones, like the effective radius, $r_{\rm e}$, and axis ratio, b/a.

Each synthetic galaxy was generated with a set of random, equally distributed, values for the former parameters and a set of values which resembles the characteristics of the fainter detected galaxies of the NICP12 field, derived from the catalog obtained running `SExtractor`, for the latter. Regarding the former parameters, each synthetic galaxy was generated with a set of random values. On the other hand, the structural parameters values had to resemble the characteristics of the fainter detected galaxies of the NICP12 field, derived running `SExtractor`.

An analysis of the characteristics of the galaxies detected in our field was performed using `SExtractor`. Specifically, the fainter detected ones were selected so as the mock galaxy could resemble them as well as possibile. The histograms in Figure 2.16 and Figure 2.17 show the distribution of the values for the effective radius, the ellipticity, and the position angle (PA) for both the i_{775} (top panel) and z_{850} -band (bottom panel) data. From the ellipticity $e = 1 - b/a$, the axis ratio b/a was derived.

In the simulations the values of the effective radius were chosen within the range $1.5 \leq r_{\rm e} \leq 11$ pixel, corresponding to 0.135" $\leq r_{\rm e} \leq$ 0.99" (see Figure 2.20), and the range spanned by the axis ratio was $0.3 \leq b/a \leq 1$. Even though the distribution of $r_{\rm e}$ we used was only slightly different from the one found by Ferguson et al. (2004) at $z \sim 5$, we tested the effect of this parameter on the simulations (Figure 2.21). In particular, the reduced chi-square statistics ($\chi^2_\nu = 1.22$) supports the idea that small differences in the $r_{\rm e}$ distribution do not significantly affect the final result.

- *Galaxy Morphology:* Regarding the morphology of the mock galaxies, we adopted a Sérsic profile. The $r^{1/n}$ law was introduced by Sérsic in 1968, as a generalization of the De Vaucouleurs $r^{1/4}$ law. The formula for the Sérsic profile is

$$I(r) = I_{\rm e} \cdot e^{-b_n[(r/r_{\rm e})^{1/n}-1]} \tag{2.14}$$

It is based on three parameters: r_e which is the effective radius which includes half the light of the galaxy, I_e which is the surface brightness at r_e, and n, that is a shape parameter used to distinguish between ellipticals and disk galaxies. The coefficient b_n, which is chosen to make r_e being the effective radius, depends on n and can be

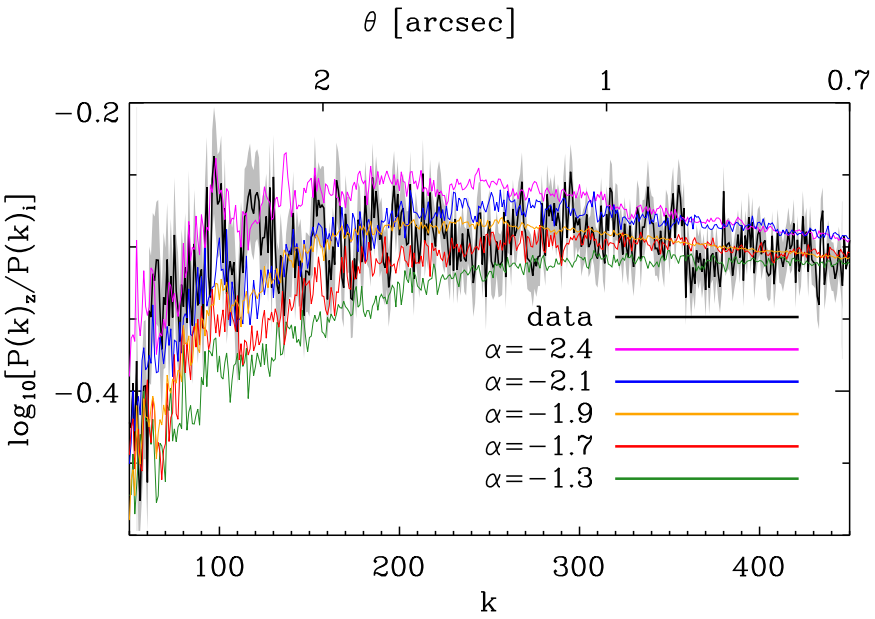


Figure 2.18: Comparison of the ratio between the simulated power spectra derived assuming different faint-end slopes. The green, red, orange, blue, and magenta lines represent the median obtained from simulations assuming $\alpha = -1.3$, $\alpha = -1.7$, $\alpha = -1.9$, $\alpha = -2.1$, and $\alpha = -2.4$, respectively. In black the ratio derived from the data is plotted with its error bars.

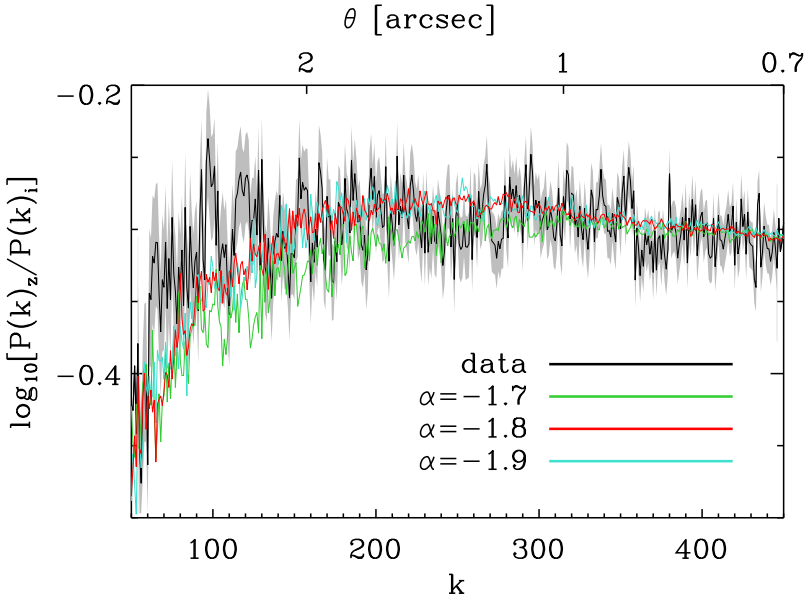


Figure 2.19: Comparison of the ratio between the simulated power spectra derived assuming different faint-end slopes and pairs of galaxies with a distance derived from the angular correlation function at $z \sim 5$ (Lee et al., 2006). The green, red, and blue lines represent the median obtained from simulations assuming $\alpha = -1.7$, $\alpha = -1.8$, and $\alpha = -1.9$, respectively. In black the ratio derived from the data is plotted with its error bars.

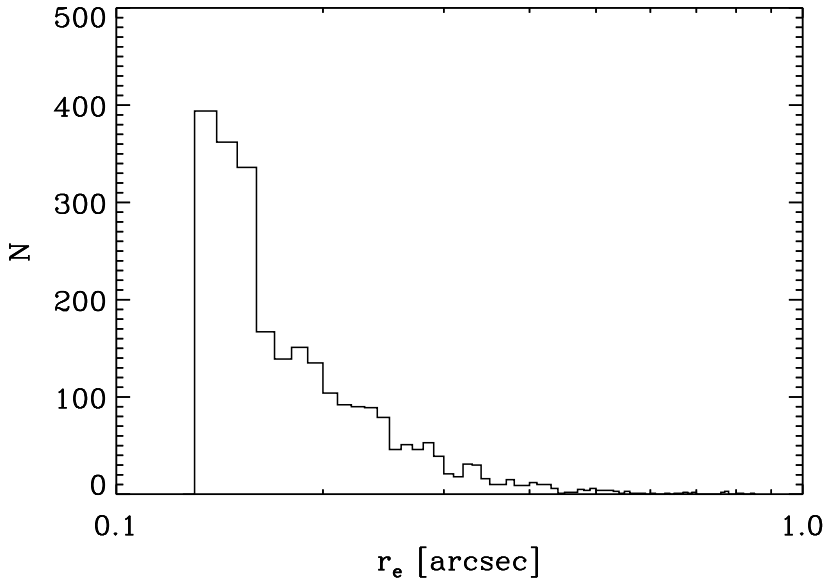


Figure 2.20: Example of distribution of the effective radius of the faint i-dropout galaxies in arcseconds.

approximate as follow

$$b_n = 1.9992 \cdot n - 0.3271 \quad 0.5 < n < 6 \tag{2.15}$$

$$b_n = 2 \cdot n - 0.332 \quad n \geq 6 \tag{2.16}$$

When $n = 4$ the Sérsic law describes a De Vaucouleurs profile, when $n = 1$ the exponential one. We choose the Sérsic profile because, simply changing the Sérsic index n, it permitted to obtain ellipticals ($n = 4$) and spirals ($n = 1$).

Since there are no clear indications regarding the distribution of different galactic types in the early Universe, we studied the effect of changing the relative number of ellipticals and spirals in our simulations. We compared the results obtained introducing 15% ellipticals (described by a de Vaucouleurs profile) and consequently 85% spirals (described by an exponential profile) in the simulations with those obtained simulating 50% ellipticals and 50% spirals and 33% ellipticals and 67% spirals. In Figure 2.22 the ratio of the power spectra obtained from the data is compared with the mean ratios and 3σ confidence ranges derived from simulations with different distributions of ellipticals and spirals. Applying the chi-square statistics to the wave number range where the peak is, i.e. $100 \lesssim k \lesssim 400$ (see Columns 1 and 2 in Table 2.3) we selected the realization characterized by the smallest χ^2 value as the best fit (see also the distribution of the residuals in the (a) panel of Figure 2.24). We found the best fit corresponding to the mix of 33% ellipticals and 67% spirals. These percentages are consistent with the findings of Somerville et al. (2001), who predicted the majority of galaxies at $z \geq 3$ to be disk dominated, as well as those of Lotz et al. (2006) and Ravindranath et al. (2006), who found the percentage of galaxies with bulge-like morphology to be $\sim 30\%$ at $z \sim 3$ and $z \sim 4$, respectively.

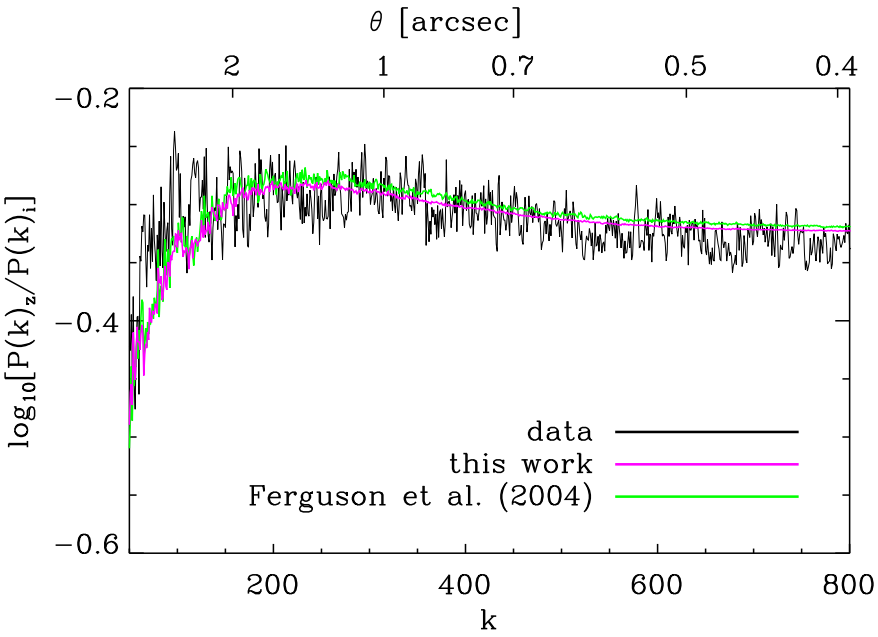


Figure 2.21: Comparison between the ratio of the power spectra obtained from the data (black line) and from the simulations assuming the effective radius distribution plotted in Figure 2.20 (magenta line) and the one by Ferguson et al. (2004, green line)

This mix was adopted to test the sensitivity of the method to the assumed value of the faint-end limit. To this aim we performed a few simulations changing the fainter magnitude limit of the mock galaxies (Figure 2.23). We used the chi-square statisticss to select the best simulation concerning the magnitude cut-off (see Columns 5 and 6 in Table 2.3). Moreover, we computed the difference between the simulations and

Percentage of E and S		Flat field residual model		Magnitude cut-off		Faint-end slope α		Grouping slope α	
sim	χ^2_ν	sim	χ^2_ν	sim	χ^2_ν	sim	χ^2_ν	sim	χ^2_ν
15%E	1.47	Gaussian	1.02	29.5 mag	1.60	α=-1.3	3.22	α=-1.7	1.60
33%E	1.02	linear	1.04	30.0 mag	1.70	α=-1.7	1.64	α=-1.8	1.06
50%E	1.49	no flat	1.04	30.5 mag	1.02	α=-1.9	1.022	α=-1.9	1.17
				31.0 mag	1.33	α=-2.1	1.47		
						α=-2.4	2.53		

Table 2.3: The result obtained from the χ^2 statistics are listed for all the realization of the simulations regarding the percentage of elliptical and spiral galaxies (Columns 1 and 2), the flat field residual model (Columns 3 and 4) introduced, the value of the cut-off in magnitude (Columns 5 and 6), and the slope of the faint-end of the LF with single sources (Columns 7 and 8) and with pairs of galaxies (Columns 9 and 10), respectively. The selected realization is the one characterized by the smallest χ^2 value. All the realizations with $\chi^2 > 1.24$, which corresponds to a confidence level of 0.3%, can be completely rejected with much more than a 3σ confidence level.

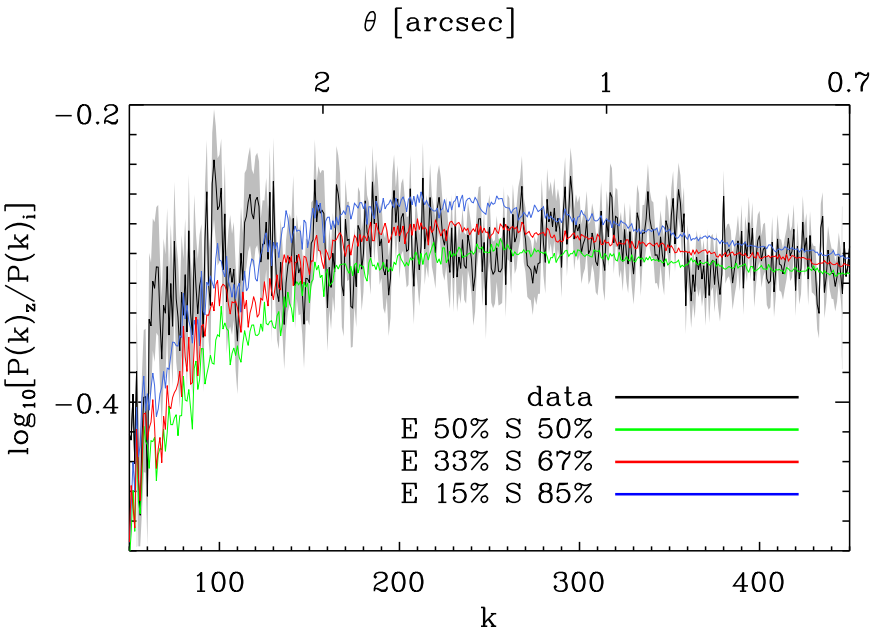


Figure 2.22: Comparison of the ratio between the z_{850} and i_{775} power spectra (black line) with the 3σ confidence ranges derived from the simulations with a different percentage of spirals and ellipticals represented by different colors, 50% ellipticals and 50% spirals in green, 33% ellipticals and 67% spirals in red, and 15% ellipticals and 85% spirals in blue.

the data and we plotted the residuals (Figure 2.24, (c) panel). On the basis of above, we decided to create galaxies with magnitudes up to 30.5 mag.

- *The Mask:* Consistent results can be achieved only by analyzing the same area in data and simulations. For this reason we used, for the simulated images, the same mask created from the ACS data. Moreover, we considered what percentage of the frame is masked and calculated the number of mock galaxies that had to be outside the masked area. We, therefore, placed them randomly on the frame checking that they were positioned outside the mask.

Finally, the frame with the mock galaxies was convolved with the HUDF PSF (FWHM $\sim$ $0.''1$, Oesch et al. 2007). Then, the sum of the simulated galaxies, the frame reproducing the large scale effects, and the random noise frames was multiplied by the mask and analyzed with the power spectrum technique.

2.3.2 Results of the Simulations

The results obtained iterating 100 times the simulations considering the 5 bins of magnitudes given in Table 2.2 and assuming one third of the galaxies to be ellipticals and the others spirals are shown in Figure 2.25. The trend of the data with the associated errors is well fitted by the mean ratio between the simulated power spectra, and all the values are in agreement with the 3σ confidence range.

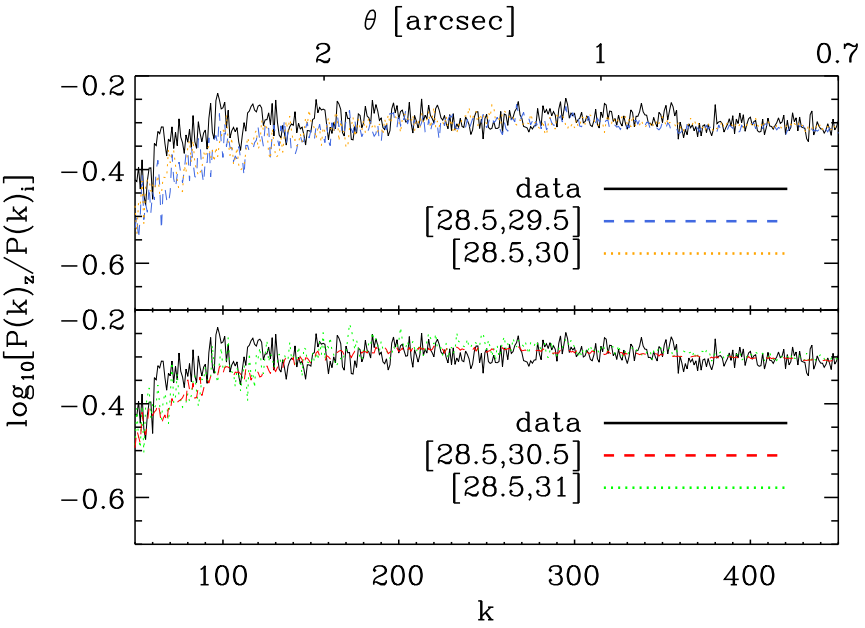


Figure 2.23: Comparison of the observed ratio between the z_{850} and i_{775} power spectra (black) and those obtained by simulating unresolved galaxies in different magnitude ranges, i.e., $28.5 < i_{775} < 29.5$ (blue dashed line), $28.5 < i_{775} < 30$ (orange dotted line), $28.5 < i_{775} < 30.5$ (red dashed line), and $28.5 < i_{775} < 31$ (green dotted line).

2.3.3 Detected i_{775} -dropouts

For the following analysis we had to derive the number of detected candidate i_{775} -dropouts galaxies existing in the NICP12 field. The selection criteria we used were derived by Bouwens et al. (2007) and Su et al. (2011) and consist in a color cut and two conditions regarding the S/N ratio which are needed to largely avoid interlopers:

$$\begin{cases} i_{775} - z_{850} & > & 1.3 \\ S/N(z_{850}) & > & 5 \\ S/N(V_{606}) & < & 2 \text{ or } V_{606} - z_{850} > 2.8 \end{cases} \tag{2.17}$$

The last two criteria, concerning the V_{606} band are equivalent, so it is possible to use either the first or the second one. In any case, to verify the last criterion it was necessary to analyze also the V_{606} band image of the field. To this aim, we rebinned it so as to have the same pixel scale of the images in the i_{775} and z_{850} band, i.e. 90 mas pix^{-1}. Then, we run `SExtractor` in dual-image mode, which permits to derive the photometry in different bands of the objects identified in the redder one. We selected the z_{850} band as the detection one, used to generate the catalog of the sources, and then we got the photometry of these objects in all the three bands.
According to the criteria listed above we obtained a catalog consisting of 84 candidate i_{775} -dropout galaxies, in accordance with the results found by Bouwens et al. (2007) for the same field.

As a complementary study we focused on the characteristic of these candidates. In particular, on the basis of the output parameters given by `SExtractor` we were able to study

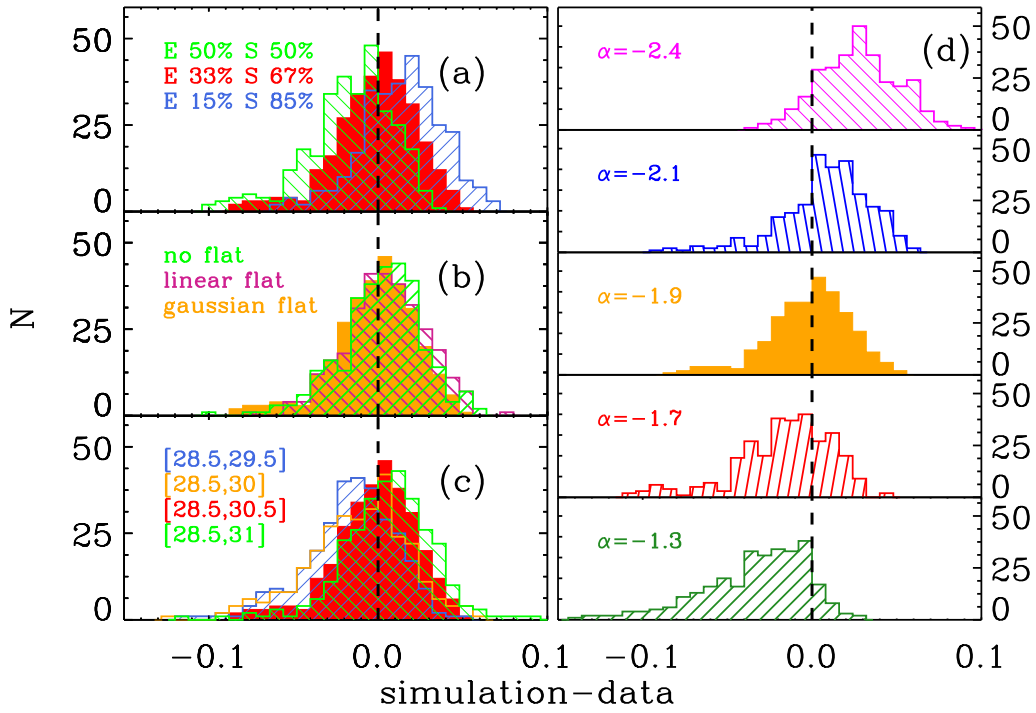


Figure 2.24: Distribution of the residuals derived computing the difference between the different realization of the simulation and the data, in terms of the ratio between power spectra. The top left panel, (a), refers to the different percentages of spirals and ellipticals, the central left one, (b), to the various flat field residual models, the bottom left one, (c), concerns the magnitude cut-off, and the right one, (d), is about the different value of the faint-end slope α. The filled histograms refer to the selected realization, the one with the lowest χ^2.

the distribution of their total magnitude in the i_{775} and z_{850} band (left panels of Figure 2.26), as well as the distribution of the values of the effective radius (right panels of Figure 2.26), the ellipticity (left panels of Figure 2.27), and the position angle (right panels of Figure 2.27). The distributions of the values obtained for the ellipticity and position angle resemble those derived considering all the detected sources in the field (Figure 2.17). Regarding the ellipticity, the values have a peak corresponding to intermediate shapes while the values corresponding to a perfect circular shape or to an extremely elongated ellipses are not common. The position angles are almost equally distributed and do not present any preferred values, meaning that `SExtractor` does not introduce any bias detecting the sources. The typical dimension of the dropout galaxies was derived from the parameter FLUX_RADIUS in the catalogs, which corresponds to the effective radius $r_{\rm e}$. Comparing the histograms of the referred to the $r_{\rm e}$ values to those derived for all the detected galaxies (Figure 2.16), we noticed that the trend of the distribution is the same.

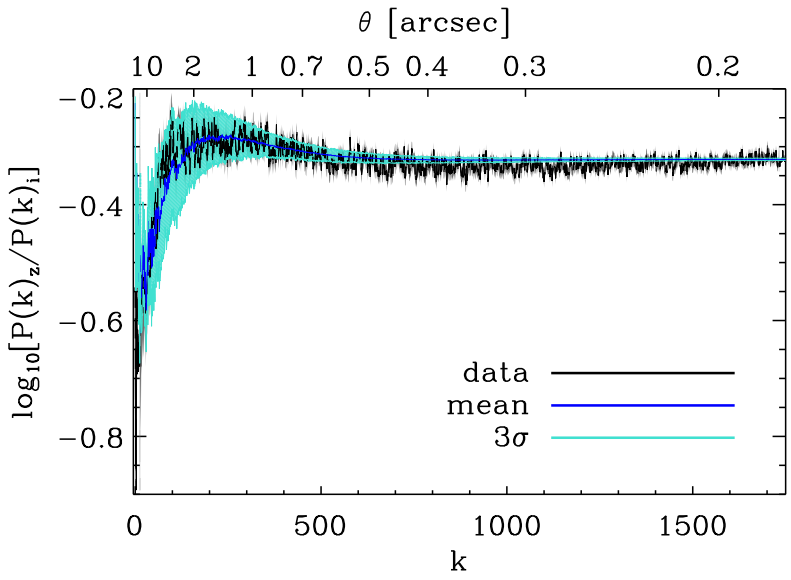


Figure 2.25: Mean ratio between the i_{775} and z_{850} simulated power spectra (blue line) and 3σ confidence range (light blue region) derived from 100 simulations compared to the results coming from the ACS data (black line).

2.3.4 The z ~ 6 Faint Galaxies Contribution to Cosmic Reionization

In order to evaluate the implications for reionization of our findings, we considered the total surface brightness contribution of the undetected galaxies following the approach of Stiavelli et al. (2004a). When this contribution is added to that of the detected i_{775} dropout galaxies, we find a value of surface brightness of 24.8 mag arcmin^{-2} when extrapolating to a magnitude $z_{850} = 29$ mag, and 23.6 mag arcmin^{-2} when extrapolating to $z_{850} = 30$ mag.

These values have to be compared to the minimum surface brightness of 27.2 mag arcmin^{-2} needed for reionization by Population II stars (Stiavelli et al., 2004b). Our results give comfortable margins of a factor 9 and 27, respectively, to accommodate the escape fraction f_c and the clumping factor C. For instance, assuming $C = 6$ and the Population II spectral energy distribution (SED) of Stiavelli et al. (2004b), the LF extrapolated down to $z_{850} = 30$ produces enough ionizing photons to reionize the Universe if $f_c \gtrsim 5\%$.

So far the implications for reionization of what we found seem to be very promising, but we are waiting for a more thorough analysis of all available fields before pushing this interpretation any further.

2.4 Conclusions

We showed that the power spectrum technique is a powerful tool in analyzing the light contribution produced by galaxies which are below the detection limit in deep and ultra deep surveys. We used this technique to estimate the contribution to cosmic reionization

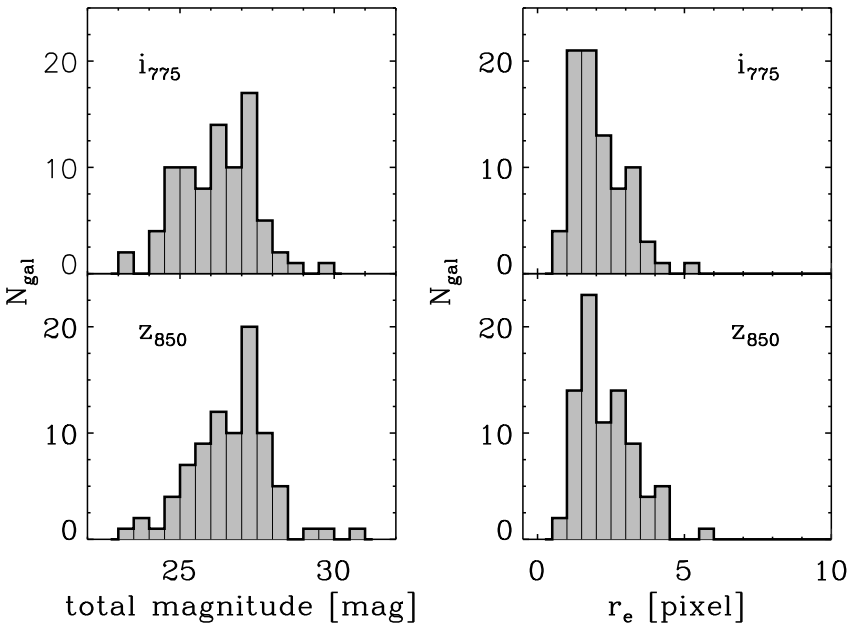


Figure 2.26: Distribution of the values of the total magnitude (left panels) and the effective radius (right panels) of the candidate i_{775} -dropouts galaxies detected in the i_{775} (top panel) and z_{850} (bottom panel) images.

from faint galaxies ($z_{850} \geq 28$ mag), which are bona fide $z \sim 6$ candidates, in the NICP12 ACS field of the UDF05.

Monte Carlo simulations were used to determine the number of faint undetected i-dropouts responsible for the peak observed in the ratio between the z_{850} and i_{775} power spectra (Figure 2.4). The data are consistent with a faint-end slope of the luminosity function of $\alpha = -1.9$ if not introducing any clustering/grouping in the i_{775} -dropouts distribution, and $\alpha = -1.8$ when considering pairs of galaxies with distances drawn accordingly to the angular correlation function by Lee et al. (2006).

Considering $\alpha = -1.9$, adding the total surface brightness contribution of these galaxies to that from the detected galaxies and comparing it to the minimum value required to reionize the Universe, we obtained a margin to model the escape fraction and the clumping factor. Adopting the clumping factor of 6 and Population II SED of Stiavelli et al. (2004b), the $z \sim 6$ undetected galaxies down to $z_{850} = 30$ mag produce enough ionizing photons to reionize the Universe assuming the escape fraction to be larger than $\sim 5\%$. The solution with all galaxies in pairs and a slope $\alpha = -1.8$ changes the cumulative surface brightness by less than 0.1 mag and implies an escape fraction larger by $\sim 10\%$. Thus, the main conclusion of this chapter that reionization is, indeed, consistent with being completed due to the contribution of faint galaxies is not affected by the presence of multiple halo occupation.

In the following chapter we will decribe the application of our technique to the ultra deep data obtained with the Wide Field Camera 3 (WFC3). The comparison of the results for different fields and instruments will permit us to better constrain the role that faint

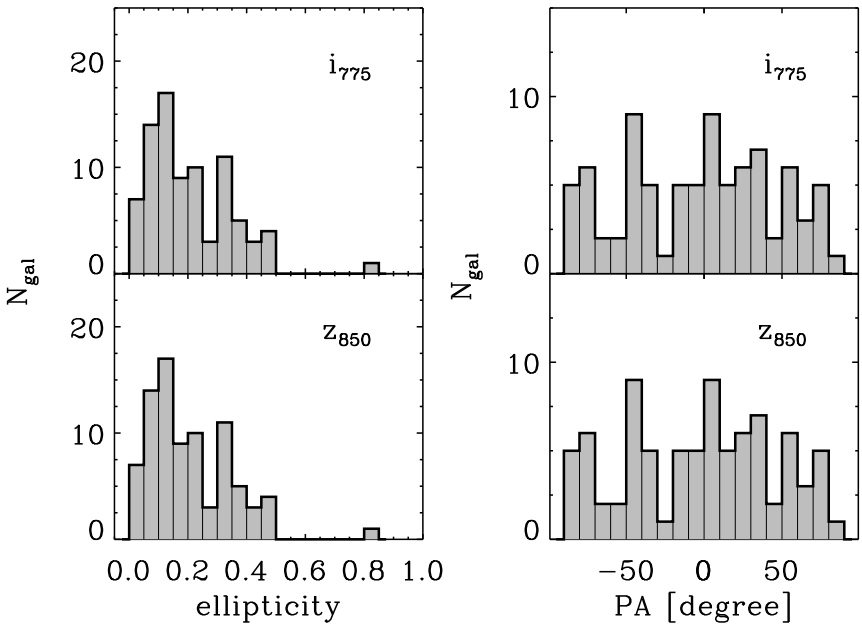


Figure 2.27: Distribution of the values of the ellipticity (left panels) and position angle (right panels) of the i_{775} -dropouts candidate galaxies detected in the i_{775} (top panel) and z_{850} (bottom panel) images.

galaxies at $z \sim 6$, $z \sim 7$, and $z \sim 8$ played in the cosmic reionization.

Chapter 3

Looking for Faint Galaxies at $z \sim 7$-8: Analysis of WFC3 Data from the HUDF09 Project

In this chapter we will present the analysis of the near-infrared deep data obtained with the Wide Field Camera 3 on HST. The choice of this set of data was a natural one, as a follow up of the work presented in the previous chapter.
Our original idea was to start working on the z_{850} and Y_{105} -band images and, then, move on comparing the Y_{105} and J_{125} -band images to constrain the contribution to reionization coming from $z \sim 7$ and $z \sim 8$ galaxies, respectively. Actually, first of all we analyzed the contribution from faint $z \sim 8$ galaxies, since the Y_{105} and J_{125} -band data were all taken with the same camera and, then, we compared the z_{850} and Y_{105} -band ones taking into account the systematics encountered when dealing with different detectors. Unfortunately, when applying the power spectrum technique we were not able to highlight any light contribution from faint galaxies below the detection limit neither at $z \sim 7$, nor at $z \sim 8$. On the basis of our analysis we realized that, to study the role of faint high-redshift objects in the context of cosmic reionization, these near-infrared dataset requires a new and better data reduction, in particular focused on the removal of all the residuals of the previous one and systematics, as was done for the ACS data obtained for the NICP12 field.

3.1 The Data

The Hubble Ultra Deep Field 2009 (HUDF09) program observations were taken in Cycle 17 (HST program 11563, P.I.: Garth Illingworth). The program consists of 192 orbits, obtained with the infrared channel of the Wide Field Camera 3 (hereafter WFC3/IR), with the aim to image the deep ACS fields that were used in the original HUDF (P.I.: Steven Beckwith) and in the HUDF05 program (P.I.: Massimo Stiavelli). The three fields observed are called HUDF09, HUDF09-1, and HUDF09-2, corresponding to the main HUDF, the NICP12 field, and NICP34 field studied in the previous programs, respectively. The position of the three fields in the sky is shown in Figure 3.2 and the coordinates are listed in Table 3.1.

The HUDF09 data were collected from August 2009 to February 2011 and the final scientific products were obtained combining the full two-year WFC3/IR datasets over the

WFC3/IR Field	ACS Field	RA (J2000)	DEC (J2000)
HUDF09	HUDF	03:32:40.440	-27:47:24.12
HUDF09-1	HUDF05P12	03:33:00.034	-27:41:50.94
HUDF09-2	HUDF05P34	03:33:04.356	-27:51:11.84

Table 3.1: Properties of the WFC3/IR fields. Col.(1): name of the field in the HUDF09 program. Col.(2): name of the field in previous programs. Col.(3): right ascension (J2000). Col.(4): declination (J2000).

Filter	HST orbits	Exposure Time [s]	Zero Point [AB mag]
HUDF09 main field			
F105W	24	67041	26.27
F125W	34	94500	26.25
F160W	53	146711	25.96
HUDF09-1			
F105W	8	21947	26.27
F125W	12	33170	26.25
F160W	13	35376	25.96
HUDF09-2			
F105W	11	30865	26.27
F125W	18	50506	26.25
F160W	19	53312	25.96

Table 3.2: Properties of the HUDF09 data (Bouwens et al., 2012a). Col.(1): filter name. Col.(2): number of HST orbits. Col.(3): exposure time in seconds. Col(4): zero point in AB magnitudes.

fields. The full two-year WFC3/IR observations consist of 192 orbits of ultra-deep WFC3/IR data over the three fields, 111 orbits for the main field HUDF09, 33 orbits for the HUDF09-1, and 48 orbits for the HUDF09-2, respectively. For all the three fields the observations were concentrated so as to point a single ~ 4.7 arcmin^2 area in the sky.
As in all the previous ultra deep surveys, the main goal was to get deep observations of the region corresponding to the main field. The 111 orbits used to image it were distributed over the three bands Y_{105} , J_{125} , and H_{160} , with 24, 34 and 53 orbits, respectively. All the pixels affected by source persistence were masked out.

The WFC3/IR observations over the other two fields, HUDF09-1 and HUDF09-2, were useful to enlarge the sample of galaxies at high-redshift. Regarding the HUDF09-1 field, 8 orbits were acquired in the Y_{105} band, 12 orbits in the J_{125} band, and 13 orbits in the H_{160} band. Finally, for the HUDF09-2 field, 11 orbits in the Y_{105} band, 18 orbits in the J_{125} band, and 19 orbits in the H_{160} band were acquired.

The dither strategy used in the HUDF09 program is similar to the one used for previous programs, so as to include

- large dithers (~ 3 arcsec) in the +y ACS/WFC direction to effectively step across the

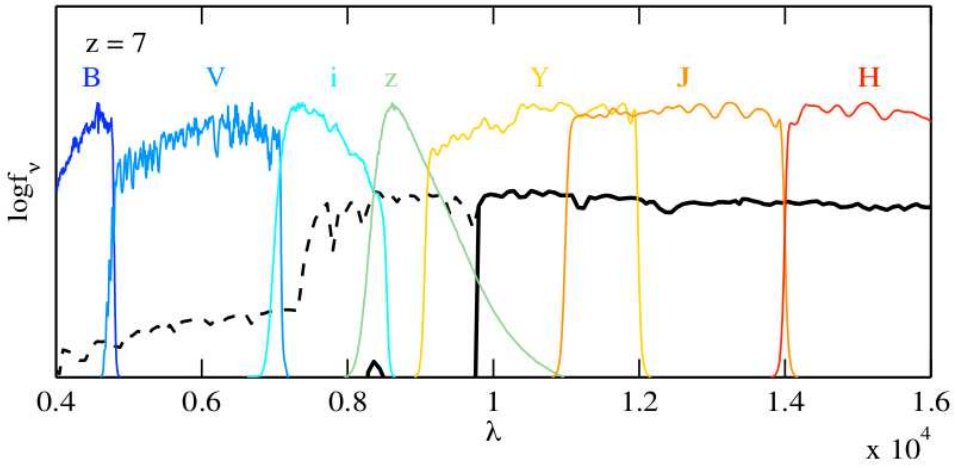


Figure 3.1: The black solid line represents the spectrum of a Lyman-break galaxy at $z \sim 7$ overlapped by the HST filters in the optical (ACS: B_{435}, V_{606} , i_{775} , and z_{850} band) and near-IR (WFC3/IR: Y_{105} , J_{125} , and H_{160} band).The dotted line indicates the part of the spectrum we can not detect due to IGM absorption. (Figure courtesy of Dr. P. Oesch.)

ACS gap (2.5 arcsec),

- smaller 4-point dithers to cope with WFC3/IR's undersampling of the HST PSF and, therefore, to improve the resolution of the final data set,
- intermediate-sized ±x dithers ($\sim 1-2$ arcsec) to maximize the range of different dither positions within the set of observations.

Bouwens et al. (2011) analyzed the HUDF09 fields to identify galaxies at $z \geq 7$. The ultra deep observations obtained in the three bands Y_{105} , J_{125} , and H_{160} allowed them to use the Lyman-Break selection technique to identify star forming galaxies at $z \gtrsim 7$. As can be seen in Figure 3.1, the combination of the Y_{105} , J_{125} , and H_{160} filters, in addition to a non detection in the bluer optical bands, permits to select $z \sim 7-8$ galaxies. The color selection criteria used are:

$$\begin{aligned} z_{850} - Y_{105} &> 0.7 \\ Y_{105} - J_{125} &< 0.45 \\ z_{850} - Y_{105} &> 1.4 \cdot (Y_{105} - J_{125}) + 0.42 \end{aligned} \tag{3.1}$$

for the sample of $z \sim 7$ z_{850} -dropout sample and

$$\begin{aligned} Y_{105} - J_{125} &> 0.45 \\ J_{125} - H_{160} &< 0.5 \end{aligned} \tag{3.2}$$

for the sample of $z \sim 8$ Y_{105} -dropout.
Combining the full two-year WFC3/IR and ACS data and considering the color selection criteria, they found 60 $z \sim 7$ and 54 $z \sim 8$ candidate galaxies (Table 3.3). Moreover, they derived the luminosity functions at $z \sim 7-8$ up to magnitudes < -18 $\mathrm{mag_{AB}}$, as well as the star formation density at the same redshifts. Combining the wide-area $z \sim 7$ search

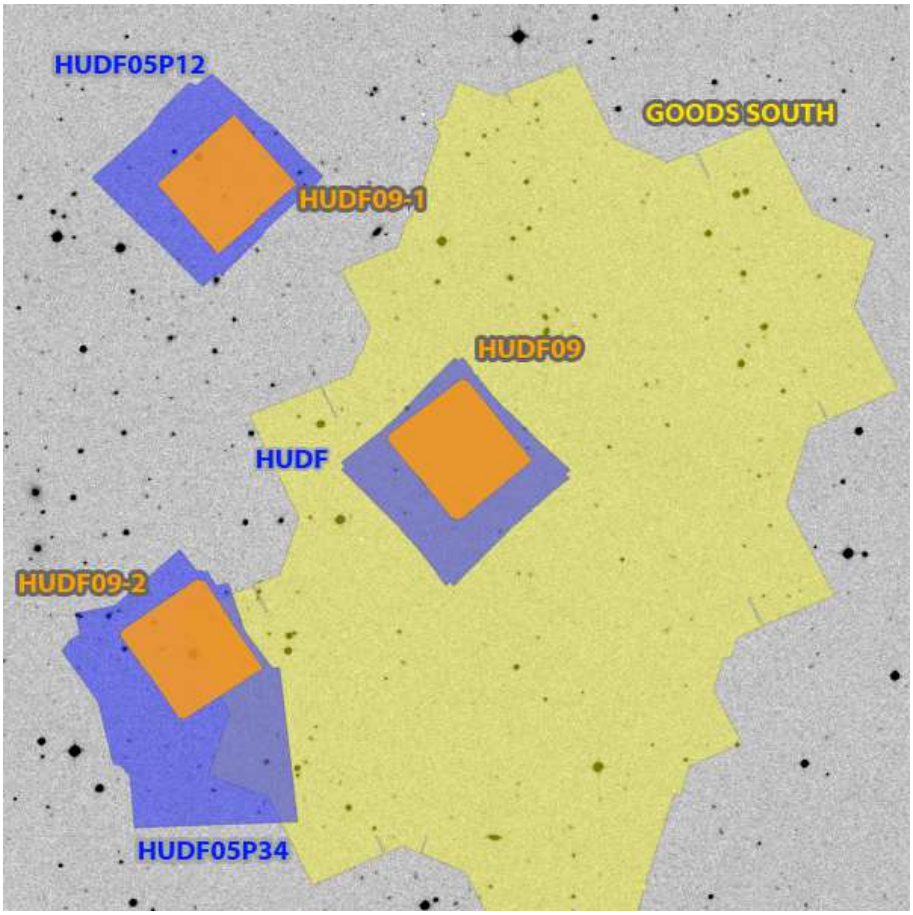


Figure 3.2: The WFC3/IR HUDF09 fields are shown in orange. The blue rectangles represent the fields imaged during the ACS HUDF and HUDF05 programs. Finally, for reference the ACS GOODS-South data are shown in yellow.

results (ERS) with the ultra-deep WFC3/IR observations, they derived a LF characterized by $M^* = -20.14 \pm 0.26$, and $\alpha = -2.01 \pm 0.21$. At $z \sim 8$ the best-fit Schechter parameters they found are $\alpha = -1.91 \pm 0.32$, $M^* = -20.10 \pm 0.52$, and $\Phi^* = 0.59^{+1.01}_{-0.37} \cdot 10^{-3}$ Mpc^{-3}.

Field	z_{850} -dropouts	Y_{105} -dropouts
HUDF09	29	24
HUDF09-1	17	14
HUDF09-2	14	16

Table 3.3: Number of z_{850} (Col. 1) and Y_{105} (Col. 2) candidate dropout galaxies detected using WFC3/IR in each of the three fields imaged in the HUDF09 program (Bouwens et al., 2011).

From the data in Bouwens et al. (2011) we derived the surface density of z_{850} (Figure 3.3) and Y_{105} -dropouts (Figure 3.4) as a function of the magnitude in the Y_{105} and J_{125} -

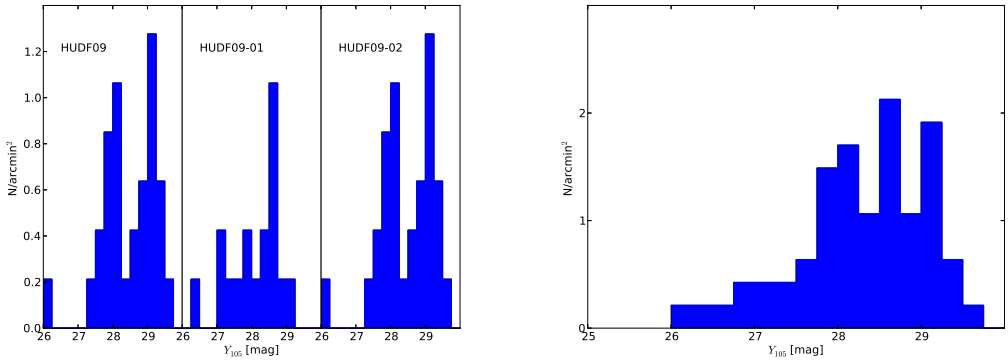


Figure 3.3: Top panels: surface density of z_{850} -dropouts in each of the HUDF09 program field as a function of the magnitude in the Y_{105} band using 0.5 mag bins. Bottom panel: total surface density of Y_{105} -dropouts derived considering all the three fields of the HUDF09 program.

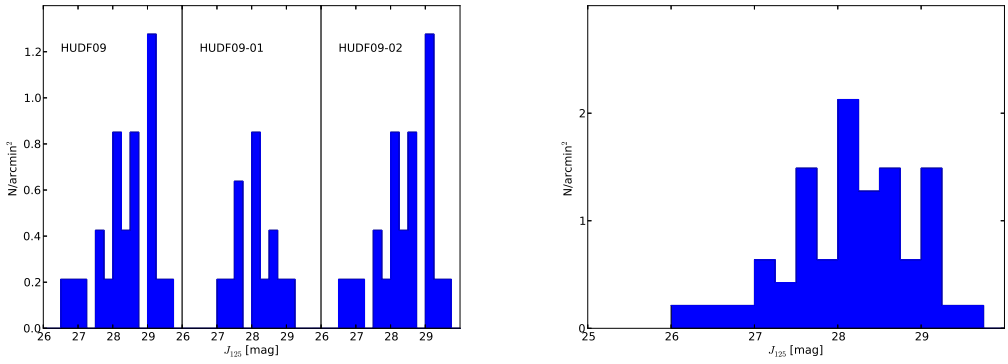


Figure 3.4: Top panels: as in Figure 3.3, but for Y_{105} -dropouts as a function of J_{125} magnitude. Bottom panel: total surface density of z_{850} -dropouts.

band, respectively, for each of the HUDF09 fields. Moreover, we obtained the distribution of the total surface density of z_{850} and Y_{105} -dropouts as a function of the magnitude in the Y_{105} and J_{125} -band, respectively, from all the HUDF09 program fields (bottom panels of Figures 3.3 and 3.4). These distributions were used to derive the value to which anchor the extrapolation of the LF towards fainter magnitudes needed to run Monte Carlo , as we did and fully described in the previous chapter. It should be noted that, according to Bouwens et al. (2012b), the luminosity function at $z \sim 7-8$ could have a shape which is different from the regular Schechter one we used in the previous chapter, but which can be described using a power-law plus an exponential cut-off towards bright magnitudes.

3.2 Data Reduction

The WFC3/IR data are characterized by several problems, such as persistence, IR blobs, and other defects that will be fully described in the following Sections (Sec. 3.2.1 and 3.2.2).

Moreover, these data were obtained using dithering and, then, combined with `MultiDrizzle`. For sake of clarity both dithering and `MultiDrizzle` are described in Sections 3.2.3 and 3.2.4, respectively.

All the detector defects must be considered when masking and analyzing the images, if not they could determine spurious signals and modify the results. At the same time, the dithered strategy should be known and taken into account when running `MultiDrizzle`, which could affect the final result derived with the power spectrum technique.

3.2.1 Persistence

The IR detector of the WFC3 camera is characterized by image persistence. Persistence is caused by traps that exist in the active regions of diodes that make up the pixels of the detector. When the diodes are exposed to light, voltage levels within the diode change slightly, allowing free electrons and holes to reach these traps. Discharging the diodes, electrons and holes that were previously trapped escape the traps slowly over time and cause after images (Smith et al., 2008). The number of traps increases with the saturation of the detector, so persistence is affecting the data, in particular, during and after each observation of a saturated target. The brightness of the source and the exposure time determine the amount of image persistence observed. To good approximation, persistence appears to decay roughly as a power law with time, at least for delay times greater than about 100 s, suggesting that the decay is faster at lower levels of saturation. The following equation, developed by the WFC3/IR team at STScI, describes the effect of persistence with time

$$P_{ij} = N_{ij} \left(\frac{1}{e^{(x-x_0)/\delta x} + 1} \right) \left(\frac{x}{x_0} \right)^{\alpha} \left(\frac{t}{1000\ s} \right)^{-\gamma} \tag{3.3}$$

The parameters used in the equation are:

- P_{ij} is the persistence in the ij^{th} pixel
- N_{ij} is the normalization factor at 1000 s, which depends on the position
- x represents the maximum depth
- x_0 (Fermi Energy) and $\delta(x)$ define the midpoint and the width of the region where the persistence is rising rapidly
- α indicates the law index of the slow increase in persistence in the case of high saturation level
- t is the time passed since the pixel was filled
- γ is the power low slope that defines the decay time.

All bright sources (close to full well or greater) exhibit some persistence, but bright stars, which saturate or nearly saturate an observation, represent the most common source of image persistence.

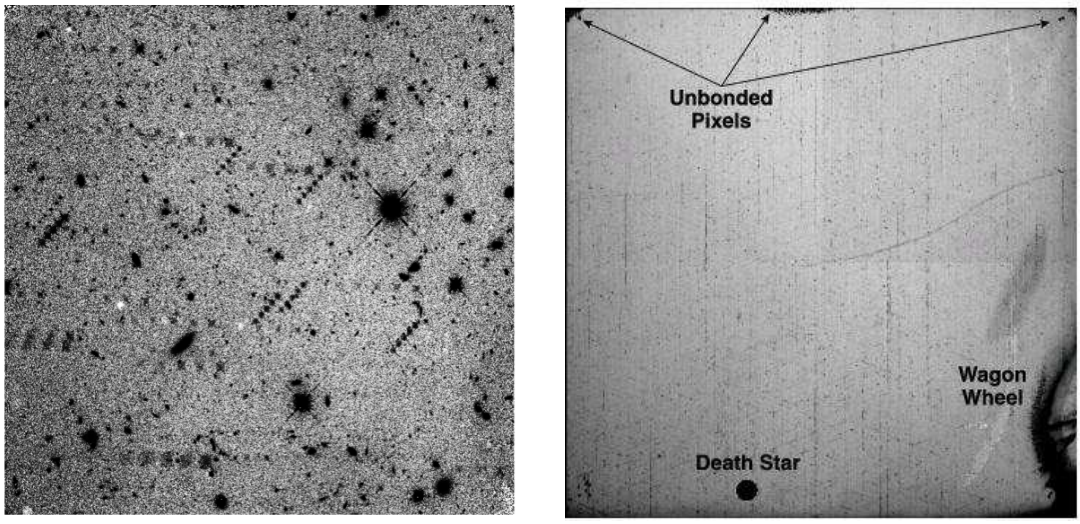


Figure 3.5: Left panel: WFC3/IR image showing the effect of persistence. The dither pattern used in these sets of observations is clearly visible since the persistence spots are aligned. Right panel: flat field image, obtained using the F140W filter, showing some of the most prominent features which affect the cosmetic of the IR detector (Bushouse, 2008). In particular the so called "Death star" and the arc-shaped structure called "Wagon Wheel" are clearly identifiable. (Images from the WFC3/IR Handbook by Dressel 2012)

Effects of Persistence on Power Spectra

Persistence is of particular concern because of the possibility of introducing an artificial coherent signal in the images, consequently leading to the identification of false candidates (Trenti et al., 2011).

To test the effect of persistence on the power spectrum we created new images starting from the original ones. In particular, we obtained mock persistence images flipping the data with respect to the x and y-axis and dimming the sources by dividing the original image by 100, 50, and 20, respectively. The choice of these values is due to the fact that persistence does not have a fixed effect but it varies with time, specifically diminishing the effects as time goes by. Adding mock persistence images to the original data and multiplying by the mask we got a new dataset mimicking the effect of persistence with results clearly verifiable and comparable to the case with no persistence.

The effect of persistence is different in the Y_{105} and J_{125} -band since the observations were not taken at the same time, so we used both the original images to build up two different models that were subsequently added to the corresponding image.

We checked both the single band power spectra (Figure 3.6) and the ratio between the Y_{105} and J_{125} -band ones (Figure 3.7) looking for a signal clearly identifiable in a well defined wavenumbers range. Ideally the effect of persistence should be restricted to an interval of wave numbers and should have a shape similar to what Tonry and Schneider (1988) found since it is mostly due to bright stars, except in the case of persistence from grism observations (as noticed in CANDELS data).

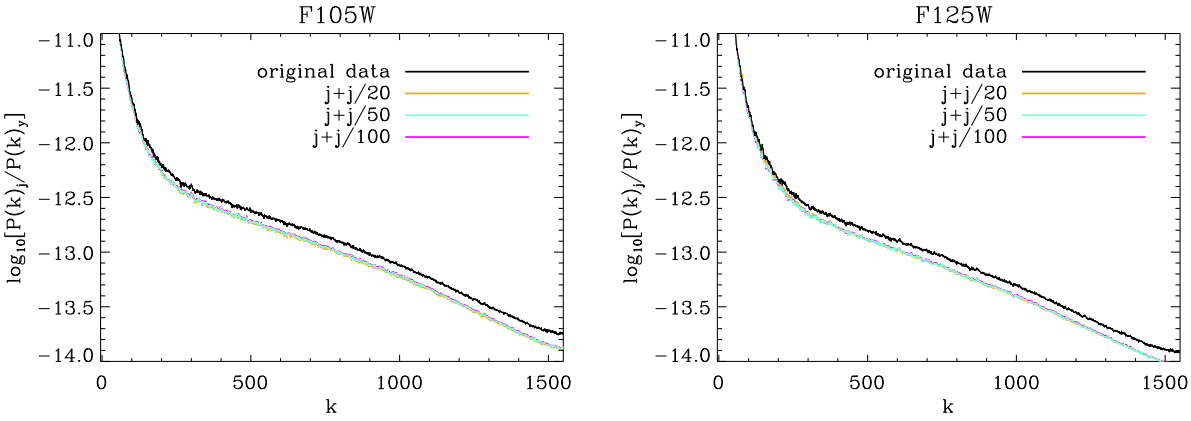


Figure 3.6: Effect of mock persistence images on the single band power spectra (Y_{105} on the left, J_{125} on the right). The different colors refer to different intensity of the effect.

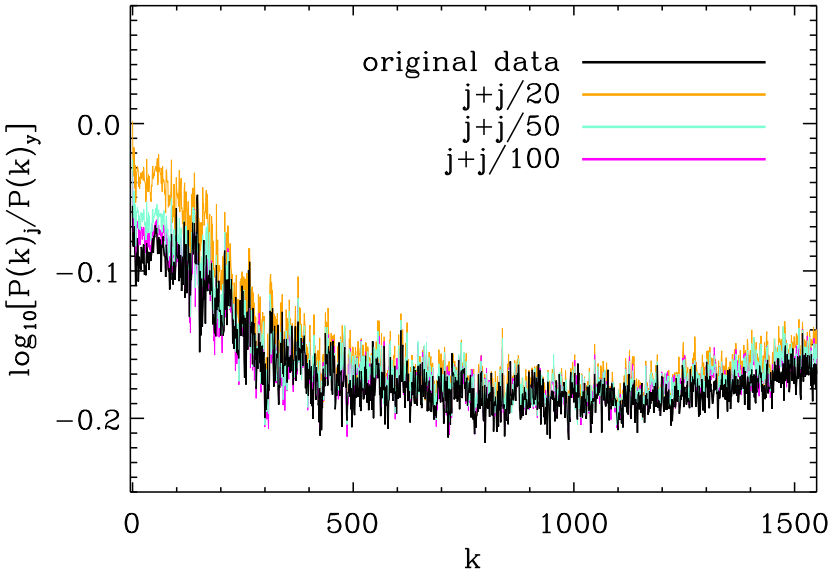


Figure 3.7: Effect of mock persistence images on the ratio between Y_{105} and J_{125} power spectra. The different colors refer to different intensity of the effect.

3.2.2 Cosmetic of the WFC3/IR Detector

Besides the effects of persistence, some of the images obtained using the IR channel are affected by blemishes which have up to 10-15% lower count rates than the surrounding areas. These features are called "blobs" and were not detected during ground testing of the camera. They appeared progressively after July 2009 and they are mostly visible in either images containing a large uniform object, or images with high background levels (Pirzkal et al., 2010). Once a blob appeared, it is detectable in all subsequent observations as long as these were deep enough (i.e. exposure time $>$ 300 s) to contain a significant

number of IR background photons. Blobs are not due to artifacts on the detector itself, but originate from particles that have been sticking to the mirror of the Channel Select Mechanism. These IR blobs can be a problem when running `MultiDrizzle` because they are strong enough to significantly affect the final products. This is the reason why it is necessary to mask manually all the regions affected by them.

Moreover, the IR detector is characterized by several large scale areas of lower than normal, or zero, quantum efficiency, which are usually irregularly shaped (Bushouse, 2008). These features are shown in the right panel of Figure 3.5. The circular region of unresponsive pixels near the bottom edge of the detector is known as the "Death Star" and has a diameter of ~ 45 pixels. The arc-shaped structure at the bottom right of the field is the so called "Wagon Wheel". This area shows typically a quantum efficiency which is from 25% to 35% below the regular level. The "Wagon Wheel" is the most color dependent region of the detector, with a variation that is on the order of $\sim 2\%$.

In our analysis of the background signal we took into account all these cosmetic-related issues and made sure they did not affect the final power spectrum. To this aim we checked carefully each image before and after applying the mask to make sure all the affected areas were properly masked.

3.2.3 Dithering

Many astronomical observations involve multiple exposures of the same field. The data are called dithered if there are small shifts introduced in the pointing of the telescope between the exposures. During the data reduction process the dithered frames are co-added so as to produce a single output frame of improved S/N.
As a general rule dithering can provide benefits to observations since it can:

- reduce the effects of pixel-to-pixel errors in the flat field or spatially varying detector sensitivity,
- remove the small scale detector defects such as hot pixels, bad columns, and charge traps from the image (integer shifts),
- allow the reconstruction of the information lost because of the spatial undersampling by pixels that are not small compared to the point spread function (sub-pixel shifts).

Obviously, dithering comes also with some tradeoffs that must be taken into account. In particular, it requires short exposures and this implies increasing the readout noise. Moreover, combining dithered data requires a special reduction procedure and more work. The variations in the intra-pixel sensitivity can, then, affect the final result. Finally, dithering leads to a smaller final field, and the cosmic ray rejection can be more difficult. For most HST observing programs the drawbacks to dithering are outweighed by the scientific benefits. However, in specific instances, such as programs with few orbits, it might be possible that the drawbacks are dominating and dithering must be avoided.

3.2.4 MultiDrizzle

The common way to combine HST images obtained using dithering is based on the Drizzle algorithm by Fruchter and Hook (2002). This is mostly done using the software called `MultiDrizzle` (Koekemoer et al., 2003).

Drizzling is a forward method unlike typical interpolation methods. All pixels of the original input images are mapped into pixels in the subsampled output image. In mapping the pixels it is necessary to take into account shifts and rotations between images, in addition to the optical distortion of the camera. The user can decide to shrink the pixels before creating the output image using the *pixfrac* parameter so as to avoid to convolve the image with the large pixel footprint of the camera. The shrunken pixels, also called drops, rain down upon the output image as shown in the left panel of Figure 3.8. The final size of the drop is, then, adjusted by the code to consider the geometric distortion introduced by the camera. The user is also allowed to specify the size of the output pixels using the parameter *scale*. The flux in each input pixel is divided up into the output pixels with weights proportional to the area of overlap between the drop and each output pixel. The parameter *kernel* permits to specify which kernel function should be used to distribute the flux onto the separate output images. The options available are:

- Square, the original classic drizzling kernel.
- Point, which forces each input pixel to contribute only to the single pixel which is the closest to the output position. This kernel is equivalent to a square one with $pixfrac \to 0$ and it is very fast.
- Gaussian, which is a circular Gaussian with $FWHM = pixfrac$, measured in input pixels.
- Turbo, which is similar to the square kernel, but with the box always with the same shape and size on the output grid and always aligned with the x and y-axis. It can increase significantly the speed of the process.
- Tophat, with a circular top hat shape of width equal to *pixfrac*. Only output pixels within $pixfrac/2$ of the output position are affected.
- Lanczos3, which is a Lanczos style kernel extending 3 pixels from the center. The Lanczos kernel is a damped, bounded form of the sinc interpolator and is very effective for resampling single images when $scale = pixfrac = 1$. It leads to less resolution loss than the other kernels, and also less correlated noise in outputs, but it is much slower. It should never be used for $pixfrac \neq 1.0$.

When the drop size is too small or the point kernel is used, it is possible that not all output pixels receive data from each of the input images. This could be dangerous especially if the number of frames combined is small. One should, therefore, choose a drop size that is small enough to avoid convolving the image with a too large input pixel footprint, and, at the same time, large enough to have signal in all the output pixels.

One of the drawbacks of the Drizzle algorithm is the fact that the output pixels in the final drizzled image are not independent of the others, causing the noise in the output image to be correlated to some degree. In particular, the noise in adjacent pixels will be correlated. This depends on splitting the power from a single input pixel between several output pixels. This effect needs to be quantified properly for estimating the statistical errors when drizzled images are analyzed using SExtractor (Bertin and Arnouts, 1996). This correlated noise implies an underestimation of the noise on larger scales in the output image. This is true for all the kernels, but one. In fact, the point kernel does not spread the

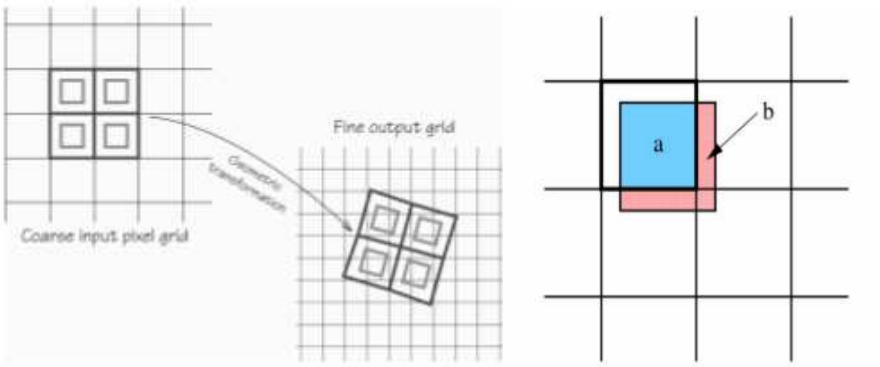


Figure 3.8: On the left panel it is shown how Drizzle maps the input pixels onto the output image. On the right panel there is a scheme of the distribution of noise from a single input pixel (a+b colored area) between neighboring pixels of the output image.

flux of one input pixel into several output pixels, but associates it only to one single output pixel, without introducing any correlated noise in the final product.

`MultiDrizzle` carries out a series of completely automated steps:

- **Static Mask.** In this step it examines all the images to identify negative bad pixels, and to include them in the data-quality array.
- **Sky Subtraction.** It subtracts the sky from each input image.
- **Driz_Separate.** It drizzles the input images onto separate, registered outputs using shifts computed from the headers.
- **Median.** It combines the separate drizzled images to create a median frame.
- **Blot.** It "blots" the median image back to each original input image. It performs the inverse operation of Drizzle, i.e. it converts an undistorted image back into the original distorted one. It is mostly used to identify the cosmic rays in the original image.
- **Driz_cr.** It creates a derivative image using each blotted image and, then, it computes the cosmic ray masks.
- **Driz_Combine.** On the basis of the cosmic ray masks previously created, it does the final drizzle combination.

The final output image created by `MultiDrizzle` is characterized by the suffix **_drz.fits* and is a multi-extension FITS file, unless specifying *built=no*, which determines the creation of separate output files. Each **_drz.fits* file contains the science image (SCI) in the first extension, the weight image (WHT) in the second, and the context image (CTX) in the third one.

The science (SCI) image is corrected for distortion and dither-combined, if applicable.

The weight (WHT) image gives the relative weight of the output pixels. It can be considered an effective exposure time map, if *final_wht_type* = *EXP*. Otherwise, setting *final_wht_type*

= *ERR* the output weight image is in units of inverse variance, calculated using the error arrays existing from the beginning in the **flt.fits* file. Finally, if *final_wht_type=IVM*, `MultiDrizzle` looks for inverse variance files provided by the user. In this case the user have to create an input file, containing two file names per line, i.e., the name of the **flt.fits* file and of the corresponding **ivm.fits* file.

The context (CTX) image encodes information about which input image contributes to a specific output pixel.

3.3 A Customized Version of the HUDF09 Dataset

Contrary to the ACS data we analyzed in the previous Chapter, nowadays only version 1.0 of the WFC3/IR data exists, and no more accurately reduced versions are available. Since a new data reduction of the entire dataset was not feasible, we tried to improve the quality of the images running `MultiDrizzle` by ourselves. In particular we aimed to get images with the same scale as the UDF05 dataset used in the previous Chapter and to reduce as much as possible the correlated noise using a tailored choice of kernel and drop size.

First of all we downloaded from the archive the **flt.fits* and **ivm.fits* files that were released to the community. The **flt.fits* files were the output from `WFC3RED`, a software which corrects for instrumental effects and creates calibrated products. The **flt.fits* files were corrected for bias, dark current and flat field, but not for distortion, and still contained cosmic rays. Together with the **ivm.fits* files, they formed the input for `MultiDrizzle`. An input file was needed, with each line containing the name of the **flt.fits* image and of the associated **ivm.fits* file.

To customize the final image we had to specify the scale of the output in units of arcsec pixel^{-1}, the kernel to be used, and the size of the drops. To be consistent with the ACS data, the scale was set to be 0.09 arcsec pixel-1. Regarding the kernel to be used, unfortunately the number of **flt.fits* files to be combined was so small to prevent the use of the point kernel. The main risk in using it even with very few **flt.fits* files is to create an image with holes or characterized by areas with a lower signal-to-noise. Since the power spectrum analysis requires uniform data, it was necessary to come to a compromise using the square kernel associated with a size of the drop not too small and, at the same time, not too big. If not, it would cause the same troubles of the point kernel, if too small, and, introduce noise correlation, if too big. After a series of tests and the comparison between the different outputs, the best choice turned out to be setting *pix_frac* equal to 0.5. Then, the coordinates RA and DEC of the center of the field were specified, as well as the number of pixels in the x and y-axis, which determine the dimension of output image.

As we mentioned above, the infrared data are affected by persistence (see Section 3.2.1), and they show regions characterized by low quantum efficiency and IR blobs. Setting *driz_sep_bits* =0 it was possible to exclude all the flagged pixels and, consequently, to mask out the IR blobs.

The set up of all the parameters used by `MultiDrizzle` for the three different fields took a lot of time. In fact, all the output images obtained using different parameters had to be checked and compared to the others so as to find out the best output possible needed to perform the following steps.

3.3.1 Median Subtraction

Unfortunately, the results obtained using the public release of **flt.fits* and **ivm.fits* were not as good as expected. In particular we noticed gradients clearly identifiable in both the Y_{105} and J_{125} -band final **_drz.fits* frames and these gradients were affecting the resulting power spectra. Therefore, the ratio between the power spectra turned out to be so noisy to not permit the identification of any signal.

To remove these artifacts we performed our specific data reduction. In particular we obtained sky flat fields directly from the images instead of using those provided by the pipeline, as done for the public release. As a general rule a sky flat field should be more accurate and reduce possible patterns in the final stack. Moreover, we obtained and subtract a median frame from all the images in each band even though this could reduce the background fluctuations we were looking for.

To clean as much as possible the final data we combined the **flt.fits* files median subtracted. Therefore, it was necessary to create the median stack of the **flt.fits* as they were, without any shift. The underlying idea is that any uncorrected detector artifact is located at the same exact (x,y) position in each **flt.fits* file, so the median stack should capture them. Subtracting the median frame from each individual **flt.fits* file, the artifacts are removed, but, at the same time, the background signal is corrupted. In this way the **flt.fits* files are characterized by the best cleaning level possible without re-performing the data reduction from scratch.

Since the **flt.fits* frames are characterized by the presence of objects and by the dither pattern (i.e. objects shifting in x/y on the **flt.fits* images), there will be some regions near detected objects where the median image will be noisier because of less non-zero data in the stack. In any case these regions did not affect our results since we carefully masked all the sources and the surrounding halos. Last but not least, we did not include in the median stack pixels whose value was zero, otherwise the median would have been biased.

We derived our median subtracted dataset running `MultiDrizzle` using the same parameter set of the one provided by the HUDF09 team and available on the archive. Then, we applied the power spectrum technique to images in the Y_{105} and J_{125} -band and plotted the ratio between the power spectra (Figure 3.9). Clearly the median subtraction gets rid of most of the spurious signals affecting the data. In detail, at high wavenumbers the public release does not show the expected flat trend due to white noise, but our dataset does.

In the following we will use our own version of the Y_{105} and J_{125} -band images rather than the public dataset to avoid as much as possible contamination and spurious effects. It should be noted that the HUDF09 dataset is characterized by a scale of 0.06 arcsec pixel^{-1} and so was the one we used for the comparison above, but for the following analysis we chose the same scale of the UDF05 NICP12 images used in the previous Chapter, i.e. 0.09 arcsec pixel^{-1}.

3.4 Data Analysis: Looking for Faint Galaxies at $z \sim 8$

The data we focused on, obtained running `MultiDrizzle`, are characterized by a scale of 90 mas pixel^{-1} and a size of 2100×2100 pixel^2. We used the weight maps associated to the data to exclude the regions where only few exposures have been combined with `MultiDrizzle` and to select the central, and the deepest, area of the final frame.

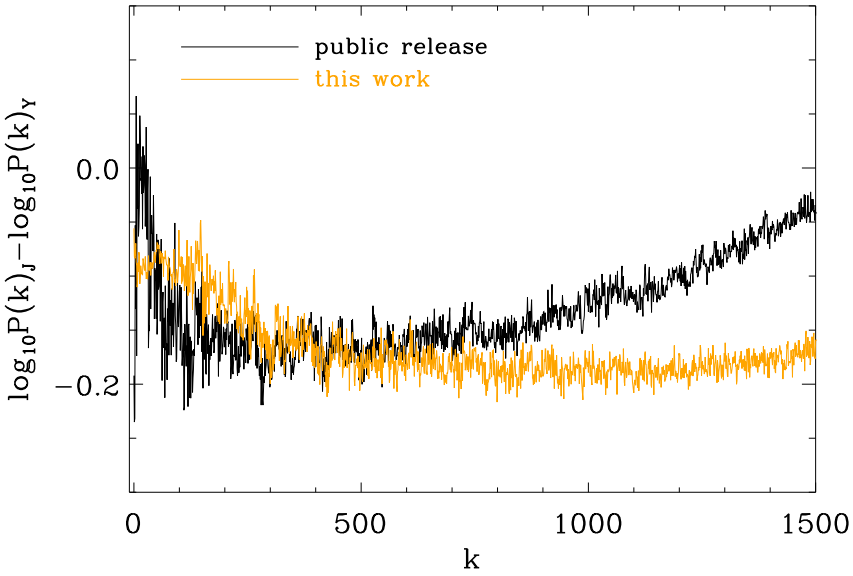


Figure 3.9: Ratio between the J_{125} and Y_{105} -band power spectra obtained for the public released dataset (in black) and the one we created (in orange) for the HUDF09 main field. Thanks to the median subtraction most of the spurious effects are gone and we can see the trend typical of white noise at high wavenumbers.

3.4.1 Identification of the Sources and Creation of the Mask

As we did for the i_{775} and z_{850} -band images, the first step in the analysis of the WFC3/IR data was the identification of all the detected galaxies, that was required to consequently mask the light coming from them, as well as the contribution from their bright halos, residuals of cosmic rays, and bad pixels. To this aim we used the version 2.5.0 of the SExtractor photometry package. The input parameters used were chosen to maximize the number of detections of galaxies as well as to minimize the detection of spurious sources. The detection threshold (DETECT_MINAREA) was set to be a minimum of 9 connected pixels with an intensity of 0.7σ (DETECT_THRESH) above the background (Bradley et al., 2012).

As discussed in the previous Chapter, since our goal was to analyze the contribution from galaxies which are too faint to be individually detected, but that produce a relevant overall light contribution, we had to mask all the reliably detected galaxies and to focus on the total contribution of the fainter ones. To this aim, out of the output list of identified sources we considered as real detections, and then masked, only those galaxies characterized by $S/N > 5$ (Bouwens et al., 2010).

In Figures 3.10 and 3.11 we plotted the histograms of the distribution of the total magnitude, the effective radius, the ellipticity, and the position angle (PA) respectively, derived for all the reliable galaxies detected in the HUDF09-main field (shaded histograms) and in the HUDF09-1 field (gray solid histograms). We compared the distribution of the effective radius and the ellipticity obtained for the NICP12 field and the HUDF09-1 one, which are mostly overlapped. Even though they were imaged using two different instruments and, consequently, a different filter set, we did not notice any relevant difference. The dis-

tributions are consistent with each other, and this implies that the characteristics of the detected high-z galaxies are similar in the optical and near-IR. This is consistent with the findings by Oesch et al. (2010b) who studied the morphology of $z \sim 7-8$ galaxies in detail. According to them on average $z \sim 7-8$ galaxies are very symmetric and compact and, at fixed luminosities, sizes evolve only a little from $z \sim 4$ to $z \sim 8$.

The following steps were followed for the analysis of both the HUDF09 main field and the HUDF09-1 field. Starting from the segmentation maps, obtained as outputs from SExtractor, we created a mask for each band to reject all the detected sources existing in the field. Then, to mask also the bright halos surrounding the detected stars and galaxies, we convolved both the Y_{105} and J_{125} -band masks with a Gaussian filter with $\mathrm{FWHM} = 30 \text{ pixel} = 2.''7$ to enlarge the masked area. The final mask was obtained merging the two single masks and, then, convolving again to reduce as much as possible the probability of spurious signals contaminating the result. The masked area corresponds to 26% and 31% of the image for the HUDF09-1 field and the HUDF09 main field, respectively.

3.4.2 Power Spectra

As we described in Section 2.2.2, the power spectra of the masked images in the Y_{105} and J_{125} -band were obtained using the IDL fast Fourier Transform. Figures 3.12 and 3.13 show the power spectra, in units of $\mathrm{nW}^2/\mathrm{m}^4/\mathrm{sr}$, derived from both the Y_{105} and J_{125} -band images of the HUDF09-1 field and the HUDF09 main field. Then, to highlight a possible contribution from undetected galaxies at $z \sim 8$ we plotted the ratio between the J_{125} and Y_{105} -band power spectra (Figures 3.14 and 3.15). As can be noticed from these plots, we were not able to detect any signal from faint $z \sim 8$ galaxies. The reason why no light contribution from these galaxy population was found probably lies in the data reduction. The entire WFC3/IR dataset was aligned to the ACS one using a routine called `superalign` that is part of the reduction package `WFC3RED`. At the moment we have no guesses on the possible effect of `superalign` on the power spectra, in particular whether or not it introduces some kind of correlated noise. Moreover, as discussed in Section 2.2.4 the cleaning level affects significantly the results of our technique. In particular, when the data reduction process is not tailored for the background signal analysis the images are often affected by residuals. Spurious signals can modify or dominate the resulting power spectrum. Moreover, the median subtraction step could possibly have removed, either partially or in total, fluctuations lying in the background and responsible for a substantial bump in the power spectrum.

We are waiting for the v2.0 of the WFC3/IR data, which will be characterized by a better cleaning level to apply again the power spectrum technique and to look for the contribution of $z \sim 8$ faint galaxies without any contamination. The contribution to cosmic reionization from this population is crucial to constrain the duration of the process and when the bulk of ionizing photons was produced.

Unfortunately, till now the HUDF09 set of images do not allow any claim on the role of galaxies at $z \sim 8$ since no signal excess can be disentangled from the noise.

As shown in Trenti et al. (2012) if a correlation between galaxy luminosity and the mass of the dark matter halo (Vale and Ostriker, 2004; Cooray, 2005; Trenti et al., 2010) is assumed, a natural prediction of dark-matter clustering is that at $z > 6$ the brightest galaxies should be surrounded by an overdensity of fainter ones at similar redshift (Muñoz and Loeb, 2008) and that these proto-clusters are as rare as bright high-z galaxies. Assuming that the same is valid even at fainter magnitudes the fact we are not detecting any signal

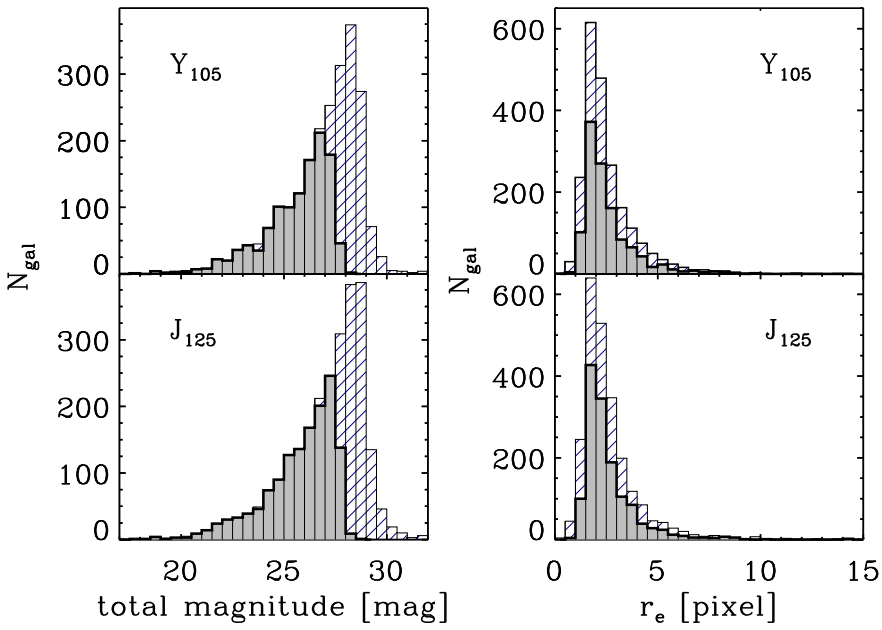


Figure 3.10: Distribution of the total magnitude (left panels) and the effective radius $\mathrm{r_e}$ (right panels) of the detected galaxies in the Y_{105} (top panels) and J_{125} -band (bottom panels) images. The gray solid histograms refer to the galaxies detected in the HUDF09-01 field, the shaded ones to the galaxies detected in the HUDF09 main field.

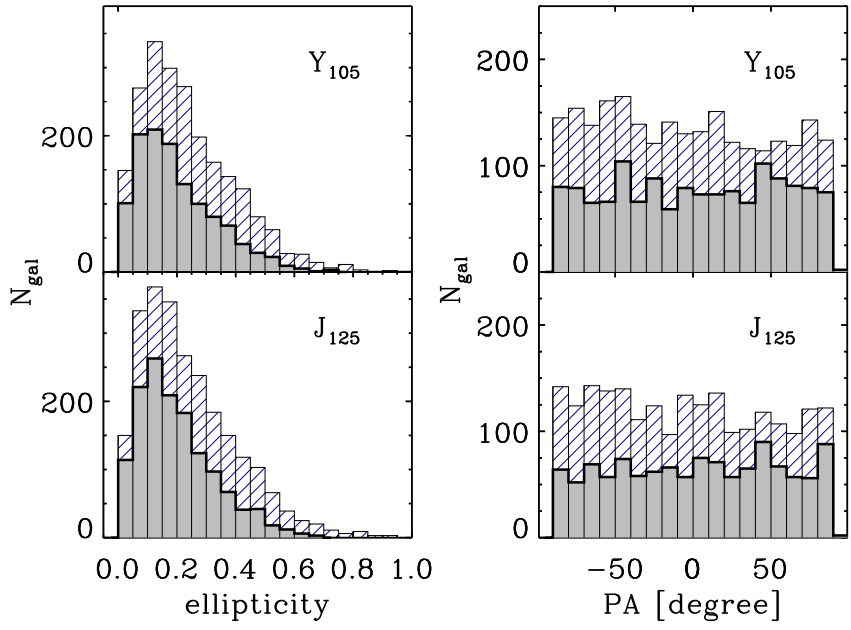


Figure 3.11: Distribution of the values of the ellipticity (left panels) and the position angle PA (right panels) of the galaxies detected in the Y_{105} (top panels) and J_{125} -band (bottom panels) images. The gray solid histograms refer to the galaxies detected in the HUDF09-01 field, the shaded ones to the galaxies detected in the HUDF09 main field.

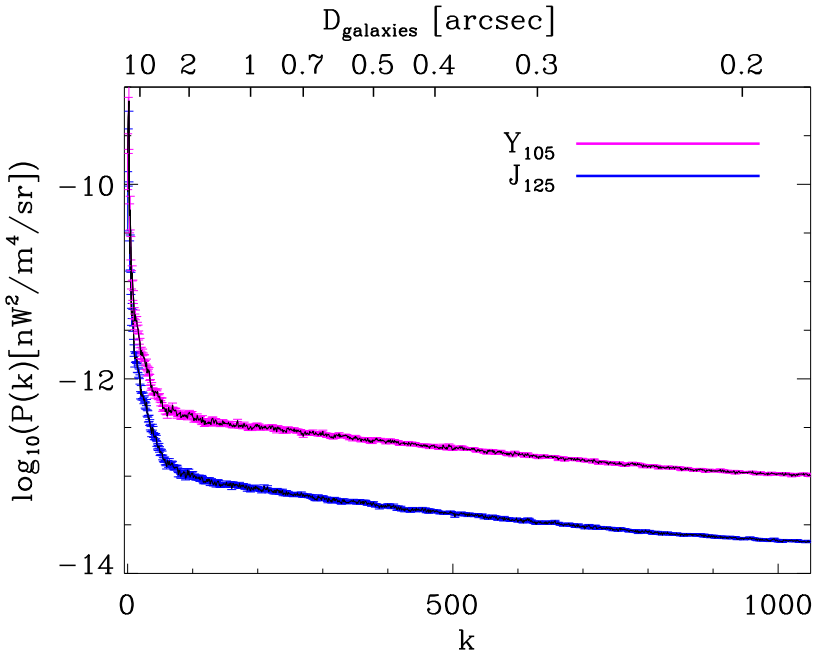


Figure 3.12: Power spectra with error bars of the background signal of the Y_{105} (top line with magenta error bars) and J_{125} -band images (bottom line with blue error bars) obtained for the HUDF09-01 field.

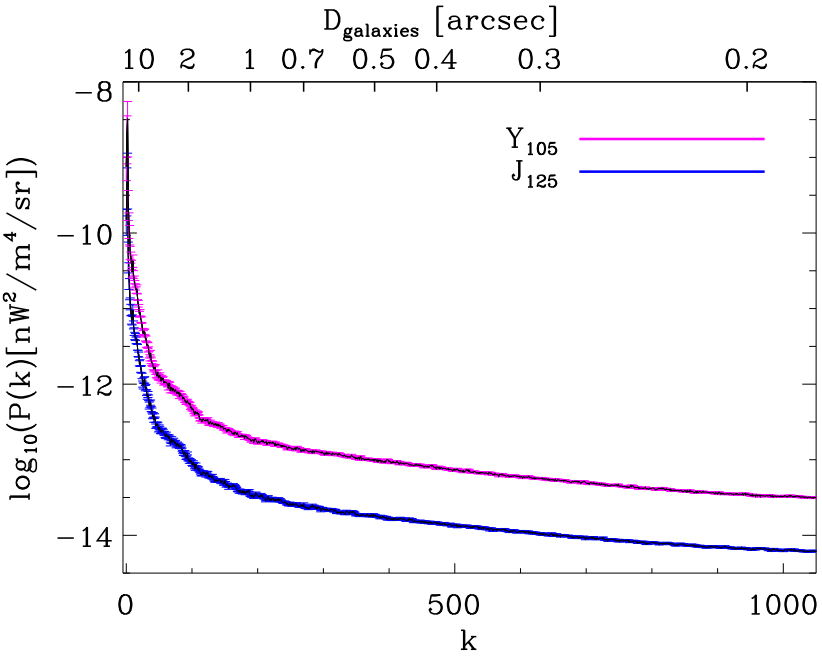


Figure 3.13: Power spectra with error bars of the background signal of the Y_{105} (top line with magenta error bars) and J_{125} -band images (bottom line with blue error bars) obtained for the HUDF09 main field.

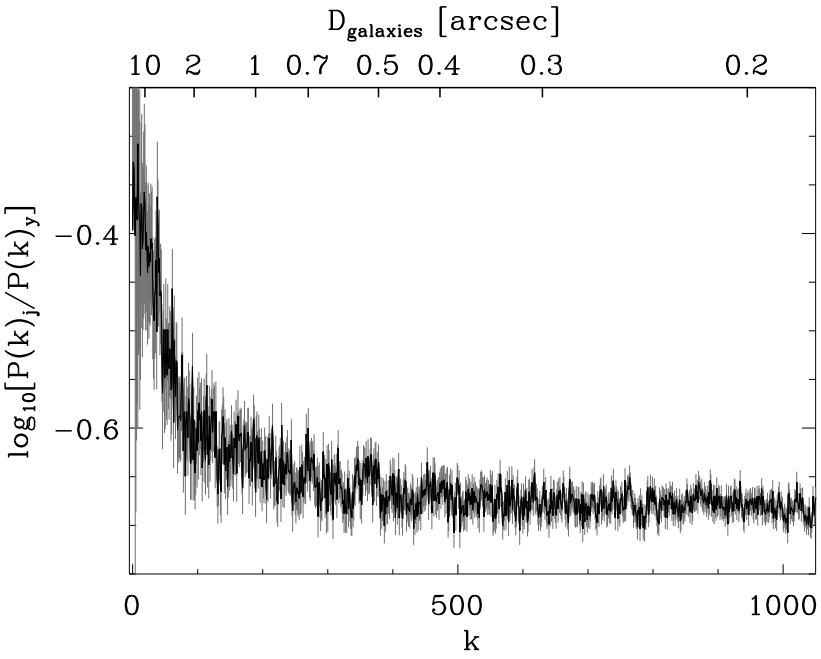


Figure 3.14: Ratio between the Y_{105} and J_{125} -band power spectra with gray error bars for the HUDF09-01 field.

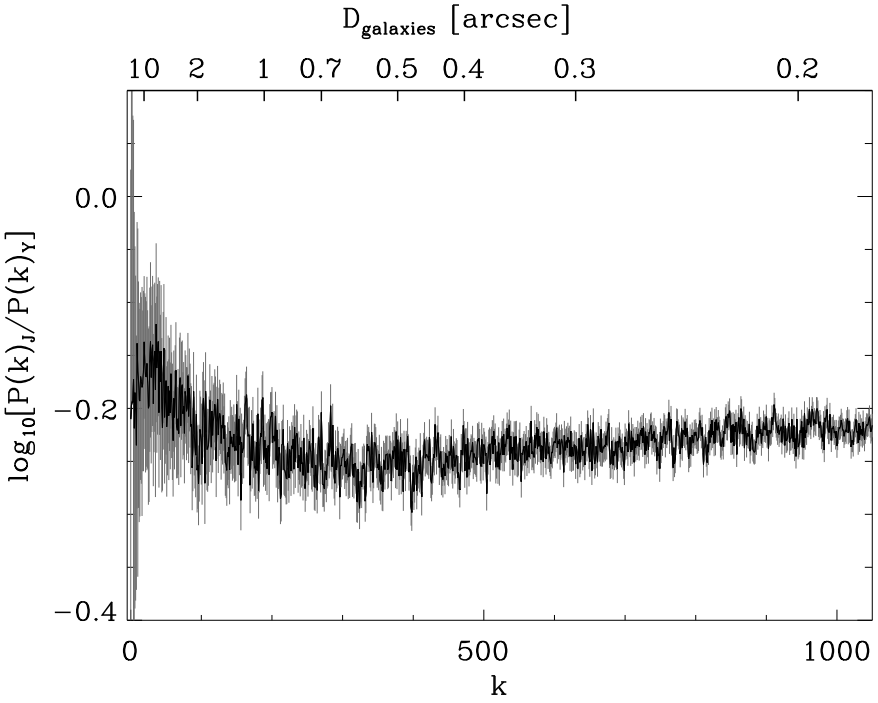


Figure 3.15: Ratio between the Y_{105} and J_{125} -band power spectra with gray error bars for the HUDF09 main field.

from $z \sim 8$ galaxies could be explained simply taking into account the underdensity of $z \sim 8$ galaxies in the field. This issue is one of the reason that led Trenti et al. (2011) to design the Brightest-of-Reionizing Galaxies (BoRG) survey that imaged the sky on random lines of sight at high Galactic latitudes in four filters. Of course this observation is not ruling out the need for an accurate data reduction and a strong attention to systematics.

3.4.3 Effect of the Source Masking

According to Arendt et al. (2010) it is necessary to derive the effect of source masking on the power spectra of background signal. To this aim we built simulated frames, i.e. a frame containing just random noise, one with noise and a flat field residual, and another one with noise and simulated galaxies. Then, we compared the power spectra obtained from these frames with those derived from the same frames multiplied by the mask (Figures 3.16, 3.17, and 3.18). For convenience we used the same mask derived from the HUDF09-01 field. As done by Arendt et al. (2010) in the bottom panels we plotted also the power spectra of the masked imaged normalized by those obtained from the unmasked simulations to highlight the differences.

When masking a frame characterized by a smooth signal, such as the case of random noise or noise and flat field, the normalized plots show no structures. The masked power spectrum and the unmasked one are basically the same, just shifted because of the different number of non rejected pixels. Adding mock galaxies changes the result and we can notice a shift of power from the higher wave numbers to the lower ones. In any case, as can be inferred from the lower panel of Figure 3.18, this shift determines only a difference of 2%.

Since we are dealing with a background signal which is supposed to be due to a large number of faint galaxies and, consequently, much closer to the case of a mock frame with noise and flat field rather than to the one with bright isolated galaxies, the effect of the mask can be ignored. Moreover, the possible shift of power towards higher wave numbers seems to not affect the range of intermediate wave numbers k we are interested in, being limited to the highest and smallest ones. Even if it affects the range we are looking to, the effect would be simply an underestimation of the signal from the faint galaxy population and, consequently, of the number of undetected sources.

3.4.4 Simulations

Even though the current quality of the data did not allow us to derive a clear detection of the signal from the undetected $z \sim 8$ galaxies, we were able to derive an upper limit for the number of faint Y_{105} dropouts galaxies in the fields of the HUDF09 project.

To this aim we built a simulation which was able to reproduce the Y_{105} and J_{125} frames, and permitted us the infer the number of faint objects per magnitude bin, as we did for the i_{775} and z_{850} band images in the previous chapter. The FWHM of the HUDF09 data is $\sim 0.16''$ (Bouwens et al., 2011).

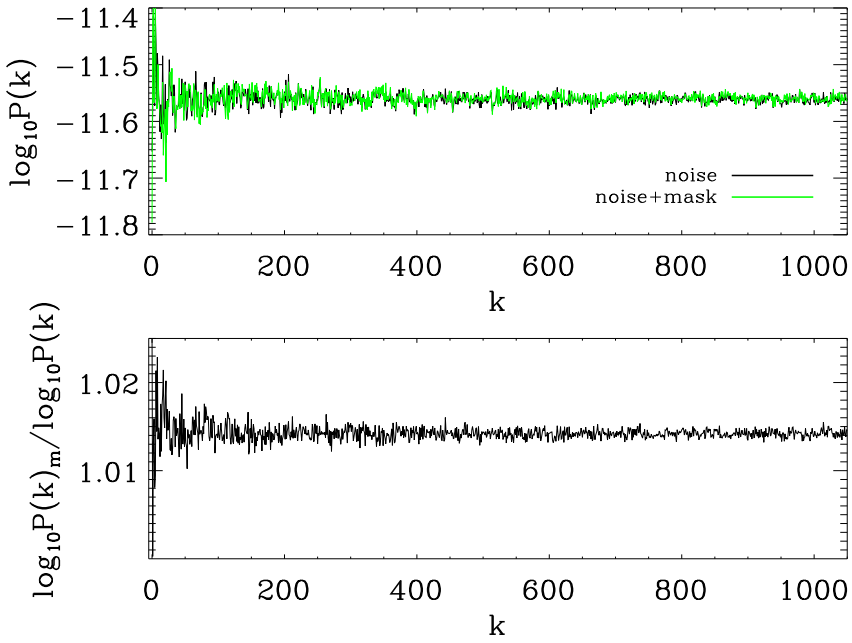


Figure 3.16: Comparison between the power spectra obtained from a simulated image containing only noise with (green line) and without (black line) applying the mask derived from the data. The lower panel reveals more details by normalizing the masked power spectrum by the unmasked one.

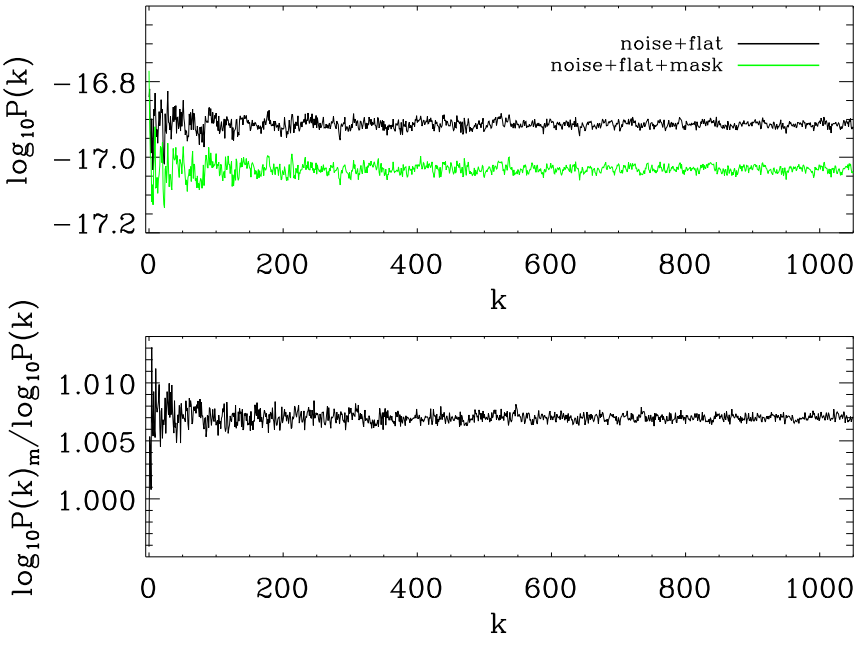


Figure 3.17: Comparison between the power spectra obtained from a simulated image containing noise and a flat field with (green line) and without (black line) applying the mask derived from the data. The lower panel reveals more details by normalizing the masked power spectrum by the unmasked one.

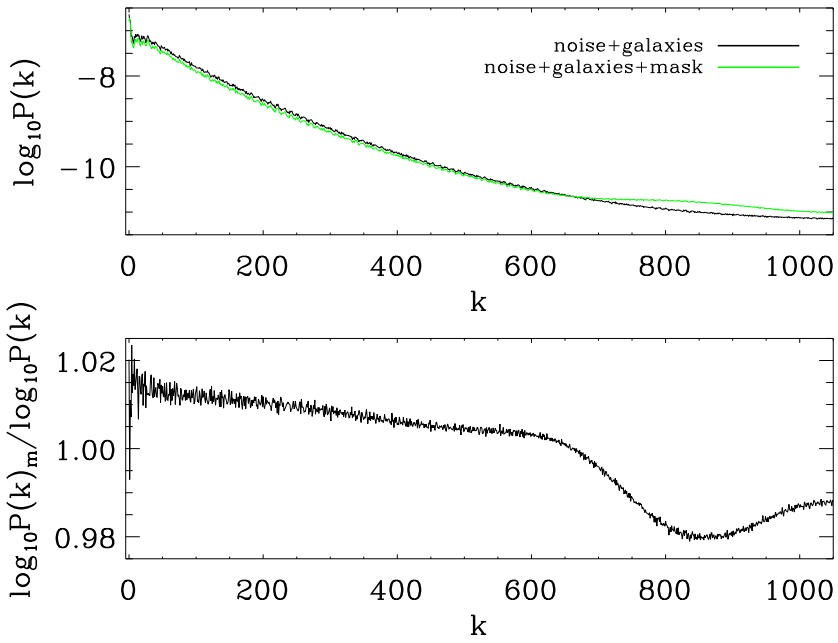


Figure 3.18: Comparison between the power spectra obtained from a simulated image containing noise and 15 simulated galaxies with (green line) and without (black line) applying the mask derived from the data(black line). The lower panel reveals more details by normalizing the masked power spectrum by the unmasked one.

3.5 Data Analysis: Looking for Faint Galaxies at $z \sim 7$

Taking into account the drop in the star formation rate density from $z \sim 6$ to $z \sim 7$ and beyond suggested by Bouwens et al. (2004b) and confirmed by Oesch et al. (2009), it is expected that the contribution in terms of photoinizing photons is more relevant at $z \sim 7$ than at $z \sim 8$. In this sense the analysis we carried out comparing z_{850} and Y_{105} -band data turned out to be more substantial than the one based on just WFC3/IR images.

Looking for the contribution to reionization from faint galaxies at $z \sim 7$ we had to deal with the contemporaneous analysis of the z_{850} -band ACS image and the Y_{105} -band WFC3/IR one To compare the power spectra obtained from different detectors, first of all it we trimmed the z_{850} -band image to the same dimension of the Y_{105} -band one because the fields observed with WFC3/IR as part of the HUDF09 project are smaller than the ones observed with ACS. As a matter of fact, it is not trivial to compare the results obtained from images with the same pointing, the same pixel scale (0.09 arcsec pixel^{-1}), but a different size (3500 pixels versus 2100 pixels per side in the z_{850} and Y_{105} -band images, respectively). We used sources easily recognizable in both the images to align the 2 images, then we trimmed the exceeding regions of the z_{850} -band image.

Of course there are several issues associated to this analysis since the two different detectors are characterized by systematics that can not be erased simply considering the ratio of the power spectra. The understanding of all WFC3/IR related problems is not as good as the ACS related one and, moreover, the IR detectors are less clean than the correspondent

optical ones.

For the analysis we present in this Section we made use of data imaging the sky area covered by the parallel field NICP12. This choice is due to the better quality of the optical dataset for this field among the 3 imaged by ACS. Our ACS dataset has a 0.09 arcsec pixel^{-1} scale and was obtained using the point kernel. We tried to match the same characteristics for the WFC3/IR image, in particular we set `MultiDrizzle` to create the final image with the same pixel scale. Regarding the type of kernel, we choose to use both the square one with a drop_size=0.5 and the point one.
Once we had both the z_{850} and Y_{105} -band images characterized by the same scale (0.09 arcsec pixel^{-1}) and dimension (2100 $\times$ 2100 pixels), we run SExtractor several times tuning the configuration until we got the best parameters to be used for the detection in both the images.
We selected a total of 32 deblending subthresholds (DEBLEND_NTHRESH), a contrast parameter of 0.03 (DEBLEND_MINCONT), and a Full Width Half Maximum (SEEING_FWHM) of $0''.09$. We set the WEIGHT_TYPE to MAP_RMS and WEIGHT_IMAGE to be the rms map associated to each image. We changed the threshold for detection (DETECT_THRESH) and analysis (ANALYSIS_THRESH), as well as the minimum number of connected pixels (DETECT_MINAREA). In Figure 3.19 we compared the single band power spectra, as well as the ratio between them, obtained assuming three different setups:
DETECT_MINAREA = 2, DETECT_THRESH = 0.1,
DETECT_MINAREA = 5, DETECT_THRESH = 0.7,
DETECT_MINAREA = 35, DETECT_THRESH = 3.
The best choice turned out to ask the source to satisfy the 0.7 detection threshold within an area of at least 5 contiguous pixels. The choice of parameters in SExtractor configuration file directly reflects on the mask we are applying before focusing on the background fluctuations. As discussed in Chapter 1 and described in Viero et al. (2013) the mask plays a relevant role when studying the signal hidden in the background. In particular, an aggressive mask can make the signal from faint galaxies to low to put a signature on the power spectrum. On the other hand, relaxed constraints on the mask can lead to non-masked halos dominating the flux.

In any case, even using different combination of parameters for the detection and subsequent mask we found that the different systematics affecting the two detectors are ruling out any chance to highlight an excess of signal ascribable to galaxies below the detection limit (Figure 3.20). Even the ratio between power spectra derived from images obtained both with the point kernel could not show any clear signature of flux from the fain galaxies population (Figure 3.21). In this case the systematics were reduced since the correlated noise should be the lowest possible, but the use of the point kernel was risky since the number of frames combined to get the Y_{105} -band image was not such to guarantee a perfect outcome.

3.6 Summary and Conclusions

In this Chapter we presented the application of the power spectrum technique, introduced in Chapter 2 to the deep near-IR dataset obtained with WFC3/IR as part of the HUDF09 project. Our analysis was a natural follow up of the work presented in the previous chapter with the aim of getting constraints on the faint-end of the luminosity function at $z \sim 7$ and

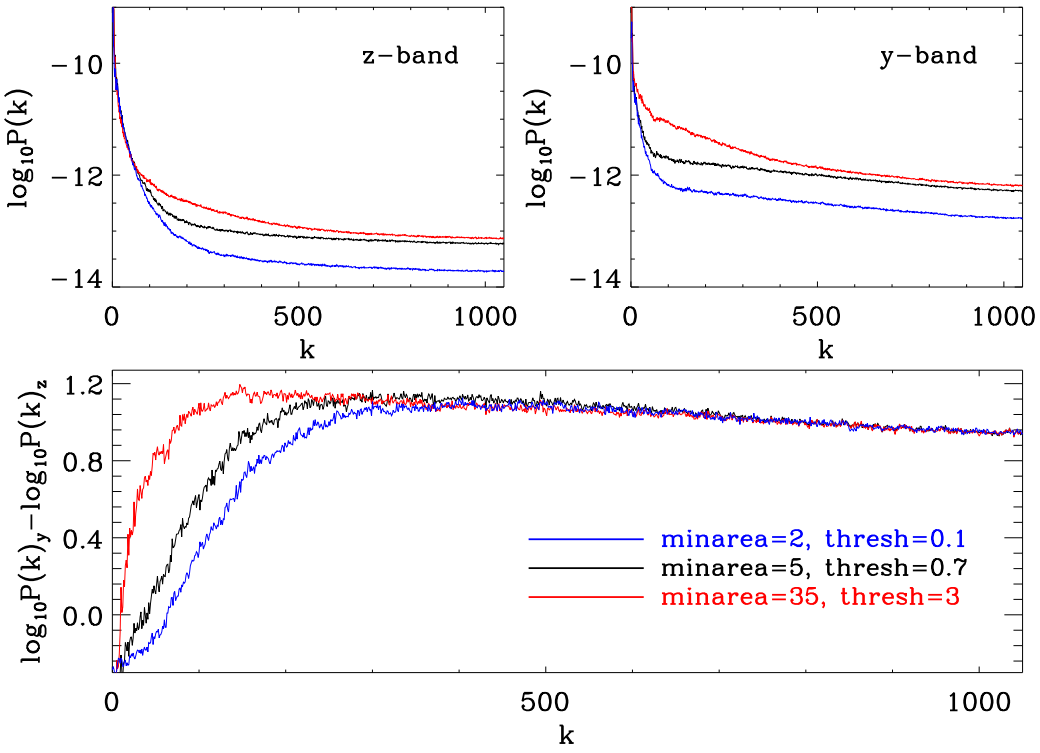


Figure 3.19: Comparison of the effect of the mask on the power spectra in the z_{850} (top right panel) and Y_{105} -band (top left panel), as well as on the ratio between the two power spectra (bottom panel). The color code refers to different choices for the detection threshold and the minimum area used to identify sources to be masked. An aggressive mask (in blue) cuts down the background signal, while a poor masking (in red) can introduce contaminant signal.

$z \sim 8$.

We compared the power spectrum of the background signal from the Y_{105} and J_{125} -band data looking for the light contribution from undetected galaxies at $z \sim 8$. We used both the released version of the data and our own, tailored to get rid of spurious effects and residuals. In neither one of the cases we were able to disentangle the feature due to the flux produced by faint galaxies. A better data reduction, possibly aimed at getting a smooth background with high quality and no prominent residuals, is definitely required to purse our analysis.

Later on, we compared the z_{850} and Y_{105} -band ones taking into account the systematics encountered when dealing with different detectors. Even matching all the characteristics of the two images we were not able to unveil any contribution from faint galaxies. One of the bigger issues is that different detectors have very different systematics and, moreover, all ACS-related issues are more well-known than the WFC3/IR ones since ACS has been taking data for a much longer period of time.

We can conclude asserting that the study of the photoinizing flux from faint galaxies at $z \sim 7-8$ is essential to understand how reionization unfold. Unfortunately, since the quality of the currently available deep images does not allow a proper study of background

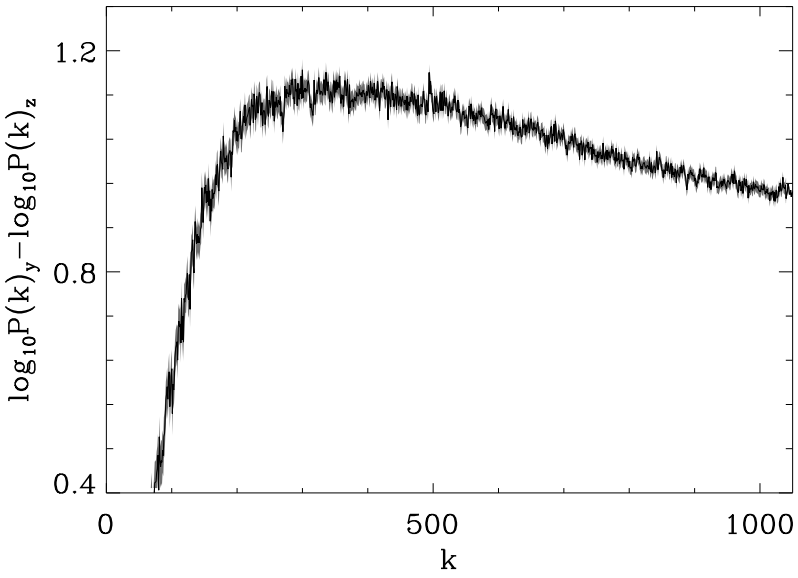


Figure 3.20: Ratio between the power spectra obtained from the z_{850} and Y_{105} -band images of the HUDF09-1 field using the dataset from the archive. The error bars are plotted in gray.

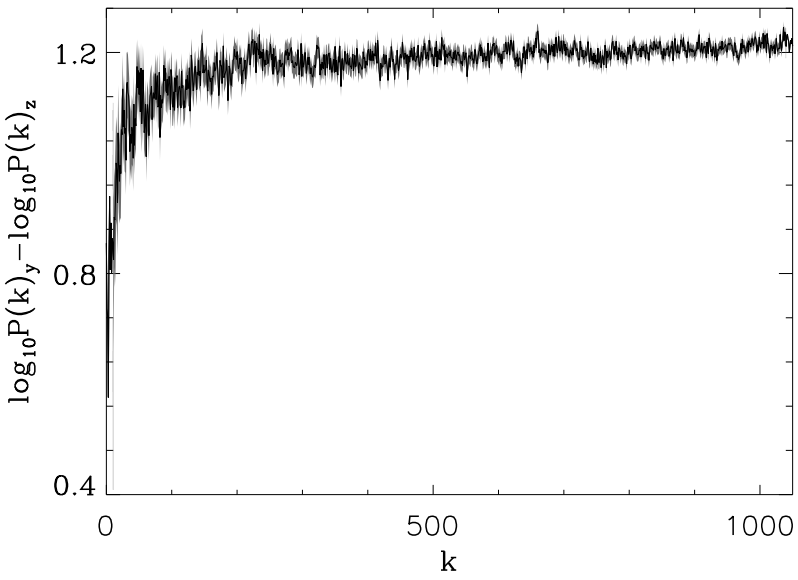


Figure 3.21: Ratio between the power spectra obtained from the z_{850} and Y_{105} -band images of the HUDF09-1 field we created running `MultiDrizzle` and setting the point kernel. The error bars are plotted in Gray.

fluctuations, we should either wait for the v2.0 of the dataset, characterized by a more careful data reduction, and try to indirectly constraint the faint-end slope of the luminosity function, or we should wait for JWST to image the early phases of the Universe.

Chapter 4

HUDF Main Field Data Reduction

In collaboration with Dr. Massimo Stiavelli, Dr. Larry Bradley, and Ray Lucas

In this Chapter we will present how we reprocessed the ACS data for the HUDF main field using improved data reduction algorithms and adding new data obtained during the HUDF09 campaign and other campaigns that happened to observe that particular area in the sky. We obtained an improved version of the XDF dataset with the aim to extend the power spectrum analysis to this field and to verify our findings on the faint-end slope of the luminosity function at $z \sim 6$ (see Chapter 2). We are still working on the dataset the create really the cleanest version ever of the deepest image of the Universe, in particular we are now modelling the electronic ghosts.
The new version is intended to be used for other investigations as well. In particular our group is carrying out a study on the surface brightness of dropout galaxies comparing the results obtained from our new data-set to those derived from GOODS data (for more details see Chapter 5).

4.1 Analysis of the XDF data

With the aim of getting an independent constraint on the faint part of the luminosity function at $z \sim 6$ the analysis of the HUDF main field was mandatory.

Recently Illingworth et al. (2013) made public the entire dataset they built up combining 10 years (from 2002 to 2012) of optical and infrared data obtained with both ACS and WFC3[1]. Images from 19 different HST programs were combined, specifically 460 and 700 frames to obtain the i_{775} and z_{850} stacked images with a median exposure time of a 377.8 and 421.6 ks, respectively.

Summary of the Data Reduction

Illingworth et al. (2013) downloaded the *flt.fits* files from the Mikulski Archive for Space Telescopes archive (hereafter MAST archive). The default data reduction performed by `calacs` (bias correction, dark subtraction, flat-field correction, and cosmic ray rejection) was assumed to be good and not improved, but all images were visually inspected looking

[1] http://archive.stsci.edu/prepds/xdf/

for any significant flaw. They rejected those characterized by not correctable issues and, in the meanwhile, identified and flagged in the DQ array satellite trails and ghosts. The CTE correction was applied as well to data taken after HST SM4.
The data-set was, then, reduced using the pipeline called APSIS (ACS pipeline science investigation software, Blakeslee et al. 2003), which is written in Python and basically performs a drizzle-blot-drizzle cycle like what MultiDrizzle does.
Finally, all the images in a single band were stacked in order to get a supermedian frame, which was subsequently subtracted from each single exposure frame to enhance the pixel-by-pixel S/N and to correct for any residual imperfections in flat field or dark frames.
The final *flt.fits* files were drizzled to 60 (ACS and WFC3 frames) and 30 mas (ACS frames) pixel^{-1} scale.

Power Spectrum Analysis

ACS dataset

Out of the two pixel scales we chose the 30 mas pixel^{-1} one because it is more suitable for ACS according to the smaller native pixels and the better PSF at shorter wavelengths. As we fully described in Chapter 2, to constrain the LF of faint i-dropout galaxies it is necessary to apply the power spectrum technique to i_{775} and z_{850} -band data to relay on the Lyman-Break technique.

Following exactly the same procedure as for the NICP12 field, we identified and masked all the sources characterized by $S/N > 4.5$ and looked for surface brightness fluctuations applying the fast Fourier transform to get the power spectrum. The power spectra we obtained for each band are plotted in Figure 4.1 and show spurious features either at low k and at intermediate ones. The ratio between the power spectra, as plotted in Figure 4.2, does not allow any detection of an overall background light coming from faint objects. In principle, this result was totally predictable because, according to the description of the data reduction in Illingworth et al. (2013), the background has been distorted during the processing operations. In fact, to remove any excess background emission on individual images a super-median frame was created combining all the data in each filter and subtracted from each single image modifying the background signal, in particular deleting traces of any possible surface brightness fluctuation.

Therefore, to pursue the search for the overall signal from faint galaxies a different data reduction, not altering the signal itself but, at the same time, taking into account all the different issues (such as herringbone effect, electronic ghosts, CTE correction, etc), was needed.

WFC3/IR Dataset

The analysis performed on the HUDF09 WFC3/IR dataset for both the main field and the HUDF09-1 parallel field did not allow us to derive any result to constrain the total light contribution from galaxies at $z \sim 8$ (Sec. 3.4). One of the main drawbacks in using the HUDF09 data were the spurious signals affecting the background. For this reason we created a median frame and subtracted it from each single image. Unfortunately this procedure basically cancelled the fluctuations existing in the background and, consequently,

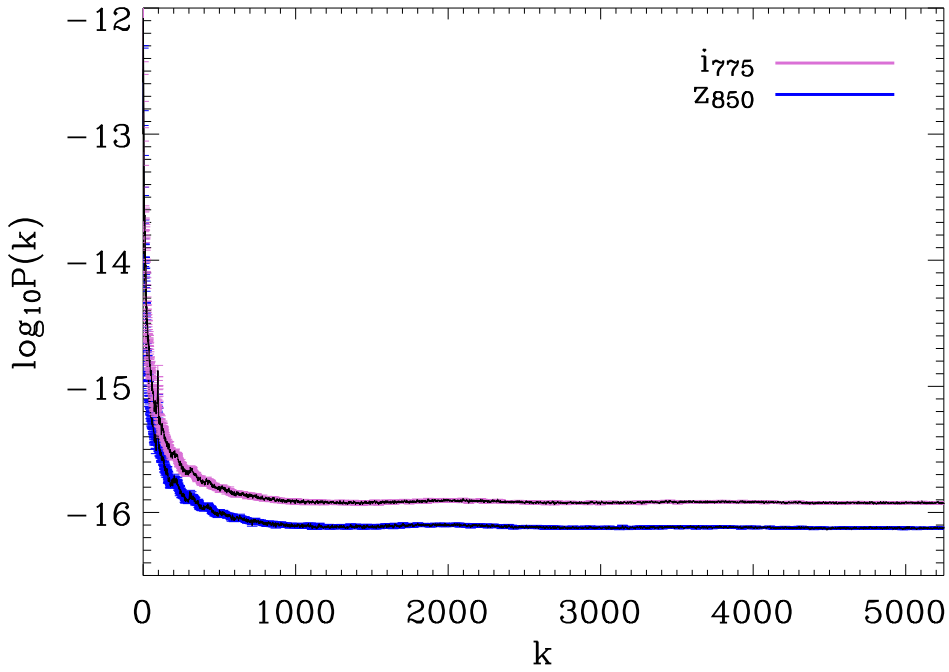


Figure 4.1: Power spectra with error bars of the background signal of the i_{775} (purple top line) and z_{850} -band images (blue bottom line) of the XDF.

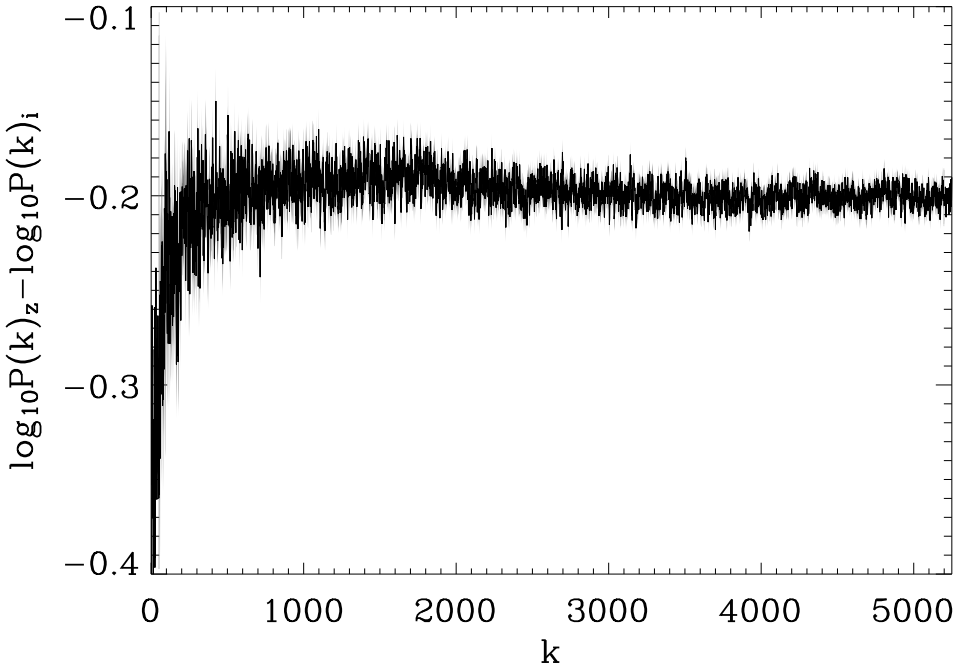


Figure 4.2: Ratio between the z_{850} and i_{775} -band power spectra of the XDF. The shaded area represents the errors associated to the ratio.

this prevented us to detect any light contribution from faint galaxies using the power spectrum technique.

When creating the IR stack of the XDF, the XDF team followed a procedure similar to ours since they subtracted twice a median frame to correct for imperfections and increase the pixel-to-pixel S/N

At first they masked the sources detected in each image and median stacked all the observations in the 4 IR bands. After subtracting the median from each frame they obtained the final drizzled image and detected the sources in that deep image. Using the mask created on the basis of the final stack they got a second median image, more accurate than the first one and characterized by almost zero contamination and subtracted that from all the frames before drizzling.

We used the images with a scale of 0.06 arcsec pixel^{-1} and applied the power spectrum technique to check if the median subtraction is actually preventing us from detecting any background signal as we claimed in the previous Chapter. As expected, the ratio between the power spectra for the HUDF09 main field (Figure 4.3) does not show any light excess ascribable to a faint galaxy population. In particular, the plot highlights that there are still feature that need to be corrected since the high-k part of the ratio is not smooth and flat as should be if just due to white noise. Moreover, in general the plot is much noisier than what we need for a proper analysis of background signals.

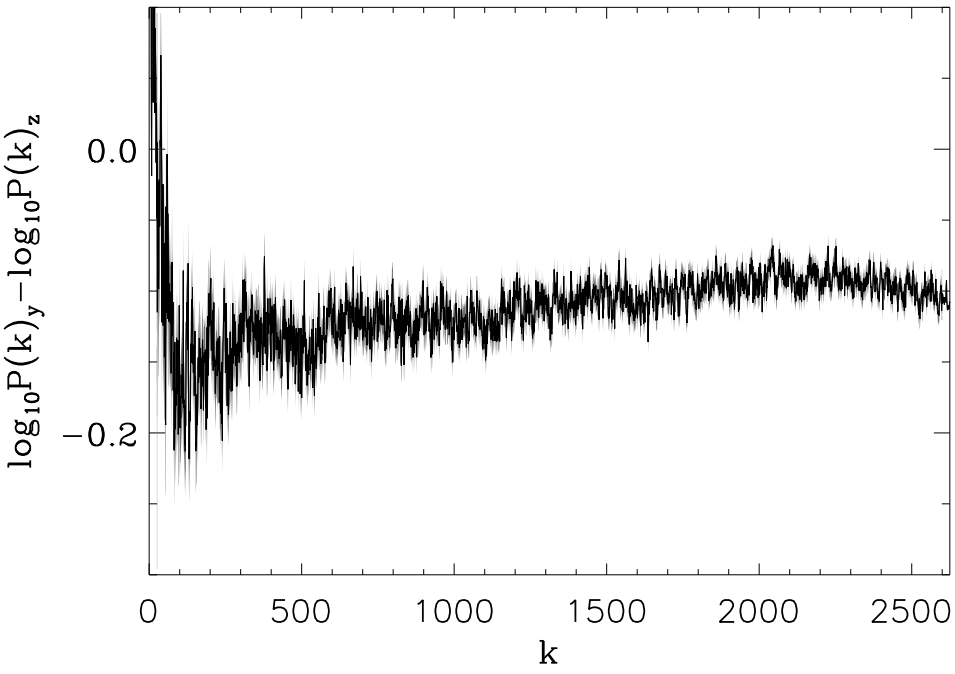


Figure 4.3: Ratio between the Y_{105} and J_{125} -band power spectra of the XDF. The shaded area represents the errors associated to the ratio.

4.2 Description and Aim of the Reprocessing Procedure

As fully discussed in Section 2.2.4 a careful data reduction and the consequent cleanliness of the data are key points that can substantially affect the identification of all high-redshift galaxies, not only the faint ones we are interested in. Since the only field characterized by an improved data reduction is the NICP12 one we used for the analysis described in Chapter 2 and since it would be interesting to confirm our results on the LF at $z \sim 6$ performing the same analysis on a different field with analogous level of data reduction, we decided to work on the HUDF main field collecting all the data available and reprocessing them to obtain the v2.0. To this aim we considered all the observations of that particular sky area taken with ACS from 2002 to 2011 and we reprocessed them the same way as previously done for the NICP12 field and described in Oesch et al. (2007).

We started from raw data not calibrated so we had to process from the very beginning all the images, as well as the reference files, meaning bias frames, darks, and flat fields.

The herringbone artifact was removed from bias and dark frames obtained before July 2006 by processing the images with a Fourier filter. All images were affected by CTE, which was corrected as well.

We obtained hyperbiases and hyperdarks by combining hundreds of single bias and dark frames to increase the S/N in the stack. Then, we derived the sky-flats from the data after masking all the detected sources existing in the field and used them to get a correction at low frequencies to be added to regular flats existing in the archive. We masked all the satellite trails existing in the data and, finally, we carefully aligned the frames and combined them to get the final drizzled frames.

4.3 Raw Data

All the raw data were downloaded from the MAST archive[2] using the following criteria:
R.A. = 03:32:39.0
DEC = -27:47:29.0
RADIUS = 3 arcmin
IMAGERS = ACS
APERTURES = WFC/WFC-CENTER/WFC-FIX
FILTERS/GRATINGS= F775W;CLEAR2L (or F850LP;CLEAR2L)
OBSERVATIONS= science
The data we selected and used for the following analysis come from different proposals and were imaged during several years (see Table 4.1 for details). Even though some proposals, such as those by Riess and Giavalisco, were not aimed to specifically study the HUDF, they were included in our analysis because they permitted to increase the depth of the final images. In total we collected 359 raw frames in the i_{775} band and 443 in the z_{850} one from 10 different proposals.

Each **raw.fits* file includes a raw uncalibrated image from a single exposure in data numbers. The suffix **asn.fits*, **spt.fits*, and **trl.fits* indicates the association file for an observation set, the telemetry and engineering data, and the trailer file containing `calacs` processing comments, respectively.

[2] http://archive.stsci.edu/hst/search.php

ACS image data are stored in multi-extension FITS files as well as the associated data quality information. Raw and calibrated WFC images contain six data extensions in addition to the global header which can be always found in extension [0]. The science image (SCI), error array (ERR), and data quality image (DQ), are stored in extension [1], [2], and [3] for chip 2 (WFC2), and [4], [5], and [6] for chip 1 (WFC1), respectively.

The name of each **.fits* file encodes all the pieces of information of the images itself. If we consider the file named **j8m848grq.fits** as example:

- **j** indicates that the camera used for the observation is ACS,
- **8m8** refers to the observing program,
- **48** indicates the visit number within the program,
- **gr** is the exposure identifier,
- **q** is always the conclusive letter in the name.

Usually the processing of ACS data is carried out by two separate packages: `calacs`, which corrects for instrumental effects and generates calibrated products, and MultiDrizzle, which corrects for geometric distortion, and performs cosmic ray rejection on combined images.

Proposal ID	P.I.	HST Cycle	Frames i_{775}	Frames z_{850}
9352	Riess	11	0	28
9425	Giavalisco	11	24	52
9488	Ratnatunga	11	2	3
9575	Sparks	11	16	0
9978	Beckwith	12	276	272
10086	Beckwith	12	12	16
10189	Riess	13	6	6
10340	Riess	13	5	12
11563	Illingworth	17	18	34
12060	Faber	18	0	20

Table 4.1: Characteristic of the data used for the reprocessing of the HUDF main field. Col. 1: Proposal identification number. Col. 2: Last name of the principal investigator (P.I.). Col. 3: HST cycle during which the observations were taken. Col. 4: Number of frames in the i_{775} band. Col. 5: Number of frames in the z_{850} band.

4.4 Calibration Reference Files

First of all, to reprocess the HUDF data it is necessary to create new calibration reference files, in particular bias frames first, and then darks and flat fields.

4.4.1 Bias

For WFC on ACS two bias frames per day are obtained, the first one using the default gain setting (i.e. CCDGAIN=1 or CCDGAIN=2) and the other one using the higher gain setting (i.e. CCDGAIN=3 or CCDGAIN=4). Every week all the daily biases are combined to obtain a superbias frame for each gain.

Since we wanted to reduce the data as well as possible, for each observing run we collected all the uncalibrated bias frames taken within one year from the last observation and we downloaded them from the MAST archive as well. Furthermore, we selected only those biases with the CCD Amplifier Readout Configuration 'ABCD' (CCDAMP='ABCD') and the gain (CCDGAIN) corresponding to the one of the raw images the bias will correct (i.e. CCDGAIN=1.0 for data obtained before 2006 and CCDGAIN=2.0 for those obtained after 2006). The goal was to combine the biases into a single frame, a so called hyperbias.

Briefly what we did was to consider all the raw biases, to change the header keyword BLEVCORR from OMIT to PERFORM and to run `calacs`. This procedure subtracts the bias level of the overscan region from the bias itself, trims off these overscan regions, and gets the **flt.fits* file. We, then, re-inserted the overscan regions setting them to zeros. The units of the **flt.fits* files are ELECTRONS so we used the GAIN to go back to COUNTS since the hyperbiases used for the calibration are required to be in unit of COUNTS and to include the overscan region. It should be noted that, with the option to perform CTE-correction, `calacs` creates an additional output frame termed **flc.fits*, namely the CTE-corrected version of the **flt.fits*. From here on we will focus just on the **flc.fits* files for any further step in the data reduction.

The **flc.fits* files were, then, combined using `IMCOMBINE` in `IRAF`. In detail we set:

- combine=median
- reject=sigclip
- lsigma=3
- hsigma=3.

These choices were based on the fact that we wanted the procedure to use the sigma clipping method to reject all the cosmic rays (i.e. all pixels with values out from the 3σ interval) and, then, to compute the median of the clean frames. It is important to note that each chip was median stacked separately and that we created a single multi-extension *.fits* file later.

ACS suffered component failures in the electronics in June 2006 and January 2007. These failures prevented respectively the operations of WFC and HRC cameras. WFC was restored to operation after the Servicing Mission 4 (SM4) in May 2009. The readout noise, linearity, pixel full-well depth, and amplifier cross-talk of the restored WFC are as good or better than the pre-failure levels and the dark current, hot-pixel fraction, and charge transfer efficiency (hereafter CTE) have degraded to the levels expected after extended exposure to a trapped radiation environment (Gonzaga, 2011). For this reason all the observations taken before June 2006 should be corrected for the herringbone artifact that arises when the data are compressed while the CCD is reading out. This introduces a very low level electronic pattern in the data that appears as a herringbone effect. The science images were not compressed, but biases and darks were, so the final science product showed this

effect, which was fully understood and modeled by Eddie Bergeron (Oesch et al., 2007), who wrote an IDL code to model the artifact according to the compression factor.

To use this code it was necessary to check the header of each **spt.fits* file associated to a bias looking for the keyword BUFSIZE which indicated the compressed block size. Moreover, each chip is affected differently by the herringbone, so it is necessary to build up two models, one per chip, for each value of the compressed block size. The suitable model was, then, subtracted from each bias frame before computing a median stack of all the corrected bias frames to remove cosmic rays.

Since all the observations carried out before 2006 were characterized by the same electronics, we ended up creating only one hyperbias for all those observation and we did the same for those taken between 2006 and 2009 and for those obtained after 2009 as well.

4.4.2 Dark

For ACS calibration four dark frames with 1000s exposure are obtained daily in addition to some extra frames obtained before and after each annealing operation. Usually a daily superdark is produced and, on a bi-weekly basis, all the daily super dark are collected to produce a basedark, characterized by a lower Poisson noise compared to the daily ones, with an $\mathrm{rms} = 0.0026\ \mathrm{e^-/pix/sec}$ versus the daily bias $\mathrm{rms} = 0.0037\ \mathrm{e^-/pix/sec}$ (Mutchler et al., 2004).

For each observing run we selected all the dark frames taken during the 6 months preceding the last observation and we downloaded the uncalibrated files from the MAST archive, as well. We chose to collect dark frames for just 6 months, instead of one year because dark frames are 4 per day and not one per day as the biases.

Each raw dark frame was processed using `calacs` using the corresponding hyperbias we previously obtained. The final output is a dark frame with a dimension of 4096x2048 pix, i.e. with overscan regions trimmed off, in units of $\mathrm{e^-/sec}$ It should be noted that since the darks obtained before 2006 are affected by the herringbone artifact as well as the biases, for those files it was necessary to re-add to the hyperbias the corresponding herringbone model according to the compressed block size and to use this hyperbias when running `calacs`. In this way we removed the artifact from the darks as well.

To take into account the CTE we switched on the corresponding correction when running `calacs` setting PCTECORR='PERFORM' for all data. For the following analysis we focused just on the CTE-corrected **flc.fits* files, discharging the **flt.fits* ones.

The **flc.fits* files were, then, combined using `IMCOMBINE` in `IRAF`. In detail we set:

- combine=average
- reject=sigclip
- lsigma=3
- hsigma=3

This choice was based on the fact that we wanted to use the sigma clipping method to reject all the cosmic rays and, then, combine the clean frames. As we did for the biases, each chip was median stacked separately and we created a single multi-extension fits file later. We used the sigma image created by `IMCOMBINE` as the ERR frame corresponding to each SCI one. The final output hyperdark frame has unit of $\mathrm{e^-/sec}$

The major problem with darks is due to hot pixels whose number is increasing with time. For this reason a monthly annealing process, consisting in warming up and cooling down the detector right after, was introduced. Once a month the CCDs of ACS are heated for 10-16 hours by 100 degrees above the typical operating temperature, which ranges around -80 C. This is intended to fix the damaged latties, reducing the population of hot pixels without affecting the others (Riess et al., 2002). Hot pixels, i.e. those characterized by a dark current $> 0.08\ \mathrm{e}^-/\mathrm{sec}$, and warm pixels, i.e. those with values that exceed, by 5σ [3], the normal distribution of the pixels in a dark frame, up to the threshold of the hot pixels, need to be flagged in hyperdark DQ arrays, i.e. extensions [3] and [6].

To identify the hot/warm pixels, we had to create a stack of the dark frames spanning the period of time from right before the scientific observation back to the date of the previous anneal. In particular, we run `IMCOMBINE` to create a stack associated to each day of observation, we identified on the stacked image all those pixels characterized by an high dark current value, and, then, we flagged them in the DQ array associated to the stack with flag 16. Finally we updated the DQ array of the hyperdark derived combining all the dark frames taken over the 6 months period in order to obtain an hyperdark per each observing day. So, all these hyperdarks are characterized by the same SCI and ERR arrays and differ only in the DQ ones.

After creating the hyperbias and the hyperdark frames it was necessary to update the corresponding keywords in the header of the raw images so as `calacs` could use the new calibration files and could create the **flt/flc.fits* files.

4.4.3 Flat Field

The flat field correction varies according to the band used to carry out the observations, so a different hyperflat reference file needs to be created for each band. To derive the hyperflat necessary for the calibration we used the data themselves, i.e. we created sky .

Survey programs such as GOODS and UDF are suitable for creating sky flats because the data consist mostly of sparsely populated images with relatively uniform sky. Stacking the images after removing cosmic rays and masking all of the sources permits to quantify the pixel-to-pixel variation of the instrumental response (Gonzaga, 2011) and to look for low frequencies residuals not correct by the pipeline flat fields.

In principle it is possible to obtain the sky flats with or without using the FLATCORR in `calacs`. The difference is that with FLATCORR = 'PERFORM', the output is a correction to the CDBS flat, basically a second order correction, which needs to be applied to the CDBS flat to make the master flat, while setting FLATCORR = 'OMIT', an absolute flat field is created. One of the possible drawbacks in getting the absolute sky flat is that if a given pixel has a source in all frames in the stack it is masked in all of them, and hence there are no information there and it is necessary either to rely on the CDBS flat or to get the median values from the surrounding pixels to fill in any such holes.

Our sky flats were created by median-combining the pipeline reduced **flc.fits* files after removing cosmic rays and masking all of the sources. As we did for biases and darks, we `IMCOMBINE` with the sigma clipping procedure to get the stack. To prevent residuals of source not properly masked to contaminate the flat field we made an aggressive choice of the lower and upper thresholds, i.e. assumed lsigma=2 and hsigma=2. At the very end,

[3] As a reference, the 5σ threshold was 0.022 $\mathrm{e}^-/\mathrm{sec}$ in July 2009

anyway, the skyflat turned out to be characterized by a lower signal-to-noise with respect to the CDBS ones, in detail the S/N differs by a factor of 10 which basically lead us to decide to rely on the CDBS flat fields for a first accurate correction of the pix-to-pix variations and to get just a second order correction starting from the data.

4.5 Satellite Trails

Another important issue which need to be taken into account is how to mask satellite trails existing in the data. To do that we used the `SATMASKS.PY` script, developed at STScI with the essential goal of automating the satellite trails and ghost images masking procedure. Please note that, to run this Python[4] script, SAOImage DS9[5] display program is required. The script will search the directory in which it is located for the **flc.fits* files and, then, it will generate masks for all the satellite trails existing in the images.

First of all, the user has to check all the **flc.fits* images one by one, dividing those showing a trail from the remaining ones. Then, for each satellite trail identified it is necessary to check the start and end coordinates and to take note of its width. These are the input that the script requires to generate the mask.
The steps taken by the script are as follows:

1. drizzle the **flc.fits* files,
2. create a single science image showing both chips at the same time (Figure 4.4),
3. run the masking script. When the science image is displayed the user needs to provide the start and end coordinates of the satellite trail by clicking on the image, then the user should enter the width of the trail in pixels,
4. blot the image,
5. create two single **.fits* files, one per chip, containing the mask of the trail for that particular chip.

After checking one by one all the frames we were dealing with, we obtained a final list of images affected by satellite trails consisting in 10 frames in the i_{775} -band and 5 in the z_{850} one (see Table 4.2 for details). Mostly the single science image show only one trail, but a few are affected by two trails.

After getting all the masks we flagged all the pixels affected by the trails in each DQ array (flag=16384[6]) so as to prevent the trail from contributing to the final drizzled image.

4.6 Alignment

A careful alignment of the frames is fundamental to obtain a proper final frame when combining all the data available. `DrizzlePac` (Gonzaga and et al., 2012) is a software package for combining and aligning images that replaced the `Dither` package in `STSDAS` and that

[4]Python Software Foundation. Python Language Reference, version 2.7. Available at http://www.python.org

[5]http://hea-www.harvard.edu/RD/ds9/ref/

[6]http://www.stsci.edu/hst/acs/analysis/reference_files/data_quality_flags.html

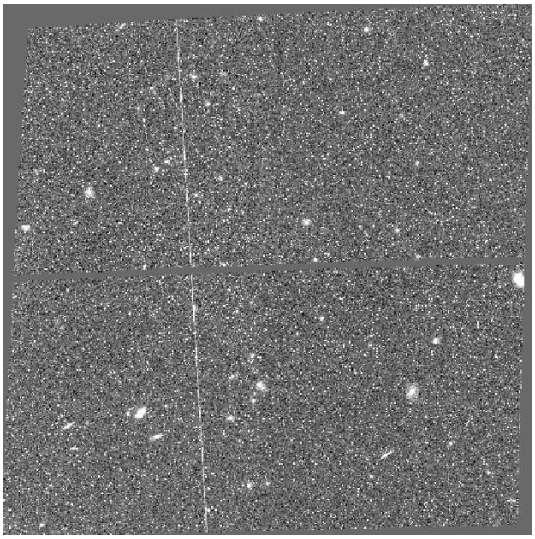


Figure 4.4: ACS i_{775} image showing a satellite trail affecting both chips.

provides a routine for the alignment of HST images named `tweakreg`. `Tweakreg` allows the user to align sets of images to each other and/or to an external astrometric reference frame/image. All the procedure is based on the identification of the same sources in different frames.

`Tweakreg` is based on a daofind-like algorithm looking for stellar objects to perform the image alignment. Unfortunately the sky areas usually selected for deep surveys are characterized by the lack of stellar sources and so `SExtractor` can be used to generate catalogs of objects suitable for the image alignment, as described in Lucas and Hack (2013).
Briefly the steps we followed for the alignment of the ACS datasets are:

1. Update the WCS coordinates for all the **flc.fits* images,
2. Obtain a cosmic ray free image for each **flc.fits* frame,
3. Run `SExtractor` to identify all the sources existing in each chip of every image and, then, select the brightest that, at the same time, do not show any elongated feature (i.e. have a roundish shape) and are not too extended,
4. Divide the frames on the basis of the proposal id and orientation,
5. Run `tweakreg` using the catalogs of sources identified in the previous to align the **flc.fits* images with the same characteristics,
6. Run `AstroDrizzle` to combine the frames aligned in the previous step,
7. Run `SExtractor` on the stack images and create a catalog of sources suitable for the alignment, following the same prescription as for the single science images,
8. Run again `tweakreg` to align the stack obtained from different proposals and/or orientation,

i_{775} -band			z_{850} -band		
Image	chip 1	chip 2	Image	chip 1	chip 2
j8fnebj2q_flc.fits	1	0	j8m819bbq_flc.fits	1	1
j8m80dcbq_flc.fits	1	1	j8m820m3q_flc.fits	1	1
j8m82fymq_flc.fits	1	1	j8m871otq_flc.fits	1	1
j8m881d9q_flc.fits	0	1	j8m899bbq_flc.fits	1	1
j8fnegrbq_flc.fits	1	0	j8m8a7rsq_flc.fits	1	1
j8m825usq_flc.fits	1	0	j8m8cbn8q_flc.fits	1	1
j8m845atq_flc.fits	0	1	j8m8d1djq_flc.fits	1	1
jb5x51l6q_flc.fits	0	2	jb5x66u9q_flc.fits	1	1
j8m818zsq_flc.fits	1	0	jbetaqryq_flc.fits	1	0
j8m837hdq_flc.fits	1	1			
j8wc7er1q_flc.fits	2	2			
j8m810ijq_flc.fits	0	1			
j8m831goq_flc.fits	0	1			
j8m8b0l0q_flc.fits	1	1			

Table 4.2: List of the images affected by satellite trails in the i_{775} (left side) and z_{850} -band (right side) and corresponding number of trails. For each frame the number of trails existing in chip 1 and 2 are specified.

9. Run `tweakback` to apply the WCS solution recorded in drizzled image to all the input images used by `AstroDrizzle`,

10. Run `AstroDrizzle` on all the **flc.fits* frames to get the final drizzled image.

The tricky part of the alignment was dealing with those proposals that were not designed to directly image the sky area occupied by HUDF main field, whose frames cover only a fraction of the main HUDF fields. Since some of these frames do not overlap but for a small region, the number of sources in common is low.
The following subsections will describe more in detail each step in the alignment procedure.

4.6.1 WCS Coordinates Update

The last version of HST data available in the MAST archive was designed to work with `AstroDrizzle` and, hence, includes by default groups that were not built in the version we originally downloaded. For this reason, first of all we had to update the WCS coordinates in all the headers of our frames running the task `updatenpol` which is included in the `DrizzlePac` software.

`Updatenpol` performs an update of the header of ACS input files with the names of new NPOLFILE and D2IMFILE reference files required for use with `AstroDrizzle`.
In Pyraf the syntax to be used is:

```
import drizzlepac
from drizzlepac import updatenpol
updatenpol.update('*flc.fits')
```

The original structure of each **flc.fits* files consists in 7 groups as show in the following table:

EXT #	FITS NAME	FILE NAME	DIMENSION	BITPI OBJECT
0	j8wc83deq_flc	j8wc83deq_flc.fits		16
1	IMAGE	SCI	4096x2048	-32
2	IMAGE	ERR	4096x2048	-32
3	IMAGE	DQ	4096x2048	16
4	IMAGE	SCI	4096x2048	-32
5	IMAGE	ERR	4096x2048	-32
6	IMAGE	DQ	4096x2048	16

After the WCS are updated through the `updatenpol` task, the structure becomes like this:

EXT #	FITS NAME	FILE NAME	DIMENSION	BITPI OBJECT
0	j8wc83deq_flc	j8wc83deq_flc.fits		16
1	IMAGE	SCI	4096x2048	-32
2	IMAGE	ERR	4096x2048	-32
3	IMAGE	DQ	4096x2048	16
4	IMAGE	SCI	4096x2048	-32
5	IMAGE	ERR	4096x2048	-32
6	IMAGE	DQ	4096x2048	16
7	IMAGE	D2IMARR	4096x1	-32
8	IMAGE	D2IMARR	4096x1	-32
9	IMAGE	WCSDVARR	65x33	-32
10	IMAGE	WCSDVARR	65x33	-32
11	IMAGE	WCSDVARR	65x33	-32
12	IMAGE	WCSDVARR	65x33	-32
13	BINTABLE	WCSCORR	24Fx14R	

4.6.2 Cosmic Ray Rejection and SExtractor Catalogs

As stated above, deep surveys lack of bright stellar sources that `tweakreg` usually uses for the triangulation and the user is required to provide his/her own catalogs of sources.
To identify sources suitable for the alignment process it is necessary to deal with images not affected by cosmic rays. `AstroDrizzle` performs a cosmic ray rejection relying on the original header astrometry.
The syntax in Pyraf is the following:

```
import drizzlepac
from drizzlepac import astrodrizzle
unlearn astrodrizzle
astrodrizzle.AstroDrizzle('*flc.fits', drizz_cr_corr='yes',driz_combine='no')
```

The output is a **crclean.fits* file for each chip of the input **flc.fits* image.

We run `SExtractor` on each **crclean.fits* to get a catalog of all the sources in the image. As a matter of fact the best objects to perform a careful alignment are stars, in particular the bright ones. Since our images lack of these preferential sources, we selected in our catalogs those objects that more resemble stars. Indeed, we selected for the alignment only those objects satisfying at the same time cuts in magnitude, shape, and dimension. Objects that are faint, elongated or covering a large area were discarded since the quality of the alignment relies on an accurate estimate of the centroid for each matched object. Typically in each chip the number of suitable sources satisfying these selection criteria

ranges from 30 to 50.

4.6.3 Tweakreg

`Tweakreg` allows the user to align sets of images to each other and/or to an external astrometric reference frame or to a specific image.

When the user provides external catalogs of sources, `tweakreg` requires two input: the list of all the frames to align and a "catfile" in which the catalogs associated to each image are listed as follows:

j8wc7cioq_flc.fits j8wc7cioq_crclean_ext1_sel.cat j8wc7cioq_crclean_ext4_sel.cat

It should be noted that the easiest way is to provide catalogs with x and y position of the selected sources in the first two columns. Otherwise, when running `tweakreg` the user should specify the columns containing the x and y coordinates using the XCOL and YCOL parameters.
The other parameters that need to be tuned to match the objects in the different frames are:

- MINOBJ, i.e. the minimum number of identified objects from each input image to be used in the matching. The default value is 15.
- SEARCHRAD, i.e. the radius within which a match is searched. The default value is 1.0 arcsec.
- SEARCHUNITS, either arcsec or pixels.
- TRESHOLD, i.e. the threshold above the local background in units of sigma used for object detection.
- CONV_WIDTH, i.e. the width of the convolution kernel in scale units. Usually the value is twice the FWHM of the point source, that means 3.5 pixels for ACS/WFC.

If no reference image is specified by the user, the task automatically performs the alignment using the first input of the list as reference.

Usually the user runs `tweakreg` a first time setting UPDATEHEADER = False to check the accuracy of the alignment. This is done on the basis of 3 different plots produced interactively by the task for each image in the input list:

- The two-dimensional histogram shows an initial guess for the offset between the images. When the task succeeds in matching a reasonable number of sources in the two images to be aligned the histogram shows a well defined peak.
- The astrometric residual plot is divided into 4 panels. Each panel shows the distribution of the x and y components of the residuals as a function of the x-axis and y-axis position. In the case of a good fit the values in each plot should be around 0 and show no trend. The NCLIP and SIGMA parameters could be tuned to improve the residual plots and, consequently the RMS value given as output.
- The vector plot shows the magnitude and direction of residuals as a function of location. If the alignment is good the different vector should not show any trend, but be distributed randomly.

As described above, at first we run `tweakreg` on small samples of images within the same proposal and, if possible, the same visit. Under these conditions the alignment was easy since the orientation of the frames was only slightly different. In this case the parameter we used were very conservative, mostly the default ones:
MINOBJ = 15
SEARCHEAD = 1.0
NCLIP = 5
SIGMA = 2
We assumed the alignment to be correct when the typical RMS was below 0.1 in both x-axis and y-axis, but for most of the images the values were below 0.05, resulting in a better alignment.

4.6.4 Drizzling and Second Alignment

The entire alignment process turned out to be an iterative one. After aligning the frames within the same proposal and visit, we run `AstroDrizzle` getting a **drc.fits* file for each sub-group of images. This image is deeper than the single *flc.fits* frames it is derived from and it permits to increase the sample of sources suitable for the subsequent alignment steps.
Since we were not interested in the stack images for scientific purposes we used the default configuration parameters in `AstroDrizzle`. Then, we run `SExtractor` on each **drc.fits* getting the catalogs from which we, then, derived those including only sources satisfying the cuts in magnitude, shape, and size we mentioned before. As done before, afterwards we used these catalogs of selected sources to run `tweakreg` to align the **drc.fits* frames to each other.

4.6.5 Tweakback

When the drizzled images were aligned to each other, we run `tweakback` on each **drc.fits* files. The task `tweakback` propagates the updated WCS information back to all the images that were used to create the input **drc.fits* file.
The syntax to run the task is:

```
from drizzlepac import tweakback
tweakback.tweakback('*drc.fits')
```

Thanks to this task more **flc.fits* images were aligned compared to the previous step and it was possible to go with the iterative process.

4.6.6 Drizzling, Alignment, and Final Combination

Combining a bigger number of frames with `AstroDrizzle` it was possible to get deeper images, consequently increasing the number of sources suitable for the alignment. In this way we managed the align even the images from the proposals not centered on the HUDF, but just marginally overlapping that sky area.
Once again, after aligning drizzled images, we had to propagate the changes in the coordinates through `tweakback` so as to have each single **flc.fits* frame aligned to the other.

Finally we were able to get the final science image running `AstroDrizzle` for the last time. The combination of more than 300 frames requires the user to set the appropriate

memory usage through `setenv PYFITS_USE_MEMMAP 0` before running `Pyraf` and `AstroDrizzle`. The parameters we used for the final drizzling procedure are the following:

```
num_cores = 1
skystat = 'mode'
skylower = 0.0
skyclip = 20
driz_combine = True
final_wht_type ='IVM'
final_kernel='point'
final_wcs = True
final_rot = 0
final_scale = 0.03
final_outnx = 10500
final_outny = 10500
final_ra = 53.1625
final_dec = -27.791417
```

The number of cores used is determined by memory issues and cannot be more than one when so many frames are combined together. The parameters related to the background determination were tuned to get a better estimate than the standard one, so as to avoid features or gradients in the final output.

We asked the final weight image to be an inverse variance map (IVM) since that is the type of weight map that `SExtractor` requires during the source detection process. The point kernel was used for both the passbands when drizzling the data to avoid correlated noise that could bias the study of background fluctuations. The final scale (0.03 $\mathrm{arcminpixel}^{-1}$ and dimension of the frame (10500 by 10500 pixels) were chosen to be consistent with the previously released XDF images.

4.7 Correcting for the Bias Jump

The final science images that we got for both the i_{775} and z_{850} -band showed the clear signature of the quadrant-to-quadrant bias jump that is known to affect ACS data. Briefly, each amplifier of WFC/ACS is characterized by a different level and structure of the bias that determines the jump noticeable in the **raw.fits* files. The issue is that even calibrated **flt/flc.fits* images obtained running the standard pipeline show a jump in the background level between the two quadrants of each chip. This residual offset existing between the quadrants A and B or C and D of the amplifier is due to uncertainties in the bias level that is subtracted by `CALACS` (Sirianni et al., 2003). The offsets vary from amplifier to amplifier, showing random variations that can be as large as 3.5 DN.

There are three ways to deal with this issue. The first one is to measure the background level in each quadrant and, then, subtract the sky level separately for each quadrant. If so, the user have to get the maximum value of the sky subtracted and add a new keyword to the header of each **flc.fits* file to store this value. Then, he/she has to provide `AstroDrizzle` with the name of this new keyword through the SKY_USER parameter. In this way `Astrodrizzle` is not going to subtract the sky from the single frames but it will use the background level provided by the user. The second option is normalizing the background to the highest level among the 4 quadrants to have a uniform value all over the two

chips and, then, let `AstroDrizzle` determine and subtract the sky level. The third option is to use the `skyfile` parameter in `AstroDrizzle`. This means the user has to create an input file associating to each **.fits* file, corrected for the chip-to-chip jump but not sky subtracted, an overall sky level (either the mean between the values of the quadrants or the maximum value). `AstroDrizzle` is, then, going to combine all the frames without computing again the background level, but using the input value instead.

In all cases the background in each chip is computed after masking all the cosmic rays and sources and using a 3σ clipping to avoid contamination from bright pixels.

We performed the combination using all the three different methods to check which one was the best. We encounterd some troubles when running `AstroDrizzle` on the sky-subtracted images, but this is a know issue that should be fixed soon. Regarding the other two options, the differences in the final outputs are very small, especially since we required `Astrodrizzle` to iterate 20 times the sigma clipping before estimating the final sky value. Anyway, the images obtained using the `skyfile` keyword are slightly better because to determine the background level we first masked all the bright sources and cosmic rays and then we performed a sigma clipping, while `AstroDrizzle` performs the sigma clipping starting from images still containing cosmic rays, stars, and galaxies.

4.8 Electronic Crosstalk Ghosts

Due to some kind of electronic crosstalk (Giavalisco, 2004) sources imaged in one of the ACS/WFC quadrant produce a dark ghost image of themselves in the other three quadrants. Electronic ghosts are faint structures that mirror real images recorded in another quadrants.
As a general rule crosstalk is a small effect in absolute terms, i.e. a few electrons/pixel within a wide range of fluxes of the sources, but it shows some kind of noise, which means that the effect changes strength even for a fixed flux. In general, as described in Giavalisco (2004) the gain setting can help reducing the effect of these ghosts if set to 2 e^-/DN, but there is no way to get completely rid of them. In particular, when drizzling multiple frames together, the final effect is a smeared ghost image. According the recent version of the ACS Handbook[7] after the servicing mission 4 (SM4) crosstalk due to sources with low signal can be ignored, but our dataset is mostly made of images that were taken way before SM4 so we need to take the electronic ghost into account.

E. Bergeron has been studying the electronic crosstalk and its effect on deep images for a long time and he ended up finding a way to get the ghost images needed to correct for the crosstalk effect. Considering how the readout directions in the 4 quadrants of ACS are oriented (Figure 4.5) the ghost images and real sources move in opposite directions along the x-axis in the same direction along the y-axis. On the basis of above, if we flip the quadrants for each image and, then, run `AstroDrizzle` on these files the ghost images will be coadded while the sources will not, being dithered in random directions. Applying a cut on the final output it is possible to get images of the faint electronic ghosts with high S/N. Finally, these ghost images should be subtracted from the final drizzled image. A more accurate correction for the electronic ghosts would be building up a model of the effect, but so far we did not have the chance to implement it.

[7] http://www.stsci.edu/hst/acs/documents/handbooks/currentDHB/acs_cover.html#1513

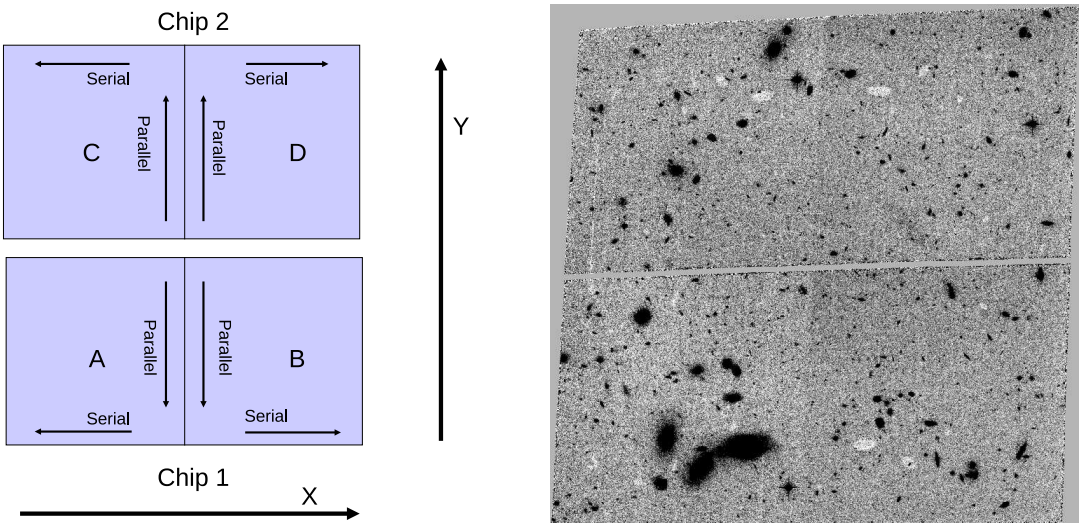


Figure 4.5: Left panel: scheme of the readout directions of chip 1 (quadrants A and B) and chip 2 (quadrants C and D) of ACS (courtesy of E. Bergeron). Right panel: ACS image showing evidences of crosstalk. Electronic ghosts are the bright structures clearly noticeable in all the four quadrants.

We are still working on this issue and we plan to correct our images for the electronic ghost as soon as possible.

4.9 Preliminary Results on the XDF v2.0 Dataset

As we said before, we are still working to get models of the electronic ghosts affecting our images, so the final dataset is still under construction. Anyway, the alignment and the overall quality of the combined images seem to be very promising.
Halfway to the end of this massive effort we checked the photometry, comparing the results to the previous HUDF and XDF datasets and we applied the power spectrum technique to show what the achievement in the background smoothness with respect to the the v1.0 of the XDF.

4.9.1 Photometry Quality Check

To test the quality of our final images, we compared our photometry to the HUDF and XDF ones. First of all we checked the alignment of the three datasets and mask to zero level those pixels that lie outside the sky area covered by the main HUDF field. Once we had all the images perfectly matching we run `SExtractor`. In particular we used the dual image mode in each band using our image for the detection and, consequently, performing the photometric analysis on the same pixels in all images. We chose the parameter set stated in Beckwith et al. (2006), tailored for the detection of high-z sources in the original HUDF images. The threshold for detection and analysis (DETECT_THRESH and

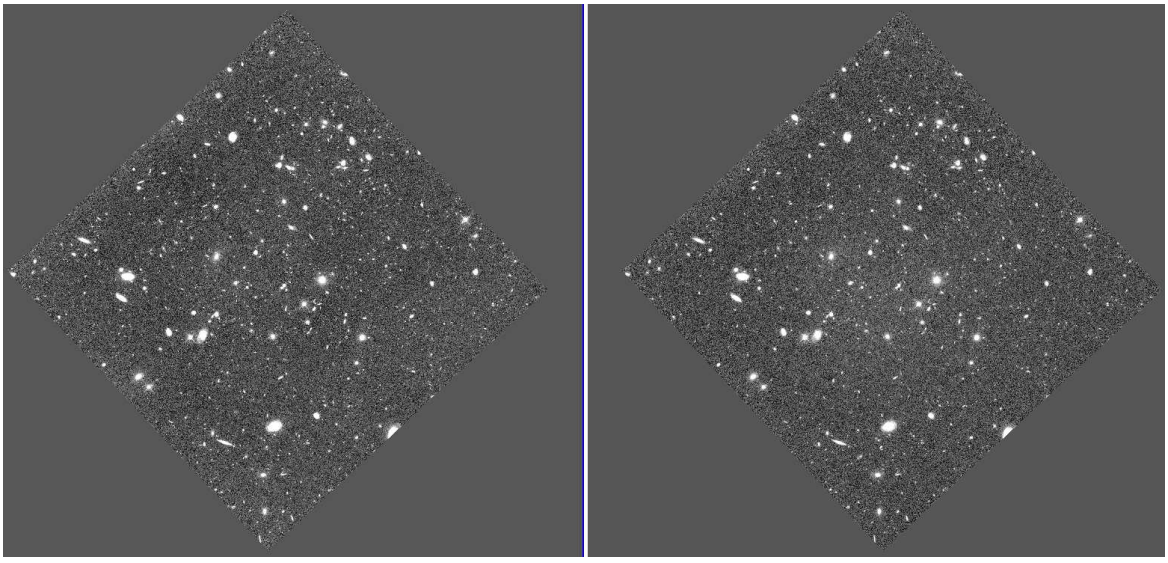


Figure 4.6: The new version of the XDF images in the i_{775} (left panel) and z_{850} -band (right panel) obtained combining 355 and 440 frames, respectively, and performing a data reduction aimed at getting the cleanest image possible.

ANALYSIS_THRESH) was set to be 0.61 within an area of at least 9 contiguous pixels (DETECT_MINAREA). We selected a total of 32 deblending subthresholds (DEBLEND_NTHRESH), a contrast parameter of 0.03 (DEBLEND_MINCONT), and a Full Width Half Maximum (SEEING_FWHM) of $0''.03$.

We compared the total magnitude, derived from the MAG_AUTO parameter, and the S/N of sources with at least a 3σ detection in each image. The plots in Figures 4.7 and 4.8 show that we are perfectly in agreement with both HUDF and XDF down to magnitude $\sim$ 27 mag in both the i_{775} and z_{850} -band images, but at fainter magnitudes it seems that we are able to better recover the objects. On average we get brighter magnitudes for the faint population of galaxies as well as a higher S/N. All the plots support the fact that our tailored dataset is comparable to the other two and the further steps we will perform soon will likely make it slightly better than the previous ones when looking for faint objects.

4.9.2 Power Spectra

Even though we are still working on the improved version of the XDF, we applied the power spectrum technique to our dataset to check if we could already notice any achievement on the quality of the background signal. As shown in Figure 4.9 the difference between the v1.0 released on the archive and our v2.0 is evident and we can see a feature resembling the light excess we used in Chapter 2 to constrain the faint-end slope α of the LF at $z \sim 6$. Moreover, the improvement in the quality of the background signal is suggested by the flatter trend in the high k regime when comparing the improved version to the original XDF.
Of course, since the data reduction process is not yet completed we could not use the bump for any claim, but the comparison with the public dataset is very promising.

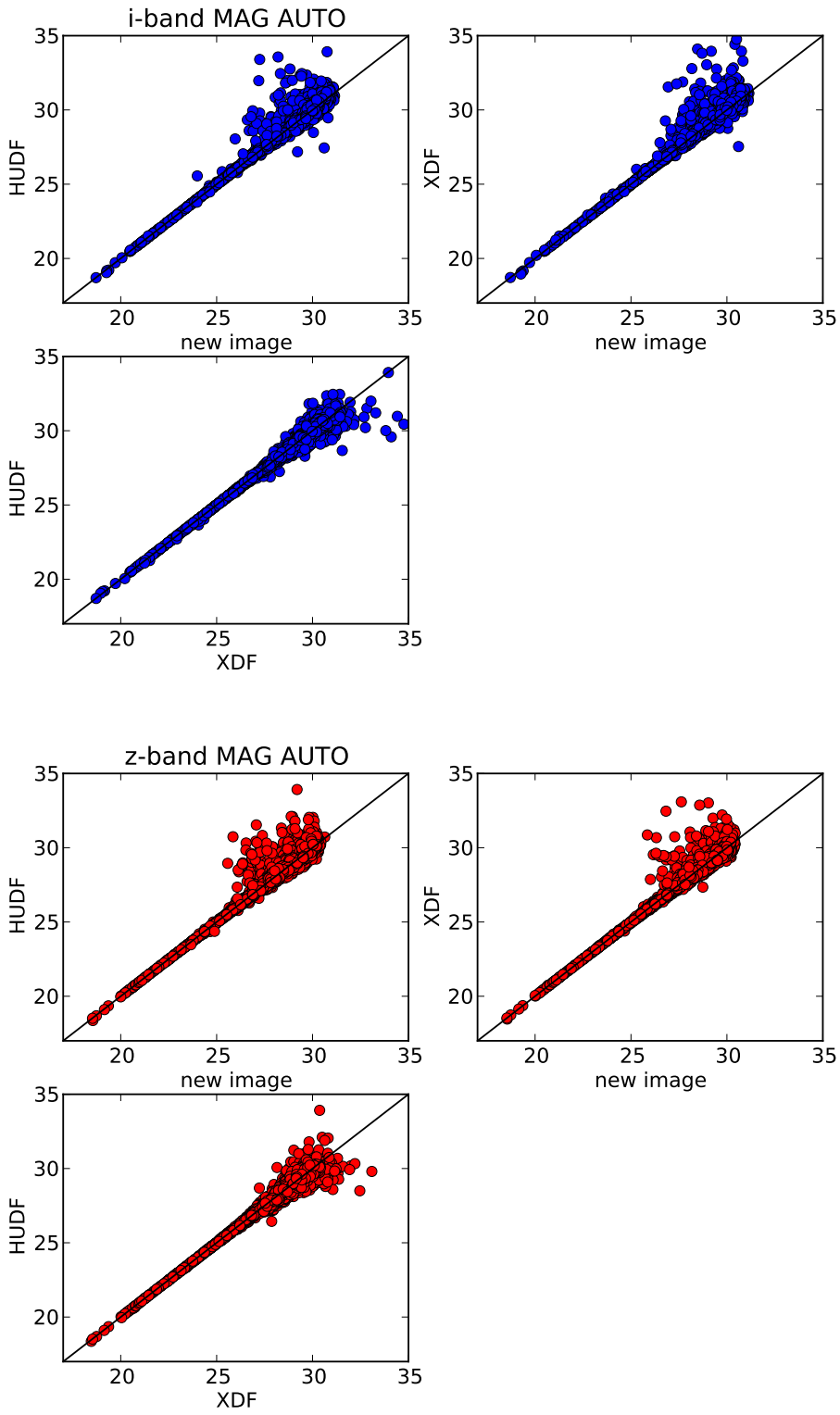


Figure 4.7: Comparison of the total magnitude, derived from SExtractor MAG_AUTO parameter, for galaxies detected in HUDF, XDF, and our images in the i_{775} (top panels, in blue) and z_{850} -band (bottom panels, in red). The distribution of the data points suggests that our dataset is slightly deeper than the other two and can better recover the magnitude of faint objects.

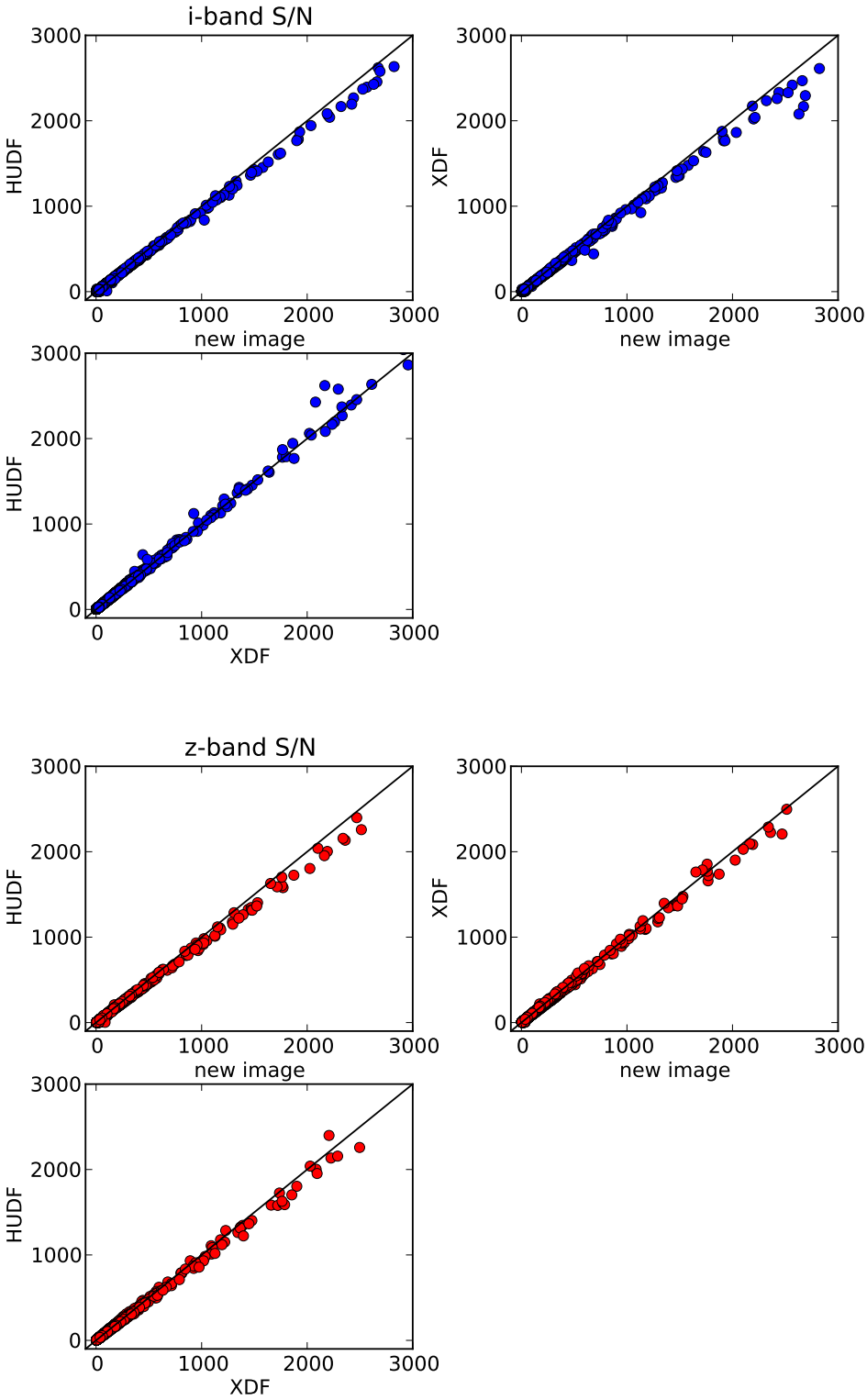


Figure 4.8: Comparison of the S/N for galaxies detected in HUDF, XDF and our image in the i_{775} (top panels, in blue) and z_{850} -band (bottom panels, in red). A clear trend shows that our new dataset is characterized by a better quality in the detection.

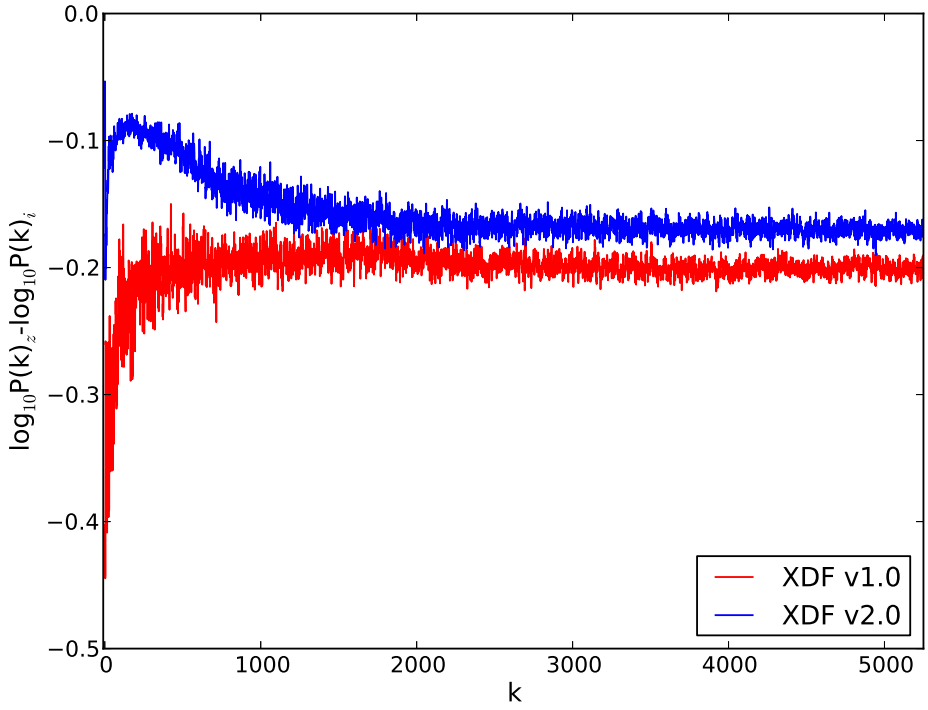


Figure 4.9: Ratio between the powerspectra in the z_{850} and i_{775} -band obtained from the v1.0 (in red) and v2.0 (in blue) of the XDF dataset. The improvement in the quality of the background signal is suggested by the flatter trend in the high k regime and by the presence of the bump in the curve relative to the improved version of the data (v2.0).

4.10 Summary and Conclusions

In this Chapter we presented all the reprocessing of the ACS dataset we performed to obtain the version 2.0 of the eXtreme Deep Field program, i.e. the deepest optical images of the Universe ever obtained. As for the XDF, we stacked together all the images covering (totally or partially) the HUDF main field sky area and we used improved data reduction algorithms.

Briefly, we obtained hyperbiases and hyperdarks covering a baseline much longer (1 year and 6 month, respectively) than the usual one (2 weeks), used for regular bias and dark frames applied by the standard pipeline.
We downloaded all the raw images from the MAST archive and run `CALACS` applying our own reference files and performing the CTE correction.
In particular we paid close attention to correct for the herringbone effect, a compression related issue affecting the early ACS data. All the hot and warm pixels were flagged in the data quality array so as to not affect the final product. Satellite trails affecting some of the images were flagged in the DQ arrays as well.
Regarding the flat field, we used the pipeline flats to get a first correction and then, we derived a second order correction comparing a smoothed version of the sky flats obtained

from the data themselves to a smoothed version of the standard flats.
Once we obtained all the **flc.fits* frames we carefully aligned them using `SExtractor` catalogs as input for `tweakreg`. Finally we got the images in the i_{775} and z_{850} -band combining all the frames together using `AstroDrizzle`.

The main goal of this work was to extend the power spectrum analysis to the HUDF main field to verify our findings on the faint-end slope of the luminosity function at $z \sim 6$, as well as to go slightly deeper than the original XDF in the search for high-z faint galaxies. The preliminary analysis on the photometry of our dataset shows a promising trend in the total magnitude and S/N with respect to the HUDF and XDF, but we are still working on the images to get the best data reduction ever. In particular, we are now focusing on modelling the electronic ghosts due to crosstalk.
Moreover, we found a few new frames within the HUDF sky area that could, in principle, further improve the S/N of our final products.
Anyway, as shown in Figure 4.9, even the halfway results obtained applying the power spectrum technique suggest we could isolate a signal excess from undetected galaxies. Therefore, as soon as we are done with the data reduction, we plan to get the power spectra in the i_{775} and z_{850} -band and to use Monte Carlo simulations with the aim to check whether or not the value of $\alpha = -1.9$ we found using the NICP12 field images is confirmed.

Chapter 5

Surface Brightness Dimming Effect

Based on **V. Calvi**, *M. Stiavelli, L. Bradley, and A. Pizzella, 2014, submitted to ApJ*

In this Chapter we will present a quantitative study on the surface brightness (SB) of high-redshift galaxies selected on the basis of the Lyman-Break technique comparing the results obtained from different ultra deep images taken with ACS. Our empirical strategy relies on the comparison of the total flux detected for the same sources in surveys characterized by different depth.

Cosmological surface brightness dimming of the form $(1+z)^{-4}$ affects all sources (Tolman, 1930, 1934). In particular, since for every extended source the surface brightness I is no longer redshift-independent as in an Euclidean Universe, there is an additional term that should be taken into account. The contribution of time dilation, redshift, and curvature makes high-z galaxies progressively difficult to detect according to

$$I_0 = \frac{I_e}{(1+z)^4} \tag{5.1}$$

where I_0 and I_e are the observed and emitted surface brightness values, respectively.

The strong dependence of SB dimming on redshift z suggests the presence of a selection bias when searching for high-redshift galaxies, i.e. we tend to detect only those galaxies with a high surface brightness (SB). Briefly, cosmological dimming makes low-SB features hard to detect at high redshifts. However, unresolved knots of emission are not affected by SB dimming, thus allowing, in principle, a way to test the clumpiness of high-z galaxies.

Our strategy relies on the comparison of the total flux detected for the same source in surveys characterized by different depth. We made use of the GOODS, HUDF, and XDF Hubble Space Telescope legacy datasets to study the effect of SB dimming on low-SB features of high-z galaxies in the sample and we compared it to the prediction for smooth sources. We selected a sample of Lyman-break galaxies at $z \sim 4$ (i.e. B_{435} -band dropouts) detected in all of the datasets and found no significant trend when comparing the total magnitudes measured from images with different depth. Then, we made use of Monte Carlo simulations to derive the expected trend for galaxies with different SB profiles and the comparison to the data hints to a compact distribution of most of the rest-frame ultraviolet light emitted from high-redshift galaxies.

The study of cosmological SB dimming is also important since it could affect our prediction of what JWST can observe at higher redshifts, where younger galaxies may exhibit

a larger fraction of clumpiness. Our direct comparison shows that galaxies detected in GOODS do not become significantly brighter in the HUDF or XDF. This suggest that most of their light is compact and hints to the fact that JWST will likely not find diffuse star forming components.

5.1 The Data

To the aim of a quantitative study on the surface brightness of high-redshift galaxies selected on the basis of the Lyman-break technique (Steidel et al., 1999; Giavalisco, 2002) we compared results obtained from different ultra-deep images taken with the Advanced Camera for Survey (ACS) installed in HST.

In the following analysis we will consider the Hubble Ultra Deep Field (HUDF) main field dataset, the XDF field and the same sky area as imaged by ACS in the Great Observatories Origins Deep Survey (GOODS). In particular, we made use of the optical images taken in the 4 ACS bands: F435W (B_{435}), F606W (V_{606}), F775W (i_{775}), and F850LP (z_{850}). We will briefly describe the GOODS and HUDF dataset hereafter, but more details can be found in Giavalisco et al. (2004), Beckwith et al. (2006). Regarding the XDF, see Section 4.1 or the reference paper Illingworth et al. (2013).

5.1.1 The Hubble Ultra Deep Field

After the successful HDF-N and HDF-S programs, Steven Beckwith decided to invest the Director Discretionary time during cycle 12 in 2006 with the aim of getting the deepest image ever in the optical bands using ACS. The selected field is an 11 arcmin^2 sky area centered on R.A. = $3^{\rm h}32^{\rm m}39^{\rm s}$, decl. = $-27^\circ47'29''1$ (J2000.0) and resides within the Chandra Deep Field South (CDF-S), a 15' 15' field centered on $\alpha = 3^{\rm h}30^{\rm m}$ and $\delta = -28^\circ.0$.

The observations, taken from September 2003 to January 2004 under two programs (IDs 9978, 10086) for a total of 400 orbits, were divided among the F435W (B_{435}), F606W (V_{606}), F775W (i_{775}), and F850LP (z_{850}) filter to get a limiting magnitude of ~ 29 $\rm m_{AB}$ for point sources. The characteristics of the dataset are listed in Table 5.1.

Filter	# of orbits	# of exposures	Exposure Time [s]	Zero Point [AB mag]
B_{435}	56	112	134880	25.673
V_{606}	56	112	135320	26.486
i_{775}	144	288	347110	25.654
z_{850}	144	288	346620	24.862

Table 5.1: Properties of the HUDF dataset. Col.(1): filter name. Col.(2): number of HST orbits. Col.(3): number of exposures. Col.(4): exposure time in seconds. Col.(5): zero point in AB magnitudes.

5.1.2 The Great Observatories Origins Deep Survey

The Great Observatories Origins Deep Survey (hereafter GOODS) is an astronomical survey combining deep multi-wavelength observations to create a dataset with the aim to explore

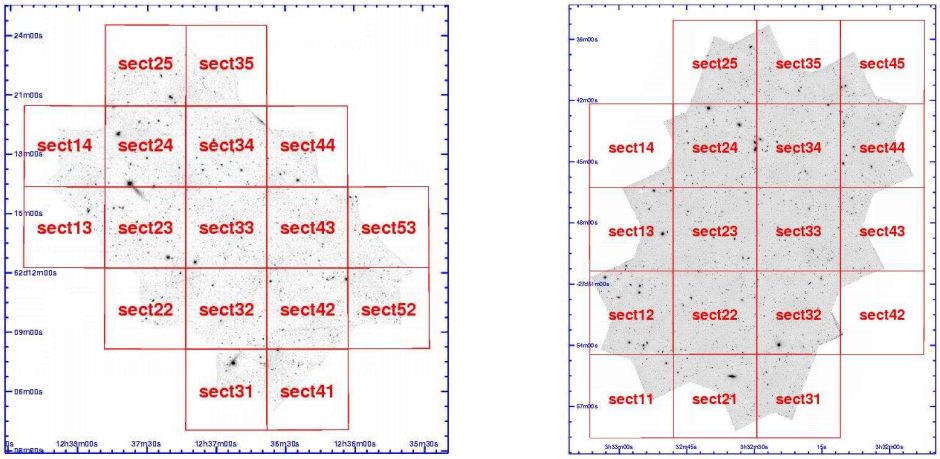


Figure 5.1: GOODS sections layout for the HDF-N and CDF-S field data, respectively on the left and right panel.

the high-z Universe. This survey includes data taken with the Hubble Space Telescope, the Spitzer Space Telescope, and the Chandra X-ray Observatory along with data from other space-based telescopes, such as XMM Newton, and some ground-based telescopes. The survey covers a total of ~ 320 square arcminutes in two fields centered on the Hubble Deep Field North (HDF-N) and the Chandra Deep Field South (CDF-S). The central coordinates for both the fields are listed in Table 5.2.

Field	R.A.(J2000.0)	decl.(J2000.0)
HDF-N	12h36m55s	+62d14m15s
CDF-S	3h32m30s	-27d48m20s

Table 5.2: Central coordinates for the Hubble Deep Field North (HDF-N) and the Chandra Deep Field South (CDF-S) as observed with ACS for GOODS (Giavalisco et al., 2004).

In August 2003 the GOODS Teams released the version v1.0 of the reduced, calibrated, stacked, and mosaiced images acquired with HST/ACS as part of the GOODS ACS Treasury program. This data release includes data acquired as part of the original GOODS HST/ACS program (HST Cycle 11, program IDs 9425 and 9583; Giavalisco et al. 2004) and consists of a full, multi-epoch stacked mosaics of the GOODS data obtained with ACS in 4 bands (F435W, F606W, F775W, and F850LP) for both the HDF-N and CDF-S fields. Each field is divided into sections, i.e. images 8192 x 8192 pixels in size. A total of 17 and 18 sections cover the HDF-N and CDF-S field, respectively. These sections are labeled according to their position in the field. There are two digits number that identify them (Figure 5.1, panel a and b), the first one represents the position of the section in the field along the x-axis, the second one the position along the y-axis. The section located in the lower left corner is section 11, the one in the upper right one is section 45 or section 55 in the CDF-S or HDF-N field, respectively.

Version 2.0 of the data includes the additional data acquired on the GOODS fields during HST Cycles 12 and 13 looking for high-redshift Type Ia supernovae (Program IDs 9727, 9728, 10339, 10340, Riess et al. 2007).

The UDF GOODS dataset we used for the following analysis consists of all the data obtained from the GOODS v1.0 data release located within the UDF pointing[1].
Since a separate mosaic with GOODS v2.0 depth covering just the HUDF sky area did not exist, we combined the single tiles of the v2.0 data release to create the image we needed. In particular, we downloaded from the MAST archive[2] the following images:

h_sb_sect13_v2.0_drz_img.fits	h_si_sect13_v2.0_drz_img.fits
h_sb_sect14_v2.0_drz_img.fits	h_si_sect14_v2.0_drz_img.fits
h_sb_sect24_v2.0_drz_img.fits	h_si_sect24_v2.0_drz_img.fits
h_sb_sect23_v2.0_drz_img.fits	h_si_sect23_v2.0_drz_img.fits
h_sb_sect33_v2.0_drz_img.fits	h_si_sect33_v2.0_drz_img.fits
h_sb_sect34_v2.0_drz_img.fits	h_si_sect34_v2.0_drz_img.fits
h_sv_sect13_v2.0_drz_img.fits	h_sz_sect13_v2.0_drz_img.fits
h_sv_sect14_v2.0_drz_img.fits	h_sz_sect14_v2.0_drz_img.fits
h_sv_sect24_v2.0_drz_img.fits	h_sz_sect24_v2.0_drz_img.fits
h_sv_sect23_v2.0_drz_img.fits	h_sz_sect23_v2.0_drz_img.fits
h_sv_sect33_v2.0_drz_img.fits	h_sz_sect33_v2.0_drz_img.fits
h_sv_sect34_v2.0_drz_img.fits	h_sz_sect34_v2.0_drz_img.fits

and the corresponding **wht.fits* files, i.e. the maps of the inverse variance for each pixel. It should be noted that the tile named "sect23" covers most of the UDF field while the others just give a marginal contribution.

To get the image covering the UDF sky area we started checking the alignment of each tile with respect to the GOODS v1.0 version image of the UDF field, then we cut the regions we were interested in from tiles named "sect13", "sect14", "sect24", "sect33", and "sect34" and inserted them, as well as the entire image of "sect23", in an empty frame with the same dimension of the v1.0 image, i.e. 10500 by 10500 pixels. We cut and combined the weight maps as well so as to get a final image and weight map in each band. Since the exposure times are not uniform along the field, we created an exposure map associated with each image. We made available these images in the MAST archive as High-Level Science Products[3].

The main reason to choose the UDF GOODS dataset instead of the individual sections is that the UDF GOODS dataset relies on a common astrometric grid defined between the UDF and GOODS and is characterized by the same astrometric properties as the HUDF.
Both the v1.0 and the v2.0 GOODS images of the UDF field are characterized by a 30 $\mathrm{mas\ pixel}^{-1}$, have a size of 10500 x 10500 pixels and are North up oriented. We used all the 4 bands available, i.e. B_{435} , V_{606} , i_{775} , and z_{850} , so as to be able to look for LBGs at different redshifts. The corresponding zero points are listed in Table 5.3.

For the following analysis we will consider only the v2.0 of the GOODS dataset since it is slightly deeper than the previous version in the V_{606} , i_{775} , and z_{850} -band.

[1]The UDF GOODS v1.0 data release is available at ftp://archive.stsci.edu/pub/hlsp/udf/goods/

[2]GOODS v2.0 data are available at http://archive.stsci.edu/pub/hlsp/goods/v2/

[3]The UDF GOODS v2.0 data release by V. Calvi is available at ftp://archive.stsci.edu/pub/hlsp/udf/goods2/. For each band the science image (**drz.fits*), weight image (**wht.fits*), and exposure map (**exp_map*) are available.

Filter	Exp Time v1.0 [s]	Exp Time v2.0 [s]	Zero Point [AB mag]
B_{435}	7200	7200	25.65288
V_{606}	5120	5450	26.49341
i_{775}	5120	7028	25.64053
z_{850}	10520	18232	24.84315

Table 5.3: Typical exposure time, in seconds, for the v1.0 and v2.0 GOODS-South dataset and photometric zero points, in AB magnitudes, in the 4 ACS bands as listed in http://archive.stsci.edu/pub/hlsp/goods/v2/h_goods_v2.0_rdm.html.

5.2 Rescaling the RMS Maps

When the images are characterized by variable noise, that is the case of our data, for a proper analysis `SExtractor` (Bertin and Arnouts, 1996) needs weight maps in input, i.e. frames having the same size as the science images that describe the noise intensity at each pixel. `SExtractor` can handle several types of weight-maps through the WEIGHT TYPE configuration keyword. The available options are:

- NONE. If this option is selected, no weighting is applied.
- BACKGROUND. In this case a variance map is computed internally by `SExtractor` from the science image itself.
- MAP RMS. This is the case when the image specified by the WEIGHT IMAGE parameter is a weight map in units of absolute standard deviations.
- MAP VAR. In this case the input WEIGHT IMAGE is a weight map in units of relative variance and `SExtractor` scales it by comparing it the an internal absolute variance map derived directly from the data.
- MAP WEIGHT. This is the case when the **fits* file specified by WEIGHT IMAGE parameter is a weight map in units of relative weights.

We made use of the inverse-variance weight images (**wht.fits*) produced by `MultiDrizzle` to generate the rms maps. In principle the rms map is simply $\frac{1}{\sqrt{wht\ image}}$ but, since the errors on fluxes and magnitudes in `SExtractor` depend directly on the kind of weight image given as input, we carefully checked the rms maps and derived a rescaling factor for each of them. This rescaling factor is needed to take into account the correlated noise introduced during the drizzling process when a kernel different from the "point" one is used (Casertano et al., 2000).

We derived the rescaling factors for each dataset separately. We selected 100 square regions with an area of 7 x 7 pixel^2 with no sources, i.e. just with background signal, from the z_{850} -band science image of each dataset. Then, for all the bands, we computed the mean signal characteristic of a $49\ \mathrm{pixel}^2$ area and the associated rms error. Then, on the same 100 regions, we derived the mean value in the corresponding rms map. For each region we, finally, divided the root mean square of the signal in the science image by the the mean value of the signal in the rms map getting the rescaling factor for the rms map.

Dataset	B_{435}	V_{606}	i_{775}	z_{850}
GOODS v2.0	2.055	3.150	2.900	3.345
HUDF	1.220	2.220	2.280	1.670

Table 5.4: Final rescaling factors in the 4 ACS bands for the GOODS and HUDF datasets. The first estimate for the rescaling factors was derived comparing the typical rms associated to the background signal in the science images and the mean signal derived from the rms map over 100 selected 7x7 pixels regions. Then, we derived the final values following an iterative approach based on the comparison between the typical rms associated to a background region and the `SExtractor` ISOFLUX_ERR derived for objects with the same total area.

The rescaled rms maps were then used as input weight images when running `SExtractor` to get the catalogs of sources in the different fields. As a further check we consider again the 100 background regions selected before and computed the mean flux characteristic of a 7 x 7 pixels sky area and the rms associated to it. Then, we selected from the catalogs all the sources with an ISO_AREA in the range 47-51 pixels^2 and averaged the associated ISO_FLUXERR values.

The two values were supposed to be in good agreement, but we found they were not. This means that `SExtractor` modifies, in some way, even the kind of weight maps that are not supposed to be adjusted.

To overcome this discrepancy and ensure that the `SExtractor` errors for each object are not underestimated we operated iteratively changing the rescaling factor to get a new rms map, re-running `SExtractor` with this new rms map and, finally, comparing again the two values until we got a difference between them that was less than 1%. The final rescaling factors for each filter and dataset are recorded in Table 5.2. Note that a rescaling is needed also for the HUDF i_{775} and z_{850} -band images that were obtained with the point kernel and are, thus, in principle safe from correlated noise components. This was noticed already by Oesch et al. (2007).

5.3 Source Selection

To study the possible effect of cosmological dimming we need to compare the magnitude of the same high-redshift sources derived from surveys that imaged the same sky area but that are characterized by different depth. Among our choice of data, the GOODS dataset is the shallowest and essentially determines the the kind of sources we could study, as well as their number, since we are interested in objects detected in all the surveys.

High-z galaxies in deep surveys are usually selected on the basis of the Lyman-break technique. Dropout galaxies show no flux at wavelengths shortward of the rest-frame Lyα line due to strong absorption by intergalactic Hydrogen. This is strictly true only for galaxies at $z \gtrsim 6$, while in the range $3 \lesssim z \lesssim 6$ there is still some flux in the region of the Ly-α forest. At $z \sim 4$ this results in a lack of flux in the B_{435} -band, i.e. no detection, associated with a clear detection in the redder bands, in our case V_{606} , i_{775} , and z_{850} .

We ended up considering only the B_{435} -dropout population, i.e. candidate high-z galaxies at $z \sim 4$, and we discarded the analysis on V_{606} and i_{775} -dropouts because the

two samples included only two and one candidate, respectively, and were, indeed, not statistically significant.

A comprehensive study on LBGs at $z \sim 4-5$, including the determination of the fraction of interloper entering the sample, was conducted by Huang et al. (2013) who studied the bivariate size-luminosity relation using HUDF and GOODS data.

We selected our sources by running `SExtractor` version 2.8.6 (Bertin and Arnouts, 1996) in dual image mode within each dataset using the z_{850} -band image for the detection and performing the photometry on all the four bands. We used the same `SExtractor` parameters stated in Beckwith et al. (2006) since they were optimized for the pixel scale and PSF typical of the datasets. The threshold for detection and analysis (DETECT_THRESH and ANALYSIS_THRESH) was set to be 0.61 within an area of at least 9 contiguous pixels (DETECT_MINAREA). We selected a total of 32 deblending subthresholds (DEBLEND_NTHRESH), a contrast parameter of 0.03 (DEBLEND_MINCONT), and a Full Width Half Maximum (SEEING_FWHM) of $0''.09$. Moreover we used the *goods.conv*[4] filter (FILTER_NAME) optimized for GOODS data.
We set the WEIGHT_TYPE to MAP_RMS and WEIGHT_IMAGE to be the final rms map derived applying the rescaling factors listed in Table 5.2. Since the detection is affected by the rescaling factor as well, within each dataset we rescaled the 0.61 value we selected as DETECT_THRESH and ANALYSIS_THRESH according to the rescaling factor in the z_{850} -band. We made use of the following color selection criteria, as stated in Beckwith et al. (2006):

$$\begin{aligned} B_{435} - V_{606} &> (1.1 + V_{606} - z_{850}) \\ B_{435} - V_{606} &> 1.1 \\ V_{606} - z_{850} &> 1.6 \\ (\frac{S}{N})_{V_{606}} &> 5 \\ (\frac{S}{N})_{i_{775}} &> 3 \end{aligned} \tag{5.2}$$

The colors are calculated from the MAG_ISO parameter in each filter and the detection S/N is derived from the FLUX_ISO and FLUX_ERR_ISO parameters according to Stiavelli (2009)

$$\frac{S}{N} = \frac{FLUXISO}{FLUXERRISO} \tag{5.3}$$

These criteria permit to select candidates that lie in the redshift range $3.4 \lesssim z \lesssim 4.4$ and, according to Huang et al. (2013), the fraction of lower redshift interlopers is below 8%.

After obtaining the catalogs of B_{435} -dropouts existing in each dataset we visually inspected all the source in these lists and rejected the spurious detections. The final lists contain 119, 74, 428, ,222 dropout galaxies for the GOODS v1.0, GOODS v2.0, HUDF, and XDF dataset, respectively. It should be noted, anyway, that the area covered by the GOODS v1.0 and GOODS v2.0 images of the HUDF sky area is slightly bigger that the HUDF and the XDF sky areas, so part of the dropouts lie in regions of the image where there is no signal in the HUDF or the XDF images.

[4]The filter named *goods.conv* is a 9x9 pixel^2 convolution mask of a Gaussian PSF with FWHM = 5.0 pixels.

5.4 Results

There are 67 matching sources in the GOODS v1.0 and GOODS v2.0 data and 173 in HUDF and XDF images. We used these samples of matched B_{435} -dropouts to check the photometry performed by `SExtractor` since we did not expect any significant variation between data that have a similar depth.

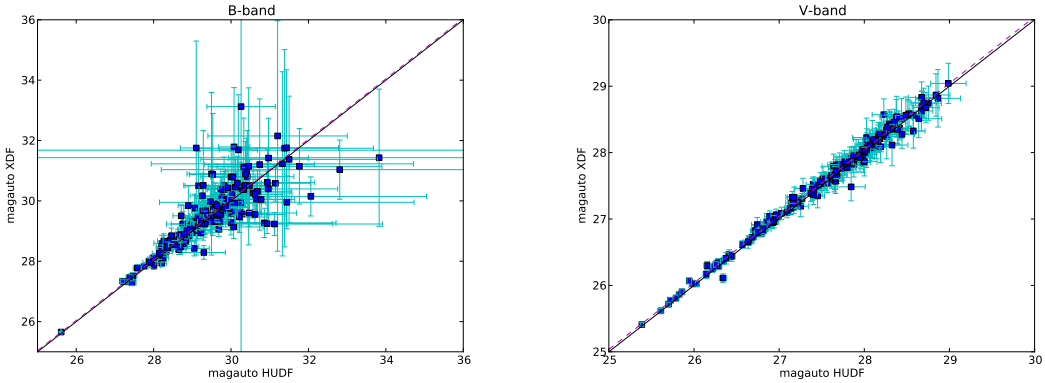


Figure 5.2: Total magnitude in the B_{435} (left panel) and V_{606} -band (right panel) derived from the HUDF and XDF dataset for a sample of 173 B_{435} -dropouts detected in both the surveys. The error bars associated with the B_{435} -band magnitudes are huge because the B_{435} -band is the dropout band, that is expected to show very little flux from the dropout galaxies.

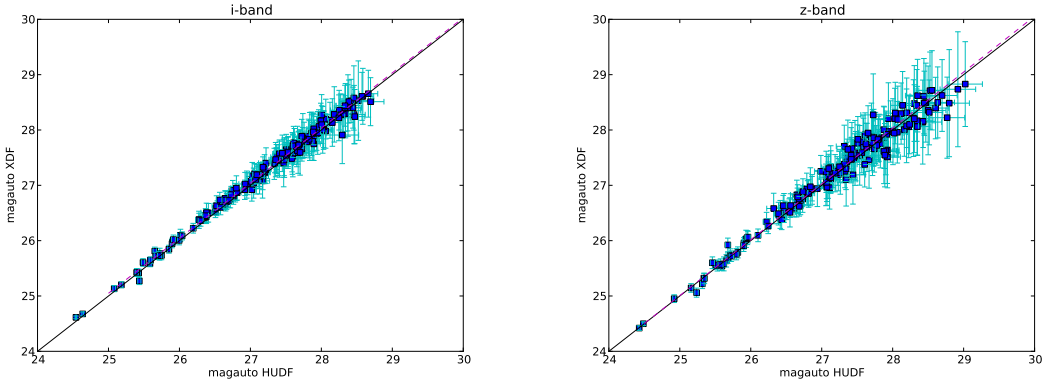


Figure 5.3: Total magnitude in the i_{775} (left panel) and z_{850} -band (right panel) derived from the HUDF and XDF dataset for a sample of 173 B_{435} -dropouts detected in both the surveys.

After testing the photometry we moved trough the proper analysis. Comparing the HUDF data with GOODS v2.0 we obtained a sample of 21 matching objects within the HUDF footprint that were selected as B_{435} -dropout candidates independently in the two datasets. Considering the XDF data the number of matches was 24 for the GOODS v1.0 and 20 for the GOODS v2.0 dataset. It should be noted that the number of matching objects is much smaller than the one in GOODS because the sky area covered by the GOODS image is about

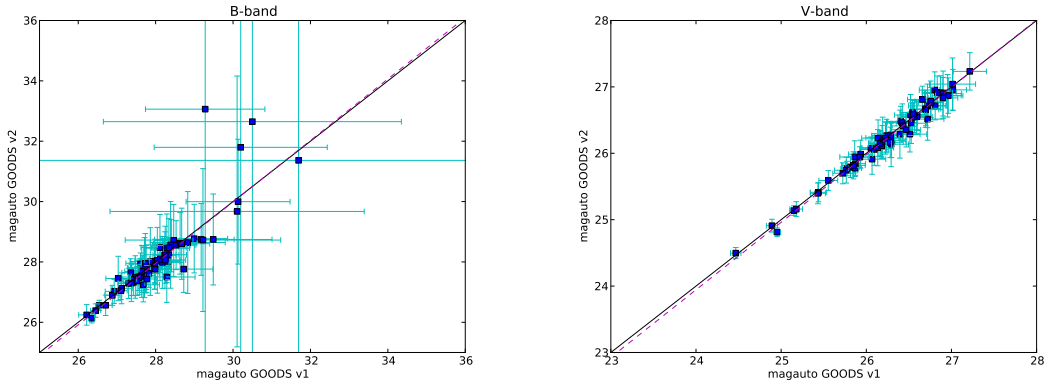


Figure 5.4: Total magnitude in the B_{435} (left panel) and V_{606} -band (right panel) derived from the GOODS v1.0 and v2.0 for a sample of 67 B_{435} -dropouts detected in both the surveys. The error bars associated with the B_{435} -band magnitude are huge because the B_{435} -band is the dropout band, that is expected to show very little flux from the dropout galaxies.

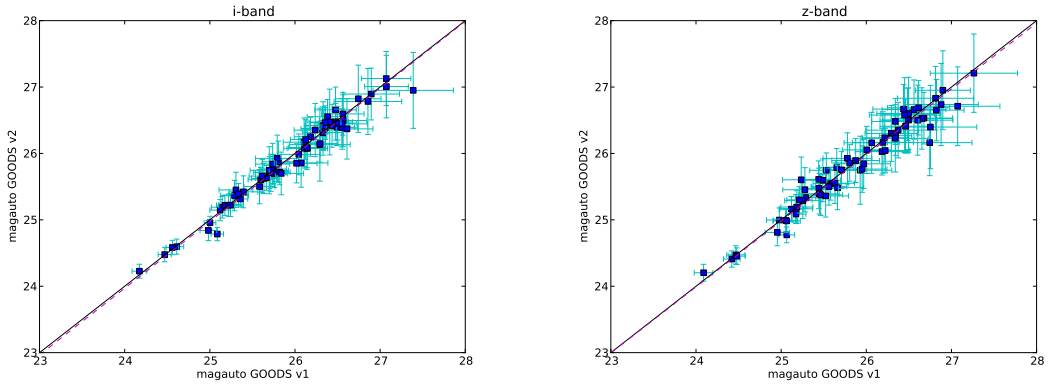


Figure 5.5: Total magnitude in the i_{775} (left panel) and z_{850} -band (right panel) derived from the GOODS v1.0 and v2.0 dataset for a sample of 67 B_{435} -dropouts detected in both the surveys.

twice the size of the HUDF footprint and most of B_{435} -dropout candidates lie outside the HUDF sky region.
We plotted the total magnitudes and errors, given by `SExtractor` MAG_AUTO and MAGERR_AUTO parameters, respectively, and derived the best linear least-squares fit in the form

$$y = A + B \cdot x \tag{5.4}$$

considering the error-bars in both coordinates using the `IDL`[5] `fitexy.pro` routine. In Equation 5.4 x and y are MAG_AUTO values in GOODS-depth and HUDF-depth, respectively. The best fit parameters are tabulated in Table 5.6 and 5.7.

[5]Interactive Data Language is distributed by Excelis Visual Information Solutions

As a further check, we investigated the behavior of all the B_{435} -dropouts selected in HUDF, with no regards for their identification in GOODS v2.0. In particular, we selected all the B_{435} -dropouts detected in HUDF with a total magnitude in the z_{850} -band brighter than $z_{850} = 26.8$ mag, value that defines the completeness of the GOODS dataset (50% complete). 79 B_{435} -dropouts satisfy this criterion and we focused in them for the following analysis. Discarding the cuts in S/N for the GOODS data, there are 75 objects selected as B_{435} -dropouts in HUDF that have a detection in GOODS v2.0, two of them were discarded since each of them is clearly detected as two distinct sources in GOODS. Out of the 73 remaining objects we did a separate analysis considering different cuts in S/N, in details $S/N > 2$, $S/N > 3$, $S/N > 4$, $S/N > 5$. For each selection we derived the best fit, as done for the sample of matching B_{435} -dropouts.
The left panels of Figure 5.6 summarize our results plotting the total magnitudes in GOODS and HUDF on the basis of the S/N cut. As can be seen from Table 5.4 there is no significant change in the trend of the linear fit, even including objects with lower S/N detection, showing no effect due to cosmological dimming. When including the objects with very poor or no detection in at least one of the bands used to identify objects there could be an effect due to SB dimming, but we are not able to disentangle this effect from the one simply due to the detection limit of GOODS.

Selection	V_{606}			i_{775}			z_{850}		
Criteria	# obj	A	B	# obj	A	B	# obj	A	B
$S/N > 5$ + color	21	-0.88	1.03	21	-0.28	1.01	21	-0.45	1.02
$S/N > 5$	32	-1.04	1.04	36	-0.59	1.02	20	-1.13	1.04
$S/N > 4$	46	-1.13	1.04	45	-0.46	1.02	38	-0.48	1.02
$S/N > 3$	58	-0.93	1.03	57	-0.17	1.01	51	-0.67	1.03
$S/N > 2$	65	-0.86	1.03	68	0.30	0.99	66	-0.51	1.02
All	73	-0.58	1.02	73	0.56	0.98	73	-0.03	1.00

Table 5.5: Parameters for the best linear least-square fit in the form $y = A + B \cdot x$. A and B were derived comparing the total magnitudes in the V_{606} , i_{775} , and z_{850} -band for the sample of matching sources taking into account the error bars associated to the magnitude values. Different rows refer to different selection criteria and, consequently, the number of matching objects varies.

Datasets	B_{435}		V_{606}	
	A	B	A	B
GOODS v1 - GOODS v2	-0.55	1.02	-0.45	1.02
GOODS v2 - HUDF	-7.68	1.29	-0.93	1.034
GOODS v2 - XDF	-10.40	1.39	-0.80	1.03
HUDF - XDF	0.025	1.00	0.00	1.00

Table 5.6: Parameters for the best linear least-square fit, $y = A + B \cdot x$, derived comparing the total magnitudes in the B_{435} and V_{606} -band for the sample of matching sources in different datasets.

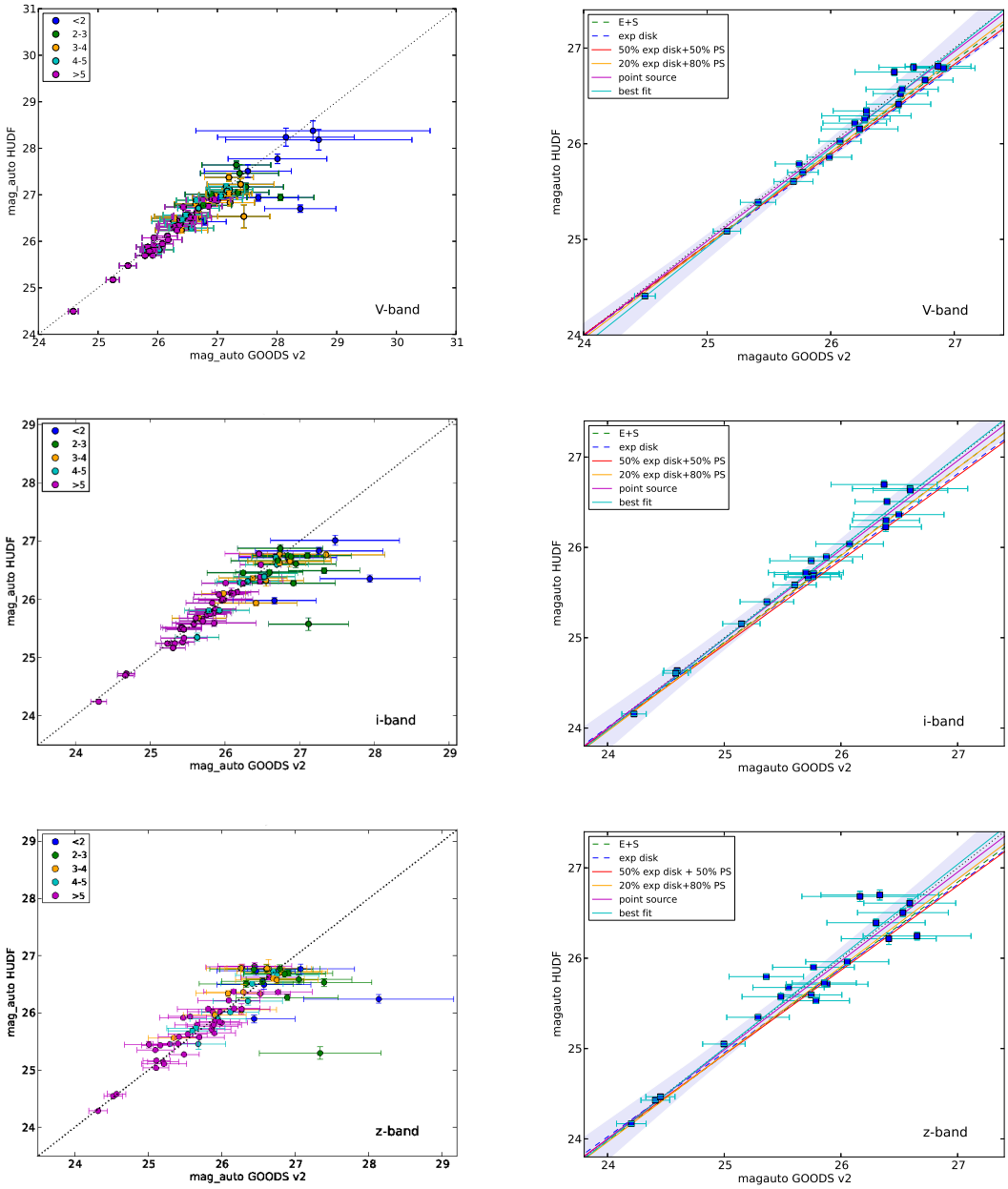


Figure 5.6: Total magnitude derived from SExtractor MAG_AUTO parameter in the V_{606} (top panels), i_{775} (central panels), and z_{850} -band (bottom panels) from the GOODS v2.0 dataset compared to the one derived from the HUDF. Left panels refer to all the objects selected in HUDF as B_{435} -dropouts that have a detection in GOODS v2.0. As tabulated in the legend, different colors refer to different cuts in S/N in GOODS if no color selection is applied in GOODS. Right panels show the sample, consisting of 21 B_{435} -dropouts, satisfying all the selection criteria in both HUDF and GOODS v2.0. The colored lines indicate the trend if cosmological dimming is acting depending on the way the flux is distributed among the galaxy, while the dotted black line shows the 1:1 trend. The shaded area shows the 95% confidence band around the best fit (solid cyan line).

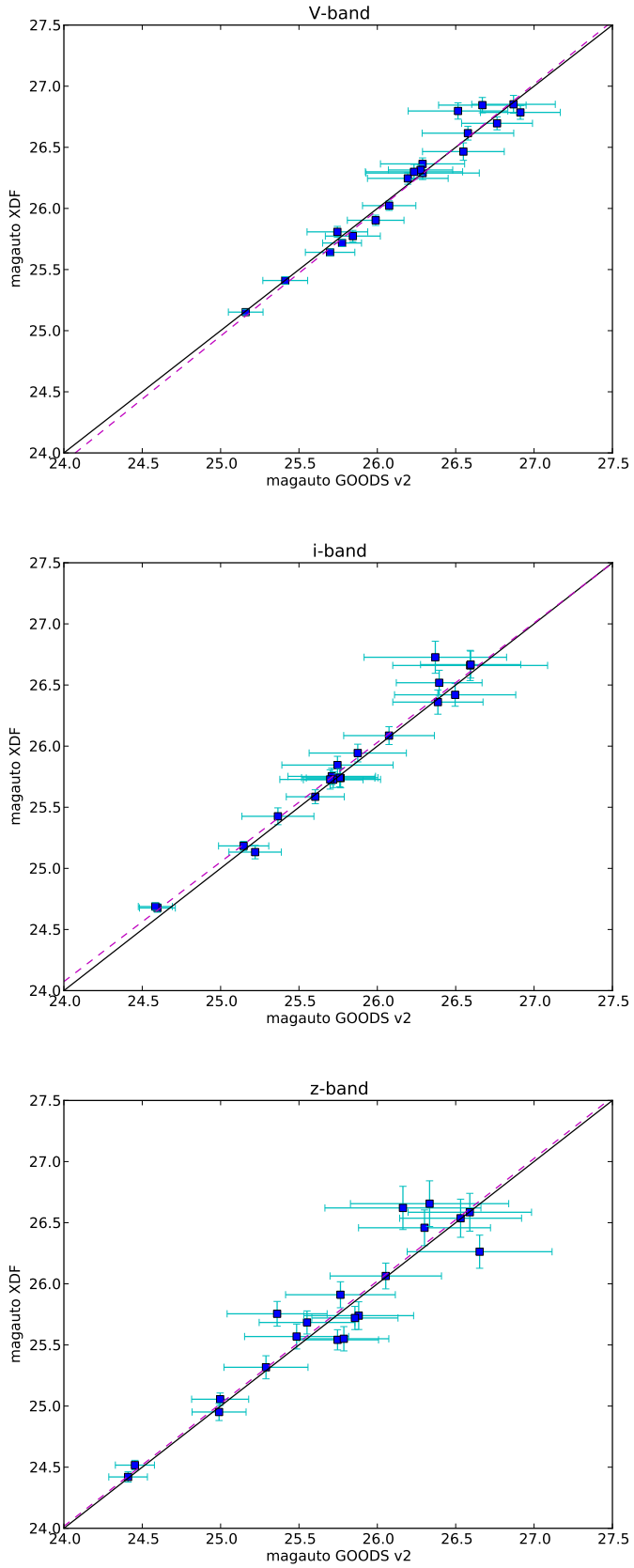


Figure 5.7: Total magnitude in the V_{606} (top panel), i_{775} (central panel), and z_{850} -band (bottom panel) derived from the GOODS v2.0 and XDF datasets for a sample of 20 B_{435} -dropouts detected in both the surveys.

Datasets	i_{775}		z_{850}	
	A	B	A	B
GOODS v1 - GOODS v2	-0.15	1.00	0.145	0.99
GOODS v2 - HUDF	-0.57	1.02	-1.25	1.05
GOODS v2 - XDF	0.60	0.98	-0.05	1.00
HUDF - XDF	0.15	1.00	-0.24	1.01

Table 5.7: Parameter for the best linear least-square fit, $y = A+B\cdot x$, derived comparing the total magnitudes in the i_{775} and z_{850} -band for the sample of matching sources in different datasets. In the fit equation x and y are MAG_AUTO values in GOODS-depth and HUDF-depth, respectively.

5.5 Simulations

Band	GOODS v1.0	GOODS v2.0	Simulations
V_{606}	0.0019	0.0018	0.00185
i_{775}	0.0016	0.0014	0.0015
z_{850}	0.0009	0.0008	0.0085

Table 5.8: Typical background rms values for GOODS v1.0 (Col. 2) and v2.0 (Col. 3) derived using the `IRAF IMEXAMINE` routine on the ACS scientific images in the V_{606} , i_{775} , and z_{850} -band. In Col. 4 we listed the rms values used to reproduce the random noise component in our Monte Carlo simulations.

Our results suggest no effect due to cosmological dimming when comparing deep sky surveys with different depth. To complete our study we wanted to show that the difference in depth between our datasets would be enough to detect Tolman dimming effects if galaxies were characterized by a smooth light distribution at the resolution of HST. To this aim we ran Monte Carlo simulations reproducing the B_{435} -dropout sample in order to derive the expected trend in the magnitude-magnitude plot in the case of cosmological dimming affecting the SB of the sources.

In detail, we created a 4000 x 4000 pixels image where we inserted 1000 mock galaxies with apparent magnitudes in each band ranging from 24.5 to 28.5 mag, axis ratio b/a, and effective radius $\mathrm{r_e}$ distributed according to the findings of Ferguson et al. (2004), position angle PA and x,y position randomly drawn from a flat distribution. We used the `noao.artdata.mkobject` package in `IRAF`[6] to make the artificial galaxies.

In the first set of simulations we reproduced the images in each band assuming all the galaxies to have just one component. First we created a mock image containing only disk-like, i.e. late-type, with a surface brightness exponential profile. According to Ferguson et al. (2004) among the B_{435} -dropout sample 78% of the galaxies are classified as late-type. We convolved each mock frame with the corresponding PSF obtained running `Tiny Tim`[7]. This point spread function modeling tool permits to model the PSF specific to a chosen

[6]IRAF is distributed by the National Optical Astronomy Observatories, which are operated by the Association of Universities for Research in Astronomy, Inc., under cooperative agreement with the National Science Foundation.

[7]The Tiny Tim web interface is available at http://www.stsci.edu/hst/observatory/focus/TinyTim

channel for any of the HST imagers and bands. We obtained three different PSF models, i.e for the V_{606} , i_{775} , and z_{850} -band. Finally we added a random noise frame with the same rms as the one measured on the GOODS science image in the corresponding frame. In particular the rms used for the simulation is averaged over the rms values obtained from the v1.0 and v2.0 in each band listed in Table 5.8 and obtained from the `IRAF IMEXAMINE` tool.

We also created a set of mock images assuming 25% of the galaxies to have a De Vaucouleurs profile (i.e. $r^{1/4}$), and the remaining an exponential one. This is broadly compatible with Ferguson et al. (2004). As previously done, we convolved all mock objects with the PSF by `Tiny Tim` and added the random noise. In both cases, i.e. exponential and mixed profiles, the effect of surface brightness dimming is expected to be large since the objects are extended and the flux is distributed in extended structures.

In the third set of simulations we inserted mock galaxies that have a double component. Specifically, we used an exponential disk and a central point source. We used the same PSF we got from `Tiny Tim` to convolve the frames as the model for the point source and we convolved only the exponential disks with the typical ACS PSF. Then, as done for the simulations previously described we added the random noise reproducing the background of the science image in the selected band. In detail, we simulated galaxies with 80% of the flux enclosed in the central point source and 20% in the extended exponential disk or with the flux equally splitted among the two components.

Finally, we run the last set of simulations modeling each mock galaxy with the PSF by `Tiny Tim`, i.e. we created point sources that are not expected to be affected by cosmological dimming. Of course in this case we did not convolve the synthetic frame with the PSF before adding the background noise.

For each set of simulation we performed 50 iterations for each band. Then, we run `SExtractor` on each mock image with the same parameters used for the detection in the real data and we compared the output catalogs with the characteristics of the mock galaxies inserted in the corresponding mock frame. Each realization of the simulation lead to a line describing the expected trend, then we averaged the A and B values over all the 50 realizations in order to derive the best fit parameters listed in Table 5.5.

To determine whether or not the trends derived from the simulations are in agreement with our data we computed the normalized residuals for each simulation. In detail, we derived the distance between each data point and the model and divided it by the corresponding error. For compact sources, for which no significant dimming is expected, we predicted the distribution of the residuals to resemble the one obtained with respect to the 1:1 line. We performed a Kolmogorov-Smirnov non-parametric statistical test comparing the distribution of the residuals derived from each set of simulations and the one from the data. The histograms showing the distribution of the residuals are plotted in Figure 5.8. The values of the maximum vertical deviation between the two cumulative curves, namely the D parameter, and the corresponding probability p[8] are listed in Table 5.5. If the p value is small, the two groups were sampled from populations with different distributions. We can reject the null hypothesis (i.e. the two distributions are the same) with a 95% confidence for all the realization characterized by $p < 0.05$. As can be seen the expected trend in case of flux distributed in extended sources ($r^{1/4}$ and exponential profile) can be

[8]In the Kolmogorov-Smirnov test p is the probability of two random samples having a D parameter as large or larger than observed.

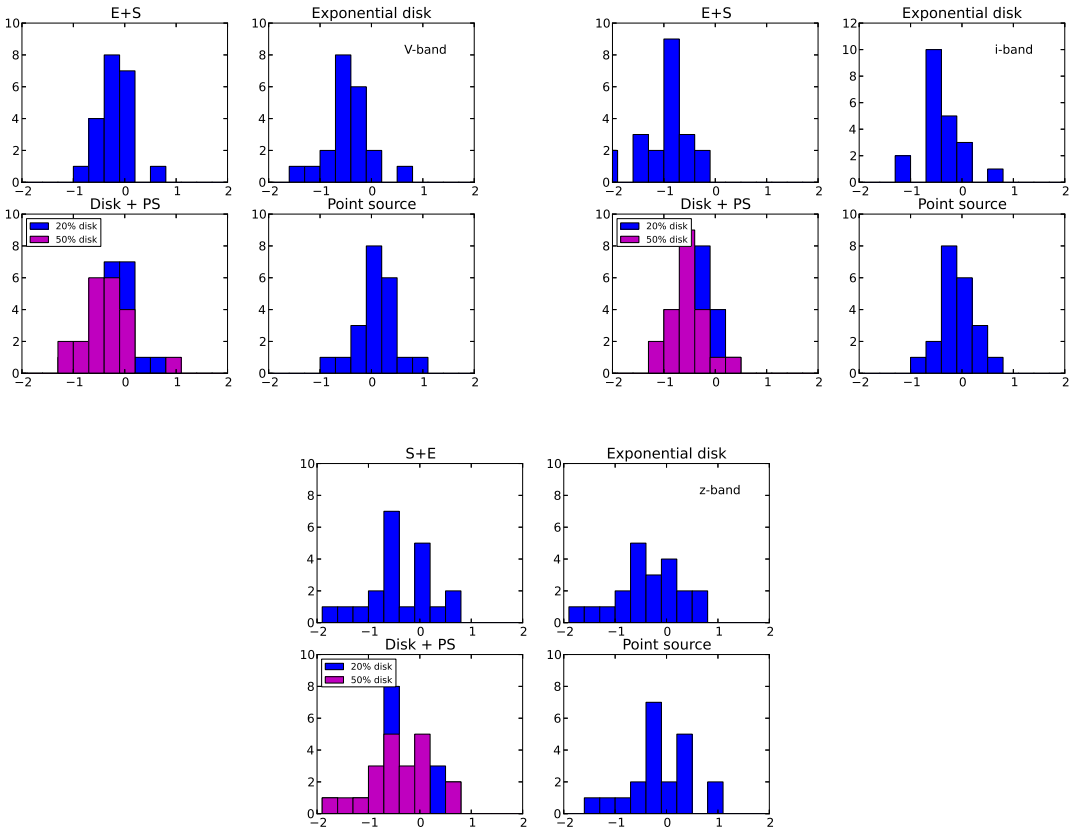


Figure 5.8: The blue histograms show the distribution of the normalized residuals for each simulation (labeled on top of each panel) for the B_{435} -dropout sample. For each data point we computed the distance from the model, then, it was divided by the corresponding error. The three figures refer to the V_{606} , i_{775} , and z_{850} -band respectively. In the case of mixed profiles the blue histogram refers to the case with 80% of the flux enclosed in the central point source and 20% in the extended exponential disk and the magenta one to the case where flux is equally splitted among the two components.

easily ruled out, while the trend for point sources is in agreement with our findings, suggesting that the light from high-z sources mostly comes from compact objects rather than extended ones. A compact structure at high-z, at least in the bluer bands, is in agreement with the results obtained by Williams et al. (2014) for $z \sim 3$ LBGs. It should be noted that diffuse models are ruled out to lower significance in the z_{850} -band and this could hint at a diffuse component starting to be more significant in the redder bands.

As a further result, the simulations involving point sources permitted us to verify the findings of Ashby et al. (2013) who claimed that the total magnitude recovered by `SExtractor` is, on average, 0.05 mag fainter than the real value. We found all the best fit lines derived for the point source simulations to be slightly shifted in the x direction and that, introducing the the 0.05 mag correction stated by Ashby et al. (2013), the fit was closer to the expected trend determined by the 1:1 line.

	Mix (E+S)		exp profile		point source		20% exp profile		50% exp profile	
Band	A	B	A	B	A	B	A	B	A	B
V_{606}	0.70	0.97	1.48	0.94	0.20	0.99	0.59	0.97	1.39	0.94
i_{775}	0.74	0.97	1.54	0.94	0.33	0.99	0.64	0.97	1.49	0.94
z_{850}	1.01	0.96	1.67	0.93	0.39	0.98	0.69	0.97	1.75	0.93

Table 5.9: Best linear least-square fit parameters derived for each of the three detection bands averaging over 50 realizations of Monte Carlo simulations. We considered different scenarios on the basis of the distribution of the total flux: a mix of galaxies (25% with a De Vaucouleurs $r^{1/4}$ SB profile and 75% described by an exponential profile), an exponential disk, a point source modeled using the PSF image created using `Tiny Tim`, and, finally, multi-component objects with the flux divided into an exponential disk and a point source modeled using `Tiny Tim` PSF. In particular we considered two cases: either the light equally splitted among the disk and point source components or 20% of the light in the exponential disk and 80% in the central point source.

	V_{606} -band		i_{775} -band		z_{850} -band	
Profile	D	p	D	p	D	p
Mix (E+S)	0.57	1.08e-03	0.43	0.03	0.42	0.03
Exponential disk	0.67	7.13e-05	0.57	1.08e-03	0.33	0.15
50% Disk + 50% PS	0.62	2.90e-04	0.66	7.10e-05	0.43	0.03
20% Disk + 80% PS	0.52	3.59e-03	0.48	0.01	0.38	0.07
Point source	0.24	0.53	0.29	0.30	0.14	0.97

Table 5.10: Results from the Kolmogorov-Smirnov test obtained comparing the distribution of the residuals derived from each set of simulations and the one from the data. The values of the D parameter and the corresponding probability p are listed for all the models and bands, as labeled. When $p < 0.05$ the distribution of the residuals differs from the expected one with a 95% confidence, so the model can be ruled out as a possible fit for our data.

5.6 Conclusions

To study the effects of surface brightness dimming on high-redshift galaxies we compared catalogs of candidate B_{435} -dropouts obtained running `SExtractor` on deep images taken with ACS as part of the GOODS, HUDF, and XDF projects. In particular, considering the total magnitude, given by the MAG_AUTO parameter, we found no systematic differences between GOODS and the deeper images. In detail, our sample of B_{435} -dropouts suggests that there is no trend in the magnitude of galaxies becoming brighter in the HUDF, exactly as stated by Beckwith et al. (2006).

If most of the light of faint dropout galaxies was in compact knots one would expect little or no dimming. In contrast if a substantial fraction of the light is in a diffuse component such components would be detected in part in the UDF and missed in GOODS. Our direct comparison shows that galaxies detected in GOODS do not become significantly brighter in the UDF. This suggests that most of their light is compact , as claimed by Bouwens et al.

(2004a) who found that the principal effect of depth in a galaxy survey is to add galaxies at fainter magnitudes, and not to significantly add to the sample galaxies with bigger sizes.

Our findings are in agreement with those by Williams et al. (2014) who focused on $z \sim 3$ LBGs as progenitors of early type galaxies (ETGs). Studying the morphology of these galaxies in the rest-frame UV and in the optical they found a very compact main structure and basically no diffuse component even in deep stacks.
Moreover, comparing the data to simulations reproducing the effect of cosmological dimming for different SB profiles we found that our trends can not be ascribed to extended structures, such as $r^{1/4}$ profiles or exponential disks, but diffuse models are ruled out to lower significance in the reddest band, i.e. z_{850} -band. This could hint at a diffuse component starting to be more significant.

The study of cosmological SB dimming is also important since it could affect our prediction of what JWST can observe at higher redshifts, where younger galaxies may exhibit a larger fraction of clumpiness. Our direct comparison shows that galaxies detected in GOODS do not become significantly brighter in the HUDF. This suggest that most of their light is compact and hints to the fact that JWST will likely not find extensive diffuse star-forming components. This is also promising as it suggests that JWST will not be confusion limited at least in the shortest bands.

Whether or not galaxies have a red diffuse component remains an open question and will require further investigations.

Chapter 6

A Preliminary Study on the Lower-redshift Contaminant Population Entering the Dropout Samples at $z \sim$ 4-5-6

In collaboration with Dr. Michele Trenti and Dr. Pascal Oesch

Each sample of dropout galaxies identified through the Lyman-break technique is affected by contaminants. There are two classes of interlopers that can enter any color-based high-redshift selection: cool galactic stars and galaxies at lower redshift (Stanway et al., 2008), but any HST-based selection is mostly affected by low-redshift galaxies entering the color-color selection box due to photometric scatter.
Surprisingly, even though lower-redshift interlopers are a well known issue when dealing with the Lyman-break selection, very little has been done till now to characterize these contaminants and in the vast majority of papers a contaminant fraction based just on Monte Carlo simulations, as those presented in Pirzkal et al. (2013), is assumed.

The goal of this Chapter is to lay the foundation to characterize the luminosity function of low-redshift objects that contaminate the samples of high-redshift dropouts. To this aim we made use of two different multi-wavelength catalogs and applied tailored color selection criteria to identify dropout galaxies at $z \sim 4-5-6$. Spectroscopic follow-ups, like those presented in Vanzella et al. (2009) and Stark et al. (2010), showed that a two-color selection is very efficient in identifying large numbers of high-redshift galaxies, but a purely color-based selection is not contaminant free, as fully discussed in Stanway et al. (2008).

In the following, we will focus on the GOODS-South field looking for galaxies that could enter the high-z galaxy sample due to photometric scattering. In particular, we will get catalogs of reliable dropout candidates as well as of these two kinds of interlopers:

1. sources that satisfy the dropout color-color selection criteria, but show a (faint) detection in bands blueward of the Lyman-break;
2. objects that lie in the region next to the color-color selection box.

We would like to investigate whether or not the LFs of the contaminants and dropout galaxies are different and how the number of interlopers evolves with Lyman-break selection at

different redshifts.

6.1 CANDELS GOODS-S Multi-wavelength Catalog

For our first attempt to estimate the LF of contaminants we made use of the CANDELS GOODS-S Multi-wavelength Catalog (Guo et al., 2013)[1] that contains a total of 34930 sources. We focused only on HST observations obtained in the F435W, F606W, F775W, F850LP (ACS data on GOODS and CANDELS fields), F105W, F125W, and F160W (WFC3/IR data on CANDELS and HUDF fields) bands. The source detection with `SExtractor` was performed in the H_{160} -band and then they obtained the photometry running `SExtractor` in dual-image mode on PSF-matched HST images.

Regarding the detection in the H_{160} -band, we rejected all the sources with quality flags (i.e. objects detected on star spikes or halos, at the image edges or on the artifacts) or missing data in one or more bands. In this way, we got an homogeneous sample of objects consisting in 21817 sources.

6.1.1 Dropouts Selection

The dropout selection was performed applying the color selection criteria by Bouwens et al. (2014) and, then, requiring a detection with $S/N > 5$ or $S/N > 8$ in the H_{160} -band. The requirements for V_{606} -dropouts are:

$$\begin{aligned} V_{606} - i_{775} &> 1.2 \\ z_{850} - H_{160} &< 1.3 \\ V_{606} - i_{775} &> 0.8 \cdot (z_{850} - H_{160}) + 1.2 \\ (S/N)_{B_{435}} &< 2 \end{aligned} \tag{6.1}$$

The color selection criteria to select the i_{775} -dropouts are, instead, the following:

$$\begin{aligned} i_{775} - z_{850} &> 1.0 \\ Y_{105} - H_{160} &< 1.0 \\ i_{775} - z_{850} &> 0.777 \cdot (Y_{105} - H_{160}) + 1.0 \\ (S/N)_{B_{435}} &< 2 \\ (S/N)_{V_{606}} &< 2 \end{aligned} \tag{6.2}$$

We, also, required all the sources in each dropout catalog to not satisfy the color-color criteria for the sample just at higher redshift.

As a further step, since galaxies lying just next to the color selection box are likely to enter the dropout sample due to photometric scatter, we enlarged the cuts by 0.2 mag for both the $S/N > 5$ and $S/N > 8$ selected samples. The number of candidate dropouts and contaminants is listed in Table 6.1.1. The distribution of candidate high-z galaxies and of the two classes of contaminants in the color-color plot are shown in Figures 6.1 and 6.2. Most of the contaminants lie next to the boundaries of the selection box, supporting the idea that they could meet the selection criteria due to photometric scatter.

[1] http://archive.stsci.edu/prepds/candels/GOODS-S.html#gdsscat

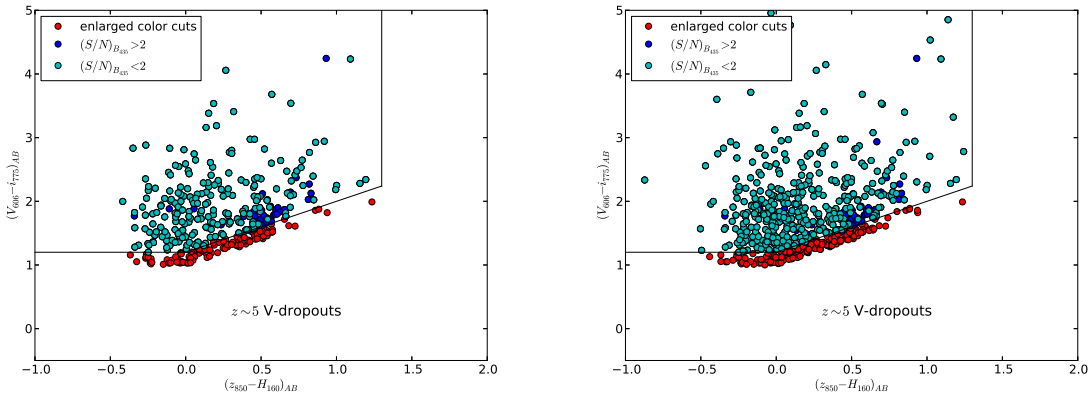


Figure 6.1: Location of the candidate V_{606} -dropout galaxies (cyan dots) and contaminants with respect to the color-color selection box defined by Bouwens et al. (2014). Blue dots indicate those contaminants that lie within the selection box, red dots those just outside the selection box that are likely to meet the dropout selection criteria due to photometric scattering. Left and right panel refer to objects selected on the basis of a different S/N cut in the H_{160} -band, i.e. $S/N > 8$ or $S/N > 5$, respectively.

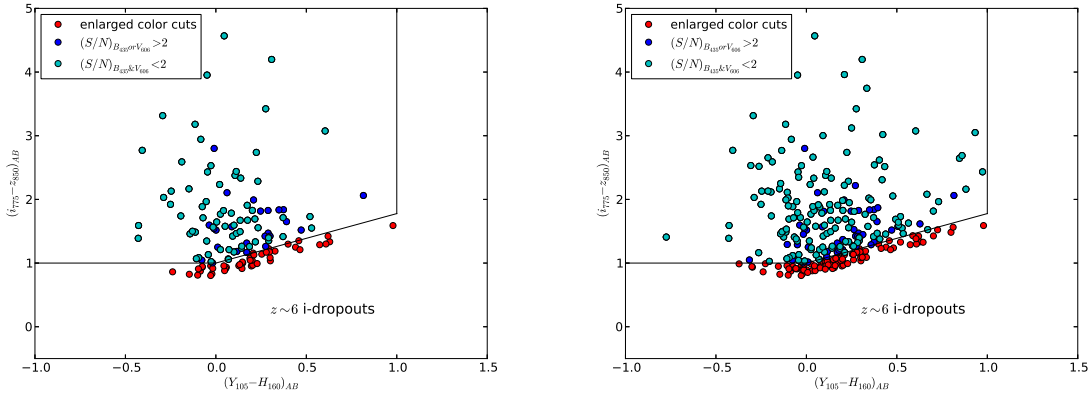


Figure 6.2: The same as in Figure 6.1, but for candidate i_{775} -dropouts. Left and right panel refer to objects selected on the basis of $S/N > 8$ or $S/N > 5$ in the H_{160} -band, respectively.

	$(S/N)_{H_{160}} > 8$			$(S/N)_{H_{160}} > 5$		
Dropouts	Enlarged color cuts	No S/N cut	Blueward S/N cut	Enlarged color cuts	No S/N cut	Blueward S/N cut
V_{606} -dropouts	372	262	223	617	435	392
i_{775} -dropouts	136	92	67	285	193	149

Table 6.1: Number of dropouts identified in the catalog requiring a $S/N > 8$ or $S/N > 5$ in the H_{160} -band and using the color cuts by Bouwens et al. (2014) with or without the S/N cut ($S/N < 2$) in the bands blueward of the Lyman-break.

6.1.2 Preliminary Results

We investigated the behavior of the two classes of contaminants previously introduced, i.e. sources showing a non-negligible detection in the bands blueward of the Lyman-break feature and sources lying next to the color-color selection box.
Figures 6.3 and 6.4 show the results of our study considering the dropouts samples obtained assuming $S/N > 8$ or $S/N > 5$ in the H_{160} -band, respectively. It should be noted that the $S/N > 8$ requirement is more conservative since, as shown by Schmidt et al. (2014) the noise in science images is not distributed according to a Gaussian distribution. On the basis of that, a source detected with a $S/N = 5$ does not imply a real 5σ detection.
The number of contaminants seems to be similar when selecting both V_{606} and i_{775} -dropout galaxies and the number counts distribution of contaminants is flatter compared to the dropouts one.
The number counts suggest that the luminosity functions are different for low-z contaminants and real dropouts and that the LF of the contaminants is slowly varying across the two selections, while the dropouts counts vary by over a factor of 3.
Regarding the two different classes of contaminants we found that:

- their number and distribution are quite similar in the i_{775} -dropouts sample and significantly different from the one of real dropouts;
- in the V_{606} -dropouts sample interlopers showing an optical detection are less than contaminants outside the color-color selection box.

6.2 3D-HST Catalogs

For an independent study on contaminants we did a similar analysis using the 3D-HST catalogs v4.1[2]. We chose the GOODS-SOUTH field catalog consisting in 50507 sources presenting photometric data in HST, ground-based, and Spitzer/IRAC bands.
We built up our sub-catalog including only photometric data in the following bands: U38, U[3], F435, F606, F775, F814W, F850, F125, F140, F160.
First of all, we selected all the sources satisfying usephot = 1, a flag that permits to have a sample of objects characterized by uniform quality in the photometry (Skelton et al., 2014).

[2]http://3dhst.research.yale.edu/Data.php
[3]U38 filter from the Wide Field Imager (WFI) at the MPG/ESO 2.2m telescope in La Silla, U filter from the VIsible MultiObject Spectrograph (VIMOS) at the Very Large Telescope (VLT) in Paranal.

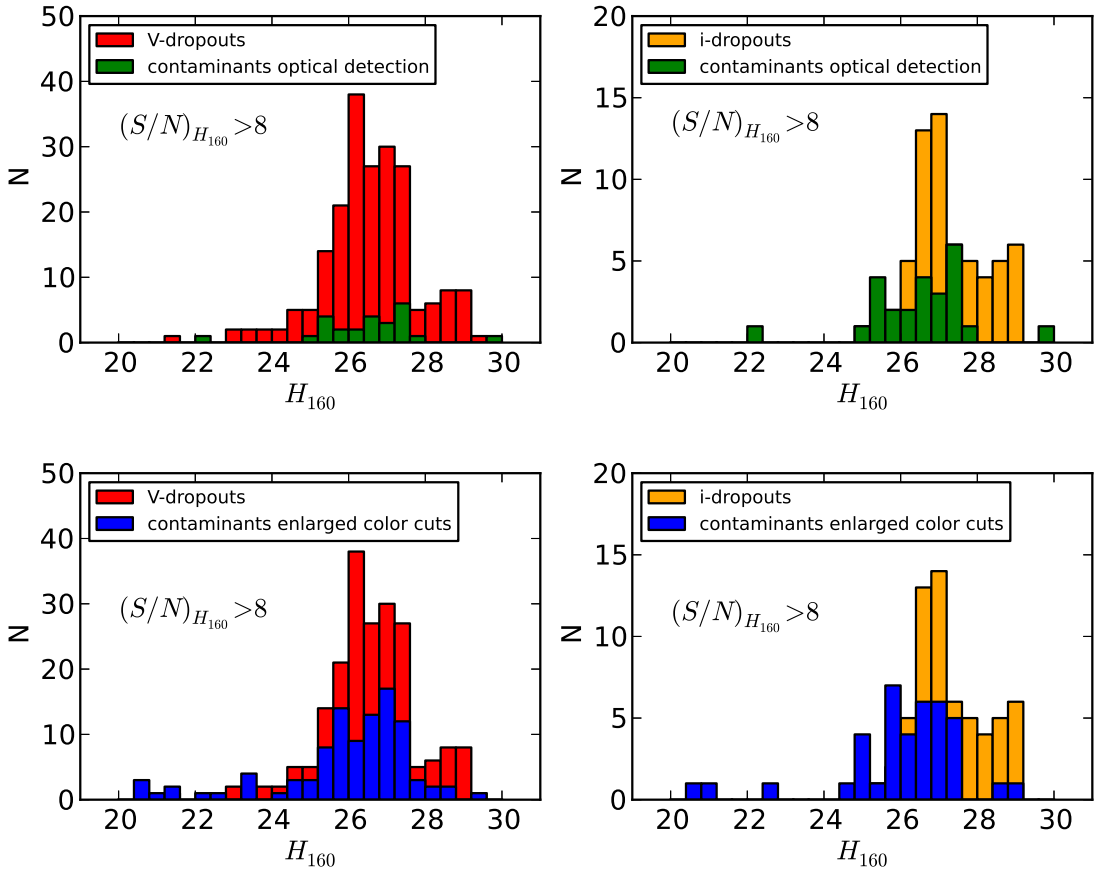


Figure 6.3: Top panels: number counts as a function of the H_{160} magnitude for the contaminants showing optical detection (green) compared to the samples of V_{606} -dropouts (red, left panel) and i_{775} -dropouts (orange, right panel) selected with $S/N > 8$ in the H_{160} -band. Bottom panels: the same as in the top panels for the contaminants lying next to the color-color selection box (blue). To built all the histograms we used 0.4 mag bins.

For the V_{606} and z_{850}-band the catalogs provide photometry from both GOODS and CANDELS. To be consistent with the other HST optical bands we used the GOODS values. Only in the case of non detection in GOODS we used the CANDELS values, if existing.
As described in Section 6.1.1 the dropout selection relies on the color-color criteria by Bouwens et al. (2014). In particular, due to the lack of any Y_{105} -band data, the 3D-HST catalogs permitted only the study of candidate dropout galaxies at $z \sim 4-5$. At $z \sim 4$ B_{435} -dropout galaxies are required to meet the following criteria:

$$\begin{aligned} B_{435} - V_{606} &> 1 \\ i_{775} - J_{125} &< 1 \\ B_{435} - V_{606} &> 1.6 \cdot (i_{775} - J_{125}) + 1 \end{aligned} \tag{6.3}$$

Regarding the V_{606} -dropout selection please refer to Equation 6.1.

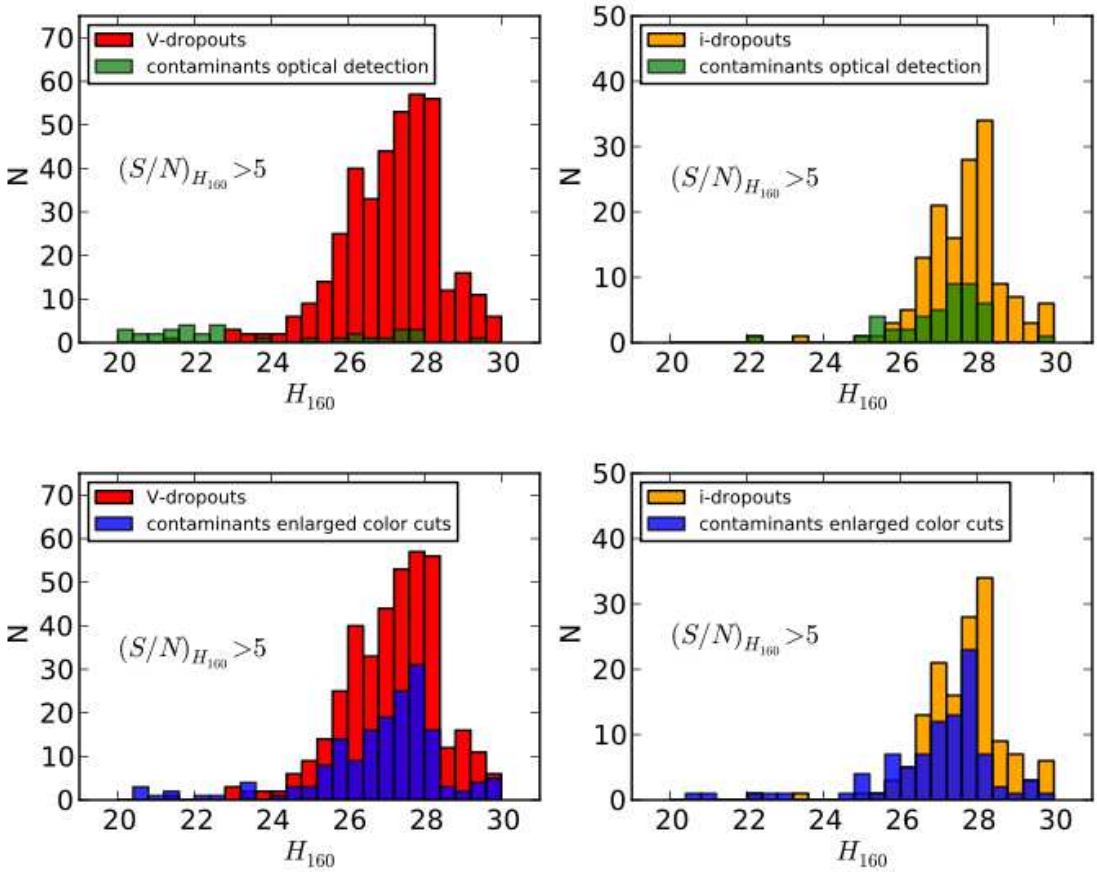


Figure 6.4: The same as in Figure 6.3 for the samples of V_{606} and i_{775} -dropouts selected with $S/N > 5$ in the H_{160} -band assuming enlarged color cuts.

Furthermore, we required B_{435} and V_{606} -dropouts to have a clear detection redward of the Lyman-break feature, i.e. we selected only sources with $S/N > 5$ or $S/N > 8$ in the i_{775} -band. The analysis on B_{435} -dropouts using CANDELS GOODS-S Multi-wavelength Catalog was prevented by the lack of data in bands bluer than the B_{435} one. On the contrary, 3D-HST provides the user with two different kinds of observations in the U-band. Please note that data blueward of the Lyman-break are strictly required to make sure the dropout selection is not including interlopers in the final sample.

Regarding the $z \sim 4$ sample, 2544 objects satisfy the $S/N > 5$ cut in the i_{775} -band, out of them 2055 are candidates B_{435} -dropouts with $S/N < 2$ in U38 and 489 are contaminants.
1530 objects satisfy the $S/N > 8$ cut in the i_{775} -band. Out of this sample 1164 are candidates B_{435} -dropouts with $S/N < 2$ in U38 and 366 are contaminants.
On the other hand, for the $z \sim 5$ sample we selected 401 objects that satisfy the $S/N > 5$ cut in the i_{775} -band, 313 are candidate V_{606} -dropouts with $S/N < 2$ in U38 and B435 and 88 are contaminants. 184 objects satisfy the $S/N > 8$ cut in the i_{775} -band. Out of this sample

	color cuts + $(S/N)_i$	real dropouts		contaminants with blue-band detection	
	total #	total #	%	total #	%
B_{435} $(S/N)_i > 5$	2544	2055	81%	489	19%
B_{435} $(S/N)_i > 8$	1530	1164	76%	366	24%
V_{606} $(S/N)_i > 5$	401	313	78%	178	22%
V_{606} $(S/N)_i > 8$	184	140	76%	83	24%

Table 6.2: Number of candidate dropouts and contaminants for the B_{435} and V_{606} -dropout samples with the two different cuts in S/N. The percentages were derived with respect to the total number of objects entering the color-color selection box.

	color cuts + $(S/N)_i$	real dropouts		contaminants next to the selection box		contaminants with blue-band detection	
	total #	total #	%	total #	%	total #	%
B_{435} $(S/N)_i > 5$	3043	2055	68%	499	0.16%	489	16%
B_{435} $(S/N)_i > 8$	1854	1164	63%	324	0.17%	366	20%
V_{606} $(S/N)_i > 5$	579	313	54%	88	0.15%	178	31%
V_{606} $(S/N)_i > 8$	267	140	52%	44	0.17%	83	31%

Table 6.3: The same as in Table 6.2, but considering the enlarged color cuts (color-color selection loosen by 0.2 mag).

140 are candidates V_{606} -dropouts with $S/N < 2$ in U38 and B435 and 44 are contaminants. Figures 6.5 and 6.6 show the distribution of candidate dropouts and contaminants in the color-color plot at $z \sim 4$ and $z \sim 5$, respectively. The contaminants (red dots) are mostly located close to the boundaries of the selection box.
The comparison between the number counts of candidate dropouts and contaminants are shown in Figures 6.7 and 6.8 as a function of the magnitude in the i_{775} and H_{160} -band. In general, the number of contaminants with a non-negligible detection in the bands blueward of the Ly-α line is much lower than the number of candidate high-z galaxies. In both redshift ranges, the number counts distributions are well separated.
As a further check we selected all the candidate dropouts and contaminants that have photometric redshift (photo-z) with a quality parameter (Q_z, Brammer et al. 2008) below 5 to get rid of possible catastrophic values of the photo-z. Then, we plotted the distribution of the photo-z for candidates high-z galaxies and contaminants (Figures 6.9 and 6.10). The histograms of candidates B_{435} and V_{606} -dropouts have a well defined peak at $z \sim 4$ and $z \sim 5$, respectively, as expected. The contaminants entering the sample of $z \sim 4$ LBGs are mainly objects with $z < 2$, but at $z \sim 5$ the distribution shows a double peak.

6.2.1 Preliminary Results

In addition to the work presented in Section 6.1 on V_{606} and i_{775} -dropouts, we made use of the 3D-HST catalogs to investigate the interlopers entering the selection of $z \sim 4$ and $z \sim 5$ galaxies, i.e. B_{435} and V_{606} -dropouts. Besides the multi-wavelength photometric data for each galaxy in the sample, 3D-HST catalogs provide the user with photometric redshifts that can be a match point when studying the population of objects contaminating the real

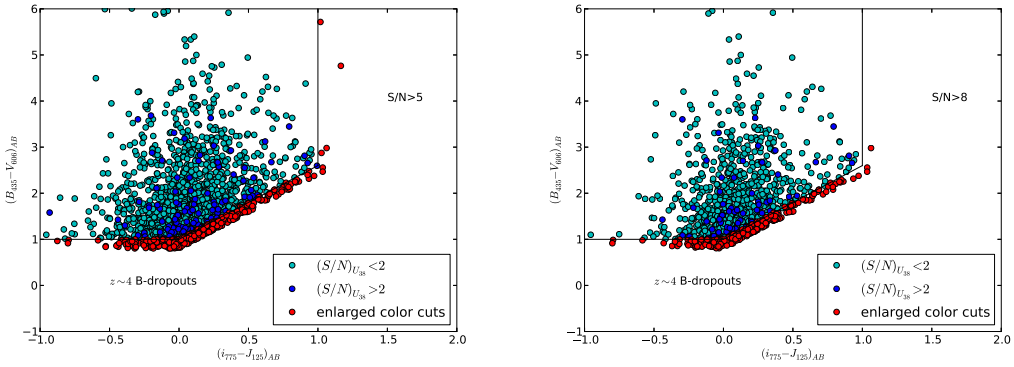


Figure 6.5: Distribution of B_{435} -dropouts (blue dots) and contaminants (red dots) in the color-color plot (color cuts by Bouwens et al. 2014) assuming two different cuts in the S/N in the i_{775} -band: $S/N > 5$ (left panel) and $S/N > 8$ (right panel).

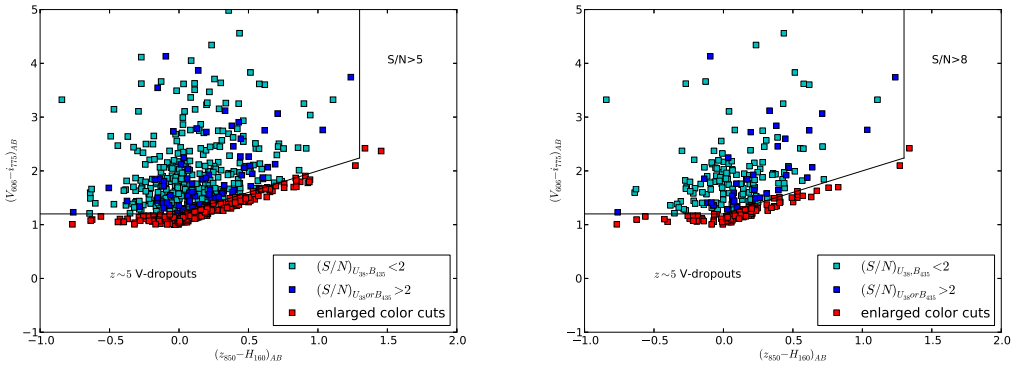


Figure 6.6: The same as in Figure 6.5, but for V_{606} -dropouts.

high-z dropout sample.
We enlarged the color-color selection box and studied both contaminants lying just outside the selection box and those with a faint detection blueward of the Lyman-break. Our results are pretty similar to those presented in the previous Section, i.e. interlopers tend to lie close to the boundaries of the selection box. This can easily be explained taking into account photometric scattering. Moreover, the availability of photometric redshift information permitted us a more accurate analysis on dropouts and contaminants. In particular, contaminants entering the sample of $z \sim 4$ LBGs are mainly objects with $z < 2$, but at $z \sim 5$ the distribution shows a double peak.

We plan to reproduce the analysis on 3D-HST using the U-band from VLT/VIMOS instead of the U38-band since VIMOS data are characterized by a higher S/N, but we do not expect our results to change significantly.

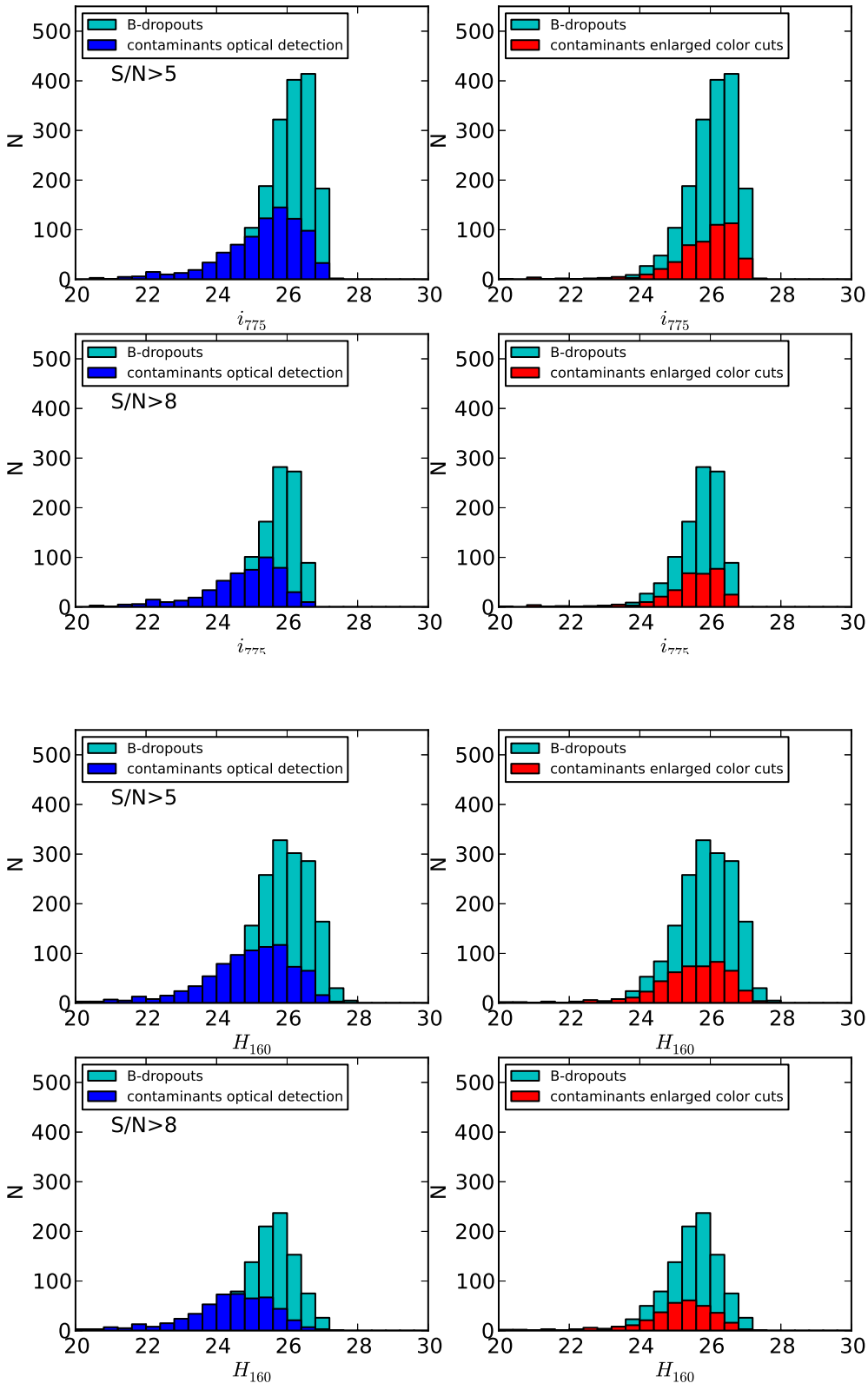


Figure 6.7: Number counts of B_{435} -dropouts (cyan histograms) and contaminants (in blue those showing optical detection and in red those lying next to the color-color selection box) as function of magnitude in the i_{775} (top Figure) or H_{160} -band (bottom Figure). In each Figure top and bottom panels refer to the samples satisfying the following cut in signal-to-noise in the i_{775} -band: $S/N > 5$ and $S/N > 8$, respectively.

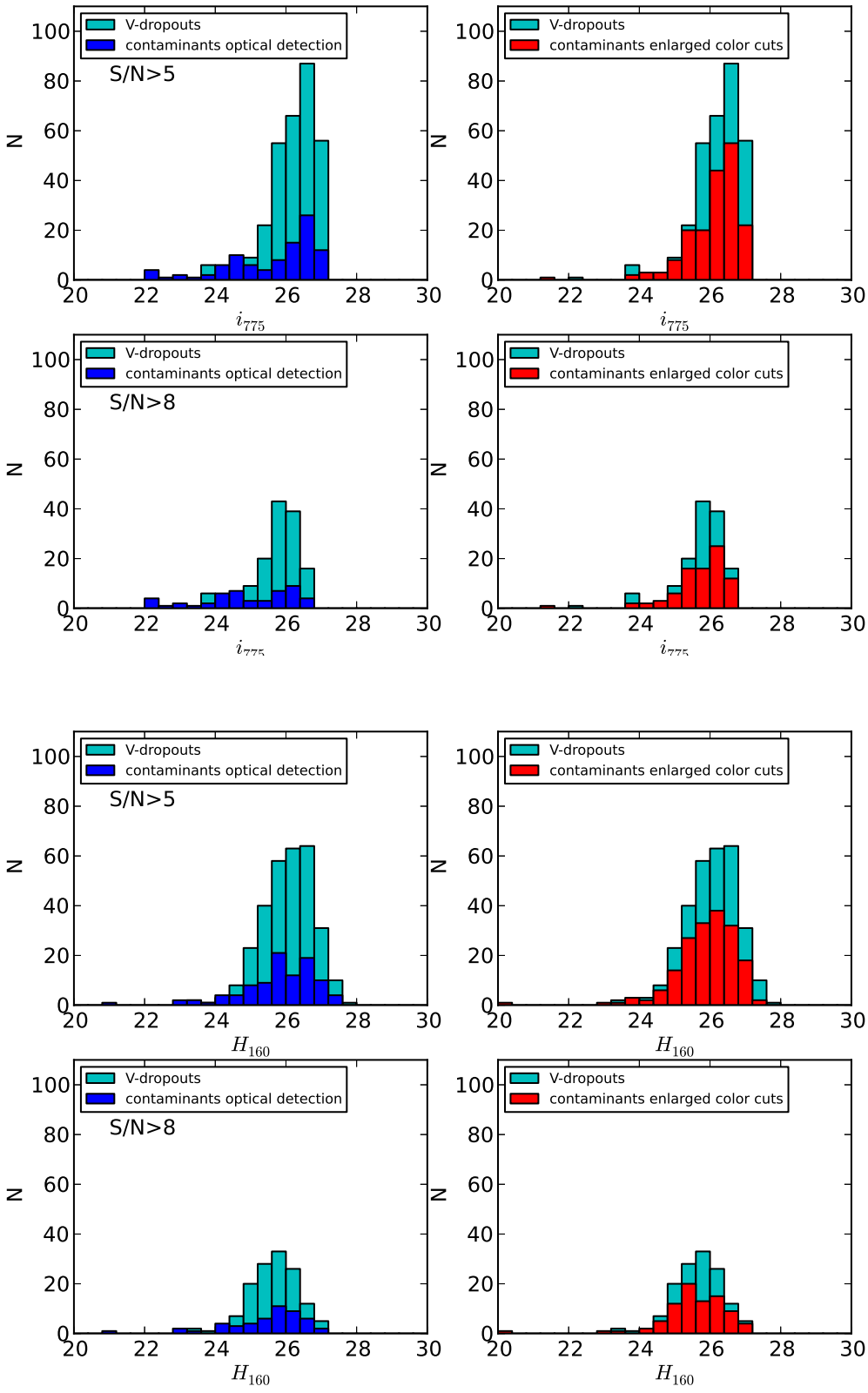


Figure 6.8: The same as in Figure 6.7, but comparing the number counts of V_{606} -dropouts (cyan histograms) and the corresponding two classes of contaminants.

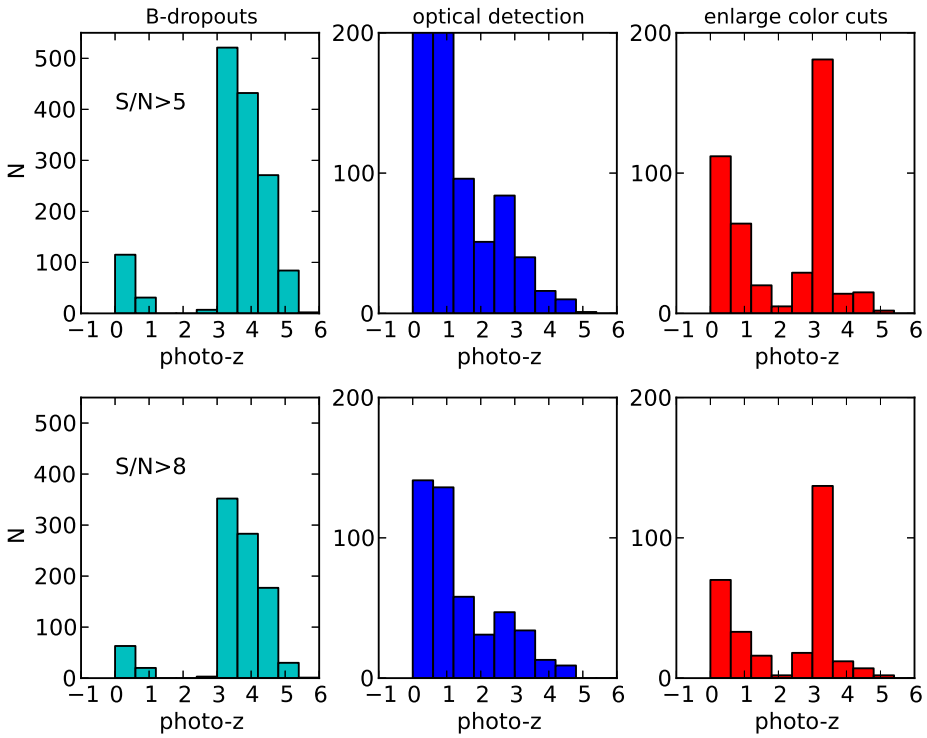


Figure 6.9: Distribution of photometric redshifts derived using the Eazy code (Brammer et al., 2008) for the B_{435} -dropouts (left panel) and the two classes of contaminants: those within the color-color selection box, but showing some flux blueward of the Lyman-break (central panel) and those right next to the selection box (right panel). A cut on the quality parameter ($Q_z < 5$) was applied to get rid of catastrophic photometric redshift values.

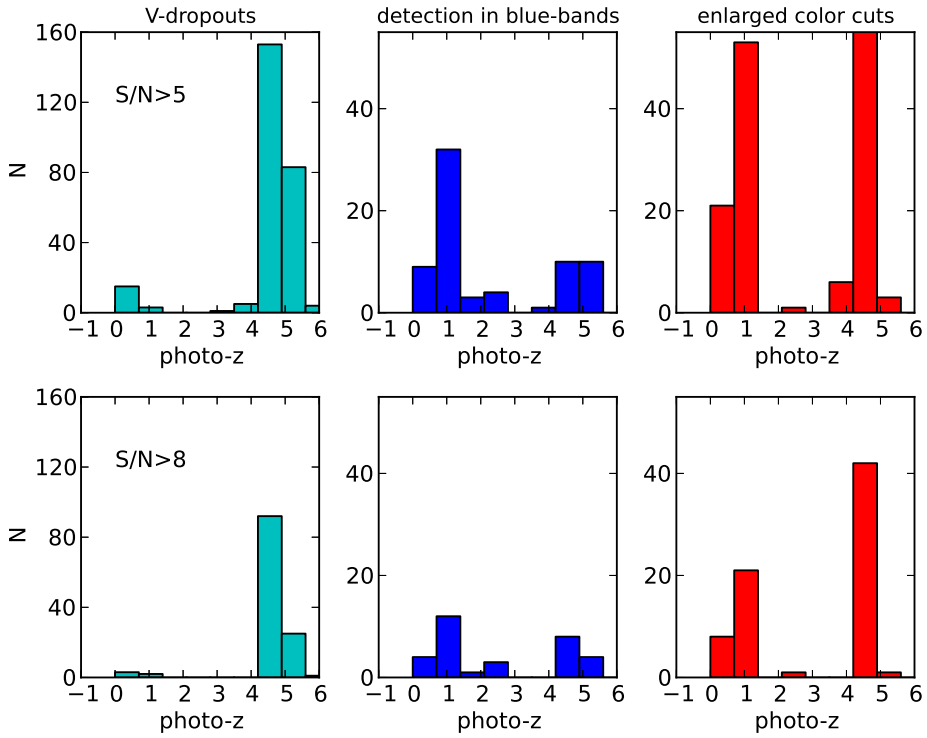


Figure 6.10: The same as in Figure 6.9, but for the V_{606} -dropouts.

6.3 Next Steps

First of all, as we said in the previous Section, we plan to extend our analysis on 3D-HST catalogs including the U-band data from VLT/VIMOS. A further check on the detection/non-detection in the bands blueward of the Lyman-break will be very helpful in disentangling candidate high-z galaxies from low-z contaminants.
Moreover, we are going to enlarge the sample using the GOODS-North catalogs as well. Tentatively, in this way we will be able to better study the photometric redshift distribution of the two classes of contaminants and to check if the color selection criteria by Bouwens et al. (2014) can be tailored even more on the search of dropout galaxies from $z \sim 4$ up to $z \sim 6$.

Anyway, it should be noted that our preliminary results were obtained using publicly available catalogs not tailored for our specific purpose. Moreover, we lacked of completeness simulations.
The next step will be to move from number counts to LFs. To this aim we need to:

- use catalogs tailored for dropout selection
- resort to completeness simulations.

In detail, we would like to characterize the two families of contaminants focusing on their LFs in the different z samples (tentatively $z \sim 4-7$) and comparing them to those of the candidate high-z galaxies selected as dropouts.

A constraint on the LF of contaminants could be crucial for the identification of all dropouts at different redshifts and should be taken into account when building up all dropouts catalogs.

Chapter 7

Summary and Conclusions

Of the 13.7 billion years of evolution of our Universe up to now we have only been able to properly investigate the last 12.8 Gyr or so, i.e. up to $z \sim 6$, starting to study the tail of the era of reionization.

To unveil the early stages in the history of the Universe high-redshift galaxies play a relevant role and so, a comprehensive analysis on them is mandatory. This thesis is addressed to the study of this population of objects using data from different deep surveys carried out with the Hubble Space Telescope. In the last two decades HST permitted to push our limits beyond that redshift mainly thanks to the capabilities of its instruments, ACS and WFC3, that allowed the detection of several objects up to $z \sim 11$ (Coe et al., 2013). Anyway, the detection of the first galaxies is limited to the tiny sky areas observed by deep surveys.

Currently several research groups focus on early galaxies, mainly getting luminosity functions (Su et al., 2011; Oesch et al., 2012; Bradley et al., 2012; Oesch et al., 2013; Bouwens et al., 2014) and studying the properties, such as radius, star formation rate, mass, and luminosity (Huang et al., 2013; Labbé et al., 2013; Oesch et al., 2014) from the photometry of relatively bright sources. However, there is a general agreement on the relevance of studying the faint galaxy population at high-z. Until the launch of the James Webb Space Telescope the study of galaxies below the current detection limit will be possible only relying on indirect methods and using datasets characterized by the best data reduction possible. The main goal of this thesis was to present and tune a technique to constrain the overall light contribution from faint primordial galaxies. In parallel we showed that the quality of the data is of the utmost importance to push our knowledge down to the detection limit of the current facilities and beyond.

The two main questions we wanted to answer are about the origin of the ionizing radiation that caused cosmic reionization and the effect of cosmological dimming on the detectability of galaxies in deep imaging surveys. In particular, since the detection of early galaxies is magnitude and surface brightness-dependent, our aim was to analyze the effect of these two parameters in deep surveys and how they could affect the derived contribution to the reionizing flux from the early galaxies.

To complete our analysis we performed, also, a preliminary study on the interlopers, most probably low-redshift galaxies with a strong break at about 4000 Å that resembles the Lyman-break. This feature makes them enter the color-color selection box of dropout galaxies. The number counts of contaminants at $z \sim 4-5-6$ suggest different luminosity functions and refer to separate populations.

7.1 The Contribution to Reionization from Faint Galaxies

In Chapter 2 we presented a new technique we developed and tuned with the aim of quantifying the overall light contribution coming from high-redshift galaxies that are below the current detection threshold of imaging data. The technique was applied to HST/ACS i_{775} and z_{850} -band images of the parallel field NICP12. The aim of this analysis was to extend by a few magnitudes the faint-end of the luminosity function at $z \sim 6$, getting constraints on the α slope. The data showed a light excess that can be ascribed to $z \sim 6$ faint galaxies and Monte Carlo simulations were used to derive a value for the α slope from that light excess. Whether we are considering a specific kind of clustering or not, the faint-end slope fitting the data is really steep ($\alpha = -1.8$ or $\alpha = -1.9$ with/without clustering). These values tell us that the ionizing flux produced by undetected galaxies at $z \sim 6$, when added to the contribution from bright galaxies at the same redshift, could have driven cosmic reionization giving a margin to model the escape fraction and the clumping factor that are still unknown.

Chapter 3 was devoted to the application of the power spectrum technique introduced in Chapter 2 to the IR dataset obtained with WFC3/IR as part of the HUDF09 program. The aim was to constrain the faint-end slope of the LF at $z \sim 7-8$, but the quality of the dataset was not suitable for our search for surface brightness fluctuations. Briefly, we compared the power spectra of the background signal obtained for the Y_{105} and J_{125} -band images looking for the light contribution from undetected galaxies at $z \sim 8$. Later on, we compared the z_{850} and Y_{105} -band ones taking into account the systematics encountered when dealing with different detectors. In none of the cases we were able to unveil any light excess ascribable to photons emitted from faint galaxies. One of the bigger issues we faced was the fact that different detectors are characterized by very different systematics and, moreover, all ACS-related issues are more well-known than WFC3/IR ones. It is undeniable that the study of the photoinizing flux from faint galaxies at $z \sim 7-8$ is essential to understand how reionization unfold, but a proper analysis of background fluctuations is not feasible with the currently available deep IR images and considering our lack of knowledge on WFC3/IR-related issues.

Since the IR datasets did not allow us to push our studies on faint galaxies towards the earlier stages of the Universe we decided to get an independent constraint on the α slope at $z \sim 6$ using the the i_{775} and z_{850} -band images from the eXtreme Deep Field (XDF). Once again our analysis was denied by spurious signals and residuals still existing in the background of the images.

In order to get rid of these effects in Chapter 4 we described all the data-reduction procedure we performed to get a new version of the XDF in the two bands we were interested in. We reprocessed the ACS data for the HUDF main field using improved data reduction algorithms, adding new data obtained during the HUDF09 program and other campaigns that happened to observe that particular area in the sky. We obtained an improved version of the XDF dataset with the aim to extend the power spectrum analysis to this field and to verify our findings on the faint-end slope of the luminosity function at $z \sim 6$. This work turned out to be a massive one and we are still working on the final products. In particular we would like to model the effects of the electronic crosstalk so as to be able to remove ghost images of bright sources appearing in other quadrants. Anyway, a very first attempt to apply the power spectrum technique suggested an hint of signal from faint galaxies swamped by background. Moreover, running `SExtractor` on the HUDF, the origi-

nal XDF, and our enhanced version of the XDF, we found that the photometry is in perfect agreement and that detected sources in our images seem to have, on average, a slightly higher signal-to-noise.

7.2 The Effect of Cosmological Dimming on Deep Survey

The goal of Chapter 5 was to quantify the effects of cosmological surface brightness dimming in the selection of galaxies at high-redshift. Cosmological surface brightness dimming of the form $(1+z)^{-4}$ has a strong dependence on redshift and suggests a selection bias when searching for high-redshift galaxies, i.e. we tend to detect only those galaxies with a high surface brightness.
The empirical strategy we used was based on the comparison of the total flux detected for the same source in surveys characterized by different depth. Our choice of datasets confined our analysis to $z \sim 4$ galaxies that are detected on the basis of a flux drop in the B_{435} -band. The sample of B_{435} -dropouts detected in GOODS, HUDF, and XDF suggests basically no effects due to cosmological dimming since the best fit of the total magnitude in the shallow image versus the one in the deep image is comparable to the bisector. So, according to our direct comparison, galaxies detected in GOODS do not become significantly brighter in deeper surveys.

Our observational results on little or no dimming are confirmed by Monte Carlo simulations as well. We derived the expected trends when assuming different surface brightness profiles and found that models of extended sources can be easily ruled out. From this result we could infer that the light distribution in high-redshift galaxies is concentrated and that star formation is likely to be clumpy. Therefore, most of the light seems to be in compact structures and this hints to the fact that JWST will likely not find extensive diffuse star forming components.

7.3 Lower-redshift Contaminants Entering the Dropout Selection

The main issue affecting all catalogs and detections of high-redshift sources is the presence of contaminants that interlope the color-color selection. Since the characteristic of the interlopers have not been studied in details yet, in Chapter 6 we presented the preliminary results of the study we conducted to constrain the luminosity function of lower-redshift galaxies that enter the samples of high-z objects selected applying the Lyman-break technique.
We made use of two different multi-wavelength catalogs (CANDELS GOODS-South Multi-wavelength Catalog, Guo et al. 2013, and 3D-HST, Skelton et al. 2014) and applied tailored color selection criteria (Bouwens et al., 2014) to identify dropout galaxies at $z \sim 4-5-6$. Contaminants are mainly $z \sim 1-2$ galaxies that show a strong 4000 Å break. This feature causes their colors to be like the dropouts ones. Due to photometric scatter these objects could satisfy the color selection criteria being erroneously identified as high-z dropouts.

On the basis of the number counts trend, we could infer that the luminosity functions are different for low-z contaminants and real dropouts. Moreover, the number of the contaminants is slowly varying across the redshift selection, while the dropouts counts vary by over a factor of 3.

According to the photometric-redshift distribution, contaminants within the color-color selection box are more likely at low redshift, while the sample lying just outside the selection box could include objects with the right redshift. A more detailed study on this could possibly lead to reconsider the color cuts applied in the selection of dropout galaxies.

7.4 Conclusions

Nowadays, the study of the early phases in the history of the Universe is very appealing and the search for galaxies at the highest redshift is at the forefront. Imaging plays a more important role than spectroscopy in searching for these high-redshift galaxies because it permits to observe more objects at the same time, better managing the telescope time. Deep surveys, obtained by observing the same sky area for several days, are the answer to the need for massive detections of high-redshift galaxies.

The Hubble Space Telescope, launched in 1990 is one the most successful and long-lasting science missions of NASA. Deep surveys carried out with HST in the last 2 decades changed astronomy and permitted to unveil the early phases in the history of the Universe, reaching up to $z \sim 11$ (Coe et al., 2013).

In this thesis we made an extensive use of datasets from HST deep surveys obtained in both the optical and IR wavelength range. In particular, thanks to GOODS, HUDF, UDF05, HUDF09, and XDF images we studied high-redshift galaxies at the tail of the reionization process and beyond. We fully explored the effects of the detection limit, the surface brightness cosmological dimming, and the presence of interlopers in the detection of primordial galaxies.

In particular, we developed a new technique permitting us to isolate the overall light contribution from faint galaxies that lie below the current detection limit. This is a very promising tool that permits to extend the knowledge on the faint-end of the high-redshift luminosity functions if associated with Monte Carlo simulations. Most of questions regarding the reionization epoch are associated with determining the amount of photoionizing radiation produced by the high-redshifts galaxy population and with the role played by faint galaxies that can not be detected one by one. Therefore putting constraints on α is a fundamental step to go deeply into reionization.

The only drawback is that the analysis of background fluctuations relies on a perfect data reduction that is rarely achieved. For this reason, up to now we have only been able to constrain the α slope at $z \sim 6$ using the NICP12 dataset (Calvi et al., 2013). We are now working to create a new version of the XDF with the aim of getting an independent α value at the same redshift, testing at the same time our technique and the previous result.

The reprocessing of the XDF field required a long and careful work and we faced several different issues affecting the ACS detector. This is mainly the reason why we are still working on the dataset, especially since we aim to create the cleanest image possible.

In the meanwhile we explored the effect of cosmological dimming on deep observations trying to understand if and, eventually how, high-redshift searches are biased by this effect. Comparing the photometry of B_{435} -dropouts galaxies detected in GOODS and deeper datasets, i.e. HUDF and XDF, we found no evidences of dimming. Making use of Monte Carlo simulations assuming different surface brightness profiles, our findings suggest a clumpy distribution of star-forming regions within the first galaxies.

Last, but not least we started to investigate the role of low-redshift interlopers enter-

ing the dropout galaxies selection criteria. The astronomical community totally lack in a comprehensive study on contaminants, in particular on their luminosity function and its evolution. Once again, deep HST surveys turned out to be of utmost importance. Indeed, in addition to ground based observations, they permitted to the build multi-wavelengths catalogs such as the 3D-HST (Skelton et al., 2014) and the CANDELS GOODS-SOUTH (Guo et al., 2013) ones, that represent a great tool for any study on the high-redshift Universe.
We made use of these catalogs to study the characteristics of galaxies entering the sample of high-redshift dropouts, but showing a detection blueward of the Lyman-break. Most likely the colors of these population of sources mimic those of high-redshift galaxies due to a strong 4000 Å break. Similarly, galaxies that lie next to color-color selection box could reasonably enter the dropout selection due to photometric scattering and so deserve a detailed study as well.
Up to now, we have noticed that interlopers and dropouts show different number counts and that the amount of contaminants is fairly constant regardless of what selection in redshift we perform.

On the basis of above, we would like to stress the key role played by the Hubble Space Telescope in the search for primeval galaxies and that deep survey, carried out with this telescope in the last two decades, permitted the astronomical community to push further our limits, deep into the reionization epoch and towards the very first sources of light.

7.5 Future Perspective

As soon as possible we would like to complete the data-processing to obtain the v2.0 of the XDF dataset. In particular, we are focused on modelling the electronic ghosts and getting rid of them. Moreover, new frames imaging that sky area came available. Even though they cover mostly just corners of the combined image, we plan to process and include them in our work to improve the quality of the final products in the i_{775} and z_{850} -band.

In the meanwhile, we are are going to complete the work introduced in Chapter 6 conducting a more accurate study on the interlopers using catalogs tailored for dropouts selection. Our main goal is to use catalogs with associated completeness simulations to focus on real luminosity functions instead of number counts.

Most of the work presented in this thesis can be considered as a pathfinder for the capabilities of the upcoming James Webb Space Telescope. In particular, considering the capabilities of JWST in exploring the early Universe in the IR domain, soon we would be able to investigate the primordial Universe using new deep survey data. JWST will permit us to go beyond the current detection limit, observing the faint galaxy population and shedding light on the contribution to reionization from galaxies at $z \sim 6$ and higher. New deep surveys will permit us to verify our findings on the α slope of the luminosity function at $z \sim 6$ and, in principle, it would be possible to apply the power spectrum technique to go a few magnitudes below the detection limit of JWST new instruments.
Finally, JWST will test our prediction on the surface brightness dimming effect, verifying whether or not high-z galaxies have a clumpy star formation.

Appendix A

Gas-phase Metallicity of 27 Galaxies at Intermediate Redshift

Based on Morelli L., **Calvi V.***, Cardullo A., Pizzella A., Corsini E. M., and Dalla Bontà E., 2012, A & A: 544, A74*

A.1 Introduction

The gas-phase metallicity is one of the most important observational diagnostics of the current evolutionary state of galaxies. Accurate gas-phase metallicity measurements are crucial for investigating the interplay between fundamental processes, such as star formation, gas accretion, gas flows, and supernova-driven winds occurring during the life of galaxies (Moustakas et al., 2011; Cresci et al., 2012; Sommariva et al., 2012).

The galaxy gas-phase metallicity correlates with many properties, such as the star formation rate (Mannucci et al., 2010; Lara-López et al., 2010), morphological type (Edmunds and Pagel, 1984), surface mass density (Ryder, 1995; Garnett et al., 1997), and maximum rotation velocity (Dalcanton, 2007). In particular, the metallicity Z most strongly correlates with the mass M and B-band luminosity L of the galaxy (McClure and van den Bergh, 1968; Skillman et al., 1989; Richer and McCall, 1995). Adopting the Sloan Digital Sky Survey (SDSS, Abazajian et al. 2009), Tremonti et al. (2004) and Gallazzi et al. (2005) derived the $M - Z$ and $L - Z$ relations in the local Universe for a large sample of galaxies on the basis of the gas-phase metallicity. These relations are very tight with a scatter of less than 0.1 dex and show that the more luminous galaxies have higher gas metallicities than their fainter counterparts, which has a great impact on the theoretical models of galaxy formation.

Another crucial piece of information to help us to understand the assembly history of galaxies comes from the study of the evolution of the $M - Z$ and $L - Z$ relations across cosmic time.

This topic has been investigated by measuring the gas-phase metallicity of galaxies as a function of redshift up to $z \sim 3.5$. While these relations show clear evidence of galactic evolution at high z, because at a given mass higher-z galaxies have lower metallicities (Shapley et al., 2004; Erb et al., 2006; Maiolino et al., 2008; Mannucci et al., 2009; Rodrigues et al., 2012), at intermediate redshift ($z < 1$) it is still debated whether there has been any evolution. Carollo and Lilly (2001) and Kobulnicky and Kewley (2004) found that the $L - Z$ relation at intermediate redshift is consistent with the local one. In contrast, Maier et al. (2005),

Savaglio et al. (2005), and Zahid et al. (2011) measured lower metallicities for intermediate-redshift objects than for local galaxies, supporting a scenario in which the $L-Z$ and $M-Z$ relations also evolve over the range of redshifts between 0.5 and 1.

In this research note, we therefore present new measurements of the equivalent widths for the $[\mathrm{O_{II}}]\lambda3727$, H_β, $[\mathrm{O_{III}}]\lambda4959$, and $[\mathrm{O_{III}}]\lambda5007$ emission-lines and present our derived metallicity-sensitive emission lines ratios R_{23}, ionization-sensitive emission-lines ratios O_{32}, and gas-phase oxygen abundances $12+\log(\mathrm{O/H})$ for 27 intermediate-redshift ($0.35 \leq z \leq 0.52$) galaxies. The data that we analised for all the sample galaxies are available from the SDSS archive, hence our measurements are a valuable supplementary resource for the astronomical community.

A.2 Observations and Data Reduction

As often happens, the data acquired for a particular aim can potentially contain information useful for different purposes. For this work, we used observations carried out in 2006 at the Very Large Telescope (VLT) of the European Southern Observatory (ESO) at Paranal Observatory using the VIsible Multi-Object Spectrograph (VIMOS). The main goal of these observations was to find satellites in the region of $\sim 500 \times 500$ kpc^2 surrounding seven nearby isolated spiral galaxies (Yegorova et al., 2011).

VIMOS was used in Multi Object Spectroscopy (MOS) mode with the grism "HR orange" and the order sorting-filter GG435. The grism is characterized by a reciprocal dispersion of 0.6 Å pixel^{-1} and a spectral resolution of $R = 2150$ for the adopted 1" slit. In the high-resolution configuration that we used for the observations, the spectral range strongly depends on the position of the slit in the field of view varying from 4550-6950Å to 6000-8400Å. The full dataset includes MOS observations of 7 different fields for a total exposure time of ~ 20 hours. This granted a typical exposure time of ~ 3 hours for each target field. Each of the four quadrants of VIMOS observed ~ 50 objects so as to have ~ 200 galaxies within each field of view of 16' $\times$ 18'.

The observed sample consists of 1450 objects with SDSS $r-$band magnitudes ranging from 17 mag to 24 mag and a modal value of 22 mag. As expected, only a small fraction of the target galaxies were found to be satellites. We focused on the remaining 1347 background and foreground objects to measure their gas-phase metallicities. Their spectra were bias-subtracted, corrected for flat-field effect, cleaned of cosmic rays, and wavelength-calibrated by means of the ESO VIMOS data reduction pipeline. A more detailed description of the data reduction is reported in Yegorova et al. (2011).

A.3 Sample Selection

To estimate the gas-phase metallicity, it is necessary to know the electron temperature provided by the comparison between auroral and nebular emission lines (Osterbrock, 1989). However, since auroral lines are generally weak, it is difficult to observe them in distant galaxies. To overcome this issue, empirical relations between the intensity of the nebular lines and metallicity have been used to derive the gas-phase metallicity. The most widely used metallicity indicator is the R_{23} emission lines ratio parameter (Pagel et al., 1979) de-

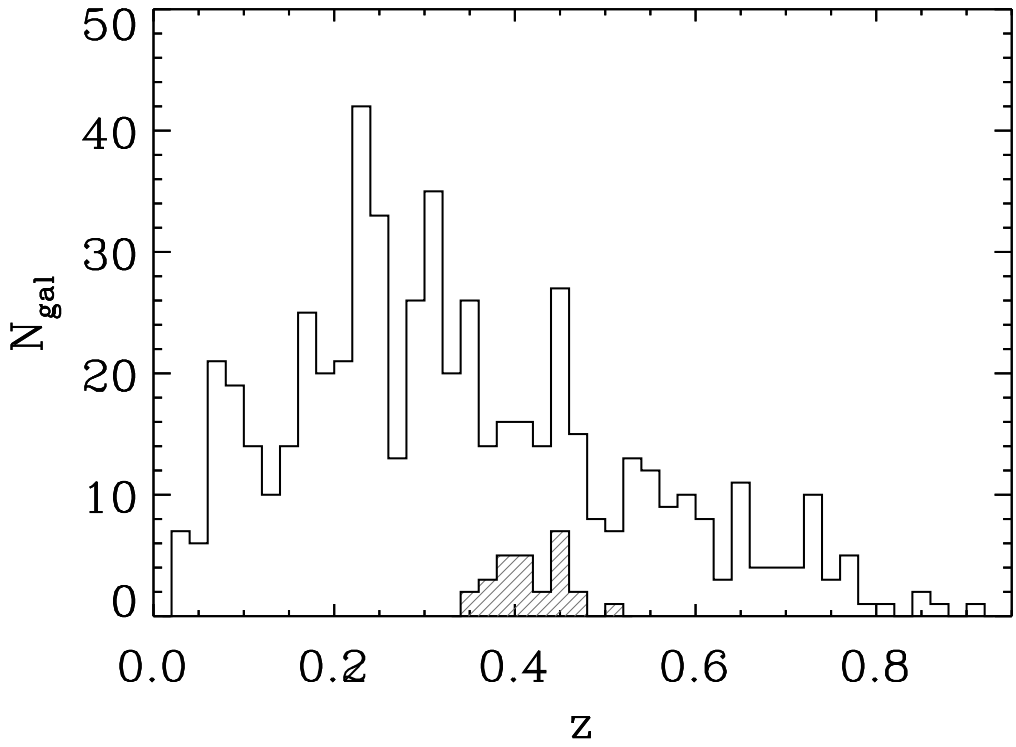


Figure A.1: Distribution of redshifts for the 571 galaxies with measured emission lines. The shaded histogram shows the distribution of the serendipitous sample of 27 objects presented in this work.

fined as:

$$R_{23} = \frac{[O_{II}]\lambda 3727 + [O_{III}]\lambda 4959 + [O_{III}]\lambda 5007}{\mathrm{H}_\beta}. \tag{A.1}$$

For our analysed dataset the signal-to-noise ratio (S/N) of the spectra allowed us to measure redshifts for only 571 objects out of the observed 1347 galaxies. We fitted all the available emission lines with Gaussians. The $[\mathrm{O_{II}}]\lambda 3727$ doublet was modelled with two Gaussians of the same full width at half maximum. The galaxy redshift was computed as the average of the redshifts derived from the central wavelengths of the measured emission lines.

We detected 177 galaxies in the redshift range between 0.35 and 0.6. However, most of their spectra did not map all the $[\mathrm{O_{II}}]\lambda 3727$, H_β, and $[\mathrm{O_{III}}]\lambda\lambda 4959, 5007$ emission-lines we needed to compute the metallicity-sensitive emission-lines ratio R_{23}, ionization-sensitive emission lines ratio O_{32}, and gas-phase oxygen abundance $12 + \log(\mathrm{O/H})$, because of their actual position in the VIMOS field of view. All the relevant emission lines fall in the observed spectral range only for 27 objects, which represent our final sample of galaxies. They are listed in Table A.1 and the distribution of their measured redshifts is shown in Figure A.1.

Galaxy name SDSS (1)	$[O_{II}]\lambda3727$ [Å] (2)	H_β [Å] (3)	$[O_{III}]\lambda4959$ [Å] (4)	$[O_{III}]\lambda5007$ [Å] (5)	z (6)
J003845.31+000006.8	2.95 ± 0.45	1.43 ± 0.49	10.01 ± 0.66	26.35 ± 0.45	0.451
J003838.68+000225.1	92.41 ± 0.68	22.22 ± 0.31	15.58 ± 0.72	46.25 ± 0.69	0.408
J003806.59+000421.9	208.77 ± 7.16	15.75 ± 1.07	9.98 ± 2.50	39.45 ± 1.30	0.474
J134235.94+014424.1	3.89 ± 0.41	3.78 ± 0.54	1.61 ± 0.46	9.17 ± 0.68	0.411
J134229.37+015244.5	58.21 ± 0.58	20.63 ± 0.48	17.33 ± 0.77	43.93 ± 0.69	0.372
J134156.84+014241.0	6.09 ± 0.47	4.79 ± 1.57	3.74 ± 0.87	13.11 ± 0.82	0.390
J145221.51+043448.8	23.01 ± 0.55	9.14 ± 0.15	-0.12 ± 0.45	5.17 ± 0.35	0.450
J145150.75+044111.6	7.09 ± 0.37	3.23 ± 0.83	2.74 ± 0.78	2.76 ± 0.40	0.425
J145147.24+043654.2	30.29 ± 1.13	6.28 ± 1.68	4.61 ± 1.35	6.12 ± 1.17	0.517
J145156.20+043321.7	7.92 ± 0.61	2.16 ± 0.12	4.13 ± 0.25	14.93 ± 0.32	0.416
J145147.77+043244.5	11.84 ± 0.61	1.03 ± 0.79	2.70 ± 0.75	4.18 ± 0.81	0.414
J145149.20+043501.6	14.03 ± 1.09	2.90 ± 0.50	1.57 ± 0.84	8.16 ± 0.67	0.443
J152638.18+034353.8	15.42 ± 0.36	9.75 ± 0.46	4.80 ± 0.79	15.49 ± 0.51	0.407
J152630.18+035110.8	19.58 ± 0.57	5.93 ± 0.45	3.27 ± 1.01	4.39 ± 0.60	0.458
J152559.83+035249.3	18.67 ± 0.45	3.24 ± 0.59	3.30 ± 0.67	8.50 ± 0.56	0.475
J152606.03+034530.7	34.36 ± 0.54	9.08 ± 0.65	6.12 ± 0.44	30.29 ± 0.36	0.386
J154933.50-004255.0	39.71 ± 0.57	14.98 ± 0.27	14.95 ± 0.68	51.32 ± 0.46	0.358
J154931.48-003406.1	45.92 ± 0.59	21.36 ± 0.62	28.59 ± 0.42	107.08 ± 0.36	0.386
J154925.24-003717.8	18.21 ± 0.45	4.76 ± 0.63	2.99 ± 0.32	13.41 ± 0.32	0.386
J154914.49-004010.2	30.77 ± 0.65	9.05 ± 0.43	4.52 ± 0.89	11.21 ± 0.51	0.441
J154925.12-003645.4	18.54 ± 0.34	5.50 ± 0.45	1.90 ± 0.19	7.83 ± 0.37	0.445
J222013.98-074355.8	6.41 ± 0.47	2.06 ± 0.20	1.00 ± 0.32	1.29 ± 0.24	0.399
J222014.86-073333.7	23.12 ± 1.01	7.10 ± 0.88	1.51 ± 1.52	10.07 ± 1.57	0.440
J221934.90-073825.9	7.73 ± 0.89	3.44 ± 0.29	8.27 ± 0.42	26.36 ± 0.35	0.433
J221941.85-074703.8	14.91 ± 0.58	2.52 ± 0.57	2.88 ± 0.52	2.38 ± 0.34	0.369
J221939.28-074744.5	1.72 ± 0.37	0.88 ± 0.14	0.31 ± 0.34	0.62 ± 0.32	0.351
J221941.90-074825.7	29.10 ± 0.60	7.12 ± 0.49	6.26 ± 0.63	15.38 ± 0.21	0.375

Table A.1: Measured properties of the sample galaxies. Col. (1): SDSS name of the galaxy. Col.(2-5): Equivalent widths of the $[O_{II}]\lambda3727$, H_β, $[O_{III}]\lambda4959$, and $[O_{III}]\lambda5007$ emission lines, respectively. Col.(6): Measured redshift.

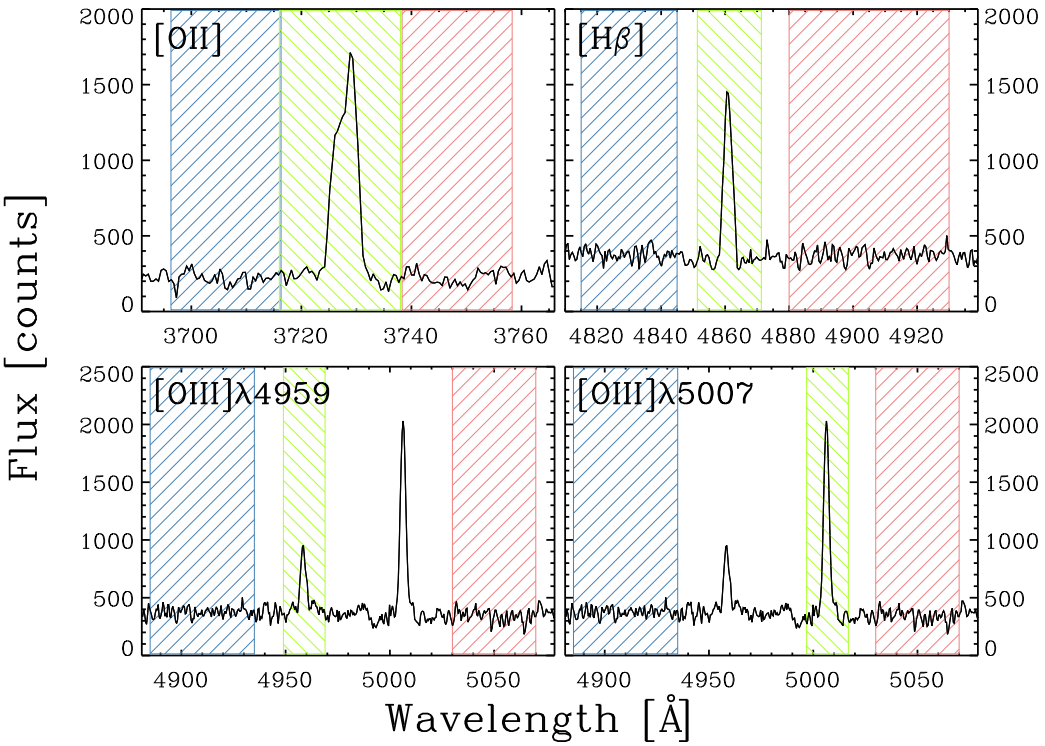


Figure A.2: Central (green dashed region) and continuum bandpasses (blue and red dashed regions) adopted for measuring the equivalent widths of the $[O_{II}]\lambda 3727$ (top left panel), H_β (top right panel), $[O_{III}]\lambda 4959$ (bottom left panel), and $[O_{III}]\lambda 5007$ (bottom right panel) emission lines.

A.4 Data Analysis and Gas-phase Metallicity

The apparent g-band Petrosian magnitudes of the sample galaxies range between 24.9 and 20.8 mag, as they were derived from the SDSS DR7. Since the S/N of the spectra was too low to derive the radial distribution of the gas metallicity, we rebinned the spectra along the spatial direction to obtain a $S/N \geq 15$ per resolution element.

Since no spectrophotometric standard star had been observed, the galaxy spectra were not flux-calibrated. Furthermore, the observed spectral range does not include the H_α emission line and thus a reliable reddening estimation was not possible. For the above reasons, we followed the approach suggested by Kobulnicky and Phillips (2003) by replacing the flux of the emission-lines with their equivalent width when measuring the gas-phase metallicity. This method does not need flux-calibrated spectra and has the further advantage of being insensitive to reddening. In addition, Kobulnicky and Phillips (2003) verified that the error associated with the use of the emission-line equivalent widths is smaller than both the typical error in the flux measurement and systematic error in the R_{23} and $12 + \log(O/H)$ calibrations. Thus, we measured the equivalent widths of the $[O_{II}]\lambda 3727$, H_β, and $[O_{III}]\lambda\lambda 4959, 5007$ emission lines in the-rest frame spectra. For each emission line, we considered a central bandpass covering the feature of interest and two adjacent bandpasses, at the red and blue side, tracing the local continuum. The continuum level un-

Emission line (1)	Central bandpass [Å] (2)	Continuum bandpasses [Å] (3)	Reference (4)
$[O_{II}]\lambda3727$	$3716.30 - 3738.30$	$3696.30 - 3716.30$	1
		$3738.30 - 3758.30$	1
H_β	$4851.32 - 4871.32$	$4815.00 - 4845.00$	2
		$4880.00 - 4930.00$	2
$[O_{III}]\lambda4959$	$4948.92 - 4968.92$	$4885.00 - 4935.00$	2
		$5030.00 - 5070.00$	2
$[O_{III}]\lambda5007$	$4996.85 - 5016.85$	$4885.00 - 4935.00$	2
		$5030.00 - 5070.00$	2

Table A.2: Definition of the bandpasses for the emission lines. Col. (4): 1 = Fisher et al. (1998), 2 = Gonzalez (1993)

derlying the emission line was estimated by interpolating a straight line in the continuum bandpasses. The bandpasses were defined following Fisher et al. (1998) for $[O_{II}]\lambda3727$ and Gonzalez-Gonzalez (1993) for H_β, $[O_{III}]\lambda4959$, and $[O_{III}]\lambda5007$. The bandpasses are listed in Table A.2 and shown in Figure A.2. The errors associated with the measured equivalent widths were derived from photon statistics and CCD read-out noise, and calibrated by means of Monte Carlo simulations.

A well-known problem in measuring the equivalent width of the H_β emission line in galaxies is the contamination by the H_β absorption line. We tried to address this issue by using GANDALF (Gas AND Absorption Line fitting; Sarzi et al. 2006) to fit the galaxy spectra with synthetic stellar population models, as done by Morelli et al. (2008, 2012). A linear combination of template stellar spectra from the ELODIE library by Prugniel and Soubiran (2001) was convolved with the line-of-sight velocity distribution and fitted to the observed galaxy spectrum by χ^2 minimization in pixel space. However, the low S/N of the stellar continuum of the galaxy spectra prevented us from deriving a reliable combination of stellar templates to fit the galaxy stellar component. Therefore, to account for the H_β absorption-line contamination we followed Kobulnicky and Kewley (2004) by applying a correction of 2Å(Kobulnicky and Phillips, 2003) to the measured equivalent width of the H_β emission line used to derive the gas-phase metallicity.

All the equivalent widths measured for the galaxies in our sample are reported in Table A.1. They were used to compute the metallicity-sensitive emission-line ratio R_{23}, as well as the emission-line ratio O_{32} defined by Kobulnicky and Kewley (2004) as

$$O_{32} = \frac{[\mathrm{OIII}]\lambda4959 + [\mathrm{OIII}]\lambda5007}{[\mathrm{OII}]\lambda3727} \tag{A.2}$$

which is mostly sensitive to the ionization (Nagao et al., 2006). All the values of $\log(R_{23})$ and $\log(O_{32})$ are listed in Table A.3.

For the sample galaxies, we could not derive the $[N_{II}]\lambda6583\ /H_\alpha$ ratio, thus it was impossible to break the R_{23} degeneracy. Following Kobulnicky and Zaritsky (1999) and Kobulnicky et al. (2003), we assumed that all the galaxies lie in the upper branch $(12+\log(O/H)\geq 8.4)$ of the R_{23}-O/H relation. We then checked the consistency of this assumption by adopting the [OIII]/[OII] diagnostic lines ratio as proposed by Maiolino et al. (2008). As done by

Kobulnicky and Kewley (2004), we derived the gas-phase oxygen abundance by averaging the values obtained from the calibrations for the upper branch given in McGaugh (1991) and Kewley and Dopita (2002). The values of $12 + \log(\mathrm{O/H})$ and their uncertainties, owing to the statistical measurement errors in the equivalent widths, are listed in Table A.3. The systematic errors due to the uncertainties in the photoionization models ($0.2 - 0.5$ dex; Kennicutt et al., 2003; Garnett et al., 2004) were not taken into account.

In Figure A.3, we plot the oxygen abundance as a function of the absolute magnitude in the B-band M_{B} for the sample galaxies. We derived M_{B} from the rest-frame g-band Petrosian magnitude given in the SDSS DR7, using the transformation $B = g + 0.327\,(g - r) + 0.216$ (Chonis and Gaskell, 2008). The resulting values are given in Table A.3. We also plotted the data for local (Nearby Field Galaxy Sample, Jansen et al. 2000; SDSS, Tremonti et al. 2004) and intermediate-redshift galaxies (Canada-France Redshift Survey, Lilly et al. 2003; GOODS-North Field, Kobulnicky and Kewley 2004).

The dashed magenta line in Figure A.3 represents the linear fit to the sample galaxies. J154925.24-003717.8 and J003806.59+000421.9 are two outliers in the $L - Z$ distribution. The most prominent outlier is the faintest galaxy of the sample, for which the assumption of an upper R_{23}-O/H branch could be incorrect (Skillman et al., 1989). Adopting the calibration of Kewley and Dopita (2002) for the lower branch, we obtained a value of $12 + \log(\mathrm{O/H}) = 8.14$. However only the measurement of the $[\mathrm{N_{II}}]\lambda6583$ $/\mathrm{H}_\alpha$ ratio could firmly help us to discriminate between the two branches. The second outlier has an extremely low value of the gas-phase metallicity. This is probably due to an underestimation of the H_β absorption-line correction. The slope and zero-point of the fitted $L-Z$ relation to our data points (calculated after excluding the two outliers) are consistent within the errors with the slope and zero-point obtained by Kobulnicky and Kewley (2004) for their sample of galaxies in the redshift range $0.2-0.4$. This was expected since 23 of our 27 galaxies have a redshift in the range $0.35 - 0.45$, and confirms their findings of an evolution in the $L - Z$ relation since intermediate redshift.

Finally, we estimated the star formation rate (SFR) of the sample galaxies from their B-band absolute magnitudes and H_β equivalent widths according to the relations of Kennicutt et al. (2003) and Kobulnicky and Kewley (2004)

$$\mathrm{SFR}\ (\mathrm{M_\odot yr^{-1}}) = \frac{2.8 \times 5.49 \cdot 10^{31} \times 2.5^{M_{\mathrm{B}}} \cdot EW_{\mathrm{H}_\beta}}{1.26 \cdot 10^{41}}. \tag{A.3}$$

The derived SFRs are reported in Table A.3.

A.5 Conclusions

Our original purpose was to explore the evolution of the $L - Z$ relation of Tremonti et al. (2004) to the intermediate-redshift range using a large number of background galaxies in very deep VIMOS data. After combining all the constraints needed to perform our analysis, the number of objects dramatically decreased from 1347 to 27, preventing us from being able to draw any conclusion about the possible evolution of the $L - Z$ relation with redshift. Here, we have reported the measured equivalent widths of the $[\mathrm{O_{II}}]\lambda3727$, H_β, $[\mathrm{O_{III}}]\lambda4959$, and $[\mathrm{O_{III}}]\lambda5007$ emission lines, the values of the metallicity-sensitive emission-line ratio R_{23}, the ionization-sensitive emission-line ratio O_{32}, and the gas-phase oxygen abundances $12+\log(\mathrm{O/H})$ derived for the serendipitous sample of 27 galaxies with the aim

Galaxy name SDSS (1)	$M_{\rm B}$ [mag] (2)	$\log(R_{23})$ (3)	$\log(O_{32})$ (4)	$12+\log({\rm O/H})$ (5)	SFR $[{\rm M}_\odot\ {\rm yr}^{-1}]$ (6)
J003845.31+000006.8	-22.06	$1.06^{+0.01}_{-0.01}$	$1.09^{+0.06}_{-0.07}$	8.37 ± 0.03	1.05
J003838.68+000225.1	-21.32	$0.80^{+0.02}_{-0.02}$	$-0.17^{+0.01}_{-0.02}$	8.63 ± 0.02	8.33
J003806.59+000421.9	-21.46	$1.16^{+0.01}_{-0.01}$	$-0.62^{+0.02}_{-0.03}$	7.72 ± 0.06	6.71
J134235.94+014424.1	-20.09	$0.40^{+0.02}_{-0.02}$	$0.44^{+0.05}_{-0.06}$	8.98 ± 0.02	0.45
J134229.37+015244.5	-19.91	$0.72^{+0.01}_{-0.01}$	$0.02^{+0.01}_{-0.03}$	8.75 ± 0.01	2.11
J134156.84+014241.0	-19.58	$0.53^{+0.02}_{-0.02}$	$0.44^{+0.04}_{-0.04}$	8.93 ± 0.02	0.36
J145221.51+043448.8	-20.97	$0.40^{+0.01}_{-0.01}$	$-0.65^{+0.04}_{-0.05}$	8.98 ± 0.03	2.46
J145150.75+044111.6	-21.01	$0.38^{+0.03}_{-0.03}$	$-0.10^{+0.06}_{-0.08}$	8.99 ± 0.04	0.90
J145147.24+043654.2	-20.94	$0.69^{+0.02}_{-0.02}$	$-0.45^{+0.06}_{-0.08}$	8.75 ± 0.06	1.65
J145156.20+043321.7	-21.18	$0.81^{+0.01}_{-0.01}$	$0.38^{+0.03}_{-0.03}$	8.67 ± 0.02	0.70
J145147.77+043244.5	-20.36	$0.79^{+0.02}_{-0.03}$	$-0.23^{+0.06}_{-0.08}$	8.64 ± 0.07	0.16
J145149.20+043501.6	-19.77	$0.68^{+0.02}_{-0.02}$	$-0.15^{+0.05}_{-0.06}$	8.78 ± 0.05	0.26
J152638.18+034353.8	-20.14	$0.48^{+0.01}_{-0.01}$	$0.11^{+0.02}_{-0.02}$	8.94 ± 0.01	1.23
J152630.18+035110.8	-20.85	$0.53^{+0.02}_{-0.02}$	$-0.40^{+0.06}_{-0.07}$	8.90 ± 0.04	1.44
J152559.83+035249.3	-20.98	$0.76^{+0.01}_{-0.01}$	$-0.19^{+0.03}_{-0.03}$	8.68 ± 0.03	0.88
J152606.03+034530.7	-19.73	$0.80^{+0.03}_{-0.02}$	$0.02^{+0.04}_{-0.01}$	8.64 ± 0.01	0.78
J154933.50-004255.0	-19.45	$0.79^{+0.01}_{-0.02}$	$0.22^{+0.01}_{-0.01}$	8.68 ± 0.02	1.00
J154931.48-003406.1	-19.41	$0.89^{+0.02}_{-0.03}$	$0.47^{+0.02}_{-0.02}$	8.57 ± 0.03	1.38
J154925.24-003717.8	-16.62	$0.71^{+0.01}_{-0.01}$	$-0.04^{+0.01}_{-0.01}$	8.76 ± 0.01	0.02
J154914.49-004010.2	-20.49	$0.62^{+0.01}_{-0.01}$	$-0.29^{+0.02}_{-0.03}$	8.83 ± 0.02	1.57
J154925.12-003645.4	-20.70	$0.58^{+0.01}_{-0.01}$	$-0.28^{+0.02}_{-0.02}$	8.87 ± 0.01	1.16
J222013.98-074355.8	-21.21	$0.33^{+0.02}_{-0.03}$	$-0.44^{+0.07}_{-0.09}$	9.01 ± 0.05	0.69
J222014.86-073333.7	-21.31	$0.58^{+0.02}_{-0.03}$	$-0.30^{+0.07}_{-0.09}$	8.87 ± 0.06	2.62
J221934.90-073825.9	-20.85	$0.89^{+0.01}_{-0.01}$	$0.65^{+0.04}_{-0.05}$	8.59 ± 0.02	0.83
J221941.85-074703.8	-19.77	$0.65^{+0.01}_{-0.01}$	$-0.45^{+0.05}_{-0.05}$	8.80 ± 0.04	0.22
J221939.28-074744.5	-21.06	$0.04^{+0.08}_{-0.09}$	$-0.26^{+0.19}_{-0.34}$	9.11 ± 0.13	0.25
J221941.90-074825.7	-20.39	$0.74^{+0.01}_{-0.01}$	$-0.12^{+0.02}_{-0.01}$	8.71 ± 0.01	1.13

Table A.3: Derived properties of the 27 galaxies in our sample. Col.(2): Absolute B-band magnitude from SDSS DR7 derived adopting the measured redshift. Col.(3): Metallicity-sensitive R_{23} parameter. Col.(4): Ionization-sensitive O_{32} parameter. Col.(5): Gas-phase oxygen abundance. Col.(6): Star formation rate as derived using the $\rm H_\beta$ uncorrected for $\rm H_\beta$ stellar absorption line.

of making them available to the community. Our measurements are consistent with those of Lilly et al. (2003) and Kobulnicky and Kewley (2004) for galaxies in the redshift range $0.3 < z < 1.0$.

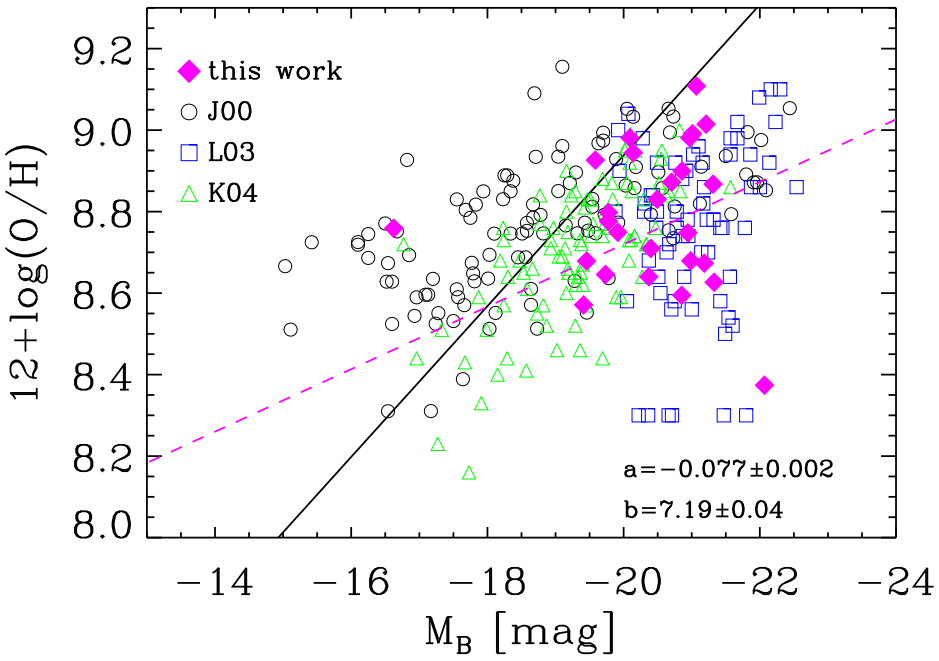


Figure A.3: Plot of our data (solid diamonds) on the $L - Z$ plane. Open squares represent the galaxies in the redshift range $0.47 < z < 0.92$ from the Canada-France Redshift Survey (Lilly et al., 2003), open triangles mark the galaxies with $0.3 < z < 0.6$ from the GOODS-North Field (Kobulnicky and Kewley, 2004), and open circles refer to the local galaxies of the Nearby Field Galaxy Sample (Jansen et al., 2000). The solid line indicates the $L - Z$ relation obtained by Tremonti et al. (2004) for local galaxies. The magenta dashed line represents the linear fit ($12 + \log(\mathrm{O/H})$= a·M_{B} + b) through our data points except for J154925.24-003717.8 and J003806.59+000421.9.

Appendix B

Stellar Population Properties for a Sample of Hard X-ray AGNs

Based on Morelli L., **Calvi V.***, Masetti N., Parisi P., Landi R., Maiorano E., Minniti D., and Galaz G., 2013, A & A: 556, A135*

B.1 Introduction

The massive work carried out by Masetti et al. (2004, 2006d,c,b,a, 2008, 2009, 2010, 2012), Landi et al. (2007), Parisi et al. (2009, 2012), and Maiorano et al. (2011) (hereafter, Papers I-XIII) supplied the astronomical community with a catalog[1] of hard X-ray sources for which the identification and the main physical parameters, computed using the multiwavelength information available in the literature, are provided. In total, more than 250 objects in the 20-200 keV range observed during the *INTEGRAL* (Winkler et al., 2003) and *Swift* (Gehrels et al., 2004) missions were studied and classified with ∼200 of them using the IBIS instrument (Ubertini et al., 2003) and with ∼60 using the BAT instrument (Barthelmy, 2004). In particular, an optical follow-up to unveil the nature of most of these objects was mandatory, since only optical spectra permit an accurate source classification and provide fundamental parameters. In Papers I-XIII, we measured the flux of the most important emission lines existing in the optical part of the spectrum with the main aims of identifying the object and of investigating the properties of the host galaxies. For example we measured properties such as, the Compton nature of these objects and an estimate of the mass of the central black hole in broad emission line AGNs. Nevertheless, information regarding the stellar population of the galaxy hosting the AGNs is missing in the catalog. Therefore, stellar populations are new, important, and complementary pieces of information, which should be included in the catalog and which shed light on the properties of the AGNs and their host galaxies.

In the last decade, several authors focused their attention on the relation between the nuclear activity and the star formation rate (Ivanov et al., 2000; González Delgado et al., 2001; Joguet et al., 2001; Ho et al., 2003; La Mura et al., 2009; Draper and Ballantyne, 2011; Cracco et al., 2011; Vaona et al., 2012). This relation can give important clues about the fate

[1]The up-to-date version of this catalog is available on http://www.iasfbo.inaf.it/∼masetti/IGR/main.html

of the gas that fuels the central black hole (Hopkins and Hernquist, 2006) and its influence on the central part of the host galaxy (Sarzi et al., 2005).

According to simulations (Di Matteo et al., 2005; Springel et al., 2005; Fontanot et al., 2011), a merging episode involving galaxies rich in gas induces radial gas inflows, which feed the black hole, and consequently, enhance the central star formation. Then, the AGN feedback wipes out basically all the remaining gas and dust (Hopkins and Hernquist, 2006; Rigopoulou et al., 2009), halting the star formation. However, the proposed scenarios in the past years became even more complicated. The model by Novak et al. (2011) showed that the AGN is a cyclic process. To account for the observational properties of this class of objects, it is necessary to consider different processes acting on different scales, such as: mechanic feedback that is a few pc from the nucleus, radiative feedback and consequent cooling flow of gas on a scale of few kpc, and SN winds, which are relevant on scales of tens of kpc (Ciotti et al., 2010).

Some observational works (Tremonti et al., 2007; Feruglio et al., 2010; McKernan et al., 2010) stated that AGNs could have played a relevant role in halting the star formation in massive host galaxies. Schawinski et al. (2007) suggested that the feedback could stop any star formation activity in the nucleus of the hosting galaxy and even beyond if the accretion of material onto the central black hole is powerful enough. As a consequence of this process, the stellar populations in the AGN host galaxies are old, and as suggested by Schawinski et al. (2007) and Faber et al. (2007), this causes the host galaxies to move from the blue cloud to the red sequence of the galaxy optical color diagram. Moreover, Bluck et al. (2011) computed the average energy output per galaxy due to AGN showing that this is at least 35 times larger than the binding energy of a typical massive galaxy.

Since an invaluable piece of information to understand the processes of formation and evolution of galaxies is imprinted in their stellar populations, some studies in recent years aimed to investigate the properties of the stellar populations of the host galaxy have been conducted. By studying the stellar populations in the nuclear region and their radial profile of a sample of AGNs that are mostly Seyfert 2 (hereafter Sy2), it was found that about one third of the galaxies in the sample shows an old bulge-like stellar population in the center (Ho et al., 2003; Cid Fernandes et al., 2004; Chen et al., 2009) and that the number of objects with an old stellar population increases to about two-thirds when the outer regions of low luminosity AGN hosts are investigated (Cid Fernandes et al., 2004).

In a recent paper, La Mura et al. (2012) investigated the connection between stellar population and mass in a large sample of type 1 and 2 AGNs. They found that the mass of the stellar component is a key ingredient to study the star formation history (Mannucci et al., 2010) of galaxies and taking this into account, they suggested an evolutionary sequence moving from starburst galaxies to AGNs (Davies et al., 2007; Schawinski et al., 2007).

Essentially, the observations show a variety of results even when they are restricted to the analysis of only Sy2 galaxies for which broad lines are weak or absent (Lawrence et al., 1987). This could be due to the difficulties in choosing homogeneous samples of objects and to an intrinsic complexity in interpreting the results, especially considering that the AGN theoretical model is still debated.

In this scenario, our results will be useful for understanding the properties of AGN hosts and the consequences of the AGN feedback on them, as well as to help in testing the predictions of theoretical models. The aim of this paper is, therefore, to present the stellar populations properties for a sample of AGN hosting galaxies, which were identified as the optical counterparts of hard X-ray emitting sources in Papers I-XIII.

In Section 2, we described the sample selection and the characteristics of the data. In Section 3, we illustrated the tools we used and the procedure we followed to obtain the final results. In Section 4, we presented and discussed our results concerning the stellar populations, in particular for the age and the metallicity derived from the stellar population fitting procedure. Finally, we summarized results and conclusions in Section 5.

B.2 Sample Selection and Data Properties

B.2.1 The Sample

The aim of this paper is to complete the catalog of optical counterparts of hard X-ray sources detected in the massive survey performed by our collaboration (papers I-XIII), for which we have already detected and measured the emission line features with the study of the stellar population properties of the hosting galaxy. Consequently, complete information in the X-ray and optical range of the spectrum will be available for a subsample of objects in the catalog. To choose the galaxy sample, we started selecting all those objects identified as AGNs in Papers I-XIII. In detail, these hard X-ray sources were selected among those with an unidentified nature that belonged to the *INTEGRAL* and *Swift* surveys (e.g., Bird et al. 2010; Cusumano et al. 2010). The selection method consisted in choosing sources containing a single soft X–ray object within their arcmin-sized hard X–ray error circle. According to Stephen et al. (2006) this is the lower-frequency counterpart of the high-energy emission with a very high degree of probability. Given the arcsecond precision with which the position of soft X-ray sources is available, this technique reduces the sky area for the search for the optical counterpart by a factor $>10^3$. This easily allows us to pinpoint the actual optical counterpart on which we could, eventually, perform optical spectroscopy to determine its nature. Our survey detected 158 AGNs of a different type. Since the study of the stellar populations was not the principal aim of the investigation when the spectra were acquired, many of them are characterized by a short exposure time that does not guarantee either a quality or a signal-to-noise ratio (S/N) suitable for performing a reliable analysis of their stellar population properties. Out of the initial sample of optical counterparts of AGNs, we considered only the extracted spectra with $S/N \geq 20$ for which the important absorption lines (i.e. H_β, Mg, and Fe) were not strongly contaminated by broad-band emission lines or residual sky subtraction for the following analysis. Therefore, the final sample was comprised of 20 objects, classified mostly as Sy2, with soft and hard X-ray emission detected, and with gas emission lines measured. All these galaxies lie in the redshift range of $0.008 \leq \mathrm{z} \leq 0.3$, and their properties are listed in Table B.1.

B.2.2 The data

The spectroscopic observations of the sample galaxies were carried out with a variety of setups using different telescopes. These were as follow:

- the 3.58m Telescopio Nazionale Galileo (TNG) in La Palma, Spain;
- the 2.1m telescope of the Observatorio Astronomico Nacional in San Pedro Martir, Mexico;
- the 1.5m at the Cerro Tololo Interamerican Observatory (CTIO), Chile; and

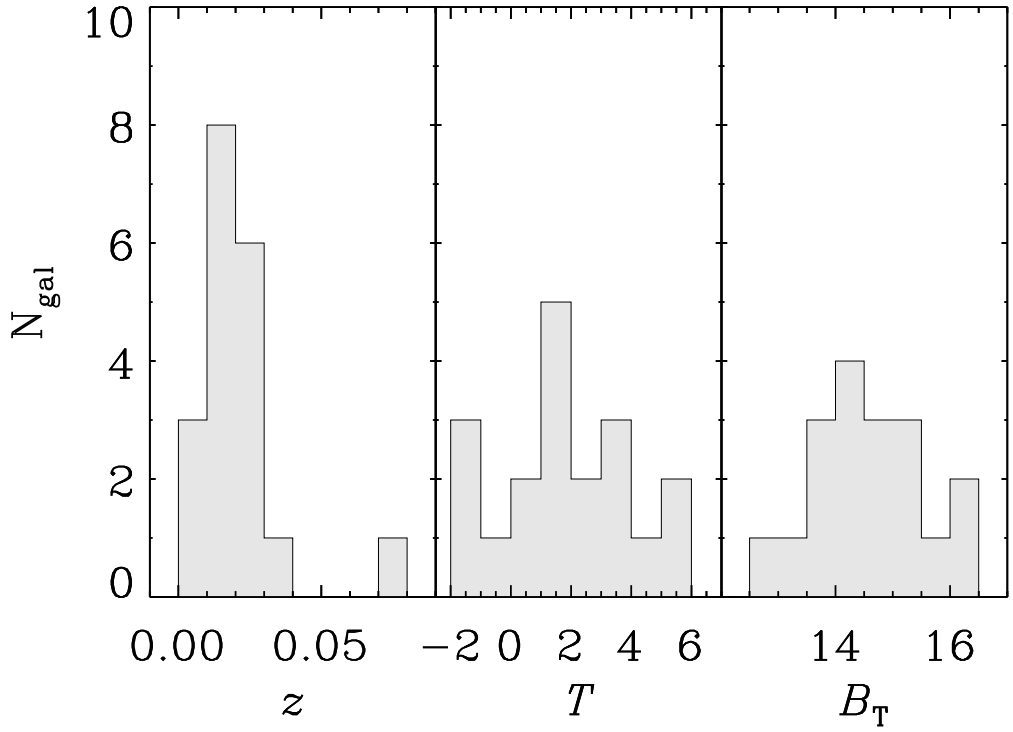


Figure B.1: Histograms showing the distribution of the main properties of the galaxies in our sample. From left to right, we show redshift z, morphological type T, and total magnitude in the B-band B_T.

- the 1.52m Giandomenico Cassini telescope of the Astronomical Observatory of Bologna in Loiano, Italy.

Some spectra were also retrieved from the Sloan Digital Sky Survey (SDSS) archive (Adelman-McCarthy et al., 2006, 2008) and from the Six-degree Field Galaxy Survey (6dFGS) archive (Jones et al., 2004).

We refer to Papers I-XIII for detailed explanations on the observing setup, data reduction, calibration, and previous analysis. In column 6 of Table B.1, the reference paper for each object is listed. In this paper, we just summarized the basic useful properties of the spectra used in the following analysis.

The sample consists of 20 spectra of galaxies acquired in long slit mode. The wavelength range between 3800 and 7500 Å was covered with a reciprocal dispersion between ~ 0.8 and ~ 5.7 Å pixel^{-1} after pixel binning. This corresponds to an instrumental velocity dispersion within the range $61 \text{ km s}^{-1} \lesssim \sigma_{\rm inst} \lesssim 300 \text{ km s}^{-1}$ at 5500 Å.

B.3 Measurements of the Stellar Populations

The stellar population properties, namely age and metallicity, were measured mainly from the following absorption features: H_α line (λ 6563Å), H_β line (λ 4861Å), H_γ line (λ 4340Å), H_δ line (λ 4102Å), Mg I line triplet ($\lambda\lambda$ 5164, 5173, 5184Å), and Fe lines ($\lambda\lambda$ 5270, 5335Å).

With few exceptions, it was not possible to obtain information either from the blue part ($\lesssim 3800$Å) of the spectrum because of the low efficiency of the optics, or from the red part ($\gtrsim 7500$Å) due to the residuals of the strong emission lines of the sky in this region.

As done by Onodera et al. (2012), we applied the penalized pixel fitting (`pPXF`; Cappellari and Emsellem 2004[2]), including the linear regularization of the weights (Press et al., 1992) and the Gas AND Absorption Line Fitting (`GANDALF`, Sarzi et al., 2006) `IDL`[3] packages which are adjusted for the sample spectra, to derive both the distribution of the mass fraction in different age and metallicity bins.

Even if the featureless continuum and broad lines are weak or absent in Sy2 galaxies (Lawrence et al., 1987), we allowed the code to use also broad band components in the fitting procedure. To account for the effect of dust and possible residuals of the data reduction procedure, we adopted a low order multiplicative polynomial in the template fitting. This has the advantage to make our method more sensitive to the absorption lines than to the continuum shape and, therefore, less sensible to the effects of reddening. However, we could not completely rule out the possibility of underestimating the weight of very reddened young stellar components in deriving the composite stellar populations.

For each spectrum, we fitted a linear combination of 156 template stellar spectra (Vazdekis et al., 2010) from the MILES (Sánchez-Blázquez et al., 2006) library ($\mathrm{FWHM} = 2.54$ Å spectral resolution Beifiori et al. 2011) to the observed galaxy spectrum by performing a χ^2 minimization in pixel space. Since the resolution of the sample spectra spans a wide range of values (always lower than the template stellar spectra) for each galaxy, it was necessary to convolve each template with the line-of-sight velocity distribution (LOSVD) and to rebin both the template and the galaxy spectrum to match their dispersions before running the fitting code.

We adopted the Salpeter initial mass function (Salpeter, 1955), the 26 ages ranging from 1 to 17 Gyr, and six metallicities $[M/\mathrm{H}]$ from -1.71 to 0.22. Simultaneously, we fitted the observed spectra using emission lines in addition to the stellar templates.

The extracted spectra on which we performed the measurement were the same used to derive the emission line properties in Papers I-XIII. This guarantees the analysis to be in the same spatial regions used for the detection and measurements of the emission lines. The linear scale on which we were measuring galaxy properties depends on the combination of redshift and slit aperture and ranges from 200 pc to 2 kpc. Swift J0811.5+0937 is the only exception, having a linear scale of 6 kpc. Based on the above, the region we were investigating was large enough to allow a study of the stellar populations of the bulge central region surrounding the AGN in almost all the cases.

The example of the fitting procedure shown in Figure B.2 (left panel) refers to galaxy PBC J0954.8+3724 and proves the good quality of the fit we obtained for all the galaxies. From the fitted single stellar population (SSP), we, then, derived the stellar mass fraction within each age and metallicity interval (right panel).

B.4 Results and Discussion

We applied the procedure described in Section B.3 to all the galaxies in our sample. In general, the distribution of the age and metallicity templates used by the code was, smooth

[2]Program available on http://www-astro.physics.ox.ac.uk/ mxc/idl/

[3]Interactive Data Language is distributed by Exelis Visual Information Solutions.

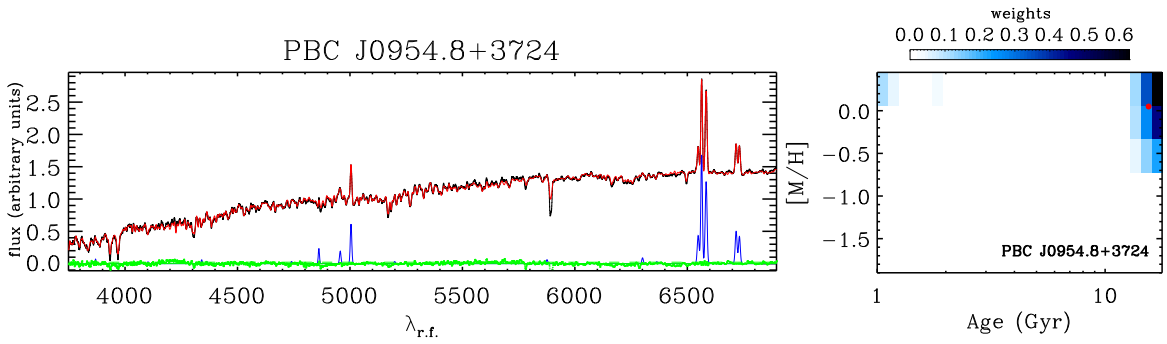


Figure B.2: Left panel: The galaxy spectrum (black line) is compared to the fitted linear combination of template stellar spectra (red line) and emission lines (blue line). The green dots show the residuals. Right panel: Age and metallicity obtained from the spectral fitting. The color scale refers to the mass fraction in each bin of age and metallicity. The red dot represents the mass-weighted age $\langle t/Gyr \rangle_M$ and the mass-weighted metallicity $\langle\, [M/\mathrm{H}] \,\rangle_M$.

for most of the galaxies and indicated that the bulk of stars in the considered region tended to be old.

To quantify this effect, we derived the mass-weighted age $\langle t/Gyr \rangle_M$ and the mass-weighted metallicity $\langle\, [M/\mathrm{H}] \,\rangle_M$ of its stellar population for each galaxy. These values are listed in Table B.2 and plotted in Figures B.3, B.4, and B.5. The color scale refers to the mass fraction in each bin of age and metallicity. The errors on the age and metallicity given in Table B.2 were obtained from photon statistics and CCD readout noise, and they were calibrated through a series of Monte Carlo simulations.
In detail, almost all galaxies are characterized by an old stellar population with an age range from 8.4 Gyr to 15.7 Gyr. There are only two galaxies (PBC J1546.5+6931 and PBC J0041.6+2534) with $\langle t/Gyr \rangle_M < 7$. However, PBC J1546.5+6931 shows a clear bimodality in the distribution and, for this reason, $\langle t/Gyr \rangle_M$ and $\langle\, [M/\mathrm{H}] \,\rangle_M$ are not properly describing its stellar population. Therefore, we independently derived the typical parameters of the two distinct stellar populations for this object. The younger one is characterized by $\langle t/Gyr \rangle_M = 3.6$ and $\langle\, [M/\mathrm{H}] \,\rangle_M = 0.18$, and it contributes to $\sim 70\%$ of the total mass of the galaxy. The remaining $\sim 30\%$ of the galaxy mass is due to a older population ($\langle t/Gyr \rangle_M = 14.8$) with subsolar metallicity ($\langle\, [M/\mathrm{H}] \,\rangle_M = -0.27$).
Even though this is the most representative example of bimodality in the stellar population distribution, there are another two galaxies, namely PBC J0919.9+3712 and PBC J1345.4+4141, that show a second stellar population accounting for more than 10% of the total galaxy mass. The final results for the three galaxies with two separate stellar populations are listed in Table B.3. As can noted in Table B.3, the less massive component is a young one for both PBC J0919.9+3712 and PBC J1345.4+4141, which accounts for 12% and 15% of the total mass, respectively. The linear scales for these two objects are the smallest in our sample (~ 200 pc), and this could have increased our sensibility in detecting a young component existing in the nuclear region of the galaxy.

It is interesting to note that an additional young component was also found by Cid Fernandes et al. (2004) in $\sim 30\%$ of the galaxies in their sample. Their percentage, which is higher than what we found, could be ascribed in that we were observing a larger region of

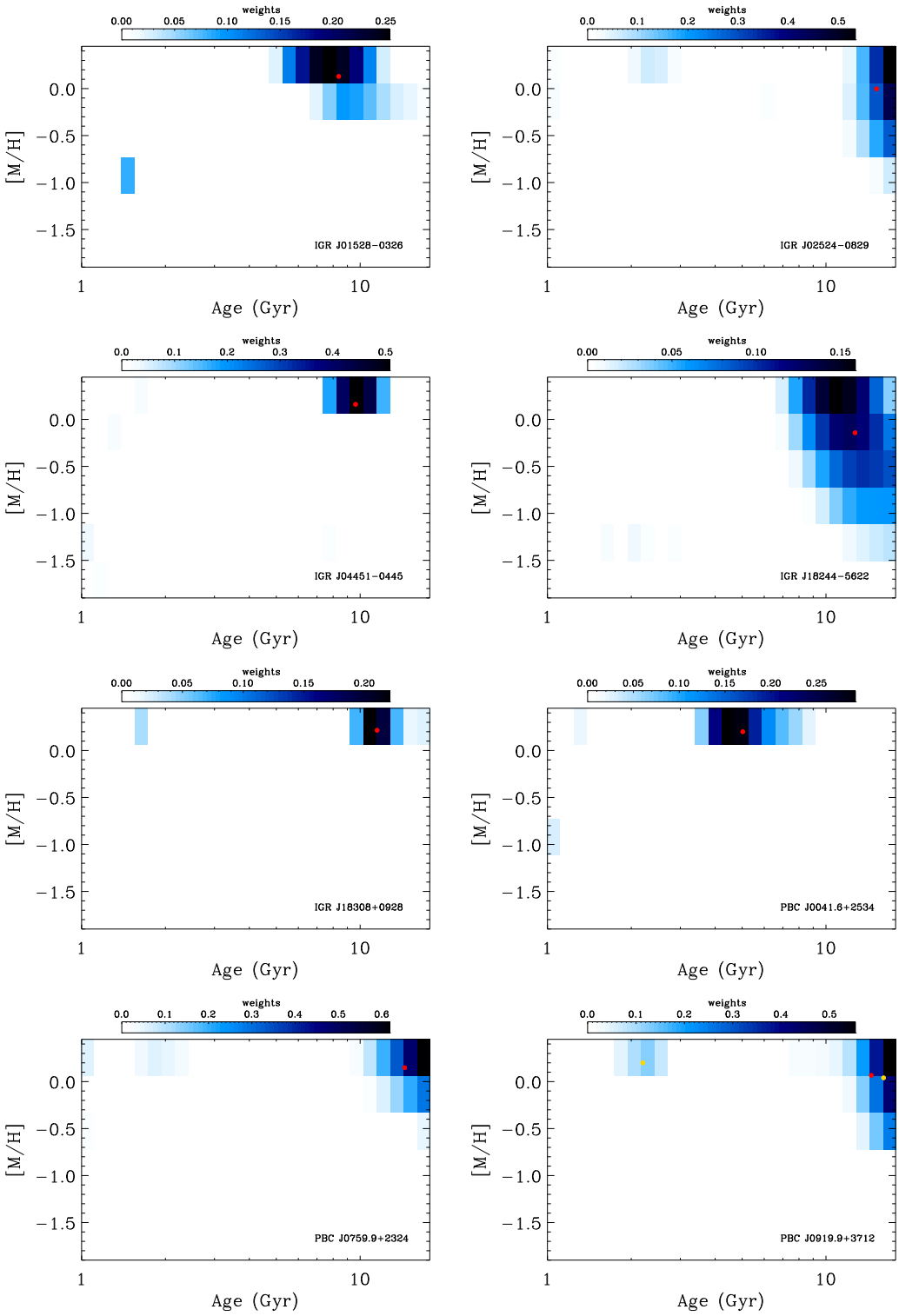


Figure B.3: Age and metallicity obtained from the spectral fitting for IGR J01528-0326, IGR J02524-0829, IGR J04451-0445, IGR J18244-5622, IGR J18308+0928, PBC J0041.6+2534, PBC J0759.9+2324, and PBC J0919.9+3712. The color scale refers to the mass fraction for each bin of age and metallicity. The red dot represents the mass-weighted age $\langle t/Gyr \rangle_M$ and the mass-weighted metallicity $\langle\, [M/\mathrm{H}]\, \rangle_M$ for each galaxy. The yellow dots represent the mass-weighted age $\langle t/Gyr \rangle_M$ and the mass-weighted metallicity $\langle\, [M/\mathrm{H}]\, \rangle_M$ in the case of two distinct stellar populations considered for the young and old component (see Sect. B.4).

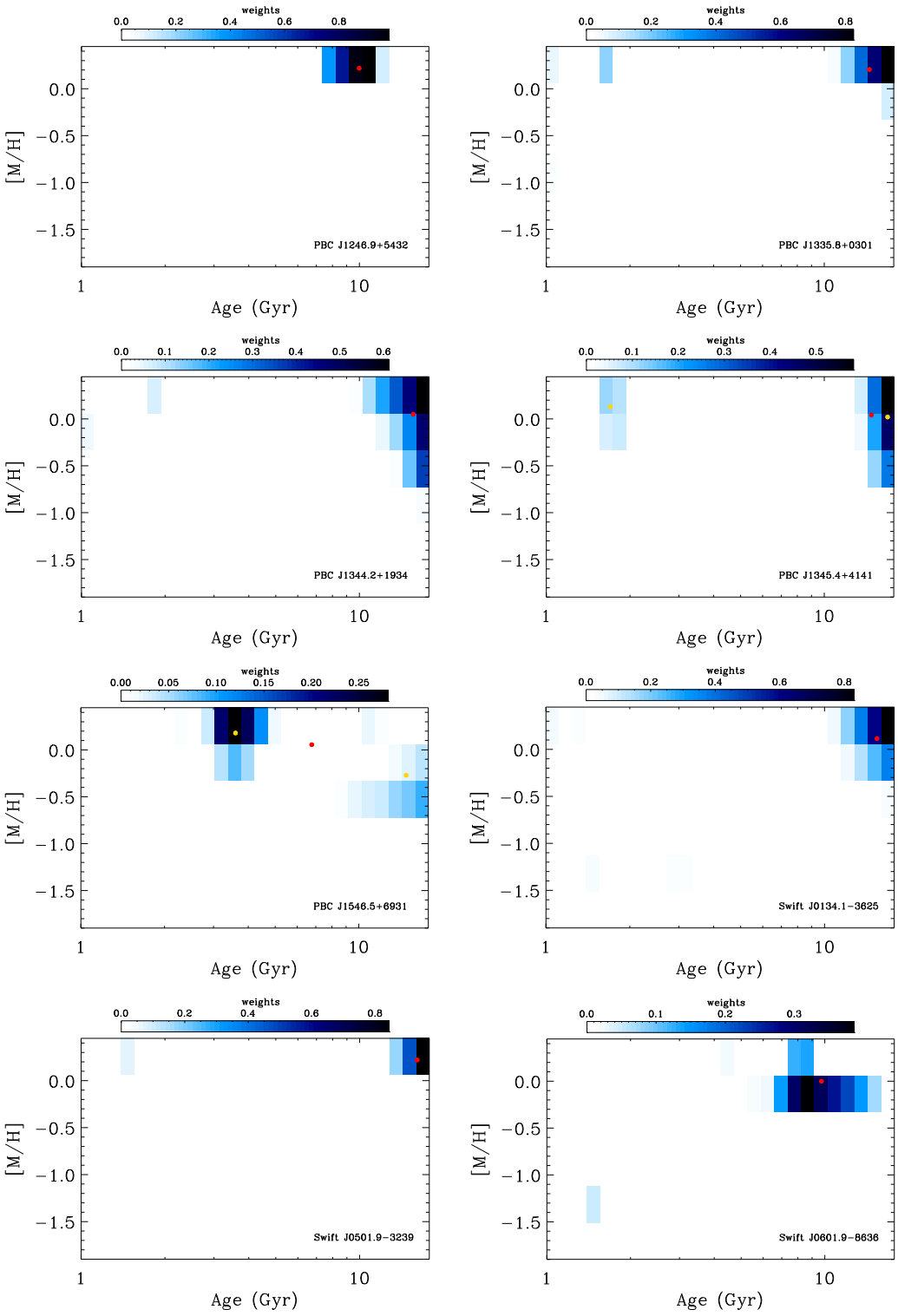


Figure B.4: The same as in Figure B.3 for PBC J1246.9+5432, PBC J1335.8+0301, PBC J1344.2+1934, PBC J1345.4+4141, PBC J1546.5+6931, Swift J0134.1-3625, Swift J0501.9-3239, and Swift J0601.9-8636.

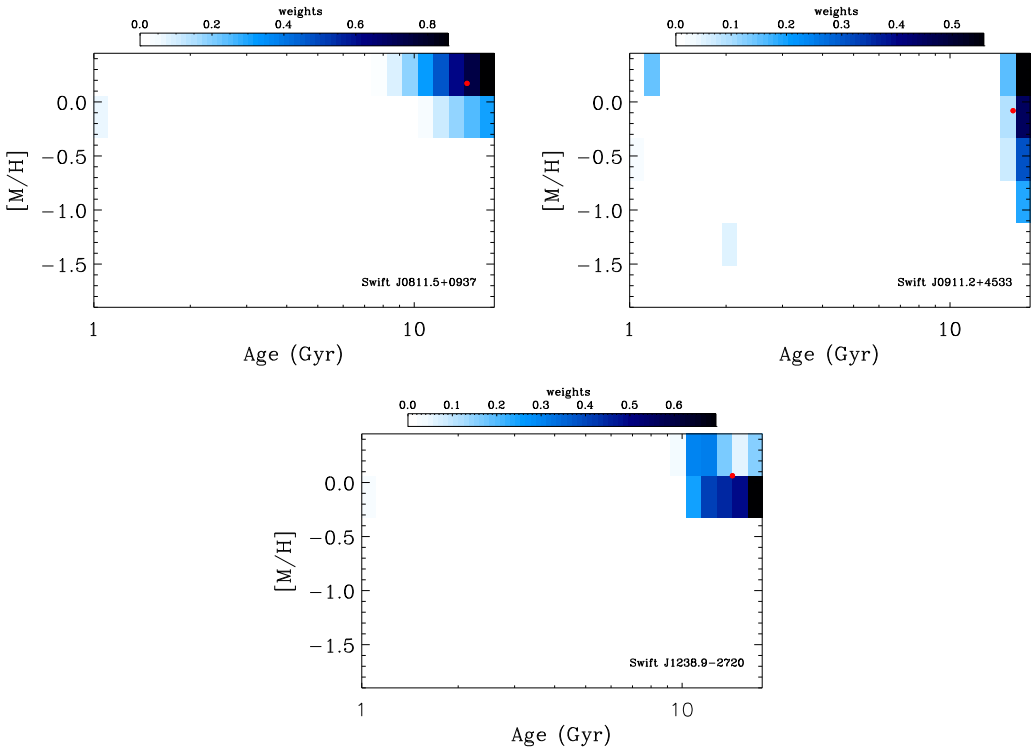


Figure B.5: The same as in Figure B.3 for Swift J0811.5+0937, Swift J0911.2+4533, and Swift J1238.9-2720.

the galaxy compared to what had been performed. For this reason, we were less sensitive to the young stellar populations that are possibly increasing towards the nucleus of the galaxy (Cid Fernandes et al., 2005). This could be the reason for the young component detected in PBC J0919.9+3712 and PBC J1345.4+4141. On the other hand, this is not the case for PBC J1546.5+6931, which shows a dominant young population and whose linear scale is bigger (~ 2 kpc). An explanation for the global young stellar population of this galaxy must be determined from its formation and evolution history.

Finally, the old ages that we derived for the majority of the galaxies in the sample agree with the results obtained for a large sample of infrared selected AGNs by Chen et al. (2009). Those objects, spanning different spectral classes and luminosities, all clearly show the old stellar population dominating the total mass and no relevant contributions from the young one. These results indirectly confirm that Sy2s reside very close to or even lie in the red sequence of galaxies (Schawinski et al., 2007).

The values of the mass-weighted age and metallicity that we obtained when considering only one stellar population (i.e. the one dominating the mass) are shown in Figure B.6 on the left and right panel, respectively. The large majority of the sample galaxies are characterized by a slightly supersolar mass-weighted metallicity (Figure B.6, right panel). The number distribution has a median value of $[M/\mathrm{H}] = 0.08$ and spreads from super ($[M/\mathrm{H}] = 0.2$) to sub-solar values ($[M/\mathrm{H}] = -0.2$). However, it should be noted that the fit used several different metallicities for the SSP, suggesting the possible existence of more metallicity components. Even when considering two distinct stellar populations for

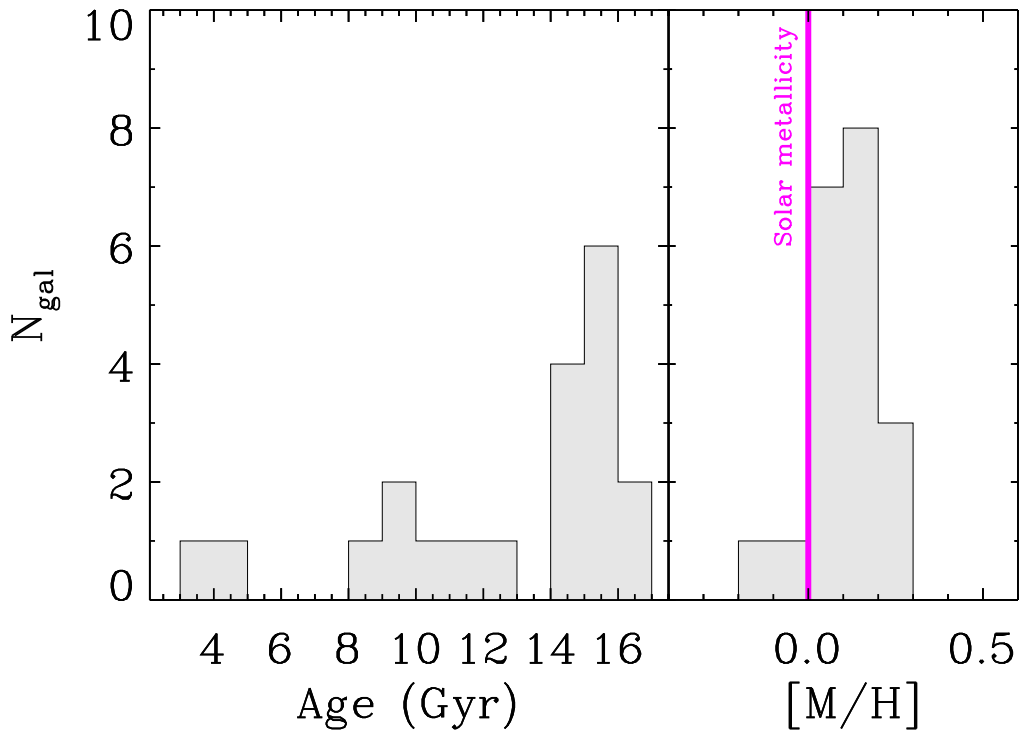


Figure B.6: Distribution of mass-weighted ages (left panel) and metallicities (right panel) for the dominant stellar component in the sample of galaxies. The magenta vertical line indicates the value of the solar metallicity.

PBC J0919.9+3712, PBC J1345.4+4141, and PBC J1546.5+6931, the values of the metallicity remain inside this range (Table B.3). Once again, this is consistent with the analysis done by Chen et al. (2009) and La Mura et al. (2012), who claimed increasing metallicity from starburst toward Seyferts and LINERs.

As a further step, we converted the mass-weighted ages and metallicities to the corresponding luminosity-weighted values. To perform the conversion, we adopted the M/L ratios tabulated for the SDSS g filter by Maraston (2005). As we expected, the luminosity-weighted ages are slightly younger than the mass-weighted ones, but they remain, with few exceptions, globally old spanning the range between 6 and 12 Gyr. In Figure B.7, we plotted the histogram of the luminosity-weighted ages compared to those derived for the high surface brightness (HSB) sample in Morelli et al. (2008), which are characterized by a similar morphological type distribution.

Even though the number of galaxies in our sample does not allow us to trace a firm statistical conclusion, it is interesting to note that the bulges in our sample are globally older than those hosted in normal spirals. The metallicity of the bulges of LSB discs spans a large range of values from high ($[Z/\mathrm{H}] = 0.30$ dex) to sub-solar ($[Z/\mathrm{H}] = -0.2$ dex) with a peak around slightly super-solar values (Figure B.7, right panel). The distribution of the metallicity in the galaxy sample is, instead, similar to the one derived for the bulges of HSB galaxies (Morelli et al., 2008).

We performed an additional analysis by looking for a possible correlation between the

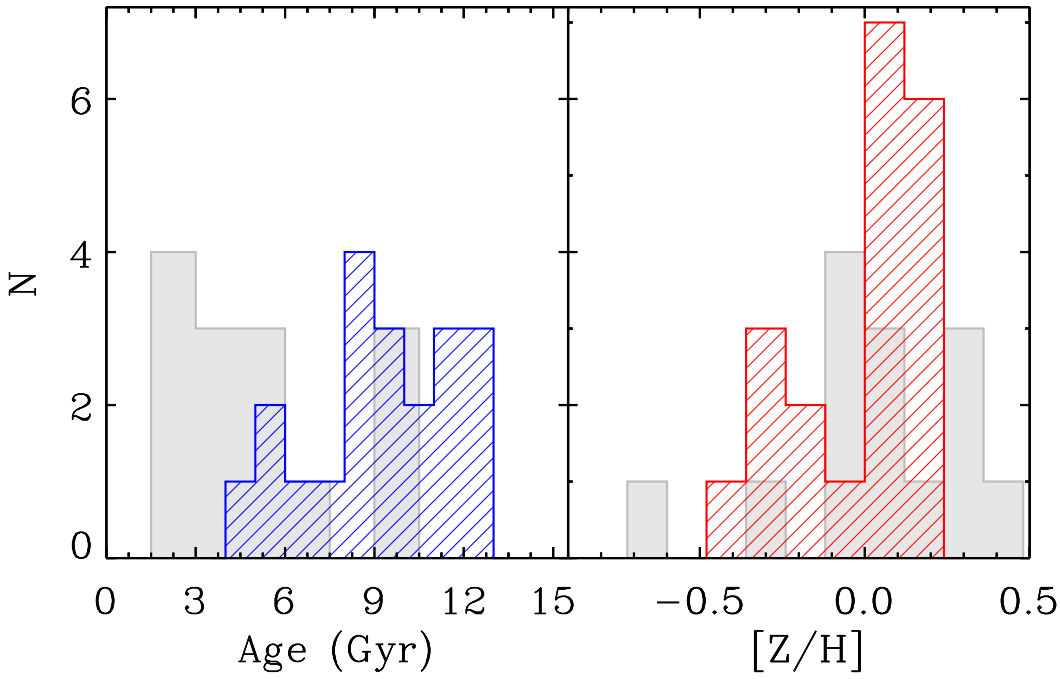


Figure B.7: The shaded histograms show the distribution of luminosity-weighted ages (left panel) and metallicities (right panel) for our sample of galaxies compared to the corresponding values derived by Morelli et al. (2008) for a similar sample of spirals (grey solid histogram).

mass-weighted ages, metallicities, and the morphological type of our galaxies. Cid Fernandes et al. (2004) did not find any relevant correlation between the host morphology and the stellar population in the nuclear region of Sy2s, but the situation is less clear in the case of non-active galaxies . Studying a sample of spiral galaxies, Thomas and Davies (2006) and Morelli et al. (2012) did not observe any correlation between the age and metallicity of the stellar population in the central region of the bulge and galaxy morphology, whereas Ganda et al. (2007) and Morelli et al. (2008) found a mild correlation with the early-type galaxies ($T < 0$) being older and more metal rich than spirals ($T \geq 0$). The galaxies in our sample span a wide range of T type values ($-2 \leq T \leq 5$) that are homogeneously distributed without any decreasing trend in number when going from the early to the late type, as observed in the Storchi-Bergmann et al. (2001) sample.

We did not observe any relevant trends between the galaxy morphological type and the age or metallicity of our galaxies (Figure B.8). In some way, it is tempting to say that the AGN feedback acts homogenizing the stellar populations in the central (few kpc) region of galaxies with different morphological type. However, the low number of galaxies with $T \leq 0$ and the shallow relation between morphology and stellar populations showed by non-active galaxies, prevented us from claiming any strong conclusion on this aspect.

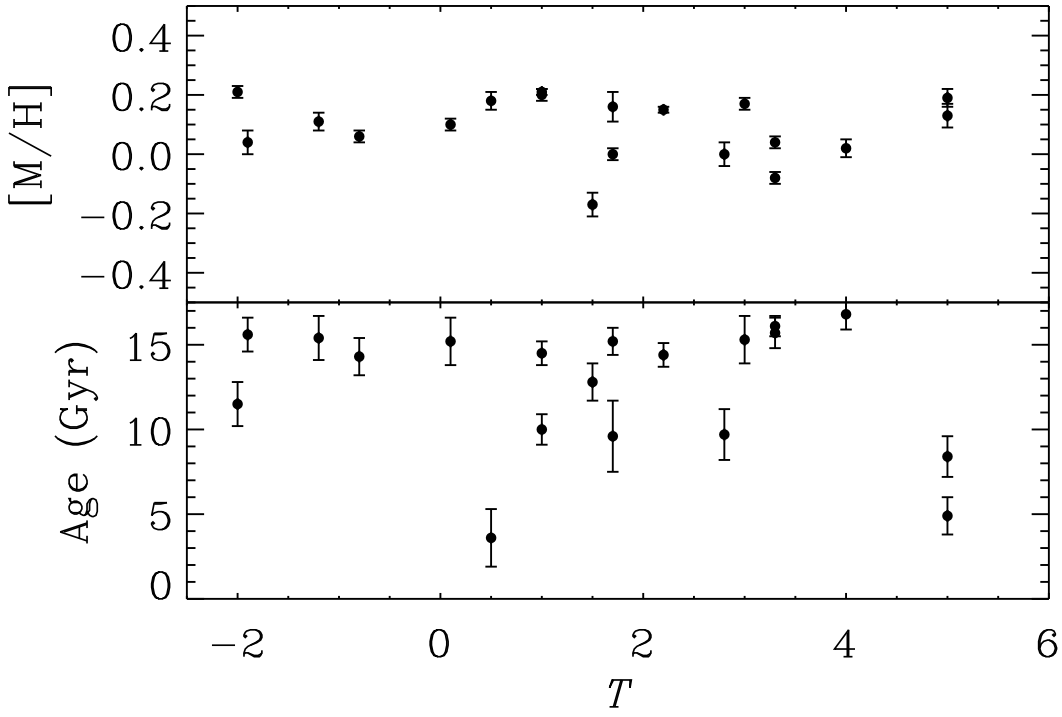


Figure B.8: Correlation between morphological type and stellar population properties. Top panel: the values for the mass-weighted metallicities are plotted as a function of the morphological type T. Bottom panel: same as in top panel for the mass-weighted ages in Gyr.

B.5 Summary and Conclusions

In the past years, our collaboration unveiled the nature of more than 250 X-ray emitting sources; 158 of which were identified as AGNs. In this paper, we presented the stellar population analysis of the host galaxy, which was performed on a carefully selected sample of 20 objects with the goal of understanding the still debated connection between the central AGN and the properties of the host galaxy, and the effect of the AGN feedback. The detailed analysis of the stellar population can give more important constraints on this topic.

The area on which we measured the galaxy properties ranges from 200 pc to 2 kpc. This allowed us to measure the stellar populations of the bulge close to the center of the galaxy, where the effects of the AGN feedback on the host are expected to be relevant.

The spectral fitting method, based on `PpXF` and `GANDALF`, was applied to the galaxy spectra to measure the stellar populations properties of the sample. In particular, we obtained a mass-weighted age and a mass-weighted metallicity for each galaxy. The values of the mass-weighted metallicity span the range of $-0.2 \leq [M/\mathrm{H}] \leq 0.2$ with a median value of $[M/\mathrm{H}] = 0.08$. The large majority of our objects (19 galaxies out of 20, or 95% of the galaxies) show an old stellar population with ages older than 8 Gyrs. Three of them are characterized by a bimodal distribution with a non-negligible contribution from young stars. We found that PBC J1546.5+6931 is dominated by a young stellar population which account for $\sim 70\%$ of the total mass, while, the young stellar component is less massive than the old one in the case of PBC J0919.9+3712 and PBC J1345.4+4141. For the former galaxy the

nature of the young stellar component is probably related to the formation and evolution of the galaxy itself. Regarding the latter galaxies, there would be a possible star-forming activity in their central region. Even after investigating the luminosity-weighted ages for the galaxies in our sample, the old nature of their stellar population is confirmed. The comparison of their ages with those obtained for a similar sample in terms of morphological type showed that bulges hosted in AGN are globally older than those hosted in non-active counterparts. However, it should be noted that, as expected, the contribution of the young component to the total light of the galaxy is greater than that to the total mass.

Our results suggest that AGN feedback acts on the first kpc of the galaxy, decreasing the efficiency of its star formation through different processes. The combination of truncation and suppression (Schawinski et al., 2009) could be responsible for disrupting the gas starforming reservoir (Davis et al., 2012) in early times, when the gas is wiped out from the strong AGN emission in the center (Di Matteo et al., 2005; Springel et al., 2005; Fontanot et al., 2011). In recent times, the phase of the AGN regulates and quenches the residual star formation of the host galaxy (Sturm et al., 2011), maintaining the global red color and old stellar populations that are also observed in this work.

Radiative feedback and cooling flow of gas to the center could also, as proposed by Ciotti et al. (2010) and Novak et al. (2011), exhaust the gas in the few central kpcs region of the galaxy and decrease the star formation. The final consequence of this process is that the bulk of the observed stellar populations in this region is, with very few exceptions, old.

These scenarios are indirectly supported by the lack of relations between the stellar population properties and the morphological ones, in the sense that in the central region the existence of an AGN influences the properties of the galaxy much more than the formation and evolution, as described by the morphological type.

Galaxy name (1)	Galaxy alt. name (2)	Type (3)	T (4)	Class (5)	z (6)	$B_{\rm T}$ (mag) (7)	Source (8)
IGR J01528-0326	PGC 6966	SA(s)c?	5.0	likely Sy2	0.017	14.11	Masetti et al. (2008)
IGR J02524-0829	LEDA 10875	Sa?	1.7	Sy2	0.017	15.03	Masetti et al. (2009)
IGR J04451-0445*	LEDA 1053623	S?	1.7	likely Sy2	0.076	17.10	Masetti et al. (2010)
IGR J18244-5622	IC 4709	Sa	1.5	Sy2	0.017	14.42	Masetti et al. (2006b)
IGR J18308+0928	LEDA 1365707	E?	-2.0	Sy2	0.019	15.06	Masetti et al. (2010)
PBC J0041.6+2534	NGC 0214	SAB(r)c	5.0	Sy2/LINER	0.015	12.94	Parisi et al. (2012)
PBC J0759.9+2324	MCG+04-19-017	Sab	2.2	Sy2	0.029	14.80	Parisi et al. (2012)
PBC J0919.9+3712	IC 2461	Sb	3.3	Sy2	0.008	14.63	Parisi et al. (2012)
PBC J0954.8+3724	IC 2515	Sb	3.0	Sy2	0.019	15.02	Parisi et al. (2012)
PBC J1246.9+5432	LEDA 43101	Sa	1.0	Sy2	0.017	13.60	Parisi et al. (2012)
PBC J1335.8+0301	NGC 5231	SBa	1.0	Sy2	0.022	14.29	Parisi et al. (2012)
PBC J1344.2+1934	PGC 048674	E?	-1.9	Sy2/LINER	0.027	15.58	Parisi et al. (2012)
PBC J1345.4+4141	NGC 5290	Sbc?	4.0	Sy1.9	0.009	13.30	Parisi et al. (2012)
PBC J1546.5+6931	PGC 2730634	S?	0.5	Sy1.9	0.037	16.08	Parisi et al. (2012)
Swift J0134.1-3625	LEDA 5827	SA0	-1.2	Sy2	0.029	14.03	Parisi et al. (2009)
Swift J0501.9-3239	LEDA 17103	SB0/a?(s)	0.1	Sy2	0.013	13.87	Parisi et al. (2009)
Swift J0601.9-8636	LEDA 18394	Sb?	2.8	Sy2	0.006	13.51	Landi et al. (2007)
Swift J0811.5+0937	USNO-A2.0	-	-	XBONG	0.286	-	Parisi et al. (2009)
Swift J0911.2+4533	LEDA 2265450	S?	3.3	Sy2	0.027	16.47	Parisi et al. (2009)
Swift J1238.9-2720	ESO506G027	S0	-0.8	Sy2	0.024	14.66	Landi et al. (2007)

Table B.1: Properties of the galaxies in the sample. Col. (1): name of the galaxy. Col. (2): alternative name of the galaxy. Col. (3): morphological type derived from the Third Reference Catalogue of Bright Galaxies (de Vaucouleurs et al., 1991) and from the Hyperleda database (Paturel et al., 2003). Col. (4): numerical morphological type T from the Hyperleda database. Col. (5): classification of the AGN. Col. (6): redshift of the object. Col. (7): total apparent magnitude in the B-band derived from the Hyperleda database. Col. (8): reference source for the galaxy. *The identification of this object is still not confirmed.

Galaxy name (1)	Age (Gyr) (2)	Metallicity [M/H] (3)
IGR J01528-0326	8.4±1.2	0.13±0.04
IGR J02524-0829	15.2±0.8	0.01±0.02
IGR J04451-0445	9.6±2.1	0.16±0.05
IGR J18244-5622	12.8±1.1	-0.17±0.04
IGR J18308+0928	11.5±1.3	0.21±0.02
PBC J0041.6+2534	5.0±1.1	0.19±0.03
PBC J0759.9+2324	14.4±0.7	0.15±0.01
PBC J0919.9+3712	14.5±0.6	0.06±0.02
PBC J0954.8+3724	15.5±1.4	0.04±0.02
PBC J1246.9+5432	10.0±0.9	0.21±0.01
PBC J1335.8+0301	14.5±0.7	0.20±0.02
PBC J1344.2+1934	15.6±1.0	0.04±0.04
PBC J1345.4+4141	14.8±0.9	0.05±0.03
PBC J1546.5+6931	6.8±1.7	0.05±0.03
Swift J0134.1-3625	15.4±1.3	0.11±0.03
Swift J0501.9-3239	15.2±1.4	0.10±0.02
Swift J0601.9-8636	9.7±1.5	0.00±0.04
Swift J0811.5+0937	14.6±1.1	0.17±0.03
Swift J0911.2+4533	15.7±0.9	-0.08±0.02
Swift J1238.9-2720	14.3±1.1	0.06±0.02

Table B.2: Mass-weighted age and metallicity measured for the galaxies in our sample. Col. (1): object name. Col. (2): mass-weighted age in Giga-years derived from the stellar population fitting. Col. (3): mass-weighted metallicity derived from the stellar population fitting.

Galaxy	Age		Metallicity		Mass
name	Young	Old	Young	Old	Old
	(Gyr)		[M/H]		%
PBC J0919.9+3712	2.2	16.1	0.22	0.04	88
PBC J1345.4+4141	1.7	16.8	0.13	0.02	85
PBC J1546.5+6931	3.6	14.8	0.18	-0.27	30

Table B.3: Mass-weighted age and metallicity measured for the young and old stellar populations considered separately. Col. (1): object name. Col. (2): mass-weighted age of the young and old stellar component. Col. (3): mass-weighted metallicity of the young and old stellar component. Col. (4): percentage of the total mass ascribed to the old stellar population.

Bibliography

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., Allam, S. S., Allende Prieto, C., An, D., Anderson, K. S. J., Anderson, S. F., Annis, J., Bahcall, N. A., and et al. (2009). The Seventh Data Release of the Sloan Digital Sky Survey. *ApJS*, 182:543–558.

Adelman-McCarthy, J. K., Agüeros, M. A., Allam, S. S., Allende Prieto, C., Anderson, K. S. J., Anderson, S. F., Annis, and et al. (2008). The Sixth Data Release of the Sloan Digital Sky Survey. *ApJS*, 175:297–313.

Adelman-McCarthy, J. K., Agüeros, M. A., Allam, S. S., Anderson, K. S. J., Anderson, S. F., Annis, J., Bahcall, and et al. (2006). The Fourth Data Release of the Sloan Digital Sky Survey. *ApJS*, 162:38–48.

Arendt, R. G., Kashlinsky, A., Moseley, S. H., and Mather, J. (2010). Cosmic Infrared Background Fluctuations in Deep Spitzer Infrared Array Camera Images: Data Processing and Analysis. *ApJS*, 186:10–47.

Ashby, M. L. N., Stanford, S. A., Brodwin, M., Gonzalez, A. H., Martinez-Manso, J., Bartlett, J. G., Benson, B. A., Bleem, L. E., Crawford, T. M., Dey, A., Dressler, A., Eisenhardt, P. R. M., Galametz, A., Jannuzi, B. T., Marrone, D. P., Mei, S., Muzzin, A., Pacaud, F., Pierre, M., Stern, D., and Vieira, J. D. (2013). The Spitzer South Pole Telescope Deep Field: Survey Design and Infrared Array Camera Catalogs. *ApJS*, 209:22.

Barkana, R. and Loeb, A. (2001). In the beginning: the first sources of light and the reionization of the universe. *Phys. Rep.*, 349:125–238.

Barthelmy, S. D. (2004). Burst Alert Telescope (BAT) on the Swift MIDEX mission. In Flanagan, K. A. and Siegmund, O. H. W., editors, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 5165 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 175–189.

Baum, W. A. (1962). Photoelectric Magnitudes and Red-Shifts. In McVittie, G. C., editor, *Problems of Extra-Galactic Research*, volume 15 of *IAU Symposium*, page 390.

Becker, G. D., Bolton, J. S., Haehnelt, M. G., and Sargent, W. L. W. (2011). Detection of extended He II reionization in the temperature evolution of the intergalactic medium. *MNRAS*, 410:1096–1112.

Becker, R. H., Fan, X., White, R. L., Strauss, M. A., Narayanan, V. K., Lupton, R. H., Gunn, J. E., Annis, and et al. (2001). Evidence for Reionization at $z \sim 6$: Detection of a Gunn-Peterson Trough in a $z = 6.28$ Quasar. *AJ*, 122:2850–2857.

Beckwith, S. V. W., Stiavelli, M., Koekemoer, A. M., Caldwell, J. A. R., Ferguson, H. C., Hook, R., Lucas, R. A., Bergeron, and et al. (2006). The Hubble Ultra Deep Field. *AJ*, 132:1729–1755.

Beifiori, A., Maraston, C., Thomas, D., and Johansson, J. (2011). On the spectral resolution of the MILES stellar library. *A&A*, 531:A109.

Benson, A. J., Bower, R. G., Frenk, C. S., Lacey, C. G., Baugh, C. M., and Cole, S. (2003). What Shapes the Luminosity Function of Galaxies? *ApJ*, 599:38–49.

Bertin, E. and Arnouts, S. (1996). SExtractor: Software for source extraction. *A&AS*, 117:393–404.

Bird, A. J., Bazzano, A., Bassani, L., Capitanio, F., Fiocchi, M., Hill, A. B., Malizia, A., McBride, and et al. (2010). The Fourth IBIS/ISGRI Soft Gamma-ray Survey Catalog. *ApJS*, 186:1–9.

Blakeslee, J. P., Anderson, K. R., Meurer, G. R., Benítez, N., and Magee, D. (2003). An Automatic Image Reduction Pipeline for the Advanced Camera for Surveys. In Payne, H. E., Jedrzejewski, R. I., and Hook, R. N., editors, *Astronomical Data Analysis Software and Systems XII*, volume 295 of *Astronomical Society of the Pacific Conference Series*, page 257.

Bluck, A. F. L., Conselice, C. J., Almaini, O., Laird, E. S., Nandra, K., and Grützbauch, R. (2011). On the co-evolution of supermassive black holes and their host galaxies since $z = 3$. *MNRAS*, 410:1174–1196.

Bolton, J. S., Becker, G. D., Raskutti, S., Wyithe, J. S. B., Haehnelt, M. G., and Sargent, W. L. W. (2012). Improved measurements of the intergalactic medium temperature around quasars: possible evidence for the initial stages of He II reionization at $z \sim 6$. *MNRAS*, 419:2880–2892.

Bolton, J. S. and Haehnelt, M. G. (2007). The observed ionization rate of the intergalactic medium and the ionizing emissivity at $z \geq 5$: evidence for a photon-starved and extended epoch of reionization. *MNRAS*,

382:325–341.

Bouwens, R. J., Illingworth, G. D., Blakeslee, J. P., Broadhurst, T. J., and Franx, M. (2004a). Galaxy Size Evolution at High Redshift and Surface Brightness Selection Effects: Constraints from the Hubble Ultra Deep Field. *ApJ*, 611:L1–L4.

Bouwens, R. J., Illingworth, G. D., Franx, M., and Ford, H. (2007). UV Luminosity Functions at $z \sim 4$, 5, and 6 from the Hubble Ultra Deep Field and Other Deep Hubble Space Telescope ACS Fields: Evolution and Star Formation History. *ApJ*, 670:928–958.

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., Franx, M., Labbe, I., Trenti, M., van Dokkum, P., Carollo, C. M., Gonzalez, V., and Magee, D. (2012a). UV-continuum slopes at $z \sim 4-7$ from the HUDF09+ERS+CANDELS observations: Discovery of a well-defined UV-color magnitude relationship for $z \geq 4$ star-forming galaxies. *ApJ*, 754:83.

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., Labbé, I., Trenti, M., van Dokkum, P., Franx, M., Stiavelli, M., Carollo, C. M., Magee, D., and Gonzalez, V. (2011). Ultraviolet Luminosity Functions from 132 $z \sim 7$ and $z \sim 8$ Lyman-break Galaxies in the Ultra-deep HUDF09 and Wide-area Early Release Science WFC3/IR Observations. *ApJ*, 737:90.

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., Stiavelli, M., van Dokkum, P., Trenti, M., Magee, D., Labbé, and et al. (2010). Discovery of $z \sim 8$ Galaxies in the Hubble Ultra Deep Field from Ultra-Deep WFC3/IR Observations. *ApJ*, 709:L133–L137.

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., Trenti, M., Labbe', I., Bradley, L., Carollo, M., van Dokkum, P. G., Gonzalez, V., Holwerda, B., Franx, M., Spitler, L., Smit, R., and Magee, D. (2014). UV Luminosity Functions at redshifts z˜4 to z˜10: 11000 Galaxies from HST Legacy Fields. *ArXiv e-prints*.

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., Trenti, M., Labbe, I., Franx, M., Stiavelli, M., Carollo, C. M., van Dokkum, P., and Magee, D. (2012b). Lower-Luminosity Galaxies could reionize the Universe: Very Steep Faint-End Slopes to the UV Luminosity Functions at $z \geq 5-8$ from the HUDF09 WFC3/IR Observations. *ApJ*, 752:L5.

Bouwens, R. J., Thompson, R. I., Illingworth, G. D., Franx, M., van Dokkum, P. G., Fan, X., Dickinson, M. E., Eisenstein, D. J., and Rieke, M. J. (2004b). Galaxies at $z \sim 7-8$: z_{850}-Dropouts in the Hubble Ultra Deep Field. *ApJ*, 616:L79–L82.

Bradley, L. D., Trenti, M., Oesch, P. A., Stiavelli, M., Treu, T., Bouwens, R. J., Shull, J. M., Holwerda, B. W., and Pirzkal, N. (2012). The Brightest of Reionizing Galaxies Survey: Constraints on the Bright End of the ~ 8 Luminosity Function. *ApJ*, 760:108.

Bradley, L. D., Zitrin, A., Coe, D., Bouwens, R., Postman, M., Balestra, I., Grillo, C., Monna, A., Rosati, P., Seitz, and et al. (2013). CLASH: A Census of Magnified Star-Forming Galaxies at $z \sim 6-8$. *ArXiv e-prints*.

Brammer, G. B., van Dokkum, P. G., and Coppi, P. (2008). EAZY: A Fast, Public Photometric Redshift Code. *ApJ*, 686:1503–1513.

Brandt, W. N. and Hasinger, G. (2005). Deep Extragalactic X-Ray Surveys. *ARA&A*, 43:827–859.

Bunker, A. J., Wilkins, S., Ellis, R. S., Stark, D. P., Lorenzoni, S., Chiu, K., Lacy, M., Jarvis, M. J., and Hickey, S. (2010). The contribution of high-redshift galaxies to cosmic reionization: new results from deep WFC3 imaging of the Hubble Ultra Deep Field. *MNRAS*, 409:855–866.

Bushouse, H. (2008). WFC3 IR Ground P-Flats. Technical report.

Calvi, V., Pizzella, A., Stiavelli, M., Morelli, L., Corsini, E. M., Dalla Bontà, E., Bradley, L., and Koekemoer, A. M. (2013). Constraining the luminosity function of faint undetected i-dropout galaxies. *MNRAS*, 432:3474–3484.

Capak, P., Mobasher, B., Scoville, N. Z., McCracken, H., Ilbert, O., Salvato, M., Menéndez-Delmestre, K., Aussel, and et al. (2011). Spectroscopy of Luminous $z > 7$ Galaxy Candidates and Sources of Contamination in $z > 7$ Galaxy Searches. *ApJ*, 730:68.

Cappellari, M. and Emsellem, E. (2004). Parametric Recovery of Line-of-Sight Velocity Distributions from Absorption-Line Spectra of Galaxies via Penalized Likelihood. *PASP*, 116:138–147.

Carollo, C. M. and Lilly, S. J. (2001). The Metallicity of $0.5 < z < 1$ Field Galaxies. *ApJ*, 548:L153–L157.

Casertano, S., de Mello, D., Dickinson, M., Ferguson, H. C., Fruchter, A. S., Gonzalez-Lopezlira, R. A., Heyer, I., Hook, and et al. (2000). WFPC2 Observations of the Hubble Deep Field South. *AJ*, 120:2747–2824.

Cassinelli, J. P., Cohen, D. H., Macfarlane, J. J., Drew, J. E., Lynas-Gray, A. E., Hoare, M. G., Vallerga, J. V., Welsh, B. Y., and et al. (1995). EUVE spectroscopy of epsilon Canis Majoris (B2 II) from 70 to 730 A. *ApJ*, 438:932–949.

Castellano, M., Fontana, A., Paris, D., Grazian, A., Pentericci, L., Boutsia, K., Santini, P., Testa, and et al. (2010). The bright end of the $z \sim 7$ UV luminosity function from a wide and deep HAWK-I survey. *A&A*, 524:A28.

Cen, R. (2003). The Universe Was Reionized Twice. *ApJ*, 591:12–37.

Chen, X. Y., Liang, Y. C., Hammer, F., Zhao, Y. H., and Zhong, G. H. (2009). Stellar population analysis on local

infrared-selected galaxies. *A&A*, 495:457–469.
Chonis, T. S. and Gaskell, C. M. (2008). Setting UBVRI Photometric Zero-Points Using Sloan Digital Sky Survey ugriz Magnitudes. *AJ*, 135:264–267.
Ciardi, B., Bianchi, S., and Ferrara, A. (2002). Lyman continuum escape from an inhomogeneous interstellar medium. *MNRAS*, 331:463–473.
Ciardi, B., Ferrara, A., Governato, F., and Jenkins, A. (2000). Inhomogeneous reionization of the intergalactic medium regulated by radiative and stellar feedbacks. *MNRAS*, 314:611–629.
Cid Fernandes, R., González Delgado, R. M., Storchi-Bergmann, T., Martins, L. P., and Schmitt, H. (2005). The stellar populations of low-luminosity active galactic nuclei - III. Spatially resolved spectral properties. *MNRAS*, 356:270–294.
Cid Fernandes, R., Gu, Q., Melnick, J., Terlevich, E., Terlevich, R., Kunth, D., Rodrigues Lacerda, R., and Joguet, B. (2004). The star formation history of Seyfert 2 nuclei. *MNRAS*, 355:273–296.
Ciotti, L., Ostriker, J. P., and Proga, D. (2010). Feedback from Central Black Holes in Elliptical Galaxies. III. Models with Both Radiative and Mechanical Feedback. *ApJ*, 717:708–723.
Coe, D., Zitrin, A., Carrasco, M., Shu, X., Zheng, W., Postman, M., Bradley, L., Koekemoer, A., Bouwens, R., Broadhurst, T., and et al.. (2013). CLASH: Three Strongly Lensed Images of a Candidate $z \sim 11$ Galaxy. *ApJ*, 762:32.
Cole, S. (1991). Modeling galaxy formation in evolving dark matter halos. *ApJ*, 367:45–53.
Cooray, A. (2005). A brighter past: galaxy luminosity function at high redshifts. *MNRAS*, 364:303–318.
Cooray, A., Amblard, A., Beichman, C., Benford, D., Bernstein, R., Bock, J. J., Brodwin, and et al. (2009). A New Era in Extragalactic Background Light Measurements: The Cosmic History of Accretion, Nucleosynthesis and Reionization. In *astro2010: The Astronomy and Astrophysics Decadal Survey*, volume 2010 of *Astronomy*, page 54.
Cooray, A., Sullivan, I., Chary, R.-R., Bock, J. J., Dickinson, M., Ferguson, H. C., Keating, B., Lange, A., and Wright, E. L. (2007). IR Background Anisotropies in Spitzer GOODS Images and Constraints on First Galaxies. *ApJ*, 659:L91–L94.
Cracco, V., Ciroi, S., di Mille, F., Vaona, L., Frassati, A., Smirnova, A. A., La Mura, G., Moiseev, A. V., and Rafanelli, P. (2011). The origin of gas in extended narrow-line regions of nearby Seyfert galaxies - I. NGC 7212. *MNRAS*, 418:2630–2641.
Cresci, G., Mannucci, F., Sommariva, V., Maiolino, R., Marconi, A., and Brusa, M. (2012). The metallicity properties of zCOSMOS galaxies at $0.2 < z < 0.8$. *MNRAS*, 421:262–269.
Cusumano, G., La Parola, V., Segreto, A., Ferrigno, C., Maselli, A., Sbarufatti, B., Romano, P., Chincarini, and et al. (2010). The Palermo Swift-BAT hard X-ray catalogue. III. Results after 54 months of sky survey. *A&A*, 524:A64.
Dahlen, T., Mobasher, B., Dickinson, M., Ferguson, H. C., Giavalisco, M., Grogin, N. A., Guo, Y., Koekemoer, and et al. (2010). A Detailed Study of Photometric Redshifts for GOODS-South Galaxies. *ApJ*, 724:425–447.
Dalcanton, J. J. (2007). The Metallicity of Galaxy Disks: Infall versus Outflow. *ApJ*, 658:941–959.
Davies, R. I., Müller Sánchez, F., Genzel, R., Tacconi, L. J., Hicks, E. K. S., Friedrich, S., and Sternberg, A. (2007). A Close Look at Star Formation around Active Galactic Nuclei. *ApJ*, 671:1388–1412.
Davis, T. A., Krajnović, D., McDermid, R. M., Bureau, M., Sarzi, M., Nyland, K., Alatalo, K., Bayet, and et al. (2012). Gemini GMOS and WHT SAURON integral-field spectrograph observations of the AGN-driven outflow in NGC 1266. *MNRAS*, 426:1574–1590.
de Marchi, G., Sirianni, M., Gilliland, R., Bohlin, R., Pavlovsky, C., Jee, M., Mack, J., van der Marel, R., and Boffi, F. (2004). Detector Quantum Efficiency and Photometric Zero Points of the ACS. Technical report.
de Vaucouleurs, G., de Vaucouleurs, A., Corwin, Jr., H. G., Buta, R. J., Paturel, G., and Fouque, P. (1991). *Third Reference Catalogue of Bright Galaxies*.
Di Matteo, T., Springel, V., and Hernquist, L. (2005). Energy input from quasars regulates the growth and activity of black holes and their host galaxies. *Nature*, 433:604–607.
Dickinson, M. (1998). Color-Selected High Redshift Galaxies and the HDF. In Livio, M., Fall, S. M., and Madau, P., editors, *The Hubble Deep Field*, page 219.
Dijkstra, M., Haiman, Z., Rees, M. J., and Weinberg, D. H. (2004). Photoionization Feedback in Low-Mass Galaxies at High Redshift. *ApJ*, 601:666–675.
Disney, M. J. and Lang, R. H. (2012). The galaxy ancestor problem. *MNRAS*, 426:1731–1749.
Dodelson, S. and Jubas, J. M. (1992). Reionization and decaying dark matter. *Phys. Rev. D*, 45:1076–1090.
Dodelson, S. and Jubas, J. M. (1994). Cosmological Signatures of Decaying Dark Matter. *MNRAS*, 266:886.
Draper, A. R. and Ballantyne, D. R. (2011). The Young, the Old, and the Dusty: Stellar Populations of Active Galactic Nucleus Hosts. *ApJ*, 740:57.

Dressel, L. (2012). *Wide Field Camera 3 Instrument Handbook for Cycle 21 v. 5.0*.

Edmunds, M. G. and Pagel, B. E. J. (1984). On the composition of H II regions in southern galaxies. III - NGC 2997 and 7793. *MNRAS*, 211:507–519.

Efstathiou, G. (1992). Suppressing the formation of dwarf galaxies via photoionization. *MNRAS*, 256:43P–47P.

Ellis, R. S., McLure, R. J., Dunlop, J. S., Robertson, B. E., Ono, Y., Schenker, M. A., Koekemoer, A., Bowler, R. A. A., Ouchi, M., Rogers, A. B., Curtis-Lake, E., Schneider, E., Charlot, S., Stark, D. P., Furlanetto, S. R., and Cirasuolo, M. (2013). The Abundance of Star-forming Galaxies in the Redshift Range 8.5-12: New Results from the 2012 Hubble Ultra Deep Field Campaign. *ApJ*, 763:L7.

Erb, D. K., Shapley, A. E., Pettini, M., Steidel, C. C., Reddy, N. A., and Adelberger, K. L. (2006). The Mass-Metallicity Relation at $z > 2$. *ApJ*, 644:813–828.

Faber, S. M., Willmer, C. N. A., Wolf, C., Koo, D. C., Weiner, B. J., Newman, J. A., Im, M., Coil, and et al. (2007). Galaxy Luminosity Functions to $z \sim 1$ from DEEP2 and COMBO-17: Implications for Red Galaxy Formation. *ApJ*, 665:265–294.

Fan, X. (2007). The Highest Redshift Quasars: Early Black Hole Growth and the End of Reionization Epoch. In N. Metcalfe & T. Shanks, editor, *Cosmic Frontiers*, volume 379 of *Astronomical Society of the Pacific Conference Series*, pages 35–+.

Fan, X., Carilli, C. L., and Keating, B. (2006). Observational Constraints on Cosmic Reionization. *ARA&A*, 44:415–462.

Fan, X., Narayanan, V. K., Lupton, R. H., Strauss, M. A., Knapp, G. R., Becker, R. H., White, R. L., Pentericci, and et al. (2001). A Survey of $z > 5.8$ Quasars in the Sloan Digital Sky Survey. I. Discovery of Three New Quasars and the Spatial Density of Luminous Quasars at $z \sim 6$. *AJ*, 122:2833–2849.

Fan, X., Narayanan, V. K., Strauss, M. A., White, R. L., Becker, R. H., Pentericci, L., and Rix, H. (2002). Evolution of the Ionizing Background and the Epoch of Reionization from the Spectra of $z \sim 6$ Quasars. *AJ*, 123:1247–1257.

Fanidakis, N., Baugh, C. M., Benson, A. J., Bower, R. G., Cole, S., Done, C., and Frenk, C. S. (2011). Grand unification of AGN activity in the ΛCDM cosmology. *MNRAS*, 410:53–74.

Faucher-Giguère, C.-A., Lidz, A., Zaldarriaga, M., and Hernquist, L. (2009). A New Calculation of the Ionizing Background Spectrum and the Effects of He II Reionization. *ApJ*, 703:1416–1443.

Fechner, C., Reimers, D., Kriss, G. A., Baade, R., Blair, W. P., Giroux, M. L., Green, R. F., and et al. (2006). The UV spectrum of HS 1700+6416. II. FUSE observations of the HeII Lyman alpha forest. *A&A*, 455:91–106.

Ferguson, H. C., Dickinson, M., Giavalisco, M., Kretchmer, C., Ravindranath, S., Idzi, R., Taylor, E., Conselice, and et al. (2004). The Size Evolution of High-Redshift Galaxies. *ApJ*, 600:L107–L110.

Ferguson, H. C., Dickinson, M., and Williams, R. (2000). The Hubble Deep Fields. *ARA&A*, 38:667–715.

Feruglio, C., Maiolino, R., Piconcelli, E., Menci, N., Aussel, H., Lamastra, A., and Fiore, F. (2010). Quasar feedback revealed by giant molecular outflows. *A&A*, 518:L155.

Finkelstein, S. L., Papovich, C., Giavalisco, M., Reddy, N. A., Ferguson, H. C., Koekemoer, A. M., and Dickinson, M. (2010). On the Stellar Populations and Evolution of Star-forming Galaxies at $6.3 < z < 8.6$. *ApJ*, 719:1250–1273.

Fisher, D., Fabricant, D., Franx, M., and van Dokkum, P. (1998). A Spectroscopic Survey of the Galaxy Cluster CL 1358+62 at $z = 0.328$. *ApJ*, 498:195.

Fontanot, F., Cristiani, S., and Vanzella, E. (2012). On the relative contribution of high-redshift galaxies and active galactic nuclei to reionization. *MNRAS*, 425:1413–1420.

Fontanot, F., Pasquali, A., De Lucia, G., van den Bosch, F. C., Somerville, R. S., and Kang, X. (2011). The dependence of AGN activity on stellar and halo mass in semi-analytic models. *MNRAS*, 413:957–970.

Fruchter, A. S. and Hook, R. N. (2002). Drizzle: A Method for the Linear Reconstruction of Undersampled Images. *PASP*, 114:144–152.

Gallazzi, A., Charlot, S., Brinchmann, J., White, S. D. M., and Tremonti, C. A. (2005). The ages and metallicities of galaxies in the local universe. *MNRAS*, 362:41–58.

Ganda, K., Peletier, R. F., McDermid, R. M., Falcón-Barroso, J., de Zeeuw, P. T., Bacon, R., Cappellari, M., Davies, and et al. (2007). Absorption-line strengths of 18 late-type spiral galaxies observed with SAURON. *MNRAS*, 380:506–540.

Garnett, D. R., Kennicutt, Jr., R. C., and Bresolin, F. (2004). The First Measured Electron Temperatures for Metal-rich H II Regions in M51. *ApJ*, 607:L21–L24.

Garnett, D. R., Shields, G. A., Skillman, E. D., Sagan, S. P., and Dufour, R. J. (1997). Interstellar Abundance Gradients in NGC 2403: Comparison to M33. *ApJ*, 489:63.

Gehrels, N., Chincarini, G., Giommi, P., Mason, K. O., Nousek, J. A., Wells, A. A., White, and et al. (2004). The Swift Gamma-Ray Burst Mission. *ApJ*, 611:1005–1020.

Giavalisco, M. (1998). The Properties of Lyman-Break Galaxies at Redshift $z \sim 3$. In M. Livio, S. M. Fall, & P. Madau, editor, *The Hubble Deep Field*, page 121.

Giavalisco, M. (2002). Lyman-Break Galaxies. *ARA&A*, 40:579–641.

Giavalisco, M. (2004). Cross-Talk in the ACS WFC Detectors. II: Using GAIN=2 to Minimize the Effect. Technical report.

Giavalisco, M., Ferguson, H. C., Koekemoer, A. M., Dickinson, M., Alexander, D. M., Bauer, F. E., Bergeron, J., Biagetti, C., and et al. (2004). The Great Observatories Origins Deep Survey: Initial Results from Optical and Near-Infrared Imaging. *ApJ*, 600:L93–L98.

Gnedin, N. Y. (2000). Effect of Reionization on Structure Formation in the Universe. *ApJ*, 542:535–541.

Gonzaga, S., e. a. (2011). *ACS Data Handbook, Version 6.0*.

Gonzaga, S. and et al. (2012). *The DrizzlePac Handbook*.

González, V., Labbé, I., Bouwens, R. J., Illingworth, G., Franx, M., Kriek, M., and Brammer, G. B. (2010). The Stellar Mass Density and Specific Star Formation Rate of the Universe at $z \sim 7$. *ApJ*, 713:115–130.

González Delgado, R. M., Heckman, T., and Leitherer, C. (2001). The Nuclear and Circumnuclear Stellar Population in Seyfert 2 Galaxies: Implications for the Starburst-Active Galactic Nucleus Connection. *ApJ*, 546:845–865.

Gonzalez-Gonzalez, J. D. J. (1993). *Line-strength Gradients and Kinematic Profiles in Elliptical Galaxies*. PhD thesis, University of California, Santa Cruz.

Grogin, N. A., Kocevski, D. D., Faber, S. M., Ferguson, H. C., Koekemoer, and et al. (2011). CANDELS: The Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey. *ApJS*, 197:35.

Guhathakurta, P., Tyson, J. A., and Majewski, S. R. (1990). A redshift limit for the faint blue galaxy population from deep U band imaging. *ApJ*, 357:L9–L12.

Gunn, J. E. (1966). *A mathematical framework for discussing the statistical distribution of galaxies in space and its cosmological implications*. PhD thesis, California Institute of Technology.

Guo, Y., Ferguson, H. C., Giavalisco, M., Barro, G., Willner, S. P., Ashby, M. L. N., Dahlen, T., Donley, J. L., Faber, S. M., Fontana, and et al. (2013). CANDELS Multi-wavelength Catalogs: Source Detection and Photometry in the GOODS-South Field. *ApJS*, 207:24.

Guth, A. H. (1981). Inflationary universe: A possible solution to the horizon and flatness problems. *Phys. Rev. D*, 23:347–356.

Haardt, F. and Madau, P. (2012). Radiative Transfer in a Clumpy Universe. IV. New Synthesis Models of the Cosmic UV/X-Ray Background. *ApJ*, 746:125.

Harrison, E. R. (1970). Fluctuations at the Threshold of Classical Cosmology. *Phys. Rev. D*, 1:2726–2730.

Hauser, M. G. and Dwek, E. (2001). The Cosmic Infrared Background: Measurements and Implications. *ARA&A*, 39:249–307.

Hayes, M., Laporte, N., Pelló, R., Schaerer, D., and Le Borgne, J.-F. (2012). A peculiar galaxy appears at redshift 11: properties of a moderate-redshift interloper. *MNRAS*, 425:L19–L23.

Helgason, K., Cappelluti, N., Hasinger, G., Kashlinsky, A., and Ricotti, M. (2014). The Contribution of z $\lesssim 6$ Sources to the Spatial Coherence in the Unresolved Cosmic Near-infrared and X-Ray Backgrounds. *ApJ*, 785:38.

Helgason, K., Ricotti, M., and Kashlinsky, A. (2012). Reconstructing the Near-infrared Background Fluctuations from Known Galaxy Populations Using Multiband Measurements of Luminosity Functions. *ApJ*, 752:113.

Hinshaw, G., Larson, D., Komatsu, E., Spergel, D. N., Bennett, C. L., Dunkley, J., Nolta, M. R., Halpern, M., Hill, R. S., Odegard, N., and et al. (2013). Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results. *ApJS*, 208:19.

Ho, L. C., Filippenko, A. V., and Sargent, W. L. W. (2003). A Search for "Dwarf" Seyfert Nuclei. VI. Properties of Emission-Line Nuclei in Nearby Galaxies. *ApJ*, 583:159–177.

Hogg, D. W. (1999). Distance measures in cosmology. *ArXiv Astrophysics e-prints*.

Hopkins, P. F. and Hernquist, L. (2006). Fueling Low-Level AGN Activity through Stochastic Accretion of Cold Gas. *ApJS*, 166:1–36.

Huang, K.-H., Ferguson, H. C., Ravindranath, S., and Su, J. (2013). The Bivariate Size-Luminosity Relations for Lyman Break Galaxies at $z \sim 4-5$. *ApJ*, 765:68.

Illingworth, G. D., Magee, D., Oesch, P. A., Bouwens, R. J., Labbé, I., Stiavelli, M., van Dokkum, P. G., Franx, M., Trenti, M., Carollo, C. M., and Gonzalez, V. (2013). The HST eXtreme Deep Field (XDF): Combining All ACS and WFC3/IR Data on the HUDF Region into the Deepest Field Ever. *ApJS*, 209:6.

Inoue, A. K., Iwata, I., Deharveng, J.-M., Buat, V., and Burgarella, D. (2005). VLT narrow-band photometry in the Lyman continuum of two galaxies at $z \sim 3$. Limits to the escape of ionizing flux. *A&A*, 435:471–482.

Ivanov, V. D., Rieke, G. H., Groppi, C. E., Alonso-Herrero, A., Rieke, M. J., and Engelbracht, C. W. (2000). Testing

the AGN-Starburst Connection in Seyfert Galaxies. *ApJ*, 545:190–204.

Iwata, I., Inoue, A. K., Matsuda, Y., Furusawa, H., Hayashino, T., Kousai, K., Akiyama, M., Yamada, and et al. (2009). Detections of Lyman Continuum from Star-Forming Galaxies at $z \sim 3$ through Subaru/Suprime-Cam Narrow-Band Imaging. *ApJ*, 692:1287–1293.

Jansen, R. A., Fabricant, D., Franx, M., and Caldwell, N. (2000). Spectrophotometry of Nearby Field Galaxies: The Data. *ApJS*, 126:331–397.

Jarosik, N., Bennett, C. L., Dunkley, J., Gold, B., Greason, M. R., Halpern, M., Hill, R. S., Hinshaw, and et al. (2011). Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Sky Maps, Systematic Errors, and Basic Results. *ApJS*, 192:14–+.

Joguet, B., Kunth, D., Melnick, J., Terlevich, R., and Terlevich, E. (2001). Stellar populations in Seyfert 2 galaxies. I. Atlas of near-UV spectra. *A&A*, 380:19–30.

Jones, D. H., Saunders, W., Colless, M., Read, M. A., Parker, Q. A., Watson, F. G., Campbell, L. A., Burkey, and et al. (2004). The 6dF Galaxy Survey: samples, observational techniques and the first data release. *MNRAS*, 355:747–763.

Kashlinsky, A., Arendt, R., Gardner, J. P., Mather, J. C., and Moseley, S. H. (2004). Detecting Population III Stars through Observations of Near-Infrared Cosmic Infrared Background Anisotropies. *ApJ*, 608:1–9.

Kashlinsky, A., Arendt, R. G., Ashby, M. L. N., Fazio, G. G., Mather, J., and Moseley, S. H. (2012). New Measurements of the Cosmic Infrared Background Fluctuations in Deep Spitzer/IRAC Survey Data and Their Cosmological Implications. *ApJ*, 753:63.

Kashlinsky, A., Arendt, R. G., Mather, J., and Moseley, S. H. (2005). Tracing the first stars with fluctuations of the cosmic infrared background. *Nature*, 438:45–50.

Kashlinsky, A., Arendt, R. G., Mather, J., and Moseley, S. H. (2007). Demonstrating the Negligible Contribution of Optical HST ACS Galaxies to Source-subtracted Cosmic Infrared Background Fluctuations in Deep Spitzer IRAC Images. *ApJ*, 666:L1–L4.

Kennicutt, Jr., R. C., Bresolin, F., and Garnett, D. R. (2003). The Composition Gradient in M101 Revisited. II. Electron Temperatures and Implications for the Nebular Abundance Scale. *ApJ*, 591:801–820.

Kewley, L. J. and Dopita, M. A. (2002). Using Strong Lines to Estimate Abundances in Extragalactic H II Regions and Starburst Galaxies. *ApJS*, 142:35–52.

Kobulnicky, H. A. and Kewley, L. J. (2004). Metallicities of 0.3¡z¡1.0 Galaxies in the GOODS-North Field. *ApJ*, 617:240–261.

Kobulnicky, H. A. and Phillips, A. C. (2003). Measuring Global Galaxy Metallicities Using Emission-Line Equivalent Widths. *ApJ*, 599:1031–1042.

Kobulnicky, H. A., Willmer, C. N. A., Phillips, A. C., Koo, D. C., Faber, S. M., Weiner, B. J., Sarajedini, V. L., Simard, L., and Vogt, N. P. (2003). The DEEP Groth Strip Survey. VII. The Metallicity of Field Galaxies at 0.26 ¡ z 0.82 and the Evolution of the Luminosity-Metallicity Relation. *ApJ*, 599:1006–1030.

Kobulnicky, H. A. and Zaritsky, D. (1999). Chemical Properties of Star-forming Emission-Line Galaxies at z=0.1-0.5. *ApJ*, 511:118–135.

Koekemoer, A. M., Ellis, R. S., McLure, R. J., Dunlop, J. S., Robertson, B. E., Ono, Y., Schenker, M. A., Ouchi, M., Bowler, R. A. A., Rogers, A. B., Curtis-Lake, E., Schneider, E., Charlot, S., Stark, D. P., Furlanetto, S. R., Cirasuolo, M., Wild, V., and Targett, T. (2013). The 2012 Hubble Ultra Deep Field (UDF12): Observational Overview. *ApJS*, 209:3.

Koekemoer, A. M., Faber, S. M., Ferguson, H. C., Grogin, N. A., Kocevski, and et al. (2011). CANDELS: The Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey: The Hubble Space Telescope Observations, Imaging Data Products, and Mosaics. *ApJS*, 197:36.

Koekemoer, A. M., Fruchter, A. S., Hook, R. N., and Hack, W. (2003). MultiDrizzle: An Integrated Pyraf Script for Registering, Cleaning and Combining Images. In Arribas, S., Koekemoer, A., and Whitmore, B., editors, *The 2002 HST Calibration Workshop : Hubble after the Installation of the ACS and the NICMOS Cooling System, Proceedings of a Workshop held at the Space Telescope Science Institute, Baltimore, Maryland, October 17 and 18, 2002. Edited by Santiago Arribas, Anton Koekemoer, and Brad Whitmore. Baltimore, MD: Space Telescope Science Institute, 2003., p.337*, page 337.

Koekemoer, A. M., Fruchter, A. S., Hook, R. N., Hack, W., and Hanley, C. (2006). Multidrizzle and Tweakshifts: Overview and Future Plans. In Koekemoer, A. M., Goudfrooij, P., and Dressel, L. L., editors, *The 2005 HST Calibration Workshop: Hubble After the Transition to Two-Gyro Mode*, page 423.

Komatsu, E., Smith, K. M., Dunkley, J., Bennett, C. L., Gold, B., Hinshaw, G., Jarosik, N., Larson, D., Nolta, M. R., Page, L., Spergel, D. N., Halpern, M., Hill, R. S., Kogut, A., Limon, M., Meyer, S. S., Odegard, N., Tucker, G. S., Weiland, J. L., Wollack, E., and Wright, E. L. (2011). Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation. *ApJS*, 192:18–+.

Kriss, G. A., Shull, J. M., Oegerle, W., Zheng, W., Davidsen, A. F., Songaila, A., Tumlinson, J., Cowie, L. L., Deharveng, J., Friedman, S. D., and et al. (2001). Resolving the Structure of Ionized Helium in the Intergalactic Medium with the Far Ultraviolet Spectroscopic Explorer. *Science*, 293:1112–1116.

Kuhlen, M. and Faucher-Giguère, C.-A. (2012). Concordance models of reionization: implications for faint galaxies and escape fraction evolution. *MNRAS*, 423:862–876.

La Mura, G., Bindoni, D., Ciroi, S., Cracco, V., D'Abrusco, R., Rafanelli, P., and Vaona, L. (2012). The relation between nuclear activity and stellar mass in galaxies. *MNRAS*, 426:1893–1904.

La Mura, G., Di Mille, F., Ciroi, S., Popović, L. Č., and Rafanelli, P. (2009). Balmer Emission Line Profiles and Complex Properties of Broad-Line Regions in Active Galactic Nuclei. *ApJ*, 693:1437–1448.

Labbé, I., Oesch, P. A., Bouwens, R. J., Illingworth, G. D., Magee, D., González, V., Carollo, C. M., Franx, M., Trenti, M., van Dokkum, P. G., and Stiavelli, M. (2013). The Spectral Energy Distributions of z ˜ 8 Galaxies from the IRAC Ultra Deep Fields: Emission Lines, Stellar Masses, and Specific Star Formation Rates at 650 Myr. *ApJ*, 777:L19.

Landi, R., Masetti, N., Morelli, L., Palazzi, E., Bassani, L., Malizia, A., Bazzano, A., Bird, A. J., Dean, A. J., Galaz, G., Minniti, D., and Ubertini, P. (2007). The AGN Nature of 11 out of 12 Swift/RXTE Unidentified Sources through Optical and X-Ray Spectroscopy. *ApJ*, 669:109–125.

Lanzetta, K. M., Yahata, N., Pascarelle, S., Chen, H.-W., and Fernández-Soto, A. (2002). The Star Formation Rate Intensity Distribution Function: Implications for the Cosmic Star Formation Rate History of the Universe. *ApJ*, 570:492–501.

Lara-López, M. A., Bongiovanni, A., Cepa, J., Pérez García, A. M., Sánchez-Portal, M., Castañeda, H. O., Fernández Lorenzo, M., and Pović, M. (2010). Study of star-forming galaxies in SDSS up to redshift 0.4. II. Evolution from the fundamental parameters: mass, metallicity and star formation rate. *A&A*, 519:A31.

Lawrence, A., Watson, M. G., Pounds, K. A., and Elvis, M. (1987). Low-frequency divergent X-ray variability in the Seyfert galaxy NGC4051. *Nature*, 325:694–696.

Lee, K.-S., Giavalisco, M., Gnedin, O. Y., Somerville, R. S., Ferguson, H. C., Dickinson, M., and Ouchi, M. (2006). The Large-Scale and Small-Scale Clustering of Lyman Break Galaxies at $3.5 \leq z \leq 5.5$ from the GOODS Survey. *ApJ*, 642:63–80.

Leitherer, C., Ferguson, H. C., Heckman, T. M., and Lowenthal, J. D. (1995). The Lyman Continuum in Starburst Galaxies Observed with the Hopkins Ultraviolet Telescope. *ApJ*, 454:L19.

Lilly, S. J., Carollo, C. M., and Stockton, A. N. (2003). The Metallicities of Star-forming Galaxies at Intermediate Redshifts 0.47$< z <$0.92. *ApJ*, 597:730–750.

Loeb, A. and Barkana, R. (2001). The Reionization of the Universe by the First Stars and Quasars. *ARA&A*, 39:19–66.

Lotz, J. M., Madau, P., Giavalisco, M., Primack, J., and Ferguson, H. C. (2006). The Rest-Frame Far-Ultraviolet Morphologies of Star-forming Galaxies at $z \sim 1.5$ and 4. *ApJ*, 636:592–609.

Lucas, R. A. and Hack, W. J. (2013). Using SExtractor With the Drizzlepac Tweakreg Software for Aligning and Combining Sources in HST/ACS Images. In *American Astronomical Society Meeting Abstracts*, volume 221 of *American Astronomical Society Meeting Abstracts*, page 240.17.

Madau, P., Haardt, F., and Rees, M. J. (1999). Radiative Transfer in a Clumpy Universe. III. The Nature of Cosmological Ionizing Sources. *ApJ*, 514:648–659.

Maier, C., Lilly, S. J., Carollo, C. M., Stockton, A., and Brodwin, M. (2005). Near-Infrared Spectroscopy of 0.4$< z <$1.0 CFRS Galaxies: Oxygen Abundances, SFRs, and Dust. *ApJ*, 634:849–860.

Maiolino, R., Nagao, T., Grazian, A., Cocchia, F., Marconi, A., Mannucci, F., Cimatti, A., Pipino, A., and et al. (2008). AMAZE. I. The evolution of the mass-metallicity relation at $z > 3$. *A&A*, 488:463–479.

Maiorano, E., Landi, R., Stephen, J. B., Bassani, L., Masetti, N., Parisi, P., Palazzi, E., Parma, P., and et al. (2011). Searching for active galactic nuclei among unidentified INTEGRAL sources. *MNRAS*, 416:531–540.

Malhotra, S., Rhoads, J. E., Pirzkal, N., Haiman, Z., Xu, C., Daddi, E., Yan, H., Bergeron, L. E., and et al. (2005). An Overdensity of Galaxies at $z = 5.9 \pm 0.2$ in the Hubble Ultra Deep Field Confirmed Using the ACS Grism. *ApJ*, 626:666–679.

Mannucci, F., Cresci, G., Maiolino, R., Marconi, A., and Gnerucci, A. (2010). A fundamental relation between mass, star formation rate and metallicity in local and high-redshift galaxies. *MNRAS*, 408:2115–2127.

Mannucci, F., Cresci, G., Maiolino, R., Marconi, A., Pastorini, G., Pozzetti, L., Gnerucci, A., Risaliti, G., and et al. (2009). LSD: Lyman-break galaxies Stellar populations and Dynamics - I. Mass, metallicity and gas at $z \sim 3.1$. *MNRAS*, 398:1915–1931.

Maraston, C. (2005). Evolutionary population synthesis: models, analysis of the ingredients and application to high-z galaxies. *MNRAS*, 362:799–825.

Martin, C. and Bowyer, S. (1989). Evidence for an extragalactic component of the far-ultraviolet background

and constraints on galaxy evolution for Z between 0.1 and 0.6. *ApJ*, 338:677–706.

Masetti, N., Bassani, L., Bazzano, A., Bird, A. J., Dean, A. J., Malizia, A., Norci, L., Palazzi, E., and et al. (2006a). Unveiling the nature of INTEGRAL objects through optical spectroscopy. IV. A study of six new hard X-ray sources. *A&A*, 455:11–19.

Masetti, N., Mason, E., Bassani, L., Bird, A. J., Maiorano, E., Malizia, A., Palazzi, E., Stephen, J. B., and et al. (2006b). Unveiling the nature of INTEGRAL objects through optical spectroscopy. II. The nature of four unidentified sources. *A&A*, 448:547–556.

Masetti, N., Mason, E., Morelli, L., Cellone, S. A., McBride, V. A., Palazzi, E., Bassani, L., Bazzano, A., and et al. (2008). Unveiling the nature of INTEGRAL objects through optical spectroscopy. VI. A multi-observatory identification campaign. *A&A*, 482:113–132.

Masetti, N., Morelli, L., Palazzi, E., Galaz, G., Bassani, L., Bazzano, A., Bird, A. J., Dean, A. J., and et al. (2006c). Unveiling the nature of INTEGRAL objects through optical spectroscopy. V. Identification and properties of 21 southern hard X-ray sources. *A&A*, 459:21–30.

Masetti, N., Palazzi, E., Bassani, L., Malizia, A., and Stephen, J. B. (2004). Unveiling the nature of three INTEGRAL sources through optical spectroscopy. *A&A*, 426:L41–L44.

Masetti, N., Parisi, P., Jiménez-Bailón, E., Palazzi, E., Chavushyan, V., Bassani, L., Bazzano, A., Bird, A. J., and et al. (2012). Unveiling the nature of INTEGRAL objects through optical spectroscopy. IX. Twenty two more identifications, and a glance into the far hard X-ray Universe. *A&A*, 538:A123.

Masetti, N., Parisi, P., Palazzi, E., Jiménez-Bailón, E., Chavushyan, V., Bassani, L., Bazzano, A., Bird, A. J., and et al. (2010). Unveiling the nature of INTEGRAL objects through optical spectroscopy. VIII. Identification of 44 newly detected hard X-ray sources. *A&A*, 519:A96.

Masetti, N., Parisi, P., Palazzi, E., Jiménez-Bailón, E., Morelli, L., Chavushyan, V., Mason, E., McBride, V. A., and et al. (2009). Unveiling the nature of INTEGRAL objects through optical spectroscopy. VII. Identification of 20 Galactic and extragalactic hard X-ray sources. *A&A*, 495:121–135.

Masetti, N., Pretorius, M. L., Palazzi, E., Bassani, L., Bazzano, A., Bird, A. J., Charles, P. A., Dean, A. J., and et al. (2006d). Unveiling the nature of INTEGRAL objects through optical spectroscopy. III. Observations of seven southern sources. *A&A*, 449:1139–1149.

Matsumoto, T., Matsuura, S., Murakami, H., Tanaka, M., Freund, M., Lim, M., Cohen, M., Kawada, M., and Noda, M. (2005). Infrared Telescope in Space Observations of the Near-Infrared Extragalactic Background Light. *ApJ*, 626:31–43.

McClure, R. D. and van den Bergh, S. (1968). UB V Observations of Field Galaxies. *AJ*, 73:1008.

McGaugh, S. S. (1991). H II region abundances - Model oxygen line ratios. *ApJ*, 380:140–150.

McKernan, B., Ford, K. E. S., and Reynolds, C. S. (2010). Black hole mass, host galaxy classification and AGN activity. *MNRAS*, 407:2399–2410.

McLure, R. J., Cirasuolo, M., Dunlop, J. S., Foucaud, S., and Almaini, O. (2009). The luminosity function, halo masses and stellar masses of luminous Lyman-break galaxies at redshifts $5 < z < 6$. *MNRAS*, 395:2196–2209.

McLure, R. J., Dunlop, J. S., Cirasuolo, M., Koekemoer, A. M., Sabbi, E., Stark, D. P., Targett, T. A., and Ellis, R. S. (2010). Galaxies at $z = 6 - 9$ from the WFC3/IR imaging of the Hubble Ultra Deep Field. *MNRAS*, 403:960–983.

McQuinn, M. (2012). Constraints on X-ray emissions from the reionization era. *MNRAS*, 426:1349–1360.

Meiksin, A. (2005). Constraints on the ionization sources of the high-redshift intergalactic medium. *MNRAS*, 356:596–606.

Metcalfe, N., Shanks, T., Fong, R., and Roche, N. (1995). Galaxy number counts - III. Deep CCD observations to B=27.5 mag. *MNRAS*, 273:257–276.

Mitra, S., Choudhury, T. R., and Ferrara, A. (2012). Joint quasar-cosmic microwave background constraints on reionization history. *MNRAS*, 419:1480–1488.

Morelli, L., Corsini, E. M., Pizzella, A., Dalla Bontà, E., Coccato, L., Méndez-Abreu, J., and Cesetti, M. (2012). Structure and dynamics of galaxies with a low surface-brightness disc - II. Stellar populations of bulges. *MNRAS*, 423:962–982.

Morelli, L., Pompei, E., Pizzella, A., Méndez-Abreu, J., Corsini, E. M., Coccato, L., Saglia, R. P., Sarzi, M., and Bertola, F. (2008). Stellar populations of bulges in 14 cluster disc galaxies. *MNRAS*, 389:341–363.

Moustakas, J., Zaritsky, D., Brown, M., Cool, R., Dey, A., Eisenstein, D. J., Gonzalez, A. H., Jannuzi, B., and et al. (2011). Evolution of the Stellar Mass-Metallicity Relation Since z=0.75. *ArXiv e-prints*.

Muñoz, J. A. and Loeb, A. (2008). Verifying the identity of high-redshift massive galaxies through the clustering of lower mass galaxies around them. *MNRAS*, 385:2175–2180.

Mutchler, M., Sirianni, M., van Orsow, D., and Riess, A. (2004). Bias and dark calibration of ACS data. Technical report.

Nagao, T., Maiolino, R., and Marconi, A. (2006). Gas metallicity diagnostics in star-forming galaxies. *A&A*, 459:85–101.

Novak, G. S., Ostriker, J. P., and Ciotti, L. (2011). Feedback from Central Black Holes in Elliptical Galaxies: Two-dimensional Models Compared to One-dimensional Models. *ApJ*, 737:26.

Oesch, P., Bouwens, R. J., Carollo, M., Illingworth, G. D., and HUDF09 Team (2010a). Results from the First Epoch Data of the HUDF09: The Galaxy Population at $z \sim 7-8$. In *Bulletin of the American Astronomical Society*, volume 41 of *Bulletin of the American Astronomical Society*, pages 532–+.

Oesch, P. A., Bouwens, R. J., Carollo, C. M., Illingworth, G. D., Trenti, M., Stiavelli, M., Magee, D., Labbé, I., and Franx, M. (2010b). Structure and Morphologies of $z \sim 7-8$ Galaxies from Ultra-deep WFC3/IR Imaging of the Hubble Ultra-deep Field. *ApJ*, 709:L21–L25.

Oesch, P. A., Bouwens, R. J., Illingworth, G. D., Carollo, C. M., Franx, M., Labbé, I., Magee, D., Stiavelli, M., and et al. (2010c). $z \sim 7$ Galaxies in the HUDF: First Epoch WFC3/IR Results. *ApJ*, 709:L16–L20.

Oesch, P. A., Bouwens, R. J., Illingworth, G. D., Gonzalez, V., Trenti, M., van Dokkum, P. G., Franx, M., Labbé, I., Carollo, C. M., and Magee, D. (2012). The Bright End of the Ultraviolet Luminosity Function at z ˜ 8: New Constraints from CANDELS Data in GOODS-South. *ApJ*, 759:135.

Oesch, P. A., Bouwens, R. J., Illingworth, G. D., Labbé, I., Franx, M., van Dokkum, P. G., Trenti, M., Stiavelli, M., Gonzalez, V., and Magee, D. (2013). Probing the Dawn of Galaxies at z ˜ 9-12: New Constraints from HUDF12/XDF and CANDELS data. *ApJ*, 773:75.

Oesch, P. A., Bouwens, R. J., Illingworth, G. D., Labbé, I., Smit, R., Franx, M., van Dokkum, P. G., Momcheva, I., Ashby, M. L. N., Fazio, G. G., Huang, J.-S., Willner, S. P., Gonzalez, V., Magee, D., Trenti, M., Brammer, G. B., Skelton, R. E., and Spitler, L. R. (2014). The Most Luminous z ˜ 9-10 Galaxy Candidates Yet Found: The Luminosity Function, Cosmic Star-formation Rate, and the First Mass Density Estimate at 500 Myr. *ApJ*, 786:108.

Oesch, P. A., Carollo, C. M., Stiavelli, M., Trenti, M., Bergeron, L. E., Koekemoer, A. M., Lucas, R. A., Pavlovsky, C. M., and et al. (2009). The UDF05 Follow-Up of the Hubble Ultra Deep Field. II. Constraints on Reionization from z-Dropout Galaxies. *ApJ*, 690:1350–1357.

Oesch, P. A., Stiavelli, M., Carollo, C. M., Bergeron, L. E., Koekemoer, A. M., Lucas, R. A., Pavlovsky, C. M., Trenti, M., and et al. (2007). The UDF05 Follow-up of the Hubble Ultra Deep Field. I. The Faint-End Slope of the Lyman Break Galaxy Population at $z \sim 5$. *ApJ*, 671:1212–1226.

Oke, J. B. and Gunn, J. E. (1983). Secondary standard stars for absolute spectrophotometry. *ApJ*, 266:713–717.

Onodera, M., Renzini, A., Carollo, M., Cappellari, M., Mancini, C., Strazzullo, V., Daddi, E., Arimoto, N., and et al. (2012). Deep Near-infrared Spectroscopy of Passively Evolving Galaxies at $z > 1.4$. *ApJ*, 755:26.

Osterbrock, D. E. (1989). *Astrophysics of gaseous nebulae and active galactic nuclei*.

Ouchi, M., Shimasaku, K., Furusawa, H., Saito, T., Yoshida, M., Akiyama, M., Ono, Y., Yamada, T., and et al. (2010). Statistics of 207 Lyα Emitters at a Redshift Near 7: Constraints on Reionization and Galaxy Formation Models. *ApJ*, 723:869–894.

Overzier, R. A., Bouwens, R. J., Illingworth, G. D., and Franx, M. (2006). Clustering of i_{775} Dropout Galaxies at $z \sim 6$ in GOODS and the UDF. *ApJ*, 648:L5–L8.

Pagel, B. E. J., Edmunds, M. G., Blackwell, D. E., Chun, M. S., and Smith, G. (1979). On the composition of H II regions in southern galaxies. I - NGC 300 and 1365. *MNRAS*, 189:95–113.

Parisi, P., Masetti, N., Jiménez-Bailón, E., Chavushyan, V., Malizia, A., Landi, R., Molina, M., Fiocchi, M., and et al. (2009). Accurate classification of 17 AGNs detected with Swift/BAT. *A&A*, 507:1345–1358.

Parisi, P., Masetti, N., Jiménez-Bailón, E., Chavushyan, V., Palazzi, E., Landi, R., Malizia, A., Bassani, L., and et al. (2012). Accurate classification of 29 objects detected in the 39 month Palermo Swift/BAT hard X-ray catalogue. *A&A*, 545:A101.

Partridge, R. B. and Peebles, P. J. E. (1967). Are Young Galaxies Visible? *ApJ*, 147:868.

Pascarelle, S. M., Lanzetta, K. M., and Fernández-Soto, A. (1998). The Ultraviolet Luminosity Density of the Universe from Photometric Redshifts of Galaxies in the Hubble Deep Field. *ApJ*, 508:L1–L4.

Paturel, G., Petit, C., Prugniel, P., Theureau, G., Rousseau, J., Brouty, M., Dubois, P., and Cambrésy, L. (2003). HYPERLEDA. I. Identification and designation of galaxies. *A&A*, 412:45–55.

Pettini, M., Steidel, C. C., Adelberger, K. L., Kellogg, M., Dickinson, M., and Giavalisco, M. (1998). The Discovery of Primeval Galaxies and the Epoch of Galaxy Formation. In Woodward, C. E., Shull, J. M., and Thronson, Jr., H. A., editors, *Origins*, volume 148 of *Astronomical Society of the Pacific Conference Series*, page 67.

Phillipps, S., Davies, J. I., and Disney, M. J. (1990). The visibility of high-redshift galaxies. *MNRAS*, 242:235–240.

Pirzkal, N., Rothberg, B., Ryan, R., Coe, D., Malhotra, S., Rhoads, J., and Noeske, K. (2013). Real or Interloper? The Redshift Likelihoods of $z > 8$ Galaxies in the HUDF12. *ApJ*, 775:11.

Pirzkal, N., Viana, A., and Rajan, A. (2010). The WFC3 IR 'Blobs". Technical report.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (1992). *Numerical recipes in FORTRAN. The art of scientific computing*.

Prugniel, P. and Soubiran, C. (2001). A database of high and medium-resolution stellar spectra. *A&A*, 369:1048–1057.

Quinn, T., Katz, N., and Efstathiou, G. (1996). Photoionization and the formation of dwarf galaxies. *MNRAS*, 278:L49–L54.

Ravindranath, S., Giavalisco, M., Ferguson, H. C., Conselice, C., Katz, N., Weinberg, M., Lotz, J., Dickinson, M., and et al. (2006). The Morphological Diversities among Star-forming Galaxies at High Redshifts in the Great Observatories Origins Deep Survey. *ApJ*, 652:963–980.

Reach, W. T. (1997). The structured zodiacal light: IRAS, COBE, and ISO observations. In Okuda, H., Matsumoto, T., and Rollig, T., editors, *Diffuse Infrared Radiation and the IRTS*, volume 124 of *Astronomical Society of the Pacific Conference Series*, page 33.

Richard, J., Stark, D. P., Ellis, R. S., George, M. R., Egami, E., Kneib, J.-P., and Smith, G. P. (2008). A Hubble and Spitzer Space Telescope Survey for Gravitationally Lensed Galaxies: Further Evidence for a Significant Population of Low-Luminosity Galaxies beyond z=7. *ApJ*, 685:705–724.

Richer, M. G. and McCall, M. L. (1995). Oxygen abundances in diffuse ellipticals and the metallicity-luminosity relations for dwarf galaxies. *ApJ*, 445:642–659.

Riess, A., Mutchler, M., and van Orsow, D. (2002). A First Look at Hot Pixels on ACS. Technical report.

Riess, A. G., Strolger, L.-G., Casertano, S., Ferguson, H. C., Mobasher, B., Gold, B., Challis, P. J., Filippenko, A. V., Jha, S., Li, W., Tonry, J., Foley, R., Kirshner, R. P., Dickinson, M., MacDonald, E., Eisenstein, D., Livio, M., Younger, J., Xu, C., Dahlén, T., and Stern, D. (2007). New Hubble Space Telescope Discoveries of Type Ia Supernovae at $z \geq 1$: Narrowing Constraints on the Early Behavior of Dark Energy. *ApJ*, 659:98–121.

Rigopoulou, D., Mainieri, V., Almaini, O., Alonso-Herrero, A., Huang, J.-S., Hasinger, G., Rieke, G., Dunlop, J., and Lehmann, I. (2009). Spectral energy distributions of type 2 quasi-stellar objects: obscured star formation at high redshifts. *MNRAS*, 400:1199–1207.

Rodrigues, M., Puech, M., Hammer, F., Rothberg, B., and Flores, H. (2012). A decrease of the gas exchanges between galaxies and the intergalactic medium, from 12 to 6 billion years ago. *MNRAS*, 421:2888–2903.

Ryder, S. D. (1995). The local metallicity-surface brightness relationship in galactic disks. *ApJ*, 444:610–631.

Salpeter, E. E. (1955). The Luminosity Function and Stellar Evolution. *ApJ*, 121:161.

Salvaterra, R., Ferrara, A., and Dayal, P. (2011). Simulating high-redshift galaxies. *MNRAS*, 414:847–859.

Sánchez-Blázquez, P., Peletier, R. F., Jiménez-Vicente, J., Cardiel, N., Cenarro, A. J., Falcón-Barroso, J., Gorgas, J., Selam, S., and Vazdekis, A. (2006). Medium-resolution Isaac Newton Telescope library of empirical spectra. *MNRAS*, 371:703–718.

Sarzi, M., Falcón-Barroso, J., Davies, R. L., Bacon, R., Bureau, M., Cappellari, M., de Zeeuw, P. T., Emsellem, E., and et al. (2006). The SAURON project - V. Integral-field emission-line kinematics of 48 elliptical and lenticular galaxies. *MNRAS*, 366:1151–1200.

Sarzi, M., Rix, H.-W., Shields, J. C., Ho, L. C., Barth, A. J., Rudnick, G., Filippenko, A. V., and Sargent, W. L. W. (2005). The Stellar Populations in the Central Parsecs of Galactic Bulges. *ApJ*, 628:169–186.

Sasaki, S. and Takahara, F. (1994). Cosmological Reionization Caused by Structure Formation. *Progress of Theoretical Physics*, 91:699–721.

Sasaki, S. and Umemura, M. (1996). Reionization of the Universe by Early-formed Massive Black Holes. *ApJ*, 462:104.

Savaglio, S., Glazebrook, K., Le Borgne, D., Juneau, S., Abraham, R. G., Chen, H.-W., Crampton, D., McCarthy, P. J., and et al. (2005). The Gemini Deep Deep Survey. VII. The Redshift Evolution of the Mass-Metallicity Relation. *ApJ*, 635:260–279.

Schawinski, K., Lintott, C. J., Thomas, D., Kaviraj, S., Viti, S., Silk, J., Maraston, C., Sarzi, M., and et al. (2009). Destruction of Molecular Gas Reservoirs in Early-Type Galaxies by Active Galactic Nucleus Feedback. *ApJ*, 690:1672–1680.

Schawinski, K., Thomas, D., Sarzi, M., Maraston, C., Kaviraj, S., Joo, S.-J., Yi, S. K., and Silk, J. (2007). Observational evidence for AGN feedback in early-type galaxies. *MNRAS*, 382:1415–1431.

Schmidt, K. B., Treu, T., Trenti, M., Bradley, L. D., Kelly, B. C., Oesch, P. A., Holwerda, B. W., Shull, J. M., and Stiavelli, M. (2014). The Luminosity Function at z ˜ 8 from 97 Y-band Dropouts: Inferences about Reionization. *ApJ*, 786:57.

Sciama, D. W. (1988). Photino decay and the ionization of Lyman alpha clouds at large redshifts. *MNRAS*, 230:13P.

Sciama, D. W. (1994). Decaying neutrinos and the abundance of He I in Lyman limit systems and Lyman-Alpha clouds. *ApJ*, 422:L49–L50.

Sciama, D. W. (1995). Decaying Neutrinos, the Metagalactic Radiation Field, and the Ionization of Nitrogen in the Interstellar Medium. *ApJ*, 448:667.

Sethi, S. K. and Nath, B. B. (1997). On the source of ionization of the intergalactic medium at $z \sim 2.4$. *MNRAS*, 289:634–644.

Shankar, F. and Mathur, S. (2007). On the Faint End of the High-Redshift Active Galactic Nucleus Luminosity Function. *ApJ*, 660:1051–1059.

Shapiro, P. R. and Giroux, M. L. (1987). Cosmological H II regions and the photoionization of the intergalactic medium. *ApJ*, 321:L107–L112.

Shapley, A. E., Erb, D. K., Pettini, M., Steidel, C. C., and Adelberger, K. L. (2004). Evidence for Solar Metallicities in Massive Star-forming Galaxies at $z > 2$. *ApJ*, 612:108–121.

Shapley, A. E., Steidel, C. C., Pettini, M., Adelberger, K. L., and Erb, D. K. (2006). The Direct Detection of Lyman Continuum Emission from Star-forming Galaxies at $z \sim 3$. *ApJ*, 651:688–703.

Shectman, S. A. (1973). Clusters of Galaxies and the Cosmic Light. *ApJ*, 179:681–698.

Shectman, S. A. (1974). The small-scale anisotropy of the cosmic light. *ApJ*, 188:233–242.

Siana, B., Polletta, M. d. C., Smith, H. E., Lonsdale, C. J., Gonzalez-Solares, E., Farrah, D., Babbedge, T. S. R., Rowan-Robinson, M., and et al. (2008). High-Redshift QSOs in the SWIRE Survey and the $z \sim 3$ QSO Luminosity Function. *ApJ*, 675:49–70.

Silva, L., Maiolino, R., and Granato, G. L. (2004). Connecting the cosmic infrared background to the X-ray background. *MNRAS*, 355:973–985.

Sirianni, M., Martel, A. R., Jee, M. J., van Orsow, D., and Sparks, W. B. (2003). Bias Subtraction and Correction of ACS/WFC Frames. In Arribas, S., Koekemoer, A., and Whitmore, B., editors, *HST Calibration Workshop : Hubble after the Installation of the ACS and the NICMOS Cooling System*, page 82.

Skelton, R. E., Whitaker, K. E., Momcheva, I. G., Brammer, G. B., van Dokkum, P. G., Labbe, I., Franx, M., van der Wel, and et al. (2014). 3D-HST WFC3-selected Photometric Catalogs in the Five CANDELS/3D-HST Fields: Photometry, Photometric Redshifts and Stellar Masses. *ArXiv e-prints*.

Skillman, E. D., Kennicutt, R. C., and Hodge, P. W. (1989). Oxygen abundances in nearby dwarf irregular galaxies. *ApJ*, 347:875–882.

Smette, A., Heap, S. R., Williger, G. M., Tripp, T. M., Jenkins, E. B., and Songaila, A. (2002). Hubble Space Telescope Space Telescope Imaging System Observations of the He II Gunn-Peterson Effect toward HE 2347-4342. *ApJ*, 564:542–558.

Smith, R. M., Zavodny, M., Rahmer, G., and Bonati, M. (2008). A theory for image persistence in HgCdTe photodiodes. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 7021 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*.

Somerville, R. S., Primack, J. R., and Faber, S. M. (2001). The nature of high-redshift galaxies. *MNRAS*, 320:504–528.

Sommariva, V., Mannucci, F., Cresci, G., Maiolino, R., Marconi, A., Nagao, T., Baroni, A., and Grazian, A. (2012). Stellar metallicity of star-forming galaxies at $z \sim 3$. *A&A*, 539:A136.

Songaila, A., Cowie, L. L., and Lilly, S. J. (1990). Galaxy formation and the origin of the ionizing flux at large redshift. *ApJ*, 348:371–377.

Springel, V., Di Matteo, T., and Hernquist, L. (2005). Modelling feedback from stars and black holes in galaxy mergers. *MNRAS*, 361:776–794.

Stanway, E. R., Bremer, M. N., and Lehnert, M. D. (2008). On contamination and completeness in $zgt5$ Lyman-break galaxy surveys. *MNRAS*, 385:493–510.

Stark, D. P., Ellis, R. S., Bunker, A., Bundy, K., Targett, T., Benson, A., and Lacy, M. (2009). The Evolutionary History of Lyman Break Galaxies Between Redshift 4 and 6: Observing Successive Generations of Massive Galaxies in Formation. *ApJ*, 697:1493–1511.

Stark, D. P., Ellis, R. S., Chiu, K., Ouchi, M., and Bunker, A. (2010). Keck spectroscopy of faint $3 < z < 7$ Lyman break galaxies - I. New constraints on cosmic reionization from the luminosity and redshift-dependent fraction of Lyman α emission. *MNRAS*, 408:1628–1648.

Stecker, F. W., Puget, J. L., and Fazio, G. G. (1977). The cosmic far-infrared background at high galactic latitudes. *ApJ*, 214:L51–L55.

Steidel, C. C., Adelberger, K. L., Giavalisco, M., Dickinson, M., and Pettini, M. (1999). Lyman-Break Galaxies at $zgtsim4$ and the Evolution of the Ultraviolet Luminosity Density at High Redshift. *ApJ*, 519:1–17.

Steidel, C. C. and Hamilton, D. (1992). Deep imaging of high redshift QSO fields below the Lyman limit. I - The field of Q0000-263 and galaxies at $z = 3.4$. *AJ*, 104:941–949.

Steidel, C. C. and Hamilton, D. (1993). Deep imaging of high redshift QSO fields below the Lyman limit. II - Number counts and colors of field galaxies. *AJ*, 105:2017–2030.

Stephen, J. B., Bassani, L., Malizia, A., Bazzano, A., Ubertini, P., Bird, A. J., Dean, A. J., Lebrun, F., and Walter, R. (2006). Using the ROSAT catalogues to find counterparts for the second IBIS/ISGRI survey sources. *A&A*, 445:869–873.

Stiavelli, M. (2009). *From First Light to Reionization: The End of the Dark Ages*.

Stiavelli, M., Fall, S. M., and Panagia, N. (2004a). Observable Properties of Cosmological Reionization Sources. *ApJ*, 600:508–519.

Stiavelli, M., Fall, S. M., and Panagia, N. (2004b). The Possible Detection of Cosmological Reionization Sources. *ApJ*, 610:L1–L4.

Storchi-Bergmann, T., González Delgado, R. M., Schmitt, H. R., Cid Fernandes, R., and Heckman, T. (2001). Circumnuclear Stellar Population, Morphology, and Environment of Seyfert 2 Galaxies: An Evolutionary Scenario. *ApJ*, 559:147–156.

Sturm, E., González-Alfonso, E., Veilleux, S., Fischer, J., Graciá-Carpio, J., Hailey-Dunsheath, S., Contursi, A., Poglitsch, A., and et al. (2011). Massive Molecular Outflows and Negative Feedback in ULIRGs Observed by Herschel-PACS. *ApJ*, 733:L16.

Su, J., Stiavelli, M., Oesch, P., Trenti, M., Bergeron, E., Bradley, L., Carollo, M., Dahlen, T., and et al. (2011). The UDF05 Follow-up of the Hubble Ultra Deep Field. III. The Luminosity Function at $z \sim 6$. *ApJ*, 738:123.

Tegmark, M., Silk, J., and Evrard, A. (1993). Late Reionization by Supernova-driven Winds. *ApJ*, 417:54.

Thomas, D. and Davies, R. L. (2006). Rejuvenation of spiral bulges. *MNRAS*, 366:510–520.

Tolman, R. C. (1930). On the Estimation of Distances in a Curved Universe with a Non-Static Line Element. *Proceedings of the National Academy of Science*, 16:511–520.

Tolman, R. C. (1934). *Relativity, Thermodynamics, and Cosmology*.

Tonry, J. and Schneider, D. P. (1988). A new technique for measuring extragalactic distances. *AJ*, 96:807–815.

Treister, E., Schawinski, K., Volonteri, M., Natarajan, P., and Gawiser, E. (2011). Black hole growth in the early Universe is self-regulated and largely hidden from view. *Nature*, 474:356–358.

Tremonti, C. A., Heckman, T. M., Kauffmann, G., Brinchmann, J., Charlot, S., White, S. D. M., Seibert, M., Peng, E. W., and et al. (2004). The Origin of the Mass-Metallicity Relation: Insights from 53,000 Star-forming Galaxies in the Sloan Digital Sky Survey. *ApJ*, 613:898–913.

Tremonti, C. A., Moustakas, J., and Diamond-Stanic, A. M. (2007). The Discovery of 1000 km s^{-1} Outflows in Massive Poststarburst Galaxies at $z = 0.6$. *ApJ*, 663:L77–L80.

Trenti, M., Bradley, L. D., Stiavelli, M., Oesch, P., Treu, T., Bouwens, R. J., Shull, J. M., MacKenty, J. W., and et al. (2011). The Brightest of Reionizing Galaxies Survey: Design and Preliminary Results. *ApJ*, 727:L39.

Trenti, M., Bradley, L. D., Stiavelli, M., Shull, J. M., Oesch, P., Bouwens, R. J., Muñoz, J. A., Romano-Diaz, E., and et al. (2012). Overdensities of Y-dropout Galaxies from the Brightest-of-Reionizing Galaxies Survey: A Candidate Protocluster at Redshift $z \sim 8$. *ApJ*, 746:55.

Trenti, M. and Stiavelli, M. (2009). Formation Rates of Population III Stars and Chemical Enrichment of Halos during the Reionization Era. *ApJ*, 694:879–892.

Trenti, M., Stiavelli, M., Bouwens, R. J., Oesch, P., Shull, J. M., Illingworth, G. D., Bradley, L. D., and Carollo, C. M. (2010). The Galaxy Luminosity Function During the Reionization Epoch. *ApJ*, 714:L202–L207.

Ubertini, P., Lebrun, F., Di Cocco, G., Bazzano, A., Bird, A. J., Broenstad, K., Goldwurm, and et al. (2003). IBIS: The Imager on-board INTEGRAL. *A&A*, 411:L131–L139.

Vale, A. and Ostriker, J. P. (2004). Linking halo mass to galaxy luminosity. *MNRAS*, 353:189–200.

Vanzella, E., Fontana, A., Zitrin, A., Coe, D., Bradley, L., Postman, M., Grazian, A., Castellano, M., Pentericci, L., Giavalisco, M., and et al. (2013). Characterizing faint galaxies in the reionization epoch: LBT confirms two L< 0.2L* sources at $z = 6.4$ behind the CLASH/Frontier Fields cluster MACS0717.5+3745. *ArXiv e-prints*.

Vanzella, E., Giavalisco, M., Dickinson, M., Cristiani, S., Nonino, M., Kuntschner, H., Popesso, P., Rosati, P., Renzini, A., Stern, D., Cesarsky, C., Ferguson, H. C., and Fosbury, R. A. E. (2009). Spectroscopic Observations of Lyman Break Galaxies at Redshifts ˜4, 5, and 6 in the Goods-South Field. *ApJ*, 695:1163–1182.

Vaona, L., Ciroi, S., Di Mille, F., Cracco, V., La Mura, G., and Rafanelli, P. (2012). Spectral properties of the narrow-line region in Seyfert galaxies selected from the SDSS-DR7. *MNRAS*, 427:1266–1283.

Vazdekis, A., Sánchez-Blázquez, P., Falcón-Barroso, J., Cenarro, A. J., Beasley, M. A., Cardiel, N., Gorgas, J., and Peletier, R. F. (2010). Evolutionary stellar population synthesis with MILES - I. The base models and a new line index system. *MNRAS*, 404:1639–1671.

Viero, M. P., Wang, L., Zemcov, M., Addison, G., Amblard, A., Arumugam, V., Aussel, H., Béthermin, M., and et al. (2013). HerMES: Cosmic Infrared Background Anisotropies and the Clustering of Dusty Star-forming Galaxies. *ApJ*, 772:77.

White, M. (1994). Reconstructing the CMB power spectrum. *A&A*, 290:L1–L4.

Williams, C. C., Giavalisco, M., Cassata, P., Tundo, E., Wiklind, T., Guo, Y., Lee, B., Barro, G., Wuyts, S., Bell, E. F.,

and et al. (2014). The Progenitors of the Compact Early-type Galaxies at High Redshift. *ApJ*, 780:1.

Williams, R. E., Baum, S., Bergeron, L. E., Bernstein, N., Blacker, B. S., Boyle, B. J., M., B. T., Carollo, C. M., and et al. (2000). The Hubble Deep Field South: Formulation of the Observing Campaign. *AJ*, 120:2735–2746.

Williams, R. E., Blacker, B., Dickinson, M., Dixon, W. V. D., Ferguson, H. C., Fruchter, A. S., Giavalisco, M., Gilliland, R. L., and et al. (1996). The Hubble Deep Field: Observations, Data Reduction, and Galaxy Photometry. *AJ*, 112:1335–+.

Willott, C. J., McLure, R. J., Hibon, P., Bielby, R., McCracken, H. J., Kneib, J.-P., Ilbert, O., Bonfield, D. G., and et al. (2013). An Exponential Decline at the Bright End of the $z = 6$ Galaxy Luminosity Function. *AJ*, 145:4.

Winkler, C., Courvoisier, T. J.-L., Di Cocco, G., Gehrels, N., Giménez, A., Grebenev, S., Hermsen, W., Mas-Hesse, J. M., and et al. (2003). The INTEGRAL mission. *A&A*, 411:L1–L6.

Yan, H., Finkelstein, S. L., Huang, K.-H., Ryan, R. E., Ferguson, H. C., Koekemoer, A. M., Grogin, N. A., Dickinson, and et al. (2012). Luminous and High Stellar Mass Candidate Galaxies at $z \sim 8$ Discovered in the Cosmic Assembly Near-Infrared Deep Extragalactic Legacy Survey. *ApJ*, 761:177.

Yan, H. and Windhorst, R. A. (2004). The Major Sources of the Cosmic Reionizing Background at $z \sim 6$. *ApJ*, 600:L1–L5.

Yan, H.-J., Windhorst, R. A., Hathi, N. P., Cohen, S. H., Ryan, R. E., O'Connell, R. W., and McCarthy, P. J. (2010). Galaxy formation in the reionization epoch as hinted by Wide Field Camera 3 observations of the Hubble Ultra Deep Field. *Research in Astronomy and Astrophysics, Volume 10, Issue 9, pp. 867-904 (2010).*, 10:867–904.

Yegorova, I. A., Pizzella, A., and Salucci, P. (2011). Probing dark matter haloes of spiral galaxies at poorly explored distances using satellite kinematics. *A&A*, 532:A105.

Zahid, H. J., Kewley, L. J., and Bresolin, F. (2011). The Mass-Metallicity and Luminosity-Metallicity Relations from DEEP2 at $z \sim 0.8$. *ApJ*, 730:137.

Zahn, O., Reichardt, C. L., Shaw, L., Lidz, A., Aird, K. A., Benson, B. A., E., B. L., Carlstrom, J. E., and et al. (2012). Cosmic microwave background constraints on the duration and timing of reionization from the South Pole Telescope. *ApJ*, 756:65.

Zaroubi, S. (2013). The Epoch of Reionization. In Wiklind, T., Mobasher, B., and Bromm, V., editors, *Astrophysics and Space Science Library*, volume 396 of *Astrophysics and Space Science Library*, page 45.

Zeldovich, Y. B. (1972). A hypothesis, unifying the structure and the entropy of the Universe. *MNRAS*, 160:1P.

Zheng, W., Meiksin, A., Pifko, K., Anderson, S. F., Hogan, C. J., Tittley, E., Kriss, G. A., Chiu, K., and et al. (2008). Intergalactic He II Absorption in the Spectra of Quasars at Redshifts 3.5 and 3.8, Observed with the HST ACS Prism. *ApJ*, 686:195–200.

Acknowledgments

I would like to thank the people at STScI and JHU in Baltimore who were involved and interested in this work:
my advisor Massimo Stiavelli who guided me through the high-redshift Universe, Larry Bradley, for the endless help when the amount of data was overwhelming, Henry Ferguson, a precious referee for this thesis, Ray Lucas, Eddie Bergeron & Anton Koekemoer for their help,
Moreover, thanks to Antonella Nota, Stefano Casertano, Elinor Medezinski, Preethi Nair, Michele Cignoni, Matteo Correnti, Andrea Bellini, Elena Sabbi, Anastasia Alexov, Annalisa Calamida, Jay Anderson, Loretta Willers, Colin Norman, & Mario Livio for their encouragement and useful conversations.

Thanks to Michele Trenti & Pascal Oesch for the precious advice and useful discussions.

Thanks to the extragalactic group at Padua University, who involved me for the first time into the magic world of galaxies. In particular, thanks to Prof. Bertola, Prof. Pizzella, Prof. Corsini who have been mentoring me for years. Also, thanks to Lorenzo Morelli, Lucio Buson, Elena dalla Bontà, Mary Cesetti & Elisa Portaluri.

A huge “thank you” to my family, who helped me through all my life, especially when I was far away and feeling home-sick, and to all my friends spread all over the world for the endless effort of supporting and cheering me up.

Finally, I would to acknowledge the studentship granted by “Fondazione Ing. Aldo Gini” in Padova, as well as the hospitality of the “Institute of Astronomy” (IoA) in Cambridge (UK) and the “Universidad Andres Bello” (UNAB) in Santiago (Chile) where part of this thesis was written.