

Springer Theses
Recognizing Outstanding Ph.D. Research

Corentin Schreiber

A Statistical and Multi-wavelength Study of Star Formation in Galaxies



Springer

Springer Theses

Recognizing Outstanding Ph.D. Research

Aims and Scope

The series “Springer Theses” brings together a selection of the very best Ph.D. theses from around the world and across the physical sciences. Nominated and endorsed by two recognized specialists, each published volume has been selected for its scientific excellence and the high impact of its contents for the pertinent field of research. For greater accessibility to non-specialists, the published versions include an extended introduction, as well as a foreword by the student’s supervisor explaining the special relevance of the work for the field. As a whole, the series will provide a valuable resource both for newcomers to the research fields described, and for other scientists seeking detailed background information on special questions. Finally, it provides an accredited documentation of the valuable contributions made by today’s younger generation of scientists.

Theses are accepted into the series by invited nomination only and must fulfill all of the following criteria

- They must be written in good English.
- The topic should fall within the confines of Chemistry, Physics, Earth Sciences, Engineering and related interdisciplinary fields such as Materials, Nanoscience, Chemical Engineering, Complex Systems and Biophysics.
- The work reported in the thesis must represent a significant scientific advance.
- If the thesis includes previously published material, permission to reproduce this must be gained from the respective copyright holder.
- They must have been examined and passed during the 12 months prior to nomination.
- Each thesis should include a foreword by the supervisor outlining the significance of its content.
- The theses should have a clearly defined structure including an introduction accessible to scientists not expert in that particular field.

More information about this series at <http://www.springer.com/series/8790>

Corentin Schreiber

A Statistical and Multi-wavelength Study of Star Formation in Galaxies

Doctoral Thesis accepted by
Paris-Sud University, France



Springer

<i>Author</i>	<i>Supervisor</i>
Dr. Corentin Schreiber	Prof. David Elbaz
Leiden Observatory	Service d'Astrophysique
Leiden University	CEA Saclay
Leiden	Gif-sur-Yvette
The Netherlands	France

ISSN 2190-5053

Springer Theses

ISBN 978-3-319-44292-1

DOI 10.1007/978-3-319-44293-8

ISSN 2190-5061 (electronic)

ISBN 978-3-319-44293-8 (eBook)

Library of Congress Control Number: 2016947473

© Springer International Publishing AG 2016

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

This Springer imprint is published by Springer Nature

The registered company is Springer International Publishing AG

The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Supervisor's Foreword

Our understanding of the processes ruling the formation and evolution of galaxies substantially progressed over the last years mostly due to new generations of observatories on the ground and in space. Yet we are still lacking a clear picture of the reasons why galaxies experienced a peak activity of star formation about ten billion years ago and why an increasing fraction of galaxies stopped forming new generations of stars since that epoch.

The astronomical community globally expected this behavior to be driven by environment effects through galaxy interactions or mergers and over the course of large-scale structure formation such as groups and clusters of galaxies. The morphology of galaxies would shift from flat rotating disks to dispersion-dominated galaxies with elliptical shapes.

Such expectation was driven not only by past observations but also by the commonly accepted hierarchical galaxy formation paradigm in which galaxies acquired their mass through successive mergers over cosmic time.

Instead, the main picture that is coming out of these new observations shows a surprising universality of galaxy properties and in particular of their growth rate with cosmic time. Despite their different morphologies, rotation, and local environment, galaxies of the same stellar mass appear to be forming stars with the same rate at a given distance from us. If confirmed over most of the cosmic history and with reliable tracers of galaxy properties, such behavior would favor a paradigm in which galaxies grow secularly rather than by episodic bursts.

In his thesis work, Corentin Schreiber addressed this issue using the deepest available data over a wide range in wavelength for galaxies out to 12.3 billion years ago, thus covering 90 % of the cosmic history since the Big Bang. The European Space Agency's cornerstone space observatory, Herschel, is central to this work. It carried the largest mirror ever launched in space offering the possibility to image the deep far-infrared Universe with a unique precision. Because massive stars are formed in giant molecular clouds, their ultraviolet radiation is absorbed by interstellar dust and reemitted in the far infrared. To reliably trace back the star

formation history of galaxies, astronomers therefore look after the far-infrared light radiated by dust grains heated at temperatures of typically 30 K.

Despite its large telescope and great sensitivity, the diffraction pattern of the distant galaxies observed at wavelengths ranging from 100 to 500 μ limited Herschel's view on the distant Universe to the most extreme objects—not representative of the main growth path of galaxies, or to the last 8–10 billion years just after the peak of the cosmic star formation history.

Corentin Schreiber proposed a new approach that made a major difference in our ability to probe the distant Universe. Starting by selecting galaxies in intervals of stellar masses, he developed a new method to determine how strongly did galaxies differ in their star formation rate even below the Herschel detection limit, which he called “scatter stacking.” Based on his new approach, he was able to show that more than 70 % of galaxy star formation did not happen in merger-driven starbursts over 90 % of the Universe. Instead, galaxies largely grew in mass through this secular mode of star formation. To validate his method, he designed a tool now publicly available to simulate a realistic Universe with galaxies representing accurately their observed colors, shapes, and environment over the whole cosmic history of the Universe. This tool called the Empirical Galaxy Generator (EGG), turned out to be powerful to calibrate not only his method but also source detection algorithms and is now even used in the preparation of the next-generation space mission of ESA, Euclid.

Then turning to the most massive galaxies in the Universe, Corentin Schreiber was able to show that about half of these galaxies, which stop forming stars before the others, happen to do so in a slow manner—through what he calls a “slow downfall”—as opposed to the classical “quenching” of star formation that is commonly advocated but still resisting a robust observational validation. If this process is confirmed, it suggests that all galaxies did die not from violent events but also from a natural exhaustion of their gas reservoirs. How this slow downfall and the more classical quenching of star formation can be reconciled with the paradigm in which galaxies are continuously fed by infalling intergalactic matter remains an open issue.

France
June 2016

Prof. David Elbaz

Publications Related to This Thesis

First-author publications

C. Schreiber, et al. 2016, The mid- to far-infrared spectral energy distribution of star-forming galaxies from $z = 3$ to $z = 0.5$, to be submitted to A&A.

C. Schreiber, M. Pannella, R. Leiton, et al. 2016, The ALMA Redshift 4 Survey (AR4S): I. The massive end of the $z = 4$ main sequence of galaxies, submitted to A&A. <http://adsabs.harvard.edu/abs/2016arXiv160606252S>

C. Schreiber, D. Elbaz, M. Pannella, et al. 2016, EGG: hatching a mock Universe from empirical prescriptions, submitted to A&A. <http://adsabs.harvard.edu/abs/2016arXiv160605354S>

C. Schreiber, D. Elbaz, M. Pannella, et al. 2016, Observational evidence of a slow downfall of star formation efficiency in massive galaxies during the past 10 Gyr, A&A, 589, 35. <http://adsabs.harvard.edu/abs/2016A...589...35A>

C. Schreiber, M. Pannella, D. Elbaz, et al. 2015, The Herschel view of the dominant mode of galaxy growth from $z = 4$ to the present day, A&A, 575, 74. <http://adsabs.harvard.edu/abs/2015A&A...575A..74S>

Other publications related to this thesis

T. Wang, D. Elbaz, E. Daddi, A. Finoguenov, D. Liu, **C. Schreiber**, et al. 2016, Discovery of a galaxy cluster with a violently starbursting core at $z = 2.506$, accepted for publication in ApJ. <http://adsabs.harvard.edu/abs/2016arXiv160407404W>

- L. Ciesla, A. Boselli, D. Elbaz, **et al.** 2016, The imprint of rapid star formation quenching on the spectral energy distributions of galaxies, ApJ, 814, 118. <http://adsabs.harvard.edu/abs/2016ApJ...814..118T>
- A. Tomczak, R. Quadri, K.-V. Tran, **et al.** 2016, The SFR– M_* Relation and Empirical Star-Formation Histories from ZFOURGE* at $0.5 < z < 4$, ApJ, 814, 118. <http://adsabs.harvard.edu/abs/2016ApJ...814..118T>
- T. Wang, D. Elbaz, **C. Schreiber**, et al. 2015, An H and IRAC color selection of massive galaxies at $z > 3$, ApJ, 816, 84. <http://adsabs.harvard.edu/abs/2016ApJ...816..084W>
- X. Shu, D. Elbaz, N. Bourne, **C. Schreiber**, et al. 2015, Identification of $z \gtrsim 2$ *Herschel* 500 μm sources using color-deconfusion, ApJS, 222, 4. <http://adsabs.harvard.edu/abs/2016ApJS..222....4S>
- J.R. Mullaney, D.M. Alexander, J. Aird, **et al.** 2015, ALMA and Herschel reveal that AGN and main-sequence galaxies have different star formation rate distributions, MNRAS, 453, L83. <http://adsabs.harvard.edu/abs/2015MNRAS.453L..83M>
- R. Gobat, E. Daddi, M. Béthermin, **et al.** 2015, Satellite content and quenching of star formation in galaxy groups at $z \sim 1.8$, A&A, 581, 56. <http://adsabs.harvard.edu/abs/2015A>
- E. Merlin, A. Fontana, H.C. Ferguson, **et al.** 2015, T-PHOT: a new code for PSF-matched, prior-based, multiwavelength extragalactic deconfusion photometry, A&A, 590, 30. <http://adsabs.harvard.edu/abs/2016A>
- R. Leiton, D. Elbaz, K. Okomura, **et al.** 2015, GOODS-*Herschel* : identification of the individual galaxies responsible for the 80–290 μm cosmic infrared background, A&A, 579, 93. <http://adsabs.harvard.edu/abs/2015A>
- M. Béthermin, E. Daddi, G. Magdis, **et al.** 2015, Evolution of the dust emission of massive galaxies up to $z = 4$ and constraints on their dominant mode of star formation, A&A, 573, 113. <http://adsabs.harvard.edu/abs/2015A&A...573A.113B>
- M. Pannella, E. David, E. Daddi, M.E. Dickinson, H.S. Hwang, **C. Schreiber**, et al. 2015, GOODS-Herschel: star formation, dust attenuation and the FIR-radio correlation on the Main Sequence of star-forming galaxies up to $z \sim 4$, ApJ, 807, 141. <http://adsabs.harvard.edu/abs/2015ApJ...807..141P>
- S. Juneau, F. Bournaud, S. Charlot, **et al.** 2014, Active Galactic Nuclei Emission Line Diagnostics and the Mass-Metallicity Relation up to Redshift $z \sim 2$: The Impact of Selection Effects and Evolution, ApJ, 788, 88. <http://adsabs.harvard.edu/abs/2014ApJ...788..088J>

Acknowledgements

These three years have been the source of many great emotions. Everything started on June 11, 2012, by an immense joy: David, my advisor, told me that my Ph.D. will be funded by CEA.¹ This was the beginning of an adventure. I then met Maurilio. Both him and David were my *Jedi masters* all along my Ph.D. and will probably remain so for the years to come. They each taught me how research is done, in their own different and complementary way: sometimes so different that I came to think of myself as a disciple of both the *light* and *dark* sides of the force. One advocated rigorousness and patience and the other visibility and effectiveness. One would only swear by observations, while the other would sometime invite a theoretician to the discussion. This constant opposition gave me a few headaches, but also made me laugh on many more occasions. It was also an extremely useful experience, which prevented me from ever considering any point of view as “The Truth.” Consequently, David and Maurilio, I think I can say today that you taught me at least one important fact: Being sure of yourself does not come from knowing everything, but from knowing that you are not more mistaken than others. Obviously, I will not talk about all the *other important things* you made me discover—from limoncello to *el Baba*; instead, I will simply thank you for your dedication, for your time, and for all we shared together. The two of you combined were the best Ph.D. advisor one could ever hope for. Thank you for giving me this chance!

My thanks will now go to the two referees, Alberto Franceschini and David Alexander, who had to read the present thesis and validate its correctness. I was very touched by your reports, and I am very happy that you found my work worth a Ph.D. degree. On the day of the public defense, Guillaume Pineau des Forets, Véronique Buat, and Ivo Labb   were added to the jury. I was very honored by your presence, and I would like to thank all for coming and laying a critical eye on my results. I would also like to thank Springer Nature for awarding me the Springer Thesis Prize.

¹The Commissariat à l’Energie Atomique et aux Energies Alternative.

So many thanks should also be awarded to my friends and family; an extensive list can be found in the French version of this Ph.D. Mom, dad, sis: thanks. You know why, I love you.

Leiden, The Netherlands

June 2016

Corentin Schreiber

Contents

1	Introduction	1
1.1	Studying Star Formation in Galaxies: A Problem of Scales	1
1.2	The Main Questions	3
1.2.1	Are Star Formation Histories Smooth or Irregular?	4
1.2.2	Why Are Some Galaxies Forming Much More Stars than Others?	8
1.2.3	Does the Interstellar Dust Hide a Significant Portion of the Star Formation Activity in the Universe?	13
1.2.4	Why Do Galaxies Stop Forming Stars?	19
References		25
2	The Main Sequence of Star-Forming Galaxies as Seen by Herschel	29
2.1	Introduction	30
2.2	Sample and Observations	33
2.2.1	GOODS–North	33
2.2.2	GOODS–South, UDS, and COSMOS CANDELS	34
2.2.3	COSMOS UltraVISTA	36
2.2.4	Photometric Redshifts and Stellar Masses	37
2.2.5	Rest-Frame Luminosities and Star Formation Rates	38
2.2.6	A Mass-Complete Sample of Star-Forming Galaxies	39
2.2.7	Completeness and Mass Functions	42
2.3	Deriving Statistical Properties of Star-Forming Galaxies	45
2.3.1	Simulated Images	46
2.3.2	The Stacking Procedure	46
2.3.3	Measuring Flux Dispersion with <i>Scatter Stacking</i>	51
2.3.4	SFR Dispersion from <i>Scatter Stacking</i>	53

2.4	Results	56
2.4.1	The SFR of Main-Sequence Galaxies	56
2.4.2	Redshift Evolution of the sSFR: The Importance of Sample Selection and Dust Correction	57
2.4.3	Mass Evolution of the SFR and Varying Slope of the Main Sequence	58
2.4.4	Mass Evolution of the SFR Dispersion Around the Main Sequence	59
2.4.5	Contribution of the Main Sequence to the Cosmic SFR Density	61
2.4.6	Quantification of the Role of Starburst Galaxies and the Surprising Absence of Evolution of the Population	64
2.5	Discussion	70
2.5.1	Connection of the Main-Sequence Dispersion with Feedback Processes	70
2.5.2	Connection Between Starbursts and Mergers	71
2.6	Conclusions	72
2.7	Appendix: The <i>UVJ</i> Selection	73
2.8	Appendix: Tests of Our Methods on Simulated Images	74
2.8.1	Mean and Median Stacked Fluxes	76
2.8.2	Clustering Correction	78
2.8.3	Error Estimates	79
	References	84
3	Modelling the Integrated IR Photometry of Star-Forming Galaxies	87
3.1	Introduction	87
3.2	A New Far Infrared Template Library	88
3.2.1	Calibration on Stacked Photometry	92
3.2.2	Calibration on Individual Detections	94
3.3	On the Redshift and Stellar-Mass Dependence of f_{PAH}	96
3.4	Appendix: Recipe for Optimal FIR SED Fitting	98
	References	100
4	EGG: An Empirical Simulation of the Observable Universe	101
4.1	Introduction	101
4.2	Sample Description	102
4.2.1	Multi-wavelength Photometry	102
4.2.2	Redshifts and Stellar Masses	103
4.3	Stellar Properties	104
4.3.1	Redshift and Stellar Mass	104
4.3.2	Star Formation Rate and Obscuration	106
4.3.3	Optical Morphology	107

4.3.4	Optical Spectral Energy Distribution	111
4.3.5	Sky Position	113
4.4	Dust Properties	115
4.5	Generating a Light Cone	116
4.6	Results	117
4.6.1	Comparison to the Observed GOODS–South Field	117
4.6.2	Comparison to the Measured Cosmic Backgrounds.	121
4.7	Appendix: Efficiency of Monochromatic FIR Measurements	122
	References	124
5	The Downfall of Massive Star-Forming Galaxies During the Last 10 Gyr	127
5.1	Introduction	128
5.2	Sample Selection and Galaxy Properties.	131
5.2.1	Multi-wavelength Photometry	131
5.2.2	Redshifts, Stellar Masses and Star Formation Rates	132
5.2.3	CANDELS Sample for the Gas Mass Measurements at $z = 1$	133
5.2.4	HRS Sample for the Gas Mass Measurements in the Local Universe	133
5.2.5	CANDELS Sample for the Morphological Decompositions at $z = 1$	134
5.2.6	Cleaning the 24 μm Catalogs.	136
5.3	Measuring Disk Masses in Distant Galaxies	138
5.3.1	The Bulge to Disk Decomposition	138
5.3.2	Simulated Galaxies	140
5.3.3	Estimating the Disk Mass	141
5.4	Measuring Gas Masses	144
5.4.1	Dust Masses	145
5.4.2	Gas Masses	147
5.5	Results	150
5.5.1	The SFR– M_{disk} Relation at $z = 1$	150
5.5.2	Gas Fraction and Star Formation Efficiency at $z = 1$	152
5.5.3	A Progressive Decrease of the SFE with Time	156
5.6	Discussion	157
5.6.1	Quantifying the “Quenching” and “Downfall” Rates	157
5.6.2	Identifying the Actors that Regulate the SFE and the Gas Content	159
5.7	Conclusions	160
5.8	Appendix: Impact of the UVJ Selection	161
	References	164
6	Reaching the Distant Universe with ALMA	167
6.1	Introduction	167
6.2	Sample Selection	168

6.3	Description of the Observations and Data.	169
6.3.1	Notes on Interferometric Imaging	169
6.3.2	General Properties of Our Data	171
6.4	Data Reduction.	174
6.5	Flux Measurement and Detection Rate.	175
6.6	The $z = 4$ Main Sequence	179
6.6.1	Calibration of the SFR	179
6.6.2	The SFR– M_* Relation.	180
6.7	Other Galaxies in the Field of View.	182
6.8	A Massive $z = 3$ Galaxy Hidden Behind a Bright Star.	183
6.9	Discovery of Two New High-Redshift Dusty Galaxies.	188
6.9.1	Optical to NIR Photometry	188
6.9.2	MIR to FIR Photometry	190
6.9.3	A First Estimate of Their Physical Properties	191
6.9.4	Measuring the Redshift with ALMA	193
6.9.5	Potential Scientific Outcome	196
	References.	197
7	Conclusions and Perspectives	199
	References.	203
	Appendix A: <i>phy++</i>: A C++ Library for Numerical Analysis	205
	Index	217

Abbreviations

AGN	Active Galactic Nucleus
ALMA	Atacama Large Millimeter Array
AOT	ALMA Observing Tool
CANDELS	Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey
CIRB	Cosmic InfraRed Background
CMB	Cosmic Microwave Background
COSMOS	Cosmic Evolution Survey
ECDFS	Extended Chandra Deep Field South
FIR	Far InfraRed
GOODS	Great Observatory Origins Deep Survey
HerMES	<i>Herschel</i> Multi-tiered Extragalactic Survey
HST	<i>Hubble</i> Space Telescope
IGL	Integrated Galaxy Light
IGM	InterGalactic Medium
IMF	Initial Mass Function
IR	InfraRed
IRS	InfraRed Spectrograph
ISM	InterStellar Medium
ISO	Infrared Space Observatory
JCMT	James Clerk Maxwell Telescope
LIRG	Luminous InfraRed Galaxy
MAD	Median Absolute Deviation
MIPS	Multiband Imaging Photometer for <i>Spitzer</i>
MIR	Mid InfraRed
MS	Main Sequence
NIR	Near InfraRed
PACS	Photodetector Array Camera and Spectrometer
PAH	Polycyclic Aromatic Hydrocarbon
PEP	PACS Evolutionary Probe
PSF	Point Spread Function

QSO	Quasi-Stellar Object
SDSS	Sloan Digital Sky Survey
SED	Spectral Energy Distribution
SFE	Star Formation Efficiency
SFH	Star Formation History
SFR	Star Formation Rate
SLED	Spectral Line Energy Distribution
SMBH	SuperMassive Black Hole
SMC	Small Magellanic Cloud
SPIRE	Spectral and Photometric Imaging Receiver
sSFR	Specific Star Formation Rate
UDS	Ultra-Deep Survey
ULIRG	Ultra Luminous InfraRed Galaxy
UV	UltraViolet
Λ CDM	Cosmological constant (Λ) and Cold Dark Matter

Symbols

b	Birthrate parameter (-)
β	UV spectral slope (-)
f_{gas}	Fraction of the baryonic mass found in the form of hydrogen gas in the ISM (-)
f_{PAH}	Fraction of the dust mass found in PAH molecules (-)
Gyr	Billion years (time)
IR8	Ratio of total IR luminosity to 8-μm luminosity (-)
IRX	InfraRed eXcess (-)
kpc	Kiloparsec (length)
L_{\odot}	Solar luminosity, or bolometric luminosity of the Sun, 3.846×10^{26} W (energy/time)
M_{disk}	Stellar mass in the galactic disk (mass)
M_{dust}	Mass of dust particles in the ISM (mass)
M_{gas}	Mass of hydrogen gas in the ISM (mass)
Mpc	Megaparsec (length)
M_*	Stellar mass, or mass of stars (mass)
M_{\odot}	Solar mass, or mass of the Sun, 1.989×10^{30} kg (mass)
Myr	Million years (time)
$12 + \log_{10}(O/H)$	Oxygen abundance (-)
pc	Parsec, 3.087×10^{16} m (length)
SFR	Star formation rate (mass/time)
sSFR	Specific star formation rate (1/time)
t_{dep}	Depletion timescale (time)
T_{dust}	Dust temperature (temperature)
U	Intensity of the interstellar radiation field (energy/time/surface)
yr	Year (time)
Z	Metallicity (-)

Chapter 1

Introduction

1.1 Studying Star Formation in Galaxies: A Problem of Scales

When I first met David, a couple of months before I started working on this thesis, I was surprised to see how little is presently known about how galaxies and stars are born, evolve, and then fade away. My field of expertise at that time was theoretical physics (quantum mechanics, general relativity, quantum gravity) and to me it felt natural that progress in these sub-branches of physics has always been slow. Indeed, these are pioneering theoretical works, often addressing questions that are hard, if not impossible, to connect to the observable world. Extra-galactic astrophysics, on the other hand, deals with objects that, however complex in their structure, are composed of well known elementary bricks: galaxies are made of dust, gas and stars, and each of these components is itself composed of different kinds of atoms, in different proportions and different thermodynamical states. We know how these atoms interact with each other through gravity, electromagnetism, and even quantum mechanics and general relativity, when they matter. Furthermore, these systems are easy to observe: galaxies are found everywhere in the sky, and they evolve on time scales large enough that we can in principle observe even the most distant and faint ones, should we invest enough telescope time. How comes, then, that there are still so many unanswered questions?

I soon realized how wrong my perspective was.

First, this naive picture already starts to break apart if we consider what modern cosmology brings to the game: dark matter and dark energy. While the impact of the latter on individual galaxies is probably negligible, it is nowadays thought that all galaxies live in dark matter “halos” (Blumenthal et al. 1984). These are the descendants of the quantum fluctuations that were amplified during the inflation (Press and Schechter 1974; Peebles 1982), and whose imprint can be seen today on the cosmic microwave background (CMB). The nature of this dark matter, let alone its very existence (e.g., Milgrom 1983), is a matter of debate. However, it is generally assumed that these exotic particles, whatever they are, only interact with ordinary

(or “baryonic”) matter, i.e., what you and I are made of, through gravity. Since the standard model of cosmology predicts that about 84 % of the mass in the Universe is made of this dark matter (e.g., Planck Collaboration et al. 2014), this invisible component is expected to dominate completely the gravitational potential at the largest scales, i.e., above tens of kiloparsecs (kpc) where baryonic processes (e.g., hydrodynamics of the gas, stellar winds, etc.) play little role. For this reason, most people believe that it is the dark matter that shapes the large-scale structures, the so-called “cosmological context” in which the individual galaxies evolve. The best example of this is probably the web-like structure observed in the spatial distribution of galaxies around our Milky Way (e.g., Peacock et al. 2001). This dark matter can be of crucial importance, since the accretion of dark matter halos and the baryonic matter they contain can provide a regular flow of cold gas onto a galaxy, replenishing its gas reservoirs and allowing it to sustain relatively high levels of star formation activity over long periods of time (Dekel et al. 2009). I will come back to this point later.

Second, even if we knew that dark matter existed and if we understood all its properties, it would still be a challenge to predict accurately the birth and evolution of a whole galaxy. Indeed, and contrary to theoretical physics, the complexity of the problem does not arise from unknown interactions, or unknown constituents: it is a problem of *scales*. It is easy to forget this fact, especially since we all work with logarithmic units, but studying galaxy evolution requires dealing with scales that span more than ten orders of magnitude (e.g., going from a star to a galaxy¹). At the time of writing, the most ambitious numerical simulations attempting to describe a whole galaxy are only able to span about six orders of magnitude in spatial scales, reaching a resolution of about 0.1 pc (e.g., Renaud et al. 2013). Below this minimum scale, “sub-grid” recipes are used to emulate the complex physics that is supposed to take place (e.g., McKee and Ostriker 2007): cooling the gas by interaction with dust grains, then collapsing this gas to form new stars, generating stellar winds, and eventually creating super-novae. On the other hand, other numerical simulations can tackle the aforementioned processes by using better resolution in smaller volumes, but then they lack the global context of the whole galaxy, i.e., the gas flows and the associated turbulence coming from the larger scales.

Third, scales are also a problem observationally. Not so much spatial scales as *time* scales. It is very convenient that galaxies evolve on long time scales, typically of the order of millions of years, because we can re-observe the same region of the sky in intervals of several years with different instruments, and still consider that we observe the same system. But this is also a huge issue: once we observe a galaxy in a given state, we can predict what its future could be, but we will never be able to *see* this future and confirm our prediction. Or at least not in a human lifetime. It is as if a detective had to solve a crime from a single photograph, shot possibly long after the criminal was gone. While that can be an interesting source of inspiration, it is not scientifically pertinent to ask ourselves what will become of a particular galaxy,

¹One could push as far down in scales as the size of a dust grain, a fraction of microns, and up to the size of a galaxy cluster, a couple of mega-parsecs (Mpc), to span about *thirty* orders of magnitude. Fortunately, not all these scales are coupled, so it is possible to study them separately to some extent.

because that is not an observable. The only way we can constrain the evolution of galaxies is therefore by studying *populations* of objects, and establish probable causality links. Thanks to the fact that the speed of light is finite, we can also observe the Universe at various epochs and link together populations of galaxies in terms of progenitors and descendants (e.g., van Dokkum et al. 2010). We can also make the link between two properties of a given galaxy population, for example the star formation rate as a function of the stellar mass, and use models to see what we can learn from these observed relations. All of our work is therefore based on such statistical arguments.

A philosophical inconvenience emerging from this issue of scales is that we cannot make *experiments* in the scientific sense. We cannot “take” two galaxies and make them collide to see what happens. Or capture a galaxy and compress it to see if it suddenly forms more stars. Worse, we only have one Universe to study. If one day we observe the whole sky, and probe the entirety of the observable Universe, we will probably be able to find several complex enough models that will reproduce all these observations.² Having no additional data to rule them out, we will not be able to learn anything more. Fortunately, this is not going to happen any time soon. But still, strictly speaking, and much like cosmology, it can be argued that extra-galactic astrophysics is not a science.

Does it mean that it is not worth spending time and money on these issues? Of course not. In a way, astrophysics is very close to archeology, in that we try to understand our past from what we see today. The fact that we cannot really *manipulate* or *reproduce* anything does not prevent us from learning much about how our Galaxy and our world came to be. And it goes even beyond this: it is through astrophysical observations that we have made among the most exciting breakthroughs of the last century. Not only by confirming the predictions of theories, with the existence of super massive black holes or gravitational lenses, but also with completely unexpected discoveries, for example with the cosmic microwave background, the expansion of the Universe, or the need for dark matter, dark energy and/or modified gravity. Astrophysics allows us to look at ourselves in a wider context, with a broader perspective. It brings ingredients to physics that, without looking up at the sky, we would have never thought about.

These are the reasons that motivated me during the last three years, and that, hopefully, will keep on amazing me for the years to come.

1.2 The Main Questions

In this section, I introduce the specific questions I address in this thesis, what we have learned from previous studies, and what they left as unknown. I intentionally do not reveal my own results here, and instead describe the state of the art as it was before

²Much like there is always an infinity of functions that fit exactly to a finite number of points. Or much like, and this is an intended pun to my particle physicist friends, we can always explain any observations at the LHC by adding new particles to the standard model of particle physics.

the work I have done in Saclay was published. Since this is not an epistemological study, I will not attempt to follow a chronologically rigorous path, nor to report all the previous dead ends that were explored and later abandoned. In the process, I will overlook a number of studies and be unfair to many researchers, all for the sake of clarity. For this, I hope they will accept my apologies.

1.2.1 Are Star Formation Histories Smooth or Irregular?

One of the major goals of our field is to learn about the *star formation history* (SFH) of galaxies, or, in other words, the variations over time of the *star formation rate* (SFR), the rate at which each galaxy is forming new stars.

For example, it is known from detailed study of the properties of stars in our neighborhood that the Milky Way has experienced frequent variations of its star formation rate in the past, about every 500 Myr (Hernandez et al. 2000). These “bursts” seem to happen on top of a slowly varying, continuous activity that showed a peak about 3 Gyr ago (Cignoni et al. 2006), as shown in Fig. 1.1. One can also refer to the review of Wyse (2009) for further details. The mechanisms that shape this SFH are still poorly understood today. The regular bursts could be associated with merging events, i.e., the accretion of other smaller galaxies (“dwarf” satellites)

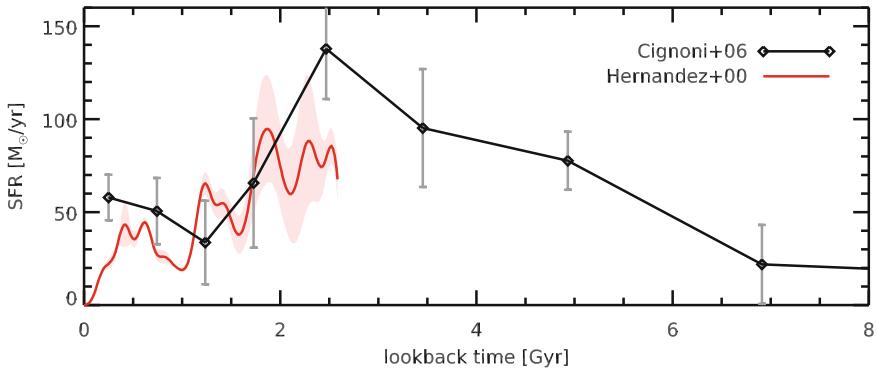


Fig. 1.1 Estimated star formation history (SFH) of the Milky Way (MW). On the y axis is the MW’s star formation rate (SFR), in units of solar masses per year (M_{\odot}/yr), and on the x axis is the lookback time in billion years, i.e., today is on the left, and the Big Bang is on the right of the plot. The *black solid line* shows data from Cignoni et al. (2006), which cover a large time window with a poor resolution, and the *red line* comes from Hernandez et al. (2000), which focus on the last 3 Gyr with a significantly higher time resolution. Both data sets were published in arbitrary units, and are here renormalized to a common reference. The SFH of Cignoni et al. (2006) is rescaled so that integrating it over time yields a total stellar mass of $6.1 \times 10^{11} M_{\odot}$ (Flynn et al. 2006), assuming no mass loss and no merger. The data of Hernandez et al. (2000) are rescaled so that the integral over time between 0 and 3 Gyr matches that of Cignoni et al. (2006). This simple approach is roughly consistent with the MW’s present-day SFR of $4 M_{\odot}/\text{yr}$, as measured by Diehl et al. (2006)

on our Milky Way. Not only will these galaxies bring additional stars, they will also briefly destabilize the gas of the Galactic disk and allow it to collapse and form stars more efficiently (this process is discussed further in Sect. 1.2.2). Another explanation which is put forward in Hernandez et al. (2000) is that, since this SFH is estimated from the solar neighborhood only, i.e., a relatively small region compared to the whole Milky Way, these bursts could correspond to the regular passage of the spiral arms. Somewhat surprisingly, the spiral arms are nothing but density waves inside the disk: they are not representative of the motion of individual stars, but emergent patterns caused by different orbits around the Galactic center (Lindblad 1960). When this density wave reaches a given region of the disk, it creates local variations of the gravitational potential and also destabilizes the gas, perhaps in a less efficient way than mergers. This means that, if we were to estimate the SFH from a larger sample of stars not limited to the solar neighborhood (something that *Gaia* will soon provide), these variations would vanish, and the SFH of the whole Galaxy would appear much smoother. However, a feature that is expected to remain would be the larger peak that happened 3 Gyr ago. This enhancement of star formation may instead be caused by a *major* merger, the collision of the Milky Way with another galaxy of similar mass, or by a more intense flow of gas coming from the intergalactic medium (IGM), through a process called “infall” (see, e.g., the discussion in Kennicutt 1983).

Indeed, our Galaxy must have received large quantities of gas from outside in its past, and probably does so even today. Its present-day star formation rate is currently estimated around $\text{SFR} = 4 \text{ M}_\odot/\text{yr}$ (Diehl et al. 2006), while the mass of gas available to form stars is of the order of $M_{\text{gas}} = 2 \times 10^9 \text{ M}_\odot$ (van den Bergh 1999). Therefore, at this rate the Milky Way would consume all its gas within 500 Myr (see van den Bergh 1957 where this problem was first reported). This latter quantity is known as the *depletion timescale*, t_{dep} . Such short timescale is not specific to the Milky Way: except for a few exceptions which are not representative of star-forming galaxies (e.g., M31 with $t_{\text{dep}} = 5 \text{ Gyr}$, Pflamm-Altenburg and Kroupa 2009), the depletion timescales in the majority of galaxies is typically no more than 1 Gyr (see, e.g., Saintonge et al. 2011). This is probably the best evidence that galaxies routinely receive gas from the intergalactic medium.

The question is then, how do galaxies actually consume this gas? Is it mainly through merger events, with episodes of intense triggered star formation? Or rather in a more peaceful but steady way, similar to the density waves created by the spiral arms in the disk? While both channels are known to generate star formation in all galaxies, it remains uncertain today which one typically *dominates* the star formation histories of galaxies. Ideally, one would transpose the studies described above from the Milky Way to other galaxies, and build a statistically meaningful sample. However, this kind of analysis requires counting individual *stars*, and that is something we can only do in our closest environment for a handful of galaxies, using the *Hubble Space Telescope (HST)*.

A key element to answer this question was brought forward by the observation of large samples of galaxies for which we could obtain good estimates of the current SFR and the stellar mass (M_*), both today (Brinchmann et al. 2004) and at earlier

epochs in the history of the Universe, e.g., 8 Gyr ago (Noeske et al. 2007; Elbaz et al. 2007). These observations aimed at studying the correlation between the SFR and the stellar mass. The connection with star formation histories becomes obvious once we consider that the stellar mass is the integral over time of the past star formation history,³ $\int_0^{t_{\text{now}}} dt \text{SFR}(t)$, while the SFR is just $\text{SFR}(t_{\text{now}})$. In fact, the quantity of interest here is the *specific* star formation rate, which is the amount of star formation rate per unit stellar mass: $\text{sSFR} \equiv \text{SFR}/M_*$. If one assumes that, at a given epoch, all galaxies have roughly the same age, then this sSFR is proportional to the *birth rate parameter*: $b \equiv \text{SFR}(t_{\text{now}})/\langle \text{SFR}(t) \rangle$ (Kennicutt 1983), where $\langle \text{SFR}(t) \rangle$ is the average of the past SFR. This parameter can be used to estimate the typical “burstiness” of star formation histories. Indeed, if all galaxies were forming stars at a constant rate, by definition they would have $b = 1$, and at a given epoch they would all have exactly the same sSFR. If on the other hand the star formation histories are very bursty, then one would expect to see a wide distribution of b (or sSFR).

What was actually observed was a correlation between the SFR and M_* (in logarithmic space), with more massive galaxies having higher rates of star formation, and most importantly with a relatively low scatter in SFR of about a factor of two at fixed stellar mass. An example is shown in Fig. 1.2. Since this scatter actually includes measurement errors, the intrinsic scatter is expected to be even lower. For this reason, this correlation was named the “Main Sequence” of star-forming galaxies (MS): as time goes, galaxies are growing in mass and “climb up” the sequence. Their star formation rate increases, until they stop forming stars and “fall down” (this process of shutting down star formation is discussed later in Sect. 1.2.4). This was a very strong step forward. The fact that, over about half the age of the present-day Universe, we observe variations of SFR of at most a factor of two from one galaxy to another places strong upper limits on the variations of the SFH within individual galaxies. This was immediately understood as a sign that these SFHs may be relatively smooth, and that mergers could only play a minor role in the whole star formation story (see in particular the discussion in the following section).

Since then, numerous studies have attempted to refine the measurement of the Main Sequence. Indeed, mostly because of the presence of dust, correctly estimating SFRs and stellar masses is not always an easy task depending on what data are available, and these estimates are often subject to systematic biases. In fact, many assumptions have to be made in order to derive these quantities, since the only thing we observe is the projected brightness of the galaxy at various wavelengths. It is reasonable to worry that these assumptions may lead to wrong conclusions because they oversimplify the situation; for example the real scatter in sSFR may actually be larger than we think. For this reason, the measurement is regularly revisited with

³Neglecting, for simplicity, the loss of stellar mass due to the death of stars. Assuming the stellar lifetimes of Bressan et al. (1993), this is 25–30 % of the total mass after 1 Gyr for a typical star formation history, and up to 40 % after 10 Gyr for an maximally old galaxy. These numbers were obtained assuming the Salpeter (1955) prescription for the mass distribution of newly born stars (the initial mass function, IMF).

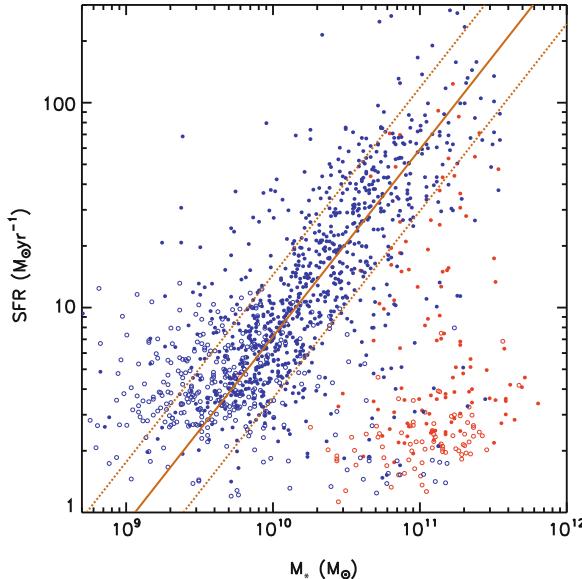


Fig. 1.2 The correlation between the star formation rate (SFR) and the stellar mass (M_*) of distant galaxies ($z = 1$), adapted from Elbaz et al. (2007). This is the so-called Main Sequence of star-forming galaxies. Points are colored according to the rest-frame ($U - B$) color of each galaxy, i.e., *blue points* for blue galaxies, and *red points* for red galaxies. This color is a proxy for the age of the stars: the redder the galaxy, the older are the stars it contains. Therefore, a red galaxy is supposed to have stopped forming new stars long ago (except if the galaxy contains dust, as discussed later in Sect. 1.2.3), while a blue galaxy must still be actively star-forming. The *solid line* indicates the best-fit power law of blue galaxies, and the *dotted lines* indicate the dispersion around this trend

deeper or more varied photometry. In particular, I present in Chap. 2 the results I have obtained during this Ph.D., building on the work of Elbaz et al. (2011), and taking advantage of the deepest *Herschel* and *Hubble* data to study the spectra of galaxies from the ultra-violet (UV) to the far-infrared (FIR). In this study, we measure the most accurate stellar masses and SFRs, looking back in time as far back as 12 Gyr ago, where the Universe was barely 1 Gyr old, and probing for the first time such a wide time window in a consistent way across such a wide range of wavelengths.

This is clearly not the end of the story though, because there are many regimes that we could not probe, especially the low-mass dwarf galaxies, or the first billion years of the Universe. This study is also not free of assumptions and biases, but I guess it is fair to say that this was, at the time, the best we could do. Further progress will surely emerge out of the new generation instruments that were recently (or will soon be) deployed: the Atacama Large Millimeter Array (ALMA), the *James Webb Space Telescope (JWST)*, *Euclid*, ... For example, by observing a region of the sky for only a few minutes, ALMA is able to detect galaxies that are up to ten times fainter

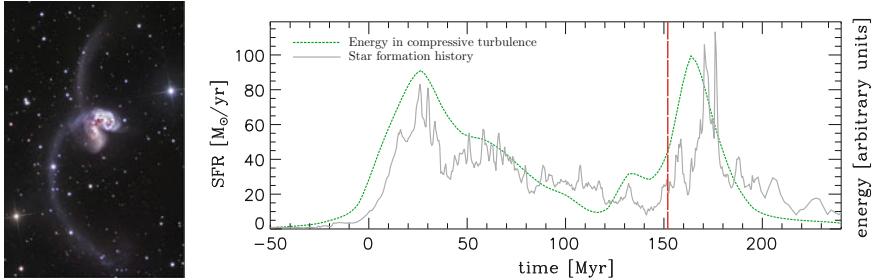


Fig. 1.3 *Left* NGC 4038 and 4039, also known as the Antennae galaxies, an on-going major merger. Both are seen here in the center of the image, and can barely be distinguished from one another. The extended arc-like features, which inspired the name of this pair of galaxies, are called *tidal tails*. These are made of gas and stars that were stripped from both galaxies through complex gravitational interactions. Copyright: SSRO, reproduced with permission. *Right* Simulated star formation history of the Antennae, adapted from Renaud et al. (2014). This data comes from a full 3D model of the merger, with a maximum spatial resolution of 1.5 pc, including gas and stars (see the description of the simulation in Renaud et al. (2015)). The evolution of the SFR with time is shown with a *solid gray line*. The *green dotted line* shows the energy in compressive turbulence, which is predicted by Renaud et al. (2014) to be the dominant way through which mergers densify the gas and trigger additional star formation. The *red line* indicates the instant where the simulation matches best the current observed state of the Antennae galaxies

than what the deepest surveys *Herschel* could achieve by observing for about 200 h.⁴ Taking advantage of this incredible sensitivity, we have created a targeted survey with ALMA to study in more depth the young Universe. The data were received in early 2015, and are described in Chap. 6.

1.2.2 Why Are Some Galaxies Forming Much More Stars than Others?

The existence of the Main Sequence does not nullify the impact that galaxy mergers can have on individual galaxies. It is clear that the most extremely star-forming galaxies in our neighborhood are actually pairs of merging galaxies (Sanders and Mirabel 1996) with very large sSFRs, indicating that they are likely short but intense phases in the lifetime of these galaxies. Indeed, in an isolated galactic disk, star formation is relatively inefficient because the gas in the disk is stabilized by the shear forces created by the *differential* rotation of the disk, i.e., the fact that the rotation speed is not the same at all radii (Toomre 1964). It was shown only recently in numerical simulations that mergers trigger additional instability of the gas through compressive tidal forces (see Renaud et al. 2014 and Fig. 1.3), effectively allowing a

⁴Note that this comparison is slightly unfair, since ALMA has a very limited field of view (about 15''), and its efficiency is substantially reduced when it comes to mapping an entire field, which is what *Herschel* was designed for.

substantial fraction of this gas to collapse and form stars in short timescales (several hundreds of Myr).

Within the Λ CDM cosmological paradigm, it is relatively straightforward to predict the rate of such merging events. Since this is a purely gravitational problem, the baryonic physics does not play an important role, and one only needs to care about dark matter which is much simpler to model. For this reason, the predictions arising from numerical simulations or choke-and-blackboard theory should be relatively robust. In fact, the general expectation is that mergers were much more frequent in the past, because the Universe was overall more homogeneous: today, most of the structures that could merge have already done so. For example, Hopkins et al. (2010) predicts that major mergers (with a mass ratio of at least 1/3) happen on average every 40 Gyr per galaxy today, while this number would go down to every 4 Gyr if we consider the Universe as it was 10 Gyr ago, i.e., ten times more frequently. Similar trends were found in observations: Kartaltepe et al. (2007) reported that the fraction of bright paired galaxies is only 0.8 % today, but was closer to 8 % about 10 Gyr ago; a similar factor of ten difference. However, linking observed pair fractions with actual merger rates is difficult. While these numbers are corrected for chance projections, without precise velocity measurements it is unknown what fraction of these pairs will actually end up merging. Perhaps even more importantly, one also need to make assumptions on the observability timescale of a merging event. These uncertainties are such that a broad range of scenarios were reported in the literature, from strong to almost no evolution of the merger rate (see, e.g., the compilation of Lotz et al. 2011), but always with a tendency for a decrease with time.

The net consequence is that we expect galaxy mergers to have played a more important role in the past. Interestingly, and as was seen in the previous section, it was observed that star formation was also globally more intense at these epochs, by about an order of magnitude. Could this be due to the larger merger rates? While this is a tempting explanation, the observed distributions of sSFRs are incompatible with this hypothesis (in the following I summarize the original discussion from Noeske et al. 2007). Assuming mergers play a negligible role today because they are rare, we can consider that the typical sSFR we observe in our neighborhood is that of isolated, non-interacting galaxies. Going back in time, as mergers were more frequent, we expect that some galaxies experience episodes of enhanced star formation, and have higher sSFR, while the rest of the population stays at the low sSFR of isolated galaxies. If mergers are predominantly responsible for the intense star formation activity in the distant Universe, a large number of galaxies should be found in this enhanced state and in the end we should see a double-peaked (or *bimodal*) distribution of sSFR: one peak created by bursty mergers with high sSFR, and another created by non-interacting galaxies with low sSFR. No such strong bimodality is observed among star-forming galaxies, as all galaxies appear to have the same sSFR within a factor of two (see previous section). Therefore, if mergers have any impact, it must be reasonably small, and in any case they cannot be responsible of setting the global star formation rate density in the Universe.

It is actually within this dispersion of a factor of two in sSFR that we think mergers may play their role. One could imagine that a fraction of the galaxies with sSFRs above the average are actually triggered by mergers. This is a path that was explored in several papers, e.g., Elbaz et al. (2011), Rodighiero et al. (2011), Sargent et al. (2012). In particular in Elbaz et al. (2011), David and his co-authors found that, on average, the galaxies that showed an excess sSFR were also showing a specific signature in their spectra (what they called IR8) that could be interpreted as an increased *compactness* of their star-forming regions. This compactness, in turn, may be a hint that a major merger event recently happened. Indeed, it is clear at least in the nearby Universe that mergers do generate very compact star-forming regions (e.g., Armus et al. 1987; Sanders et al. 1988). In the distant Universe, however, this is still a poorly explored territory.

Inspired by pioneering works done with the ISO satellite (Franceschini et al. 2001), Rodighiero et al. (2011) and Sargent et al. (2012) analyzed the sSFR distributions in a deep *Herschel* survey, and found that these distributions could be well described by a simple two-component model, dubbed “Two Star Formation Mode” (2SFM, Sargent et al. 2012). In this framework, most galaxies are in the “Main Sequence” (MS) mode, with sSFRs varying by a bit less than a factor of two, and a small fraction are in the “Starburst” (SB) mode, with a systematic enhancement of their sSFR by about a factor of a few compared to Main Sequence galaxies. In practice, although the philosophy is radically different, this is conceptually very similar to the bimodal sSFR distribution introduced in the previous paragraph (Noeske et al. 2007). The main differences are that isolated galaxies and mergers are replaced by the anonymous “Main Sequence” and “Starburst” galaxies, and that these starbursts are a clear minority, both in numbers (3 %) and star formation rate density (10 %) so that no strong bimodality emerges, consistently with the argument of Noeske et al. (2007).

This finding can be related to another scaling law, namely the Schmidt-Kennicutt (SK) law (Schmidt 1959; Kennicutt 1983), which is one of the most important building block of star formation as we know it. This scaling law tells us that the density of star formation in a given volume is directly connected to the density of hydrogen gas in the same volume.⁵ The correlation is super-linear, meaning that gas is more

⁵In practice, it is more common in the literature to use *surface* densities instead of the probably more intuitive *volume* densities. This is actually what was done since the very beginning in the original paper by Schmidt (1959), where he considered star formation inside the disk of the Milky Way. While the disk is actually made of several components, a young thin disk, and an older thick disk, Schmidt assumed that this difference was simply caused by the passing of time, and that all stars were born in a disk of non-evolving width (between 200 and 800 pc, depending on the distance to the center of the Galaxy). For this reason, he averaged his observables along a direction perpendicular to the galactic plane, leading to surface densities of star and gas. Probably by convention, it has remained the standard ever since. When studying distant galaxies, it is questionable whether this choice makes any sense, since only a fraction of these galaxies actually have a clear disk structure (see, e.g., Labb   et al. 2003). However, from a more practical point of view, surface densities are model-independent observables, while volume densities cannot be measured without knowing the extent of the object about the third dimension, an information that is often missing and has to be assumed.

efficiently converted into stars in denser environments. Recently, it was argued that this scaling law was subject to a bimodal behavior (Daddi et al. 2010; Genzel et al. 2010), with a sequence of “disks” and a sub-population of “starbursts” with greatly enhanced star formation efficiency, see Fig. 1.4. Interestingly, these outliers to the SK law are also outliers to the SFR– M_* relation, indicating that they are indeed growing in a different *mode*, which Daddi et al. (2010) also suggested to be triggered by major mergers.

Once again, if major mergers are indeed the cause of these starbursts, then the number of such starbursts should have been larger in the past, where mergers were more frequent. Some studies have already reported such an evolution (e.g., Dressler et al. 2009), however it is important to note that these results are very sensitive to the exact definition of a “starburst”. In studies focusing exclusively on the local Universe, it is not uncommon to refer to a galaxy as a starburst if its sSFR (or, worse, its SFR) is larger than a given value. This definition breaks down as soon as one looks back in time, where the SFRs were *globally* higher, as it would imply that most galaxies in the distant Universe were starbursts. Because it lacks a proper reference point to anchor itself to, this definition is not very useful. An alternative, more interesting

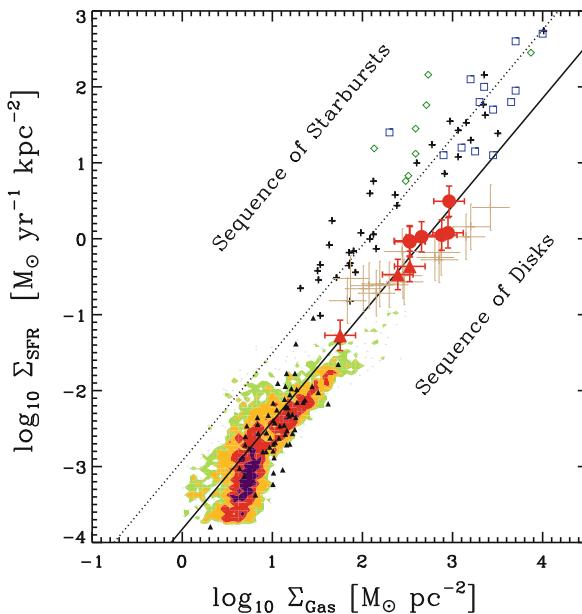


Fig. 1.4 Correlation between the surface density of star formation rate (Σ_{SFR}) and the surface density of hydrogen gas (Σ_{gas}), the so-called Schmidt–Kennicutt law, adapted from Daddi et al. (2010). The red triangles, red circles, brown crosses, black triangles, and the shaded region at the bottom represent “normal” star-forming galaxies at various epochs in the history of the Universe. The black crosses, green diamonds and blue squares are ultra-luminous galaxies, i.e., starburst galaxies (see the original paper for details). The black solid line is the best-fit power law to the normal galaxies, and the black dotted line is this same power law adapted for starbursting systems

definition uses the birth-rate parameter b , introduced in the previous section. One can define a starburst as a galaxy with $b > N$, i.e., a galaxy whose current SFR is at least N times more than its past average (with, e.g., $N = 2$ as in Heckman et al. 1990). While more physically motivated, such a definition also suffers from a bias, this time toward young galaxies, or equivalently, toward all galaxies in the young Universe. Indeed, being young, their SFR can only be rising, and their birth-rate parameter must consequently be larger than 1. This does not necessarily mean that they are evolving in a particular way, and for all we know, a young galaxy with $b > 1$ could be growing like any other young galaxy. Picking a threshold in b high enough should prevent this bias, but this precise threshold depends on the expected star formation history. For example, all star formation histories following the “delayed exponentially declining” functional shape, where $\text{SFR}(t) \propto t \exp(-t/\tau)$, start with $b = 2$ and only fall below $b = 1$ after $t \gtrsim 2\tau$, i.e., when the galaxy has already formed more than half of its final mass.

Instead, the definition I will be using in this thesis is the one that allows us to pinpoint *unusual* behaviors, galaxies whose star formation rates are different from that of other galaxies with otherwise similar properties observed at the same epoch in the history of the Universe. One way to achieve this is to define a galaxy as a “starburst” if its SFR is at least N times higher than the average SFR of galaxies of the same stellar mass, at the same epoch. Interestingly, this definition alone does not allow us to disentangle between two different scenarios, corresponding to different *duty cycles*, i.e., the time a given galaxy spends in the starburst mode. First, the enhancement of star formation could be rare, in the sense that it happens only in a few particular galaxies that will always be highly star-forming, while all the others will never experience it in all their lifetime. Second, the enhancement could be more common, but sustained over very short periods of times so that we only see a scant of starbursts at a given instant. Actually, it could very well be both at the same time. A funny picture I have in mind to illustrate this degeneracy is to consider the photograph of a pool filled with frogs, where a handful of these frogs are seen hanging in the air. What could be happening to them? We know that frogs tend to leap quite often, so a natural explanation is that the ones that are hanging in the air were just caught in the act of jumping, and that they will fall down a couple of seconds later. But that’s on Earth. Now, what do you think would happen if instead the photograph was taken on the Moon?⁶ It could very well be that most frogs preferred to stay safe in the water, while a few adventurous ones attempted to jump some time ago and remained hanging above the pool for several (long) hours, lacking sufficient gravity to fall back toward the pool. The fact is, from the picture alone, we cannot tell between these two alternatives. We need to bring additional information, i.e., on which planet the photograph was taken, to figure out what is actually going on. In the case of the starbursts, it is the depletion timescale that helps us disentangling the different scenarios: as can be seen from Fig. 1.4, at fixed gas mass, starbursts are forming stars about ten times faster, therefore their depletion timescales are very low

⁶ And if you are willing to assume, for the sake of the argument, that there are pools and frogs on the Moon.

(of the order of 100 Myr, see, e.g., Daddi et al. 2010). For this reason, we know that starbursts cannot remain starbursts for a long time, and unless their reservoirs are quickly replenished with enormous amounts of gas, their star formation activity has to fall down soon after they are observed. This is very well matching the expected behavior of a galaxy experiencing a major merger.

For this reason, in Chap. 2 (Sect. 2.4.6), I use this definition to study the time evolution of the starburst population observed in our deep *Herschel* surveys, and compare it to the trends expected for major mergers to learn more about this extravagant population.

1.2.3 Does the Interstellar Dust Hide a Significant Portion of the Star Formation Activity in the Universe?

The first estimate of the SFR density in the Universe was established by measuring the evolution of the UV luminosity of galaxies at different epochs (Lilly et al. 1996; Madau et al. 1996). Indeed, the sum of the UV light emitted by all the stars in a galaxy is a good tracer of the galaxy’s current star formation rate (e.g., Kennicutt 1998). In star-forming galaxies, the majority of the UV light is produced by very hot stars, which are at least five times more massive and several hundred times more luminous than our Sun. Because of these extreme masses, their gravitational potential is higher, the hydrogen gas they contain is heated to higher temperatures and therefore converted faster into helium. For this reason, massive stars have very short lifetimes of less than 100 Myr, compared to the estimated 10 Gyr of our Sun. Knowing this, one can use these stars as a signpost of “recent” star-formation, with a time resolution of about 100 Myr.

There is an issue though. This UV light is made of energetic photons, with wavelengths between 150 and 300 nm. This spatial scale happens to be smaller than the typical size of the *dust grains* that are present in the interstellar medium (ISM, see, e.g., Zubko et al. 2004). Therefore, whenever a UV photon intercepts the course of a dust grain, it has a non-negligible probability of being absorbed (or scattered) by this grain, and may never reach our telescope. Depending on the density of dust along the line of sight, only a fraction of the UV light of a galaxy actually manages to escape, and star formation rates can therefore be severely underestimated. An example of such a situation is shown in Fig. 1.5.

Fortunately, we know of different ways to recover this missing light. The most direct one is to look for the energy that was absorbed: since energy is always conserved, it has to come out of the dust grain one way or another. In fact, it does so through *thermal radiation*. When a grain absorbs a UV photon, in virtue of conservation of momentum, the energy it acquires is transmitted in the form of kinetic energy to the individual molecules that compose it. If the grain absorbs such photons at a high enough rate (which is the case for the biggest grains which have the largest cross-section), the grain itself thermalizes and reaches a temperature of a few tens

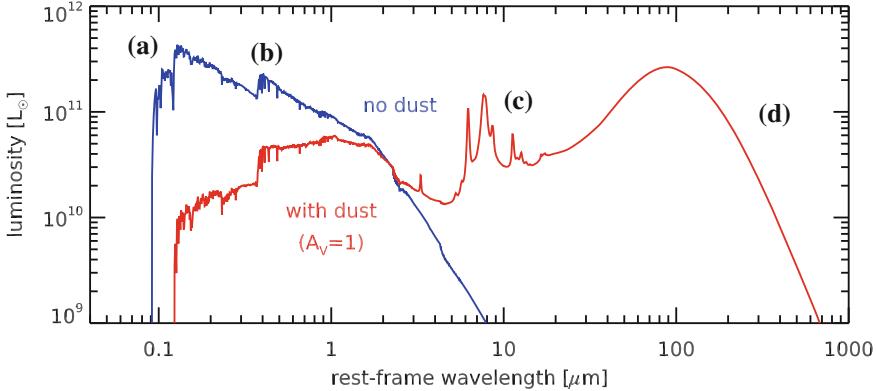


Fig. 1.5 Spectral energy distribution (SED) of a typical star-forming galaxy similar to our Milky Way. I show here the intensity of the emitted light (in units of L_{\odot} , our Sun's own luminosity) as a function of wavelength. The blue curve shows how the SED of the galaxy would look like in the absence of interstellar dust, using the Bruzual and Charlot (2003) stellar population models and a constant star formation history. The red curve shows the actual observed SED, where dust absorbs a non-negligible fraction of the stellar light (an attenuation of one magnitude in the V band, i.e., at $\lambda \simeq 0.6 \mu\text{m}$). The extinction law is taken from Calzetti et al. (2000), and the dust emission is produced using the models of Galliano et al. (2011). (a) This sharp decrease of the light intensity at $\lambda \simeq 0.1 \mu\text{m}$ is called the *Lyman break*. Photons emitted at wavelengths shorter than this threshold have enough energy to fully ionize hydrogen atoms, regardless of their excitation state. They are therefore very easily absorbed, either within the galaxy, or somewhere along the line of sight in the intergalactic medium (IGM). The net consequence is that we receive essentially zero photons shortward of the Lyman break. (b) This second break in the SED at $\lambda \simeq 0.4 \mu\text{m}$ is called the *Balmer break*. This is conceptually identical to the Lyman break, except that this time the photons just have enough energy to ionize an hydrogen atom if this atom is in its first excited state, or above. Because a good fraction of the hydrogen atoms are in their ground state, photons with wavelengths shorter than the Balmer break have a fair chance of not being absorbed. However, the Balmer break is almost coincident in wavelength with another feature, called the 4000 Å break. This break arises in the atmosphere of the stars themselves, and is the result of more complex opacity processes due to non-hydrogen atoms (e.g., calcium). (c) These prominent features in the mid-infrared between $\lambda = 5$ and $15 \mu\text{m}$ are created by a combination of numerous emission lines which are emitted by large carbonated molecules, called polycyclic aromatic hydrocarbons (PAHs). This is a peculiar type of dust grain typically found within star-forming regions. They are relatively fragile, and tend to be destroyed by too intense radiation fields. Their connection to physical processes inside galaxies are still not very well understood. (d) This is the emission of normal dust grains, re-emitting the stellar light that was absorbed. It is the sum of many gray bodies of different temperatures, ranging between a few tens to a hundred of Kelvins. I show here a typical such combination, but the shape of this part of the spectrum can vary significantly from one galaxy to another, depending on the geometry of the dust clouds and their position relative to young stars, but also on the physical composition of the dust (i.e., silicate *versus* carbonated grains, and the grain size distribution)

of Kelvins. According to Planck's law, a black body of such temperature will radiate its energy by emitting photons at wavelengths of the order of $100 \mu\text{m}$. This falls in the FIR domain, which is also commonly referred to as the “sub-millimeter” domain (the right hand side of Fig. 1.5). Therefore, if one can measure the luminosity of a

galaxy in the FIR, and since the dust is transparent to these wavelengths, one can add it back to the *observed* UV luminosity to recover the *intrinsic* UV luminosity, and eventually measure an accurate SFR. In most star-forming galaxies, dust attenuation is such that the FIR luminosity is usually much higher than the observed UV luminosity. For this reason, star formation rates obtained this way are usually dubbed “FIR-based”.

The main issue with this approach is that measuring the FIR luminosity is not always easy, and for two reasons. The most important one is that our atmosphere is not transparent between 3 and 800 μm (the atmospheric transmission is poor), precisely because its temperature makes it also radiate at these wavelengths. Some observatories (like JCMT and ALMA) allow us to observe at 300–400 μm , but the observing times needed to detect anything but the brightest nearby objects are usually prohibitive. For this reason, most of what we know of this wavelength domain comes from space telescopes like *Spitzer* and *Herschel* (to name only the two most recent ones), which are obviously not bothered by the atmosphere. However, there comes the second issue: at these wavelengths, the angular resolution is two orders of magnitude worse than in the optical, because of the increased diffraction (which is proportional to the wavelength). This means that most of the distant galaxies observed by *Hubble* are nothing more than large “blobs” in FIR images: we cannot see their detailed structure, and worse, galaxies appear so big that they tend to overlap, making it difficult (if not sometimes impossible) to robustly attribute the observed flux to the right counterpart. This is called the problem of “confusion”, and is illustrated in

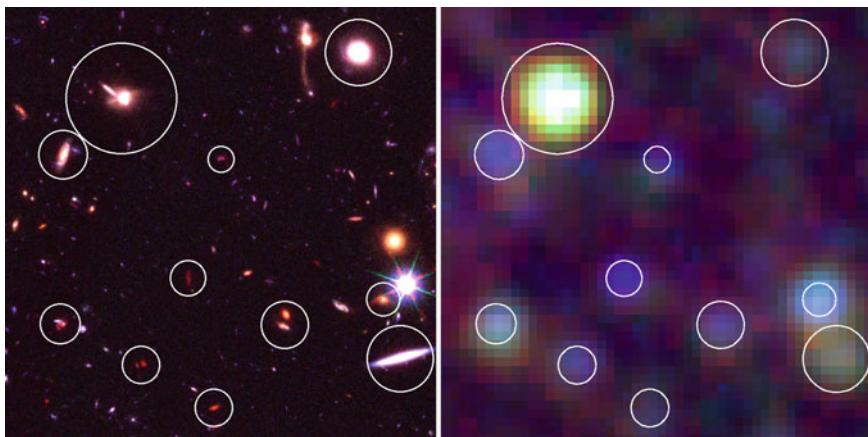


Fig. 1.6 *Left* A $30'' \times 30''$ region of the cosmological deep field COSMOS-South, as observed by *Hubble* and shown here in false colors (F606W+F850LP+F160W, i.e., green, red, and near-infrared). These very deep observations allow us to detect many galaxies at varying distances. *Right* The same region of the sky observed this time by *Spitzer* ($24 \mu\text{m}$ as blue) and *Herschel* ($100 \mu\text{m}$ as green, and $160 \mu\text{m}$ as red). The most obvious detections are pinpointed with white circles, and reported on the *Hubble* image. These two pictures give an example of the fraction of galaxies for which we have a far-infrared detection

Fig. 1.6. The only way to reduce this diffraction is to use larger mirrors.⁷ However, because of practical constraints during launch, the size of the mirror on these space telescopes is much smaller compared to that of their ground-based equivalents. For example, the mirror of *Spitzer* has a diameter of only 0.9 m, while that of *Herschel*, the largest ever launched, is 3.5 m wide. In comparison, the standard mirror size for optical telescopes nowadays is about 8 m, and up to 10 m for the largest ones. Sub-millimeter ground-based telescopes like the JCMT can reach even up to 15 m.

In the end, in a typical cosmological deep-field observed by *Hubble*, *Spitzer* and *Herschel*, we can measure the FIR luminosity of only 15 % of the galaxies with stellar mass larger than $3 \times 10^9 M_\odot$ (see Sect. 2.2.6). The other ones are too faint to be detected in the FIR, even on the deepest *Spitzer* and *Herschel* images. There are of course many things to do with these 15 %, and the study of these detections has provided a wealth of key results during the past ten years (e.g., Elbaz et al. 2007; Magnelli et al. 2009; Tacconi et al. 2010; Daddi et al. 2010; Rodighiero et al. 2011; Elbaz et al. 2011; Magdis et al. 2012). In this thesis, I present in Chap. 2 (Sect. 2.2.6) new results about the evolution of the starburst population (introduced in the previous section) which are only based on this somewhat limited sample. Yet, we would definitely like to be able to measure SFRs for the remaining 85 %, since *Herschel* and *Spitzer* detections only unveil about half of the star formation rate density of the Universe (see Sect. 2.4.5), or, equivalently, of the cosmic infrared background (CIRB, e.g., Leiton et al. 2015).

This is actually possible by interpreting in a clever way the observed UV spectrum of the galaxy (Calzetti et al. 2000). We do not know *a priori* how bright is the intrinsic spectrum, i.e., what we would see without dust, but we do have a good idea of the *shape* of this spectrum, in particular its *spectral slope* β . The spectral slope characterizes the way the light intensity varies with wavelength. A “gray” slope means that the light has the same intensity at all wavelengths, a “blue” slope indicates that the galaxy is brighter in the short wavelengths, while a “red” slope means the opposite.⁸ It turns out that the intrinsic spectral slope of a star-forming galaxy between 0.1 and 0.3 μm is fairly blue (see Fig. 1.5). Then, because the strength of the dust absorption depends on the wavelength, the light at the shortest wavelengths will be more attenuated than the light at the longest wavelengths. In the end, dust will tend to make the observed spectral slope redder (again, see Fig. 1.5). By measuring this observed slope, we can estimate how much dust is present in the line of sight, and recover the intrinsic spectrum. This is informally known as the “ β -slope” technique, and by opposition to the FIR-based SFRs introduced above, this method provides “UV-based” SFRs.

Measuring fluxes in the UV domain is much easier than in the FIR. This is especially true for distant galaxies, for which this emission is shifted by the cosmic Doppler effect into the optical domain, which is easily accessible from the ground. Thanks to this fact, we have access to the UV spectrum of essentially all the galaxies

⁷Or, equivalently, to use interferometry, which is very common in the radio domain.

⁸These names were not chosen randomly: these colors are those that our eyes would perceive with such spectra.

detected in cosmological deep fields, allowing us to derive star formation rates even for very faint galaxies. For this reason, UV-based SFR are very commonly used in the literature (e.g., Meurer et al. 1999; Steidel et al. 1999; Daddi et al. 2004a; González et al. 2010; Bouwens et al. 2011 to only cite a few of the most influential works).

The problem of this approach is that it requires quite a number of assumptions. At first sight, the most obvious one is the assumption about the spectral slope of the intrinsic spectrum. While it is true that all star-forming galaxies will have blue intrinsic spectral slopes (because their light is dominated by the young and massive stars), the precise value of the slope will depend on the star formation history of the galaxy: a recent burst will make the slope slightly bluer, while a declining star formation activity will make the slope slightly redder (Leitherer and Heckman 1995; see also Boquien et al. 2012). Other factors can have similar effects to some extent, like the stellar *metallicity* Z of the galaxy, which is the proportion of stellar baryons which are neither hydrogen nor helium (i.e., “metals”: oxygen, iron, ...). But probably more problematic are the assumptions about the dust. The transition from the observed to the intrinsic slope is made using an extinction curve, which tells us exactly how efficient is a parcel of dust at absorbing photons as a function of wavelength. Observations of our neighborhood, either within the Milky Way (Witt et al. 1984) or peering inside its satellites like the Small Magellanic Cloud (SMC, Prevot et al. 1984), have shown that this curve is not universal (see, e.g., Gordon et al. 2003). In particular, it is expected to depend on a combination of factors, among which are the distribution of dust grain sizes and their chemical properties. Finally, building the effective extinction curve over the whole galaxy requires assumptions on the geometry of the dust cloud, i.e., how is the dust spatially distributed with respect to the stars. For example, the usual assumption is that the stars are located behind a uniform dust “screen” of variable width.

Another very common and similar technique is to use a model to interpret simultaneously the whole spectrum from the UV to the near-infrared (NIR), including the stellar emission and the dust absorption, which is known as “spectral energy distribution (SED) fitting” (see Silva et al. 1998; da Cunha et al. 2008; Kriek et al. 2009; Noll et al. 2009 where some of the most commonly used codes are described). While it may help to remove some degeneracies, e.g., with respect to the star formation history or the metallicity, it essentially boils down to the same mechanism to estimate the star formation rates, and therefore requires the same set of assumptions.⁹

It turns out that, in spite of all these (sometimes crude) assumptions, the end result is on average in good agreement with the more direct estimates obtained from the FIR luminosity of local galaxies (Meurer et al. 1999; Calzetti et al. 2000), although some correction were later published (Takeuchi et al. 2012). This puzzling agreement probably shows that although each galaxy is unique in its detailed properties and structure, most of the differences are washed out when averaging quantities over the

⁹There are codes which actually interpret the photometry from the UV all the way up to the FIR in a consistent way, when FIR photometry is available. In this case, the mechanism to estimate the SFR is much closer to the FIR-based approach.

whole volume of the galaxy.¹⁰ *Most*, but not all. While these dust corrections appear to be working for the majority of Main Sequence galaxies (e.g., Daddi et al. 2007; Rodighiero et al. 2014), a number of studies have shown that these techniques tend to systematically underestimate the SFRs of the most actively star-forming galaxies (Goldader et al. 2002; Buat et al. 2005; Elbaz et al. 2007; Rodighiero et al. 2011; Wuyts et al. 2011; Penner et al. 2012; Oteo et al. 2013), i.e., the “starburst” galaxies I introduced in the previous section. In other words, when using only UV-based SFRs, one will miss a fraction of the star formation happening in the Universe. Rodighiero et al. (2011) have estimated this fraction to be of the order of 10 % when looking at the Universe as it was 10 Gyr ago, but we do not really know how this number evolves with time.

More worrisome, some recent studies have shown that the UV-based dust correction recipes described above seems to be failing *globally*, i.e., for all galaxies, in the very early Universe (more than 12 Gyr ago, Castellano et al. 2014; Pannella et al. 2015; Capak et al. 2015; Reddy et al. 2015). This illustrates the crucial fact that UV-based SFRs are only reliable in the regimes where they were demonstrated to work, i.e., by comparing them with FIR-based SFRs. This is mostly an issue for the first billion years of the history of the Universe, where too few galaxies are currently detected in the FIR.¹¹

For this reason, I have pushed the current *Herschel* surveys to their limit and provide in Chap. 2 (Sect. 2.4.2) a FIR-based measurement of the average SFR of the most distant sample of Main Sequence galaxies, to check the consistency of the published UV-based estimates. This is achieved using a special image analysis technique called “stacking” (e.g., Dole et al. 2006), which can only tell us about the statistical properties of a given sample, without individually detecting all of the galaxies it contains. The next step is to use ALMA, and in particular analyze our survey that was recently observed (Chap. 6), where we targeted the very same galaxies, this time with the aim to measure their individual star formation rates, reaching FIR luminosities an order of magnitude lower than what *Herschel* could detect. This will allow us to check the robustness of the result I present in this thesis, and further study the evolution of dust in the distant Universe.

Interestingly, these ALMA data contained by chance two new galaxies that are too far and too attenuated to be detected with *Hubble*. Their distance from us is uncertain, but most likely very large, and only ALMA can help us determine it with precision through spectroscopy. Should these galaxies be as distant as their (admittedly poor) SEDs suggest they are, they will be the most distant massive and dusty galaxies ever known, observed when the Universe was less than a billion years old, an epoch when dust is currently assumed to be mostly absent. These two galaxies are described in Chap. 6 (Sect. 6.9).

¹⁰See however Boquien et al. (2012) where this uniformity is found even in a resolved analysis.

¹¹Note that there are many other approaches to estimate SFRs without using the UV or the FIR light, e.g., hydrogen recombination lines, X-ray binaries, or radio emission from super-novae and star-forming regions. However, they are usually at least as expensive as the FIR-based SFRs observationally speaking, while suffering from more biases.

1.2.4 Why Do Galaxies Stop Forming Stars?

The existence of very massive galaxies that show little to no detectable star formation has been reported at nearly all epochs in the history of the Universe (e.g., Baldry et al. 2004; Daddi et al. 2004b; Williams et al. 2009; Straatman et al. 2014). These galaxies are therefore lying more than one order of magnitude below the Main Sequence of star-forming galaxies (Noeske et al. 2007). Although we know of several different ways of “quenching” a galaxy (i.e., making it stop forming stars), the exact physical mechanism that made these galaxies turn off is still uncertain. An example of such galaxy is shown in Fig. 1.7 (right).

The trend during the last decades has been to invoke “feedback” processes inside the galaxy. For example, when a galaxy hosts a very high density of SFR, it also harbors a large quantity of massive stars and exploding supernovae, which drive very strong *stellar winds* (Larson 1974): these stars are so massive and luminous that the light they emit is also pushing the surrounding gas away (*via* radiation pressure). By expelling and heating the hydrogen gas in which they were born, these massive stars can actually depress or even totally prevent future star formation in the area. This is called a feedback mechanism because star formation regulates itself without



Fig. 1.7 False color image of two galaxies (F606W+F814W+F160W, i.e., green, red, and near-infrared). On the *left*, J02172899-0508264, a star-forming disk galaxy in the Ultra Deep Survey (UDS) located about 190 Mpc from us ($z = 0.044$). On the *right*, J100022.0+022326, a quiescent elliptical galaxy in the Cosmic Evolution Survey (COSMOS) located about 1 000 Mpc from us ($z = 0.2206$). Both galaxies are shown here with the same color bar. While they have roughly the same angular size, the elliptical galaxy is much more distant, and would be five times larger if it was brought back at the same distance. On the other hand, one can see from this picture that a large fraction of the light (hence, of the mass) of this elliptical galaxy is located very close to its center, while the light of the disk galaxy is more evenly spread. This difference is quantified in Fig. 1.8. Figuring out how this morphological transformation takes place is also a key to understanding the mechanisms that quench a galaxy

requiring external influence. One can refer to, e.g., Hopkins et al. (2014) which describes the current state of the art in numerical simulations. However, Hopkins et al. (2014) also argue that this stellar feedback is not sufficient to prevent a *whole* galaxy from forming stars (except maybe for dwarfs, Dekel and Silk 1986). In fact, it is more commonly assumed today that this feedback mechanism only acts as a *regulator* of star formation, preventing it from being too efficient.

The hypothesis that is most commonly put forward nowadays is that the process responsible for quenching is yet another form of feedback, this time originating from supermassive black holes (SMBHs, Silk and Rees 1998). Although the existence of such black holes have only been unambiguously confirmed in the Milky Way (e.g., Gillessen et al. 2009) and a couple of other nearby galaxies, a number of indirect evidence suggest that most massive galaxies host an SMBH in their core (e.g., Hickox et al. 2014). Because black holes, by definition, do not emit any light, the only way to detect them is through the effect they have on their surroundings, for example gravitational attraction (as was done in Gillessen et al. 2009). Through these interactions, black holes can actually be the cause of extremely luminous events, and the region around them sometimes becomes brighter than the combination of all the stars present in the galaxy. Galaxies in which these events are occurring are said to harbor an active galactic nucleus (AGN). When the luminosity of the AGN becomes very large, by contrast the galaxy is almost invisible and all we can observe is a point source, looking very much like a star. For this reason, the most extreme cases are called quasi-stellar objects (QSOs, or quasars). The physical process behind these luminous events is thought to be the accretion of gas and/or stars onto the black hole (Lynden-Bell 1969). The closer this material orbits around the black hole, the stronger the friction force it feels. Right before crossing the horizon, this friction is so intense that a large fraction of the accreted mass is actually turned into thermal energy, which is then radiated away in the form of light on a large variety of wavelengths from the UV to the IR. Through this so-called “radiative-mode”, AGNs behaves very much like huge stars, in the sense that the radiation they emit can also drive away large amounts of gas (see the review of Cattaneo et al. 2009). This radiation is so intense that it should be able to channel the gas outside of the galaxy, and in fact there are recent evidence that this is indeed happening in some galaxies (Förster Schreiber et al. 2014). By expelling the gas away, the AGN is preventing further accretion onto the black hole, which is why this phenomenon is also categorized as a feedback process. But then, when a large fraction of the gas has been expelled, the accretion of matter onto the SMBH decreases, and the AGN switches off. At that point, it is only a matter of time for the expelled gas to cool down, fall back onto the galaxy, and form stars again (e.g., Gabor et al. 2011). To permanently prevent star formation in the whole galaxy, this gas and all future infall must be prevented from cooling down, and another process has to be invoked.

This is supposedly the role of a second feedback mechanism associated with black holes, which is usually called the “radio-mode” (again, see Cattaneo et al. 2009). This time, the expelled gas is concentrated in a pair of collimated *jets* escaping from both poles of the black hole (Blandford and Begelman 1999). These jets are typically observed in the radio domain, hence the name of this mechanism, but they also

emit in the X-ray and IR. Because they are mostly seen in those galaxies which are already not forming any star (also called “quiescent” galaxies, Dressel 1981), they are indeed good candidates for maintaining the hydrogen gas at high temperatures, thereby preventing further star formation (Nusser et al. 2006). The physical origin of the jet is not very well understood, and it is still not clear whether it originates from the black hole itself or from its accretion disk. Supposedly, jets appear in black holes with lower accretion rates.

Combined together, both these mechanisms have the necessary power to abruptly and permanently stop star formation in a previously gas-rich galaxy. However, these explanations currently lack direct observational support. It is true that “radiative-mode” and “radio-mode” AGNs appear to be more common among massive star-forming and quiescent galaxies, respectively, which makes them good candidates for shutting down star formation and maintaining it low, respectively. But we cannot rule out that the causality link goes on the other direction, i.e., that AGNs are merely *consequences* of whatever phenomenon acts to quench the galaxies, and that in spite of their strength they are relatively inefficient in affecting the galaxy. For example, using numerical simulations, Roos et al. (2015) recently predicted that “radiative-mode” AGNs should have no significant impact on the short-term star formation rates of gas-rich distant galaxies, where the disk is dense enough to shield itself from the radiation of the AGN (see also Gabor and Bournaud 2014).

In fact, there are other ways to prevent a galaxy from forming stars without invoking strong feedback mechanisms. The difficulty is to make galaxies stop at the right moment, before they grow too massive (indeed, we do not observe galaxies more massive than $10^{12} M_{\odot}$, see for example Naab et al. 2007 for a scenario where too massive galaxies are produced). One can cite for example the so-called “morphological quenching” (Martig et al. 2009), which is connected to the presence of a massive stellar bulge. A bulge is a large concentration of old stars which cohabit with the disk of most massive galaxies, including our Milky Way. Their radial density profiles are very similar to that of quiescent elliptical galaxies, i.e., they show a pronounced overdensity in their core, together with a low density tail that extend much further than the disk (the de Vaucouleurs profile, de Vaucouleurs 1948), see Fig. 1.8. The idea behind the morphological quenching mechanism is that the large stellar mass in the center of the bulge generates additional turbulence that, in virtue of the Toomre criterion (Toomre 1964), should make it harder for the gas to fragment into clumps within the disk (see, e.g., Lehnert et al. 2015), therefore inhibiting star formation without actually expelling the gas out of the galaxy. This is supported by the recent finding that star formation seems to stop gradually from inside out (Tacchella et al. 2015), probably concurrent to the growth of the bulge with respect to the disk. It is also interesting to note that a tight correlation has been observed between the mass of the bulge and that of the SMBH¹² (Magorrian et al. 1998; Häring and Rix 2004). While the origin of this correlation is still unknown, it could explain why

¹²Another correlation, somewhat related, was also reported between the SMBH mass and the velocity dispersion of the stars inside the bulge (Ferrarese and Merritt 2000; Gebhardt et al. 2000).

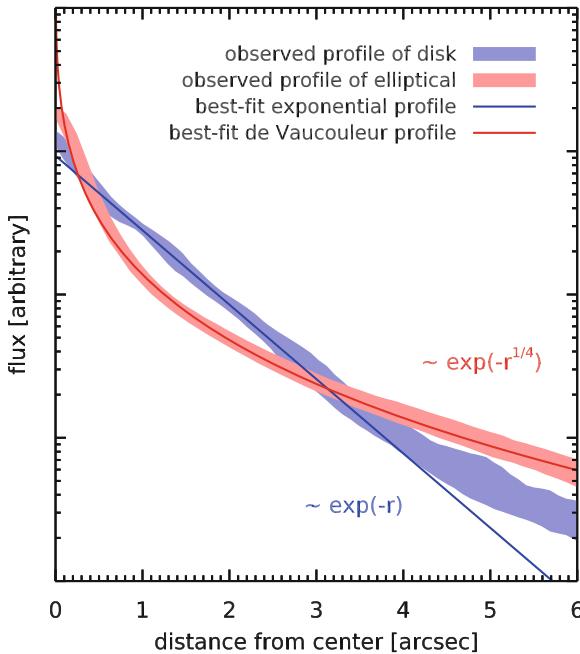


Fig. 1.8 Radial light profile of the two galaxies of Fig. 1.7. The observed light profile of the disk and the elliptical galaxies are shown in shades of *blue* and *red*, respectively. These profiles were obtained by masking pixels contaminated by other galaxies, and are derived assuming that each galaxy's center is the brightest pixel of the image. These profiles are then fitted with idealized functions, ignoring for simplicity the smearing caused by the *HST* point spread function, and assuming that both galaxies are seen face-on. The disk galaxy is fitted with an exponential profile (Freeman 1970), while the elliptical is fitted with a de Vaucouleur profile (de Vaucouleurs 1948). It is clear that the latter has a much denser core, and it is in fact similar to that of *bulges*, which are sometimes found within disk galaxies. These high concentration of stellar mass may have a significant impact on the dynamics of the gas, and therefore on star formation

AGNs seem to be connected with quenching, since according to the morphological quenching mechanism quiescent galaxies are likely to host the most massive bulges, and therefore the most massive SMBHs.

At present, it is probably fair to say that the role played by AGNs and supermassive black holes in the growth of their host galaxy is still relatively unknown. During the past years, evidence has gathered for a *coevolution* of SMBH and galaxies, i.e., that whatever fuels and then regulates star formation in a galaxy also affects the growth of the central SMBH (see, e.g., the review of Brandt and Alexander 2015). The clearest evidence probably came from the observed correlations between the mass of the SMBH and the properties (either mass or velocity dispersion) of the bulge (see above), but additional clues comes from observations in the distant Universe, where an intriguing coincidence is observed between the activity of black hole accretion and the star formation rate, either in individual galaxies (e.g., Alexander et al. 2005)

or as a whole within a given volume of the Universe (e.g., Aird et al. 2010; Ueda et al. 2014). Similarly, it was also found that AGNs follow their own “Main Sequence” in the L_X – M_* plane (where L_X measures the rate of accretion onto the black hole), mimicking closely the trend of the SFR– M_* Main Sequence (Mullaney et al. 2012). This connection between black hole accretion and star formation suggests that a fraction of the gas that fuels the galaxy manages to be channeled toward its center and fuel, in turn, the black hole. However, the sphere of influence of the SMBH is ridiculously small (e.g., a couple of parsecs in our Milky Way, Ferrarese and Merritt 2000) compared to the total size of a galaxy, and it is therefore a challenge to understand how such channeling could occur.¹³ Lastly, it is sometimes advocated that the link between activity of the SMBH and star formation in galaxies should be approached in the opposite direction, in the sense that it would be the black hole that triggers the formation of stars in its host galaxy (e.g., Croft et al. 2006; Elbaz et al. 2009; Silk 2013). Although this phenomenon appears to be rare in the present-day Universe, it may have played a more important role in the earliest stage of galaxy evolution.

The last mechanism I will describe is called “halo quenching”, or “gravitational quenching” (Birnboim and Dekel 2003; Dekel and Birnboim 2008). This approach considers that quenching is due to the accretion of gas from the intergalactic medium: when gas is channeled toward the galaxy, it cools down and form stars, while the energy that is released by the cooling process (whatever it is) is progressively transferred to the hydrogen gas that still resides in the dark matter halo. By receiving this energy, this gas progressively heats up, and it becomes harder for it to later reach the galaxy and form stars. It is expected that this quenching mechanism can only affect galaxies that receive enough accretion so that the gas in the halo heats faster than it cools down, and therefore, according to the Λ CDM cosmology, it should only happen in galaxies that live in dark matter halos typically more massive than $10^{12} M_\odot$ (corresponding to galaxies more massive than $10^{10} M_\odot$, according to Behroozi et al. 2013). This would naturally explain why most of the quiescent galaxies are among the most massive structures we observe today.

There are, indeed, many processes that can act to prevent star formation in a galaxy. The question is then to find out which of the above-listed mechanisms is preponderant. To this end, we can use the same statistical tools that I introduced in Sect. 1.2.1, namely the SFR– M_* correlation, or Main Sequence of star-forming galaxies. While this correlation is most often studied with star-forming galaxies only, the position of the quiescent galaxies with respect to the Main Sequence is also of interest. As stated above, quiescent galaxies are among the most massive galaxies in the Universe, yet they form little to no stars. Therefore, they have very small sSFR (which, I remind, is defined as SFR/M_*). The distribution of sSFR of massive galaxies, both star-forming and quiescent, can then be used to put constraints on the

¹³In fact, an interesting parallel can be drawn with dark matter halos, compared to which galaxies are themselves relatively small. Here also, is it not trivial to transport gas accreted from the intergalactic medium from the halo down to the galaxy.

quenching mechanisms, by providing information on how galaxies evolve from the sSFRs typical of the Main Sequence down to the low levels of the quiescent galaxies.

In particular, it is observed that the two populations are clearly segregated in sSFR, so that the overall distribution is strongly bimodal. Initially, it was first discovered through galaxy colors, which are simple proxies of the sSFR if attenuation by dust is neglected (e.g., Strateva et al. 2001; Baldry et al. 2004). For this reason, it is often said that star-forming and quiescent galaxies lie on the “blue” and “red cloud”, respectively (see the discussion in the previous section about the spectral slope). One can see a recent visualization of this phenomenon in Renzini and Peng (2015). The presence of this bimodality has often been used to argue that the quenching mechanism must act on short timescales: indeed, if quenching is a slow process, we should see a large number of galaxies transiting from the blue cloud down to the red cloud (traversing the so-called “green valley”). This is why, in particular, AGN quenching is a popular scenario, because we expect it to be the fastest quenching mode.

On the other hand, a recent discovery may counteract this argument. Up until now, most studies which quantified the position and evolution of the Main Sequence have reported that this correlation had a fixed *slope* in logarithmic space. Many different values of this slope were reported, owing probably to a combination of selection effects and observational uncertainties (e.g., on the dust corrections for UV-based SFRs, see for example the recent compilation of Speagle et al. 2014). Interestingly, as data are becoming more accurate and statistic is building up, a couple of recent studies actually showed that modeling the sequence with a single slope may not be appropriate, and that a more accurate description would be to consider a *varying slope* at different stellar masses (e.g., Whitaker et al. 2012, 2014; Magnelli et al. 2014). I also revisit this observation in Chap. 2 (Sect. 2.4.3).

Several explanations have been put forward to account for this observation. In particular, it is possible to make a link with the quenching process. Indeed, the evolution of the slope is such that massive star-forming galaxies tend to have lower sSFRs than low mass galaxies, and this discrepancy is growing larger with time. This is strangely mimicking the evolution of the quiescent galaxies, which are predominantly found among massive galaxies, and grow more numerous as time goes. This could be understood as the fact that the *whole* population of massive star-forming galaxies is progressively dying on long timescales (see, e.g., the discussion in Ilbert et al. 2015), which would instead favor mechanisms which are not linked to the AGN.

Alternatively, Abramson et al. (2014) have shown in the local Universe that the Main Sequence slope of massive galaxies could be brought back to that of low-mass galaxies by subtracting the mass of the bulge from M_* , i.e., not considering the relation between SFR and M_* , but between SFR and M_{disk} (the mass of the *disk*). The underlying argument is that no star formation happens in the bulge, and it is therefore not relevant to compare its stellar mass to the present rate of star formation. To further investigate this issue, I review in Chap. 5 the observation of Abramson et al. (2014), applying their methodology to the distant Universe, where the slope of the Main Sequence is found to be different. The aim is to check if their result also holds at early epochs, and draw further conclusions on the link with quenching.

In Chap. 5 I also approach the study of this varying slope from another angle, by measuring directly the gas content in these galaxies to see whether the decrease of the SFR is caused by a reduction of the *gas fraction* f_{gas} , or a lowered *star formation efficiency* (SFE). The SFE characterizes the speed at which the galaxy will exhaust all of its gas content, and it is actually defined as the inverse of the depletion timescale I introduced earlier. These represent the two possible ways for galaxies to turn off: indeed, a galaxy can become quiescent either if it ran out of gas (low f_{gas}), or if it became extremely inefficient at converting this gas into stars (low SFE). Knowing which of these two alternative is true would shed some light on the quenching mechanisms.

References

- L.E. Abramson, D.D. Kelson, A. Dressler et al., ApJ **785**, L36 (2014)
 J. Aird, K. Nandra, E.S. Laird et al., MNRAS **401**, 2531 (2010)
 D.M. Alexander, I. Smail, F.E. Bauer et al., Nature **434**, 738 (2005)
 L. Armus, T. Heckman, G. Miley, AJ **94**, 831 (1987)
 I.K. Baldry, K. Glazebrook, J. Brinkmann et al., ApJ **600**, 681 (2004)
 P.S. Behroozi, R.H. Wechsler, C. Conroy, ApJ **770**, 57 (2013)
 Y. Birnboim, A. Dekel, MNRAS **345**, 349 (2003)
 R.D. Blandford, M.C. Begelman, MNRAS **303**, L1 (1999)
 G.R. Blumenthal, S.M. Faber, J.R. Primack, M.J. Rees, Nature **311**, 517 (1984)
 M. Boquien, V. Buat, A. Boselli et al., A&A **539**, A145 (2012)
 R.J. Bouwens, G.D. Illingworth, P.A. Oesch et al., ApJ **737**, 90 (2011)
 W.N. Brandt, D.M. Alexander, A&A Rev. **23**, 1 (2015)
 A. Bressan, F. Fagotto, G. Bertelli, C. Chiosi, A&AS **100**, 647 (1993)
 J. Brinchmann, S. Charlot, S.D.M. White et al., MNRAS **351**, 1151 (2004)
 G. Bruzual, S. Charlot, MNRAS **344**, 1000 (2003)
 V. Buat, J. Iglesias-Páramo, M. Seibert et al., ApJ **619**, L51 (2005)
 D. Calzetti, L. Armus, R.C. Bohlin et al., ApJ **533**, 682 (2000)
 P.L. Capak, C. Carilli, G. Jones et al., Nature **522**, 455 (2015)
 M. Castellano, V. Sommariva, A. Fontana et al., A&A **566**, 19 (2014)
 A. Cattaneo, S.M. Faber, J. Binney et al., Nature **460**, 213 (2009)
 M. Cignoni, S. Degl'Innocenti, P.G. Prada Moroni, S.N. Shore, A&A **459**, 783 (2006)
 S. Croft, W. van Breugel, W. de Vries et al., ApJ **647**, 1040 (2006)
 E. da Cunha, S. Charlot, D. Elbaz, MNRAS **388**, 1595 (2008)
 E. Daddi, F. Bournaud, F. Walter et al., ApJ **713**, 686 (2010)
 E. Daddi, A. Cimatti, A. Renzini et al., ApJ **617**, 746 (2004a)
 E. Daddi, A. Cimatti, A. Renzini et al., ApJ **600**, L127 (2004b)
 E. Daddi, M. Dickinson, G. Morrison et al., ApJ **670**, 156 (2007)
 G. de Vaucouleurs, Annales d'Astrophysique **11**, 247 (1948)
 A. Dekel, Y. Birnboim, MNRAS **383**, 119 (2008)
 A. Dekel, Y. Birnboim, G. Engel et al., Nature **457**, 451 (2009)
 A. Dekel, J. Silk, ApJ **303**, 39 (1986)
 R. Diehl, H. Halloin, K. Kretschmer et al., Nature **439**, 45 (2006)
 H. Dole, G. Lagache, J. Puget et al., A&A **451**, 417 (2006)
 L.L. Dressel, ApJ **245**, 25 (1981)
 A. Dressler, A. Oemler, M.G. Gladders et al., ApJ **699**, L130 (2009)
 D. Elbaz, E. Daddi, D. Le Borgne et al., A&A **468**, 33 (2007)

- D. Elbaz, K. Jahnke, E. Pantin, D. Le Borgne, G. Letawe, A&A **507**, 1359 (2009)
 D. Elbaz, M. Dickinson, H.S. Hwang et al., A&A **533**, 119 (2011)
 L. Ferrarese, D. Merritt, ApJ **539**, L9 (2000)
 C. Flynn, J. Holmberg, L. Portinari, B. Fuchs, H. Jahreiß, MNRAS **372**, 1149 (2006)
 N.M. Förster Schreiber, R. Genzel, S.F. Newman et al., ApJ **787**, 38 (2014)
 A. Franceschini, H. Aussel, C.J. Cesarsky, D. Elbaz, D. Fadda, A&A **378**, 1 (2001)
 K.C. Freeman, ApJ **160**, 811 (1970)
 J.M. Gabor, F. Bournaud, MNRAS **441**, 1615 (2014)
 J.M. Gabor, R. Davé, B.D. Oppenheimer, K. Finlator, MNRAS **417**, 2676 (2011)
 F. Galliano, S. Hony, J. Bernard et al., A&A **536**, A88 (2011)
 K. Gebhardt, R. Bender, G. Bower et al., ApJ **539**, L13 (2000)
 R. Genzel, L.J. Tacconi, J. Gracia-Carpio et al., MNRAS **407**, 2091 (2010)
 S. Gillessen, F. Eisenhauer, S. Trippe et al., ApJ **692**, 1075 (2009)
 J.D. Goldader, G. Meurer, T.M. Heckman et al., ApJ **568**, 651 (2002)
 V. González, I. Labb  , R.J. Bouwens et al., ApJ **713**, 115 (2010)
 K.D. Gordon, G.C. Clayton, K.A. Misselt, A.U. Landolt, M.J. Wolff, ApJ **594**, 279 (2003)
 N. H  ring, H. Rix, ApJ **604**, L89 (2004)
 T.M. Heckman, L. Armus, G.K. Miley, ApJS **74**, 833 (1990)
 X. Hernandez, D. Valls-Gabaud, G. Gilmore, MNRAS **316**, 605 (2000)
 R.C. Hickox, J.R. Mullaney, D.M. Alexander et al., ApJ **782**, 9 (2014)
 P.F. Hopkins, K. Bundy, D. Croton et al., ApJ **715**, 202 (2010)
 P.F. Hopkins, D. Kere  , J. O  norbe et al., MNRAS **445**, 581 (2014)
 O. Ilbert, S. Arnouts, E. Le Floc'h et al., A&A **579**, A2 (2015)
 J.S. Kartaltepe, D.B. Sanders, N.Z. Scoville et al., ApJS **172**, 320 (2007)
 R.C. Kennicutt Jr., ApJ **272**, 54 (1983)
 R.C. Kennicutt Jr., ARA&A **36**, 189 (1998)
 M. Kriek, P.G. van Dokkum, I. Labb  , et al., ApJ **700**, 221 (2009)
 I. Labb  , G. Rudnick, M. Franx et al., ApJ **591**, L95 (2003)
 R.B. Larson, MNRAS **169**, 229 (1974)
 M.D. Lehnert, W. van Driel, L. Le Tiran, P. Di Matteo, M. Haywood, A&A **577**, A112 (2015)
 C. Leitherer, T.M. Heckman, ApJS **96**, 9 (1995)
 R. Leiton, D. Elbaz, K. Okumura et al., A&A **579**, A93 (2015)
 S.J. Lilly, O. Le Fevre, F. Hammer, D. Crampton, ApJ **460**, L1 (1996)
 P.O. Lindblad, Stockholms Observatoriums Annaler **21**, 4 (1960)
 J.M. Lotz, P. Jonsson, T.J. Cox et al., ApJ **742**, 103 (2011)
 D. Lynden-Bell, Nature **223**, 690 (1969)
 P. Madau, H.C. Ferguson, M.E. Dickinson et al., MNRAS **283**, 1388 (1996)
 G.E. Magdis, E. Daddi, M. B  thermin et al., ApJ **760**, 6 (2012)
 B. Magnelli, D. Elbaz, R.R. Chary et al., A&A **496**, 57 (2009)
 B. Magnelli, D. Lutz, A. Saintonge et al., A&A **561**, 86 (2014)
 J. Magorrian, S. Tremaine, D. Richstone et al., AJ **115**, 2285 (1998)
 M. Martig, F. Bournaud, R. Teyssier, A. Dekel, ApJ **707**, 250 (2009)
 C.F. McKee, E.C. Ostriker, ARA&A **45**, 565 (2007)
 G.R. Meurer, T.M. Heckman, D. Calzetti, ApJ **521**, 64 (1999)
 M. Milgrom, ApJ **270**, 365 (1983)
 J.R. Mullaney, M. Pannella, E. Daddi et al., MNRAS **419**, 95 (2012)
 T. Naab, P.H. Johansson, J.P. Ostriker, G. Efstathiou, ApJ **658**, 710 (2007)
 K.G. Noeske, B.J. Weiner, S.M. Faber et al., ApJ **660**, L43 (2007)
 S. Noll, D. Burgarella, E. Giovannoli et al., A&A **507**, 1793 (2009)
 A. Nusser, J. Silk, A. Babul, MNRAS **373**, 739 (2006)
 I. Oteo, J. Cepa, A. Bongiovanni et al., A&A **554**, L3 (2013)
 M. Pannella, D. Elbaz, E. Daddi et al., ApJ **807**, 141 (2015)
 J.A. Peacock, S. Cole, P. Norberg et al., Nature **410**, 169 (2001)

- P.J.E. Peebles, ApJ **263**, L1 (1982)
- K. Penner, M. Dickinson, A. Pope et al., ApJ **759**, 28 (2012)
- J. Pflamm-Altenburg, P. Kroupa, ApJ **706**, 516 (2009)
- Planck Collaboration, P.A.R. Ade, N. Aghanim et al., A&A **571**, A16 (2014)
- W.H. Press, P. Schechter, ApJ **187**, 425 (1974)
- M.L. Prevot, J. Lequeux, L. Prevot, E. Maurice, B. Rocca-Volmerange, A&A **132**, 389 (1984)
- N.A. Reddy, M. Kriek, A.E. Shapley et al., ApJ **806**, 259 (2015)
- F. Renaud, F. Bournaud, E. Emsellem et al., MNRAS **436**, 1836 (2013)
- F. Renaud, F. Bournaud, K. Kraljic, P. Duc, MNRAS **442**, L33 (2014)
- F. Renaud, F. Bournaud, P. Duc, MNRAS **446**, 2038 (2015)
- A. Renzini, Y.-J. Peng, ApJ **801**, L29 (2015)
- G. Rodighiero, E. Daddi, I. Baronchelli et al., ApJ **739**, L40 (2011)
- G. Rodighiero, A. Renzini, E. Daddi et al., MNRAS **443**, 19 (2014)
- O. Roos, S. Juneau, F. Bournaud, J.M. Gabor, ApJ **800**, 19 (2015)
- A. Saintonge, G. Kauffmann, J. Wang et al., MNRAS **415**, 61 (2011)
- E.E. Salpeter, ApJ **121**, 161 (1955)
- D.B. Sanders, I.F. Mirabel, ARA&A **34**, 749 (1996)
- D.B. Sanders, B.T. Soifer, J.H. Elias et al., ApJ **325**, 74 (1988)
- M.T. Sargent, M. Béthermin, E. Daddi, D. Elbaz, ApJ **747**, L31 (2012)
- M. Schmidt, ApJ **129**, 243 (1959)
- J. Silk, ApJ **772**, 112 (2013)
- J. Silk, M.J. Rees, A&A **331**, L1 (1998)
- L. Silva, G.L. Granato, A. Bressan, L. Danese, ApJ **509**, 103 (1998)
- J.S. Speagle, C.L. Steinhardt, P.L. Capak, J.D. Silverman, ApJS **214**, 15 (2014)
- C.C. Steidel, K.L. Adelberger, M. Giavalisco, M. Dickinson, M. Pettini, ApJ **519**, 1 (1999)
- C.M.S. Straatman, I. Labb  , L.R. Spitler et al., ApJ **783**, L14 (2014)
- I. Strateva, Ivezi  , v., Knapp, G.R. et al., AJ **122**, 1861 (2001)
- S. Tacchella, C.M. Carollo, A. Renzini et al., Science **348**, 314 (2015)
- L.J. Tacconi, R. Genzel, R. Neri et al., Nature **463**, 781 (2010)
- T.T. Takeuchi, F. Yuan, A. Ikeyama, K.L. Murata, A.K. Inoue, ApJ **755**, 144 (2012)
- A. Toomre, ApJ **139**, 1217 (1964)
- Y. Ueda, M. Akiyama, G. Hasinger, T. Miyaji, M.G. Watson, ApJ **786**, 104 (2014)
- S. van den Bergh, Zeitschrift fur Astrophysik **43**, 236 (1957)
- S. van den Bergh, A&A Rev. **9**, 273 (1999)
- P.G. van Dokkum, K.E. Whitaker, G. Brammer et al., ApJ **709**, 1018 (2010)
- K.E. Whitaker, P.G. van Dokkum, G. Brammer, M. Franx, ApJ **754**, L29 (2012)
- K.E. Whitaker, M. Franx, J. Leja et al., ApJ **795**, 104 (2014)
- R.J. Williams, R.F. Quadri, M. Franx, P. van Dokkum, I. Labb  , ApJ **691**, 1879 (2009)
- A.N. Witt, R.C. Bohlin, T.P. Stecher, ApJ **279**, 698 (1984)
- S. Wuyts, N.M. F  rster Schreiber, D. Lutz et al., ApJ **738**, 106 (2011)
- R.F.G. Wyse, in *Proceedings of the IAU Symposium*, vol. 258 (2009), pp. 11–22. arXiv:0907.0415
- V. Zubko, E. Dwek, R.G. Arendt, ApJS **152**, 211 (2004)

Chapter 2

The Main Sequence of Star-Forming Galaxies as Seen by Herschel

In this chapter I present an analysis of the deepest *Herschel* images in four major extragalactic fields: GOODS–North, GOODS–South, UDS, and COSMOS. These images were obtained within the GOODS–*Herschel*, CANDELS–*Herschel* and PEP programs, and provide far-infrared detections for a total of 10 497 individual galaxies. I supplement this analysis by stacking these images using a mass-complete sample of 62 361 star-forming galaxies selected from the *Hubble Space Telescope (HST)* H band-selected catalogs of the CANDELS survey, and from two deep ground-based K_s band-selected catalogs in the GOODS–*North* and COSMOS fields. The goal is to obtain one of the most accurate and unbiased understanding to date of the stellar mass growth over the cosmic history.

One can obtain a robust and complete estimation of the SFR of a galaxy by combining the direct UV luminosity obtained with, e.g., *HST*, and the far-infrared reprocessed light provided by *Herschel*. I take advantage of this approach in this work to confirm that galaxies from $z = 4$ to $z = 0$ and of all stellar masses (M_*) follow a universal scaling law, the so-called Main Sequence of star-forming galaxies. I find a universal, close-to-linear slope of the $\log_{10}(\text{SFR})$ – $\log_{10}(M_*)$ relation, with evidence for a flattening at high masses ($\log_{10}(M_*/M_\odot) > 10.5$) that becomes less prominent with increasing redshift and almost vanishes by $z \simeq 2$, where the slope becomes compatible with one. I investigate the origin of this change of slope in Chap. 5. The specific SFR ($\text{sSFR} = \text{SFR}/M_*$) of star-forming galaxies is found to decrease continuously from $z = 4$ to 0.

I introduce in this chapter a new method called “scatter stacking” and show, for the first time, that stacking also provides a powerful tool to determine the dispersion of a physical correlation. Using this tool, I measure within the Main Sequence

This chapter is essentially a reprint of a published article (Schreiber et al. 2015). Credit: C. Schreiber, A&A, 575, A74, 2015, reproduced with permission © ESO.

a nonvarying SFR dispersion of 0.3 dex: at a fixed redshift and stellar mass, about 68 % of star-forming galaxies form stars at a universal rate within a factor of two.

Finally, I discuss the implications of our findings on the cosmic SFR history and on the origin of present-day stars. Combining all these results, one can show that more than two-thirds of present-day stars must have formed in a regime dominated by the “Main Sequence” mode. As a consequence I conclude that, although omnipresent in the distant Universe, galaxy mergers had little impact in shaping the global star formation history over the last 12.5 billion years.

2.1 Introduction

Most extremely star-forming galaxies in the local Universe are heavily dust obscured and show undeniable signs of an ongoing major merger, however such objects are relatively rare (Armus et al. 1987; Sanders and Mirabel 1996). They have been historically classified as Luminous and Ultra Luminous InfraRed Galaxies, LIRGs and ULIRGs, based on their bolometric infrared luminosity over the wavelength range 8–1000 μm , by $L_{\text{IR}} > 10^{11} L_{\odot}$ and $> 10^{12} L_{\odot}$, respectively. However, they make up for only 2 % of the integral of the local IR luminosity function, the remaining fraction mainly produced by more typical isolated galaxies (Sanders and Mirabel 1996).

More recently, studies at higher redshift showed that LIRGs were the dominant population at $z = 1$ (Chary and Elbaz 2001; Franceschini et al. 2001; Le Floc’h et al. 2005), replaced by ULIRGs at $z = 2$ (Magnelli et al. 2013). This was first interpreted as an increasing contribution of gas-rich galaxy mergers to the global star formation activity of the Universe, in qualitative agreement with the predicted and observed increase of the major merger rate (e.g., Patton et al. 1997; Le Fèvre et al. 2000; Conselice 2003).

The discovery of the correlation between star formation rate (SFR) and stellar mass (M_*), also called the “Main Sequence” of star-forming galaxies (Noeske et al. 2007), at $z \simeq 0$ (Brinchmann et al. 2004), $z \simeq 1$ (Noeske et al. 2007; Elbaz et al. 2007), $z \simeq 2$ (Daddi et al. 2007; Pannella et al. 2009; Rodighiero et al. 2011; Whitaker et al. 2012) $z = 3\text{--}4$ (Daddi et al. 2009; Magdis et al. 2010; Heinis et al. 2013; Pannella et al. 2015) and even up to $z = 7$ (e.g., Stark et al. 2009; Bouwens et al. 2012; Stark et al. 2013; González et al. 2014; Steinhardt et al. 2014; Salmon et al. 2015) suggested instead a radically new paradigm. The tightness of this correlation is indeed not consistent with the idea that most stars are formed in frequent random bursts induced by processes like major mergers of gas-rich galaxies, and favors more stable star formation histories (Noeske et al. 2007).

Furthermore, systematic studies of the dust properties of the “average galaxy” at different redshifts show that LIRGs at $z = 1$ and ULIRGs at $z = 2$ bear close resemblance to normal star-forming galaxies at $z = 0$. In particular, in spite of having star formation rates (SFRs) higher by orders of magnitude, they appear to share similar star-forming region sizes (Rujopakarn et al. 2011), polycyclic aromatic hydrocarbon

(PAH) emission lines equivalent widths (Pope et al. 2008; Fadda et al. 2010; Elbaz et al. 2011; Nordon et al. 2012), [C II] to far-infrared (FIR) luminosity (L_{FIR}) ratios (Díaz-Santos et al. 2013), and universal FIR spectral energy distributions (SEDs) (Elbaz et al. 2011). Only outliers above the SFR– M_* correlation (usually called “starbursts”, Elbaz et al. 2011) show signs of different dust properties: more compact geometry (Rujopakarn et al. 2011), excess of $\text{IR8} \equiv L_{\text{IR}}/L_{8\,\mu\text{m}}$ (Elbaz et al. 2011), [C II] deficit (Díaz-Santos et al. 2013), increased effective dust temperature (Elbaz et al. 2011; Magnelli et al. 2014), and PAH deficit (Nordon et al. 2012; Murata et al. 2014), indicating that these starburst galaxies are the true analogs of local LIRGs and ULIRGs. In this paradigm, the properties of galaxies are no longer most closely related to their rest-frame bolometric luminosities, but rather to their excess SFR compared to that of the Main Sequence.

This could mean that starburst galaxies are actually triggered by major mergers, but that the precise mechanism that fuels the remaining vast majority of “normal” galaxies is not yet understood. Measurements of galactic gas reservoirs yield gas fractions evolving from about 10 % in the local Universe (Leroy et al. 2008) up to 60 % at $z \simeq 3$ (Tacconi et al. 2010; Daddi et al. 2010; Geach et al. 2011; Magdis et al. 2012; Saintonge et al. 2013; Santini et al. 2014; Genzel et al. 2015, Béthermin et al. 2014, submitted). Compared to the observed SFR, this implies gas-consumption timescales that are much shorter than the typical duty cycle of most galaxies. It is thus necessary to replenish the gas reservoirs of these galaxies in some way. Large volume numerical simulations (Dekel et al. 2009a) have shown that streams of cold gas from the intergalactic medium can fulfill this role, allowing galaxies to keep forming stars at these high but steady rates. Since the amount of gas accreted through these “cold flows” is directly linked to the matter density of the intergalactic medium, this also provides a qualitative explanation for the gradual decline of the SFR from $z = 3$ to the present day (e.g., Davé et al. 2011).

This whole picture relies on the existence of the Main Sequence. However, actual observations of the SFR– M_* correlation at $z > 2$ rely mostly on ultraviolet-derived star formation rates, which need to be corrected by large factors to account for dust extinction (Calzetti et al. 1994; Madau et al. 1998; Meurer et al. 1999; Steidel et al. 1999). These corrections, performed using the ultraviolet (UV) continuum slope β and assuming an extinction law, are uncertain and still debated. Although dust-corrected SFRs are able to match more robust estimators on average in the local Universe (Calzetti et al. 1994; Meurer et al. 1999) and beyond (e.g., Pannella et al. 2009; Overzier et al. 2011; Heinis et al. 2013; Rodighiero et al. 2014), it has been shown for example that these corrections cannot recover the full star formation rate of the most active objects (Goldader et al. 2002; Buat et al. 2005; Elbaz et al. 2007; Rodighiero et al. 2011; Wuyts et al. 2011; Penner et al. 2012; Oteo et al. 2013). More recently, several studies have pointed toward an evolution of the calibration between the UV slope and UV attenuation as a function of redshift, possibly due to changes in the interstellar medium (ISM) properties (e.g., Castellano et al. 2014; Pannella et al. 2015) or even as a function of environment (Koyama et al. 2013). It is therefore possible that using UV-based SFR estimates modifies the normalization of the Main Sequence, and/or its dispersion. In particular, it could be that the tight scatter of

the Main Sequence observed at high redshift (e.g., Bouwens et al. 2012; Salmon et al. 2015) is not real but induced by the use of such SFRs, thereby questioning the very existence of a Main Sequence at these epochs. Indeed, a small scatter is a key ingredient without which the Main Sequence loses its meaning.

Infrared telescopes allow us to measure the bolometric infrared luminosity of a galaxy (L_{IR}), a robust star formation tracer (Kennicutt 1998). Unfortunately, they typically provide observations of substantially poorer quality (both in angular resolution and typical depth) compared to optical surveys. The launch of the ISO space telescope space telescope (Kessler et al. 1996), embarking the ISOCAM instrument (Cesarsky et al. 1996), was a huge step forward, opening for the first time to door to the distant dusty Universe, and showing that dust is a key component to study the evolution of galaxies (e.g., Franceschini et al. 2001; Chary and Elbaz 2001). The consequent launch of *Spitzer* (Werner et al. 2004) built on this success and allowed the detection of moderately luminous objects at high redshifts ($z < 3$) in the mid-infrared (MIR) thanks to the MIPS instrument (Rieke et al. 2004). It was soon followed by the *Herschel* space telescope (Pilbratt et al. 2010), which provided better constraints on the spectrum of the dust emission by observing in the FIR with the PACS (Poglitsch et al. 2010) and SPIRE instruments (Griffin et al. 2010).

Nevertheless only the most luminous star-forming objects can be detected at high redshifts, yielding strongly SFR biased samples (Elbaz et al. 2011). In particular, most galaxies reliably detected with these instruments at $z \geq 3$ are very luminous starbursts, making it difficult to study the properties of “normal” galaxies at these epochs. So far only a handful of studies have probed in a relatively complete manner the Universe at $z \gtrsim 3$ with IR facilities (e.g., Heinis et al. 2014; Pannella et al. 2015) and most of what we know about normal galaxies at $z > 3$ is currently based on UV light alone (Daddi et al. 2009; Stark et al. 2009; Bouwens et al. 2012; Stark et al. 2013; González et al. 2014; Salmon et al. 2015).

Here we take advantage of the deepest data ever taken with *Herschel* in the Great Observatories Origins Deep Survey fields (GOODS, PI: D. Elbaz), covering the GOODS–North and GOODS–South fields, and the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey fields (CANDELS, PI: M. E. Dickinson) covering a fraction of the Ultra-Deep Survey¹ (UDS) and Cosmic Evolution Survey (COSMOS) fields, to infer stricter constraints on the existence and relevance of the Main Sequence in the young Universe up to $z = 4$. To do so, we first construct a mass-selected sample with known photometric redshifts and stellar masses and then isolate star-forming galaxies within it. We bin this sample in redshift and stellar mass and stack the *Herschel* images. This allows us to infer their average L_{IR} , and thus their SFRs. We then present a new technique we call “scatter stacking” to measure the dispersion around the average stacked SFR, taking nondetected galaxies into account. Finally, we cross-match our sample with *Herschel* catalogs to study individually detected galaxies.

¹This field is also known as the Subaru XMM Deep Survey (SXDS) field.

2.2 Sample and Observations

We use the ultra-deep H -band catalogs provided by the CANDELS–*HST* team (Grogin et al. 2011; Koekemoer et al. 2011) in three of the CANDELS fields, namely GOODS–South (GS Guo et al. 2013), UDS (Galametz et al. 2013), and COSMOS (Nayyeri et al. in prep.). With the GOODS–North (GN) CANDELS catalog not being finalized at the time of writing, we fall back to a ground-based K_s -band catalog. To extend our sample to rarer and brighter objects, we also take advantage of the much wider area provided by the K_s -band imaging in the COSMOS field acquired as part of the UltraVISTA program (UVISTA). In the following, we will refer to this field as “COSMOS UltraVISTA”, while the deeper but smaller region observed by CANDELS will be called “COSMOS CANDELS”.

Using either the H or the K_s band flux as the selection criterion will introduce potentially different selection effects. In practice, these two bands are sufficiently close in wavelengths that one does not expect major differences to arise: if anything, the K_s -band catalogs are potentially more likely to be mass-complete, since this band will probe the rest-frame optical up to higher redshifts. However these catalogs are ground-based, and lack both angular resolution and depth when compared to the *HST* H -band data. It is thus necessary to carefully estimate the mass completeness level of each catalog, and only consider mass-complete regimes in the following analysis.

All these fields were selected for having among the deepest *Herschel* observations, which are at the heart of the present study, along with high-quality, multi-wavelength photometry in the UV to NIR. The respective depths of each catalog are listed in Table 2.1. We next present the details of the photometry and source extraction of each field.

2.2.1 GOODS–North

GOODS–North is one of the fields targeted by the CANDELS–*HST* program, and the last to be observed. Consequently, the data reduction was delayed compared to the other fields and there was no available catalog when we started this work. We thus use the ground-based K_s -band catalog presented in Pannella et al. (2015), which is constructed from the deep CFHT WIRCAM K_s -band observations of Wang et al. (2010). This catalog contains 20 photometric bands from the NUV to IRAC 8 μm and was built using SExtractor (Bertin and Arnouts 1996) in dual image mode, with the K_s -band image as the detection image. Fluxes are measured within a 2'' aperture on all images, and the effect of varying point spread function (PSF) and / or seeing is accounted for using PSF-matching corrections. Per-object aperture corrections to total are provided by the ratio of the FLUX_AUTO as given by SExtractor and the aperture K_s -band flux. This results in a 0.8'' angular resolution catalog of 79 003 sources and a 5σ limiting magnitude of $K_s = 24.5$.

Table 2.1 Catalog depths for each field

Field	Area ^a arcmin ²	NIR ^b (5σ)	24 μm μJy (3σ) ^c	100 μm mJy (3σ) ^c	160 μm mJy (5σ) ^c	250 μm mJy (5σ) ^c	350 μm mJy (5σ) ^c	500 μm mJy (5σ) ^c
GN	168	$K_s < 24.5$	21	1.1	2.7	7.3	7.8	13
GS	184	$H < 27.4$	20	0.8	2.4	7.0	7.5	13
UDS	202	$H < 27.1$	40	1.7	3.9	10	11	13
COSMOS								
-CANDELS	208	$H < 27.4$	27–40	1.5	3.1	11	14	14
-UVISTA	1.6 deg ²	$K_s < 23.4$	27–40	4.6	9.9	—	—	—

^aThis is the sky coverage of our sample, and may be smaller than the nominal area of the detection image

^bFor the fields GS, UDS and COSMOS–CANDELS, the H -band coverage is highly nonuniform over the whole field. Here we conservatively quote the depth of the shallowest region

^cThese limiting depths are computed from the median uncertainty on the fluxes as reported in the *Herschel* catalogs of each field

The K_s -band image extends over 0.25 deg², but only the central area is covered by *Spitzer* and *Herschel*. We therefore only keep the sources that fall inside the coverage of those two instruments, i.e., 15 284 objects in 168 arcmin². We also remove stars identified either from the SExtractor flag CLASS_STAR for bright enough objects ($K_s < 20$), or using the BzK color-color diagram (Daddi et al. 2004). Our final sample consists of 14 828 galaxies, 12 317 of which are brighter than the 5σ limiting magnitude, with 3 775 spectroscopic redshifts.

The *Herschel* images in both PACS and SPIRE were obtained as part of the GOODS–*Herschel* program (Elbaz et al. 2011). The source catalog of *Herschel* and *Spitzer* MIPS 24 μm are taken from the public GOODS–*Herschel* DR1. *Herschel* PACS and SPIRE 250 μm flux densities are extracted using PSF fitting at the position of MIPS priors, themselves extracted from IRAC priors. SPIRE 350 μm and 500 μm flux densities are obtained by building a reduced prior list out of the 250 μm detections. This procedure, described in more detail in Elbaz et al. (2011), yields 2 681 MIPS and 1 039 *Herschel* detections ($>3\sigma$ in any PACS band or $>5\sigma$ in SPIRE, following Elbaz et al. 2011) that we could cross-match to the K_s -band catalog using their IRAC positions.

2.2.2 GOODS–South, UDS, and COSMOS CANDELS

In GOODS–South, UDS and COSMOS CANDELS we use the official CANDELS catalogs presented, respectively, in Guo et al. (2013) (version 121114), Galametz et al. (2013) (version 120720) and Nayyeri et al. (in prep.) (version 130701). They

are built using SExtractor in dual image mode, using the *HST H*-band image as the detection image to extract the photometry at the other *HST* bands. The ground-based and *Spitzer* photometry is obtained with TFIT (Laidler et al. 2007). The *HST* photometry was measured using the FLUX_ISO from SExtractor and corrected to total magnitudes using either the FLUX_BEST or FLUX_AUTO measured in the *H* band, while the ground-based and *Spitzer* photometry is already “total” by construction. These catalogs gather 16 photometric bands in GOODS–South, 19 in UDS, and 27 in COSMOS, ranging from the *U* band to IRAC 8 μm , for a total of 34 930 (respectively 35 932 and 38 601) sources, 1 767 (respectively 575 and 1 175) of which have a spectroscopic redshift. The *H*-band exposure in the fields is quite heterogeneous, the 5σ limiting magnitude ranging from 27.4 to 29.7 in GOODS–South, 27.1 to 27.6 in UDS, and 27.4 to 27.8 in COSMOS, but it always goes much deeper than the available ground-based photometry. These extreme depths can also become a problem, especially when dealing with sources so faint that they are significantly detected in the *HST* images only. The SED of these objects is so poorly constrained that we cannot robustly identify them as galaxies, or compute accurate photometric redshifts. To solve this issue, one would like to only keep sources that have a sufficient wavelength coverage, e.g., imposing a significant detection in at least ten UV to NIR bands, but this would introduce complex selection effects. Here we decide to only keep sources that have an *H*-band magnitude brighter than 26. This ensures that the median number of UV to NIR bands for each source (along with the 16th and 84th percentiles) is 11_{-2}^{+3} , 16_{-4}^{+3} and 21_{-5}^{+5} , respectively, as compared to 9_{-4}^{+4} , 13_{-5}^{+5} and 18_{-7}^{+7} when using the whole catalogs.

As for GOODS–North, we remove stars using a combination of morphology and BzK classification, and end up with 18 364 (respectively 21 552 and 24 396) galaxies with $H < 26$ in 184 arcmin² (respectively 202 arcmin² and 208 arcmin²).

In both UDS and COSMOS, the *Herschel* PACS and SPIRE images were taken as part of the CANDELS–*Herschel* program (PI: M. Dickinson), and are slightly shallower than those in the two GOODS fields. The MIPS 24 μm images, however, are clearly shallower, since they reach a noise level of approximately 40 μJy (1σ), as compared to the 20 μJy in GOODS. In COSMOS, however, the MIPS map contains a “deep” region (Sanders et al. 2007) that covers roughly half of the COSMOS CANDELS area with a depth of about 30 μJy .

In those two fields, sources are extracted with the same procedure as in GOODS–North (Inami et al. in prep). These catalogs provide, respectively, 2 461 and 2 585 MIPS sources as well as 730 and 1 239 *Herschel* detections within the *HST* coverage. Since the IRAC priors used in the source extraction come directly from the CANDELS catalog, no cross-matching has to be performed.

The *Herschel* images in GOODS–South come from three separate programs. The PACS images are the result of the combined observation of both GOODS–*Herschel* and the PACS Evolutionary Probe program (PEP, Lutz et al. 2011), while SPIRE images were obtained as part of the *Herschel* Multi-tiered Extragalactic Survey (HerMES, Oliver et al. 2012). The PACS fluxes are taken from the public PEP DR1 catalog (Magnelli et al. 2013), and were extracted using the same procedure as in

GOODS–North. To extract the SPIRE fluxes, B. Magnelli first produced the maps by downloading the individual level-2 data products covering the full ECDFS from the *Herschel* ESA archive² and reduced them following the same procedure as the other sets of SPIRE data from GOODS and CANDELS–*Herschel*. He then performed the source extraction as in Elbaz et al. (2011). This catalog provides 1 875 MIPS and 1 058 *Herschel* detections within the *HST* coverage, which were cross-matched to the CANDELS catalog using their IRAC positions.

2.2.3 COSMOS UltraVISTA

Only a small region of the COSMOS field has been observed within the CANDELS program. For the remaining area, we have to rely on ground-based photometry. To this end, we consider two different K_s -band catalogs, both based on the UltraVISTA DR1 (McCracken et al. 2012), and taken from Muzzin et al. (2013b) and Ilbert et al. (2013).

The first catalog, presented in Muzzin et al. (2013b), is built using SExtractor in dual image mode, with the K_s -band image as detection image. The photometry in the other bands is extracted using PSF-matched images degraded to a common resolution of $\sim 1.1''$ and an aperture of $2.1''$, except for the *Spitzer* bands and *GALEX*. Here, an alternative cleaning method is used, where nearby sources are first subtracted using the PSF-convolved K_s -band profiles (u^* band for *GALEX*), then the photometry of the central source is measured inside an aperture of $3''$. In both cases, aperture fluxes are corrected to total using the ratio of FLUX_AUTO and aperture K_s -band flux. In the end, the catalog contains 30 photometric bands ranging from *GALEX* FUV to IRAC 8 μm (we did not use the 24 μm photometry), for a total of 262 615 objects and a 5σ limiting magnitude of $K_s = 23.4$. As for the CANDELS fields, stars are excluded using a combination of morphological and BzK classification, resulting in a final number of 249 823 galaxies within 1.6 deg^2 , 168 509 of which are brighter than the 5σ limiting magnitude, with 5 532 having spectroscopic redshifts.

The second catalog, presented in Ilbert et al. (2013), is very similar in that, apart from missing *GALEX* and *Subaru* g^+ , it uses the same raw images and was also built with SExtractor. The difference lies mostly in the extraction of IRAC fluxes. Here, and for IRAC only, SExtractor is used in dual image mode, with the *Subaru* i -band image as the detection image. Since the IRAC photometry was not released along with the rest of the photometry, we could not directly check the consistency of the two catalogs, nor use this photometry to derive accurate galaxy properties. Nevertheless, the photometric catalog comes with a set of photometric redshifts and stellar masses that we can use as a consistency check. These were built using a much more extensive but private set of spectroscopic redshifts, and are thus expected to be of higher quality. A direct comparison of the two photometric redshift estimations shows a constant relative scatter of 4 % below $z = 2$. At higher redshifts, the scatter

²<http://www.cosmos.esa.int/web/herschel/science-archive>.

increases to 10 % because of the ambiguity between the Balmer and Lyman breaks. This ambiguity arises because of the poor wavelength coverage caused by the shallow depths of these surveys, but it takes place in a redshift regime where our results are mostly based on the deeper, and therefore more robust, CANDELS data. We also checked that redoing our analysis with Ilbert et al.'s catalog yielded very similar results in the mass-complete regimes.

Finally, while the *Spitzer* MIPS imaging is the same as that in COSMOS CANDELS, the *Herschel* PACS images in this wide field were taken as part of the PEP program, at substantially shallower depth (Lutz et al. 2011). The *Spitzer* MIPS and *Herschel* PACS photometry are taken from the public PEP DR1 catalog³, itself based on the MIPS catalog of Le Floc'h et al. (2009), yielding 37 544 MIPS and 9 387 PACS detections successfully cross-matched to the first K_s band catalog.

2.2.4 Photometric Redshifts and Stellar Masses

Photometric redshifts (photo- z) and stellar masses are derived by M. Pannella using the procedure described in Pannella et al. (2015). Briefly, photo- zs are computed using EAZY⁴ (Brammer et al. 2008) in its standard setup. Global photometric zero points are adjusted iteratively by comparing the photo- zs to the available spectroscopic redshifts (spec- z), and minimizing the difference between the two. We emphasize that, although part of these adjustments are due to photometric calibration issues, they also originate from defects in the adopted SED template library. To estimate the quality of the computed photo- zs , we request that the odds computed by EAZY, which is the estimated probability that the true redshift lies within $\Delta z = 0.2 \times (1 + z_{\text{phot}})$ (Benítez 2000), be larger than 0.8. A more stringent set of criteria is adopted in COSMOS CANDELS, because of the lower quality of the photometric catalog. To prevent contamination of our sample from issues in the photometry, we prefer to be more conservative and only keep odds > 0.98 and impose that the χ^2 of the fit be less than 100 to remove catastrophic fits. The median $\Delta z \equiv |z_{\text{phot}} - z_{\text{spec}}| / (1 + z_{\text{spec}})$ is respectively 3.0, 3.2, 1.8, 2.0, and 0.8 % in GOODS–North, GOODS–South, UDS CANDELS, COSMOS CANDELS, and COSMOS UltraVISTA. We stress however that the representativeness of this accuracy also depends on the spectroscopic sample. In COSMOS UltraVISTA, for example, we only have spec- zs for the brightest objects, hence those that have the best photometry. Fainter and more uncertain sources thus do not contribute to the accuracy measurement, which is why the measured value is so low. Lastly, although we use these spec- zs to calibrate our photo- zs , we do not use them afterwards in this study. The achieved precision of our photo- zs is high enough for our purposes, and the selection functions of all spectroscopic surveys we gather here are very different,

³<http://www.mpe.mpg.de/ir/Research/PEP/DR1>

⁴<http://code.google.com/p/eazy-photoz>.

if not unknown. To avoid introducing any uncontrollable systematic, we therefore decide to consistently use photo- z s for all our sample.

Stellar masses are derived using FAST⁵ (Kriek et al. 2009), adopting Salpeter (1955) IMF,⁶ the Bruzual and Charlot (2003) stellar population synthesis model and assuming that all galaxies follow delayed exponentially declining⁷ star formation histories (SFHs), parametrized by $SFR(t) \propto (t/\tau^2) \exp(-t/\tau)$ with $0.01 < \tau < 10$ Gyr. Dust extinction is accounted for assuming the Calzetti et al. (2000) law, with a grid ranging from $A_V = 0$ to 4. Metallicity is kept fixed and equal to Z_\odot . We assess the quality of the stellar mass estimate with the reduced χ^2 of the fit, only keeping galaxies for which $\chi^2 < 10$.

2.2.5 Rest-Frame Luminosities and Star Formation Rates

Star formation rates are typically computed by measuring the light of young OB stars, which emit the bulk of their light in the UV. However this UV light is most of the time largely absorbed by the interstellar dust, and re-emitted in the IR as thermal radiation. To obtain the total SFR of a galaxy, it is therefore necessary to combine the light from both the UV and the IR.

Rest-frame luminosities in the FUV (1500 Å), U , V , and J bands are computed by M. Pannella with *EAZY* by convolving the best-fit SED model from the stellar mass fit with the filter response curves. The FUV luminosity is then converted into SFR *uncorrected* for dust attenuation using the formula from Daddi et al. (2004), i.e.,

$$SFR_{\text{UV}} = 2.17 \times 10^{-10} L_{\text{UV}} [L_\odot]. \quad (2.1)$$

I computed the infrared luminosity L_{IR} , following the procedure of Elbaz et al. (2011). We fit the *Herschel* flux densities with CE01 templates, and compute L_{IR} from the best-fit template. In this procedure, photometric points below 30 μm rest-frame are not used in the fit since this is a domain that is potentially dominated by active galactic nuclei (AGNs) torus emission, and not by star formation (e.g., Mullaney et al. 2011). We come back to this issue in Sect. 2.2.6. This IR luminosity is, in turn, converted into dust-reprocessed SFR using the formula from Kennicutt (1998)

$$SFR_{\text{IR}} = 1.72 \times 10^{-10} L_{\text{IR}} [L_\odot]. \quad (2.2)$$

⁵<http://astro.berkeley.edu/string~mariska/FAST.html>.

⁶Using another IMF would systematically shift both our M_* and SFRs by approximately the same amount, and therefore would not affect the shape of the Main Sequence.

⁷Other star formation histories were considered, in particular with a constant or exponentially declining SFR. Selecting all galaxies from $z > 0.3$ to $z < 5$, no systematic offset is found, while the scatter evolves mildly from 0.12 dex at $M_* = 1 \times 10^8 M_\odot$ to 0.08 at $M_* = 3 \times 10^{11} M_\odot$.

The total SFR is finally computed as the sum of SFR_{UV} and SFR_{IR} . The above two relations are derived assuming a Salpeter (1955) initial mass function (IMF) and assume that the SFR remained constant over the last 100 Myr.

A substantial number of galaxies in this sample (50 % in the CANDELS fields, 75 % in COSMOS UltraVISTA) are detected by *Spitzer* MIPS but not by *Herschel*. Although for these galaxies we only have a single photometric point in the MIR, we can still infer accurate monochromatic SFRs using the original L_{IR} calibration of the CE01 library. This calibration is valid up to $z < 1.5$, as shown in Elbaz et al. (2011), hence we only use MIPS-derived SFRs for sources not detected by *Herschel* over this redshift range. Although there exist other calibrations that are applicable to higher redshifts (e.g., Elbaz et al. 2011; Wuyts et al. 2011), we do not know how they would impact the measurement of the scatter of the Main Sequence. We therefore prefer not to use them and discard the 24 μm measurements above $z = 1.5$. Galaxies not detected in the MIR ($z < 1.5$) or FIR have no individual SFR estimates and are only used for stacking. When working with detections alone (Sect. 2.4.6), this obviously leads to an SFR selected sample and is taken into account by estimating the SFR completeness.

Lastly, there are some biases that can affect our estimates of SFR from the IR. In particular, the dust can also be heated by old stars that trace the total stellar mass content rather than the star formation activity (e.g., Salim et al. 2009). Because of the relatively low luminosity of these stars, this will most likely be an issue for massive galaxies with low star formation activity, i.e., typically quiescent galaxies (see, e.g., Sect. 2.7 where we analyze such cases). Since we remove these galaxies from our sample, we should not be affected by this bias. This is also confirmed by the excellent agreement of IR-based SFR estimates with those obtained from the radio emission (e.g., Pannella et al. 2015), the latter not being affected by the light of old stars.

2.2.6 A Mass-Complete Sample of Star-Forming Galaxies

We finalize our sample by selecting actively star-forming galaxies. Indeed, the observation of a correlation between mass and SFR only applies to galaxies that are still forming stars, and not to quiescent galaxies. The latter are not evolving anymore and pile up at high stellar masses with little to no detectable signs of star formation. Nevertheless, they can still show residual IR emission due to the warm inter stellar medium (ISM). This cannot be properly accounted for with the CE01 library, and will be misinterpreted as an SFR tracer.

Several methods exist to exclude quiescent galaxies. The most obvious is to select galaxies based on their specific SFR ($\text{sSFR} \equiv \text{SFR}/M_*$). Indeed, quiescent galaxies have very low SFR by definition, and they are preferentially found at high M_* . Therefore, they will have very low sSFR compared to star-forming galaxies. This obviously relies on the very existence of the correlation between SFR and M_* , and removing galaxies with too low sSFR would artificially create the correlation even where it does not exist. On the other hand, selecting galaxies based on their SFR

alone would destroy the correlation, even where it exists (Rodighiero et al. 2011; Lee et al. 2013). It is therefore crucial that the selection does not apply directly to any combination of SFR or M_* . Furthermore, these methods require that an accurate SFR is available for all galaxies, and this is something we do not have since most galaxies are not detected in the mid- or far-IR. We must therefore select star-forming galaxies based on information that is available for all the galaxies in our sample, i.e., involving optical photometry only.

There are several color-magnitude or color-color criteria that are designed to accomplish this. Some, like the BzK approach (Daddi et al. 2004), are based on the observed photometry and are thus very simple to compute, but they also select a particular redshift range by construction. This is not desirable for our sample, and we thus need to use rest-frame magnitudes. Color-magnitude diagrams (e.g., $U - r$ versus r -band magnitude as in Baldry et al. 2004) tend to wrongly classify some of the red galaxies as passive, while they could also be red because of high dust attenuation. Since high mass galaxies suffer the most from dust extinction (Pannella et al. 2009), it is thus likely that color-magnitude selections would have a nontrivial effect on our sample. It is therefore important to use another color to disentangle galaxies that are red because of their old stellar populations and those that are red because of dust extinction.

To this end, Williams et al. (2009) devised the UVJ selection, based on the corresponding color-color diagram introduced in Wuyts et al. (2007). It uses the $U - V$ color, similar to the $U - r$ from the standard color-magnitude diagram, but combines it to the $V - J$ color to break the age–attenuation degeneracy. Although the bimodality stands out clearly on this diagram, the locus of the passive cloud has been confirmed by Williams et al. (2009) using a sample of massive galaxies in the range $0.8 < z < 1.2$ with little or no [O II] line emission, while the active cloud falls on the Bruzual and Charlot (2003) evolutionary track for a galaxy with constant SFR. One can then draw a dividing line that passes between those two clouds to separate one population from the other. In this chapter we use the following definition, at all redshifts and stellar masses:

$$\text{quiescent} = \begin{cases} U - V > 1.3, \\ V - J < 1.6, \\ U - V > 0.88 \times (V - J) + 0.49. \end{cases} \quad (2.3)$$

This definition differs by only 0.1 magnitude compared to that of Williams et al. (2009). Rest-frame colors can show offsets of similar order from one catalog to another, because of photometric coverage and uncertainties in the zero-point corrections. It is thus common to adopt slightly different definitions to account for these effects (see e.g., Cardamone et al. 2010; Whitaker et al. 2011; Brammer et al. 2011; Strazzullo et al. 2013; Viero et al. 2013; Muzzin et al. 2013b). In COSMOS UltraVISTA, we follow the definition given by Muzzin et al. (2013b). In Chap. 5, we will also use a slightly modified version of this diagram for $z = 0$ galaxies.

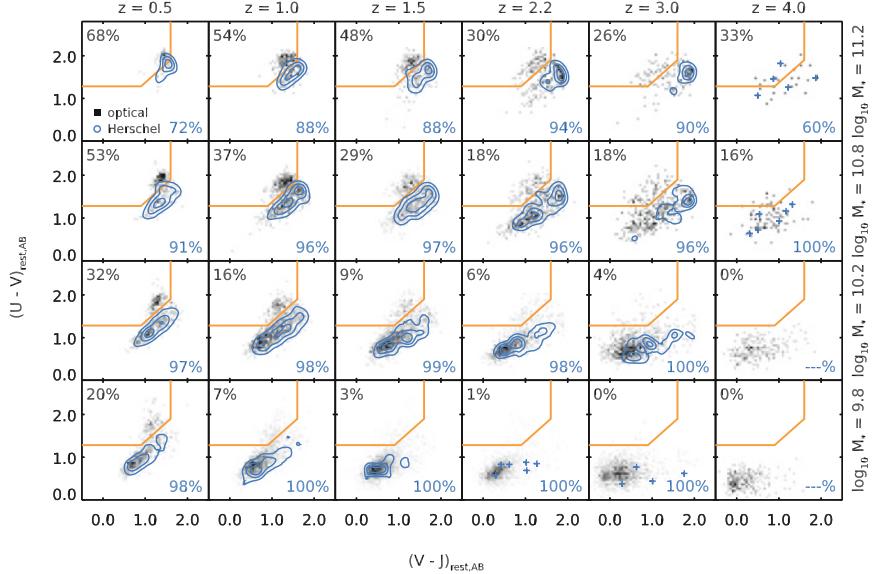


Fig. 2.1 *UVJ* diagrams in each bin of redshift (*horizontally*) and mass (*vertically*) of our CANDELS sample. The central value of the redshift and mass bins are shown at the *top* and on *right-hand side* of the figure, respectively. The dividing line between active and passive galaxies is shown as a *solid orange line* on each plot, with passive galaxies located on the *top-left* corner. We show in the background the distribution of sources from the *H*-band catalogs in *gray scale*. We also overplot the position of sources detected with *Herschel* as *blue contours* or, when the source density is too low, as individual *blue open circles*. On the *top-left* corner of each plot, we give the fraction of *H* band-selected galaxies that fall inside the quiescent region, and on the *bottom-right* corner we show the fraction of *Herschel* sources that reside in the star-forming region (color figure online)

The corresponding diagram in bins of mass and redshift for the CANDELS fields is shown in Fig. 2.1. Here we also overplot the location of the galaxies detected by *Herschel*; because of the detection limit of the surveys, the vast majority of *Herschel* detections have high SFRs. We therefore expect them to fall on the *UVJ* “active” region. This is indeed the case for the vast majority of these galaxies, even when the majority of optical sources are quiescent as is the case at $z = 0.5$ and $\log_{10}(M_*/M_\odot) > 10$. In total, only 5 % of the galaxies in our *Herschel* sample are classified as passive, and about a third of those have a probability larger than 20 % to be misclassified because of uncertainties in their *UVJ* colors. The statistics in COSMOS UltraVISTA are similar.

The number of galaxies with reliable redshifts and stellar masses (see Sect. 2.2.4) that are classified with this diagram as actively star-forming are reported in Table 2.2. These are the galaxies considered in the following analysis. As a check, we also analyze separately the quiescent galaxies in Sect. 2.7.

Finally, we do not explicitly exclude known AGNs from our sample. We expect AGNs to reside in massive star-forming galaxies (Kauffmann et al. 2003; Mullaney

Table 2.2 Number of object in our sample per field

Field	All galaxies ^a	SF ^b	Spec- <i>z</i> ^c	<i>Herschel</i> ^d
GN	6 973	5 358	2 605	867
GS	5 539	4 630	2 275	947
UDS	7 455	6 372	504	654
COSMOS				
-CANDELS	7 580	6 599	811	976
-UVISTA	58 202	39 375	3 736	7 053

^aNumber of galaxies in our mass-complete NIR sample, removing stars from the Milky Way, spurious sources, and requiring *Spitzer* and *Herschel* coverage

^bFinal subsample of good quality galaxies classified as star-forming with the *UVJ* criterion (see Sect. 2.2.6)

^cSubsample of “SF” galaxies with a spectroscopic redshift (various sources, see catalog papers for references)

^dSubsample of “SF” galaxies with a detection in any *Herschel* band, requiring $>3\sigma$ significance in PACS or $>5\sigma$ in SPIRE (following Elbaz et al. 2011)

et al. 2012; Santini et al. 2012; Juneau et al. 2013; Rosario et al. 2013). While the most luminous optically unobscured AGNs may greatly perturb the optical photometry, and therefore the measurement of redshift and stellar mass, they will also degrade the quality of the SED fitting because we have no AGN templates in our fitting libraries. This can produce an increased χ^2 , hence selecting galaxies with $\chi^2 < 10$ (see Sect. 2.2.4) helps remove some of these objects. Also, their point-like morphology on the detection image tends to make them look like stars, which are systematically removed from the sample. The more common moderate luminosity AGNs can still be fit properly with galaxy templates (Salvato et al. 2011). Therefore, several AGNs do remain in our sample without significantly affecting the optical SED fitting and stellar masses. Still, obscured AGNs will emit some fraction of their light in the IR through the emission of a dusty torus. To prevent pollution of our FIR measurements by the light of such dusty AGNs, we only use the photometry at rest-frame wavelengths larger than 30 μm , where the contribution of the AGN is negligible (Mullaney et al. 2011). Indeed, while the most extreme AGNs may affect mid-to-far IR colors, such as 24-to-70 μm color, their far-IR colors are indistinguishable from that of star-forming galaxies (Hatziminaoglou et al. 2010). By rejecting the most problematic cases, and mitigating against AGN contribution to the IR, we aim to remove severe contamination while retaining a high sample completeness.

2.2.7 Completeness and Mass Functions

The last step before going through the analysis is to make sure that, in each stellar mass bin we will work with, as few galaxies as possible are missed because of our selection criteria. The fact that we built these samples by starting from an NIR

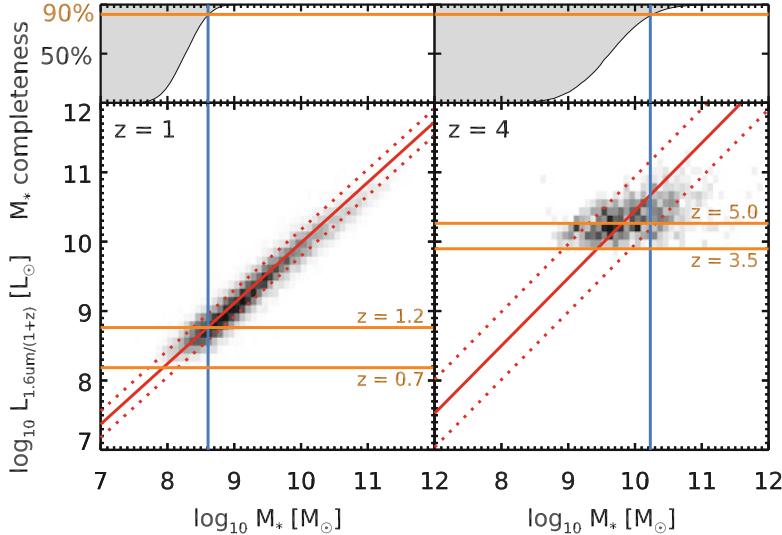


Fig. 2.2 Correlation between the stellar mass and the luminosity in the observed-frame H band at $0.7 < z < 1.2$ (left) and $3.5 < z < 5$ (right) in the three CANDELS fields GOODS-South, UDS, and COSMOS. On the bottom plots, the two horizontal orange lines show the position of the $H = 26$ limiting magnitude at $z = z_{\min}$ and $z = z_{\max}$. The red line is the best-fit relation, and the dotted lines above and below show the 1σ dispersion (0.2 and 0.5 dex, respectively). The blue vertical line shows the locus of the estimated 90 % mass completeness in each redshift bin. The top plots show the evolution of completeness (i.e., the estimated fraction of detected objects) with stellar mass, and the horizontal orange line shows the 90 % completeness level (color figure online)

selection makes it much simpler to compute the corresponding mass completeness: the stellar mass of a galaxy at a given redshift is indeed well correlated with the luminosity in the selection band (either H or K_s), as illustrated in Fig. 2.2, the scatter around the correlation being caused by differences of age, attenuation, and to some extent flux uncertainties and k -correction. From our sample, we can actually see by looking at this correlation with various bands (H , K_s , and IRAC channels 1 and 2) that this scatter is minimal (0.14 dex) when probing the rest-frame $1.7 \mu\text{m}$, but it reaches 0.4 dex in the rest-frame UV (3500\AA). While this value is of course model dependent, it stresses the importance of having high-quality NIR photometry, especially the *Spitzer* IRAC bands (observed $3\text{--}5 \mu\text{m}$).

To estimate the mass completeness, we decided to use an empirical approach, where we do not assume any functional form for the true mass function. Instead, we directly compute the completeness assuming that, at a given redshift, the stellar mass is well estimated by a power law of the luminosity (measured either from the observed H or K_s band), i.e., $M_* = C L^\alpha$, plus a Gaussian scatter in log space.

Table 2.3 $\log_{10}(M_*/M_\odot)$ above which our samples are at least 90 % complete, for each catalog

Catalog	$z = 0.5$	1.0	1.5	2.2	3.0	4.0
GN	8.9	9.3	9.8	10.1	10.5	10.7
CANDELS ^a	8.3	8.7	9.0	9.4	9.9	10.3
COSMOS	9.1	9.6	10.1	10.6	10.9	11.3
UVISTA						

^aThese values are valid for GOODS–South, UDS, and COSMOS CANDELS, keeping all sources with $H < 26$

We fit this power law and estimate the amplitude of the scatter using the detected galaxies, as shown in Fig. 2.2. Using this model (red solid and dotted lines) and knowing the limiting luminosity in the selection band (orange horizontal lines), we can estimate how many galaxies we miss at a given stellar mass, using, e.g., a Monte Carlo simulation. At a given stellar mass, we generate a mock population of galaxies with uniform redshift distribution within the bin and estimate what would be their luminosity in the selection band by using the above relation and adding a Gaussian scatter to the logarithm of the luminosity. The completeness is then computed as the fraction of galaxies that have a luminosity greater than the limiting luminosity at the considered redshift. We consider our catalogs as “complete” when the completeness reaches at least 90 %.

The same procedure is used on COSMOS UltraVISTA and GOODS–North separately, and the estimated completeness levels are all reported in Table 2.3. We compared the values obtained in GOODS–North with those reported in Pannella et al. (2015), where the completeness is estimated following Rodighiero et al. (2010) using a stellar population model. The parameters of the model chosen in Pannella et al. (2015) are quite conservative, and their method consistently yields mass limits that are on average 0.3 dex higher than ours. In COSMOS UltraVISTA, we obtain values similar to that of Muzzin et al. (2013a).

Finally, we build stellar mass functions by simply counting the number of galaxies in bins of redshift and stellar masses in the three CANDELS fields that are H -band selected, and normalize the counts by the volume that is probed. These raw mass functions are presented in Fig. 2.3 as dashed lines. Assuming that the counts follow a Schechter-like shape (Schechter 1976), i.e., rising with a power law toward low stellar mass, the incompleteness of our sample is clearly visible. We then use the estimated completeness (top panel in Fig. 2.2) to correct the stellar mass functions. Here, we limit ourselves to reasonable corrections of at most a factor two in order not to introduce too much uncertainty in the extrapolation. The resulting mass functions are shown as solid lines in Fig. 2.3, with shaded areas showing the Poisson noise. The obtained mass functions are in good agreement with those already published in the literature (e.g., Ilbert et al. 2013).

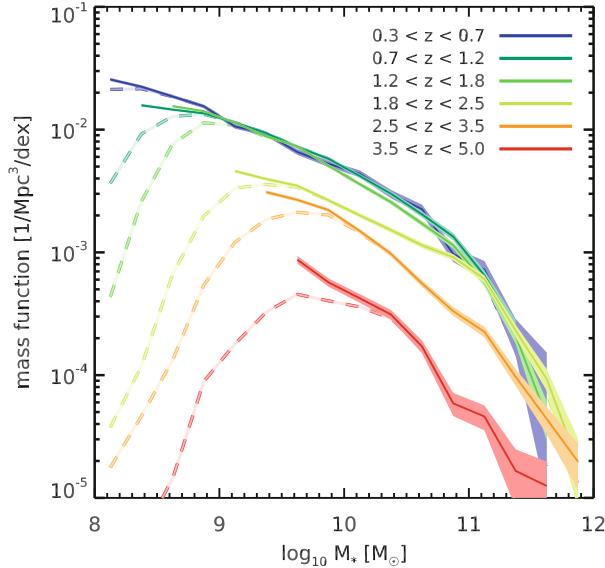


Fig. 2.3 Evolution of the star-forming galaxy stellar mass function with redshift in the three CANDELS fields GOODS-South, UDS, and COSMOS for galaxies brighter than $H = 26$. Raw, incomplete counts are shown as *dashed lines*, while *solid lines* show the corrected counts. The *shaded areas* correspond to Poissonian errors (color figure online)

2.3 Deriving Statistical Properties of Star-Forming Galaxies

Because of the limitations of the *Herschel* surveys (the result of photometric or confusion noise), we cannot derive robust individual SFRs for all the sources in our sample (see Sect. 2.2.5). Indeed, the fraction of star-forming galaxies detected in the FIR ranges from 80 % at $M_* > 3 \times 10^{10} M_\odot$ and $z < 1$, to almost 0 % for $M_* < 10^{10} M_\odot$ and $z > 1$. Above $z = 1$, the completeness in FIR detections reaches better than 60 % only above $M_* = 10^{11} M_\odot$ and up to $z = 2.5$. Below this mass and above that redshift, the FIR completeness is lower than 20–30 %.

We overcome these limitations by stacking the *Herschel* images. Stacking is a powerful and routinely used technique that combines the signal of multiple sources at various positions on the images, known from deeper surveys (see, e.g., Dole et al. 2006, where it was first applied to FIR images). This effectively increases the signal to noise ratio of the measurement, allowing us to probe fainter fluxes than can be reached by the usual source extraction. The price to pay is that we lose information about each individual source, and only recover statistical properties of the considered sample. Commonly, this method is used to determine the average flux density of a selected population of objects. We will show in the following that it can also be used

to obtain information on the flux *distribution* of the sample, i.e., not only its average flux, but also how much the stacked sources scatter around this average value.

This scatter is a crucial information. If we measure an average correlation between SFR and M_* , as has been measured in several other studies at different redshifts, this correlation cannot be called a “sequence” if the sources show a large dispersion around it.

Several studies have already measured this quantity. Noeske et al. (2007) and Elbaz et al. (2007) at $z = 1$ reported a 1σ dispersion in $\log_{10}(\text{SFR})$ of around 0.3 dex from *Spitzer* MIPS observations of a flux-limited sample. At $z = 2$, Rodighiero et al. (2011) reported 0.24 dex, using mostly UV-derived SFRs, while Whitaker et al. (2012) reported 0.34 dex from *Spitzer* MIPS observations. These two studies tested the consistency of their SFR estimator on average, but we do not know how they impact the measure of the dispersion. The variation found in these two studies suggests that this is indeed an issue (see for example the discussion in Speagle et al. 2014). On the one hand, UV SFRs have to be corrected for dust extinction. If one assumes a single extinction law for the whole sample, one might artificially reduce the dispersion. On the other hand, MIPS 24 μm at $z = 2$ probes the rest-frame 8 μm . While Elbaz et al. (2011) have shown that it correlates well with L_{IR} , this same study also demonstrates that it misses a fraction of L_{IR} that is proportional to the distance from the Main Sequence. This can also have an impact on the measured dispersion.

Here we measure for the first time the SFR– M_* Main Sequence and its dispersion with a robust SFR tracer down to the very limits of the deepest *Herschel* surveys to constrain its existence and relevance at higher redshifts and lower stellar masses.

2.3.1 Simulated Images

All the methods described in this section have been extensively tested to make sure that they are not affected by systematic biases or, if they are, to implement the necessary corrections. We conduct these tests on simulated *Herschel* images that we set up to be as close as possible to the real images, in a statistical sense. In other words, we reproduce the number counts, the photometric noise, the confusion noise, and the source clustering. The algorithms, the methodology, and the detailed results are described fully in Appendix 2.8.

2.3.2 The Stacking Procedure

We divide our star-forming galaxy sample into logarithmic bins of stellar mass and redshift, as shown in Fig. 2.4, to have a reasonable number of sources in each bin. We then go to the original *Herschel* images of each field and extract $N \times N$ pixel cutouts around each source in the bin, thus building a pixel cube. We choose $N = 41$ for all *Herschel* bands, which is equivalent to 8 times the full-width at half maximum

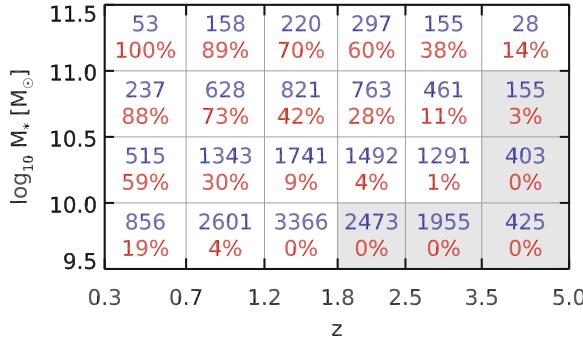


Fig. 2.4 Redshift and stellar mass bins chosen for stacking. We display in each bin (from *top* to *bottom*) the total number of star-forming H or K_s -band galaxies that are stacked in the CANDELS fields, and the fraction of galaxies individually detected with *Herschel*. The bins where we do not detect any stacked signal are shown with a *gray* background (color figure online)

(FWHM) of the PSF, and $N = 61$ for *Spitzer* MIPS ($13 \times$ FWHM), as a substantial fraction of the *Spitzer* flux is located in the first Airy ring. Since the maps were reduced in a consistent way across all the CANDELS fields, we can safely merge together all the sources in a given bin, allowing us to go deeper while mitigating the effects of cosmic variance.

In parallel, we also stack the sources of the COSMOS UltraVISTA catalog in the wider but shallower FIR images. These stacked values are mostly used as consistency checks, since they do not offer any advantage over those obtained in the CANDELS fields: the shallow *Herschel* exposure is roughly compensated by the large area, but the mass completeness is much lower.

In the literature, a commonly used method consists of stacking only the undetected sources on the *residual* maps, after extracting sources brighter than a given flux threshold. This removes most of the contamination from bright neighbors, and thus lowers the confusion noise for the faint sources, while potentially introducing a bias that has to be corrected. Detected and stacked sources are then combined using a weighted average (as in, e.g., Magnelli et al. 2009). We prefer here to treat both detected and undetected sources homogeneously in order not to introduce any systematic error tied to either the adopted flux threshold or the details of the source extraction procedure. Although simpler, this procedure nevertheless gives accurate results when applied to our simulated images. Indeed, the contribution of bright neighbors is a random process: although it is clear that each source suffers from a varying level of contamination, statistically they are all affected in the same way. In other words, when a sufficient number of sources are stacked, the contribution of neighbors tends to average out to the same value μ_{gal} on all pixels, which is the contribution of galaxies to the Cosmic InfraRed Background (CIRB). But this is only true in the absence of galaxy clustering (Béthermin et al. 2010). When galaxies are clustered, there is an increased probability of finding a neighbor close to each stacked galaxy (Chary and Pope 2010), so that μ_{gal} will be larger toward the center of

the stacked image. Kurczynski and Gawiser (2010) proposed an alternative stacking technique (implemented by Viero et al. 2013, in the `SIMSTACK` code) that should get rid of most of this bias, and that consists of simultaneously fitting for the flux of all sources within a given volume (i.e., in a given redshift bin). It is however less versatile, and in particular it is not capable of measuring flux dispersions. Béthermin et al. (2015) also show that it can suffer from biases coming from the incompleteness of the input catalog.

The next step is to reduce each cube into a single image by combining the pixels together. There are several ways to do this, the two most common being to compute the mean or the median flux of all the cutouts in a given pixel. The advantage of the mean stacking is that it is a linear operation, thus one can exactly understand and quantify its biases (e.g., Béthermin et al. 2010). More specifically, it can be shown that the mean stacked value corresponds to the covariance between the input source catalog and the map (Marsden et al. 2009). Median stacking, on the other hand, has the nice property of naturally filtering out bright neighbors and catastrophic outliers and thus produces cleaner flux measurements. On the down side, we show in Appendix 2.8.2 that this measurement is systematically biased in a nontrivial way (see also White et al. 2007). Correcting for this bias requires some assumptions about the stacked flux distribution, e.g., the dispersion. Since this is a quantity we want to measure, we prefer to use mean over median stacking. An example of a mean stacked cutout from the SPIRE 250 μm images is shown in Fig. 2.5 (left). However, in two bins at low masses and high redshifts ($z = 1.5$ and $\log_{10}(M_*/\text{M}_\odot) = 9.75$, as well as $z = 3.0$ and $\log_{10}(M_*/\text{M}_\odot) = 10.25$), the mean stacked fluxes have signal to noise ratios that are too low and thus cannot be used, while the median stacked fluxes are still robustly measured. To extend our measurement of the Main Sequence SFR, we allow ourselves to use the median stacked fluxes in these particular bins only. This is actually a regime where we expect the median stacking to most closely measure the mean flux (see Appendix 2.8.1), hence this should not introduce significant biases. Lastly, we are interested in the *mode* of the Main Sequence, which is not strictly

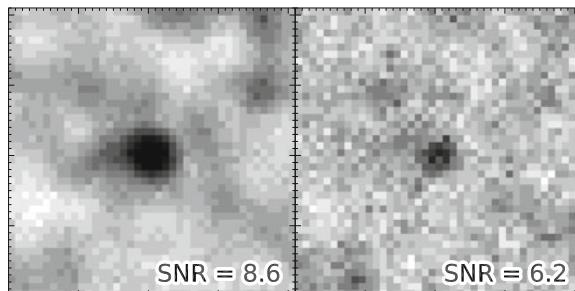


Fig. 2.5 Stack of 155 galaxies at $z = 3$ and $\log_{10}(M_*/\text{M}_\odot) = 11.3$ in the SPIRE 250 μm images. *Left* mean flux image, *Right* MAD dispersion image. Measuring the dispersion is more difficult than measuring the flux, since the signal is always fainter. 38 % of these galaxies are individually detected by *Herschel*, and only 25 % are detected in the SPIRE 250 μm channel

speaking the mean SFR we measure. We calibrated the difference between those two quantities with our simulations, and in all the following we refer to the SFR of the Main Sequence as the mode of the distribution. For example, for a log-normal distribution of $\sigma = 0.3$ dex, this difference is about 0.1 dex.

To measure the stacked flux, we choose to use PSF fitting in all the stacked bands, using a simple linear solver. In all fields, we use the same PSFs as those used to extract the photometry of individual objects, and apply the corresponding aperture corrections.⁸ This method assumes that the stacked image is a linear combination of: (1) a uniform background; and (2) the PSF of the instrument, since none of our sources is spatially resolved. The measured flux is then obtained as the best-fit normalization factor applied to the PSF that minimizes the residuals. In practice, we simultaneously fit both the flux and the background within a fixed aperture whose radius is 0.9 times the FWHM of the PSF. The advantage of this choice is that although we use less information in the fit, the background computed this way is more local, and the flux measurement is more robust against source clustering. Indeed, the amplitude of the clustering is a continuous function of angular distance: although a fraction of clustered sources will fall within a radius that is much smaller than the FWHM of the PSF and will bias our measurements no matter what, the rest will generate signal over a scale that is larger than the PSF itself, such that it will be resolved. Estimating the background within a small aperture will therefore remove the contribution of clustering coming from the largest scales.

We quantify the expected amount of flux boosting due to source physical clustering using our simulated maps. We show in Appendix 2.8.2 that it is mostly a function of beam size, i.e., there is no effect in the PACS bands but it can boost the SPIRE fluxes by up to 25 % at 500 μm . We also compare our flux extraction method to other standard approaches and show that it does reduce the clustering bias by a factor of 1.5 to 2.5, while also producing less noisy flux measurements. The value of $0.9 \times \text{FWHM}$ was chosen to get the lowest clustering amplitudes and flux uncertainties.

To obtain an estimate of the error on this measure, we also compute the standard deviation σ_{RES} of the residual image (i.e., the stacked image minus the fitted source) and multiply it by the PSF error scaling factor

$$\sigma_{\text{IMG}} = \sigma_{\text{RES}} \times \left(|P^2| - \frac{|P|^2}{N_{\text{pix}}} \right)^{-1/2}, \quad (2.4)$$

where N_{pix} is the number of pixels that are used in the fit, $|P|$ is the sum of all the pixels of the PSF model within the chosen aperture, and $|P^2|$ the sum of the squares of these pixels. This is the formal error on the linear fit performed to extract the flux (i.e., the square root of the diagonal element corresponding to the PSF in the covariance matrix), assuming that all pixels are affected by a similar uncorrelated Gaussian error of amplitude σ_{RES} . In practice, since the PSFs that we use are all sampled by roughly

⁸These PSFs are normalized to unit integral flux, but are truncated beyond a certain radius. Therefore aperture correction is necessary to recover the total flux. These corrections were derived by the GOODS–*Herschel* team using in flight observation of Vesta.

the same number of pixels (approximately twice the Nyquist sampling), this factor is always close to 0.5 divided by the value of the central pixel of the PSF. Intuitively, this comes from the fact that the error on the measured flux is the combination of the error on all the pixels that enter in the fit, weighted by the amplitude of the PSF. It is thus naturally lower than the error on one single pixel. In other words, using PSF fitting on these stacks allows for measuring fluxes that are twice as faint as those obtained when using only the central pixel of the image. Simple aperture photometry yields $\sigma_{\text{APER}} = \sigma_{\text{RES}} \times \left(\sqrt{N_{\text{pix}} + N_{\text{pix}}^2/N_{\text{bg}}} \right) / |P|$, where N_{bg} is the number of pixels used to estimate the background (e.g., within an annulus around the source). If N_{bg} is sufficiently large ($\gtrsim N_{\text{pix}}$), this error is lower than that obtained with our PSF fitting technique because the background is estimated independently of the flux. The price to pay is that this background is not local, hence the aperture flux will be most sensitive to clustering. Finally, if there is no clustering, PSF fitting will give the lowest errors of all methods, provided the full PSF is used in the fit. The optimal strategy is therefore always to use PSF fitting, varying the aperture within which the fit is performed depending on the presence of clustering.

To be conservative, we compute an alternative error estimate using bootstrapping: we randomly discard half of the sources, stack the remaining ones, measure the stacked flux, and repeat this procedure 100 times. The error σ_{BS} is then computed as the standard deviation of the measured flux in these 100 realizations, divided by $\sqrt{2}$, since we only work with half of the parent sample. Using our simulated images, we show in Appendix 2.8.3 that accurate error estimates are obtained by keeping the maximum error between σ_{IMG} and σ_{BS} . For the SPIRE bands, however, the same simulations show that both error estimates are systematically underestimated and need to be corrected by a factor of ~ 1.7 . We demonstrate in Appendix 2.8.3 that this comes from the fact that the error budget in the SPIRE bands is mostly generated by the random contribution of nearby sources rather than instrumental or shot noise. In this case, the error on each pixel is largely correlated with that of its neighbors, and the above assumptions do not hold.

We apply the above procedure to all the redshift and stellar mass bins of Fig. 2.4 and stack all the MIR to FIR images, from MIPS 24 μm to SPIRE 500 μm . Using the measured mean fluxes, we build effective SEDs⁹ in each bin, shown in Fig. 2.6. We fit the *Herschel* photometry with CE01 templates, leaving the normalization of each template free and keeping only the best-fit, and obtain the mean L_{IR} . As for the individual detections, we do not use the photometry probing rest-frame wavelengths below 30 μm (see Sect. 2.2.5). The MIPS 24 μm photometry is used as a check only. Converting the measured L_{IR} to SFR_{IR} with the Kennicutt (1998) relation and adding the mean observed SFR_{UV} (non-dust-corrected contribution), we obtain the mean total SFR in each bin.

⁹These SEDs are effective in the sense that they are not necessarily the SED of the average galaxy in the sample: they are potentially broadened by the range of redshifts and dust temperatures of the galaxies in the stacked samples. In practice, we checked that the broadening due to the redshift distribution is negligible, and the photometry is well fitted by standard galaxy templates, as can be seen in Fig. 2.6.

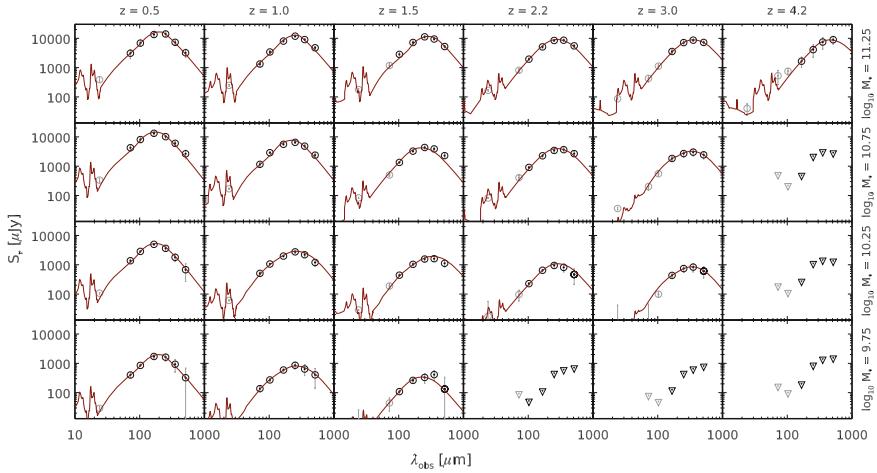


Fig. 2.6 Stacked SEDs of our star-forming mass-selected samples in bins of redshift (horizontally) and stellar mass (vertically). Stacked points are shown as *empty circles*, and the best-fit CE01 template is shown as the *solid red curve*. *Gray data points* were not used in the fit because they are probing rest-frame wavelengths below 30 μm. The data points have been corrected for the contribution of galaxy clustering (see Table 2.5). In the bins where the signal is too low (typically $< 5\sigma$), we plot 3σ upper limits as *downward triangles* (color figure online)

2.3.3 Measuring Flux Dispersion with Scatter Stacking

To measure the flux dispersion, we introduce a new method called “scatter stacking”. The idea is to come back to the pixel cube and build a *dispersion* image by measuring the scatter of each pixel around its average value. Stacked pixels away from the center measure the background fluctuations (the combination of photometric noise and random contribution from nearby sources), while pixels in the central region show enhanced dispersion due to flux heterogeneities in the stacked population, as in Fig. 2.5. In particular, if all the stacked sources had the same flux, the dispersion map would be flat.

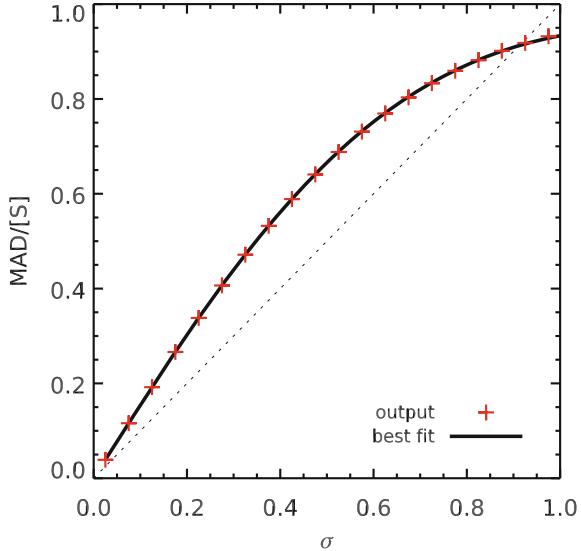
Again, this can be achieved in different ways. Computing the standard deviation of pixels is the most straightforward approach, but it suffers from similar issues as mean stacking with respect to bright neighbor contamination, in a more amplified manner because pixels are combined in quadrature. Our simulations also show that this method is not able to reliably measure high dispersion values. We thus use the median absolute deviation (MAD), which is more effective in filtering out outliers while providing the same information.

The MAD is formally defined as the half-width of the range that is centered on the median flux $\langle S \rangle$ and contains 50 % of the whole sample. In other words

$$\phi(\langle S \rangle + \text{MAD}) - \phi(\langle S \rangle - \text{MAD}) = \frac{1}{2}, \quad (2.5)$$

where ϕ is the cumulative probability distribution function of the flux.

Fig. 2.7 Median absolute deviation (MAD) computed by solving Eq. 2.5 numerically for a log-normal distribution of $\langle S \rangle = 1$ as a function of the chosen σ . The *solid line* is the best-fit of Eq. 2.7 to the numerical solutions, and the *dashed line* is the one-to-one correlation (color figure online)



To interpret this value in terms of more common dispersion indicators, we will convert the MAD to a log-dispersion σ assuming that fluxes follow a Gaussian distribution in $\log_{10}(S)$, i.e., a log-normal distribution in S . There are two reasons that justify this choice: (1) it allows for direct comparison of our measured dispersions to the data from literature that quote standard deviations of $\log_{10}(\text{SFR})$; and (2) log-normal distribution are good models for describing sSFR distributions in the regimes where we can actually detect individual sources (see, e.g., Rodighiero et al. 2011; Sargent et al. 2012; Gladders et al. 2013; Guo et al. 2013, and also Sect. 2.4.6). For this family of distributions,

$$\phi(S) = \frac{1}{2} \operatorname{erfc} \left(-\frac{\log_{10} \left(\frac{S}{\langle S \rangle} \right)}{\sqrt{2} \sigma} \right), \quad (2.6)$$

where erfc is the complementary error function. In this case there is no analytical solution to Eq. 2.5, but it can be solved numerically. It turns out that one can relate the MAD and $\langle S \rangle$ directly to σ (see Fig. 2.7) via the following equation, which was fit on the output of the numerical analysis¹⁰ (for $\sigma \in [0.05, 1.0]$ dex):

$$\frac{\text{MAD}}{\langle S \rangle} \simeq \frac{1.552 \sigma}{1 + 0.663 \sigma^2}, \quad (2.7)$$

¹⁰This analysis was performed with *Mathematica*.

with a maximum absolute error of less than 0.01. This relation can, in turn, be inverted to obtain σ . Defining the “normalized” median absolute deviation NMAD \equiv MAD/ $\langle S \rangle$, and only keeping the positive solution of Eq. 2.7, we obtain

$$\sigma \simeq \frac{1.171}{\text{NMAD}} \left(1 - \sqrt{1 - \left(\frac{\text{NMAD}}{0.953} \right)^2} \right). \quad (2.8)$$

Therefore, measuring the MAD allows us to obtain the intrinsic log-normal flux dispersion σ of the stacked sample. To do so, we perform PSF fitting on the squared images (since the dispersion combines quadratically with background noise) and fit a constant background noise plus the square of the PSF on all the pixels within a fixed radius of $0.6 \times \text{FWHM}$. Here we do not use the same $0.9 \times \text{FWHM}$ cut as for the flux extraction, since the MAD does not fully preserve the shape of the PSF when its pixels are low in signal to noise (see below). We thus restrain ourselves to a more central region to prevent being dominated by these faint pixels. Again, this value was chosen using the simulated maps in order to produce the least biased and least uncertain measurements.

Even then, the dispersion measured with this method is slightly biased toward higher values, but this bias can be quantified and corrected in a self-consistent way with no prior information using Monte Carlo simulations. For each source in the stack, we extract another cutout at a random position in the map. We then place a fake source at the center of each random cutout, whose flux follows a log-normal distribution of width σ_{MC} , and with a mean flux equal to that measured for the real sources. We apply our scatter stacking technique to measure the dispersion on the resulting mock flux cube, and compare it to σ_{MC} . We repeat this procedure for different values of σ_{MC} (from 0.1 to 0.7 dex), and derive the relation between the intrinsic and measured dispersion. Examples are shown in Fig. 2.8. To average out the measurement error, we repeat this procedure 20 times for each value of σ_{MC} . In practice, this correction is mostly negligible, except for the lowest measured mass bins at any redshift where it reaches up to 0.1 dex.

2.3.4 SFR Dispersion from Scatter Stacking

The procedure described in the previous section allows us to measure the log-normal flux dispersion, while we are interested in the dispersion in SFR.

The first step is to obtain the $\log_{10}(L_{\text{IR}})$ dispersion σ_{IR} . Using detected sources, we observe that the dispersion in L_{IR} of a population of galaxies having the same flux at a given redshift depends on the rest-frame wavelength probed, as illustrated in Fig. 2.9. The data points in this figure are produced by looking at multiple bins of redshift, and measuring the scatter of the correlation between L_{IR} , measured by fitting all available FIR bands, and the flux in each *Herschel* band converted to rest-frame luminosity ($v L_v$). By spanning a range of redshift, the five *Herschel* bands will

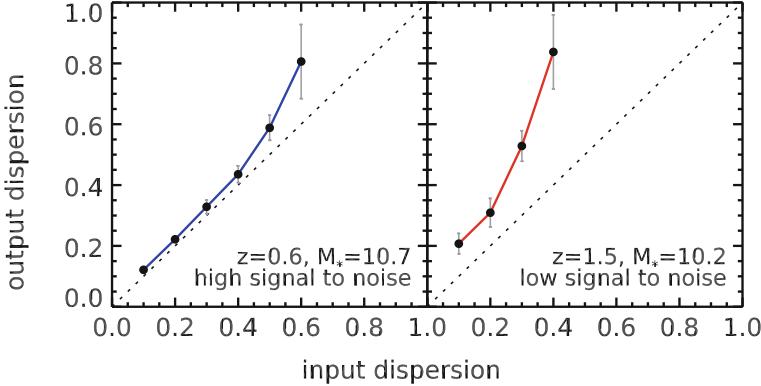


Fig. 2.8 Correction procedure for the measured dispersion. Each point is a simulated dispersion measurement with a different input value. *Error bars* show the scatter observed among the 20 realizations. The *dashed line* shows the one-to-one relation. The *plots* display two examples of simulated dispersions for the PACS 100 μm band, at $z = 0.6$ for $M_* = 3 \times 10^{10}$ (*left panel*), and at $z = 1.5$ for $M_* = 2 \times 10^{10} M_\odot$ (*right panel*). These bins were chosen to illustrate the two regimes of high and low signal to noise, respectively (color figure online)

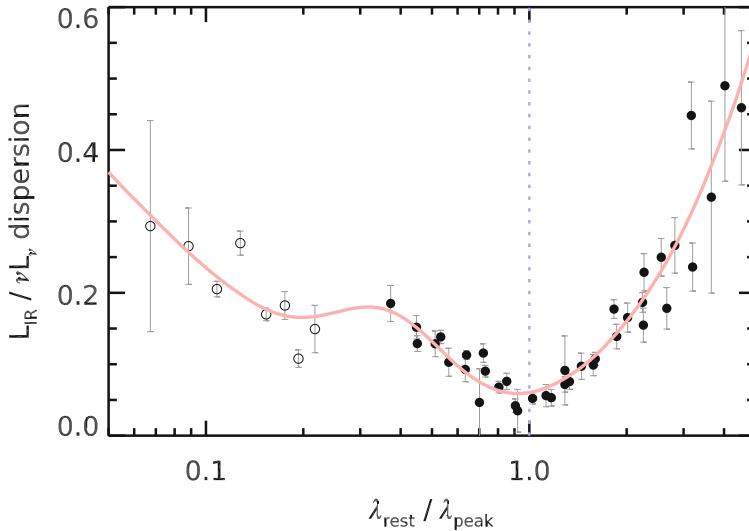


Fig. 2.9 Dispersion of the ratio L_{IR} / vL_v as a function of wavelength in bins of redshift and for the five *Herschel* bands in the four CANDELS fields. The wavelength is normalized here to the “peak” wavelength, where the FIR SED in vL_v reaches its maximum (calibrated from our stacked SEDs, Fig. 2.6). The L_{IR} is computed by fitting all the available *Herschel* bands (we require a minimum of three) together with CE01 templates, while vL_v is the flux in a single *Herschel* band converted to rest-frame luminosity. *Open symbols* denote measurements where vL_v comes from MIPS 24 μm . *Error bars* come from simple bootstrapping. The contribution of photometric errors was statistically removed. The *red line* shows a fit to the data points to guide the eye (color figure online)

probe a varying range of rest-frame wavelengths, allowing us to observe the behavior of the L_{IR} scatter with rest-frame wavelength. The smaller dispersions are found at wavelengths close to the peak of the SED, in which case the dispersion drops as low as 0.05 dex. This is due to galaxies showing a variety of effective dust emissivities and temperatures that both influence the shape of the FIR SED, respectively longward and shortward of the peak.

Therefore, to obtain σ_{IR} , we simply measure the flux dispersion of the *Herschel* band that is the closest to the peak. We thus first measure the peak wavelength λ_{peak} from the stacked SEDs (Fig. 2.6), and interpolate the measured log-normal flux dispersions at λ_{peak} . By construction, this also tends to select *Herschel* measurements with the highest signal to noise ratio.

One then has to combine the dispersion in L_{IR} with that in L_{UV} , since we combine both tracers to derive the total SFR. This is not straightforward, as the two quantities are not independent (i.e., at fixed SFR, more attenuated objects will have higher L_{IR} and lower L_{UV}). In particular, we see on individual detections that the dispersion of $\text{SFR} = \text{SFR}_{\text{IR}} + \text{SFR}_{\text{UV}}$ is actually *lower* than that of SFR_{IR} alone.

To address this issue, we choose to work directly on “SFR stacks”. First, we use our observed FIR SEDs to derive L_{IR} monochromatic conversion factors for all bands in each of our redshift and stellar mass bins. Second, in each stacked bin, we convert all cutouts to SFR_{IR} units, using the aforementioned conversion factor and the Kennicutt (1998) relation. Third, we add to each individual cutout an additional amount of SFR equal to the non-dust-corrected SFR_{UV} , as a centered PSF. Finally, to correct for the smearing due to the width of the redshift and mass bins, we also use our observed relation between mass, redshift, and SFR (given below in Eq. 2.9) and normalize each cutout to the reference mass and redshift of the sample by adding $\text{SFR}_{\text{MS}}(z_{\text{ref}}, M_{*,\text{ref}}) - \text{SFR}_{\text{MS}}(z, M_*)$. This last step is a small correction: it reduces the measured dispersion by only 0.02 to 0.03 dex.

We stack these cutouts and again run the dispersion measurement procedure, including the bias correction. Interpolating the measured dispersions in the five *Herschel* bands at λ_{peak} as described earlier, we obtain σ_{SFR} . As expected, the difference between the flux dispersion at the peak of the SED and the SFR dispersion is marginal, except for the lowest mass bins where it can reach 0.05 dex. This is mainly caused by the increasing contribution of the escaping UV light to the total SFR, as $\text{SFR}_{\text{IR}}/\text{SFR}_{\text{UV}}$ approaches unity in these bins.

A remaining bias that we do not account for in this study is the impact of errors on the photo-zs and stellar masses. As pointed out in Sect. 2.2.4, the measured few percent accuracy on the photo-zs only applies to the bright sources, and we do not know the reliability of the fainter sources. We measure statistical uncertainties on both these quantities, but this does not take systematic errors coming from the library or gaps in the photometry into account. Intuitively, one can expect these errors to increase the dispersion, but this would be true only if the true error was purely random. It could be that our SED fitting technique is too simplistic in assuming a universal IMF, metallicity, and SFH functional form for all galaxies, and as such erases part of

the diversity of the population. This could in turn *decrease* the measured dispersion (see discussion in Reddy et al. 2012). It is therefore important to keep in mind that our measurement is tied to the adopted modeling of stellar mass.

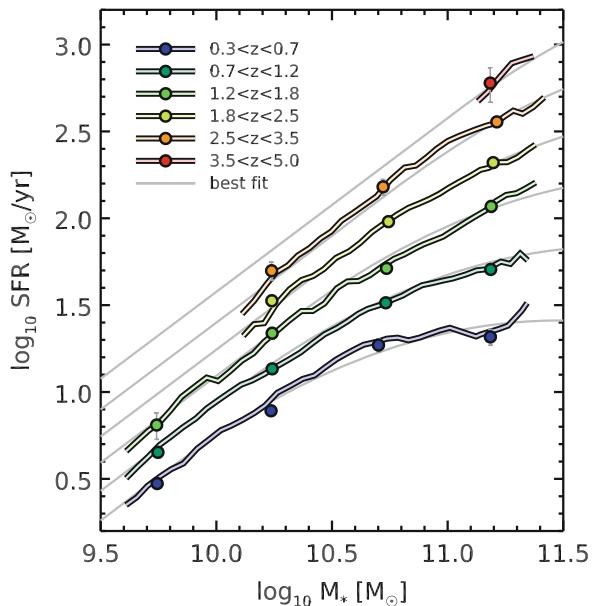
2.4 Results

2.4.1 The SFR of Main-Sequence Galaxies

The first results we present concern the evolution of the Main Sequence with redshift, as well as its dependence on stellar mass. In Sect. 2.4.2 we start by describing the redshift evolution of the $\text{sSFR} \equiv \text{SFR}/M_*$, and we then address the mass dependence of the slope of the Main Sequence in Sect. 2.4.3.

These results are summarized in Fig. 2.10 where, for the sake of visualization, we also run our full stacking procedure on sliding bins of mass, i.e., defining a fine grid of M_* and selecting galaxies within mass bins of constant logarithmic width of 0.3 dex. The data points are not independent anymore, since a single galaxy is included in the stacked sample of multiple neighboring points, but this allows us to better grasp the evolution of the Main Sequence with mass. These “sliding averages” of the SFR are displayed as solid colored lines, while the points obtained with regular mass bins are shown as filled circles.

Fig. 2.10 Evolution of the average SFR of star-forming galaxies with mass and redshift. Our results from stacking are shown as *colored filled circles*, the colors corresponding to the different redshifts as indicated in the legend. We complement these measurements by stacking sliding bins of mass (see text) for visualization purposes only to better grasp the mass dependence of the SFR. In the background, we show as *light gray curves* our best-fit relation for the Main Sequence (Eq. 2.9) (color figure online)



By fitting these points (filled circles only), we parametrize the SFR of main-sequence galaxies with the following formula, defining $r \equiv \log_{10}(1 + z)$ and $m \equiv \log_{10}(M_*/10^9 M_\odot)$:

$$\begin{aligned} \log_{10}(\text{SFR}_{\text{MS}}[\text{M}_\odot/\text{yr}]) = & m - m_0 + a_0 r \\ & - a_1 [\max(0, m - m_1 - a_2 r)]^2, \end{aligned} \quad (2.9)$$

with $m_0 = 0.5 \pm 0.07$, $a_0 = 1.5 \pm 0.15$, $a_1 = 0.3 \pm 0.08$, $m_1 = 0.36 \pm 0.3$ and $a_2 = 2.5 \pm 0.6$. The choice of this parametrization is physically motivated: we want to explicitly describe the two regimes seen in Fig. 2.10 and explored in more detail in Sect. 2.4.3, namely a sequence of slope unity whose normalization increases with redshift (first terms), and a “bending” that vanishes both at low masses and high redshifts (last term). The precise functional form however is arbitrary, and was chosen as the simplest expression that accurately reproduces the bending behavior. This SFR will be used in the following as a reference for the locus of the Main Sequence.

2.4.2 Redshift Evolution of the sSFR: The Importance of Sample Selection and Dust Correction

We show in Fig. 2.11 the evolution of sSFR ($\equiv \text{SFR}/M_*$) as a function of both redshift and stellar mass. Our results at $z \leq 3$ are in good agreement with previous estimates from the literature, showing the dramatic increase of the sSFR with redshift. At $z = 4$, we still measure a rising sSFR, reaching 5 Gyr^{-1} , i.e., a mass doubling timescale of only 200 Myr.

At this redshift, however, our measurement is substantially higher than UV-based estimates (Daddi et al. 2009; Stark et al. 2009). More recent results (Bouwens et al. 2012; Stark et al. 2013; González et al. 2014) seem to be in better agreement, but it is important to keep in mind that these studies mostly focus on relatively low mass galaxies, i.e., typically $3 \times 10^9 M_\odot$. Therefore the quoted sSFR values only formally apply to galaxies in this range, i.e., to galaxies a factor of 10 to 100 times less massive than those in our sample. Extrapolating their measurements to match the mass range we are working with requires that we know the slope of the sSFR– M_* relation. In their study, Bouwens et al. (2012) measured this slope from $M_* = 10^8$ to $10^{10} M_\odot$ at $z = 4$ and found it to be around -0.27 . Assuming that this holds for all masses, this means that we should reduce the sSFR by about 0.4 dex to be able to compare it directly to our result. This is illustrated by the gray arrow in Fig. 2.11.

Previous observations of the sSFR “plateau” (Daddi et al. 2009) could be the consequence of two key issues. First, selection effects: these studies are based either on Lyman break galaxies (LBGs) or rest-frame FUV-selected samples that, while less prone to lower redshift contaminants, are likely to miss highly attenuated and thus highly star-forming galaxies. Our sample is mass-complete, so we do not suffer from such biases. Second, failure of dust extinction correction: UV-based SFR estimates

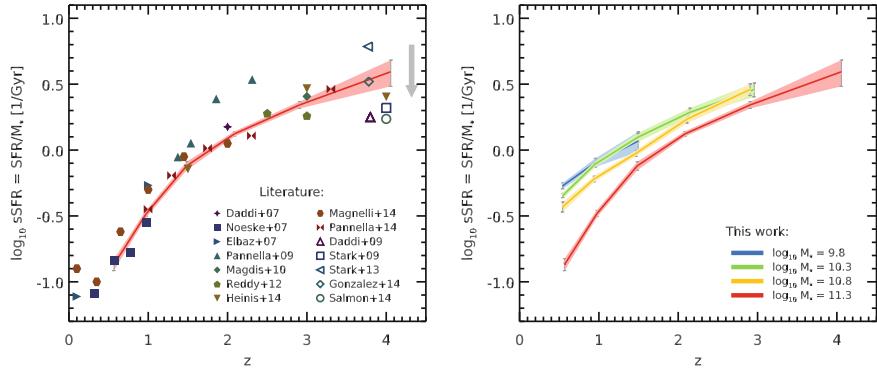


Fig. 2.11 Evolution of the average sSFR of star-forming galaxies with redshift. *Left* comparison of our results at $M_* = 2 \times 10^{11} M_\odot$ (red curve) to published values in the literature (filled and open symbols). Filled symbols compile various results that were derived from mass-complete samples with SFRs computed either from the IR (Daddi et al. 2007; Noeske et al. 2007; Elbaz et al. 2007; Magdis et al. 2010; Reddy et al. 2012; Heinis et al. 2014; Magnelli et al. 2014; Pannella et al. 2015) or the radio (Pannella et al. 2009, 2015). When possible, these were rescaled to a common stellar mass of $2 \times 10^{11} M_\odot$ using the corresponding published SFR– M_* relations. Results from stacking have been corrected by -0.1 dex to reach the mode of the Main Sequence (see discussion in Sect. 2.3.2). Open symbols show results from the literature that make use of the Lyman break selection technique (LBGs) and where the SFRs are obtained from the UV light alone (Daddi et al. 2009; Stark et al. 2009, 2013; González et al. 2014; Salmon et al. 2015). These samples are mostly composed of galaxies of much lower stellar mass, typically $3 \times 10^9 M_\odot$, so the extrapolation to $10^{11} M_\odot$ is more uncertain. We therefore simply quote the published values. The gray arrow shows how the open symbols would move if we were to apply a mass correction assuming the $z = 4$ Main Sequence slope of Bouwens et al. (2012). When necessary, data from the literature have been converted to a Salpeter IMF. *Right* same figure showing our other stacked mass bins with different colors (color figure online)

are plagued by uncertainties in dust attenuation. Most studies rely on observed correlations between UV SED features and dust attenuation that are calibrated in the local Universe, such as the $\text{IRX}-\beta$ relation (Meurer et al. 1999). Recent studies tend to show that these correlations are not universal and evolve with redshift, possibly due to subsolar metallicity (Castellano et al. 2014), ISM conditions, or dust geometry (Oteo et al. 2013; Pannella et al. 2015).

2.4.3 Mass Evolution of the SFR and Varying Slope of the Main Sequence

It is also worth noting the dependence of the SFR on stellar mass from Fig. 2.10. Low mass bins ($M_* < 3 \times 10^{10} M_\odot$) are well fit with a slope of unity. Many studies have reported different values of this slope, ranging from 0.4 to unity (Brinchmann et al. 2004; Noeske et al. 2007; Elbaz et al. 2007; Daddi et al. 2007; Santini et al. 2009;

Pannella et al. 2009; Rodighiero et al. 2011). A slope of unity can be interpreted as a signature of the universality of the star formation process, since it implies a constant star formation timescale $\tau \equiv 1/\text{sSFR}$ at all stellar masses, with $M_*(t) \sim \exp(t/\tau)$. As suggested by Peng et al. (2010), it is also a necessary ingredient for explaining the observed shape invariance of the stellar mass function of star-forming galaxies.

We find however that the SFR of the highest mass bin ($M_* \sim 2 \times 10^{11} M_\odot$) falls systematically below the value expected for a linear relation, effectively lowering the high mass slope of the SFR– M_* relation to 0.8 at high redshift, down to an almost flat relation at $z = 0.5$. Other studies obtain similar “broken” shapes for the SFR– M_* sequence (Rodighiero et al. 2010; Whitaker et al. 2012; Magnelli et al. 2014). Our results are also in very good agreement with Whitaker et al. (2014), who used a very similar approach, albeit only using MIPS 24 μm for stacking.

The reason for this bending of the slope is still unknown. Abramson et al. (2014) showed that the relation between the disk mass M_{disk} and SFR has a slope close to one with no sign of bending at $z \simeq 0$, suggesting that the bulge plays little to no role in star formation. We investigate if this explanation holds at higher redshifts in Chap. 5.

2.4.4 Mass Evolution of the SFR Dispersion Around the Main Sequence

We present in Fig. 2.12 the evolution of the measured SFR dispersion σ_{SFR} as a function of both redshift and stellar mass. We show our measurements only from stacking *Herschel* bands. *Spitzer* MIPS is more sensitive and thus allows measurements down to lower stellar masses, but it is less robust as an SFR indicator. This is mostly an issue at $z \simeq 2$, where the 24 μm is probing the rest-frame 8 μm . Elbaz et al. (2011) have shown that the 8 μm luminosity L_8 correlates very well with L_{IR} (0.2 dex scatter), except for starburst galaxies. Inferring SFR from 8 μm thus has the tendency to erase part of the starburst population, effectively reducing the observed SFR dispersion. We checked that our results are nevertheless in good agreement between MIPS and *Herschel*, with MIPS derived dispersions being smaller on average by only 0.03 ± 0.02 dex.

As a sanity check, we also show an estimation of σ_{SFR} from individual *Herschel* detections. We select all galaxies in our *Herschel* sample that fall in a given bin of redshift and mass, and compute their offset from the Main Sequence $R_{\text{SB}} \equiv \text{SFR}/\text{SFR}_{\text{MS}}$, where SFR_{MS} is the average SFR of “Main Sequence” galaxies given in Eq. 2.9. Following Elbaz et al. (2011), we call this quantity the “starburstiness”. Because of the sensitivity of *Herschel*, this sample is almost never complete, and is biased toward high values of R_{SB} : since this sample is SFR selected, all the galaxies at low mass are starbursts. To avoid completeness issues, we remove the galaxies that have $R_{\text{SB}} < 1$, i.e., galaxies that are below the Main Sequence, and compute the 68th percentile of the resulting R_{SB} distribution. By construction, this

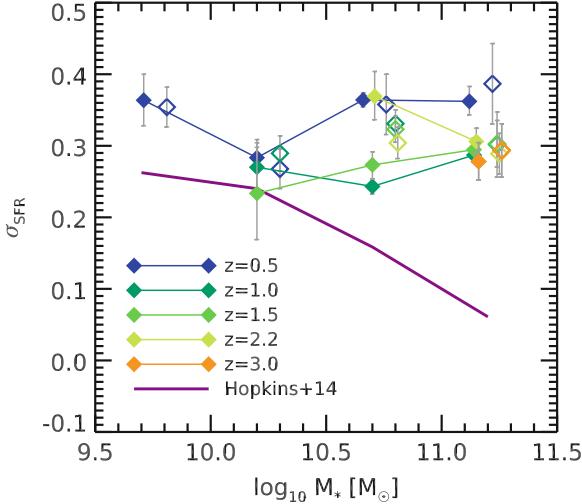


Fig. 2.12 Evolution of the $\log_{10}(\text{SFR})$ dispersion as a function of both redshift and stellar mass. Each color is showing a different redshift bin. *Filled symbols* show the result of scatter stacking, while *open symbols* show the dispersion estimated from individual *Herschel* detections above the Main Sequence (see text). The *open symbols* have been shifted up by 0.1 dex in mass for clarity. *Errors* are from bootstrapping in all cases. We compare these to the typical scatter of the SFHs in the numerical simulation of Hopkins et al. (2014) shown as a *solid purple line* (color figure online)

value does not need to be corrected for the width of the redshift and mass bins. However, it is only probing the *upper* part of the SFR– M_* correlation, while the stacked measurements also take undetected sources below the sequence into account. In spite of this difference, the values obtained are in very good agreement with the stacked values. There is a tendency for these to be slightly higher by 0.03 dex on average, and this could be due to uncertainties in the individual SFR measurements. We conclude that the SFR distributions must be quite symmetric. This however does not rule out a “starburst” tail, i.e., a subpopulation of galaxies with an excess of star formation. Indeed, simulating a log-normal distribution of R_{SB} with a dispersion of 0.3 dex and adding 3 % more sources with an excess SFR of 0.6 dex (following Sargent et al. 2012) gives a global dispersion measured with MAD of 0.309 dex, while the 68th percentile of the $R_{\text{SB}} > 1$ tail is 0.319 dex, a difference of only 0.01 dex, which is well within the uncertainties.

Implications for the Existence of the Main Sequence

Probably the most striking feature of Fig. 2.12 is that σ_{SFR} remains fairly constant over a large fraction of the parameter space we explore, only increasing for the lowest redshift bin and at high stellar masses. This increase is most likely caused by the same phenomenon that bends the sequence at high stellar mass (see Sect. 2.4.2, e.g., a substantial population of bulge-dominated objects that blur the correlation). On average, *Herschel* stacking thus gives $\sigma_{\text{SFR}} = 0.30^{+0.06}_{-0.06}$ dex, with a random error

of 0.01 dex, and can be considered almost constant. Doing the same analysis in COSMOS UltraVISTA consistently yields $\sigma_{\text{SFR}} = 0.33^{+0.03}_{-0.03}$ dex, with a random error of 0.01 dex, showing that this result is not tied to specifics of our input H -band catalogs.

More importantly, this value of 0.3 dex means that, at a given stellar mass, 68 % of actively star-forming galaxies have the same SFR within a factor of two. This confirms the existence of the Main Sequence of star-forming galaxies for all of the stellar mass range probed here and up to $z = 3$, i.e., over more than 80 % of the history of the universe. A more illustrative picture is shown later in Fig. 2.16, and we discuss the implication of this finding in Sect. 2.5.1.

2.4.5 Contribution of the Main Sequence to the Cosmic SFR Density

Using our stacked SFRs, we can infer the contribution of each of our stacked bins to the cosmic star formation rate density ρ_{SFR} (Lilly et al. 1996; Madau et al. 1996). To this end, we use the stellar mass functions described in Sect. 2.2.7 and extrapolate our results to obtain a prediction for the total ρ_{SFR} , assuming a main-sequence slope of unity for low mass galaxies, and integrating the mass functions down to $M_* = 3 \times 10^9 M_\odot$ (i.e., $\sim 0.03 M^*$). The results of this analysis are presented in Figs. 2.13 and 2.14, and compared to the literature compilation of Madau and Dickinson (2014) (where luminosity functions are integrated down to $0.03 L^*$, and should thus match our measurements to first order).

We also infer the total stellar mass density ρ_* by integrating ρ_{SFR} as a function of time. At each time step, we create a new population of stars whose total mass is given by ρ_{SFR} , and let it evolve with time. We account for stellar mass loss using the Salpeter (1955) IMF to model the population, allowing stars to evolve and die assuming the stellar lifetimes of Bressan et al. (1993) for solar metallicity. As stars die, some of the matter is left in the form of stellar remnants that are traditionally also included in ρ_* , i.e., neutron stars and white dwarfs. We parametrize the masses of these remnants following Prantzos and Silk (1998). The contribution of these remnants continuously rises with time to reach about 12 % of the stellar mass at $z = 0$. The result is presented in Fig. 2.15.

One can see from these figures that individual *Herschel* detections in the ultra-deep GOODS and CANDELS surveys (orange dash-dotted line) unveil about 50 % of the star formation budget below $z = 2$, but less than 10 % at $z = 4$. In total, and over the redshift range probed here, these galaxies have built 49 % of the mass of present day stars, and are thus to be considered as major actors in the stellar mass build up in the Universe. Stacking (purple line) allows us to go much deeper, since we reach almost 100 % of the total ρ_{SFR} at $z < 2$, and accounts for 83 % of the mass

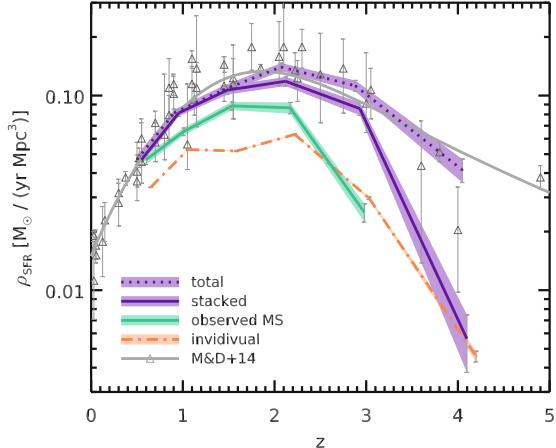


Fig. 2.13 Evolution of the cosmic star formation rate density ρ_{SFR} with redshift. The *orange dash-dotted line* traces the SFR density inferred from individual *Spitzer* MIPS (for $z < 1.5$) and *Herschel* detections alone. The *solid purple line* represents the contribution of stacked sources with significant signal ($>5\sigma$), and the *dotted line* is the extrapolation of the stacked SFR down to $M_* = 3 \times 10^9 M_\odot$ assuming constant sSFR and using the mass functions of Fig. 2.3. The *green line* shows the fraction of ρ_{SFR} in regimes where we have probed the existence of the Main Sequence. The *lines* are slightly offset in redshift for clarity. *Light shaded regions* in the background show the corresponding 1σ statistical errors. We compare these to the literature compilation of Madau and Dickinson (2014), shown as *open triangles*, with their best-fit plotted as a *solid gray line* (color figure online)

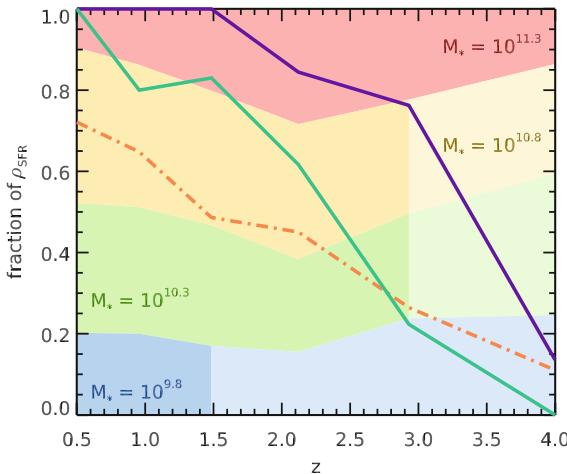


Fig. 2.14 Contribution to the total ρ_{SFR} (purple dotted line in Fig. 2.13) as a function of redshift for the various sub-samples of Fig. 2.13. *Background colors* represent how galaxies of different stellar masses contribute to the total ρ_{SFR} (from top to bottom: $\log_{10}(M_*/M_\odot) = 11.2, 10.8, 10.2$ and 9.8), *lighter colors* indicating regions where ρ_{SFR} is extrapolated. The *colored lines* are defined as in Fig. 2.13: the *solid purple line* shows the contribution of stacked sources with significant signal, the *green line* shows the contribution of galaxies in the regimes where we have probed the existence of the Main Sequence, and the *orange line* is the contribution of individually detected FIR sources (color figure online)

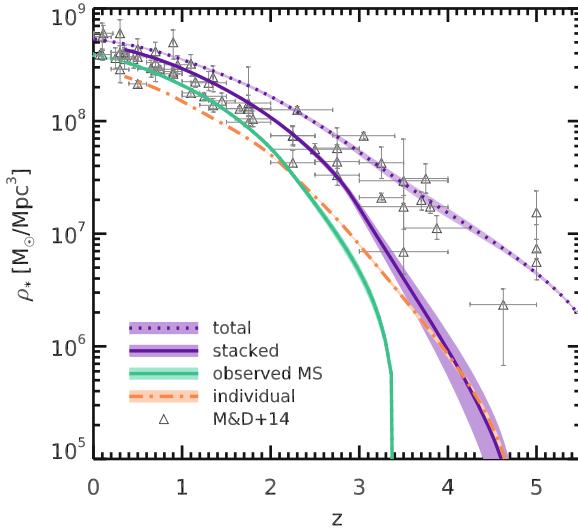


Fig. 2.15 Predicted evolution of the cosmic stellar mass density ρ_* with redshift. The *lines* show the inferred mass density by extrapolating our stacked SFRs down to $M_* = 3 \times 10^9 M_\odot$ and out to $z = 6$ using the trend from Madau and Dickinson (2014) and integrating as a function of time. Stellar lifetimes are accounted for, and the mass of stellar remnants is included in ρ_* (see text). Colors are the same as in Fig. 2.13: the *solid purple line* shows the contribution of stacked sources with significant signal, the *green line* shows the contribution of galaxies in the regimes where we have probed the existence of the Main Sequence, and the *orange line* is the contribution of individually detected FIR sources. *Shaded regions* in the background show the corresponding 1σ statistical errors. We compare these results to the literature compilation of Madau and Dickinson (2014) shown as *open triangles* (color figure online)

of present day stars. Extrapolating our observations to lower stellar masses using the mass functions and to $z = 0$ using the best-fit ρ_{SFR} of Madau and Dickinson (2014), we obtain an estimate of the total amount of star formation in the Universe (purple dotted line). Integrating it to $z = 0$ gives $\rho_*(z = 0) = (5.3 \pm 0.1) \times 10^8 M_\odot \text{ Mpc}^{-3}$, consistent with the value reported by Cole et al. (2001) and Bell et al. (2003) (our error estimate being purely statistical).

Although the range in redshift and stellar mass over which we are able to probe the existence of the Main Sequence is limited, it nevertheless accounts for 66 % of the mass of present day stars. This number climbs up to 73 % if we take other studies that have observed a tight correlation down to $z = 0$ (Brinchmann et al. 2004) into account. We show in the next section that starburst galaxies make up about 15 % of the SFR budget in all the redshift and mass bins that we probe with individual detections, and that the remaining fraction is accounted for by a single population of “Main Sequence” galaxies. Subtracting these 15 % from the above 73 %, we can say that at least 62 % of the mass of present day stars was formed by galaxies belonging to the Main Sequence. In other words, whatever physical phenomenon shapes the Main Sequence is the dominant mode of star formation in galaxies.

2.4.6 Quantification of the Role of Starburst Galaxies and the Surprising Absence of Evolution of the Population

An Overview of the Main Sequence

We summarize the previous results in Fig. 2.16. Here we show the distribution of individually detected galaxies on the SFR– M_* plane at various redshifts. The locus of our stacked SFRs (solid blue lines) may not appear to coincide with the average of the detections because of the SFR detection limit, symbolized by the horizontal

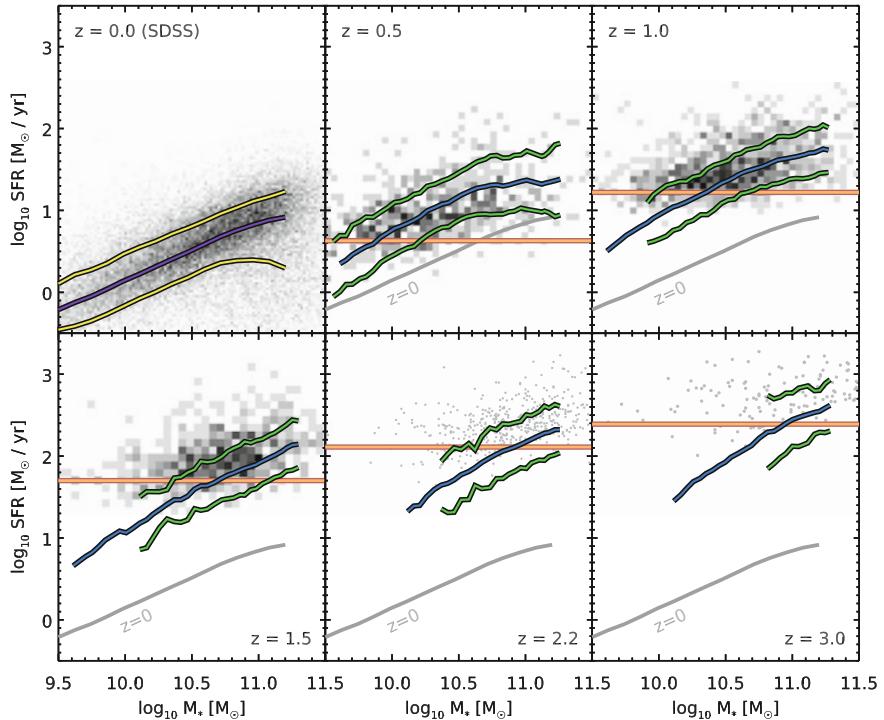


Fig. 2.16 Compilation of both detections and stacking results on the SFR– M_* plane for the CANDELS fields. The *top left panel* shows the results obtained with the Sloan Digital Sky Survey (SDSS) in the local Universe, as presented in Elbaz et al. (2007), while each subsequent panel displays our result for increasing redshifts. The *blue line* shows the average stacked SFR (Sect. 2.4.2), and the *green lines* above and below show the 1σ dispersion obtained with scatter stacking (Sect. 2.4.4). Both of these were performed on sliding bins of mass for the sake of visualization, and for this figure only. The SFR detection limit of each sample is indicated with a *solid orange line*. We also show the sliding median and percentiles of the SDSS distribution with *purple* and *yellow* lines, respectively, to emphasize that both the SFR tracer and the sample selection are different (see text). This correlation, observed in the local Universe, is reproduced as a *gray line* on each panel. The density of individual detections is shown in *gray scale* in the background, except for the two highest redshift bins where we show the individual galaxies as *gray filled circles* (color figure online)

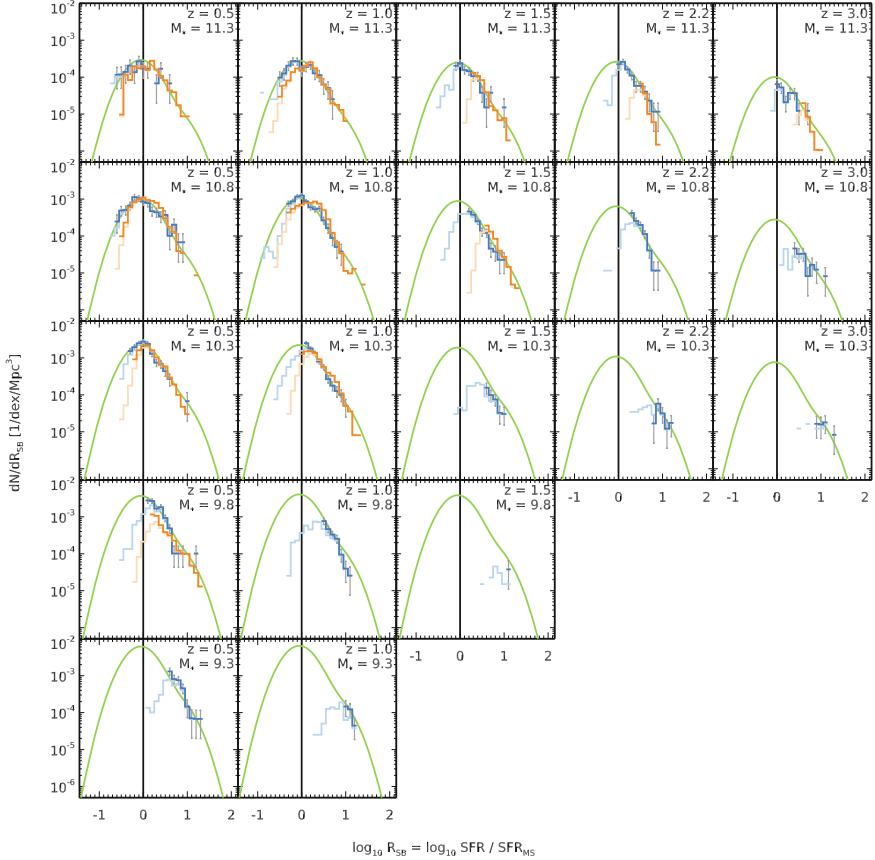


Fig. 2.17 Starburstiness ($R_{\text{SB}} \equiv \text{SFR}/\text{SFR}_{\text{MS}}$) histograms of individual *Herschel* and *Spitzer* MIPS (for $z < 1.2$) detections in each of our redshift and stellar mass bins. The *blue* and *orange* lines correspond to the counts in the CANDELS and COSMOS 2 deg² fields, respectively. We also show the incomplete counts in *light colors* in the background. The *green curve* shows our best-fit to the combined data set, and is the same for all bins except for the normalization, which is set by the mass function. The *black vertical line* shows the locus of the Main Sequence. *Error bars* indicate Poissonian noise (color figure online)

dashed line. We discuss later on (in Fig. 2.17) the distribution of these detected sources and confirm that the stacks and the detections are in perfect agreement.

We also show for reference the $z = 0$ sample taken from the Sloan Digital Sky Survey (SDSS DR4, Brinchmann et al. 2004) as presented in Elbaz et al. (2007). In this data set, actively star-forming galaxies are selected according to their rest-frame $U - V$ colors only (i.e., what is usually referred to as the “blue cloud”), and SFRs are estimated from the dust-corrected H _{α} line. These differences of observables and sample selection are likely to affect the shape of the Main Sequence. In particular, it is clear that the bending at high mass is less pronounced in the SDSS sample, and

this is likely due to the selection. Therefore, the comparison of this $z = 0$ data set with our own sample should be done with caution. This nevertheless resembles our own results quite closely and allows us to paint a consistent picture from $z = 0$ to $z = 3$.

“Starburstiness” Distributions

Although the depth of the *Herschel* surveys is limited, there is still a lot to be learned from the individually detected sources, in particular for the bright starburst galaxies. Now that we have a good definition of the Main Sequence, we can study these galaxies in more detail. Rodighiero et al. (2011) have used similar data in COSMOS and found that the distribution of star-forming galaxies on and off the Main Sequence is bimodal: a population of normal star-forming galaxies shapes the Main Sequence with a log-normal distribution of sSFR at a given mass, while another smaller population of “starbursts” boosts the high sSFR counts. Their work was restricted to $z = 2$ because of the BzK selection, so we want to extend it here to a mass-complete sample over a wider range of redshifts to see what we can learn about the starburst population.

In Fig. 2.17 we show the distributions of “starburstiness” R_{SB} , defined as the ratio between the actual SFR of each galaxy and SFR_{MS} , the SFR they would have if they were exactly following the Main Sequence defined in Eq. 2.9. We analyze these distributions in the same bins that were used for stacking, to make the comparison simpler. Since the CANDELS fields have a relatively similar depth, we group them together into a single distribution (blue curve), and following Rodighiero et al. (2011) we keep the COSMOS UltraVISTA sources apart (orange curve) where the catalog is mass-complete.

As was the case for the stellar mass functions discussed in Sect. 2.2.7, these distributions are affected by completeness issues. To correct this, we use a procedure very similar to that used for the mass functions. We assume that the total L_{IR} of a galaxy at a given redshift is well modeled from the rest-frame monochromatic luminosity in each *Herschel* band by a power law plus a Gaussian scatter in logarithmic space. In each bin of redshift and stellar mass, we select galaxies that are detected in at least three *Herschel* bands, fit this power law and measure the dispersion as in Fig. 2.2. In this case, this dispersion is mainly due to differences in dust temperature, and is found to be minimal at the peak of the FIR emission (see Fig. 2.9). Then, for each *Herschel* band, in each redshift and mass bin, we then generate a mock population of 10 000 galaxies with uniform redshift and mass distribution within the bin and attribute a starburstiness with uniform probability to each mock galaxy. We multiply this starburstiness by the SFR_{MS} of the galaxy computed from its redshift and mass, subtract the average observed SFR_{UV} in this bin (we assume no scatter in SFR_{UV} for simplicity), convert the remaining SFR_{IR} into L_{IR} , and finally the L_{IR} into monochromatic luminosity in the considered *Herschel* band, adding a random logarithmic scatter whose amplitude is given by the dispersion measured earlier. The completeness is then given as the fraction of mock galaxies with simulated monochromatic luminosity larger than the limiting luminosity at the corresponding redshift.

Since we include in our sample all sources provided that they are detected in at least one *Herschel* band, we then take the maximum completeness among all bands.

In Fig. 2.17, raw incomplete counts are shown as light curves in the background, and corrected counts are shown as darker lines. Error bars indicate Poisson noise and for clarity are only shown for the CANDELS counts.

In all fields, the low R_{SB} counts at $z < 1.2$ come from MIPS derived SFRs. Since the MIPS imaging in COSMOS UltraVISTA is only half as deep as the deepest CANDELS fields (see Sect. 2.2.3 and Table 2.1), the two curves probe almost similar ranges of R_{SB} . At $z \geq 1.2$ (i.e., starting from the bin at $z = 1.5$) MIPS is not used any more, and the difference in depth of the *Herschel* surveys becomes quite obvious. Reassuringly, we see very good agreement between the two data sets where they overlap.

Evolution of the Fraction of Starbursts

From these distributions, we can derive interesting statistical properties of our star-forming galaxy sample. In particular, Rodighiero et al. (2011) reported that only 2 to 3 % of the galaxies in their $z = 2$ sample were in a “starburst” mode, with an SFR increased by more than a factor 4 (or 0.6 dex) compared to the Main Sequence (i.e., $R_{\text{SB}} > 4$). Using our data set, we are able to measure this fraction at different redshifts and look for an evolution of this population. To do so, we select in each redshift bin all star-forming galaxies more massive than $5 \times 10^{10} M_{\odot}$ (this mass threshold is chosen to avoid SFR completeness issues), and compute the fraction of objects for which the observed SFR is at least a factor X_{SB} above the Main Sequence. Following Rodighiero et al. (2011), we choose $X_{\text{SB}} = 4$. However, to make sure that our results are not affected by this somewhat arbitrary choice, we also do this analysis with $X_{\text{SB}} = 3$ and 2.5. By lowering this threshold, the number of objects increases and the statistics become more robust, at the price of having a higher number of nonstarburst contaminants scattering from the Main Sequence. We could have overcome this problem by fitting the observed counts, decomposing the total SFR distribution as coming from two populations: a main-sequence component and a starburst component, as was done in Sargent et al. (2012). While such a deconvolution provides a more physical definition of a “starburst”, it is also dependent on the model one chooses to describe the starburst population. Also, except in a few low redshift bins, our data do not probe a wide enough range to be able to robustly perform this decomposition. We therefore choose this simpler approach of a fixed R_{SB} threshold for now, and will come back to the decomposition later. The results are presented in Fig. 2.18. Between $z = 0.5$ and $z = 4$ and for $X_{\text{SB}} = 4$, we measure a roughly constant value ranging between 2 and 4 %, and no clear trend with redshift emerges. We discuss the implication of this fact in Sect. 2.5.2.

Quantifying the Contribution of Starbursts to the Total SFR Budget

We now normalize the counts by the integral of the stellar mass function in all bins and, supported by our findings on the constant width of the Main Sequence (Fig. 2.12) and on a constant starburst fraction (Fig. 2.18), we assume that the R_{SB} distribution does not vary. With this same assumption of an unvarying distribution, Sargent et al. (2012) managed to reconstruct the IR luminosity function at various redshifts. With the increased statistics, we are now able to perform a two-component decomposition

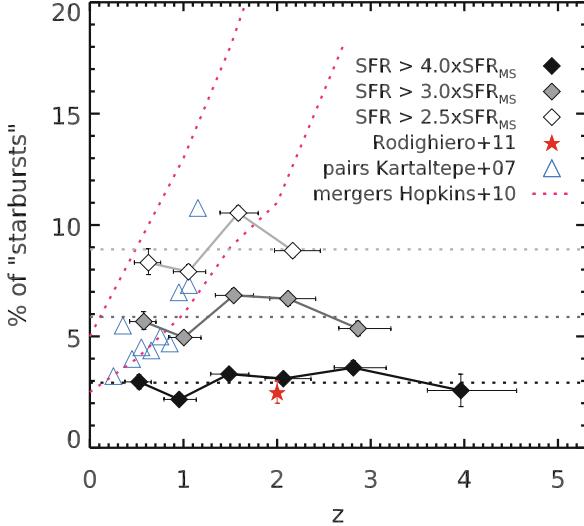


Fig. 2.18 Evolution of the observed “starburst” fraction with redshift, where starbursts galaxies are defined as having an SFR enhanced by at least a factor X_{SB} compared to the SFR on the Main Sequence. Our results are shown for $X_{\text{SB}} = 4$, 3 and 2.5 as *diamonds* (black, gray, and white, respectively), slightly offset in redshift for clarity. Only points where the starburst sample is complete are shown, and *error bars* are estimated using bootstrapping. We also show the value observed by Rodighiero et al. (2011) at $z = 2$ as a *filled red star*, which was obtained with $X_{\text{SB}} = 4$. These figures are compared qualitatively to the observed pair fraction reported by Kartaltepe et al. (2007) as *open blue triangles*, and the range of major merger fractions predicted by Hopkins et al. (2010a) is shown with *dashed purple lines*. It is clear that, both in observations and simulations, the merger fraction evolves significantly faster than the observed starburst fraction, the latter remaining almost constant regardless of the precise definition of what is a “starburst” (color figure online)

of the whole distribution. We thus fit all the counts simultaneously with a double log-normal distribution following Sargent et al. (2012). The chosen parametrization for the fit is

$$\phi_{R_{\text{SB}}}(x) = \frac{1 - f_{\text{SB}} - f_{\text{miss}}}{\sqrt{2\pi} \sigma_{\text{MS}}} \exp \left[-\frac{\log_{10}(x/x_0)^2}{2\sigma_{\text{MS}}^2} \right] + \frac{f_{\text{SB}}}{\sqrt{2\pi} \sigma_{\text{SB}}} \exp \left[-\frac{\log_{10}(x/B_{\text{SB}})^2}{2\sigma_{\text{SB}}^2} \right], \quad (2.10)$$

where σ_{MS} and σ_{SB} are the widths of the Main Sequence and starburst distributions, respectively, f_{SB} is the fraction of starbursts, and B_{SB} is the median multiplicative SFR boost of starburst galaxies. We also introduce f_{miss} as the fraction of star-forming galaxies that are neither “Main Sequence” nor “starburst” galaxies (e.g., “green valley” galaxies), and x_0 the median R_{SB} of main-sequence galaxies. By construction, the latter two parameters should be close to 0 and 1, respectively, but we allow them to vary to check for the consistency between the detections and the stacks.

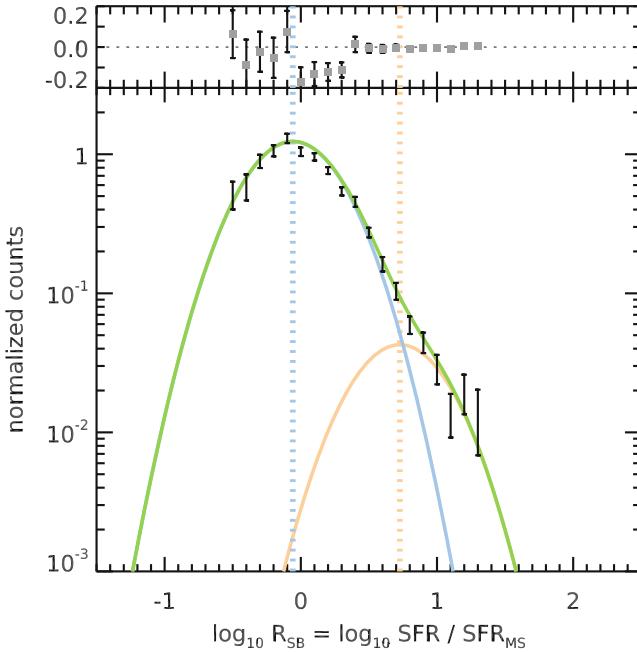


Fig. 2.19 Combined starburstiness (R_{SB}) distributions from Fig. 2.17 normalized to the total number of star-forming galaxies in each bin. The *green line* shows our best-fit model from Eq. 2.10, and the *blue* and *orange lines* show the contributions of Main Sequence and starburst galaxies, respectively. The residuals of the fit are shown at the *top* of the figure (color figure online)

The result is shown in Fig. 2.19. Leaving all parameters free, the fit of the starburst population is highly uncertain, so we decided to fix $\sigma_{\text{SB}} = \sigma_{\text{MS}}$, and fit the logarithm of the counts. We obtain $\sigma_{\text{MS}} = 0.31 \pm 0.02$ dex, $f_{\text{SB}} = 3.3\% \pm 1.5\%$, $B_{\text{SB}} = 5.3 \pm 0.4$, $f_{\text{miss}} = 0\% \pm 2\%$, and $x_0 = 0.87 \pm 0.04$.

These numbers depend heavily on the chosen parametrization of the starburst population. For example, not imposing $\sigma_{\text{SB}} = \sigma_{\text{MS}}$ would change the values of B_{SB} considerably, hence the measured values should be used with caution. The integrated contribution of the starburst population is however well constrained (Sargent et al. 2012). Taking these numbers at face value, we reach a similar conclusion as Rodighiero et al. (2011) and Sargent et al. (2012), i.e., that starbursts are rare and happen in only about 3 % of galaxies at a given instant. However, they form stars on average ~ 5 times faster than their main-sequence counterparts, and thus contribute to ~ 15 % of the SFR budget. It is worth noting that the bimodality, if any, is not clearly apparent in our data, and the high R_{SB} counts can also be fit with a single power law (with a slope close to -2). While our goal is not to demonstrate the validity of this bimodal decomposition, we want to stress that the absence of a “gap” in the distribution between the peaks of the two components does not rule out the bimodal hypothesis.

The main-sequence distribution, on the other hand, is very well constrained and both its average and the measured σ_{MS} are in agreement with the stacked value. The fact that f_{miss} is close to zero means that we are able to recover essentially *all* the star-forming galaxies with this model. More precisely, if there is another population of star-forming galaxies, we can say with 70 % probability that it can only make up for less than 2 % of the counts.

Last but not least, the accuracy of the fit in all the bins (as shown in Fig. 2.17) confirms the validity of our hypothesis of a universal R_{SB} distribution.

2.5 Discussion

2.5.1 Connection of the Main-Sequence Dispersion with Feedback Processes

The nonevolution of the main-sequence dispersion, as described in Sect. 2.4.4, is intriguing. Indeed, this dispersion can originate from several completely different processes. On the one hand, the scatter within the star formation history (SFH) of individual galaxies, i.e., bursts of star formation due to minor or major merging and feedback from AGNs or supernova winds, will naturally broaden the distribution of SFR. On the other hand, the scatter may also be due to one or more missing variables, such as age, metallicity, geometry, or environment. For example, Salmi et al. (2012) found, using 24 μm based SFRs at $z \simeq 1$, that the dispersion of the Main Sequence could artificially be reduced to about 0.15 dex by introducing the rest-frame $U - V$ color as well as z -band clumpiness as extra variables. This also shows that most of the observed scatter of the Main Sequence is physical and not due to measurement errors.

Hopkins et al. (2014) have computed the expected scatter of SFH from a set of numerical simulations, and found it to be a strong function of halo mass, and thus of stellar mass. Performing abundance matching using their $M_* - M_{\text{halo}}$ relation, one finds that they predict a variation of the SFR (averaged over 200 Myr, hence comparable to the timescale of our FIR SFR tracer) of about 0.1 dex at $M_* > 10^{11} \text{ M}_\odot$, rising up to 0.4 dex as stellar mass decreases down to 10^8 M_\odot . They also find that this evolution is coming predominantly from the rising importance of stellar feedback, and not from merging or global gravitational instabilities. Intuitively, the smaller the galaxy, the more sensitive it is to the impact of stellar winds and super novae, since the characteristic length scale over which these phenomena tend to heat and blow away the gas is more or less constant. Since there are other components that add up to the total scatter in SFR (age, environment, metallicity, etc.), this prediction should be considered as a lower limit.

The predicted values of Hopkins et al. (2014) are shown as the purple line in Fig. 2.12. The dependence of their prediction on stellar mass is clear, yet we seem to measure a constant value. Even though there are other sources of scatter at play,

it would be a strange conspiracy for them to exactly counterbalance the evolution of the scatter within the SFH to maintain a constant main-sequence scatter (see however Sparre et al. 2015). Our interpretation is thus the following.

Stellar feedback is a necessary ingredient in numerical simulations. Without it, galaxies would consume their gas too efficiently, and with the amount of infalling gas they receive from the inter-galactic medium, they would end up today with extremely high stellar masses that are not observed. The *real* strength of the stellar feedback is poorly constrained, so it is usually considered as a free parameter and fine-tuned to reproduce the local stellar mass density. However, our observations show that it cannot be arbitrarily high. Other processes can be considered to either decrease the star formation efficiency of galaxies, or reduce the amount of infalling gas they receive (e.g., Gabor and Bournaud 2014).

2.5.2 Connection Between Starbursts and Mergers

We have shown in Sect. 2.4.6 that the starburst population is not evolving, both in relative numbers and SFR excess with respect to the Main Sequence. This is intriguing in many aspects. Both observations (and references therein Le Fèvre et al. 2000; Kartaltepe et al. 2007; Lotz et al. 2011) and numerical simulations (e.g., Somerville et al. 2008; Hopkins et al. 2010a) predict an increase of the major merger rate with increasing redshift, typically proportional to $(1 + z)^m$. Although the slope m of the evolution of the merger fraction is quite uncertain (see discussion in Kampczyk et al. 2007) and sometimes found to be flat, most studies report positive values, ranging from $m \simeq 0$ up to $m \simeq 6$. For example, Kartaltepe et al. (2007) analyzed the fraction of close pairs from $z = 0$ to $z = 1.2$, and found $m = 3.1 \pm 0.1$. Their $z = 0$ value of $0.7\% \pm 0.1\%$ is comparable to our observed starburst fraction with $X_{\text{SB}} = 4$, however extrapolating this relation to $z = 2$ would predict a pair fraction of about 50 % (20 % if we consider instead the numerical simulation of Hopkins et al. 2010a). If all or a constant fraction of those pairs do lead to gas-rich major mergers, this would have a huge impact on the number of starburst, at odds with our observations.

On the other hand, Perret et al. (2014) ran several numerical simulations of mergers of $z = 2$ clumpy galaxies, and found little to no impact of the merger on star formation when compared to isolated galaxies. Their point is that by $z = 2$ star formation is already fairly active in isolated galaxies and actually close to a saturation point due to feedback processes. When the merger happens, it therefore cannot increase the total SFR by a large amount because star formation is already at its maximum. So even if mergers were more frequent in the past, they were also less efficient at triggering bursts of star formation, and this could explain why we are not seeing a huge increase in the number of starburst galaxies. This goes in the same direction as the results of Hopkins et al. (2010b) who found in their simulations that merger-driven bursts contribute to the same fraction (5–10 %) of the IR luminosity function at all redshifts, but it does not explain why the fraction of such bursts remains constant over time.

Although the most extreme starburst events are unambiguously associated with major mergers in the local Universe (e.g., Armus et al. 1987), another interpretation of our results is that the situation may be different at earlier epochs, and that some other phenomena may be responsible for such bursts of star formation, such as large scale dynamical instabilities (e.g., Dekel et al. 2009b).

2.6 Conclusions

We have put together a catalog of star-forming galaxies that is mass-complete above $2 \times 10^{10} M_{\odot}$ and extends up to $z = 4$, using the deep UV to NIR observations in the CANDELS fields. By stacking the *Herschel* images at the positions of these galaxies, using bins of mass and redshift, we measured their average star formation rates in a dust-unbiased way. We then derived a new technique called “scatter stacking” to measure the scatter in SFR around the average stacked value. We also analyzed sources individually detected on the *Herschel* images to study the SFR distribution in more detail over a more limited range of redshift and stellar mass.

We observe a continuously rising $\text{sSFR} \equiv \text{SFR}/M_*$ up to $z = 4$, with no clear sign of a saturation or plateau at the highest redshifts. Previous observations of this type of saturation are mostly based on LBG samples that lack observations in the FIR to reliably constrain the dust extinction. Earlier results are likely due to a combination of selection effects and biases in the dust extinction correction. It is therefore mandatory to have mass-complete samples and rest-frame MIR or FIR data to provide reliable constraints on the star formation activity of actively star-forming galaxies.

We find that the slope of the $\text{SFR}-M_*$ relation is close to unity, except for high mass galaxies ($M_* \gtrsim 10^{10.5} M_{\odot}$), where the slope is shallower. Furthermore, the high mass slope is evolving from ~ 0.8 at high redshifts down to almost 0 at $z \sim 0.5$. One possible explanation is the increasing contribution of the bulge to the stellar mass of these galaxies, while the star formation rates come mostly from the disk (Abramson et al. 2014).

At fixed mass and redshift, the scatter around the average SFR appears to be constant and close to 0.3 dex from $M_* = 3 \times 10^9 M_{\odot}$ to $2 \times 10^{11} M_{\odot}$, with no clear redshift dependence. We therefore confirm the existence of the “Main Sequence” of star-forming galaxies over a large range of mass and redshift with a robust star formation rate tracer. We show that at least 66 % of present day stars were formed in main-sequence galaxies. Consequently, whatever physical process produces the Main Sequence is the dominant mode of stellar growth in galaxies.

The nonevolution of the SFR scatter with mass can be connected to the expected strength of stellar feedback. State-of-the-art numerical simulations indeed predict that stellar feedback generates additional scatter in the star formation histories of galaxies, a scatter whose amplitude is strongly anticorrelated with halo mass and thus galaxy mass. Our observations provide useful constraints for numerical simulations where stellar feedback is often used as an efficient star formation regulator. We show here that it cannot be arbitrarily high.

Refining the above analysis with individual *Herschel* detections, we look for starburst galaxies whose SFRs are systematically larger than those of main-sequence galaxies. In agreement with Sargent et al. (2012) and extending their analysis to higher redshifts and more complete samples, we find that the fraction of these starburst galaxies does not evolve with time. This questions the usual interpretation of starburst as the consequence of triggering by major mergers. Several studies, both of simulations and observations, indeed show that the fraction of mergers was substantially higher in the past. An alternative explanation is that mergers may be less efficient at creating bursts of star formation within high redshift galaxies.

We have pushed *Herschel* as far as possible to study the Main Sequence of star-forming galaxies, but it is still necessary to dig deeper than that, i.e., probing higher redshifts or lower stellar masses. Most of what we know at present about the high redshift Universe ($z > 4$) comes from rest-frame UV-based studies, and we have shown here that dust extinction plays an important role even at these redshifts. Therefore it will be necessary to explore these epochs of the Universe with an independent and more robust SFR tracer to confirm the pioneering results obtained with the UV light alone. Probing lower stellar masses will also be an important challenge since, owing to their small sizes, low mass ($M_* < 3 \times 10^9 M_\odot$) galaxies are probably most sensitive to smaller scale physics, e.g., stellar or AGN feedback.

Valuable insights already come from the study of lensed galaxies. This technique allows us to observe galaxies about an order of magnitude fainter than the nominal instrument depths, either by chance in blank fields (e.g., the *Herschel* ATLAS, Eales et al. 2010), or by explicitly targeting large galaxy clusters (e.g., the *Herschel* Lensing Survey, Egami et al. 2010). Studying these regimes on statistically relevant samples and with a dust-unbiased SFR tracer will only be possible with a new generation of instruments. The most promising candidate available today for the high redshift Universe is certainly the Atacama Large Millimeter/submillimeter Array (ALMA), and interesting science is already on its way. In particular, we are now waiting for the completion of Cycle 2 observations targeting a mass-complete sample of $z = 4$ star-forming galaxies down to $\log_{10}(M_*/M_\odot) = 10.7$. With only a few minutes of on-source integration, these data will allow us to probe SFRs about five times lower than those available with the deepest *Herschel* surveys. As for the low mass galaxies, substantial progress is likely to happen in a few years thanks to the exceptional MIR capabilities of the *James Webb Space Telescope* (*JWST*).

2.7 Appendix: The *UVJ* Selection

To further test the reliability of the *UVJ* selection technique, we have separately stacked the galaxies classified as quiescent. The result is presented in Fig. 2.20. On this plot we show what the location of the quiescent galaxies would be on the SFR– M_* plane assuming that *all* their IR luminosity is coming from star formation. This is certainly wrong because in these massive galaxies dust is mostly heated by old stars, so the SFR we derive is actually an upper limit on the true star formation

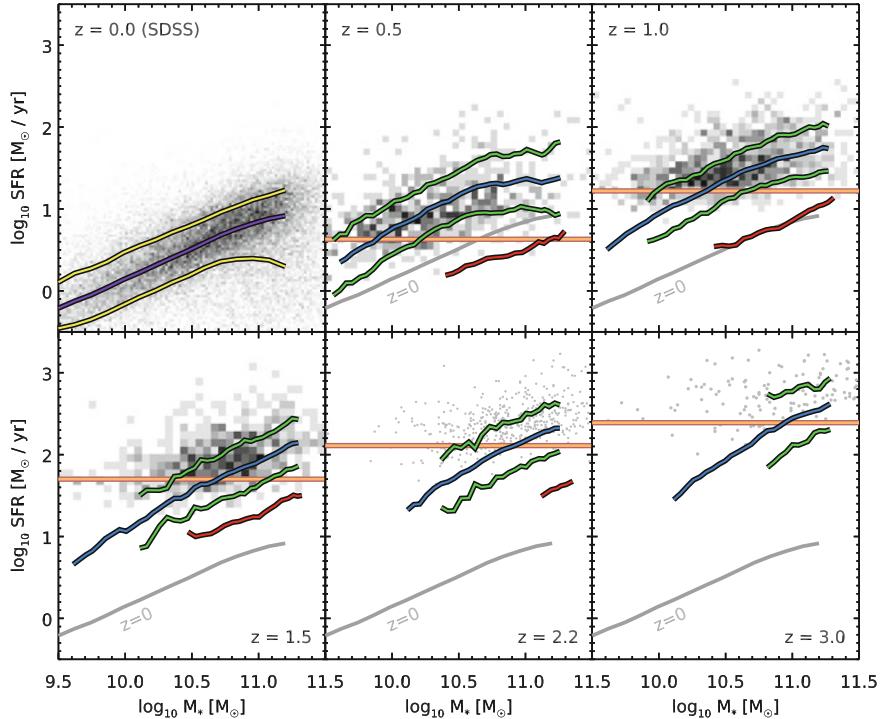


Fig. 2.20 Same as Fig. 2.16, this time also showing the location of *UVJ* passive galaxies. In each panel, the *blue line* shows the average stacked SFR (Sect. 2.4.2), and the *green lines* above and below show the 1σ dispersion obtained with scatter stacking. The *orange horizontal line* shows the detection limit of *Herschel* in SFR. The *red line* shows the stacked SFR of *UVJ* passive galaxies, naively assuming that all the IR light comes from star formation. This is a conservative upper limit, since in these galaxies dust is predominantly heated by old stars, and the effective dust temperature inferred from the FIR SED is much colder than for actively star-forming galaxies of comparable mass (color figure online)

activity of these galaxies. However, even with this naive assumption, the derived SFRs are an order of magnitude lower than that of the star-forming sample. We also observe that the effective dust temperature, inferred from the wavelength at which the FIR emission peaks, is lower and this is expected if dust is indeed mainly heated by less massive stars.

2.8 Appendix: Tests of Our Methods on Simulated Images

To test all of these procedures, we build a set of simulated images. We design these to be as close as possible to the real images in a *statistical* sense, i.e., the same photometric and confusion noise, and the same number counts.

To do so, we start from our observed *H*-band catalogs, knowing redshifts and stellar masses for all the galaxies. Using our results from stacking *Herschel* images, we can attribute an SFR to each of these galaxies. We then add a random amount of star formation, following a log-normal distribution of dispersion 0.3 dex. We also put 2 % of our sources in *starburst* mode, where their SFR is increased by 0.6 dex. Next, we assign an FIR SED to each galaxy following the observed trends with redshift (no mass dependence) and excess SFR (Magnelli et al. 2014). Starburst galaxies are also given warmer SEDs.

From these simulated source catalogs, we generate a list of fluxes in all *Herschel* bands. Given noise maps (either modeled from RMS maps assuming Gaussian noise, or constructed from the difference between observing blocks), we build simulated images by placing each source as a PSF centered on its sky position, with a Gaussian uncertainty of $0.45''$ and a maximum offset of $0.9''$. We randomly reposition the sources inside the fields using uniform distributions in right ascension and declination, to probe multiple realizations of confusion. These simulated images have pixel distribution, or $P(D)$ plots, very close to the observed images, and are thus good tools to study our methods. An example is shown in Fig. 2.21 for the GOODS–South field at $100\text{ }\mu\text{m}$.

We produce 400 sets of simulated catalogs and images, each with a different realization of photometric noise, confusion noise and SFR. We then run our full stacking procedure on each, using the same setup as for the real images (i.e., using the same redshift and mass bins), to test the reliability of our flux extraction and the accuracy of the reported errors.

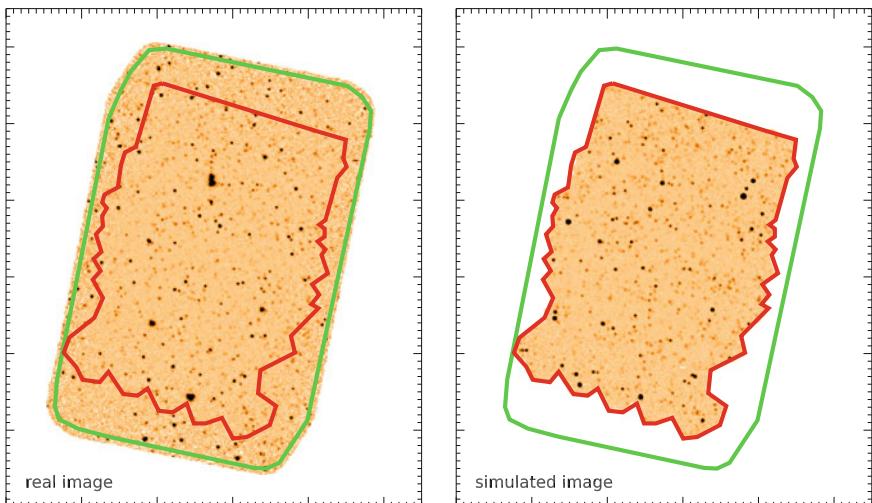


Fig. 2.21 Real *Herschel* PACS $100\text{ }\mu\text{m}$ image (left) and one of our simulations (right). The green region shows the extent of the PACS coverage, while the red region shows the *Hubble* ACS coverage, i.e., the extent of our input catalog. The two images are shown here with the same color bar (color figure online)

2.8.1 Mean and Median Stacked Fluxes

For each of the 400 realizations we compare the measured flux densities using both mean and median stacking to the expected mean and median flux densities, respectively. The results are shown in Fig. 2.22 for the PACS 100 μm band. The other bands show similar behavior.

Although less noisy, median fluxes are biased toward higher values (at most by a factor 2 here). This is because the median is not a linear operation, so it is not true in general that $\langle a + b \rangle = \langle a \rangle + \langle b \rangle$, where $\langle \cdot \rangle$ denotes the median. In particular, this means that if we compute the median of our noisy stacked image and subtract the median value of the noise, we do not exactly recover the median flux density. We will call this effect the *noise bias* in what follows. White et al. (2007) show that this bias arises when: (1) the signal to noise ratio of stacked sources is low; and (2) the distribution of flux is skewed toward either faint or bright sources. The latter is indeed true in our simulations, since we used a log-normal distribution for the SFR. Correcting for this effect is not trivial, as it requires knowledge of the real flux distribution. Indeed, Fig. 2.23 shows the amplitude of this bias for different log-normal flux dispersions, the highest dispersions producing the highest biases. White

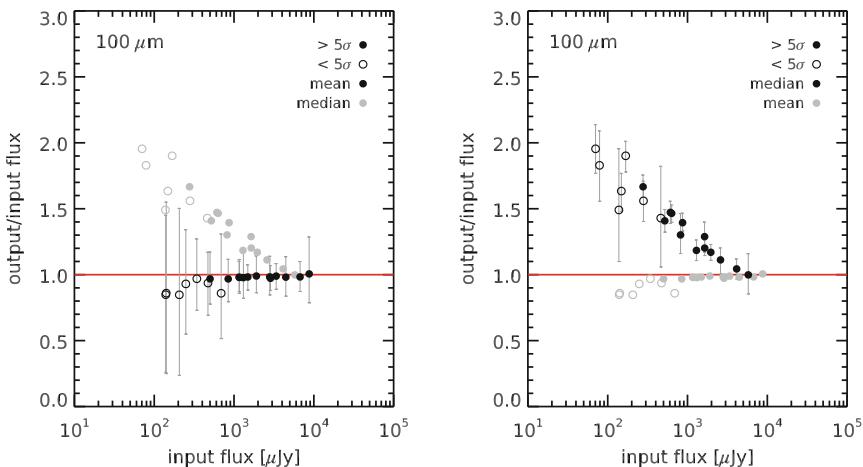
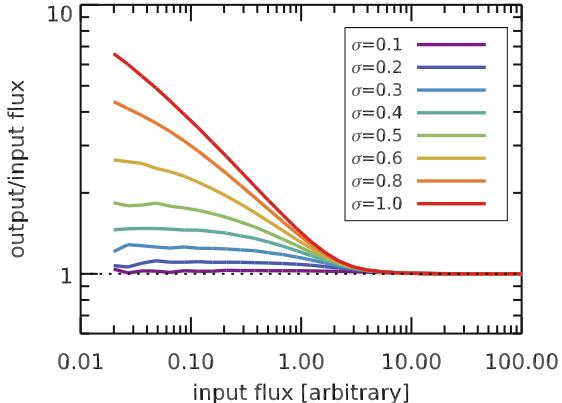


Fig. 2.22 Comparison of measured stacked flux densities from the simulated images with the real flux densities that were put into the 100 μm map (the other wavelengths behave the same). The stacked sources were binned in redshift and mass using exactly the same bins as those that were used to analyze the real images. *Left* mean stacked flux densities, *Right* median stacked flux densities. Each point shows the median $S_{\text{output}}/S_{\text{input}}$ among all the 400 realizations, while error bars show the 16th and 84th percentiles of the distribution. Filled circles indicate measurements that are individually significant at $>5\sigma$ on average, i.e., those we would actually use, while open circles indicate measurements at $<5\sigma$ to illustrate the trend. On each plot, gray circles show the values obtained with the other method (i.e., median and mean) for the sake of direct comparison. It is clear that mean fluxes are more noisy, while median fluxes exhibit a systematic bias (color figure online)

Fig. 2.23 Monte Carlo analysis showing evidence for a systematic bias in median stacking. These values have been obtained by computing medians of log-normally distributed values in the presence of Gaussian noise of fixed amplitude ($\sigma_{\text{noise}} = 1$ in these arbitrary flux units, so that the input flux is also the S/N) (color figure online)



et al. (2007) argue that the median stacked flux is still a useful quantity, since it is actually a good measure of the *mean* of the distribution, but this is only true in the limit of low signal to noise ratios. In their first example, a double normal distribution, the measured median reaches the true mean for $\text{SNR} < 0.1$, but correctly measures the true median for $\text{SNR} > 3$.

Of course these values depend on the distribution itself, as is shown in Fig. 2.23. In particular, for a log-normal distribution with 0.3 dex scatter, the mean is reached for $\text{SNR} < 0.4$, and the median for $\text{SNR} > 3$. Theoretically, the difference between the mean and the median for a log-normal distribution is $\log(10) \sigma^2 / 2$ dex. In our simulations, the typical $100 \mu\text{m}$ flux dispersion within a stacking bin is $\sim 0.45 \pm 0.1$ dex, which yields a mean-to-median ratio of $\sim 1.7^{+0.5}_{-0.2}$, in agreement with the maximum observed bias of Fig. 2.22.

To see how this affects the measured L_{IR} in practice, we list in Table 2.4 the ratio of the median to mean measured L_{IR} in each stacked bin, as measured on the real images. We showed in Sect. 2.4.4 that the dispersion in L_{IR} is about 0.3 dex. Therefore, assuming a log-normal distribution, we would theoretically expect the median-to-mean L_{IR} ratio to be close to 0.78. In fact, the L_{IR} is likely not going to follow this prediction, since we do not measure directly the mean (or median) L_{IR} , but estimate these values from a set of mean (or median) stacked fluxes, each suffering from a different bias. It is indeed clear from Table 2.4 that we do not measure this 0.78

Table 2.4 Ratio of the L_{IR} values obtained from median and mean stacking using the same sample on the real *Herschel* images

$\log_{10}(M_*/\text{M}_\odot)$	$z = 0.5$	1.0	1.5	2.2	3.0	4.0
11.2	0.79	0.95	0.84	0.88	0.82	0.86
10.8	0.63	0.90	0.92	0.94	0.77	—
10.2	0.84	0.98	0.90	0.97	—	—
9.8	0.89	0.91	—	—	—	—

ratio in practice: the median is usually (but not always) much closer to the mean than expected for a noiseless situation. Therefore, the median stacked L_{IR} are often not measuring the median or the mean L_{IR} , but something in between. Since correcting for this bias requires assumptions on the flux distribution, we prefer (when possible) to use the more noisy but unbiased mean fluxes for this study.

2.8.2 Clustering Correction

These values were obtained by computing the ratio of measured mean stacked fluxes to the expected mean fluxes in simulated images using our flux extraction method (see Sect. 2.3.2). Median stacked fluxes are affected the same way, after removing the noise bias described in Appendix 2.8.1. We also show the 16th and 84th percentiles of the bias distribution. The methods are: **A**, using our own flux extraction procedure (Sect. 2.3.2); **B**, same as A, but using the full PSF; and **C**, using only the central pixel.

Among our 400 random realizations, the measured mean fluxes do not show any systematic bias. However these simulations do not take the flux boosting caused by source physical clustering into account, because we assigned random positions to the sources in our catalog. To test the effect of clustering, we regenerate a new set of 200 simulations, this time using the real optical positions of the sources and only varying the photometric noise and the SFRs of the sources.

If galaxies are significantly clustered in the image, then the measured fluxes will be boosted by the amount of light from clustered galaxies that falls inside the beam. Since the beam size here is almost a linear function of the wavelength, we expect SPIRE bands to be more affected than PACS bands. Since the same beam at different redshifts corresponds to different proper distances, low redshift measurements ($z < 0.5$) should be less affected. However, because of the flatness of the relation between redshift and proper distance for $z > 0.5$, this should not have a strong impact for most of our sample. Indeed, we do not observe any significant trend with redshift in our simulations. No trend was found with stellar mass either, hence we averaged the clustering signal over all stacked bins for a given band, and report the average measured boost in Table 2.5 (“method A”) along with the 16th and 84th percentiles. Although we limited this analysis to fluxes measured at better than 5σ , the scatter in the measured bias is compatible with being only caused by uncertainties in flux extraction.

Table 2.5 Clustering bias in simulated *Herschel* images

Method	100 μm	160 μm	250 μm	350 μm	500 μm
A	0 % $^{+7\%}_{-7\%}$	3 % $^{+9\%}_{-8\%}$	8 % $^{+12\%}_{-8\%}$	13 % $^{+12\%}_{-10\%}$	25 % $^{+19\%}_{-18\%}$
B	0 % $^{+8\%}_{-12\%}$	3 % $^{+13\%}_{-12\%}$	19 % $^{+17\%}_{-11\%}$	33 % $^{+27\%}_{-19\%}$	58 % $^{+54\%}_{-31\%}$
C	0 % $^{+8\%}_{-7\%}$	7 % $^{+11\%}_{-9\%}$	14 % $^{+14\%}_{-9\%}$	22 % $^{+19\%}_{-14\%}$	39 % $^{+22\%}_{-23\%}$

Table 2.6 Ratio of the L_{IR} s obtained after and before applying clustering corrections listed in Table 2.5

$\log_{10}(M_*/M_\odot)$	$z = 0.5$	1.0	1.5	2.2	3.0	4.0
11.2	0.96	1.01	0.90	0.93	0.91	0.75
10.8	0.96	1.02	0.87	0.97	0.93	—
10.2	0.99	0.98	0.96	0.99	0.94	—
9.8	0.99	0.95	0.78	—	—	—

Although negligible in PACS, this effect can reach 30 % in SPIRE 500 μm data. Here we correct for this bias by simply deboosting the real measured fluxes by the factors listed in Table 2.5, band by band. The net effect on the total measured L_{IR} is reported in Table 2.6.

By construction, these corrections are specific to our flux extraction method. By limiting the fitting area to pixels where the PSF relative amplitude is larger than 10 %, we absorb part of the large scale clustering into the background level. If we were to use the full PSF to measure the fluxes, we would measure a larger clustering signal (see Sect. 2.3.2). We have re-extracted all the fluxes by fitting the full PSF, and we indeed measure larger biases. These are tabulated in Table 2.5 as “method B”. An alternative to PSF fitting that is less affected by clustering consists of setting the mean of the flux map to zero before stacking and then only using the central pixel of the stacked cutout (Béthermin et al. 2012). Because of clustering, the effective PSF of the stacked sources will be broadened, and using the real PSF to fit this effective PSF will result in some additional boosting. Therefore, by only using the central pixel, one can get rid of this effect. We show in Table 2.5 as “method C” how the figures change using this alternative method. Indeed the measured boosting is smaller than when using the full PSF, and is consistent with that reported by Béthermin et al. (2015), but our method is even less affected thanks to the use of a local background (Fig. 2.24).

2.8.3 Error Estimates

We now study the reliability of our error estimates on the stacked fluxes. We compute the difference between the observed and input flux for each realization, ΔS . We then compute the median $\langle \Delta S \rangle$, which is essentially the value plotted in Fig. 2.22, i.e., it is nonzero mostly for median stacked fluxes. We subtract this median difference from ΔS , and compute the scatter σ of the resulting quantity using median absolute deviation, i.e., $\sigma \equiv 1.48 \times \text{MAD}(\Delta S - \langle \Delta S \rangle)$. We show in Fig. 2.25 the histograms of $(\Delta S - \langle \Delta S \rangle)/\sigma$ for the mean and median stacked PACS 100 μm fluxes in each stacked bin. By construction, these distributions are well described by a Gaussian of width unity (black curve).

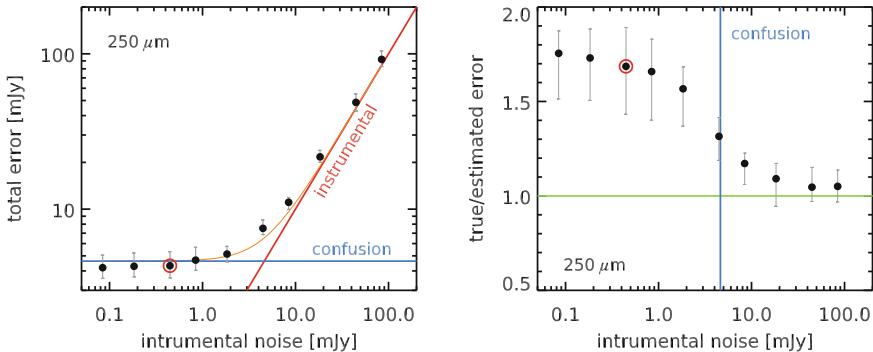


Fig. 2.24 True error σ on the stacked flux measurements as a function of the instrumental white noise level σ_{inst} that is put on the image (here normalized to a “PSF” noise in mJy, i.e., the error on the flux measurement of a point source in the absence of confusion). We generated multiple simulations of the 250 μm maps using varying levels of white noise, and compute σ from the difference between the measured fluxes and their expected values. *Left*: evolution of the average total noise per source $\sigma_{\text{tot}} = \sigma \times \sqrt{N_{\text{stack}}}$ where N_{stack} is the number of stacked sources. This is the total error when extracting the flux of a single source on the map. When the instrumental noise (red line) is high, it dominates the error budget over the confusion noise (blue line). However, when reaching too low values, the measured total noise is dominated by the confusion noise σ_{conf} . (orange line) We fit this evolution as $\sigma_{\text{tot}}^2 = \sigma_{\text{inst}}^2 + \sigma_{\text{conf}}^2$ (orange line) to obtain $\sigma_{\text{conf}} = 4.6$ mJy. The red circle marks the instrumental noise level reached in the real maps. *Right*: comparison between the estimated error from the stack residual σ_{IMG} and the true error σ . The points show the median of $\sigma/\sigma_{\text{IMG}}$, and the error bars are showing the 16th and 84th percentiles of the distribution. The green horizontal line is the line of perfect agreement, and the blue vertical line marks the confusion noise at 250 μm . The red circle marks the instrumental noise level reached in the real maps (color figure online)

We have two error estimates at our disposal. The first, σ_{IMG} , is obtained by measuring the RMS of the residual image (after the stacked fluxes have been fitted and subtracted), and multiplying this value by the PSF error factor (see Eq. 2.4). The second, σ_{BS} , is obtained by bootstrapping, i.e., repeatedly stacking half of the parent sample and measuring the standard deviation of the resulting flux distribution (again, see Sect. 2.3.2). Each of these method provides a different estimation of the error on the flux measurement, and we want to test their accuracy.

In Fig. 2.25, we show as red and blue lines the predicted error distribution according to σ_{IMG} and σ_{BS} , respectively. When the predicted distribution is too narrow or too broad compared to the observed distribution (black curve), this means that the estimated error is respectively too low or too high.

For median stacked fluxes, it appears that σ_{BS} is accurate in all cases. It tends to slightly overestimate the true error on some occasions, but not by a large amount. On the other hand, σ_{IMG} dramatically underestimates the error when the measured S/N of stacked sources is high (or the number of stacked sources is low).

The situation for mean stacked fluxes is quite different. The behavior of σ_{IMG} is the same, but σ_{BS} show the completely opposite trend, i.e., it underestimates the error at low signal to noise and high number of stacked sources. This may be caused

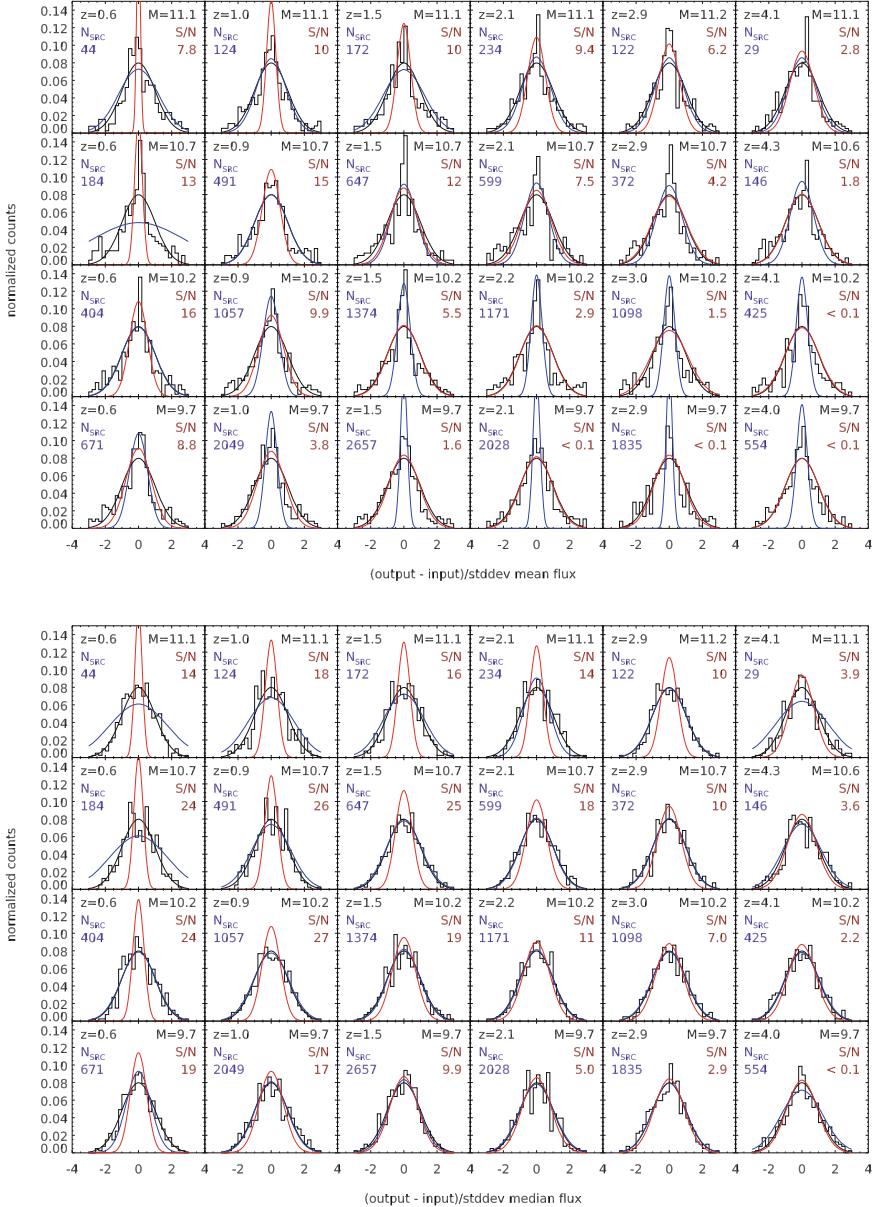


Fig. 2.25 Normalized distribution of $(\Delta S - \langle \Delta S \rangle)/\sigma$ of the mean (top) and median (bottom) stacked PACS 100 μ m fluxes in each stacked bin. The black, blue, and red curves show Gaussians of width 1, σ_{BS}/σ and σ_{IMG}/σ , respectively. The estimation of the true signal to noise ratio of the flux measurement is displayed in dark red, while the average number of stacked sources is shown in dark blue (color figure online)

by the fact that bootstrapping will almost always produce the same confusion noise, since it uses the same sources. The reason why this issue does not arise for median stacked fluxes might be because the median naturally filters out bright neighbors, hence reducing the impact of confusion noise.

The results are the same for the PACS 70 and 160 μm band. Therefore, keeping the maximum error between σ_{IMG} and σ_{BS} ensures that one has an accurate error measurement in all cases for the PACS bands.

The SPIRE fluxes on the other hand show a substantially different behavior. We reproduce the same figures in Fig. 2.26, this time for the SPIRE 350 μm band. Here, and except for the highest mass bin, the errors are systematically underestimated by a factor of ~ 1.7 , regardless of the estimator used. We therefore use this factor to correct all our measured SPIRE errors in these bins.

We believe this underestimation of the error is an effect of confusion noise. Indeed, it is clear when looking at the stacked maps at these wavelengths (e.g., Fig. 2.5) that there is a substantial amount of large scale noise coming from the contribution of the neighboring bright sources. The main issue with this noise is that it is spatially correlated. This violates one of the assumptions that were made when deriving the error estimation of Eq. 2.4, which may thus give wrong results. The reason why only the SPIRE bands are affected is because the noise budget here is (by design) completely dominated by confusion. This is clear from Fig. 2.24 (left): when putting little to no instrumental noise σ_{inst} on the simulated maps, the total error σ_{tot} on the flux measurements is completely dominated by the confusion noise σ_{conf} (blue line), and it is only by adding instrumental noise of at least 10 mJy (i.e., ten times more than what is present in the real maps) that the image becomes noise dominated. By fitting

$$\sigma_{\text{tot}} = \sqrt{\sigma_{\text{conf}}^2 + \sigma_{\text{inst}}^2}, \quad (2.11)$$

we obtain $\sigma_{\text{conf}} = 4.6$ mJy. This value depends on the model we used to generate the simulated fluxes, but it is in relatively good agreement with already published estimates from the literature (e.g., Nguyen et al. 2010, who predict $\sigma_{\text{conf.}} = 6$ mJy).

We then show in Fig. 2.24 (right) that the error underestimation in the SPIRE bands, here quantified by the ratio $\sigma/\sigma_{\text{IMG}}$, goes away when the image is clearly noise dominated, meaning that this issue is indeed caused by confusion and the properties of the noise that it generates.

Note that the confusion noise we measure here is a *global* quantity, averaged over the whole SPIRE maps. Elbaz et al. (2011) have shown that it is possible to measure fluxes below this limit if the *local* source density is low. However, limiting the stacked sample to these “clean” regions would dramatically reduce the size of our sample, hence we do not attempt it here.

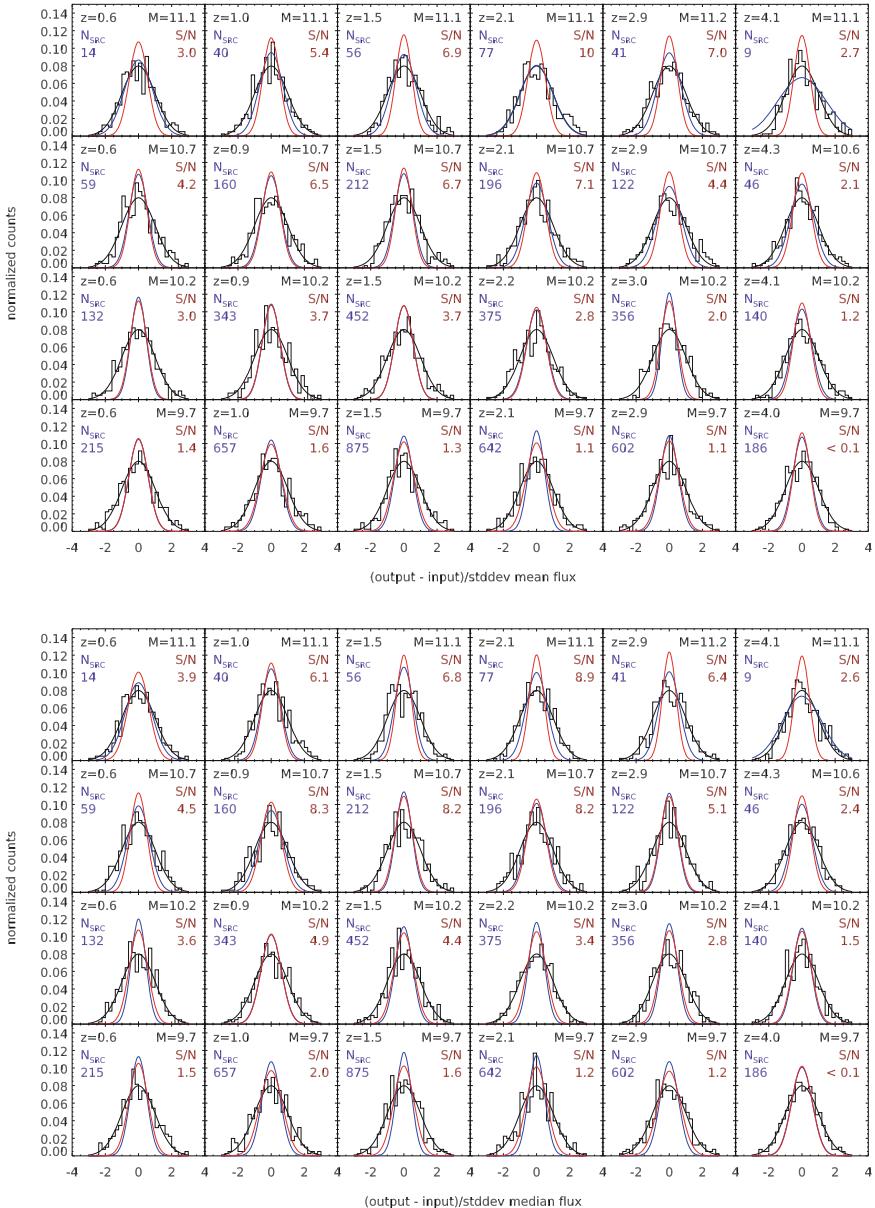


Fig. 2.26 Same as Fig. 2.25 for SPIRE 350 μm (color figure online)

References

- L.E. Abramson, D.D. Kelson, A. Dressler et al., *ApJ* **785**, L36 (2014)
- L. Armus, T. Heckman, G. Miley, *AJ* **94**, 831 (1987)
- I.K. Baldry, K. Glazebrook, J. Brinkmann et al., *ApJ* **600**, 681 (2004)
- E.F. Bell, D.H. McIntosh, N. Katz, M.D. Weinberg, *ApJS* **149**, 289 (2003)
- N. Benítez, *ApJ* **536**, 571 (2000)
- E. Bertin, S. Arnouts, *A&AS* **117**, 393 (1996)
- M. Béthermin, C. De Breuck, M. Sargent, E. Daddi, *A&A* **576**, L9 (2015)
- M. Béthermin, H. Dole, M. Cousin, N. Bavouzet, *A&A* **516**, 43 (2010)
- M. Béthermin, E. Le Floc'h, O. Ilbert et al., *A&A* **542**, 58 (2012)
- R.J. Bouwens, G.D. Illingworth, P.A. Oesch et al., *ApJ* **754**, 83 (2012)
- G.B. Brammer, P.G. van Dokkum, P. Coppi, *ApJ* **686**, 1503 (2008)
- G.B. Brammer, K.E. Whitaker, P.G. van Dokkum et al., *ApJ* **739**, 24 (2011)
- A. Bressan, F. Fagotto, G. Bertelli, C. Chiosi, *A&AS* **100**, 647 (1993)
- J. Brinchmann, S. Charlot, S.D.M. White et al., *MNRAS* **351**, 1151 (2004)
- G. Bruzual, S. Charlot, *MNRAS* **344**, 1000 (2003)
- V. Buat, J. Iglesias-Páramo, M. Seibert et al., *ApJ* **619**, L51 (2005)
- D. Calzetti, L. Armus, R.C. Bohlin et al., *ApJ* **533**, 682 (2000)
- D. Calzetti, A.L. Kinney, T. Storchi-Bergmann, *ApJ* **429**, 582 (1994)
- C.N. Cardamone, C.M. Urry, K. Schawinski et al., *ApJ* **721**, L38 (2010)
- M. Castellano, V. Sommariva, A. Fontana et al., *A&A* **566**, 19 (2014)
- C.J. Cesarsky, A. Abergel, P. Agnese et al., *A&A* **315**, L32 (1996)
- R. Chary, D. Elbaz, *ApJ* **556**, 562 (2001)
- R. Chary, A. Pope (2010). [arXiv:1003.1731](https://arxiv.org/abs/1003.1731)
- S. Cole, P. Norberg, C.M. Baugh et al., *MNRAS* **326**, 255 (2001)
- C.J. Conselice, *ApJS* **147**, 1 (2003)
- E. Daddi, F. Bournaud, F. Walter et al., *ApJ* **713**, 686 (2010)
- E. Daddi, A. Cimatti, A. Renzini et al., *ApJ* **617**, 746 (2004)
- E. Daddi, H. Dannerbauer, D. Stern et al., *ApJ* **694**, 1517 (2009)
- E. Daddi, M. Dickinson, G. Morrison et al., *ApJ* **670**, 156 (2007)
- R. Davé, K. Finlator, B.D. Oppenheimer, *MNRAS* **416**, 1354 (2011)
- A. Dekel, Y. Birnboim, G. Engel et al., *Nature* **457**, 451 (2009a)
- A. Dekel, R. Sari, D. Ceverino, *ApJ* **703**, 785 (2009b)
- T. Díaz-Santos, L. Armus, V. Charmandaris et al., *ApJ* **774**, 68 (2013)
- H. Dole, G. Lagache, J. Puget et al., *A&A* **451**, 417 (2006)
- S. Eales, L. Dunne, D. Clements et al., *PASP* **122**, 499 (2010)
- E. Egami, M. Rex, T.D. Rawle et al., *A&A* **518**, L12 (2010)
- D. Elbaz, E. Daddi, D. Le Borgne et al., *A&A* **468**, 33 (2007)
- D. Elbaz, M. Dickinson, H.S. Hwang et al., *A&A* **533**, 119 (2011)
- D. Fadda, L. Yan, G. Lagache et al., *ApJ* **719**, 425 (2010)
- A. Franceschini, H. Aussel, C.J. Cesarsky, D. Elbaz, D. Fadda, *A&A* **378**, 1 (2001)
- J.M. Gabor, F. Bournaud, *MNRAS* **441**, 1615 (2014)
- A. Galametz, A. Grazian, A. Fontana et al., *ApJS* **206**, 10 (2013)
- J.E. Geach, I. Smail, S.M. Moran et al., *ApJ* **730**, L19 (2011)
- R. Genzel, L.J. Tacconi, D. Lutz et al., *ApJ* **800**, 20 (2015)
- M.D. Gladders, A. Oemler, A. Dressler et al., *ApJ* **770**, 64 (2013)
- J.D. Goldader, G. Meurer, T.M. Heckman et al., *ApJ* **568**, 651 (2002)
- V. González, R. Bouwens, G. Illingworth et al., *ApJ* **781**, 34 (2014)
- M.J. Griffin, A. Abergel, A. Abreu et al., *A&A* **518**, L3 (2010)
- N.A. Grogin, D.D. Kocevski, S.M. Faber et al., *ApJS* **197**, 35 (2011)
- K. Guo, X.Z. Zheng, H. Fu, *ApJ* **778**, 23 (2013)
- E. Hatziminaoglou, A. Omont, J.A. Stevens et al., *A&A* **518**, L33 (2010)

- S. Heinis, V. Buat, M. Béthermin et al., MNRAS **429**, 1113 (2013)
S. Heinis, V. Buat, M. Béthermin et al., MNRAS **437**, 1268 (2014)
P.F. Hopkins, K. Bundy, D. Croton et al., ApJ **715**, 202 (2010a)
P.F. Hopkins, J.D. Younger, C.C. Hayward, D. Narayanan, L. Hernquist, MNRAS **402**, 1693 (2010b)
P.F. Hopkins, D. Kereš, J. Oñorbe et al., MNRAS **445**, 581 (2014)
O. Ilbert, H.J. McCracken, O. Le Fèvre et al., A&A **556**, 55 (2013)
S. Juneau, M. Dickinson, F. Bournaud et al., ApJ **764**, 176 (2013)
P. Kampczyk, S.J. Lilly, C.M. Carollo et al., ApJS **172**, 329 (2007)
J.S. Kartaltepe, D.B. Sanders, N.Z. Scoville et al., ApJS **172**, 320 (2007)
G. Kauffmann, T.M. Heckman, S.D.M. White et al., MNRAS **341**, 33 (2003)
R.C. Kennicutt, ARA&A, **36**, 189 (1998)
M.F. Kessler, J.A. Steinz, M.E. Anderegg et al., A&A **315**, L27 (1996)
A.M. Koekemoer, S.M. Faber, H.C. Ferguson et al., ApJS **197**, 36 (2011)
Y. Koyama, I. Smail, J. Kurk et al., MNRAS **434**, 423 (2013)
M. Kriek, P.G. van Dokkum, I. Labbé et al., ApJ **700**, 221 (2009)
P. Kurczynski, E. Gawiser, AJ **139**, 1592 (2010)
V.G. Laidler, C. Papovich, N.A. Grogan et al., PASP **119**, 1325 (2007)
O. Le Fèvre, R. Abraham, S.J. Lilly et al., MNRAS **311**, 565 (2000)
E. Le Floc'h, H. Aussel, O. Ilbert et al., ApJ **703**, 222 (2009)
E. Le Floc'h, C. Papovich, H. Dole et al., ApJ **632**, 169 (2005)
N. Lee, D.B. Sanders, C.M. Casey et al., ApJ **778**, 131 (2013)
A.K. Leroy, F. Walter, E. Brinks et al., AJ **136**, 2782 (2008)
S.J. Lilly, O. Le Fevre, F. Hammer, D. Crampton, ApJ **460**, L1 (1996)
J.M. Lotz, P. Jonsson, T.J. Cox et al., ApJ **742**, 103 (2011)
D. Lutz, A. Poglitsch, B. Altieri et al., A&A **532**, 90 (2011)
P. Madau, M. Dickinson, ARA&A **52**, 415 (2014)
P. Madau, H.C. Ferguson, M.E. Dickinson et al., MNRAS **283**, 1388 (1996)
P. Madau, L. Pozzetti, M. Dickinson, ApJ **498**, 106 (1998)
G.E. Magdis, E. Daddi, M. Béthermin et al., ApJ **760**, 6 (2012)
G.E. Magdis, D. Elbaz, E. Daddi et al., ApJ **714**, 1740 (2010)
B. Magnelli, D. Elbaz, R.R. Chary et al., A&A **496**, 57 (2009)
B. Magnelli, P. Popesso, S. Berta et al., A&A **553**, 132 (2013)
B. Magnelli, D. Lutz, A. Saintonge et al., A&A **561**, 86 (2014)
G. Marsden, P.A.R. Ade, J.J. Bock et al., ApJ **707**, 1729 (2009)
H.J. McCracken, B. Milvang-Jensen, J. Dunlop et al., A&A **544**, 156 (2012)
G.R. Meurer, T.M. Heckman, D. Calzetti, ApJ **521**, 64 (1999)
J.R. Mullaney, D.M. Alexander, A.D. Goulding, R.C. Hickox, MNRAS **414**, 1082 (2011)
J.R. Mullaney, M. Pannella, E. Daddi et al., MNRAS **419**, 95 (2012)
K. Murata, H. Matsuura, H. Inami et al., A&A **566**, A136 (2014)
A. Muzzin, D. Marchesini, M. Stefanon et al., ApJ **777**, 18 (2013a)
A. Muzzin, D. Marchesini, M. Stefanon et al., ApJS **206**, 8 (2013b)
H.T. Nguyen, B. Schulz, L. Levenson et al., A&A **518**, L5 (2010)
K.G. Noeske, B.J. Weiner, S.M. Faber et al., ApJ **660**, L43 (2007)
R. Nordon, D. Lutz, R. Genzel et al., ApJ **745**, 182 (2012)
S.J. Oliver, J. Bock, B. Altieri et al., MNRAS **424**, 1614 (2012)
I. Oteo, J. Cepa, A. Bongiovanni et al., A&A **554**, L3 (2013)
R.A. Overzier, T.M. Heckman, J. Wang et al., ApJ **726**, L7 (2011)
M. Pannella, C.L. Carilli, E. Daddi et al., ApJ **698**, L116 (2009)
M. Pannella, D. Elbaz, E. Daddi et al., ApJ **807**, 141 (2015)
D.R. Patton, C.J. Pritchett, H.K.C. Yee, E. Ellingson, R.G. Carlberg, ApJ **475**, 29 (1997)
Y.-J. Peng, S.J. Lilly, K. Kovač et al., ApJ **721**, 193 (2010)
K. Penner, M. Dickinson, A. Pope et al., ApJ **759**, 28 (2012)
V. Perret, F. Renaud, B. Epinat et al., A&A **562**, 1 (2014)

- G.L. Pilbratt, J.R. Riedinger, T. Passvogel et al., A&A **518**, L1 (2010)
- A. Poglitsch, C. Waelkens, N. Geis et al., A&A **518**, L2 (2010)
- A. Pope, R. Chary, D.M. Alexander et al., ApJ **675**, 1171 (2008)
- N. Prantzos, J. Silk, ApJ **507**, 229 (1998)
- N.A. Reddy, M. Pettini, C.C. Steidel et al., ApJ **754**, 25 (2012)
- G.H. Rieke, E.T. Young, C.W. Engelbracht et al., ApJS **154**, 25 (2004)
- G. Rodighiero, A. Cimatti, C. Gruppioni et al., A&A **518**, L25 (2010)
- G. Rodighiero, E. Daddi, I. Baronchelli et al., ApJ **739**, L40 (2011)
- G. Rodighiero, A. Renzini, E. Daddi et al., MNRAS **443**, 19 (2014)
- D.J. Rosario, P. Santini, D. Lutz et al., ApJ **771**, 63 (2013)
- W. Rujopakarn, G.H. Rieke, D.J. Eisenstein, S. Juneau, ApJ **726**, 93 (2011)
- A. Saintonge, D. Lutz, R. Genzel et al., ApJ **778**, 2 (2013)
- S. Salim, M. Dickinson, R. Michael Rich et al., ApJ **700**, 161 (2009)
- F. Salmi, E. Daddi, D. Elbaz et al., ApJ **754**, L14 (2012)
- B. Salmon, C. Papovich, S.L. Finkelstein et al., ApJ **799**, 183 (2015)
- E.E. Salpeter, ApJ **121**, 161 (1955)
- M. Salvato, O. Ilbert, G. Hasinger et al., ApJ **742**, 61 (2011)
- D.B. Sanders, I.F. Mirabel, ARA&A **34**, 749 (1996)
- D.B. Sanders, M. Salvato, H. Aussel et al., ApJS **172**, 86 (2007)
- P. Santini, A. Fontana, A. Grazian et al., A&A **504**, 751 (2009)
- P. Santini, D.J. Rosario, L. Shao et al., A&A **540**, 109 (2012)
- P. Santini, R. Maiolino, B. Magnelli et al., A&A **562**, 30 (2014)
- M.T. Sargent, M. Béthermin, E. Daddi, D. Elbaz, ApJ **747**, L31 (2012)
- P. Schechter, ApJ **203**, 297 (1976)
- C. Schreiber, M. Pannella, D. Elbaz et al., A&A **575**, A74 (2015)
- R.S. Somerville, P.F. Hopkins, T.J. Cox, B.E. Robertson, L. Hernquist, MNRAS **391**, 481 (2008)
- M. Sparre, C.C. Hayward, V. Springel et al., MNRAS **447**, 3548 (2015)
- J.S. Speagle, C.L. Steinhardt, P.L. Capak, J.D. Silverman, ApJS **214**, 15 (2014)
- D.P. Stark, R.S. Ellis, A. Bunker et al., ApJ **697**, 1493 (2009)
- D.P. Stark, M.A. Schenker, R. Ellis et al., ApJ **763**, 129 (2013)
- C.C. Steidel, K.L. Adelberger, M. Giavalisco, M. Dickinson, M. Pettini, ApJ **519**, 1 (1999)
- C.L. Steinhardt, J.S. Speagle, P. Capak et al., ApJ **791**, L25 (2014)
- V. Strazzullo, R. Gobat, E. Daddi et al., ApJ **772**, 118 (2013)
- L.J. Tacconi, R. Genzel, R. Neri et al., Nature **463**, 781 (2010)
- M.P. Viero, L. Moncelsi, R.F. Quadri et al., ApJ **779**, 32 (2013)
- W. Wang, L.L. Cowie, A.J. Barger, R.C. Keenan, H. Ting, ApJS **187**, 251 (2010)
- M.W. Werner, T.L. Roellig, F.J. Low et al., ApJS **154**, 1 (2004)
- K.E. Whitaker, I. Labb  , P.G. van Dokkum et al., ApJ **735**, 86 (2011)
- K.E. Whitaker, P.G. van Dokkum, G. Brammer, M. Franx, ApJ **754**, L29 (2012)
- K.E. Whitaker, M. Franx, J. Leja et al., ApJ **795**, 104 (2014)
- R.L. White, D.J. Helfand, R.H. Becker, E. Glikman, W. de Vries, ApJ **654**, 99 (2007)
- R.J. Williams, R.F. Quadri, M. Franx, P. van Dokkum, I. Labb  , ApJ **691**, 1879 (2009)
- S. Wuyts, I. Labb  , M. Franx et al., ApJ **655**, 51 (2007)
- S. Wuyts, N.M. F  rster Schreiber, D. Lutz et al., ApJ **738**, 106 (2011)

Chapter 3

Modelling the Integrated IR Photometry of Star-Forming Galaxies

In this chapter, I introduce a new set of template SEDs to model the dust emission of star-forming galaxies. These SEDs are based on the dust model of Galliano et al. (2011), and can be considered as an extension of the SEDs recently introduced by Magdis et al. (2012) and Béthermin et al. (2015). The advantage of this new library is two folds: first, the shape of the dust continuum is tuned to reproduce the Chary and Elbaz (2001) library (which is known to model correctly a large number of observed SEDs with only four different templates) while providing a finer control on the effective dust temperature (T_{dust}); and second, the mass fraction of PAH molecules is a free parameter, allowing a finer modeling of the MIR photometry around $8 \mu\text{m}$. This last point will be particularly important in the *James Webb Space Telescope (JWST)* era, when we will routinely observe galaxies in deep MIR surveys.

These SEDs will be published together with EGG, a code that I developed within the ASTRODEEP program and that is described in Chap. 4. The aim of this code is to simulate a patch of the sky with realistic galaxy distributions, to be able to test photometric codes and other image-based techniques like stacking. The corresponding papers are still in preparation, and the two following chapters are essentially borrowed from the current drafts.

3.1 Introduction

A number of dust SED libraries have been published during the last years, among which are the Chary and Elbaz (2001) (CE01) library, calibrated in L_{IR} from local galaxies, the Dale and Helou (2002) library, calibrated in FIR colors, or the Magdis et al. (2012) library, calibrated in intensity of the interstellar radiation field (U) (or, equivalently, in dust temperature T_{dust}).

More complex models, like that of Draine and Li (2007) or Galliano et al. (2011), provide a finer description of the dust content, allowing for example to fine tune the dust chemical composition and build composite spectrum with arbitrary temperatures distributions. However, properly constraining most of these parameters require exquisite SEDs with good wavelength sampling, which is a level of quality that is rarely achieved outside of the Local Universe. Therefore, the simpler libraries quoted above (CE01, etc.) are often preferred.

Here we are seeking for an additional level of control over the SED: we aim to be able to choose different effective dust temperatures T_{dust} and to change the relative contribution of the polycyclic aromatic hydrocarbon molecules (PAHs) f_{PAH} . There are two reasons for this choice: first, these are the two parameters that are the easiest to measure without FIR spectroscopy (which is only available for very few selected objects), and are those that affect the most the shape of the SED; and second, PAHs emit the bulk of their light around the rest-frame 8 μm , which is a domain that will be routinely accessed by the *James Webb Space Telescope* in the near future, and there will be a need for a properly calibrated library to exploit these data together with ancillary *Herschel* or *Spitzer* observations.

Therefore, we introduce in Sect. 3.2 a new SED library in which both T_{dust} and f_{PAH} are free parameters. In Sect. 3.2.1 we calibrate the redshift evolution of both parameters using the MIR to FIR stacks of Schreiber et al. (2015) (hereafter S15, and see previous chapter), to which we add stacks of the *Spitzer* IRS 16 μm imaging (Teplitz et al. 2011) to better constrain the PAH features (available in GOODS–North and South only). This calibration is revisited in Sect. 3.2.2 using individual *Herschel* detections to constrain the scatter on these parameters, and also to calibrate how they are modified for those galaxies that are offset from the Main Sequence.

Lastly, in Sect. 3.4 we further describe how the library can be used to derive infrared luminosities even when a single photometric band is available. The accuracy of such measurements is quantified in the next Chapter (Sect. 4.7).

In the following, we assume a Λ CDM cosmology with $H_0 = 70 \text{ km s}^{-1}\text{Mpc}^{-1}$, $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$ and a Salpeter (1955) initial mass function (IMF), to derive both star formation rates and stellar masses. All magnitudes are quoted in the AB system, such that $M_{\text{AB}} = 23.9 - 2.5 \log_{10}(S_\nu [\mu\text{Jy}])$.

3.2 A New Far Infrared Template Library

Since it was published, the CE01 library has been used routinely to derive infrared luminosities, and therefore star formation rates, for large samples of galaxies at various redshifts. In S15, we found that, in spite of the relatively small number of different SEDs it contains, it is able to fit relatively well our stacked *Herschel* photometry from $z = 0.5$ to $z = 4$. However, the behavior of these SEDs at 8 μm , which was calibrated from the local Universe, is peculiar. Daddi et al. (2007) found

that it is inadequate to measure star formation rates from the rest-frame 8 μm at $z = 2$, and it was later shown in Elbaz et al. (2011) that this was caused by an improper calibration of the IR8, the ratio between the total infrared luminosity L_{IR} and the luminosity at 8 μm , L_8 . This illustrates how critical it is to properly describe the 8 μm features, and in particular the emission of PAHs.

Apart from this wavelength regime, the rest of the FIR continuum of the **CE01** library is a solid reference. Therefore, when building our new library, we try to reproduce the same shape of the continuum, and only change the position of the peak of the SED to vary T_{dust} . To do so, we use the dust continuum model of Galliano et al. (2011). This model can output the mid- to far-IR spectrum emitted by a dust cloud of mass 1 M_{\odot} , composed of a mixture of carbonated and silicate grains of different sizes (split in “small” and “big” grains), under the influence of a radiation field of integrated intensity U (taken here in units of the Mathis et al. (1983) interstellar radiation field). For our templates, we use the Milky Way mass-fraction of small versus big grains and carbonated versus silicate grains (Zubko et al. 2004). We then generate the final composite templates by adding together the emission of different dusty regions, heated by different radiation intensities. Following Dale et al. (2001), we assume that the distribution of radiation intensity follows a power law in $dU/dM_{\text{dust}} = U^{\alpha}$, where M_{dust} is the mass of dust, and integrate this distribution from $U = U_{\text{min}}$ to $U = U_{\text{max}}$. The main parameter that allows us to tune the effective dust temperature here is U_{min} or, equivalently, $\langle U \rangle$ (with the U distributions we assume here, the final SED is relatively insensitive to the precise value of U_{max} , provided it is high enough). Therefore, we generate a logarithmic grid of U_{min} ranging from 10^{-1} to 10^4 with 200 samples, and take $U_{\text{max}} = 10^6$. With this model, the shape of the FIR continuum of the **CE01** library from 15 to 70 μm is reproduced using a power-law slope of the U distribution of $\alpha = 2.6$.

The model of Galliano et al. (2011) can also produce the associated PAH emission, assuming that a fraction f_{PAH} of the total mass of dust is found in PAH molecules. We will assume that these molecules are subject to the same U distribution as the other dust grains, although this choice has very little consequence since the PAH molecules are not thermalized. We also choose not to follow the Milky Way fraction of neutral versus ionized PAH molecules (50 %) and use instead a value of 10 %: this fraction will change the relative strength of the 8 versus 12 μm PAH features. When both these wavelength ranges are observed with broad band photometry (i.e., with *Spitzer* IRS 16 μm and MIPS 24 μm at $z = 1$, see Fig. 3.1), we find this choice to provide a better description of the observed data than the **CE01** library, in which the 12 μm feature is stronger. Using these parameters, we generate PAH templates with the same U distributions as for the dust continuum.

The final library is therefore composed of two sets of templates: dust continuum on one side, and PAH emission on the other. We show in Fig. 3.1 how the resulting SEDs fit to the stacks of **S15**, to illustrate the variety of SED shapes that can be reproduced with this library, and how well it is able to match the observed features. As shown in Fig. 3.2, the relative amplitude of each component can be freely adjusted,

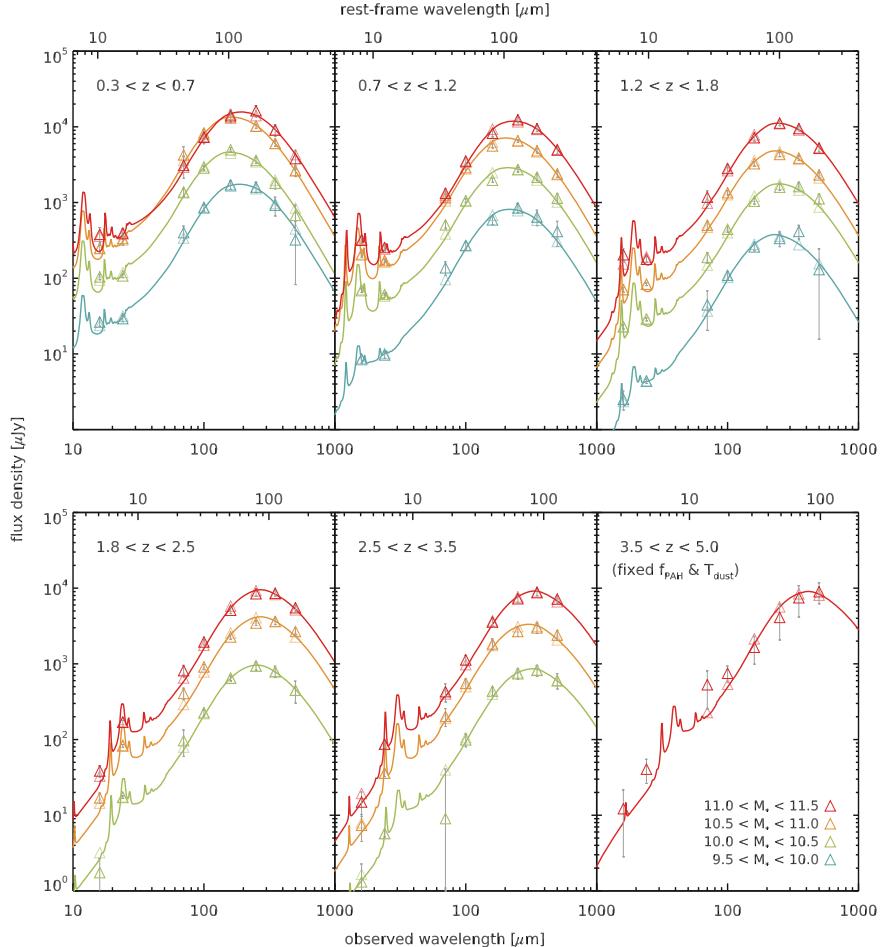


Fig. 3.1 *Spitzer* and *Herschel* stacks of S15 (open triangles) of Main Sequence galaxies at different redshifts (from left to right) and for different stellar masses (colors, see legend). We add to these data new stacks of the *Spitzer* 16 μm images in the GOODS fields, and overplot the best fit template from our library with colored solid lines. Fainter empty triangles in the background show the expected broadband flux from the best-fit template, to illustrate any offset with the observations. For the last redshift bin (bottom-right panel), we fixed f_{PAH} and T_{dust} (see Appendix 3.4) because there is no data to constrain the PAH region and the Rayleigh–Jeans tail of the dust continuum (Color figure online)

effectively changing the total dust mass M_{dust} and the PAH fraction f_{PAH} . The dust temperature T_{dust} is obtained from the best-fitting dust continuum template, and the other quantities of importance, in particular the total infrared luminosity L_{IR} , can be obtained from the combined SED (continuum+PAH).

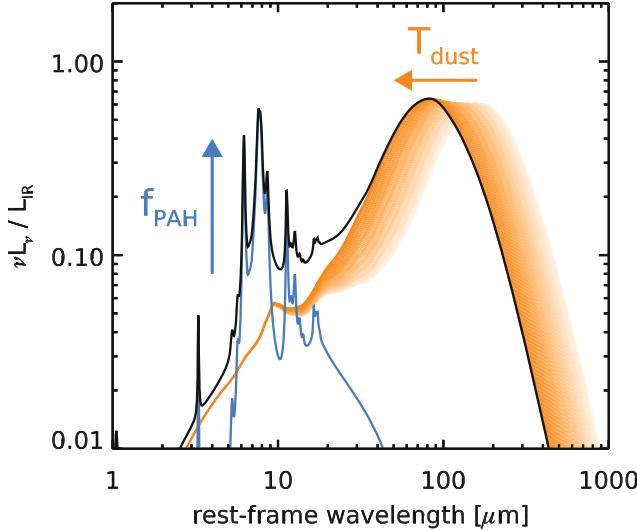


Fig. 3.2 Cartoon picture of how the FIR SED library works. The total SED is shown with a *black solid line*, while the dust continuum and PAH components are shown with a *solid orange* and *blue lines*, respectively. We also show how the shape of the SED varies with dust temperature T_{dust} by displaying several templates of different T_{dust} in *orange lines* of varying intensity. The *orange* and *blue arrows* illustrate how the SED is modified by increasing T_{dust} and f_{PAH} , respectively (Color figure online)

Using these SEDs, we have the following relations between T_{dust} , L_{IR} , M_{dust} and $\langle U \rangle$:

$$\langle U \rangle = \left(\frac{T_{\text{dust}}}{17.5 \text{K}} \right)^{5.54} \quad (3.1)$$

$$L_{\text{IR}}^{\text{cont}}[L_{\odot}] = 185 \times (1 - f_{\text{PAH}}) \times (M_{\text{dust}}[\text{M}_{\odot}]) \times \langle U \rangle \quad (3.2)$$

$$L_{\text{IR}}^{\text{PAH}}[L_{\odot}] = 307 \times f_{\text{PAH}} \times (M_{\text{dust}}[\text{M}_{\odot}]) \times \langle U \rangle^{0.98} \quad (3.3)$$

$$L_{\text{IR}} = L_{\text{IR}}^{\text{cont}} + L_{\text{IR}}^{\text{PAH}} \quad (3.4)$$

Note that we actually provide two dust temperatures per template in this library. The first, and the one we use in this thesis, is computed by applying Wien's law to each elementary template of Galliano et al. (2011) (of unique U , and taking into account the emissivity $\beta = 1.7$), then weighting the obtained values by the dust mass associated to each such template. It is therefore a mass-weighted average. The second is computed by applying Wien's law to the peak of the final dust template, and is therefore a light-weighted average. In practice, the difference between the two is simply a constant factor, with

$$T_{\text{dust}}^{\text{mass}} = 0.91 \times T_{\text{dust}}^{\text{light}} \quad (3.5)$$

but $T_{\text{dust}}^{\text{light}}$ is less stable because the summed dust template is broader, making it harder to locate accurately the position of the peak.

We also have the following relation between f_{PAH} and $\text{IR8} \equiv L_{\text{IR}}/L_8$ for $T_{\text{dust}} < 50$ K:

$$\text{IR8} = \frac{L_{\text{IR}}}{L_8} = \frac{1 + 0.627 \times f_{\text{PAH}}}{0.0429 + 4.64 \times f_{\text{PAH}}}, \quad (3.6)$$

where L_8 is the luminosity in the rest-frame IRAC channel 4 (8 μm) broadband. For $T_{\text{dust}} > 50$ K, the contribution of the dust continuum to the 8 μm luminosity becomes non-trivial and the relation is more complex.

In the next section, we analyze in more details the fits of this library to the stacks of S15 and derive redshift trends for both T_{dust} and f_{PAH} .

3.2.1 Calibration on Stacked Photometry

Compared to our previous fits with the CE01 library, we find very similar values of L_{IR} , except for the lowest redshift bin where we obtain value that are systematically 0.1 dex lower. This is caused by a peculiar feature of the adopted best-fit template from the CE01 library around the rest-frame 30 μm . This particular SED (ID 40) shows an enhanced flux in this wavelength range compared to our library. Without any data to constrain this feature, we cannot say whether it is real or not, although we tend to favor the result of the new SED library which has a consistent shape at all T_{dust} .

In Fig. 3.3 we show the best-fit values we obtain for T_{dust} and f_{PAH} on the stacked SEDs of S15. Errors bars are derived by perturbing the measured photometric points within their estimated uncertainties and redoing the fit 100 times, then computing the standard deviation of each parameter among all 100 realizations (see Sect. 3.4). Consistently with what was previously reported, e.g., by Magdis et al. (2012), Magnelli et al. (2013) or Béthermin et al. (2015), we find that the dust temperature increases continuously with redshift, at least up to $z = 3$. The lack of data points on the Rayleigh–Jeans tail of the SED at $z = 4$ prevents us from drawing any conclusion at that redshift, and there T_{dust} was fixed to the extrapolated value from lower redshifts (again, see Sect. 3.4). The trend we find is in very good agreement with that of Béthermin et al. (2015) (after converting their $\langle U \rangle$ values into T_{dust} using Eq. 3.1). At $z \geq 2$, we tend to find warmer dust temperatures, but this is mostly affecting the low-mass bins which are the most uncertain, and we prefer to trust the trend observed in the two most massive bins. We therefore calibrate the evolution of the average dust temperature with the following equation:

$$T_{\text{dust}}^{\text{MS}}[\text{K}] = \begin{cases} 20.2 \times (1+z)^{0.44} & \text{for } z \leq 2 \\ 26.3 \times (1+z)^{0.2} & \text{for } z > 2 \end{cases} \quad (3.7)$$

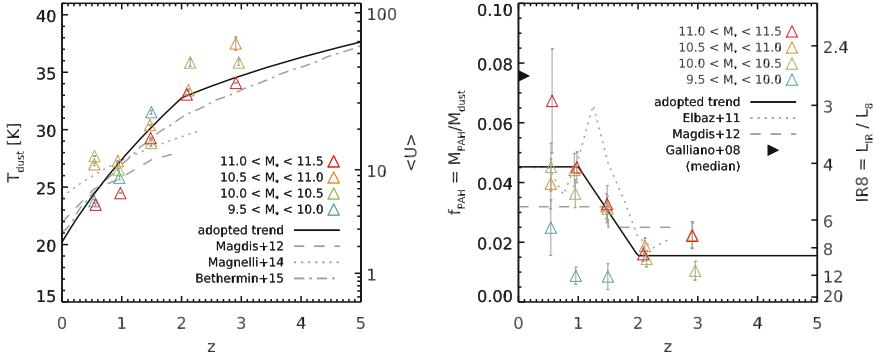


Fig. 3.3 *Left* Evolution of the effective dust temperature T_{dust} with redshift. The T_{dust} estimated from each stacked SED at different stellar masses are shown with empty triangles of different colors (see legend). The trend we adopt in this paper is illustrated with a *solid black line*. We also show the T_{dust} evolution of Magdis et al. (2012) and Béthermin et al. (2015) (both converted from $\langle U \rangle$ to T_{dust} using Eq. 3.1) as well as Magnelli et al. (2014) (corrected from light-weighted to mass-weighted using Eq. 3.5). *Right* Evolution of the PAH mass fraction f_{PAH} with redshift. The legend is the same as for the plot on the *left*, except that here we show the values obtained by Elbaz et al. (2011) and Magdis et al. (2012) (computed from their SEDs). Both were converted from IR8 to f_{PAH} using our SEDs. For reference, we also give the median f_{PAH} value measured in local galaxies by Galliano et al. (2008) (Color figure online)

This can also be compared, e.g., with the result of Magnelli et al. (2014), who found $T_{\text{dust}} = 26.5 \times (1+z)^{0.18}$. The normalization of this relation is higher than the one we report here, which may be linked to the fact that Magnelli et al. (2014) measured the T_{dust} with modified black body fits,¹ and therefore that their dust temperatures are light-weighted. Correcting for this difference using Eq. 3.5, as was done in Fig. 3.3, these values are fully consistent with the ones we measure here.

We also find a trend for f_{PAH} to decrease with redshift, and therefore for $\text{IR8} \equiv L_{\text{IR}} / L_8$ to increase. Elbaz et al. (2011) propose that a unique value of $\text{IR8} = 4.9$ holds for all Main Sequence galaxies, however it can be seen already from their data that the average IR8 is closer to 8 at $z = 2$. Nevertheless, this finding does not affect the conclusions of Elbaz et al. (2011), which we revisit in Sect. 3.2.2. Interestingly, we do find a systematic trend in the lowest mass bin ($M_* \simeq 5 \times 10^9 M_\odot$) for noticeably lower f_{PAH} , at least at $z < 2$ where we can measure it. We come back to these results in Sect. 3.3, and for now we simply parametrize the evolution of the average f_{PAH} with

$$f_{\text{PAH}}^{\text{MS}} = 0.04 + 0.035 \times (1 - 0.85 \times \text{clamp}(z, 1, 2)), \quad (3.8)$$

¹In fact, they compute the effective dust temperature of the Dale and Helou (2002) templates, and use these templates to associate a T_{dust} to each galaxy. This approach is more robust than a simple modified black body fit, but will essentially return the same result.

where “clamp” is defined as

$$\text{clamp}(x, x_0, x_1) = \begin{cases} x & \text{for } x_0 < x < x_1 \\ x_0 & \text{for } x < x_0 \\ x_1 & \text{for } x > x_1 \end{cases}. \quad (3.9)$$

3.2.2 Calibration on Individual Detections

We now use the above library to fit all the FIR-detected galaxies in the CANDELS fields. For now, we discard galaxies that have a poor wavelength coverage, i.e., those that have less than 3 photometric points to constrain the shape of the dust continuum (to measure T_{dust}), or those that are not detected simultaneously in the FIR and rest-frame 8 μm (to measure f_{PAH}). The case of these galaxies is discussed in Sect. 3.4. We also exclude fits of poor quality, by rejecting galaxies whose T_{dust} is uncertain by more than 2 K or that were fitted with a χ^2 larger than 10, indicative of issues in the photometry and/or counterpart identification.

Our goal in this section is to measure the typical scatter of both T_{dust} and f_{PAH} about the average values we obtained in Sect. 3.2.1, as well as to calibrate how these quantities are modified in starburst galaxies, i.e., those galaxies that have an excess SFR at a given stellar mass compared to the Main Sequence. To quantify this excess, we use the “starburstiness” (Elbaz et al. 2011) which is defined as $R_{\text{SB}} \equiv \text{SFR}/\text{SFR}_{\text{MS}}$, so that galaxies with $R_{\text{SB}} = 1$ are on the Main Sequence, and those with $R_{\text{SB}} > 1$ are located above the sequence.

In Fig. 3.4 (top) we show the measured dust temperatures for individual *Herschel* detections. The redshift evolution of the average temperature is well matching that seen in the stacked SEDs, although at $z > 1.5$ the sliding median of the detections is found below our stacked trend. This can be explained by a selection effect: by requiring a robust measurement of T_{dust} , we essentially require a detection in the *Herschel* SPIRE bands, which will in turn favor the inclusion of *cold* SEDs in the sample, at the expense of warmer SEDs. We then subtract this redshift-dependent average from the measured T_{dust} , and quantify how the remainder correlates with the offset from the Main Sequence. Such a trend was first observed in Elbaz et al. (2011), and later quantified by Magnelli et al. (2014) who stacked galaxies at various locations on the $\text{SFR}-M_*$ plane. They reported a linear relation between T_{dust} and $\log_{10}(R_{\text{SB}})$ (which they call $\Delta \log(\text{sSFR})$) with a slope of 6.5 K, which is very similar to the value we measure here of 6.6 K. In the right panel of Fig. 3.4 (top), we finally show the T_{dust} distribution after removing the redshift and starburstiness trends, and find a residual dispersion of 4.4 K (4.2 K after subtracting the measurement and redshift uncertainties, assuming $\Delta z/(1+z) = 5\%$), with no dependence on L_{IR} . We therefore parametrize the dust temperature of a galaxy with the following equation:

$$T_{\text{dust}}[\text{K}] = T_{\text{dust}}^{MS} + 6.6 \times \log_{10}(R_{\text{SB}}). \quad (3.10)$$

where T_{dust}^{MS} is defined in Eq. 3.7, with a Gaussian scatter of 4.2 K.

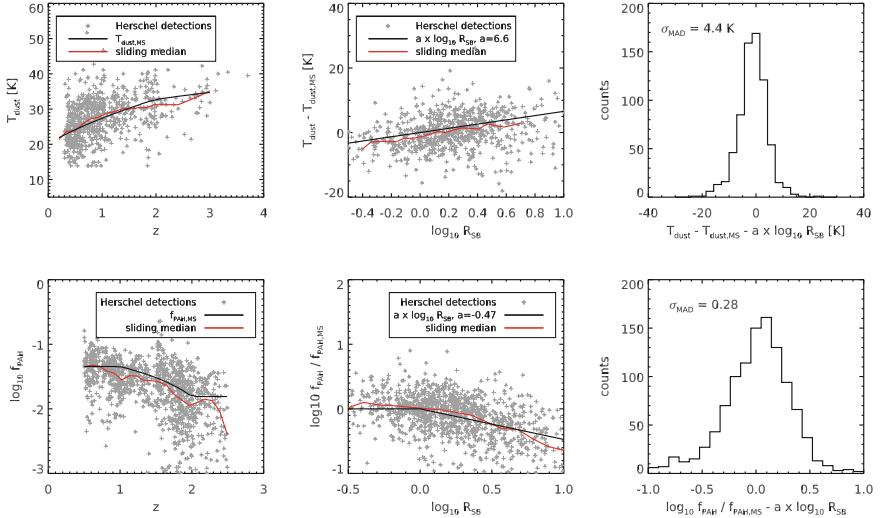


Fig. 3.4 *Left* Evolution of the dust temperature (T_{dust} , top) and PAH mass fraction (f_{PAH} , bottom) of galaxies individually detected with *Herschel* in the CANDELS fields. For the T_{dust} measurement, we consider only those galaxies with at least 3 measured photometric points in the dust continuum so that the dust temperature is relatively well constrained. For the f_{PAH} measurement, we use only those galaxies with at least one measured photometric point around the rest-frame 8 μm and at least one point in the dust continuum. We overplot the trends found in stacking (Sect. 3.2.1) with *solid black lines*, as well as the sliding medians on the detection with *solid red lines*. *Middle* Calibration of the evolution of T_{dust} (top) and f_{PAH} (bottom) with the offset from the Main Sequence (R_{SB} , see text). The legend is the same as for the plot on the *left*, except that here the *black solid line* shows our best-fit linear relation to the data. *Right* Distribution of T_{dust} (top) and f_{PAH} (bottom) after removing the redshift evolution as well as the starburstiness trend. The scatter of this distribution, measured with $1.48 \times \text{median}|\Delta X|$, where X is either T_{dust} or $\log_{10}(f_{\text{PAH}})$, is shown in inset (Color figure online)

In Fig. 3.4 (bottom) we apply exactly the same procedure for f_{PAH} . Consistently with the results of Elbaz et al. (2011), we find an anti-correlation between f_{PAH} and $\log_{10}(R_{\text{SB}})$, meaning that starburst galaxies have depressed PAH emission, which Elbaz et al. (2011) also interpret as a sign of increased compactness of the star-forming regions (we discuss this interpretation in Sect. 3.3). We find that the f_{PAH} of individual galaxies can be well described by

$$f_{\text{PAH}} = f_{\text{PAH}}^{\text{MS}} \times R_{\text{SB}}^{-0.47}, \quad (3.11)$$

where $f_{\text{PAH}}^{\text{MS}}$ is defined in Eq. 3.8, with a log-normal scatter of 0.28 dex.

3.3 On the Redshift and Stellar-Mass Dependence of f_{PAH}

It is known, at least in the Local Universe, that the strength of the PAH features is well correlated with the gas-phase metallicity, with more metal-rich galaxies having more pronounced PAH emission (e.g., Galliano et al. 2003; Ciesla et al. 2014). One plausible explanation is that a metal-poor ISM blocks less efficiently the UV radiation of young stars, and makes it harder for PAH molecules to survive (e.g., Galliano et al. 2003). Other scenarios have been put forward, suggesting either that low metallicity objects are just too young to host enough carbon grains to form PAH complexes (Galliano et al. 2008), or that this is instead caused by a different filling factor of molecular clouds in metal poor environments (Sandstrom et al. 2012). Metallicity, in turn, is positively correlated with the stellar mass through the mass-metallicity relation (Lequeux et al. 1979; Tremonti et al. 2004), and this relation has been found to evolve with time, so that galaxies were more metal-poor in the past (e.g., Erb et al. 2006). One therefore expects to find the strongest PAH features within massive low-redshift galaxies, which is indeed what we observe in Fig. 3.3.

To ease the comparison, we show in Fig. 3.5 the relation between our stacked f_{PAH} measurements and the average metallicity of the stacked galaxies. Because we do not have individual metallicity measurements for all these galaxies, we follow

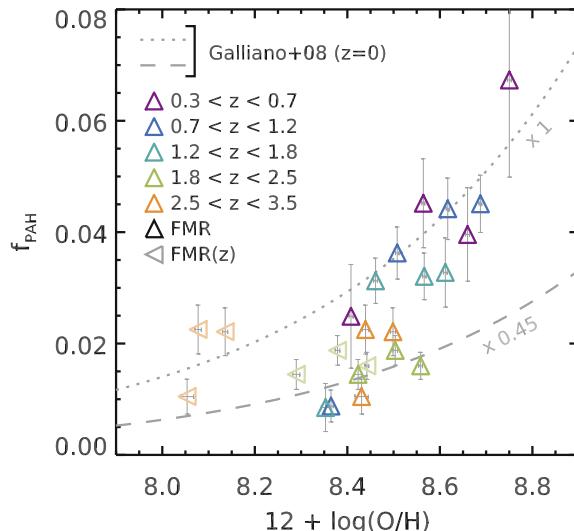


Fig. 3.5 Relation between the PAH mass fraction f_{PAH} observed in stacked *Spitzer* and *Herschel* photometry, and the gas-phase metallicity (given here in terms of oxygen abundance, $12 + \log_{10}(O/H)$, where the solar value is 8.69, as given in Allende Prieto et al. 2001), which is estimated either using the Fundamental Metallicity Relation (FMR, Mannucci et al. 2010) (upward facing triangles) or a broken FMR (Béthermin et al. 2015) (leftward facing triangles). The $z = 0$ relation obtained by Galliano et al. (2008) is shown for reference with a dotted gray line, and the same relation rescaled by a factor 0.45 is shown with a dashed gray line

Magdis et al. (2012) and estimate the average metallicity (more precisely, the oxygen abundance $12 + \log_{10}(O/H)$) using the Fundamental Metallicity Relation (FMR, Mannucci et al. 2010) and the average measured stellar masses and star formation rates. We then compare our results to the Local Universe relation reported by Galliano et al. (2008),

$$\log_{10}(f_{\text{PAH}})_{z=0} = -1.3 + 0.8 \times \log_{10}(Z/Z_{\odot}), \quad (3.12)$$

where we assume $Z_{\odot} = 0.017$ (Grevesse and Sauval 1998) and $12 + \log_{10}(O/H)_{\odot} = 8.69$ (Allende Prieto et al. 2001).

We first consider the galaxies more massive than $10^{10} M_{\odot}$. The values we measure in the two lowest redshift bins ($0.3 < z < 1.2$) are in agreement with the $z = 0$ trend. A tension starts to appear at $1.2 < z < 1.8$, and at $z > 1.8$ we find values that are systematically below the fiducial relation by a factor of about two. Such a discrepancy could be caused by uncertainties in the metallicity of high-redshift galaxies, since most metallicity measurements at $z > 2$ are based on indirect tracers like N II (e.g., Erb et al. 2006). However, the mass fraction we measure at $z = 2$, $f_{\text{PAH}} = 1\%$, would imply an oxygen abundance of 7.8 using the relation of Galliano et al. (2008), as compared to the 8.4 we inferred from the FMR. Such an offset of 0.6 dex is substantially larger than the known systematics in the metallicity calibration (e.g., Kewley and Ellison 2008), and we can therefore rule out this interpretation. That being said, it could also be that the FMR is not applicable to $z > 2$ galaxies. This possibility was recently explored by Béthermin et al. (2015), who introduce a “broken” FMR with a redshift dependence. This alternative prescription predicts metallicities at $z > 1.7$ that are systematically lower than the original FMR, and increasingly so with increasing redshift. As can be seen from Fig. 3.5, this modification goes indeed in the right direction, as our $z = 3$ measurements become consistent with the $z = 0$ trend using this alternative FMR. On the other hand, the $z = 2$ values remain systematically offset, suggesting the existence of a real physical difference at $z = 2$.

This can be put in perspective with the work of Elbaz et al. (2011), who argue that the large IR8 (and therefore the low f_{PAH}) observed in starbursting galaxies is caused by an increased compactness of the star-forming regions. PAH molecules are indeed expected to survive and shine mostly within H II regions. If the star-forming regions are more tightly packed, the H II regions tend to percolate, and the volume in which PAHs can survive is reduced. The same interpretation can be invoked for our $z = 2$ measurements, in particular since high-redshift galaxies are known to be more compact than their low-redshift equivalents (e.g., Ferguson et al. 2004). Observed differences in the dust attenuation properties of high-redshift Main Sequence galaxies also suggest such a change of geometry (Pannella et al. 2015).

However, one important caveat of the present analysis is that f_{PAH} is in fact degenerate with two other unrelated phenomena. First, buried AGNs are known to emit the bulk of their light in the mid-IR (e.g., Mullaney et al. 2011), and can therefore largely perturb the measurement of f_{PAH} . Bright AGNs are not very common and can be easily identified, e.g., from their IRAC colors (Donley et al. 2012) or X-ray luminosity. Elbaz et al. (2011) found however that these AGNs have the same IR8

distribution as normal galaxies. Lower luminosity AGNs could be more numerous, and significantly harder to detect, but their impact on the MIR luminosity, if any, should be even more subtle. Second, another way to increase the $8\text{ }\mu\text{m}$ luminosity is to increase the number of very small dust grains, which are heated at several hundreds to thousands of Kelvins.

Both these phenomena could contribute to some extent to the $8\text{ }\mu\text{m}$ continuum. With MIR spectroscopy, as was provided by *Spitzer* IRS, the contribution of PAHs can be determined accurately, but unfortunately very few galaxies in our surveys have IRS spectra. More progress will be possible as soon as *JWST* is launched.

3.4 Appendix: Recipe for Optimal FIR SED Fitting

Contrary to the standard FIR libraries from the literature (e.g., CE01), the one we introduce here has one more degree of freedom: the PAH mass fraction f_{PAH} . This parameter can only be constrained if observations in both the rest-frame $6\text{--}15\text{ }\mu\text{m}$ and $20\text{--}200\text{ }\mu\text{m}$ are available. If this criterion is not satisfied, the fit is degenerate, and the best we can do is to fix f_{PAH} to its most probable value. On the other hand, T_{dust} can only be constrained if the dust continuum is probed with a sufficiently large wavelength range, e.g., with at least 3 *Herschel* bands, and especially if the available photometry covers the peak of the FIR emission (which, as can be seen in Fig. 3.1, is not the case at $z \geq 4$). Similarly, the best thing we can do in this case is to fix this parameter to our most accurate prediction. Using only the knowledge of the galaxy's redshift, the procedure we recommend is described in the following.

Selection of Free Parameters

- If the number of $S/N > 3$ measurements probing the rest-frame $15\text{ }\mu\text{m}\text{--}3\text{ mm}$ is less than 3, or if the measurements do not cover both $\lambda < 80\text{ }\mu\text{m}$ and $\lambda > 80\text{ }\mu\text{m}$, fix $T_{\text{dust}} = T_{\text{dust}}^{\text{MS}}$ from Eq. 3.7. In all other cases, let T_{dust} vary to all the values permitted by the library.
- If no $S/N > 3$ measurement probes the rest-frame $5\text{--}15\text{ }\mu\text{m}$, or if all measurements are within this range, fix $f_{\text{PAH}} = f_{\text{PAH}}^{\text{MS}}$ from Eq. 3.8. Else, let f_{PAH} vary from 0 to 1.

Fitting the Observed Photometry

- If all parameters are free, iterate over all the $\langle U \rangle$ in the library. For each value, fit to the observed photometry a linear combination of the dust continuum and PAH template associated to this given $\langle U \rangle$:

$$S_v^{\text{model}} = \textcolor{blue}{M}_{\text{dust}}^{\text{cont}} S_v^{\text{cont}} + \textcolor{blue}{M}_{\text{dust}}^{\text{PAH}} S_v^{\text{PAH}}, \quad (3.13)$$

where the two free parameters are highlighted in blue, and compute the χ^2 . Among all the templates in the library, pick as the best-fit solution the $\langle U \rangle$ value which

produced the smallest χ^2 . If this solution has a negative $M_{\text{dust}}^{\text{cont}}$ or $M_{\text{dust}}^{\text{PAH}}$, start over by fixing $f_{\text{PAH}} = f_{\text{PAH}}^{\text{MS}}$ and then $f_{\text{PAH}} = 0$ and pick the best χ^2 . Else, the dust mass is computed as $M_{\text{dust}} = M_{\text{dust}}^{\text{cont}} + M_{\text{dust}}^{\text{PAH}}$. The PAH mass fraction is computed as the ratio between $M_{\text{dust}}^{\text{PAH}}$ and M_{dust} . The dust temperature can be read directly from the library using the index of the best-fit $\langle U \rangle$. Finally, the infrared luminosity is computed by summing up the L_{IR} associated to both the dust continuum and PAH templates in the library (which are given per unit M_{dust}), weighted by their respective dust mass:

$$L_{\text{IR}} = M_{\text{dust}}^{\text{cont}} L_{\text{IR}}^{\text{cont}} + M_{\text{dust}}^{\text{PAH}} L_{\text{IR}}^{\text{PAH}}. \quad (3.14)$$

- If f_{PAH} is fixed, iterate over all the $\langle U \rangle$ in the library. For each value, compute the combined template by summing together the dust continuum and the PAH templates associated to this $\langle U \rangle$, fit this single template to the observed photometry:

$$S_{\nu}^{\text{model}} = \textcolor{blue}{M_{\text{dust}}} \left[(1 - f_{\text{PAH}}) \times S_{\nu}^{\text{cont}} + f_{\text{PAH}} \times S_{\nu}^{\text{PAH}} \right], \quad (3.15)$$

where the only free parameter is highlighted in blue, and then compute the χ^2 . Similarly to the procedure above, choose as the best-fit solution the value of $\langle U \rangle$ that produced the smallest χ^2 . Then let $M_{\text{dust}}^{\text{cont}} = (1 - f_{\text{PAH}}) M_{\text{dust}}$ and $M_{\text{dust}}^{\text{PAH}} = f_{\text{PAH}} M_{\text{dust}}$, and derive the other quantities as described above.

- If both T_{dust} and f_{PAH} are fixed, locate the $\langle U \rangle$ that corresponds to the chosen T_{dust} in the library. Combine both dust continuum and PAH templates as described above, and follow the rest of the procedure by only considering the unique template.
- In all cases, if the χ^2 of the best-fit is larger than 10 and the fit used photometry below 30 μm rest-frame, then remove these measurements, fix $f_{\text{PAH}} = f_{\text{PAH}}^{\text{MS}}$ or $f_{\text{PAH}} = 0$, do the fit again, and pick the best χ^2 . If this new χ^2 is smaller than the old one, consider this new fit as the best. Else, fall back to the original fit. One would expect the χ^2 to be systematically lower, because there are fewer points to fit, but it should also be taken into account that fixing f_{PAH} will remove one degree of freedom, and can thus make the fit worse. This step takes care of contamination by AGNs, and also uncertain redshifts that make PAH emission lines appear where the library doesn't expect them to.

Computing Uncertainties on Best-Fit Parameters

- Add a random perturbation to each measured photometric point, as a Gaussian scatter of amplitude set by the estimated measurement uncertainty. Then redo the fit as described above.
- Repeat the procedure 100 times and, for each fit parameter, the uncertainty is given by the RMS of the difference between the best-fit value measured on the real photometry and the best-fit value of each 100 noise realizations.

If only a single photometric band is available, the accuracy of monochromatic L_{IR} or M_{dust} measurements is quantified using a simulated galaxy catalog in Chap. 4, Sect. 4.7.

References

- M. Béthermin, C. De Breuck, M. Sargent, E. Daddi, A&A **576**, L9 (2015)
- R. Chary, D. Elbaz, ApJ **556**, 562 (2001)
- L. Ciesla, M. Boquien, A. Boselli et al., A&A **565**, A128 (2014)
- E. Daddi, D.M. Alexander, M. Dickinson et al., ApJ **670**, 173 (2007)
- D.A. Dale, G. Helou, A. Contursi, N.A. Silbermann, S. Kolhatkar, ApJ **549**, 215 (2001)
- D.A. Dale, G. Helou, ApJ **576**, 159 (2002)
- J.L. Donley, A.M. Koekemoer, M. Brusa et al., ApJ **748**, 142 (2012)
- B.T. Draine, A. Li, ApJ **657**, 810 (2007)
- D. Elbaz, M. Dickinson, H.S. Hwang et al., A&A **533**, 119 (2011)
- D.K. Erb, A.E. Shapley, M. Pettini et al., ApJ **644**, 813 (2006)
- H.C. Ferguson, M. Dickinson, M. Giavalisco et al., ApJ **600**, L107 (2004)
- F. Galliano, S.C. Madden, A.P. Jones et al., A&A **407**, 159 (2003)
- F. Galliano, E. Dwek, P. Chanial, ApJ **672**, 214 (2008)
- F. Galliano, S. Hony, J. Bernard et al., A&A **536**, A88 (2011)
- N. Grevesse, A.J. Sauval, Space Sci. Rev. **85**, 161 (1998)
- L.J. Kewley, S.L. Ellison, ApJ **681**, 1183 (2008)
- J. Lequeux, M. Peimbert, J.F. Rayo, A. Serrano, S. Torres-Peimbert, A&A **80**, 155 (1979)
- G.E. Magdis, E. Daddi, M. Béthermin et al., ApJ **760**, 6 (2012)
- B. Magnelli, P. Popesso, S. Berta et al., A&A **553**, 132 (2013)
- B. Magnelli, D. Lutz, A. Saintonge et al., A&A **561**, 86 (2014)
- F. Mannucci, G. Cresci, R. Maiolino, A. Marconi, A. Gnerucci, MNRAS **408**, 2115 (2010)
- J.S. Mathis, P.G. Mezger, N. Panagia, A&A **128**, 212 (1983)
- J.R. Mullaney, D.M. Alexander, A.D. Goulding, R.C. Hickox, MNRAS **414**, 1082 (2011)
- M. Pannella, D. Elbaz, E. Daddi et al., ApJ **807**, 141 (2015)
- C.A. Prieto, D.L. Lambert, M. Asplund, ApJ, **556**, L63 (2001)
- E.E. Salpeter, ApJ **121**, 161 (1955)
- K.M. Sandstrom, A.D. Bolatto, C. Bot et al., ApJ **744**, 20 (2012)
- C. Schreiber, M. Pannella, D. Elbaz et al., A&A **575**, A74 (2015)
- H.I. Teplitz, R. Chary, D. Elbaz et al., AJ **141**, 1 (2011)
- C.A. Tremonti, T.M. Heckman, G. Kauffmann et al., ApJ **613**, 898 (2004)
- V. Zubko, E. Dwek, R.G. Arendt, ApJS **152**, 211 (2004)

Chapter 4

EGG: An Empirical Simulation of the Observable Universe

4.1 Introduction

The following work was undertaken as part of the ASTRODEEP collaboration. Briefly, ASTRODEEP is an FP7–SPACE project lead by the PI A. Fontana (INAF, Rome) and the Co-Is D. Elbaz, J. Dunlop (Royal Observatory, Edimbourg), S. Derrière (CDS, Strasbourg), M.E. Dickinson (NOAO, Tucson), H. Ferguson (STScI, Baltimore) and S. Faber (UCO-Lick, Santa Cruz). One of the main goal of ASTRODEEP is to design tools to robustly measure the photometry of galaxies in a variety of images, ranging from the UV-optical to the far-IR. For example, the first product of this collaborative effort is the T-PHOT code (Merlin et al. 2015). This program is meant to be a replacement for TFIT (Laidler et al. 2007), which was used, e.g., to extract the UV-to-NIR photometry in all the CANDELS fields (Guo et al. 2013; Galametz et al. 2013). T-PHOT is more versatile though, since it can also be used to extract FIR fluxes on the *Herschel* images (see, in particular, Wang et al. in prep.).

The performance and accuracy of such a code has to be properly characterized before applying it to real images, to check not only the robustness of the flux measurements, but also the quality of the error estimates. Furthermore, the main challenge when extracting photometry in FIR images is not so much the details of the flux extraction itself, but rather the proper choice of the prior selection strategy. For this reason, a second goal of ASTRODEEP is to provide the astrophysics community with realistic simulations of the sky at different wavelengths, and with different angular resolutions, so that we can test our procedures and tools and quantify their efficiency.

A first rendition of these simulations was based on the *SkyMaker*¹ program (E. Bertin), in order to produce realistic high resolution “*HST-like*” images. In input, this program requires a simulated galaxy catalog, containing morphologies and fluxes, which can be produced by the *Stuff*² program (also created by E. Bertin).

¹<http://www.astromatic.net/software/skymaker>.

²<http://www.astromatic.net/software/stuff>.

The quality of the simulated catalogs generated by *Stuff* is not optimal though. In particular, the distribution of the simulated fluxes in some bands (in particular the *U* band, but not only) differ substantially from those that are observed, leading to simulated images that are not representative of the real products we are working on. Unfortunately, both *SkyMaker* and *Stuff* are poorly documented, and we cannot easily remedy this problem by editing the codes.

For this reason, I have extended the simulations that are described in Chap. 2 (Sect. 2.8) and developed a new tool to generate simulated galaxy catalogs, called EGG³ (the Empirical Galaxy Generator), that I designed and wrote in C++ using the *phy_++* library (see Appendix A). The main ideas behind the procedure are summarized in this chapter, and a paper was just submitted to present the code (Schreiber et al. 2016).

This new tool can generate catalogs in the format required by *SkyMaker*, and therefore can be used as a “drop-in” replacement for *Stuff*. Using this tool I am able not only to generate fluxes in all the photometric bands from 3000 Å to 8 μm, like *Stuff*, but I also merge in my technique to simulate far-IR fluxes from 8 μm to 3 mm (see Chap. 2, Sect. 2.8), essentially covering, in a single tool, the whole wavelength range where stellar and dust emission dominate.

The quality of the generated catalogs has greatly improved compared to original catalogs built with *Stuff*. As can be seen in Sect. 4.6, we are now able to produce flux distributions that are indistinguishable from the real, observed flux distributions, in all bands from *U* to SPIRE 500 μm. The simulated images, both at *Hubble*- and *Herschel*-like resolution, have very good statistical properties. This will allow us to perform more accurate tests of our methods, and also to deliver high quality simulations to the community.

4.2 Sample Description

We base this analysis on the sample and data described in Schreiber et al. (2015) (hereafter S15, see Chap. 2). In this section, we make a brief summary of these observations.

4.2.1 Multi-wavelength Photometry

The catalogs we use in this work are based on the CANDELS (Grogin et al. 2011; Koekemoer et al. 2011) *Hubble Space Telescope* (*HST*) WFC3 *H* band images in the four CANDELS fields that are covered by deep *Herschel* PACS and SPIRE observations, namely GOODS–North (Barro et al. in prep.), GOODS–South (Guo et al. 2013), UDS (Galametz et al. 2013) and COSMOS (Nayyeri et al. in prep.).

³<https://github.com/cschreib/egg>.

Each of these fields is about 150 arcsec² and they are evenly distributed on the sky to mitigate cosmic variance.

The ancillary photometry varies from one field to another, being a combination of both space- and ground-based imaging from various facilities. The UV to near-IR wavelength coverage typically goes from the *U* band up the *Spitzer* IRAC 8 μm, including at least the *HST* bands F606W, F814W, and F160W and a deep *K* (or *K_s*) band, and all these images are among the deepest available views of the sky. These catalogs therefore cover most of the important galaxy spectral features across a wide range of redshifts, even for intrinsically faint objects.

We complement these catalogs with mid-IR photometry from *Spitzer* MIPS and far-IR photometry from *Herschel* PACS and SPIRE taken as part of the GOODS–*Herschel* (Elbaz et al. 2011) and CANDELS–*Herschel* programs (PI: M. Dickinson, Inami et al. in prep.).

4.2.2 Redshifts and Stellar Masses

Photometric redshifts and stellar masses are computed by M. Pannella following Pannella et al. (2015). We use EAZY (Brammer et al. 2008) to derive the photometric redshifts from the CAsDELS catalogs, allowing slight adjustments of the photometric zero points by iteratively comparing our photo-zs against the available spec-zs. The stellar masses are then computed using FAST (Kriek et al. 2009) by fixing the redshift to the best-fit photo-z and fitting the observed photometry up to the IRAC 4.5 μm band using the Bruzual and Charlot (2003) stellar population synthesis model, assuming a Salpeter (1955) IMF and a Calzetti et al. (2000) extinction law.

Galaxies with an uncertain photometric redshift (redshift odds less than 0.8) or bad SED fitting (reduced χ^2 larger than 10) are excluded from our sample. We estimated in S15 the evolution of the stellar mass completeness (90 %) of these catalogs at all redshifts, and in the present study we only consider galaxies above this limit. For example, at $z = 1$ the completeness is as low as $5 \times 10^8 M_{\odot}$.

Lastly, the rest-frame *U*, *V* and *J* magnitudes are computed by M. Pannella for each galaxy using EAZY, by integrating the best-fit galaxy template from the photo-z estimation. These colors are used, following Williams et al. (2009), to separate galaxies that are “quiescent” from those that are “star-forming”. We use the same selection criteria as those described in S15, i.e., a galaxy is deemed quiescent if its colors satisfy

$$\text{quiescent} = \begin{cases} U - V > 1.3, \\ V - J < 1.6, \\ U - V > 0.88 \times (V - J) + 0.49, \end{cases} \quad (4.1)$$

otherwise the galaxy is considered as star-forming.

4.3 Stellar Properties

4.3.1 Redshift and Stellar Mass

The initial purpose of EGG is to simulate a deep field similar to the GOODS–*South* field. Therefore, we compute the stellar mass function in this field only, in order to most closely mimic its properties (including, in particular, cosmic variance). To do so, we use the procedure described in S15, which we now briefly recall.

The whole GOODS–*South* catalog is cut at $H < 26$ to ensure high quality photometry for all galaxies and to reduce the effect of the Eddington bias. We then make multiple redshift bins from $z = 0.3$ to $z = 4.5$, and within each of these bins, we compute the mass distribution of both sub-samples of star-forming and quiescent galaxies separately, according to the UVJ color-color selection (see Sect. 4.2.2). We apply completeness corrections as estimated from the observed scatter in the M_* to $L_{H/(1+z)}$ ratio. Then, we fit a double Schechter law to each distribution:

$$\frac{d^2N(z)}{d \log_{10} M_* dV} = S(M_*, \phi_1^*, M_1^*, \alpha_1) + S(M_*, \phi_2^*, M_2^*, \alpha_2),$$

$$S(M_*, \phi^*, M^*, \alpha) \equiv \log(10) \phi^* \left(\frac{M_*}{M^*}\right)^{\alpha+1} \exp\left(-\frac{M_*}{M^*}\right). \quad (4.2)$$

The results are shown in Fig. 4.1, and the best-fit parameters are summarized in Tables 4.1 and 4.2. Our goal is only to find a functional form that describes well the observed data. We thus attribute no physical origin to each component of the double Schechter law, and because the fit is quite degenerate, we allow ourselves to arbitrarily fix some of the fit parameters. These are surrounded by brackets in the tables.

We estimated that our catalog is not complete to assess the mass function of $z = 4$ passive galaxies, and therefore do not attempt to fit it. Instead, we use the same parameters as that obtained at lower redshifts and only adjust ϕ^* to have a fraction of passive galaxies equal to 15 % (for $M_* > 4 \times 10^{10} M_\odot$), the extrapolation of the trend we observe at lower redshifts. This is consistent with what was previously reported by, e.g., Muzzin et al. (2013). However, Straatman et al. (2014) suggested recently that this fraction could be substantially higher, since they found 34 % of passive galaxies at $z = 3.7$ using ZFOURGE, a deep medium-band NIR survey. In any case, this will not change dramatically the quality of our simulated catalogs, because the number density of these objects is still very low, and also because they would only be detectable in the NIR images, where they would have a negligible impact. In contrast, this is not true for $z \geq 4$ star-forming galaxies, which can be among the brightest sources in a sub-mm image.

To reach higher redshifts, we use the recent results of Grazian et al. (2015) for $4.5 < z < 7.5$. Since their stellar mass functions are not split between star-forming and quiescent galaxies, here we assume that the double Schechter function we observe

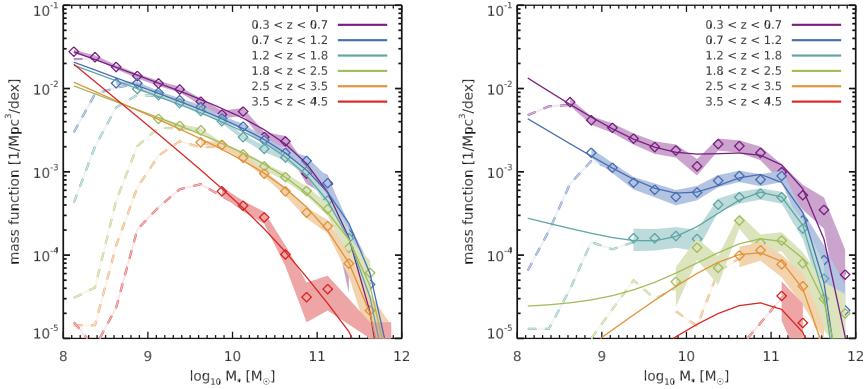


Fig. 4.1 Conditional stellar mass function at different redshifts for star-forming (*left*) and quiescent (*right*), selected with $H < 26$. The *dashed lines* in the background indicate the raw mass functions, before completeness corrections are applied. The *solid colored regions* show the completeness-corrected estimate of the mass function, and the width of the region indicates the statistical uncertainty on the measurement (i.e., Poisson noise)

Table 4.1 Double Schechter function parameters for the star-forming galaxy population. Parameters that were chosen manually are enclosed in brackets

z	ϕ_1^* $\text{dex}^{-1} \text{Mpc}^{-3}$	$\log_{10}(M_1^*)$ $\log_{10}(M_\odot)$	α_1	ϕ_2^* $\text{dex}^{-1} \text{Mpc}^{-3}$	$\log_{10}(M_2^*)$ $\log_{10}(M_\odot)$	α_2
0.3–0.7	1.04×10^{-3}	[11]	[-1.37]	0	[11]	[0.5]
0.7–1.2	7.77×10^{-4}	[11]	[-1.37]	1.72×10^{-4}	[11]	[0.5]
1.2–1.8	7.14×10^{-4}	[11]	[-1.37]	6.56×10^{-5}	[11]	[0.5]
1.8–2.5	3.87×10^{-4}	[11]	[-1.37]	1.2×10^{-4}	[11]	[0.5]
2.5–3.5	2.77×10^{-4}	[11]	-1.44	4.95×10^{-5}	[11]	[0.5]
3.5–4.5	3.4×10^{-5}	[11]	-1.83	1.01×10^{-5}	[11]	[0.5]

Table 4.2 Double Schechter function parameters for the quiescent galaxy population. Parameters that were chosen manually are enclosed in brackets

z	ϕ_1^* $\text{dex}^{-1} \text{Mpc}^{-3}$	$\log_{10}(M_1^*)$ $\log_{10}(M_\odot)$	α_1	ϕ_2^* $\text{dex}^{-1} \text{Mpc}^{-3}$	$\log_{10}(M_2^*)$ $\log_{10}(M_\odot)$	α_2
0.3–0.7	1.84×10^{-4}	[11]	[-1.50]	1.83×10^{-3}	10.91	-0.14
0.7–1.2	4.67×10^{-5}	[11]	[-1.60]	3.73×10^{-3}	10.86	0.07
1.2–1.9	3.55×10^{-5}	[11]	[-1.35]	1.76×10^{-3}	10.85	0.26
1.9–2.5	0	[11]	[-1.35]	1.28×10^{-4}	11.01	-0.33
2.5–3.5	0	[11]	[-1.35]	1.34×10^{-4}	10.94	-0.26
3.5–4.5	0	[11]	[-1.35]	$[1.10 \times 10^{-5}]$	[11]	[-0.30]

at $z = 4$ also holds at $z > 4$, and we simply decrease the integrated stellar mass density to mimic the evolution observed by Grazian et al. (2015), keeping a constant quiescent galaxy fraction of 15 %. We then extrapolate these trends to reach up to $z = 9$. On the other hand, the $z = 0$ mass function is adapted from Baldry et al. (2012), but this should be of little importance since we are aiming for pencil-beam surveys containing almost no local galaxies. Extrapolating these combined mass functions toward the low-mass end, assuming that the low-mass slope is not varying, we can generate galaxies of all stellar masses in an arbitrary volume between $z = 0$ and $z = 9$.

4.3.2 Star Formation Rate and Obscuration

Given the redshift and the stellar mass, we can attribute a star formation rate (SFR) to each galaxy by following the Two Star Formation Mode model (2SFM, Sargent et al. 2012). This model is based on the Main Sequence of star-forming galaxies, i.e., the observed correlation between the SFR and the stellar mass of actively star-forming systems (Noeske et al. 2007; Elbaz et al. 2007). This approach has been applied in Béthermin et al. (2012) to successfully reproduce the observed flux distribution from the MIR to the sub-mm and even the radio. Building on this strength, we use here a similar prescription where the model parameters are updated to our latest measurements.

Using the $\text{SFR}(z, M_*)$ equation published in S15, we associate a “Main Sequence” star formation rate to each star-forming galaxy with

$$\log_{10}(\text{SFR}_{\text{MS}}[\text{M}_\odot/\text{yr}]) = m - 9.5 + 1.5 r - 0.3 [\max(0, m - 9.36 - 2.5 r)]^2, \quad (4.3)$$

where $m \equiv \log_{10}(M_*[\text{M}_\odot])$, and $r \equiv \log_{10}(1 + z)$

We then apply a log-normal scatter of 0.3 dex to reproduce the observed width of the Main Sequence, which was found in S15 to be constant both as function of stellar mass and redshift. In addition, 3 % of the galaxies are randomly chosen and placed in a “starburst” mode, where their SFR is enhanced by a factor of ~ 6 , following the observed distribution of SFRs about the Main Sequence in S15. Sargent et al. (2012) showed that this last step is necessary to correctly capture the bright-end of the IR luminosity functions. We parametrize both the starburst component and the scatter of the Main Sequence using the “starburstiness” R_{SB} :

$$R_{\text{SB}} \equiv \frac{\text{SFR}}{\text{SFR}_{\text{MS}}} = \begin{cases} 1 & \text{for Main Sequence galaxies} \\ 5.24 & \text{for starburst galaxies} \end{cases}, \quad (4.4)$$

to which we add a log-normal scatter of 0.3 dex.

For quiescent galaxies, we use the IR stacks presented in the Appendix of S15 (Chap. 2, Sect. 2.7), where it was reported that quiescent galaxies do show some IR emission, typically a factor of ten fainter than star-forming galaxies of the same mass. This light may be caused either by residual star formation, or by dust heated by old stars, or by incorrect classification of some star-forming galaxies. Although this is an interesting question, its answer is irrelevant for our purposes, and we choose to model this faint emission by interpreting it as residual star-formation. Therefore, quiescent galaxies are attributed an SFR following

$$\log_{10}(\text{SFR}_{\text{QS}}[\text{M}_\odot/\text{yr}]) = 0.5m + r - 6.1, \quad (4.5)$$

to which we add a log-normal scatter of 0.45 dex.

To prepare the ground for the treatment of dust, we then decompose these SFRs into a dust-obscured component, which re-emerges in the FIR, and dust-free component, which emerges in the UV. To do so, we use the observed relation between stellar mass and dust obscuration (e.g., Pannella et al. 2009; Buat et al. 2012; Heinis et al. 2014), which we calibrate here in terms of the infrared excess $\text{IRX} \equiv \log_{10}(L_{\text{IR}}/L_{\text{UV}})$ (Meurer et al. 1999) following Heinis et al. (2014). Using the stacked L_{IR} from S15, we find that the relation between IRX and M_* can be described by

$$\text{IRX} = (0.45 \min(3.0, z) + 0.35) \times (m - 10.5) + 1.2. \quad (4.6)$$

This formula is very similar to that reported by Heinis et al. (2014), save from the fact that our relation is found to be redshift dependent. This is consistent with the finding of Pannella et al. (2015), who found that the typical MS galaxy at $z = 2$ is sensibly different from its analog at $z \leq 1$, which they argue is because of modifications in the geometry of the star-forming regions. We also add a scatter of 0.4 dex to this relation: although it has a negligible impact on the generated IR luminosities, Bernhard et al. (2014) showed that this is a necessary ingredient to properly reproduce the bright-end of the UV luminosity function.

4.3.3 Optical Morphology

Following the approach of *Stuff* and *SkyMaker*, we consider here that galaxies are mostly made of two components: a bulge (Sérsic index $n = 4$) and a disk (Sérsic index $n = 1$). Each of these components is described by a number of morphological parameters, including the position angle θ , the axis ratio b/a , and the half-light radius R_{50} . Also, the fraction of the stellar mass that goes into one or the other component is dictated by the bulge-to-total ratio (B/T). In the following, we present how we calibrate each of these parameters.

The bulge-to-total ratio is estimated following the results of Lang et al. (2014) who conveniently measured the average B/T as a function of stellar mass for both *UVJ* quiescent and star-forming galaxies in the CANDELS fields at different redshifts.

While they found the bulge fraction to increase with stellar mass for both populations, they did not observe any significant difference with redshift between $z = 1$ and $z = 2$, so we chose to make the B/T simply depend on mass following

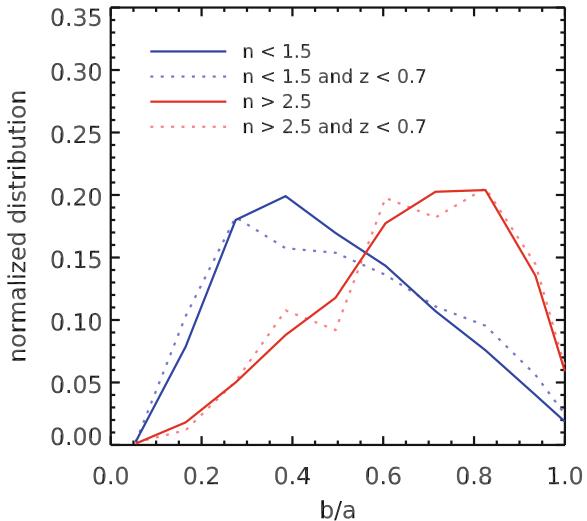
$$\log_{10}(B/T)_{\text{active}} = -0.7 + 0.27 \times (m - 10) \text{ and} \\ \log_{10}(B/T)_{\text{passive}} = -0.3 + 0.1 \times (m - 10), \quad (4.7)$$

to which we add a log-normal scatter of 0.2 dex in order to reproduce the width of the distribution reported by Lang et al. (2014). The final B/T is limited to be no more than 1. Note that this value is a *mass-weighted* bulge-to-total ratio, therefore we can directly use it to compute the stellar mass inside the disk and the bulge. Estimating the contribution of each component to the *light* of the galaxy is done in Sect. 4.3.4.

Then, we attribute a uniformly random position angle to each galaxy, and assign this same angle to both the bulge and the disk components. To calibrate the other morphological parameters of both bulges and disks, we use the morphological catalogs of van der Wel et al. (2012) who fit single Sérsic profiles of varying index n to all galaxies in the CANDELS fields using the *Galfit* software (Peng et al. 2002) on the *HST* *H*-band images. In the following, we will consider two sub-samples: first, galaxies with $n < 1.5$ and $M_* > 10^9 M_\odot$, second, galaxies with $n > 2.5$ and $M_* > 3 \times 10^{10} M_\odot$. The cut in stellar mass is used to select galaxies bright enough that the Sérsic fit are reliable, and to prevent our trends to be dominated by the numerous low-mass galaxies. We use these sub-samples to calibrate the morphology of the disk and bulges, respectively. Indeed, for galaxies with $n < 1.5$ the presence of a bulge can be neglected so that the measured properties can be attributed to the disk alone (see, e.g., the Appendix of Lang et al. 2014), and conversely for $n > 2.5$. This latter sample of $n > 2.5$ galaxies is probably less pure though, since high Sérsic indices can be produced either by a dominant bulge, or by a minor bulge that has a much smaller half-light radius than the disk, as shown in the Appendix of Lang et al. (2014). However these extreme cases are relatively rare, and the majority of $n > 2.5$ galaxies are indeed bulge dominated.

For each sub-sample, we start by measuring the axis ratio distribution (Fig. 4.2). We find, as expected, that bulge-dominated galaxies have more circular shapes, their b/a distribution peaking at 0.8, while disk-dominated galaxies are more elongated, with a peak b/a of 0.3. We consider that these distributions hold for all masses and all redshifts. van der Wel et al. (2014) reported that the b/a distribution of *UVJ* star-forming galaxies at $z = 1.7$ shows a clear mass evolution from 10^9 to $10^{11} M_\odot$: while the low-mass distribution is very similar to our disk-dominated distribution, the high-mass distribution is found to be bimodal. Without attempting to demonstrate it, we argue here that this trend is likely the result of the increase of the B/T with stellar mass among star-forming galaxies (Lang et al. 2014). On the one hand, low-mass galaxies are preferentially bulgeless, and should therefore follow the trend of pure-disks of Fig. 4.2. On the other hand, high-mass galaxies are more complex systems with a varying mixture of bulges and disks; among those, we expect to find both bulge- and disk-dominated systems, and this would explain the bimodal distribution observed

Fig. 4.2 Observed axis-ratio distribution of disk-dominated ($n < 1.5$, blue) and bulge-dominated ($n > 2.5$, red) galaxies. These values are taken from the Sérsic profile fits produced by van der Wel et al. (2012). We also show the distributions for the sub-samples at $z < 0.7$, in dotted lines of the same colors



by van der Wel et al. (2014). However, by comparing their $z = 1.7$ result to a similar analysis in the Sloan Digital Sky Survey (SDSS, $z = 0$), van der Wel et al. (2014) showed that these distributions are also redshift dependent. One possible explanation for this would be that the redshift invariance of the $B/T - M_*$ relation found by Lang et al. (2014) may not hold at $z < 1$, indicating that galaxies in the local Universe have more prominent bulges at fixed stellar mass. As shown in Fig. 4.2, we do find a similar trend at $z < 0.7$ (blue dotted line), but the difference is small enough that it can be safely neglected for our purposes.

The next step is the calibration of the half-light radius. It is known that the proper size of a galaxy correlates with its stellar mass (i.e., the mass–size relation), and also that galaxies were overall smaller (hence more compact) in the past (e.g., Ferguson et al. 2004). For this reason, we bin our two sub-samples in stellar mass and observe the evolution of the median half-light radius with redshift. The observed trends are reported in Fig. 4.3. We parametrize these relations with the following equations, for disks:

$$\log_{10}(R_{50,\text{disk}}[\text{kpc}]) = 0.45 + 0.2 \times (m - 9.35) + F_z, \quad (4.8)$$

$$\text{with } F_z = \begin{cases} -0.3 \times r & \text{for } z \leq 1.7 \\ -0.7 \times r + 0.17 & \text{for } z > 1.7 \end{cases},$$

and for bulges:

$$\log_{10}(R_{50,\text{bulge}}[\text{kpc}]) = 0.9 + 0.56 \times (m - 11.25) - 1.3 \times r, \quad (4.9)$$

to which we add a log-normal scatter of 0.25 dex. Although this latter value is smaller than what was reported, e.g., by Shen et al. (2003) or Dutton et al. (2011), we find

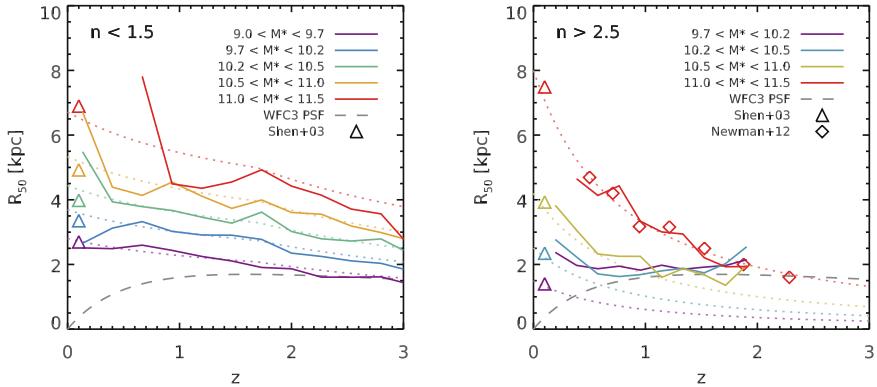


Fig. 4.3 Observed relation between the half-light radius R_{50} and redshift of disk-dominated galaxies (left) and bulge-dominated galaxies (right). Different stellar mass bins are shown with different colors as indicated in the legend. The median values over all CANDELS fields are shown with *solid colored lines*, and the prescription adopted in this work is displayed with a *dotted line* in the background. Empty triangles at $z = 0.1$ are the values obtained by Shen et al. (2003) in the SDSS, converted from the Kroupa to Salpeter IMF, and multiplied by a factor 1.4 to correct for the fact that their radii are measured in circularized apertures (Dutton et al. 2011). We also show how the size of the *HST H*-band PSF ($0.2''$) translates into proper distance with a *long dashed line*. Measuring sizes below this line is difficult, and we expect this to cause biases in the measured median radii (as is probably happening for the low-mass bulge-dominated galaxies). Finally, for bulge-dominated galaxies, we also display the size measurements of Newman et al. (2012), which were obtained by selecting passive galaxies based on their sSFR from SED fitting. Their values are reported as $R_{50}/M_{11}^{0.57}$, which we renormalize to the stellar mass of our highest mass bin

that it is sufficient to reproduce the observed scatter in the mass–size relation, where the actual half-light radius of the *whole* galaxy (i.e., combining both bulge and disk components) is approximated as

$$R_{50,\text{total}} = R_{50,\text{disk}} \times (1 - (B/T)^\alpha) + R_{50,\text{bulge}} \times (B/T)^\alpha, \quad (4.10)$$

where $\alpha = 0.8$ for $(R_{50,\text{bulge}}/R_{50,\text{disk}}) < 1$, and $\alpha = 2$ otherwise. This empirical relation was obtained by computing numerically the half-light radius of simulated double Sérsic profiles ($n = 1$ and $n = 4$) of varying sizes and relative flux.

To preserve the normalization of the mass–size relation in composite systems, we use the *total mass* M_* to derive each component’s respective size. However, although the above $R_{50,\text{bulge}}$ provides a good description of bulge-dominated objects, we find that using this same prescription for the bulges of disk-dominated galaxies leads to total sizes that are systematically smaller than what is observed. Instead, adopting the prescription of $R_{50,\text{disk}}$ for these bulges gives a perfect fit. This would mean that the bulges found in disk-dominated galaxies are substantially different, and in particular less compact, than those found in bulge-dominated galaxies. Alternatively, it could also mean that the reported bulge fractions among star-forming galaxies reported by Lang et al. (2014) are overestimated.

4.3.4 Optical Spectral Energy Distribution

Once the main physical properties are generated, we can associate an optical SED to both the disk and bulge component of every galaxy. Instead of basing our approach solely on physical arguments, e.g., stellar age and dust content, we choose a simpler effective prescription where the SED is chosen based on the position of the galaxy on the UVJ diagram. Indeed, this color-color diagram is a good way to describe a wide range of spectral types, in particular “blue and star-forming”, “red and dead” and “red and dust-obscured”.

Our first goal is therefore to find a recipe to generate $U - V$ and $V - J$ colors for every galaxy, with realistic statistical properties. Since galaxies are segregated into two groups (or “clouds”) of quiescent and star-forming galaxies, we will treat each population separately, starting from the star-forming galaxies.

To this end, we consider all the UVJ star-forming galaxies in the CANDELS catalogs and split this sample in bins of redshift and stellar mass. In each of these bins, we compute the median $U - V$ and $V - J$ colors, and display the resulting tracks on the UVJ diagram in Fig. 4.4 (left). One can see that, at a given stellar mass, both colors go from blue to red as redshift decreases, and same trends could already be seen in S15. We interpret this as a combination of varying dust content (see previous section and Pannella et al. 2015) and age. Interestingly, we find here that all the tracks seem to follow a single straight line that we call the “ UVJ sequence” (see also Labb   et al. 2007 where such a sequence is found among blue galaxies in a color-magnitude diagram). In Fig. 4.4 (right) we show the projection of the tracks on

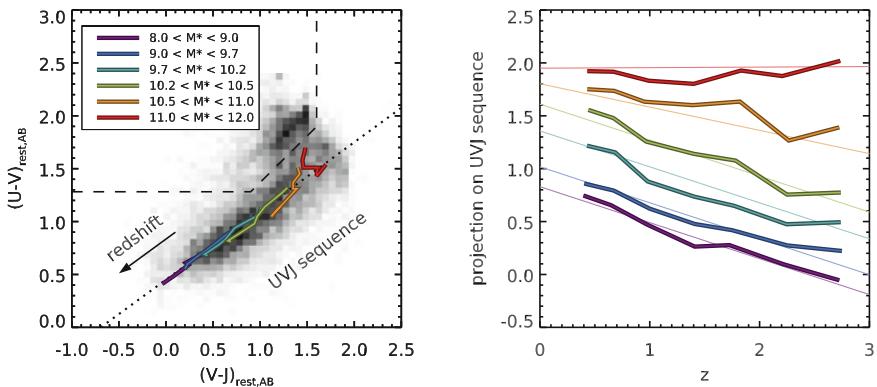


Fig. 4.4 *Left* UVJ diagram of all galaxies in the CANDELS fields more massive than $10^{10} M_\odot$ (background gray scale). The redshift evolution of the median $U - V$ and $V - J$ colors of star-forming galaxies in different stellar mass bins is shown with *colored lines*. They all fall along a single line we dub the “ UVJ sequence”, which is illustrated by a dotted line. Finally, the adopted dividing line between quiescent and star-forming galaxies is shown with a *long dashed line*. *Right* projection of each individual track on the UVJ sequence. Low values indicate bluer colors. We show in the background the prescription adopted in this work

this line, to illustrate more clearly the trend with redshift. Massive galaxies seem to always show the same very red colors, while less massive galaxies were substantially bluer in the past.

We parametrize the evolution both with mass and redshift of the projection A using the following equations:

$$A = A_0 + A_s \times z, \quad (4.11)$$

$$\text{with } A_0 = 0.58 \times \text{erf}(m - 10) + 1.39,$$

$$\text{and } A_s = -0.34 + 0.3 \times \max(0, m - 10.35),$$

and add a Gaussian scatter of 0.1 magnitudes. The resulting value is limited to be at most 2 to prevent nonphysical extremely red colors. Finally, to recover the $U - V$ and $V - J$ colors, we use the definition of the UVJ sequence:

$$(V - J)_{\text{active}} = A \times \cos(\theta), \\ (U - V)_{\text{active}} = 0.45 + A \times \sin(\theta). \quad (4.12)$$

with $\theta = \arctan(0.65)$, and we add a Gaussian scatter of 0.12 magnitudes to reproduce a wider variety of colors than what is allowed by this simple prescription. This scatter is most likely caused by variations of star formation histories or specific star formation rates (see, e.g., Arnouts et al. 2013), but we do not attempt to dissect its origin here.

For quiescent galaxies, the prescription is relatively simpler, since they are mostly found within a small region of the UVJ diagram, the so-called red cloud. We use the same approach as for star-forming galaxies, this time selecting the UVJ quiescent galaxies in the CANDELS catalog, and compute the median $U - V$ and $V - J$ colors in bins of redshift and mass. We find no significant redshift trend, and choose to only model the observed distribution with a mass dependence and some random scatter. We consider that all quiescent galaxies cluster around $U - V = 1.85$ and $V - J = 1.25$, and that within the red cloud, more massive galaxies are redder (probably because they are older). This mass dependence is encoded into the reddening B defined as

$$B = 0.1 \times (m - 11), \quad (4.13)$$

to which we add a Gaussian scatter of 0.1 magnitudes, and which is clamped to the range $B \in [-0.1, 0.2]$. The final UVJ colors are obtained using

$$(V - J)_{\text{passive}} = 1.25 + B, \quad (4.14)$$

$$(U - V)_{\text{passive}} = 1.85 + 0.88 \times B, \quad (4.15)$$

combined with a Gaussian scatter of 0.1 magnitudes.

We use these relations to derive the colors of the disk and bulge components of each galaxy. To do so, we consider that all disks are “active” and obtain their colors from Eq. 4.12. Similarly, we consider that all bulges of bulge-dominated galaxies

$(B/T > 0.6)$ are “passive” and described by Eq. 4.15. However, bulges in other galaxies are randomly chosen to be “active” or “passive” with uniform probability to simulate both bulges and pseudo-bulges.

The last step is to associate a detailed SED to each UVJ color, i.e., build an optical SED library. To reach this goal, we bin the UVJ plane into small buckets of 0.1 mag, and compute the average rest-frame SED of all the observed galaxies that fall inside each bucket, regardless of their redshift and stellar mass. The rest-frame SED of each galaxy is taken to be the best-fit SED produced by M. Pannella with *FAST* when fitting for the stellar mass (Sect. 4.2.2). We discard the buckets containing less than 10 galaxies, and end up with a library of 345 SEDs, all normalized per unit stellar mass, and each corresponding to a given position in the UVJ diagram. Because it is built out of observed galaxies, this library does not cover the whole UVJ plane. Therefore, if a simulated galaxy has colors that fall outside of the covered region (which will be rare by construction, but can still happen), it is attributed the SED of the closest non-empty bucket.

Since these SEDs are given per unit stellar mass, the final optical SED of each component is obtained by multiplying its stellar mass to the chosen SED from the UVJ diagram.

4.3.5 Sky Position

The simplest approach to generate the position on the sky of each galaxy is to draw these positions uniformly on the sphere, within the region of the sky that is covered by the simulated survey. The stellar mass functions we used in Sect. 4.3.1 ensure that we will get a correct sky density of object over the whole simulated area.

However, within the Λ CDM cosmology, we expect galaxies to form large-scale structures by following the merging history of their dark matter halos. In other words, galaxies tend to *cluster* on the sky, and we need to simulate this effect to generate realistic sky positions. In S15 (see also Béthermin et al. 2010), we showed that clustering can have a significant impact on the statistical properties of confused, long-wavelength images from *Spitzer* and *Herschel*: it will tend to increase the contrast compared to a uniform position distribution, i.e., creating overdense and underdense regions within the survey area. On the other hand, we expect this to be no more than a cosmetic change for the high-resolution *HST* images, which do not suffer from confusion.

The procedure we use here is to aim at reproducing the observed angular two-point correlation function, i.e., the excess probability of finding a galaxy at a given distance from another, as compared to a uniform position distribution. The first step is therefore to measure this two-point correlation function in the real GOODS-South field. To do so, we bin the whole catalog in redshift slices, and only two mass bins because the statistics is limited ($M_* = 10^9$ to $3 \times 10^{10} M_\odot$, and $M_* = 3 \times 10^{10}$ to $10^{12} M_\odot$), and we do not attempt to further refine the sample by separating different galaxy types. We then use the Landy and Szalay (1993) estimator to compute the

two-point correlation function of each sample, and observe a significant clustering signal between $1''$ and $100''$ which is well described by a single power law of index -0.5 . As in S15, we find no significant trend with redshift between $z = 0.3$ and $z = 4$, consistent with the results of Béthermin et al. (2015), but we do find that massive galaxies are on average twice more clustered.

Before proceeding to generating the positions, it is important to note that the measured two-point correlation function is affected by the uncertainties on the photometric redshifts (photo-zs). Indeed, within each adopted redshift bin, there is a chance that we miss some galaxies that scattered *out* of the bin, and another chance that we are contaminated by some galaxies scattering *into* the bin. The net result is that we observe a clustering amplitude which is lower than the intrinsic one. This effect can be simulated (and we do so in the following) once the uncertainty on the redshift is known. To measure this uncertainty, we cross-matched our GOODS-South catalog to that provided by 3DHST (DR1, Skelton et al. 2014). While the two catalogs are based on the same raw observations, the data reduction and photometry are completely independent. On the other hand, the photo-zs are estimated with the same code, so we will likely underestimate the true uncertainty. We chose not to use spectroscopic redshifts (spec-zs) for this experiment for two reasons: first, the photo-zs have been “trained” to most closely match the available spec-zs, so the agreement may be good only for this particular set of galaxies while not being representative of the true uncertainty; and second, spec-zs are only available for relatively bright and therefore biased samples. We thus measure the distribution of redshift differences between the two catalogs, and take into account that what we observe is the combination of uncertainties coming from *both* catalogs (i.e., assuming they are independent, $\sqrt{2}$ higher than that of a single catalog). We find that the redshift uncertainty in $\Delta z/(1+z)$ is well described by the combination of two zero-mean Gaussians: a first distribution of width 2 % that describes 80 % of galaxies, and a second distribution of width 7 % that describes the remaining 20 %.

To produce sky positions that resemble these observations, I interacted with H. Fergusson and C. White who advised me to use the Soneira and Peebles (1978) algorithm, which is one of the few known algorithms that are able to produce a two-point correlation function with a power-law shape. The algorithm is designed so that the slope of the power law can be chosen easily: we use $L = 4$ and choose η and λ to match both the requested number of simulated positions N_{sim} and the power law index $\gamma = 0.5$, i.e.,

$$\eta = N_{\text{sim}}^{1/L}, \quad (4.16)$$

$$\lambda = \eta^{1/(2-\gamma)}. \quad (4.17)$$

To prevent instabilities in the algorithm when N_{sim} is too small, we generate twice more positions than needed (with a minimum of 1 000), and randomly pick among the generated positions the ones we need, which preserves the two-point correlation function.

Using this method, we can produce a catalog of clustered positions with the right power-law slope. However, we still have to tune the *amplitude* of this clustering. We chose here a simple approach where we use the Soneira and Peebles (1978) algorithm only for a given fraction f of the simulated galaxies, and use uniformly distributed positions for the remaining fraction. We choose this fraction by first generating a set of positions with $f = 100\%$, apply the above procedure to measure the correlation function, and compare it to the observed one. The difference of amplitude then tells us by how much we need to reduce the simulated clustering. We stress that it is important here to take into account the redshift uncertainties that affect the observed relation. To do so, we measured the two-point correlation function in the simulation using “wrong” redshifts, which were taken from the “true” redshifts of the simulation and then perturbed within the uncertainty described above. After taking this into account, we find that $f = 40\%$ for $M_* < 3 \times 10^{10} M_\odot$, and $f = 70\%$ for more massive galaxies.

To double check, we also compute the angular correlation function of the whole catalog above $M_* > 10^{10} M_\odot$, mixing all redshifts all together. Doing so, we get rid of the issue of the redshift uncertainty, and find also a very good agreement with the observations.

4.4 Dust Properties

In Sect. 4.3.2, we have generated SFRs for all the galaxies in the simulation, and we have estimated what fraction of the associated light is supposed to come out in the FIR. Knowing this, all we need to predict FIR fluxes is a suitable FIR SED library with various adjustable parameters that can be used to reproduce accurately the observed counts. A number of such SED libraries have already been published, among which are the Chary and Elbaz (2001) (CE01) library, calibrated in L_{IR} from local galaxies, the Dale and Helou (2002) library, calibrated in FIR colors, or the Magdis et al. (2012) library, calibrated in $\langle U \rangle$ (or, equivalently, in dust temperature T_{dust}).

Here we are seeking for an additional level of control over the SED: we aim to be able to choose different effective dust temperatures T_{dust} and to change the relative contribution of the polycyclic aromatic hydrocarbon molecules (PAHs) f_{PAH} . There are two reasons for this choice: first, these are the two parameters that are the easiest to measure without FIR spectroscopy (which is only available for very few selected objects), and are those that affect the most the shape of the SED; and second, PAHs emit the bulk of their light around the rest-frame 8 μm , which is a domain that will be routinely accessed by the *James Webb Space Telescope* in the near future, and there will be a need for a properly calibrated library to exploit these data together with ancillary *Herschel* or *Spitzer* observations.

Therefore, we introduced in Chap. 3 a new SED library in which both T_{dust} and f_{PAH} are free parameters. We calibrate the redshift evolution of both parameters using the MIR to FIR stacks of S15, to which we add stacks of the *Spitzer* IRS 16 μm imaging

(Teplitz et al. 2011) to better constrain the PAH features (available in GOODS–*North* and *South* only). This calibration was refined in Sect. 3.2.2 using individual *Herschel* detections to constrain the scatter on these parameters, and also to calibrate how they are modified for those galaxies that are offset from the Main Sequence. We use these prescriptions to generate the FIR SED of each galaxy in the simulation, which then allows us to predict their MIR to submm fluxes.

The parametrization for T_{dust} is given by:

$$T_{\text{dust}}^{\text{MS}}[\text{K}] = \begin{cases} 20.2 \times (1+z)^{0.44} & \text{for } z \leq 2 \\ 26.3 \times (1+z)^{0.2} & \text{for } z > 2 \end{cases}$$

and

$$T_{\text{dust}}[\text{K}] = T_{\text{dust}}^{\text{MS}} + 6.6 \times \log_{10}(R_{\text{SB}}),$$

with a Gaussian scatter of 4.1K, and where $R_{\text{SB}} = \text{SFR}/\text{SFR}_{\text{MS}}$ (see Sect. 4.3.2).

The parametrization for f_{PAH} is:

$$f_{\text{PAH}}^{\text{MS}} = 0.04 + 0.035 \times (1 - 0.85 \times \text{clamp}(z, 1, 2)),$$

and

$$f_{\text{PAH}} = f_{\text{PAH}}^{\text{MS}} \times R_{\text{SB}}^{-0.47},$$

with a log-normal scatter of 0.2 dex.

4.5 Generating a Light Cone

With all these recipes, we can now generate a complete catalog of galaxies, each with its own UV-to-submm SED. In this section, we summarize the procedure that is implemented in EGG to produce a final flux catalog.

Given the area of the mock survey, the first step is to choose the number of galaxies that will be generated. Since we use the stellar mass function as a starting point, this amounts to choosing the lowest stellar mass that we will generate. This threshold can be chosen to be constant, e.g., down to $M_* = 10^8 \text{ M}_\odot$, but this is in fact quite inefficient: in the real GOODS–*South* field, which is a flux limited survey, we do detect galaxies that are less massive than 10^8 M_\odot at low redshifts, while the smallest measured stellar mass at $z > 2$ is closer to 10^9 M_\odot . Therefore, this approach can result either in a catalog that is incomplete (if the mass threshold is too high and we miss detectable galaxies at low redshift), or bloated (if the threshold is too low and we generate galaxies that will never be observed).

A more efficient approach is to use a redshift dependent threshold, so that galaxies are generated down to very low stellar mass at low redshifts, and then increase this threshold to generate fewer and fewer galaxies at higher redshifts. To do so, we first choose a “selection band”, e.g., the *HST* F160W or the VISTA K_s band, and

a magnitude limit, e.g., $H < 29$, above which the catalog will be at least 90 % complete. We then build a redshift grid, and for each redshift in that grid we compute the distribution of mass-to-light ratios in the selection band for all the optical SEDs in the library. We pick the 10th percentile of this distribution, and use it to compute the minimum stellar mass that is associated to the chosen magnitude limit at this redshift.

Once the stellar mass and the redshift are generated from the mass functions, we use the method described in Sect. 4.3.5 to place these galaxies on the sky. Note that, at this point, it is also possible to feed EGG with an existing catalog of redshifts, stellar masses, star-forming/quiescent flags, and positions (e.g., coming from a real catalog).

We then apply all the above recipes to generate the SFR, the L_{IR} and other dust related parameters (T_{dust} and f_{PAH}), the UVJ colors, and the morphological parameters (B/T , R_{50} , b/a).

Then, the optical SED is chosen based on the stellar mass and the generated UVJ colors (Sect. 4.3.4), and the FIR SED is chosen from the L_{IR} , T_{dust} and f_{PAH} . The two SEDs are co-added to form a single, panchromatic SED that ranges from the FUV up to the submm (the radio domain is not yet implemented). The last step is then to integrate this SED multiplied by the response curve of each broadband filter for which we want to generate the flux.

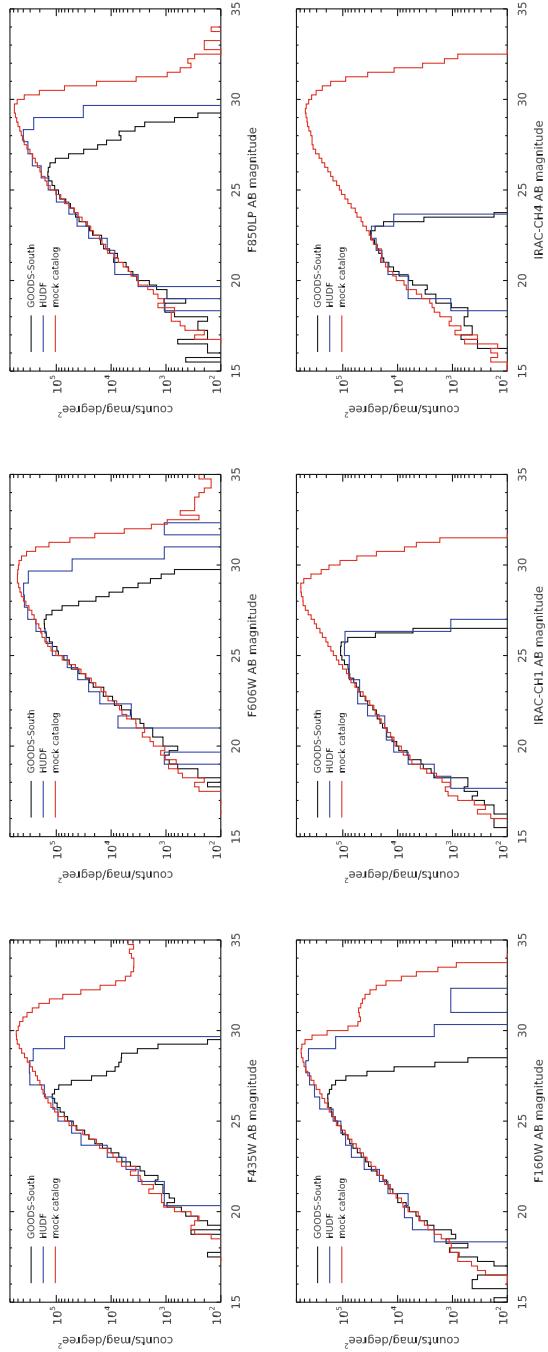
4.6 Results

4.6.1 Comparison to the Observed GOODS–South Field

Figure 4.5 is showing the total (bulge+disk) magnitude distributions in multiple bands as produced by the simulation. These are compared to the observed distributions in GOODS–South, splitting the field into two parts: the HUDF, which is deeper, and the rest of the field. The agreement is found to be very good in the NIR. Since these wavelengths are most closely correlated to the stellar mass of the galaxies, and since the mock catalog was built to reproduce exactly the stellar mass function in GOODS–South, this should not come as a surprise. Still, this shows that the procedure works well. Generating the optical (F435W and F606W) fluxes is more complex, because these bands actually trace the emerging UV light coming from star formation. Nevertheless, the agreement here is also very good.

Figure 4.6 (left) shows instead the FIR fluxes distributions. We jointly analyze in Fig. 4.7 the pixel histogram distribution of the simulated maps against the observed maps. This second test is important because of the blending, which sometimes pollutes the measured flux catalogs (two sources are combined into a single one), which tends to produce more bright fluxes than there actually is in the real Universe. By analyzing the map statistics directly, one gets rid of this issue of the counter part identification. This comparison also takes into account the clustering, which will tend to

Fig. 4.5 Observed magnitude distribution from the *B*-band up to *Spitzer* IRAC 8 μm . The simulated fluxes (red histogram) come from a mock field of $10' \times 10'$ that is complete down to $H < 29$. These are compared to the observed fluxes in the *Hubble* Ultra Deep Field (HUDF, blue) and the rest of the GOODS—South field (shallower, in black)



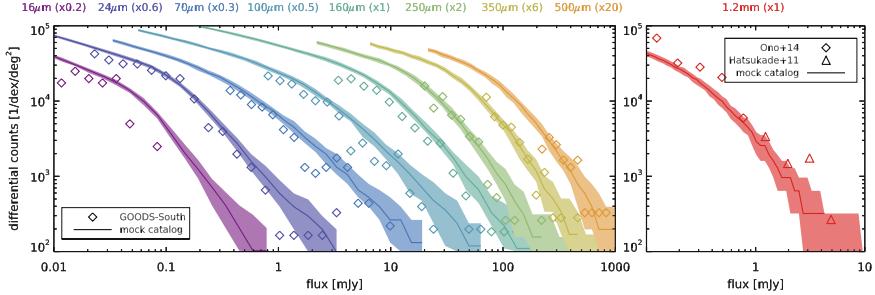


Fig. 4.6 *Left* Number counts in the MIR and FIR for various bands (different colors, see legend at the top). For the sake of clarity, the fluxes in each band were scaled by an arbitrary factor reported in the legend. The observed counts in GOODS—South are reported as open diamonds, and we compare these to the median histogram of 100 simulated catalogs using our prescriptions, shown with a solid line of the same color. In the background, the shaded area show the range covered between the 16th and 84th percentile of 100 simulated catalogs, to illustrate how much scatter one should expect simply due to cosmic variance. Note that the observed 16 μm fluxes below 40 μJy were taken from the deep region covering the HDFS, while counts above 40 μJy come from the rest of the field. *Right* Number counts at 1.2 mm. Observations from Ono et al. (2014) using ALMA 1.2 mm and Hatsukade et al. (2011) using AzTEC 1.1 mm are reported with open diamonds and triangles, respectively. Note that these observations were not done in GOODS—South. The counts predicted by our simulation are shown with a solid line, similarly to the plot on the *left*. Contrary to the other bands, the 1.2 mm is indeed a *prediction*, since our library and recipes were not explicitly calibrated to match these observations

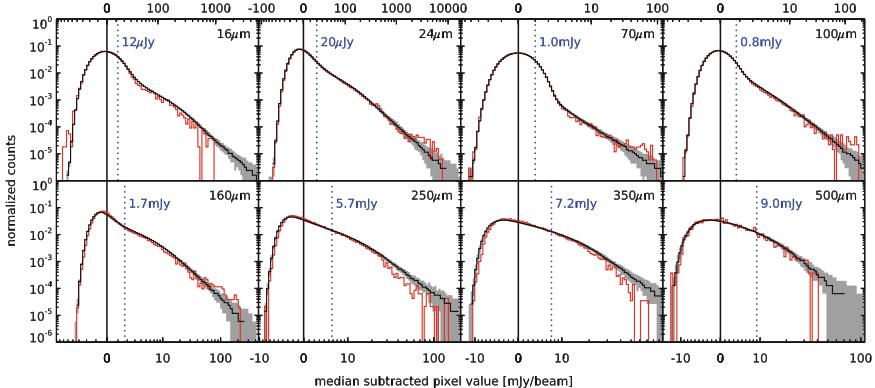


Fig. 4.7 Pixel value distributions of the 16–500 μm maps, in μJy/beam for 16 and 24 μm and mJy/beam otherwise. We show the observed distribution in GOODS—South in red, and compare this reference to 100 simulated catalogs generated with different random realizations. The median of these 100 realizations is shown with a solid black line, while the range covered by the 16th and 84th percentiles is shaded in gray in the background. Each map is median-subtracted, and the pixel values displayed here are scaled using the hyperbolic arcsine function (asinh). This is very similar to a logarithmic scale, except that it behaves linearly close to zero, allowing proper treatment of negative pixel values. We show the location of the median of the map with a vertical solid black line, and the 3σ point-source detection limit with a vertical blue dotted line (Color figure online)

increase the contrast of the map without actually changing the fluxes of individual galaxies. The downside is that the bright pixel counts are very sensitive to statistical fluctuations, and a single very bright (but usually rare) object can drastically impact the measured distribution. The agreement with the observed maps and counts is very good, and validates the robustness of our recipes.

Furthermore, in Fig. 4.6 (right), we show our *predicted* number counts at 1.2 mm, which is a wavelength domain for which our FIR SEDs and recipes are not calibrated. The agreement with published number counts from recent ALMA and single dish AzTEC observations is also good, reinforcing the validity of our approach. Some observed data points are found above our prediction, however, suggesting that we may slightly underestimate the sub-millimeter fluxes. This could be caused by the choice of T_{dust} , which may be too warm in the simulation (possibly if the starburstiness calibration is excessive) and will be investigated in the future.

Finally, we show in Fig. 4.8 an excerpt from the simulation to give a sense of the fidelity of the generated data.

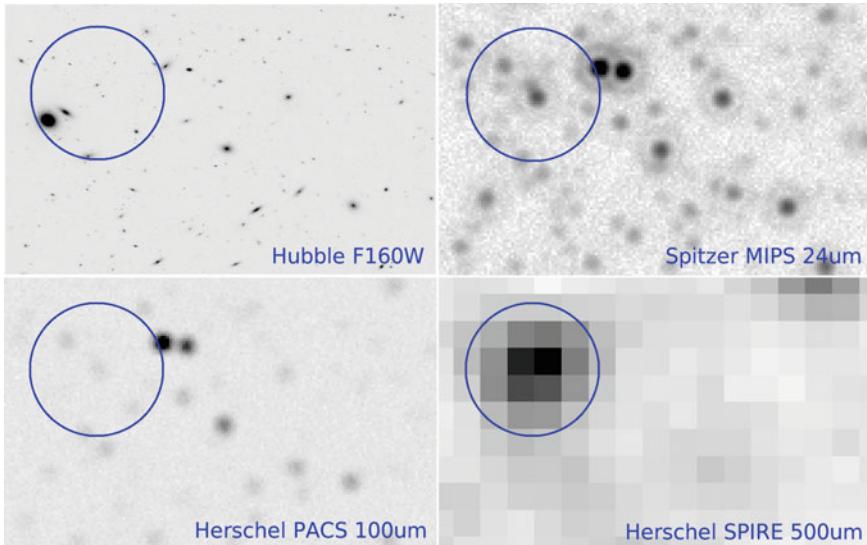


Fig. 4.8 Example of a simulated region of the sky seen at different wavelengths by different instruments, in the *HST H* band (top left), *Spitzer* MIPS (top right), *Herschel* PACS (bottom left) and SPIRE (bottom right). The *blue circle* indicates the region with bright SPIRE emission, to identify the counterparts on the other images

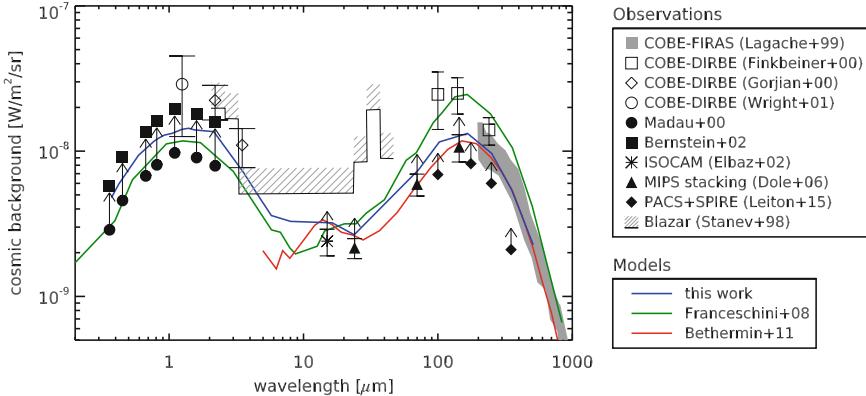


Fig. 4.9 The panchromatic cosmic background predicted by the simulation (blue solid line). We compare this against extragalactic background light measurements from the literature: the COBE-FIRAS observations from (Lagache et al. 1999, gray shaded region), COBE-DIRBE from (Finkbeiner et al. 2000, open squares), the COBE-DIRBE measurements of (Gorjian et al. 2000, open diamonds) and (Wright 2001, open circle). We display with a striped histogram the upper limits provided by Stanev and Franceschini (1998) by modeling the photon–photon interaction in the line of sight of the blazar MKN501. Lastly, we also show the summed light of resolved galaxies in the UV-to-optical from (Madau and Pozzetti 2000, filled circles) and the revision of these measurements by (Bernstein et al. 2002, filled squares), the ISOCAM measurement at 16 μm from (Elbaz et al. 2002, asterisk), the stacked *Spitzer* MIPS measurements at 24, 70 and 160 μm of (Dole et al. 2006, filled triangles) and the *Herschel* measurements at 100, 160, 250, 350 and 500 μm of individual detections from (Leiton et al. 2015, filled diamonds). Note that the values at 160 μm of Dole et al. (2006) and Leiton et al. (2015) have been slightly offset for clarity. All the measurements in this last series are to be considered at lower limits, since only the light of individually detected galaxies is taken into account. Two other galaxy models are shown for comparison: the panchromatic model of (Franceschini et al. 2008, green solid line) and the infrared model of (Béthermin et al. 2011, red solid line) (Color figure online)

4.6.2 Comparison to the Measured Cosmic Backgrounds

As a second test of the realism of these simulations, we show in Fig. 4.9 the cosmic background light we predict at various wavelengths from the UV to the sub-millimeter. To do so, a new simulation was created, covering $1.4^\circ \times 1.4^\circ$ down to $H < 27$ and containing 1.6 millions of galaxies.⁴ For each broadband, the simulated flux of all individual galaxies in the simulation were summed (with $0.01 < z < 9$) and simply divided by the area covered by the simulated catalog.

The comparison with observations from COBE, ISO, and Integrated (IGL, the summed contribution of resolved galaxies) from UV-to-NIR surveys shows that our simulation matches the observed cosmic background at all wavelengths. Our

⁴Generating this catalog took no more than 15 minutes on a modern laptop, the most time consuming part being the integration of the SEDs to produce the broadband fluxes. There is room for further optimization though, by precomputing the integrated fluxes of each template SED rather than doing it for each galaxy.

estimation in the optical is also close to that of Franceschini et al. (2008), within 50 %, although we predict a different far-infrared background by a factor of two. However, in this wavelength regime, we are in better agreement with the model of Béthermin et al. (2011) and the COBE-FIRAS observations of Lagache et al. (1999).

4.7 Appendix: Efficiency of Monochromatic FIR Measurements

The FIR SED library we introduced in Chap. 3 has more degrees of freedom than the other libraries released in the literature, and is therefore capable of reproducing a larger variety of real SEDs. However, the critical point is actually when too few measurements are available (typically, only one band), and we need to fix either f_{PAH} or T_{dust} (or both) to their redshift-average value (see Sect. 3.4). In this case, the accuracy of the derived quantities (i.e., L_{IR} and M_{dust}) depends largely on the proper choice of both f_{PAH} and T_{dust} , but assessing this accuracy with real galaxies is difficult without additional data. See, for example, the “mid-infrared excess” problem that was found in Daddi et al. 2007 when extrapolating L_{IR} from the MIPS 24 μm flux at $z = 2$. It was shown several years later in Elbaz et al. 2011 to be caused essentially by an inadequate choice of f_{PAH} for $z = 2$ galaxies (i.e., the f_{PAH} of local ULIRGs chosen in the CE01 library is not appropriate for $z = 2$ ULIRGs).

Assuming that the right average value is known, another issue that is seldom discussed or quantified is the uncertainty on the fit parameters resulting from the arbitrary choice of the SED (i.e., here, the choice of f_{PAH} and T_{dust}). For example, if we have a sample of $z = 2$ galaxies, each with a single measured flux in the PACS 160 μm , and if we assume a unique $T_{\text{dust}} = 37\text{K}$ to derive their L_{IR} , what is the uncertainty on this L_{IR} ? In this section, we take advantage of the simulated catalogs introduced earlier in this Chapter to quantify this uncertainty.

Indeed, because our simulation is built to reproduce most of the observed sources of scatter in the parameters that affect the shape of the FIR SED, we can use it to estimate a lower limit on the accuracy of a monochromatic L_{IR} measurement. In Fig. 4.10 (top left) we show the optimal uncertainty on the derived L_{IR} and M_{dust} if only a single photometric point is available, assuming: no error on the measured flux; perfect knowledge of the best average conversion factor between the flux and L_{IR} (i.e., perfect knowledge of the average f_{PAH} and T_{dust} at a given redshift); perfect subtraction of the stellar component (for 16 and 24 μm at high redshifts); and no contamination from AGN torus emission (affecting mostly rest-frame wavelengths below 30 μm , e.g., Mullaney et al. 2011). These relations were obtained with a simulated catalog of 10 deg² and selecting galaxies more massive than $M_* = 10^{10} \text{ M}_\odot$ at all redshifts.

There are several interesting features that come out of this figure. First, one can see that, when measuring fluxes on the dust continuum, the L_{IR} is best measured when the photometric point is close to the peak of the FIR SED, i.e., around 80 μm rest-frame: the optimal band goes from 100 μm at $z < 0.1$, 160 μm at $0.1 < z < 1.5$,

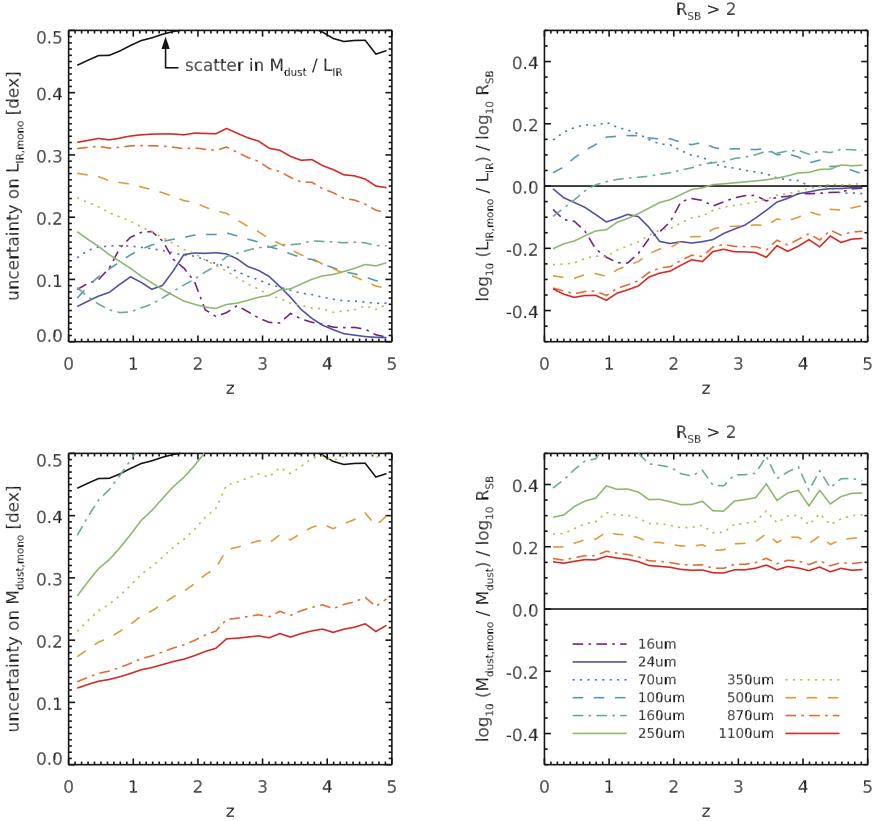


Fig. 4.10 Top left Predicted evolution of the uncertainty in $L_{\text{IR}}^{\text{mono}}$, i.e., the L_{IR} inferred from a single broadband photometric measurement, each band corresponding to a different color and line style (see legend). This uncertainty is derived by measuring the standard deviation of the difference between the *true* L_{IR} that was put in the simulation and the observed rest-frame luminosity. This is an *optimal* uncertainty, assuming (1) no error on the measured flux, (2) knowledge of the best average L_{IR}/L_{ν} conversion factor, (3) perfect subtraction of the stellar component (which only matters for 16 and 24 μm at high-redshift), and (4) no contamination from AGNs. We also show with a black solid line the scatter between L_{IR} and M_{dust} . Top right Predicted systematic error on the L_{IR} of starburst galaxies (selected here with $R_{\text{SB}} > 2$) normalized to the galaxies' offset from the Main Sequence. In other words, a value of x on this plot means that the L_{IR} will be wrong on average by a factor R_{SB}^x . Bottom left & right Same as top, but for $M_{\text{dust}}^{\text{mono}}$

250 μm at $1.5 < z < 3$, and 350 μm at $z > 3.5$. This is caused by variations of T_{dust} , which at fixed L_{IR} leave the 80 μm flux almost unchanged. Leftward of the peak, the uncertainty rises significantly when probing the rest-frame ~ 8 μm, i.e., for 16 μm at $0.5 < z < 2$ and 24 μm at $1.5 < z < 3$. This, in turn, is caused by variations of f_{PAH} . However, outside of these ranges, the shortest wavelengths are almost always the best tracers of L_{IR} . While this accuracy obviously depends on the contribution of the very small grains to the continuum, this would be a priori very useful for studying high-

redshift galaxies. In practice however, the typical fluxes are very low and out of reach of even the deepest *Spitzer* surveys. A second drawback is that we also expect AGNs and the stellar continuum to contribute significantly at these wavelengths, which is a problem that we do not address here. Finally, rightward of the peak, the uncertainty rises continuously as the wavelength increases, since rest-frame wavelengths beyond 250 μm are rather tracing the dust mass (see below).

Using a single photometric point, and therefore fixing f_{PAH} and T_{dust} to their redshift-average, one expects to be systematically biased against those galaxies which have unusual f_{PAH} and / or T_{dust} compared to other galaxies at the same redshift. As shown in Sect. 3.2.2, this is in particular the case for starburst galaxies. For this reason, we also display on Fig. 4.10 (top right) the predicted value of this systematic bias for galaxies with $R_{\text{SB}} > 2$ (i.e., at least one sigma away from the Main Sequence). The trend is for measurements leftward of the peak to barely overestimate the L_{IR} by no more than $R_{\text{SB}}^{0.2}$, while measurements rightward of the peak can reach systematic underestimation by a factor of $R_{\text{SB}}^{0.4}$. Similarly, bands which are dominated by emission of PAH molecules can underestimate the L_{IR} by about $R_{\text{SB}}^{0.2}$ (because PAH emission is depressed in starburst galaxies, Daddi et al. 2007; Elbaz et al. 2011). While these systematic errors will tend to bring starburst galaxies closer to the Main Sequence than what they are in reality, they are not sufficient to “erase” completely the starburst nature of these galaxies: for this, one would need to reach a bias close to R_{SB}^{-1} . However, this will affect any attempt at measuring the dispersion of the Main Sequence with a single MIR or FIR band.

We also present similar figures for M_{dust} in Fig. 4.10 (bottom panel) for submillimeter bands. The most striking fact to take out of this plot is that the uncertainty rises steadily width redshift, for all bands. This is simply caused by the redshift itself, which moves the observed wavelengths closer to the peak of the dust emission. As we showed above, this wavelength domain is measuring L_{IR} very accurately, and therefore is not suitable for measuring M_{dust} . This is only partly compensated for by the increase of the dust temperature with redshift (Eq. 3.7), which shifts the peak toward shorter wavelengths, but this shift is not large enough to counter-balance the effect of redshift. On the other hand, the systematic bias for starburst galaxies is relatively constant with redshift, and most importantly never reaches zero. For these galaxies, we predict a systematic overestimation of the dust masses by a factor of $R_{\text{SB}}^{0.1}$ at best.

References

- S. Arnouts, E. Le Floc'h, J. Chevallard et al., A&A **558**, 67 (2013)
- I.K. Baldry, S.P. Driver, J. Loveday et al., MNRAS **421**, 621 (2012)
- E. Bernhard, M. Béthermin, M. Sargent et al., MNRAS **442**, 509 (2014)
- R.A. Bernstein, W.L. Freedman, B.F. Madore, ApJ **571**, 56 (2002)
- M. Béthermin, C. De Breuck, M. Sargent, E. Daddi, A&A **576**, L9 (2015)
- M. Béthermin, H. Dole, M. Cousin, N. Bavouzet, A&A **516**, 43 (2010)
- M. Béthermin, H. Dole, G. Lagache, D. Le Borgne, A. Penin, A&A **529**, A4 (2011)

- M. Béthermin, E. Le Floc'h, O. Ilbert et al., *A&A* **542**, 58 (2012)
G.B. Brammer, P.G. van Dokkum, P. Coppi, *ApJ* **686**, 1503 (2008)
G. Bruzual, S. Charlot, *MNRAS* **344**, 1000 (2003)
V. Buat, S. Noll, D. Burgarella et al., *A&A* **545**, A141 (2012)
D. Calzetti, L. Armus, R.C. Bohlin et al., *ApJ* **533**, 682 (2000)
R. Chary, D. Elbaz, *ApJ* **556**, 562 (2001)
E. Daddi, M. Dickinson, G. Morrison et al., *ApJ* **670**, 156 (2007)
D.A. Dale, G. Helou, *ApJ* **576**, 159 (2002)
H. Dole, G. Lagache, J. Puget et al., *A&A* **451**, 417 (2006)
A.A. Dutton, F.C. van den Bosch, S.M. Faber et al., *MNRAS* **410**, 1660 (2011)
D. Elbaz, C.J. Cesarsky, P. Chanial et al., *A&A* **384**, 848 (2002)
D. Elbaz, E. Daddi, D. Le Borgne et al., *A&A* **468**, 33 (2007)
D. Elbaz, M. Dickinson, H.S. Hwang et al., *A&A* **533**, 119 (2011)
H.C. Ferguson, M. Dickinson, M. Giavalisco et al., *ApJ* **600**, L107 (2004)
D.P. Finkbeiner, M. Davis, D.J. Schlegel, *ApJ* **544**, 81 (2000)
A. Franceschini, G. Rodighiero, M. Vaccari, *A&A* **487**, 837 (2008)
A. Galametz, A. Grazian, A. Fontana et al., *ApJS* **206**, 10 (2013)
V. Gorjian, E.L. Wright, R.R. Chary, *ApJ* **536**, 550 (2000)
A. Grazian, A. Fontana, P. Santini et al., *A&A* **575**, A96 (2015)
N.A. Grogin, D.D. Kocevski, S.M. Faber et al., *ApJS* **197**, 35 (2011)
K. Guo, X.Z. Zheng, H. Fu, *ApJ* **778**, 23 (2013)
B. Hatsukade, K. Kohno, I. Arétxaga et al., *MNRAS* **411**, 102 (2011)
S. Heinis, V. Buat, M. Béthermin et al., *MNRAS* **437**, 1268 (2014)
A.M. Koekemoer, S.M. Faber, H.C. Ferguson et al., *ApJS* **197**, 36 (2011)
M. Kriek, P.G. van Dokkum, I. Labb  , et al., *ApJ* **700**, 221 (2009)
I. Labb  , M. Franx, G. Rudnick et al., *ApJ* **665**, 944 (2007)
G. Lagache, A. Abergel, F. Boulanger, F.X. D  s  r  t, J. Puget, *A&A* **344**, 322 (1999)
V.G. Laidler, C. Papovich, N.A. Grogin et al., *PASP* **119**, 1325 (2007)
S.D. Landy, A.S. Szalay, *ApJ* **412**, 64 (1993)
P. Lang, S. Wuyts, R.S. Somerville et al., *ApJ* **788**, 11 (2014)
R. Leiton, D. Elbaz, K. Okumura et al., *A&A* **579**, A93 (2015)
P. Madau, L. Pozzetti, *MNRAS* **312**, L9 (2000)
G.E. Magdis, E. Daddi, M. B  thermin et al., *ApJ* **760**, 6 (2012)
E. Merlin, A. Fontana, H.C. Ferguson et al., *A&A* **582**, A15 (2015)
G.R. Meurer, T.M. Heckman, D. Calzetti, *ApJ* **521**, 64 (1999)
J.R. Mullaney, D.M. Alexander, A.D. Goulding, R.C. Hickox, *MNRAS* **414**, 1082 (2011)
A. Muzzin, D. Marchesini, M. Stefanon et al., *ApJ* **777**, 18 (2013)
A.B. Newman, R.S. Ellis, K. Bundy, T. Treu, *ApJ* **746**, 162 (2012)
K.G. Noeske, B.J. Weiner, S.M. Faber et al., *ApJ* **660**, L43 (2007)
Y. Ono, M. Ouchi, Y. Kurono, R. Momose, *ApJ* **795**, 5 (2014)
M. Pannella, C.L. Carilli, E. Daddi et al., *ApJ* **698**, L116 (2009)
M. Pannella, D. Elbaz, E. Daddi et al., *ApJ* **807**, 141 (2015)
C.Y. Peng, L.C. Ho, C.D. Impey, H. Rix, *AJ* **124**, 266 (2002)
E.E. Salpeter, *ApJ* **121**, 161 (1955)
M.T. Sargent, M. B  thermin, E. Daddi, D. Elbaz, *ApJ* **747**, L31 (2012)
Schreiber, C., Elbaz, D., Pannella, M., et al. 2016, ArXiv e-prints. [arXiv:1606.05354](https://arxiv.org/abs/1606.05354)
C. Schreiber, M. Pannella, D. Elbaz et al., *A&A* **575**, A74 (2015)
S. Shen, H.J. Mo, S.D.M. White et al., *MNRAS* **343**, 978 (2003)
R.E. Skelton, K.E. Whitaker, I.G. Momcheva et al., *ApJS* **214**, 24 (2014)
R.M. Soneira, P.J.E. Peebles, *AJ* **83**, 845 (1978)
T. Stanev, A. Franceschini, *ApJ* **494**, L159 (1998)
C.M.S. Straatman, I. Labb  , L.R. Spitler et al., *ApJ* **783**, L14 (2014)
H.I. Teplitz, R. Chary, D. Elbaz et al., *AJ* **141**, 1 (2011)

- A. van der Wel, E.F. Bell, B. Häussler et al., *ApJS* **203**, 24 (2012)
A. van der Wel, Y. Chang, E.F. Bell et al., *ApJ* **792**, L6 (2014)
R.J. Williams, R.F. Quadri, M. Franx, P. van Dokkum, I. Labb , *ApJ* **691**, 1879 (2009)
E.L. Wright, *ApJ* **553**, 538 (2001)

Chapter 5

The Downfall of Massive Star-Forming Galaxies During the Last 10 Gyr

In this chapter I discuss in more detail one of the observation that was reported in Chap. 2 (Schreiber et al. 2015). By measuring the correlation between the SFR and the stellar mass (M_*) (a.k.a. the Main Sequence) of a mass-complete sample of galaxies, I found that the slope of the correlation was evolving with both redshift and stellar mass. At high redshift, this relation is well described by a single power law of slope unity, i.e., a linear correlation. The same is true at low redshift, but only for galaxies of relatively low stellar mass (typically less than a few $10^{10} M_\odot$). Above this stellar mass threshold, the SFR– M_* correlation flattens: as the mass increases, the corresponding enhancement of SFR becomes less and less pronounced. Because this study was based exclusively on star-forming galaxies, i.e., I have removed from the sample those galaxies that show little to no sign of recent star formation activity, this change of slope suggests that this is in fact the whole population of massive star-forming galaxies that is forming stars at a decreased rate. This observation has been reported in several independent studies (Magnelli et al. 2013; Whitaker et al. 2014; Lee et al. 2015), using different star formation rate indicators, different data sets, or different fields, and can therefore be considered as robust.

One way to interpret this fact is to consider that we are witnessing a *slow downfall* of star formation in this population. In other words, that there is a physical process that either decreases the mass of gas available to form stars (e.g., outflows), or lowers the efficiency with which this gas is transformed into stars. Another interpretation was put forward in Abramson et al. (2014), in which the authors argue that it is not really the SFR that is going down, but rather that the stellar mass is abnormally large owing to the presence of a quiescent *bulge* at the center of most massive galaxies. This has been recently revisited by Guo et al. (2015), who reached different conclusions.

In this chapter, I investigate these hypotheses using the data set introduced in my first paper. In collaboration with Maurilio Pannella, I perform a bulge-to-disk decomposition of the light profile of thousands of galaxies at $z = 1$, using high-resolution NIR *HST* imaging from the CANDELS program. To validate the robustness of our results, I compare the efficiency of two well known codes for morphological analysis:

I was in charge of using GALFIT, while M. Pannella used GIM2D. Using a large set of simulated galaxies that I created, we test our respective tool by trying to recover the intrinsic profile of the simulated galaxies (in particular, M. Pannella did not know how the simulation was built, nor the assumptions I used to build them). I find that the results of GIM2D tend to be more reliable, although the code is substantially slower. I therefore use the result of this tool to analyze the contribution of bulges to the stellar mass, and test the hypothesis of Abramson et al. (2014). I find that the bulge masses of $z = 1$ galaxies are overall too small to fully explain the change of slope of the Main Sequence, and argue that another process is required.

I then use the *Herschel* stacked photometry obtained in Schreiber et al. (2015) to measure the dust masses in different stellar mass bins, and infer from these values the mass of hydrogen gas, assuming that a fixed fraction of the metals are locked into dust grain (as in, e.g., Magdis et al. 2012). I find that the shallow slope of the Main Sequence at high stellar mass is mostly the result of a lower star formation efficiency, and further confirm this trend using local galaxies from the *Herschel* Reference Survey (HRS), kindly provided by Laure Ciesla with their stellar masses, star formation rates and dust masses.

I end this chapter by quantifying the amount of star formation that is lost because of this reduced star formation efficiency, and conclude that this is a major effect at $z < 1.5$, reaching similar level as the growth of the quiescent population. This provides evidence for a slow downfall of the star formation activity in massive Main Sequence galaxies, acting in parallel with the rapid quenching process that builds the red sequence.

This work was published in Schreiber et al. (2016).

5.1 Introduction

The discovery of a relation between the star formation rate (SFR) and the stellar mass (M_*) of galaxies, also called the “Main Sequence” of star-forming galaxies (Noeske et al. 2007), at $z \simeq 0$ (Brinchmann et al. 2004; Elbaz et al. 2007), $z \simeq 1$ (Noeske et al. 2007; Elbaz et al. 2007), $z \simeq 2$ (Daddi et al. 2007; Pannella et al. 2009a; Rodighiero et al. 2011; Whitaker et al. 2012) $z = 3\text{--}4$ (Daddi et al. 2009; Magdis et al. 2010; Heinis et al. 2013; Schreiber et al. 2015; Pannella et al. 2015) and even up to $z = 7$ (e.g., Stark et al. 2009; Bouwens et al. 2012; Stark et al. 2013; González et al. 2014; Steinhardt et al. 2014; Salmon et al. 2015) suggested a radically new paradigm for star formation. The tightness of this correlation is indeed not consistent with the frequent random bursts induced by processes like major mergers of gas-rich galaxies, and favors more stable, long-lasting episodes of star formation (Noeske et al. 2007).

Most studies focusing on this Main Sequence have measured the slope (in logarithmic space) of this correlation, and many different values were reported. A thorough compilation was recently published in Speagle et al. (2014), summarizing most measurements obtained so far. In particular, we can distinguish three kinds of

measurements. First, measured slopes close to unity (e.g., Elbaz et al. 2007; Daddi et al. 2007; Pannella et al. 2009a; Peng et al. 2010). Second, slopes *shallower* than unity, typically 0.8, and as low as 0.6 (e.g., Noeske et al. 2007; Karim et al. 2011; Rodighiero et al. 2011; Bouwens et al. 2012; Steinhardt et al. 2014; Speagle et al. 2014; Pannella et al. 2015). And finally, more recently a third group of studies actually advocate a broken power-law shape, or continuously varying slopes, where low-mass galaxies are well fitted with a slope of unity, and high mass galaxies exhibit much shallower (if not flat) slopes (e.g., Whitaker et al. 2012, 2014; Magnelli et al. 2014; Ilbert et al. 2015; Schreiber et al. 2015; Lee et al. 2015; Gavazzi et al. 2015). This latter, more refined description could actually explain the diversity of slope measurements that were obtained so far. Indeed, depending on the stellar mass range covered by the sample, which is usually limited, as well as the chosen redshift window, fitting a single power law will yield different best-fit slopes.

A tempting interpretation of this broken power law is that low mass galaxies evolve with a unique star formation efficiency, as shown by their universal specific SFR ($\text{sSFR} \equiv \text{SFR}/M_*$) (see, e.g., the discussions in Ilbert et al. 2015; Lee et al. 2015). Higher mass galaxies on the other hand depart from this universal relation and show a reduced star formation activity, probably gradually declining toward a quiescent state. This picture is somehow in contradiction with the idea that massive galaxies must *quench* rapidly (e.g., Peng et al. 2010), a process that often involves violent episodes in the lifetime of the galaxy, e.g., strong active galactic nucleus (AGN) feedback (Silk and Rees 1998). Instead, such a slow decline toward the red cloud could be more consistent with less abrupt processes like “radio-mode” AGN feedback (Croton et al. 2006; Bower et al. 2006) or “halo quenching” (Gabor and Davé 2012), where the infalling gas is heated up and prevented from forming stars. One can also invoke the “morphological quenching” mechanism (Martig et al. 2009), where the drop of efficiency is caused by the presence of a massive and dense stellar bulge that increases differential rotation within the disk, and prevents gas from fragmenting.

Recently, Abramson et al. (2014) put forward another possible explanation for this “bending” of the Main Sequence. They claim that this change of slope is not due to a reduced star formation efficiency. Instead, because of the presence of a bulge, they argue that the total stellar mass has become a poor proxy for the mass of gas available. Therefore, their argument is that one should rather expect the star formation rate to correlate with the mass of the *disk* instead, since this is where the star-forming gas is located. To support their claim, they used bulge-to-disk decompositions of the observed light profiles of thousands of local galaxies in the Sloan Digital Sky Survey (SDSS), and estimated their disk masses. They found indeed that the slope of the Main Sequence was put back to unity at all masses (at least for $M_* > 10^{10} M_\odot$) if the disk mass was substituted to the total stellar mass (see, however, Guo et al. 2015 where a conflicting result is obtained using the same data set).

In Schreiber et al. (2015) (hereafter S15), we have reported that the high-mass slope of the Main Sequence is gradually increasing with increasing redshift, approaching unity at $z > 2$ (see also Whitaker et al. 2014; Lee et al. 2015). In particular, at $z = 1$ we observed a less pronounced (but still significant) bending than

what is reported at lower redshifts. Our goal in this paper is to test if the bending of the Main Sequence disappears when the disk mass is substituted to the total stellar mass at $z = 1$, similarly to what was found by Abramson et al. (2014) at $z = 0$.

Thanks to the very high angular resolution provided by the *Hubble* Advanced Camera for Surveys (ACS) imaging, it is possible to perform the morphological analysis of the stellar profile of distant galaxies out to $z = 1$, either through non-parametric approaches (e.g., Abraham et al. 1996; Conselice 2003; Ferguson et al. 2004; Lotz et al. 2004), profile fitting (e.g., Bell et al. 2004; Ravindranath et al. 2004; Barden et al. 2005; McIntosh et al. 2005; Pannella et al. 2006; Pannella et al. 2009a; Häussler et al. 2007), or decomposition of this profile into multiple components (e.g., Simard et al. 1999, 2002; Stockton et al. 2008). The advent of the WFC3 camera on board *Hubble* has recently allowed studying the rest-frame near-IR (NIR) and optical stellar profiles toward higher redshifts (e.g., van der Wel et al. 2012; Newman et al. 2012; Bruce et al. 2012, 2014; Lang et al. 2014). In particular, Bruce et al. (2012) have performed bulge-to-disk decomposition on the CANDELS H -band imaging in the UDS field, focusing of massive galaxies ($M_* > 10^{11} M_\odot$) from $z = 1$ to $z = 3$, and finding a clear trend of decreasing bulge-to-total ratio (B/T) with redshift. However, later on Lang et al. (2014) pushed the analysis down by one order of magnitude in stellar mass in all five CANDELS fields. By fitting stellar-mass maps estimated through resolved SED-fitting, they derived the relation between M_* and B/T for active and passive galaxies, and found very little evolution of this relation with redshift. Both these observations are contradictory, and would potentially lead to different conclusions when trying to link the bulge mass to the Main Sequence bending.

In this paper, we therefore revisit the bulge-to-disk decomposition, carefully computing disk masses of $z = 1$ galaxies in Sect. 5.3.3, and analyzing the change of slope between the SFR– M_* and SFR– M_{disk} relations in Sect. 5.5.1.

In parallel, we explore an alternative route where we directly quantify the mass of gas present in these galaxies (M_{gas}), to see if the bending is caused by a variation of gas fraction or a variation of the star formation efficiency (SFE $\equiv \text{SFR}/M_{\text{gas}}$). To this end, we follow the approach of Magdis et al. (2012) and Magnelli et al. (2012b) and employ the far-infrared (FIR) stacks of S15 to measure dust masses in Sect. 5.4. Assuming that a fixed fraction of the metals ($\sim 26\%$, as discussed in Sect. 5.4.2) condenses to form dust grains, and with the knowledge of the gas-phase metallicity, one can infer the gas mass from the dust mass (Franco and Cox 1986) and derive the SFE. This approach has been used extensively in the recent literature to measure gas masses in a wide variety of samples from $z = 0.3$ to $z = 4$ (e.g., Magdis et al. 2012; Santini et al. 2014; Scoville et al. 2014; Béthermin et al. 2015). We apply it in Sect. 5.5.2 to look for an evolution of both the gas fraction (f_{gas}) and the SFE along the Main Sequence at $z = 1$, and complement this analysis with local galaxies drawn from the *Herschel* Reference Survey (HRS, Boselli et al. 2010).

In the following, we assume a Λ CDM cosmology with $H_0 = 70 \text{ km s}^{-1}\text{Mpc}^{-1}$, $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$ and, unless otherwise specified, a Salpeter (1955) initial mass function (IMF) to derive both star formation rates and stellar masses. All magnitudes are quoted in the AB system, such that $M_{\text{AB}} = 23.9 - 2.5 \log_{10}(S_\nu [\mu\text{Jy}])$.

5.2 Sample Selection and Galaxy Properties

In this work we investigate the change of slope of the Main Sequence from two different angles. On the one hand, we measure the gas content inside Main Sequence galaxies to look for a decrease of either the gas fraction or the star formation efficiency. To do so, we use the stacked *Herschel* SEDs of S15 at $z = 1$ in the CANDELS fields (see Sect. 5.2.3), and complement the analysis with a $z = 0$ sample of Main Sequence galaxies from the HRS (see Sect. 5.2.4). On the other hand, we extract a subsample of massive galaxies from our $z = 1$ sample and perform the morphological decomposition of the *HST* light profile. Among these, we will also consider the galaxies with an individual IR detection in order to derive robust SFRs. The description of this subsample is given in Sect. 5.2.5.

5.2.1 Multi-wavelength Photometry

The $z = 1$ catalogs we use in this work are based on the CANDELS (Grogin et al. 2011; Koekemoer et al. 2011) *Hubble Space Telescope* (*HST*) WFC3 H -band images in the four CANDELS fields that are covered by deep *Herschel* PACS and SPIRE observations, namely GOODS–North (Barro et al. in prep.), GOODS–South (Guo et al. 2013), UDS (Galametz et al. 2013) and COSMOS (Nayyeri et al. in prep.). Each of these fields covers about 150 arcsec² and they are evenly distributed on the sky to mitigate cosmic variance.

The ancillary photometry varies from one field to another, being a combination of both space- and ground-based imaging from various facilities. The UV to near-IR wavelength coverage typically goes from the U band up the *Spitzer* IRAC 8 μm , including at least the *HST* bands F606W, F814W, and F160W and a deep K (or K_s) band, and all these images are among the deepest available views of the sky. These catalogs therefore cover most of the important galaxy spectral features across a wide range of redshifts, even for intrinsically faint objects.

We complement these catalogs with mid-IR photometry from *Spitzer* MIPS and far-IR photometry from *Herschel* PACS and SPIRE taken as part of the GOODS–*Herschel* (Elbaz et al. 2011) and CANDELS–*Herschel* programs (PI: M. E. Dickinson, Inami et al. in prep.).

The UV to NIR photometry for the HRS galaxies is compiled from various sources, and this dataset is fully described in Boselli et al. (2010). The *Herschel* PACS and SPIRE observations were taken as part of the *Herschel* Reference Survey and the fluxes were extracted by Ciesla et al. (2012) for SPIRE and Cortese et al. (2014) for PACS.

5.2.2 Redshifts, Stellar Masses and Star Formation Rates

Photometric redshifts and stellar masses for our $z = 1$ sample are computed by M. Pannella following Pannella et al. (2015). He uses EAZY (Brammer et al. 2008) to derive the photometric redshifts from the CANDELS catalogs, allowing slight adjustments of the photometric zero points by iteratively comparing our photo- z 's against the available spec- z 's. The stellar masses are then computed using FAST (Kriek et al. 2009) by fixing the redshift to the best-fit photo- z and fitting the observed photometry up to the IRAC 4.5 μm band using the Bruzual and Charlot (2003) stellar population synthesis model, assuming a Salpeter (1955) IMF and a Calzetti et al. (2000) extinction law.

Galaxies with an uncertain photometric redshift (redshift “odds” less than 0.8) or bad SED fitting (reduced χ^2 larger than 10) are excluded from our sample. We estimated in S15 that the stellar mass 90 % completeness at $z = 1$ was as low as $5 \times 10^8 \text{ M}_\odot$, i.e., almost one order of magnitude below the lowest stellar mass used in the present study ($2 \times 10^{10} \text{ M}_\odot$ for the morphological decomposition, 10^9 M_\odot for stacking).

I estimate star formation rates (SFRs) of individual galaxies are estimated only for the galaxies with a MIR or FIR detection. I fit the observed MIR to FIR photometry is fit with Chary and Elbaz (2001) templates, and the IR luminosity L_{IR} (from 8 to 1000 μm) is obtained from the best-fit SED. Since our study focuses exclusively on the $z \sim 1$ Universe (see next section), galaxies only detected by *Spitzer* MIPS 24 μm are also used in this analysis. For these objects, I use the original Chary and Elbaz (2001) L_{IR} calibration. I then use the Kennicutt (1998) and Daddi et al. (2004) relations to convert this L_{IR} and the observed L_{UV} (1500 Å, non-dust-corrected) into SFR_{IR} and SFR_{UV} , respectively. The total SFR of a galaxy is then computed as the sum $\text{SFR} = \text{SFR}_{\text{IR}} + \text{SFR}_{\text{UV}}$, although for all our galaxies with a FIR detection the contribution of SFR_{UV} is negligible.

Lastly, the rest-frame U , V and J magnitudes are computed by M. Pannella for each galaxy using EAZY, by integrating the best-fit galaxy template from the photo- z estimation. These colors are used to separate star-forming from quiescent galaxies using the UVJ classification scheme as introduced in Williams et al. (2009). This classification will be used in the following to study separately the behavior of both populations.

For HRS galaxies, stellar masses and star formation rates are derived by L. Ciesla using CIGALE (Noll et al. 2009, with the modifications of Burgarella et al. and Boquien et al. in prep.) which fits template SEDs to the available UV to FIR photometry simultaneously, in a consistent way. Since all galaxies of the HRS have *Herschel* coverage, the resulting SFRs are therefore based on both the observed far-UV and far-IR fluxes. These fits are described in more detail in Ciesla et al. (in prep.). She then computes the U -, V - and J -band rest-frame magnitudes from the best-fit template.

5.2.3 CANDELS Sample for the Gas Mass Measurements at $z = 1$

For the gas mass measurements at $z = 1$, we use the stacked *Herschel* photometry presented in S15. In this work, we showed that the bending of the Main Sequence is more pronounced at lower redshifts, and almost absent by $z > 2$. To study the origin of this bending, we therefore need to focus on low redshifts, where the bending is most significant. On the other hand, the area covered by the CANDELS fields is relatively small, and consequently we cannot afford to reach too low redshifts, say $z < 0.5$, without being affected by limited statistics and cosmic variance. Furthermore, our estimation of the gas mass is based on the dust mass (see Sect. 5.4.2), and at $z > 1.5$ *Herschel* does not probe the Rayleigh–Jeans tail of the dust SED ($\lambda_{\text{rest}} > 250 \mu\text{m}$), which would prevent accurate determination of the dust mass (Scoville et al. 2014).

For these reasons we choose to base our analysis on galaxies at $0.7 < z < 1.3$, and use the same sample as in S15, namely selecting all the galaxies in this redshift window that are classified as *UVJ* star-forming:

$$UVJ_{\text{SF}} = \begin{cases} U - V < 1.3, \text{ or} \\ V - J > 1.6, \text{ or} \\ U - V < 0.88 \times (V - J) + 0.49. \end{cases} \quad (5.1)$$

This selection is shown in Fig. 5.5. As discussed in S15, more than 85 % of the *Herschel* detections are classified as *UVJ* star-forming. The *UVJ* selection is therefore an efficient tool to pinpoint star-forming galaxies, even when MIR or FIR detections are lacking. However, it affects more strongly the galaxies at high stellar mass. In particular, between 10^{11} and $3 \times 10^{11} \text{ M}_\odot$, about half of our galaxies are classified as *UVJ* quiescent. Since the precise definition of Eq. 5.1 can affect our results, we discuss its impact a posteriori in Sect. 5.8.

5.2.4 HRS Sample for the Gas Mass Measurements in the Local Universe

For the $z = 0$ sample, we define the dividing line between “star-forming” and “quiescent” galaxies as follows:

$$UVJ_{\text{SF}}(\text{HRS}) = \begin{cases} U - V < 1.6, \text{ or} \\ V - J > 1.6, \text{ or} \\ U - V < 0.88 \times (V - J) + 0.79. \end{cases} \quad (5.2)$$

In practice, this is equivalent to making a cut in $\text{sSFR} > 6 \times 10^{-3} \text{ Gyr}^{-1}$, i.e., about one dex below the $z = 0$ Main Sequence. Different *UVJ* dividing lines have been adopted in the literature, reflecting a combination of both zero point offsets in the

photometry and physical evolution of the colors caused by the evolution of the sSFR. For example, Williams et al. (2009) used different *UVJ* classifications depending on the redshift, with a $0 < z < 0.5$ criterion that is different from Eq. 5.2 by only 0.1 magnitudes, and a $1 < z < 2$ criterion identical to our Eq. 5.1.

In the following, we use all the galaxies from the HRS survey that satisfy the *UVJ* criterion given above, regardless of their morphological type. In practice, the *UVJ* selection naturally filters out all the early-type galaxies (E-S0-S0/Sa), and about half of the H I -deficient galaxies (Boselli et al. 2010).

However, it is important to note that, although the HRS is a purely K -band selected sample, the volume it spans is relatively small and this field is thus subject to cosmic variance. Furthermore, because one of the science goals of the HRS is to study the influence of the environment on the star formation activity, the sample also contains the Virgo cluster, a strong overdensity that encloses 46 % of the galaxies in the whole HRS (and 39 % of *UVJ* star-forming galaxies). This is a very biased environment, and although clusters are more common in the Local Universe, the HRS is known to be particularly deficient in gas mass, likely because of the inclusion of Virgo (Boselli et al. 2010). To ease the comparison with our $z = 1$ sample described in the previous section, we therefore exclude from the HRS all the galaxies that belong to Virgo (149 galaxies out of 323). Combined with the *UVJ* selection, this excludes 80 % of the H I -deficient galaxies, and yields a final sample of 131 galaxies. We note however that our results would be essentially unchanged if we were to keep the Virgo galaxies in our sample.

5.2.5 CANDELS Sample for the Morphological Decompositions at $z = 1$

For the morphological analysis, we consider the same redshift window as for the gas mass measurement at $z = 1$, following the same motivations. In addition, limiting ourselves to $z = 1$ ensures that the *HST H* band probes the rest-frame i band, where mass-to-light ratios are weakly varying (e.g., de Jong 1996). However, to obtain reliable morphological decompositions, we further select galaxies more massive than $2 \times 10^{10} \text{ M}_\odot$, corresponding roughly to an H -band limited sample at these redshifts, with no galaxy fainter than $H = 22.5$ (see Sect. 5.3.2 where we justify this choice using simulated images). Unfortunately, this stellar mass cut will prevent us from performing the morphological decomposition in the regime where the Main Sequence is linear. However, it is known that disk-dominated galaxies dominate the low-mass galaxy population, both in the Local Universe (e.g., Bell et al. 2003) and at higher redshifts (e.g., Pannella et al. 2009a; Lang et al. 2014; Bluck et al. 2014). Therefore we will assume in the following that most galaxies below our mass threshold are disk-dominated, with $M_* \simeq M_{\text{disk}}$, and only consider changes in Main Sequence slope above this threshold. We also remove 6 IRAC power law AGNs (Donley et al. 2012).

To prevent systematic effects in the morphological analysis due to strong galaxy blending (either due to mergers or chance projections), we also need to remove from our sample the galaxies that have too close bright neighbors in the H -band image. Deblending can be done, to some extent, by fitting the profiles of multiple objects simultaneously, e.g., with GALFIT (Peng et al. 2002), but this is often adding more instability in the fit, and should be done with great caution. We will not attempt it here. Therefore, I flagged the galaxies that have at least one companion within $2''$ with a total flux that is no less than 10 % fainter. This flags out ~ 410 galaxies, and our final “ H -sample” consists of $\sim 2\,440$ galaxies. The impact of these selections on the stellar mass distribution of our sample is shown in Fig. 5.1.

Then, among these, we also consider the galaxies with a MIR or FIR detection ($>5\sigma$), i.e., with a robust SFR estimate coming from *Spitzer* or *Herschel* observations. To do so, I start from the same IR catalogs as those introduced in Chap. 2, but here I further revisit the catalogs to solve an issue that, although irrelevant to the results I presented earlier, can have important consequences for the present study. Briefly, I flag the *Spitzer* MIPS detections that are wrongly associated to UVJ passive H -band counterparts because of the adopted source extraction procedure. The details of this procedure are described in the next section. In total I flag no more than 5 % of the

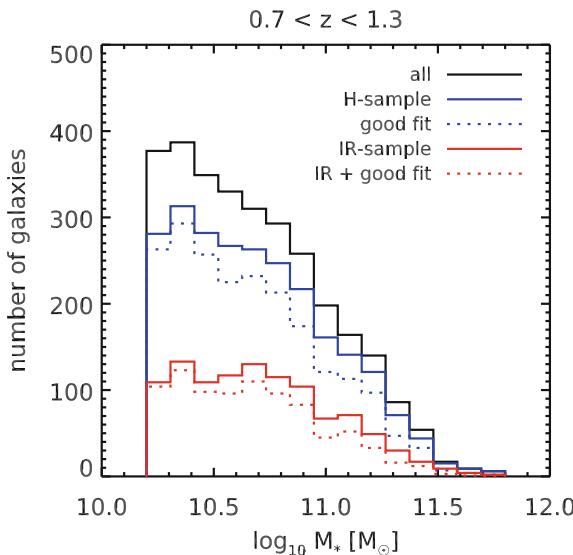


Fig. 5.1 Stellar mass distribution of the various samples at $z = 1$ that we consider for the morphological decomposition. The *black solid line* shows the distribution of our parent sample, as used in S15, and containing all the galaxies at $0.7 < z < 1.3$ with $M_* > 2 \times 10^{10} M_\odot$. The *blue solid line* is our “ H -sample”, after removing close pairs and IRAC power-law AGNs from the parent sample. The *red solid line* is our “IR-sample”, requiring a clean *Spitzer* MIPS or *Herschel* detection. The *dotted lines* indicate the number of galaxies that we manage to correctly decompose with GIM2D within each sample (color figure online)

MIPS detections as wrong or uncertain associations.¹ Two thirds of these are *UVJ* passive galaxies.

The final “IR-sample” contains $\sim 1\,010$ galaxies, and therefore about 44 % of the galaxies of the “*H*-sample” have a robust SFR estimation (see Fig. 5.1). This number rises to 63 % if we only consider *UVJ* star-forming galaxies. For consistency checks, we do perform the morphological detection on the whole *H*-sample, but only use the IR-sample to derive the slope of the Main Sequence, meaning that we will work with a sample that is both mass and SFR selected. This is not an issue for our purposes: even with an SFR selection, the change of slope of the Main Sequence can be seen as long as the SFR detection limit is low enough, which is the case here.

5.2.6 Cleaning the 24 μm Catalogs

I focus here on the association of a *Spitzer* MIPS 24 μm flux to the galaxies in the *H*-band catalog. The procedure that was used to build the 24 μm flux catalog (see Magnelli et al. 2009) is based on IRAC 3.6 μm position priors: sources are extracted on the 24 μm map (and then, sequentially on the *Herschel* images) at the position of bright 3.6 μm sources. If two priors are too close to be deblended on the MIPS image, only the brightest 3.6 μm source is kept in the prior list. Because the IRAC bands are good tracers of the stellar mass, and because the stellar mass correlates with the star formation rate, this approach is very effective for extracting reliably the vast majority of the MIR and FIR sources. But it will fail in a few rare cases that will be particularly important for our study (see also Mancini et al. 2015). Indeed, one expects the method to be biased as soon as some objects deviate from the SFR– M_* correlation. For example, it will happen that a massive, quiescent galaxy lies within a few arcseconds of a smaller mass (or slightly higher redshift) star-forming galaxy. The quiescent galaxy, being very massive, is most likely the brightest emitter in the IRAC 3.6 μm image, however it is not expected to shine much in the MIR because it is not forming any stars. The nearby star-forming galaxy on the other hand can be fainter in the IRAC image, but will contribute to most, if not all, of the MIR emission. In this situation, the typical outcome is that the star-forming galaxy is removed from the prior list, since it has the faintest IRAC flux, while the quiescent galaxy is given all the IR flux. The end result is that we do have in our catalogs a few massive quiescent galaxies with bright 24 μm emission that are obvious mismatches. I emphasize that the issue does not affect the 24 μm fluxes listed in the published catalogs, but rather the association of these fluxes to counterparts in the higher-resolution *HST* images.

I therefore eyeballed every galaxy of the *H*-sample that was attributed a counterpart in the MIPS image, looking for this kind of problematic cases. To identify quiescent galaxies, I rely on the *UVJ* classification introduced in the previous section. In total, I find 40 clearly wrong associations over the four CANDELS fields, based

¹If I had not previously removed close galaxy pairs from the parent *H*-sample, this number would rise to 8 %.

on a combination of the UVJ classification and the presence of a likely star-forming candidate nearby, or by significant off-centering of the MIPS emission. Because this approach is hard to replicate and translate to other surveys, I introduce here a systematic and objective procedure to identify this kind of issues that does not require eyeballing every galaxy. It also allows me to further refine the flagging and discard not only galaxies that are clearly wrong associations, but also those that are uncertain, so that we work with a sample that is as clean as possible.

For each UVJ star-forming galaxy in the H -sample, I derive their expected “Main Sequence” star formation rate from their redshift and stellar mass, i.e., the SFR they would have if they were exactly following the Main Sequence as defined in this chapter. From this SFR I subtract the observed, non-dust-corrected SFR_{UV} , and use the Kennicutt (1998) relation to convert the remaining obscured SFR into L_{IR} . I then use the best-fit IR SEDs of Chap. 3 to estimate their $24\mu m$ flux. For UVJ passive galaxies, I follow a similar procedure where the total SFR is instead taken from the stacking of UVJ passive galaxies, as described in the Appendix of this chapter. This SFR is typically a factor of ten below the Main Sequence at all stellar masses.²

Using this procedure I am able to obtain a rough prediction of the MIR output of all the galaxies in the H -band parent sample. Then, for each galaxy with a $24\mu m$ detection, I estimate the reliability of the MIR association. To do so, I take all the galaxies that (1) lie within $4''$ of the detection, (2) have a *predicted* $24\mu m$ flux that is at least a tenth of that predicted for the detection, and (3) have no measured $24\mu m$ (or below 3σ) in the catalog. I then sum all their fluxes, weighted by the MIPS PSF amplitude at their corresponding distance, and divide this sum by the predicted flux of the detection. The resulting value gives an estimation of the fraction of the measured flux that can be contaminated by neighboring sources that were excluded from the prior list.

As expected, the vast majority of the sources in the MIPS catalog are classified as robust identifications: 80 % of them have an estimated contamination of zero. In the following, we only use the individual SFRs of galaxies for which this contamination fraction is below 30 %. This criterion recovers 27 of the 40 wrong associations I identified by eye, the remaining 13 galaxies are either not properly deblended on the HST image, or their neighbors have wrong photometric redshifts and their contamination is underestimated. I therefore also exclude these 13 galaxies from our sample.

Note that this flagging does not apply to the sample we use to make the gas mass measurements. Indeed, the gas masses are measured by stacking H -band selected galaxies, and therefore do not rely on the $24\mu m$ catalogs.

²This may sound surprisingly high, but it should be noted that this stacked “SFR” of quiescent galaxies also includes, for a large fraction, some L_{IR} coming from the dust headed by old stars, and not actual star formation. Therefore this prescription allows us to take into account both residual star formation and dust headed by old stars at the same time. See also Fumagalli et al. (2014) where this was done in more details.

5.3 Measuring Disk Masses in Distant Galaxies

5.3.1 *The Bulge to Disk Decomposition*

The bulge-to-disk decomposition is performed by M. Pannella, following Pannella et al. (2009b). He uses the software GIM2D (Simard et al. 2002) on the HST H -band images ($0.06''/\text{pixel}$ resolution). To carry out a proper parametric modeling of the galaxy two-dimensional light distribution, it is of fundamental importance to obtain a careful estimate of the local background level. An extended disk or the low surface brightness wings of a high Sérsic index galaxy can easily fool the fitting code and hence retrieve the wrong galaxy model (e.g., Häussler et al. 2007; Pannella et al. 2009a; Barden et al. 2012). In order to avoid this issue, we run SExtractor (Bertin and Arnouts 1996) on the public CANDELS H -band images in “cold” mode. This allows us to better minimize the artificial source splitting and maximize the number of pixels assigned to each object. Our newly extracted H -band catalog is then cross-matched to the original CANDELS photometric catalog so that every entry is assigned a redshift and a stellar mass. Less than 10 % of the original sample is actually not retrieved by our cold source extraction. For the most part, these are blended objects for which a bulge-to-disk decomposition would be both impractical and uncertain, and we do not consider these in the following. For every galaxy, we then extract a cutout in both the original image and our SExtractor segmentation map, the size of which depends on the actual galaxy angular dimensions. This ensures that GIM2D is able to properly fit for the image background and recover accurate galaxy parametric modeling.

Using these image and segmentation cutouts, we fit a combination of two Sérsic profiles: an exponential disk ($n = 1$) and a de Vaucouleur profile ($n = 4$), both convolved with the “hybrid” WFC3 PSFs from van der Wel et al. (2012). An example of such decomposition is given in Fig. 5.2.

Although the fit generally settles to physically reasonable solutions, there are cases where the effective radius of either component converges to zero, meaning that the component is essentially unresolved. In this case, there is no way to disentangle an exponential disk from a de Vaucouleur profile, and this unresolved component could be either an AGN, a nuclear starburst, or just the badly-fit core-component of a bulge. Fortunately such cases are rare (5 % of our sample), so we decided to consider them as bad fits and exclude them from the following analysis.

When defining our sample, we took care to exclude close galaxy pairs that would cause blending issues (see previous section). However, while analyzing the results of the decomposition, we also found that there are a few galaxies which are not even properly deblended in the CANDELS catalogs to begin with, e.g., because the two galaxies are too close and SExtractor considered the pair as a single object. These galaxies cannot be fitted with our procedure, and typically show large χ^2 . To filter out these catastrophic failures, we therefore impose a maximum value of $\chi^2 < 2$. This also removes remaining catastrophic fit failures, and galaxies with too irregular morphologies. This cut excludes 10 % of the sample. Finally, we also

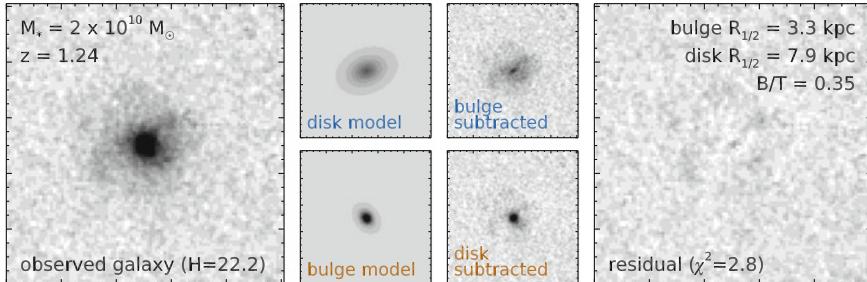


Fig. 5.2 Example bulge-to-disk decomposition of an $H = 22.2$ galaxy from the GOODS–South field, which is among the faintest galaxy in our sample. The *first column* shows the observed HST WFC3 image of the galaxy, and we also provide in the *top-left corner* its main physical properties. The *second column* shows the best-fit disk (*top*) and bulge (*bottom*) components as extracted by GIM2D. The *third column* shows the residual of the image after subtraction of the bulge (*top*) and the disk (*bottom*), to visualize the profile of the other component. Finally, the *fourth column* shows the residual image after both components are subtracted. The best-fit parameters are given in the *top-right corner*

exclude galaxies that are fit with extremely small component sizes, i.e., less than a fifth of a pixel, indicating that the code would have rather fitted a point source instead of an extended component. Because we cannot reliably attribute this flux either to the disk or the bulge, we choose not to use these fits in the present analysis (4 % of the sample).

To make sure that our results are not strongly biased by our decomposition approach, I also run in parallel the same decomposition using GALFIT (Peng et al. 2002). The same images and segmentations are used, the only difference is that I can allow for some small position offset between the bulge and the disk. The minimization procedure is also different between both codes, and therefore different results are usually obtained for the same data, providing an estimation of the uncertainty on the decomposition. Since GALFIT requires an initial guess of the fit parameters, I used the single-component morphological parameters measured by van der Wel et al. (2012) who fit a single Sérsic profile to the H -band image of each galaxy in the CANDELS catalogs of GOODS–South, UDS and COSMOS. I complement these measurements by running myself similar fits in GOODS–North.³ These parameters are used to set the initial size, axis ratio and position angle of both the disk and the bulge components, while the initial flux of each component is set to half the total flux of the galaxy (i.e., an initial $B/T = 0.5$). I then run GALFIT, leaving free every parameter including the position of each component, with a maximum offset between both components of 10 pixels (in practice, the results are essentially the same if I do not allow for such offsets).

We have checked that our conclusions are not affected if we only keep the galaxies for which the two codes agree (variation of B/T smaller than 0.15), or if we used only

³As a consistency check, I also refit the galaxies in the other fields with single Sérsic profile, and find that I recover accurately the same results as those published by van der Wel et al. (2012).

the decomposition provided by GALFIT. In the end, we prefer to used the results provided by GIM2D since this code does not require choosing starting conditions, which are known to influence strongly the final result of GALFIT owing to the presence of local minima in the χ^2 (Lang et al. 2014).

We do not further select galaxies based on their measured morphological parameters. Abramson et al. (2014) only used face-on galaxies in their $z = 0$ analysis (axis ratio larger than 0.8), arguing that the decomposition is less reliable for edge-on objects. We could not find any such trend in our simulations (see Sect. 5.3.2), and we also checked that no systematic trend emerges in the real data if we only use face-on galaxies. Because our sample is much smaller to start with, and since only 22 % of our galaxies pass this axis ratio cut, we therefore decide to use all galaxies regardless of their inclination.

For each galaxy that was properly fit, we now have an estimation of how the H -band flux is distributed between the disk and the bulge. From this decomposition, we can compute a light-weighted B/T , and we assess in the next section the robustness of this estimation. We will discuss in Sect. 5.3.3 how to convert this value into a mass-weighted ratio, to finally obtain the stellar mass of the disk.

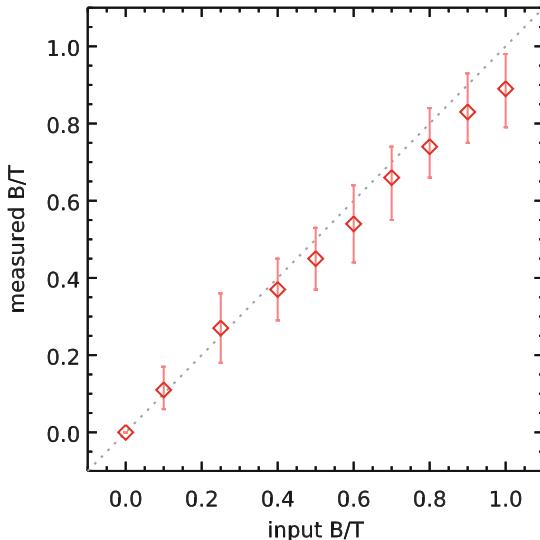
5.3.2 Simulated Galaxies

To test the robustness and quality of our morphological decomposition, we create a large set of simulated galaxies of known profiles and B/T , and try to measure their properties in the presence of photometric noise. To do so, we use GALFIT to model 5 000 idealized double Sérsic profiles ($n = 1$ and $n = 4$) of varying sizes, axis ratios, position angles, and fluxes, and place these models on empty regions of the real *HST* images. We then run both GALFIT and GIM2D trying to find back the input parameters.

We find that the total magnitude of the galaxy is always well recovered, except in the case of some catastrophic failures which happen almost exclusively with GALFIT. Enforcing that the measured total magnitude is close to that chosen in input effectively gets rid of most of these poor fits. For the real galaxies, we choose to compare the measured total magnitude to that quoted in the CANDELS catalogs, and discard GALFIT runs for which the difference is more than 0.5 magnitudes.

We also find that the bulge-to-disk decomposition is usually hopeless at $H > 23$, as the measured B/T are either very noisy or systematically biased toward roughly equal partition of the flux. For galaxies brighter than $H = 23$, we show in Fig. 5.3 the comparison between the B/T we put in the simulation, and the ones that are recovered by GIM2D. We find that the code is able to identify disk dominated galaxies with great accuracy, while bulge-dominated galaxies and intermediate systems show a slight systematic underestimation: given the choice, GIM2D will tend to put more flux in the disk component than in the bulge. This effect is small however, and we checked that our conclusions are not affected if we correct for it by adding 0.05 to the $B/T > 0.5$. We also observe that the uncertainty on the flux of the disk depends

Fig. 5.3 Comparison between the simulated B/T and that measured by GIM2D, for galaxies with $H < 22.5$. The median measured B/T are shown with *empty red diamonds*, and the error bars give the 16th and 84th percentiles of the distribution. The *dotted line* in the background gives the expected one-to-one relation (color figure online)



on B/T , with brighter bulges leading to more uncertain disk fluxes. For example, assuming constant mass-to-light ratio, for $M_{\text{disk}} \simeq 2 \times 10^{10} M_{\odot}$, the error on M_{disk} is 0.04 dex for $B/T \simeq 0$, and 0.07 dex for $B/T > 0.3$. It should be noted that these simulations are only able to capture the ability of the codes to recover what was put on the simulated image, i.e., idealized profiles with realistic photometric noise and neighbor contamination, but it does not allow us to say how reliable is the decomposition in the case of perturbed, irregular or clumpy galaxies, nor does it hint about actually measuring a disk *mass* (which is done in the next section), e.g., it does not contain varying mass to light ratios. Therefore the real uncertainties on the measurements are probably larger. Still, even doubled, the errors we estimate here are low enough for our purposes.

5.3.3 Estimating the Disk Mass

Once the flux of both the bulge and disk are measured, the last step is to measure the stellar mass of the disk. Both components have different mass-to-light ratios, since bulges are mostly made of old stars and will typically have higher mass-to-light ratios compared to the star-forming disks. In practice, since we are doing the decomposition in the H band (rest-frame i band at $z = 1$), the variation in mass-to-light ratio is supposed to be minimal (e.g., de Jong 1996). Yet, to prevent any bias in our results, we will nevertheless correct for this effect. The ideal way to treat this issue is to perform the decomposition on multiple photometric bands, and use the colors to infer accurate mass-to-light ratios as in Abramson et al. (2014), or even

complete SED fitting similar to what was done in Bruce et al. (2014). However this is only possible for the brightest objects (e.g., Bruce et al. 2014 only focused on galaxies with $M_* > 10^{11} M_\odot$).

Here we use a simpler approach where we assume an average mass-to-light ratio for the bulge components, infer the bulge masses, and subtract them from the total stellar masses. Doing so, we do not make any assumption on the mass-to-light ratio of the disk, and take best advantage of the robust stellar masses obtained by fitting the total photometry.

To determine the average mass-to-light ratio of bulges, we build a sample of “pure bulge” galaxies ($B/T > 0.8$) and compare their $1.6 \mu\text{m}$ (observer frame) luminosity against the stellar mass that was measured on the whole multi-wavelength photometry. Since these galaxies are clearly bulge-dominated, we can neglect the disk mass and assume that the observed mass-to-light ratio is representative of that of a bulge. The corresponding relation is shown in Fig. 5.4 (right). We derive the average trend by performing a linear fit to the running median in logarithmic space and obtain

$$\frac{M_{\text{bulge}}}{M_\odot} = \left(\frac{\nu L_{\nu, \text{bulge}}}{3.25 L_\odot} \right)^{1.09}, \quad (5.3)$$

with a constant residual scatter of about 0.1 dex. We then use this relation for all the other galaxies that are not bulge-dominated to estimate M_{bulge} , and subtract this value from M_* to obtain M_{disk} . The main advantage of this approach is that, although we perform the bulge-to-disk decomposition in a single band, we take advantage of

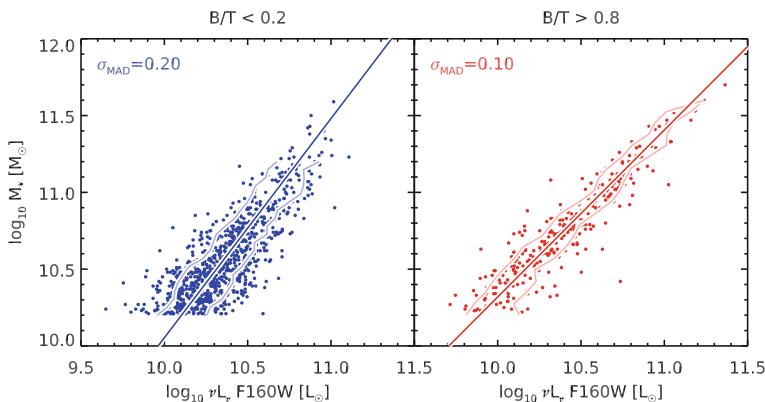


Fig. 5.4 Relation between the total stellar mass (M_*) and the measured luminosity from the *HST* H band flux (without k -correction) for a sample of disk-dominated galaxies ($B/T < 0.2$, left) and bulge-dominated galaxies ($B/T > 0.8$, right). Individual galaxies are shown with filled colored circles. The best-fit relation is shown with a straight line, and the dispersion around this relation is shown with light solid lines on each side. The global dispersion is given in the top-left corner of each plot, and is computed from the median absolute deviation (MAD) using $1.48 \times \text{MAD}(\Delta M_*)$ (color figure online)

the accurate mass-to-light ratio that was derived by fitting the total photometry of the galaxy, using a large number of photometric bands.

However, we rely here on the low scatter of the mass-to-light ratio in bulges. It is true that this ratio is less variable in bulges than in star-forming disks (see, e.g., Fig. 5.4, left), because the latter can display a wider variety of star formation histories. Still, bulges are expected to show some variation of their dust content and metallicity, and this will not be taken into account here. In particular, one possibility we cannot account for is that bulges in composite or disk-dominated galaxies may have different colors than pure bulges. Lastly, another downside of this empirical approach is that, since we do not measure the colors of each individual bulge, we cannot flag out the “blue bulges”, which are *not* bulges but likely compact nuclear starbursts. These are supposed to be rare though, and if anything, this population would end up substantially above the Main Sequence in the SFR– M_{disk} relation and bias the slope toward higher values.

To make sure that our results are not significantly biased by the adopted mass-to-light ratio calibration, we have tried several other methods for estimating the disk mass, e.g., assuming the same mass-to-light ratio for the bulge and the disk, or measuring also the average mass-to-light ratio in star-forming pure disks (Fig. 5.4, left) and combine it with the bulge mass-to-light ratio to estimate a mass-weighted B/T . These alternative estimations did not change our conclusions. It should be noted however that the typical dispersion observed when comparing these different disk masses is of the order of 0.2 dex. The crudest approach would be to assume the average M/L ratio of disks and apply it to the measured disk luminosity, without using the information provided by the total M_* . In this case the scatter rises to 0.3 dex, suggesting that this is a poor approach. Regardless, a typical scatter of 0.2 dex means that there is little hope of seeing the *dispersion* of the Main Sequence becoming smaller by using the disk mass, because the latter is too uncertain. However, the absence of systematic shifts in the derived stellar masses suggests that any modification of the slope of the Main Sequence will be correctly captured.

In Fig. 5.5, we show on the UVJ diagram the location of galaxies that are either disk-dominated ($B/T < 0.2$), intermediate ($0.2 < B/T < 0.6$), and bulge-dominated ($B/T > 0.6$) according to our mass-weighted bulge-to-total ratios. Reassuringly, the disk-dominated galaxies populate preferentially the UVJ active branch, while the bulge-dominated galaxies pile up in the passive cloud, although there is some overlap between the two populations close to the dividing line. Intermediate objects are preferentially in the passive region, but are also widely spread in the tip of the active branch. It should be noted that the relations we find between total stellar mass and B/T for UVJ star-forming and quiescent galaxies are consistent with those derived in Lang et al. (2014).

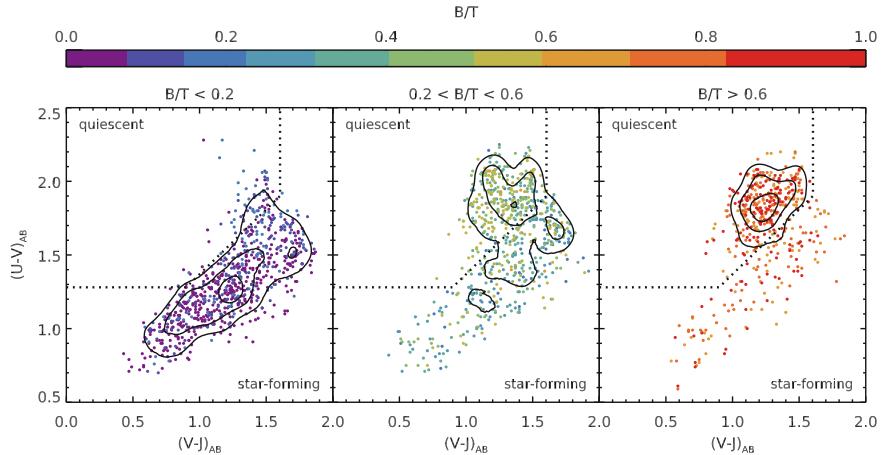


Fig. 5.5 Location of $z = 1$ and $M_* > 2 \times 10^{10} M_\odot$ galaxies with varying B/T on the UVJ diagram (left $B/T < 0.2$, middle $0.2 < B/T < 0.6$, right $B/T > 0.6$), using the total magnitudes of each galaxy. The dotted line shows the dividing line between the star-forming and quiescent populations, as defined in S15 and adapted from Williams et al. (2009). It is clear that both bulge- and disk-dominated galaxies occupy very different regions of the diagram, illustrating the good agreement between the colors and the morphology. However, intermediate galaxies with roughly equal mass in the disk and bulge (middle panel, $\langle B/T \rangle = 0.4$) are spread over the two regions, with a tendency for being preferentially in the quiescent region

5.4 Measuring Gas Masses

The star formation efficiency (SFE) is defined as the galaxy's current star formation rate divided by the mass of hydrogen gas found within the galaxy (M_{gas}). While we have robust estimates of the SFRs, measuring gas masses is notoriously difficult, especially among distant galaxies. We choose here to infer the gas masses from the dust masses (M_{dust}), which themselves can be measured from the dust continuum emission in the FIR. This approach has been used extensively in the recent literature to constrain the SFE of distant galaxies (e.g., Magdis et al. 2011, 2012; Magnelli et al. 2012a; Santini et al. 2014; Scoville et al. 2014; Béthermin et al. 2015) and is based on the observed anti-correlation between the gas-to-dust ratio $M_{\text{gas}}/M_{\text{dust}}$ and the metallicity Z in the Local Universe (e.g., Leroy et al. 2011; Rémy-Ruyer et al. 2014).

In this section, we describe the measurement of dust masses (Sect. 5.4.1) from the FIR to submm photometry, and then detail the derivation of the associated gas masses (Sect. 5.4.2).

5.4.1 Dust Masses

Accurate dust masses can only be derived from FIR measurements down the Rayleigh–Jeans tail of the dust continuum, meaning at $z = 1$ that we need to measure the observer-frame emission of galaxies at $\lambda \geq 400 \mu\text{m}$. While *Herschel* does provide deep imaging at $500 \mu\text{m}$, the poor angular resolution prevents measuring the $500 \mu\text{m}$ flux of most galaxies, since finding the right counterpart to the fluxes measured on these maps is challenging (see, e.g., Shu et al. in prep.). This issue can be avoided by stacking the image, since the contribution from neighboring sources averages out to form a constant background. However, there still remain a source of uncertainty which is the contribution of galaxy clustering (e.g., Béthermin et al. 2010). In the presence of clustering, the contribution of neighboring sources will not average out to a uniform value, and instead will tend to produce more flux toward to the position of the stacked galaxies. In S15, we implemented an empirical correction to remove this flux boosting, which was derived from a set of realistic simulated images. The stacked $500 \mu\text{m}$ fluxes in the simulation were found to be boosted by 20 % on average, and we therefore de-boosted the observed fluxes by that same amount.⁴ After this factor is taken into account, no remaining bias was found in the stacked fluxes. We also considered stacking longer wavelength sub-millimeter data from AzTEC or LABOCA, however these are only available for a few fields⁵ hence reducing significantly the number of stacked sources. Combined with the fact that, at $z = 1$, the expected flux in these bands is fairly low, we could not detect any significant signal. These upper limits are consistent with the rest of *Herschel* photometry at the 1 to 2σ level.

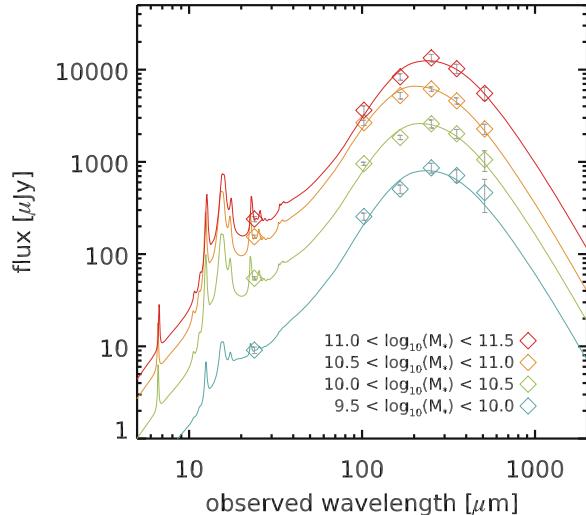
For our $z = 1$ sample, we therefore use the stacked SEDs of S15, which are reproduced here in Fig. 5.6. These SEDs were built by stacking all the *UVJ* star-forming galaxies in the four CANDELS fields at $0.7 < z < 1.3$ and in four bins of stellar mass: $\log_{10}(M_*/M_\odot) = 9.5$ to 10, 10 to 10.5, 10.5 to 11 and 11 to 11.5. As described above, a correction for clustering is also applied.

We then fit the stacked photometry with a library of template SEDs built from the amorphous carbon dust model of Galliano et al. (2011). This new library is presented in Chap. 3, and is introduced to extend the Chary and Elbaz (2001) SED library (hereafter CE01), with the aim to provide a wider and finer grained range of dust temperatures (or, equivalently, $L_{\text{IR}}/M_{\text{dust}}$) and finer control on the PAH mass-fraction (or, equivalently, $\text{IR8} \equiv L_{\text{IR}}/L_8$). If the contribution of PAHs is neglected (n.b.: they represent only 4 % of the total dust mass), the following relation links together the dust mass M_{dust} , the total infrared luminosity L_{IR} , the mass-weighted average dust temperature T_{dust} , and the mass-weighted average intensity of the radiation field $\langle U \rangle$:

⁴Note that these correction factors depend greatly on the way the fluxes are measured, as shown in the Appendix of S15.

⁵AzTEC in GOODS–North and LABOCA in GOODS–South. Both are also covering COSMOS at shallower depth.

Fig. 5.6 Mean stacked FIR SEDs of $z = 1$ galaxies in our sample, split in four mass bins. The broadband photometry (open diamonds) is taken from Chap. 2. The fit to the stacked measurements is performed using the dust models of Galliano et al. (2011) (see also Chap. 3)



$$\frac{L_{\text{IR}}}{L_{\odot}} = 185 \frac{M_{\text{dust}}}{M_{\odot}} \left(\frac{T_{\text{dust}}}{17.5 \text{K}} \right)^{5.54} = 185 \frac{M_{\text{dust}}}{M_{\odot}} \langle U \rangle. \quad (5.4)$$

Each SED in the library is calibrated per unit M_{dust} , and therefore the dust mass is trivially obtained from the normalization of the best-fit template. Here, we allow the dust temperature to vary between 15 and 50 K, while the PAH mass-fraction is left free to vary between 0 and 1. The best-fit values we obtain are referenced in Table 5.2.

The infrared luminosities we derive with this library are in perfect agreement with those obtained in S15 using the CE01 library. As a cross check, we also fit this photometry with the CIGALE SED fitting code, using the Draine and Li (2007) dust SED library. We recover identical L_{IR} , but M_{dust} values that are systematically higher by a factor of two. Systematic differences in the dust masses are typically found by comparing the results of two different approaches, e.g., comparing the results from the Draine and Li (2007) library against simple modified black bodies (as is shown in Magdis et al. 2012 and Magnelli et al. 2012a), or different chemical compositions of dust grains within the same model (e.g., graphite and silicate *versus* amorphous carbon grains, as in Galliano et al. 2011). The factor of two we observe here is consistent with the value reported by Galliano et al. (2011), who argue that dust masses derived by models using graphite (like, e.g., the models of Draine and Li 2007) instead of amorphous carbon grains are overestimated by a factor of 2.6. They also claim that this overestimation creates a tension with the measured metallicity of the Large Magellanic Cloud by violating the element abundances, and therefore advocate instead the use of amorphous carbon grains in dust models. Independently of this choice, we do not expect that galaxies at different stellar masses host dust grains of radically different chemical composition, hence we argue that if our measurements are biased because of the assumption on dust composition, this bias only affects

our dust mass estimates globally. This is of no consequence for the present work, since it will not affect the relative evolution of the SFE. On the other hand, it also emphasizes that without precise knowledge of the detailed chemical composition of dust, the *absolute* value of the dust masses should be taken with a grain of salt.

For galaxies in the HRS, angular resolution is not an issue, and the *Herschel* photometry of each galaxy can be obtained and fitted individually. The dust mass is estimated directly by CIGALE, when fitting the photometry to obtain the stellar mass and the SFR (see Ciesla et al. 2014 and Ciesla et al. in prep.). As written above, CIGALE uses the Draine and Li (2007) SEDs to model the dust emission. To homogenize this sample with our $z = 1$ dust mass measurements that are obtained with the models of Galliano et al. (2011), we therefore correct the dust masses given by CIGALE down by a factor of two.

5.4.2 Gas Masses

The idea behind the conversion from M_{dust} to M_{gas} is that a universal fraction f_{d} of all the metals in the galaxy are locked into dust grains, while the remaining fraction remains mixed with the gas (Franco and Cox 1986). With this assumption and a measurement of the dust mass, one just needs to know the gas-phase metallicity (Z) to infer the gas mass:

$$M_{\text{gas}} = \frac{1}{Z} \frac{1 - f_{\text{d}}}{f_{\text{d}}} M_{\text{dust}}. \quad (5.5)$$

The value of f_{d} is unknown, but it can be inferred empirically from observations where both the dust and the gas masses are known. In these cases, the gas mass is usually inferred by adding together 21 cm measurements of the neutral atomic hydrogen, and estimates of the molecular hydrogen mass, which are typically obtained from the carbon monoxide (CO) emission lines (since, indeed, molecular hydrogen is extremely hard to observe directly). This latter step implies yet another uncertainty on the conversion factor from CO intensity to molecular gas mass (α_{CO}). To alleviate this problem, Leroy et al. (2011) performed a resolved analysis of local galaxies, inferring jointly the gas-to-dust ratio and α_{CO} from combined dust, 21 cm and CO observations. Assuming that the gas-to-dust ratio remains constant throughout each galaxy, they observed a relation between $M_{\text{gas}}/M_{\text{dust}}$ and metallicity, and found a dependence that is consistent with Eq. 5.5.

Once the dust masses are measured (see previous section), the second step is therefore to estimate the metallicity. Since only half of the galaxies in the HRS have individual metallicity measurements (Hughes et al. 2013), and almost none of the galaxies in our $z = 1$ sample, we need to use empirical recipes to estimate the metallicities. Following Magdis et al. (2012), Santini et al. (2014) and Béthermin et al. (2015), we estimate the metallicity from the Fundamental Metallicity Relation (FMR, Mannucci et al. 2010 Eq. 5)

$$(12 + \log_{10}(\text{O/H}))_{\text{KD02}} = \begin{cases} 8.9 + 0.47 (\mu_{0.32} - 10) & \text{for } \mu_{0.32} < 10.4 \\ 9.07 & \text{for } \mu_{0.32} > 10.4 \end{cases}, \quad (5.6)$$

with $\mu_{0.32} \equiv \log_{10}(M_* [\text{M}_\odot]) - 0.32 \times \log_{10}(\text{SFR} [\text{M}_\odot/\text{yr}])$, and where both M_* and SFR are converted to the Chabrier (2003) IMF (i.e., divided by 1.8 from the Salpeter values). For our $z = 1$ sample, we use the average stellar mass and SFR obtained in the stacks, and for the $z = 0$ HRS galaxies without metallicity measurement we use their respective M_* and SFR. We checked that using this prescription or estimating the metallicity from the mass–metallicity relation (e.g., Zahid et al. 2011) would not change our conclusions (+0.12 dex metallicity shift at $z = 1$, after accounting for the different calibration). It is also worth noting that Béthermin et al. (2015) argue for an additional redshift-dependence of the FMR, i.e., that Eq. 5.6 may not hold in the distant Universe. However, this is not an issue for the present study since, first, the difference proposed by Béthermin et al. (2015) is a constant shift of the metallicity at all stellar masses, and second, it only takes place at $z > 1.7$. On the other hand, Kewley and Ellison (2008) showed that there exists substantial systematic differences of metallicity measurements, depending both on the available observables used to derive the oxygen abundance, and the calibration that is used. For example, the FMR was derived using the Kewley and Dopita (2002) (KD02) calibration, while the metallicities of Magdis et al. (2012) are obtained with the prescription of Pettini and Pagel (2004) (PP04). According to Kewley and Ellison (2008), the difference between these two metallicity estimates is roughly constant and equal to about 0.25 dex (at least in the metallicity range considered in this paper), with a scatter of only 0.05 dex: it is only a global shift of the absolute metallicity, and will not affect the relative trends. To derive accurate dust-to-gas ratios, it is nevertheless important to make sure that the same metallicity calibration is used consistently in all calculations. For this reason, since we are going to use the data of Magdis et al. (2012), we convert the FMR metallicities to the Pettini and Pagel (2004) “[N II]” scale, following the calibration proposed by Kewley and Ellison (2008):

$$(12 + \log_{10}(\text{O/H}))_{\text{PP04}} = 569.4927 - 192.5182x + 21.91836x^2 - 0.827884x^3, \quad (5.7)$$

with $x \equiv (12 + \log_{10}(\text{O/H}))_{\text{KD02}}$. As written above, in practice for the galaxies we consider in this study these “PP04” abundances are systematically lower by 0.3 dex compared to the original “KD02” values (this constant shift holds within 0.05 dex for all $12 + \log_{10}(\text{O/H})_{\text{KD02}} > 8.5$).

The measured metallicities of the HRS galaxies are already in this scale, and needed no conversion. For galaxies with a metallicity measurement, comparing the latter to the metallicity derived from the FMR, we find a median offset of 0.08 dex and a scatter of 0.1 dex. Since these latter values are low, and to avoid mixing together metallicities that are directly observed and those that are inferred from the FMR, we

decide to use the FMR-based metallicities for all galaxies in the HRS. We checked that our results are not affected by this choice.

The last missing ingredient to estimate gas masses is the gas-to-dust ratio or, equivalently, f_d in Eq. 5.5. Here we use the gas-to-dust ratios measured by Leroy et al. (2011), that we multiply by 2 to account for systematic differences in the dust mass measurements between the dust model that we used and that of Draine and Li (2007) (see previous section). Then, to relate these measurements to metallicity, we refer to Magdis et al. (2012) who have conveniently converted all the measurements of Leroy et al. (2011) to a uniform metallicity scale (PP04), and found a best-fit relation of $\log_{10}(M_{\text{gas}}/M_{\text{dust}}) = 10.54 - 0.99 \times (12 + \log_{10}(\text{O/H}))$, i.e., with a metallicity dependence very close to that of Eq. 5.5. Taking into account the systematic difference in the dust masses, and re-fitting the data by assuming the functional form of Eq. 5.5 (i.e., using a slope of -1 for the metallicity), we get

$$\log_{10}\left(\frac{M_{\text{gas}}}{M_{\text{dust}}}\right) = (10.92 \pm 0.04) - (12 + \log_{10}(\text{O/H}))_{\text{PP04}}, \quad (5.8)$$

Assuming a solar oxygen abundance of $(12 + \log_{10}(\text{O/H}))_{\odot} = 8.69$ (Allende Prieto et al. 2001) and a solar metallicity of $Z_{\odot} = 0.017$ (Grevesse and Sauval 1998), this leads to the equivalent expression

$$\frac{M_{\text{gas}}}{M_{\text{dust}}} = (170 \pm 16) \times \frac{Z_{\odot}}{Z}, \quad (5.9)$$

which is consistent with the gas-to-dust ratio of the Milky Way ($M_{\text{gas}}/M_{\text{dust}})_{\text{MW}} = 158$ (Zubko et al. 2004). This prescription is therefore equivalent to assuming that 26 % of the metals are locked into dust.⁶ For our $z = 1$ sample, this yields gas-to-dust ratios between 145 and 387 (the precise values we obtain are listed in Table 5.2), while it ranges from 145 to 494 for the $z = 0$ HRS galaxies (which cover a wider metallicity range).

Applying Eq. 5.8 to the measured dust masses, we can infer the total gas mass in each stacked bin at $z = 1$, and for each HRS galaxy.

To check if our results depend on the way redshifts, stellar masses and *UVJ* classifications were derived in S15, we also run the same analysis using the “official” photometric redshifts and stellar masses of the CANDELS team, which were obtained by combining together the results of different fitting codes (Dahlen et al. 2013; Santini et al. 2015), as well as the 3DHST catalogs (Skelton et al. 2014). We find that using the CANDELS fits yield the same conclusions, but using the 3DHST catalogs changes substantially the measured SFEs. To investigate this issue, we analyze the intersection of our sample and that of 3DHST, i.e., galaxies that satisfy the selection criteria in both catalogs simultaneously. This reduces the analysis to about half of the initial sample, and yields SFEs that are comparable to those presented in this paper. We therefore conclude that our results are robust against catalog changes, and that there

⁶Using the dust masses from the Draine and Li models would increase this fraction to 41 %.

is probably an issue in the 3DHST catalogs. Investigating this latter issue any further goes out of the scope of this paper.

Lastly, as a consistency check for the HRS, we compare our gas masses against those estimated from the combination of 21 cm and CO emission line fluxes (using data from Boselli et al. 2014), with a constant $\alpha_{\text{CO}} = 3.6 \text{ M}_\odot / (\text{K km/s/pc}^2)$ (Strong et al. 1988). The latter are found to be systematically larger by 30 %, with a scatter of 0.2 dex. Since the vast majority (90 %) of the $M_* > 10^{10} \text{ M}_\odot$ star-forming galaxies are detected in both atomic and molecular surveys, we also do the following analysis with these alternative gas mass estimates. We find that our conclusions remain unchanged, save for this global shift of the gas masses by a factor of 1.3. In the end, we prefer to use the dust-based estimates in order to preserve the homogeneity of our analysis.

5.5 Results

5.5.1 The SFR– M_{disk} Relation at $z = 1$

Having measured the disk masses, we can now see if the SFR– M_{disk} relation is universal and linear by comparing the slopes of the Main Sequence using either the total stellar mass M_* or the disk mass M_{disk} . To be able to measure this slope on our whole sample at once, and because our redshift window is relatively large, we correct for the redshift evolution of the Main Sequence by renormalizing the SFR of each galaxy to a common redshift of $z = 1$. To do so, we use the redshift evolution measured in S15, taking the trend of low-mass galaxies where the bending of the Main Sequence is negligible. This correction is typically of the order of 0.05 dex, and no more than 0.1 dex.

In Fig. 5.7, we show the resulting SFR– M_* (top) and SFR– M_{disk} (bottom) relations of our sample. Each panel focuses on a different range of B/T , starting from disks-dominated galaxies on the left, then increasing progressively the contribution of the bulge. In the rightmost panels, we show all galaxies regardless of their B/T . We show the running median on the measurements in each plot, either considering all the galaxies (purple lines) or only the UVJ star-forming galaxies (blue lines), and compare them to the stacked Main Sequence of S15. In the top-rightmost panel, this running median overlaps with the stacked relation, which indicates that we are not strongly affected by the SFR selection of our sample. However, we can see from the top-leftmost panel that disk-dominated galaxies do not populate a particularly different region of the SFR– M_* diagram: they cluster around the stacked relation of S15, and follow a sequence of slope 0.67 ± 0.07 (from $M_* = 3 \times 10^{10}$ to $3 \times 10^{11} \text{ M}_\odot$). Even after subtracting the bulge mass, which is by definition very low in these systems, the measured slope is 0.65 ± 0.08 , i.e., clearly not unity. For the other galaxies, we do find a trend for some of the lowest sSFR objects to be brought back toward the Main Sequence by removing the bulge mass, but they constitute a very

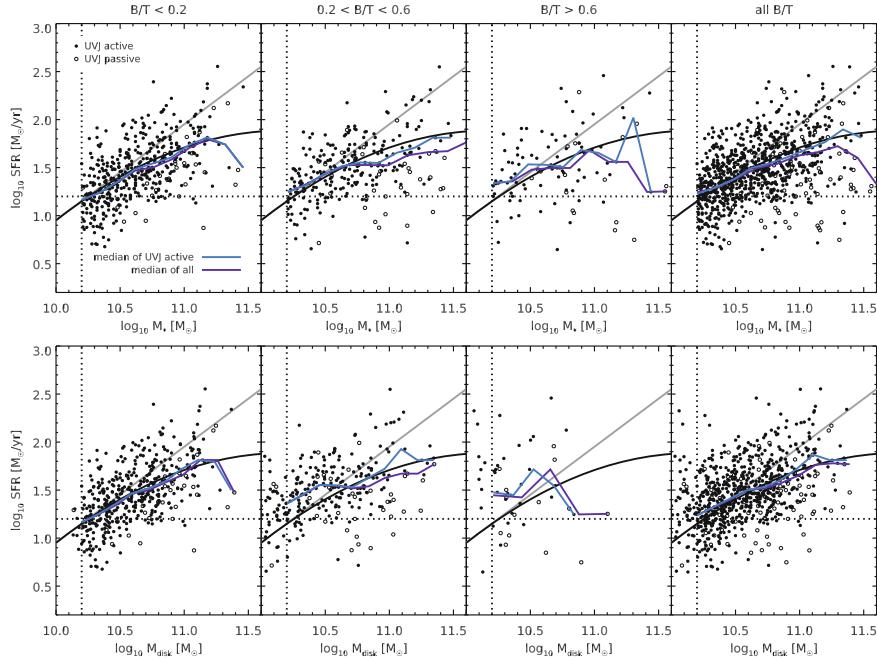


Fig. 5.7 **Upper panel** Location of galaxies with varying B/T on the $SFR-M_*$ plane, using the mass and star formation rate (IR+UV) of the whole galaxy. On all plots, the *vertical dotted line* shows our adopted stellar mass cut, the *horizontal dotted line* is the 90 % completeness in SFR, and the *solid black line* shows the locus of the $z = 1$ Main Sequence as observed through stacking in **S15**, while the *solid gray line* shows the extrapolation of the low-mass trend assuming a slope of unity, as observed at lower stellar masses. In each column, galaxies of different B/T are plotted, either as plain circles for UVJ active galaxies, or open circles for UVJ passive galaxies. In the rightmost panel, we show all galaxies regardless of their B/T . The *solid colored lines* show the running median of the sample, either taking all galaxies (*purple line*) or only the UVJ active ones (*blue line*). **Lower panel** Same as upper panel, but on the $SFR-M_{\text{disk}}$ plane (color figure online)

small fraction of the whole sample (in fact, a good fraction of these are classified as UVJ passive), and cannot counterbalance the bending observed in disk-dominated galaxies. In the end, the slope of the $SFR-M_{\text{disk}}$ relation as measured on the whole sample (bottom-rightmost panel) is 0.60 ± 0.05 , and therefore we do not find that the $SFR-M_{\text{disk}}$ relation is universal.

In their $z = 0$ study, Abramson et al. (2014) only considered galaxies with $B/T < 0.6$, arguing that galaxies above this threshold cannot be fitted reliably (we show indeed in Sect. 5.3.2 that disk masses measured in bulge-dominated galaxies are the most uncertain). We therefore tried to reject galaxies with $B/T > 0.6$, and did not find any significant difference. In fact, most of them do not show any measurable IR emission (83 %, compared to 46 % for galaxies with $B/T < 0.6$), and are likely genuine bulge-dominated objects.

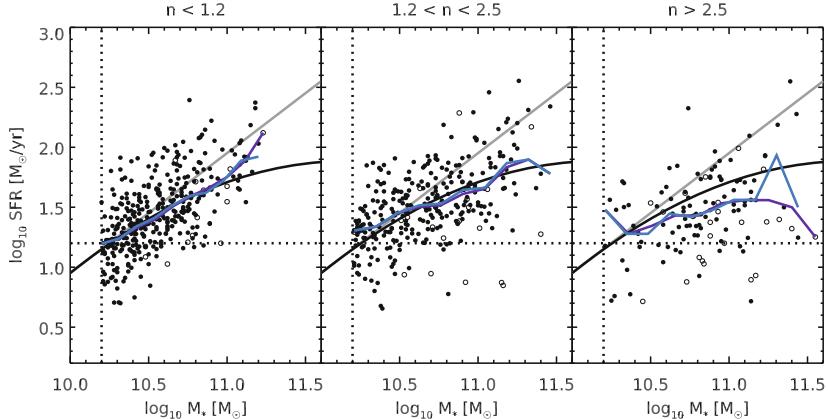


Fig. 5.8 Same as the upper panel of Fig. 5.7, but this time varying the Sérsic index n

Table 5.1 Measured slopes of the SFR– X relation, where X is either M_* or M_{disk} . All slopes were obtained by fitting a straight line (in logarithmic space) to the running median shown in Figs. 5.7 and 5.8, considering only star-forming galaxies with $10.2 < \log_{10}(X) < 11.3$. Uncertainties are estimated by bootstrapping

	All B/T	$B/T < 0.2$	$n < 1.2$
SFR– M_*	0.54 ± 0.05	0.67 ± 0.07	0.75 ± 0.05
SFR– M_{disk}	0.60 ± 0.05	0.65 ± 0.08	—

To make sure that our results are not caused by an uncertain bulge-to-disk decomposition, we show in Fig. 5.8 how the SFR– M_* diagram is populated by galaxies of varying effective Sérsic index n (van der Wel et al. 2012 and our own fits in GOODS–North, see Sect. 5.3.1). While the Sérsic index alone is not well suited for measuring the disk masses of composite systems, it is a robust way of identifying disk-dominated galaxies. Indeed, the fit is intrinsically simpler and therefore more stable, and the presence of a significant bulge component will rapidly make the effective Sérsic index depart from 1, the nominal value for pure disks (see Appendix A of Lang et al. 2014). We find that disk-dominated galaxies ($n < 1.2$) follow a slightly steeper slope of 0.75 ± 0.05 , consistent with that found in Salmi et al. (2012), but this is still not unity. These slope measurements are summarized in Table 5.1.

5.5.2 Gas Fraction and Star Formation Efficiency at $z = 1$

We show in Fig. 5.9 (left) the behavior of the SFE as a function of the stellar mass in our stacked $z = 1$ sample. These values are also reported in Table 5.2. From this

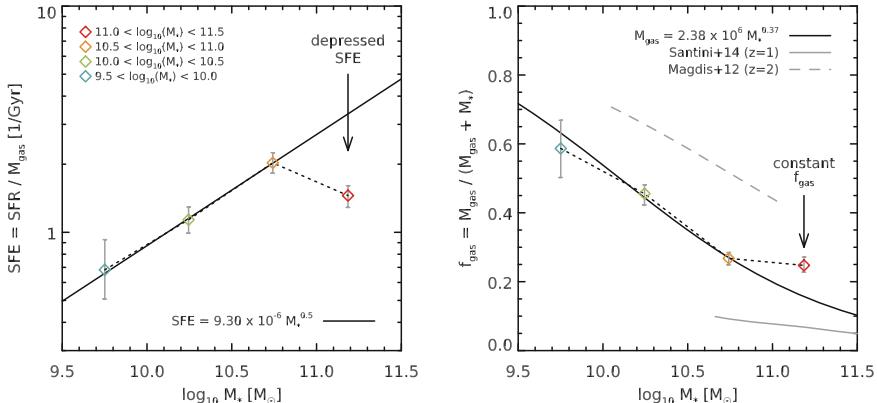


Fig. 5.9 **Left** Relation between the SFE = SFR/M_{gas} and the gas mass (M_{gas}) for Main Sequence galaxies at $z = 1$. Colored diamonds show the measured SFRs and M_{gas} of our sample, the color being associated to the stellar mass as in Fig. 5.6. The best-fit power law to our measurements, excluding the most massive point, is given with a black solid line (Eq. 5.10). **Right** Gas fraction ($f_{\text{gas}} \equiv M_{\text{gas}} / (M_{\text{gas}} + M_*)$) as a function of the stellar mass (M_*) for Main Sequence galaxies at $z = 1$. The legend is the same as in the left figure, and here the solid black line gives the value of f_{gas} computed using the best-fit $M_{\text{gas}} - M_*$ relation, also excluding the most massive point in the fit. The resulting expression of f_{gas} is given in Eq. 5.11. We also show the measured gas fractions by Magdis et al. (2012) at $z = 2$ with a dashed gray line (color figure online)

figure, one can see that the SFE of galaxies at $M_* < 10^{11} \text{ M}_\odot$ rises steadily with stellar mass, following

$$\text{SFE [1/Gyr]} = \frac{\text{SFR}}{M_{\text{gas}}} = 9.30 \times 10^{-6} \left(\frac{M_*}{\text{M}_\odot} \right)^{0.5}. \quad (5.10)$$

However, our data point with the highest gas mass, i.e., corresponding to the stellar mass of $2 \times 10^{11} \text{ M}_\odot$ where the bending of the Main Sequence is most pronounced, has an SFE that is a factor of 2 lower than that predicted from this scaling law. Our data clearly favor two regimes of SFE: low stellar mass galaxies follow a universal relation, and high stellar mass galaxies drop below this trend.

In contrast, the gas fraction (Fig. 5.9, right) is found to decrease continuously with stellar mass (similarly to what was found in Magdis et al. 2012 and Santini et al. 2014). This is the expected behavior if the Main Sequence has a linear (or sublinear) slope while the SFR– M_{gas} law (the so-called integrated Schmidt–Kennicutt law) is superlinear with a power-law slope of $n > 1$ (e.g., Daddi et al. 2010; Sargent et al. 2014; Santini et al. 2014). Indeed, if $\text{SFR} \sim M_*$ and $\text{SFR} \sim M_{\text{gas}}^n$, then $M_{\text{gas}} \sim M_*^{1/n}$ and the gas fraction has to decrease. By fitting the $M_{\text{gas}} - M_*$ relation for galaxies with $M_* < 10^{11} \text{ M}_\odot$, we get

Table 5.2 Average physical properties of the galaxies in the stacked $z = 1$ sample. The quoted errors indicate the uncertainty on the average, not the intrinsic spread of the population. These uncertainties are derived through bootstrapping half of the full sample, recomputing all quantities for each bootstrap realization separately, then measuring the standard deviation among all realizations. The gas-to-dust ratio is randomized within the allowed statistical uncertainty (Eq. 5.8). The resulting values are then divided by $\sqrt{2}$ to take into account that only half of the initial sample was used in each bootstrap realization

$M_* 10^{10} M_\odot$	$M_{\text{dust}} 10^7 M_\odot$	$L_{\text{IR}} 10^{10} L_\odot$	$T_{\text{dust}} (\text{K})$	$f_{\text{PAH}} (\%)$	SFR M_\odot/yr
0.56	$2.1^{+0.9}_{-0.5}$	$2.4^{+0.2}_{-0.2}$	$24.5^{+1.3}_{-1.4}$	$0.8^{+0.9}_{-0.5}$	$5.5^{+0.3}_{-0.4}$
1.8	$5.2^{+0.8}_{-0.5}$	$8.7^{+0.3}_{-0.3}$	$26.1^{+0.3}_{-0.7}$	$4.5^{+0.2}_{-0.2}$	$16.7^{+0.4}_{-0.5}$
5.5	$10.2^{+0.7}_{-0.9}$	$23.0^{+0.9}_{-0.8}$	$27.7^{+0.6}_{-0.5}$	$4.9^{+0.3}_{-0.3}$	$40.9^{+1.5}_{-1.4}$
16	$34.7^{+4.1}_{-3.2}$	$41.7^{+2.3}_{-2.1}$	$24.5^{+0.4}_{-0.5}$	$4.4^{+0.3}_{-0.3}$	$73.3^{+3.8}_{-3.7}$
$M_* 10^{10} M_\odot$	$12 + \log_{10}(\text{O/H})$ (PP04 [N II])	$M_{\text{gas}}/M_{\text{dust}}$	$M_{\text{gas}} 10^{10} M_\odot$	SFE 1/Gyr	$f_{\text{gas}} (\%)$
0.56	8.33	387^{+24}_{-22}	$0.8^{+0.4}_{-0.2}$	$0.68^{+0.26}_{-0.18}$	$58.7^{+7.8}_{-8.8}$
1.8	8.47	284^{+20}_{-19}	$1.5^{+0.2}_{-0.2}$	$1.14^{+0.14}_{-0.15}$	$45.5^{+3.1}_{-3.1}$
5.5	8.62	197^{+10}_{-15}	$2.0^{+0.2}_{-0.2}$	$2.03^{+0.25}_{-0.20}$	$26.8^{+2.2}_{-2.0}$
16	8.76	145^{+9}_{-10}	$5.0^{+0.8}_{-0.6}$	$1.45^{+0.21}_{-0.19}$	$24.7^{+2.8}_{-2.3}$

$$\frac{M_{\text{gas}}}{M_\odot} = 2.38 \times 10^6 \left(\frac{M_*}{M_\odot} \right)^{0.37},$$

$$f_{\text{gas}} = \frac{1}{1 + \left(\frac{M_*}{1.32 \times 10^{10} M_\odot} \right)^{0.63}}. \quad (5.11)$$

For galaxies with $M_* > 3 \times 10^{10} M_\odot$, we measure a constant value of $f_{\text{gas}} \equiv M_{\text{gas}}/(M_{\text{gas}} + M_*) \sim 26\%$, so that galaxies with $M_* > 10^{11} M_\odot$ actually have larger gas fractions than expected from the above trend. This can be explained if these galaxies also had lower SFEs in the past, suggesting that we are witnessing a process that acts on long timescales.

We also find that the overall decrease of gas fraction cannot be explained solely from the growing mass of the bulges. Indeed, if we substitute the disk mass to the total stellar mass, using the average B/T measured in each mass bin and assuming that galaxies of $M_* < 10^{10} M_\odot$ are pure disks, the gas fraction in the disk is also found to decrease, albeit with a slightly shallower slope. Similar results are obtained if we use the $B/T - M_*$ relations of Lang et al. (2014).

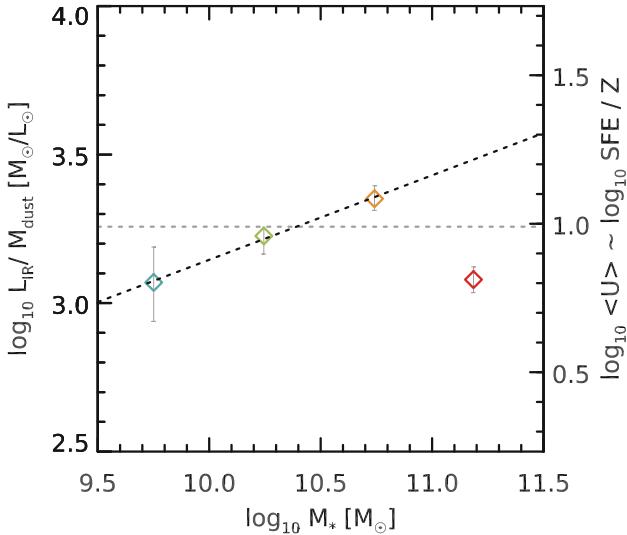


Fig. 5.10 Ratio between the dust mass (M_{dust}) and the total infrared luminosity (L_{IR}) as a function of the stellar mass for stacked galaxies at $z = 1$. Colors are the same as in Fig. 5.9. We overplot a linear fit (in log space) of the first three mass bins with a *dotted black line*. We also perform a second fit by imposing a flat slope, shown here with a *gray dotted line*, and following the trend observed by Magdis et al. (2012). This shows that the data at low stellar mass is roughly consistent with being flat, as reported in Magdis et al. (2012), in which case the drop in the highest mass bin would be less pronounced but still significant (4σ)

It should be noted that the SFE and f_{gas} we measure in high-mass galaxies are consistent with the $z = 1$ value reported by Béthermin et al. (2015), who applied the same methodology to a single mass bin around $M_* \sim 10^{11} M_\odot$ using galaxies from the larger COSMOS field. On the other hand, similar measurements were performed in Santini et al. (2014), in the same field as Béthermin et al. (2015), finding smaller gas masses by about a factor of 3. The discrepancy appears to come from different calibrations of the dust-to-gas ratio, and therefore should only result in a systematic shift. In any case, owing to the shallow depths of the COSMOS survey, Santini et al. (2014) could only focus on galaxies more massive than $3 \times 10^{10} M_\odot$, i.e., they do not probe the linear Main Sequence regime (as is illustrated in Fig. 5.9, right).

Lastly, to see how the assumptions about metallicity and gas-to-dust ratio affect our result, we show in Fig. 5.10 the $L_{\text{IR}}/M_{\text{dust}}$ ratio, which is a direct observable. The behavior of this quantity is very similar to that of the SFE, namely there is a steady rise with stellar mass, and then a sudden drop at $M_* > 10^{11} M_\odot$. This should not come as a surprise, knowing that our estimated gas-to-dust ratio ends up being a simple power law of the stellar mass (see Sect. 5.4), and that the SFRs in this sample are largely dominated by the dust-obscured, IR-luminous component.

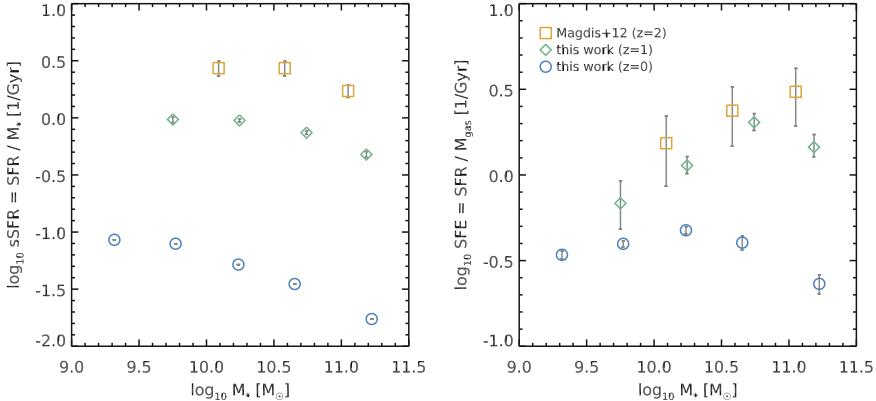


Fig. 5.11 **Left** Relation between the specific SFR ($\text{sSFR} = \text{SFR}/M_*$), and the stellar mass (M_*), at various redshifts. Our $z = 1$ stacked measurements from S15 are shown with *empty diamonds*, and the average values of the star-forming HRS galaxies are shown with *empty circles* (see Ciesla et al. in prep.). The associated error bar is the error on the mean, not the dispersion of the sample. We compare these measurements to the $z = 2$ values obtained by Magdis et al. (2012) for star-forming BzK galaxies. **Right** Same as left, but replacing the sSFR by the star formation efficiency (SFE = $\text{SFR}/M_{\text{gas}}$). The *diamonds* and *circles* use the gas mass estimated in this paper, while the empty squares come from Magdis et al. (2012), and were computed with the same method

5.5.3 A Progressive Decrease of the SFE with Time

In Fig. 5.11 (right) we put together our SFR and M_{gas} measurements at both $z = 1$ (previous section) and $z = 0$ using galaxies from the HRS survey to display the evolution of the SFE with stellar mass and redshift. The values in the HRS are obtained by binning galaxies in stellar mass, and computing the mean SFE in each bin, since all the HRS star-forming galaxies are individually detected by *Herschel*, and therefore have individual gas masses estimates. These results are compared to that of Magdis et al. (2012), who performed a similar analysis in the GOODS fields, stacking galaxies in different bins of stellar mass from $M_* = 10^{10}$ to $3 \times 10^{11} \text{ M}_\odot$, but focusing on $z = 2$ BzK galaxies.⁷ The selection effects inherent to the BzK classification are not very well understood, and it is known that this selection tends to affect the shape of the Main Sequence (Speagle et al. 2014). With this caveat in mind, we proceed comparing these results to our data at $z = 0$ and $z = 1$.

Similarly to our $z = 1$ sample, the most massive galaxies in the HRS ($M_* > 10^{10} \text{ M}_\odot$) are also found to have a reduced SFE, thereby confirming the trend observed

⁷They did stacked galaxies at $z = 1$, but did not separate them in different stellar mass bins. Also, since the BzK selection only selects star-forming galaxies at $z = 2$, they had to use another method to discard quiescent galaxies at $z = 1$. To do so, they used a cut in Sérsic index of $n < 1.5$ (see e.g., Wuyts et al. 2011, and Fig. 5.8). Because the associated selection effects are not obvious to determine, we prefer not to consider this data point in the present analysis, although the gas fraction they report is compatible with the one we measure here.

in the previous section. However, Magdis et al. (2012) observe a fairly different picture than the one we present here, since their galaxies of all stellar mass are found to lie on the same SFR– M_{gas} relation, i.e., following a universal star formation law.

In fact, this is fully consistent with the observed evolution of the high-mass slope of the Main Sequence (see, e.g., the comprehensive analysis of Gavazzi et al. 2015), since at $z = 2$ the SFR– M_* relation is found to be almost linear (see S15 and Fig. 5.11, left), indicating that whatever process drives this change of slope has not yet taken place. On the other hand, at $z = 0$ the bending of the Main Sequence is more pronounced, in agreement with the more pronounced drop of SFE that we observe for the HRS galaxies.

5.6 Discussion

5.6.1 Quantifying the “Quenching” and “Downfall” Rates

We find that the bending of the Main Sequence cannot be caused by abnormally low gas fractions, but is instead resulting from a progressive decrease of the star formation efficiency, as shown in Figs. 5.9 and 5.11. These observations converge toward a “slow downfall” of star formation, where massive galaxies gradually decrease their star formation activity while staying on the Main Sequence. While staying on the Main Sequence, these galaxies become gradually less efficient in their star formation activity instead of abruptly turning off though a “fast quenching”. Because the SFE is going down with time, these galaxies do not grow too massive by $z = 0$, as shown in Leja et al. (2015) who simulate the evolution of the observed stellar mass function using a Main Sequence of varying slope. The downfall of the star formation rate in massive Main Sequence galaxies may lead to the death of galaxies if, e.g., the gas surface density falls below the critical density that is necessary to switch on the Schmidt–Kennicutt relation, but our analysis does not allow us to make any firm claim favoring or disfavoring a scenario in which this downfall feeds the red sequence. Instead, we propose here to quantify the “downfall rate” of this slow process, and compare it to the fast quenching rate associated with the growth of the red sequence.

As shown, e.g., in Muzzin et al. (2013) and Tomczak et al. (2014), the stellar mass density of UVJ quiescent galaxies increases monotonously with time, illustrating the progressive buildup of the red sequence. The time derivative of this quantity, neglecting stellar mass loss and residual star formation, is a measure of the quenching rate of galaxies (see, e.g., Peng et al. 2010). Here, we make the hypothesis that all the UVJ quiescent galaxies were quenched by a *fast* process, and set

$$\rho_{\text{quench}} = \frac{d\rho_*^Q}{dt}, \quad (5.12)$$

where ρ_*^Q is the stellar mass density of UVJ quiescent galaxies. We parametrize this latter quantity by fitting the redshift evolution reported in the CANDELS fields by Tomczak et al. (2014), accounting for the different choice of IMF:

$$\rho_*^Q \left[\text{M}_\odot / \text{Mpc}^3 \right] = (2.6 \pm 0.7) \times 10^8 \exp(-z). \quad (5.13)$$

To estimate the downfall rate associated to the *slow* process that lowers the SFE of massive star-forming galaxies, we compute the difference between the *observed* SFR density (ρ_{SFR}) and the density that *would be observed* if there was no drop of SFE, therefore if the Main Sequence had a slope of unity at all stellar masses ($\rho_{\text{SFR}}^{\text{unity}}$). This is a measure of the amount of star formation that was lost because of the reduced SFE within the Main Sequence. We estimate both SFR densities using the stellar mass functions of star-forming galaxies introduced in S15 (that we complement toward $z = 0$ using the mass function from Baldry et al. 2012), and integrate these mass functions weighted by the SFR. For the observed ρ_{SFR} , we use the SFR– M_* relation given in S15. Defining $r \equiv \log_{10}(1 + z)$ and $m \equiv \log_{10}(M_*/10^9 \text{ M}_\odot)$, this relation reads

$$\begin{aligned} \log_{10}(\text{SFR}_{\text{MS}} [\text{M}_\odot / \text{yr}]) &= m - m_0 + a_0 r \\ &\quad - a_1 [\max(0, m - m_1 - a_2 r)]^2, \end{aligned} \quad (5.14)$$

with $m_0 = 0.5 \pm 0.07$, $a_0 = 1.5 \pm 0.15$, $a_1 = 0.3 \pm 0.08$, $m_1 = 0.36 \pm 0.3$ and $a_2 = 2.5 \pm 0.6$. For $\rho_{\text{SFR}}^{\text{unity}}$ we use this same equation excluding the last term (which is used to describe the bending), i.e.:

$$\log_{10}(\text{SFR}_{\text{MS}}^{\text{unity}} [\text{M}_\odot / \text{yr}]) = m - m_0 + a_0 r. \quad (5.15)$$

Since these equations were not calibrated at $z < 0.5$ in S15, we use the observed Main Sequence from the HRS galaxies for these redshifts.

The downfall rate is then defined simply as

$$\rho_{\text{downfall}} = \rho_{\text{SFR}}^{\text{unity}} - \rho_{\text{SFR}}. \quad (5.16)$$

The resulting evolution of both ρ_{quench} and ρ_{downfall} is shown in Fig. 5.12. One can see from this figure that the fast quenching mode clearly dominates at all $z > 1.5$, while the slow downfall rapidly catches up to reach similar rates from $z = 1.5$ to the present day, i.e., over $\sim 70\%$ of the history of the Universe.

Two conclusions can be drawn from this observation. First, the fact that both the quenching and downfall rates reach similar values at all $z < 1.5$ implies that the downfall is a quantitatively important effect that should be considered alongside the growth of the red sequence. Second, it is clear that the two modes act at different epochs in the history of the Universe. While the fast quenching appears to hold a steady rate all the way from $z = 4$ to the present day, the slow downfall becomes a significant source of SF suppression only at $z < 2$. This suggests that the buildup of

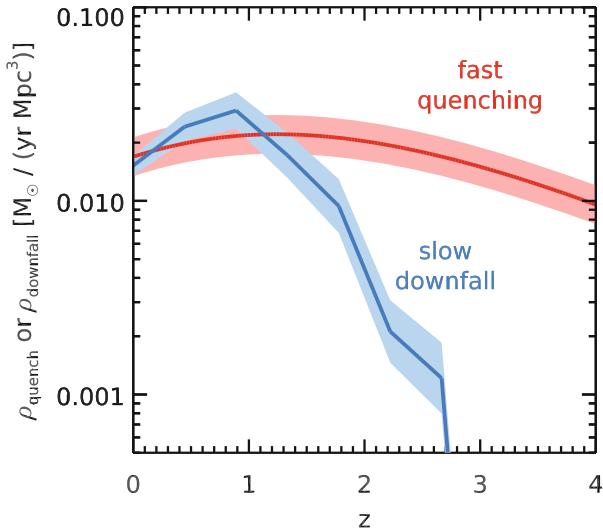


Fig. 5.12 Evolution of the mass-weighted quenching and downfall rate densities with redshift. The *red* curve shows the time derivative of the stellar mass density of *UVJ* quiescent galaxies, which we assume are produced by a “fast quenching” mechanism. The *blue* curve shows the star formation density that is lost because of the lowered SFE in massive galaxies, which we call the “slow downfall” rate. The *shaded regions* in the background give the uncertainty on both measurements (color figure online)

the red sequence and the change of slope of the Main Sequence are in fact related to two separate physical processes. This is discussed further in the next section.

5.6.2 Identifying the Actors that Regulate the SFE and the Gas Content

The idea that the specific star formation rate of galaxies is universal when computed over the disk rather than total mass of galaxies (as proposed by Abramson et al. 2014) is natural since bulges do not form stars. Yet, it would contradict another concept linked to the Main Sequence, namely the fact that galaxies are fed by the infall of extragalactic matter, which is in turn proportional to the total mass of galaxies, including dark matter (e.g., Dekel et al. 2013). Hence the fact that our results from Sect. 5.5.1 refute this mechanism may not be surprising, and possibly even expected when accounting for the large-scale context of infall. This echoes the result obtained more recently in the SDSS by Guo et al. (2015), who also found a sublinear slope for $z = 0$ pure disk galaxies, in conflict with the result of Abramson et al. (2014).

We observe instead in Sect. 5.5.2 that the star formation efficiency is decreasing in massive galaxies, leading to a slow downfall of star formation. This suggests

the existence of an *active* process that impacts the star formation activity, although the question remains to figure out exactly what this process could be. We cannot definitely address this question with the present data alone, but we review in the following the known mechanisms in light of our results.

We may already state that feedback from supernovae is not the favored solution, for it would affect more efficiently galaxies with a low gravitational potential, and therefore with low stellar masses, oppositely to our finding. Interestingly, the range in redshift and galaxy mass where the Main Sequence flattens corresponds to the regime where theory predicts group formation to be most effective, hence suggesting that structure formation or the membership to massive haloes may affect the rate of gas infall and the energetics regulating star formation (disk rotation and turbulence, see, e.g., Hennebelle and Chabrier 2008). Gravitational heating (Birnboim and Dekel 2003; Dekel and Birnboim 2008), i.e., the injection of energy into the dark matter halo from gas accretion itself, only depends on the mass of this halo, and can therefore act also in isolated galaxies. According to Dekel and Birnboim (2008), this can completely stop star formation in halos more massive than $\sim 6 \times 10^{12} M_{\odot}$ (corresponding to a stellar mass of $\sim 2 \times 10^{11} M_{\odot}$, Behroozi et al. 2013). This halo mass is the threshold above which natural cooling cannot counterbalance the energy brought into the halo by accretion, but in fact this energy is always there, even below this mass threshold, and can affect less massive halos more moderately. Lastly, we cannot rule out the action of the “radio-mode” AGN feedback, where jets heat the gas in the surroundings of galaxies, that may also be more common in massive galaxies.

Over the last years, the emphasis was put mostly on violent quenching mechanisms to explain the low baryonic fraction per unit dark matter halo mass, switching off the growth of galaxies by supernovae and AGNs at low and high masses, respectively (see, e.g., Silk and Mamon 2012; Behroozi et al. 2013; Behroozi and Silk 2015). We present here evidence that a slow downfall of the star formation efficiency should also be considered as a key mechanism.

5.7 Conclusions

We addressed here the origin of the change of slope of the Main Sequence of star-forming galaxies at $z < 1.5$, where high-mass galaxies exhibit a lower $sSFR \equiv SFR/M_*$ than what one would extrapolate from low-mass galaxies (e.g., Whitaker et al. 2012, 2014; Magnelli et al. 2014; Ilbert et al. 2015; Schreiber et al. 2015; Lee et al. 2015; Gavazzi et al. 2015).

It was reported in the Local Universe that the $SFR - M_{\text{disk}}$ relation is linear, suggesting that it is the bulge that creates most of the change of slope of the Main Sequence (Abramson et al. 2014). This claim was recently questioned by Guo et al. (2015) at $z = 0$, who reported that the slope of the $SFR - M_{\text{disk}}$ relation is in fact sublinear.

We performed the bulge-to-disk decomposition of a sample of $\sim 1\,000$ galaxies at $z = 1$ in the CANDELS fields, with robust SFRs measured from their mid- to far-IR photometry. We find that, as for the $SFR - M_*$ relation, the high mass slope of the

SFR– M_{disk} relation remains substantially shallower than unity. Such shallow slope is also observed among pure disk galaxies, selected either from their decomposed bulge-to-total ratio, or from their effective Sérsic index (see also Salmi et al. 2012 for a similar result at $z = 1$). This implies the existence of a physical mechanism at play even within the disks of massive galaxies, uncorrelated to the presence or absence of a bulge.

We then used *Herschel* stacking to derive jointly the average SFR and dust mass of star-forming galaxies in four bins of stellar mass in the same redshift range. Deriving the gas-phase metallicity from the Fundamental Metallicity Relation, we inferred the total gas mass, assuming that a fixed fraction of the metals are locked into dust, and analyzed the relation between the SFE $\equiv \text{SFR}/M_{\text{gas}}$ and the gas fraction in bins of stellar mass. We found that the most massive galaxies with $M_* > 2 \times 10^{11} M_\odot$ show a significantly reduced SFE by about a factor of 2 to 3 when compared to extrapolations from lower stellar masses, while the gas fraction remains constant. We measured gas masses in Local galaxies from the *Herschel* Reference Survey and found a similar behavior, reinforcing this finding. There, the drop of SFE happens at lower stellar masses, in agreement with the redshift evolution of the slope of the Main Sequence (see S15).

Combined together, these results point toward the existence of a *slow downfall* mechanism that impacts the SFE of massive star-forming galaxies. We showed that this phenomenon is quantitatively important at $z < 1.5$, and is likely disconnected from the fast quenching phenomenon that builds the red sequence. We argue that both mechanisms should be considered on the same footing when exploring the latest stages of galaxy evolution.

Leads for future research include studying the variation of the SFE above and below the Main Sequence, at fixed stellar mass. In this paper we show evidence that variations of SFR at high stellar masses are caused by variations of the SFE rather than gas mass. Since we have only been able to probe this through stacking and with relatively uncertain selection effects at $z = 1$, it would certainly be interesting to confirm these trends for individual objects. This kind of analysis can only be accomplished using a statistically complete sample of SFR and dust mass measurements at different stellar masses (ideally with direct metallicity estimates from emission lines). While SFRs and metallicities are currently within our reach, ALMA observations remain the only way to derive individual dust mass measurements for non-starbursting systems. A statistical sample with such measurement can be obtained either through dedicated pointed observations, or using a blind continuum survey, which will soon become possible with ALMA.

5.8 Appendix: Impact of the UVJ Selection

It has been shown that the properties of the SFR– M_* relation, i.e., its slope but also its scatter, are very sensitive to the sample selection (e.g., Speagle et al. 2014). In the present paper, we have used the standard UVJ color–color diagram to isolate

quiescent galaxies, and although this selection has been widely used in the recent literature (e.g., Whitaker et al. 2012, 2014; Muzzin et al. 2013; Bruce et al. 2014; Lang et al. 2014; Straatman et al. 2014; Pannella et al. 2015), its reliability can still be questioned. Indeed, while the quiescent and star-forming clouds can be easily identified on this diagram (see, e.g., Fig. 5.5), there is a non-negligible amount of galaxies in between, populating what is often referred to as the “green valley”. The dividing line defined by Williams et al. (2009) goes arbitrarily through this population, and it would be unwise to consider blindly that a “green valley” galaxy slightly above that line is quiescent, and that a similar galaxy slightly below the line is star-forming.

One way to circumvent this issue is not to apply any selection of star-forming galaxies in the first place, and identify the Main Sequence as the ridge (or mode) of the distribution of galaxies on the SFR– M_* plane. This was done, e.g., in Magnelli et al. (2014) and Renzini and Peng (2015). However this approach is only feasible in samples that are not SFR-selected. Building such a sample requires using SFRs that are not fully based on the FIR, and that are therefore potentially unreliable (one exception is the deep H α data of the SDSS, as in Renzini and Peng (2015) but translating this study to the distant Universe is currently out of our reach). Of course, this is also not applicable to stacking analyses, for which the SFR is only determined *a posteriori*.

Coming back to the *UVJ* selection, there are two ways our study could be affected by this arbitrary dividing line. On the one hand, the selection may be too strict, and we could actually discard from our sample some galaxies that are still forming stars at non-negligible rates, but have colors similar to that of quiescent galaxies because of peculiar combination of star formation history and dust content. On the other hand, the selection may be too loose, and our “star-forming” sample could actually contain a number of quiescent galaxies. We expect both effects to take place mostly for the most massive galaxies, where dust is more abundant and where most quiescent galaxies are found. The first alternative can be addressed by looking at Fig. 5.7, on which we show the position of both *UVJ* star-forming and *UVJ* quiescent galaxies in the SFR– M_* plane. One can see that there are indeed a few genuinely star-forming galaxies that are classified as *UVJ* quiescent. However, as can be seen from the running median, these galaxies tend to have systematically lower star formation rates compared to *UVJ* star-forming galaxies. Therefore, including these mistakenly identified galaxies in our sample would likely flatten the Main Sequence even more. Consequently, it is also unlikely that this would change dramatically the average SFE. The second alternative is probably more worrisome, as the drop of the SFE we observe in massive galaxies could be created by quiescent galaxies polluting our sample. One interesting observation to make out of Fig. 5.7 (and that can be made more quantitatively by studying the distribution of SFR around the median value, Ilbert et al. 2015; Schreiber et al. 2015) is that the mode of the SFR distribution at a given stellar mass (approximated here by the running median) coincides with the average value obtained from the stacked measurements. This means that, although our sample is SFR-selected, the amount of galaxies below our SFR detection limit is small enough that their impact on the average trend is marginal. In fact, for galaxies more massive than $5 \times 10^{10} M_\odot$, where the bending of the sequence is most

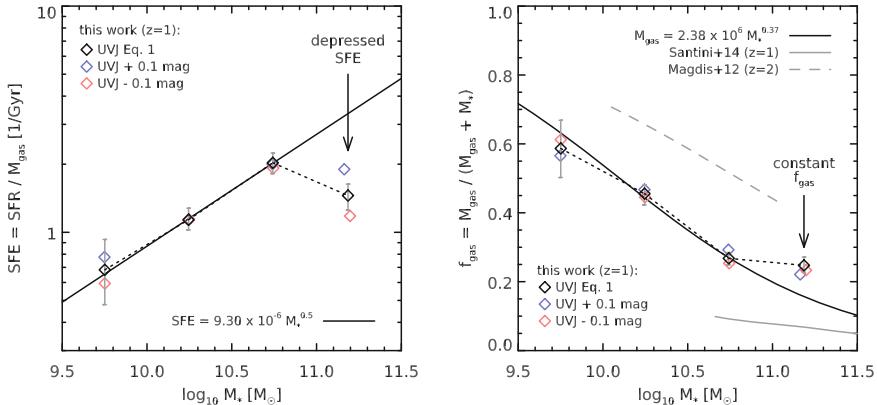


Fig. 5.13 Same as Fig. 5.9, but here black diamonds show the measured SFRs and M_{gas} of our chosen sample, while blue (respectively red) diamonds show how these values change if we shift the UVJ dividing line toward the star-forming (respectively quiescent) region by 0.1 magnitude (color figure online)

pronounced, 79 % of the UVJ star-forming galaxies are detected in the FIR. Therefore, the contamination of genuinely quiescent galaxies to the UVJ star-forming sample in this stellar mass range must be reasonably small (i.e., a maximum of 20 %).

Nevertheless, in an attempt to quantify how our results are influenced by the choice of the UVJ dividing line, we replicate our SFE measurements by stacking two different additional samples which are built by slightly shifting the UVJ dividing line by ± 0.1 magnitude. The resulting SFE and f_{gas} are shown in Fig. 5.13. As can be seen from this figure, moving the dividing line further into the quiescent cloud (red points) or further into the star-forming cloud (blue points) does not impact f_{gas} in any statistically significant way. In both cases, we still observe a drop of SFE, although the amplitude of this drop does vary, in this case mostly because of a change of SFR.

This can be put in perspective with the work of Arnouts et al. (2013), who found that the sSFR of a galaxy could be inferred from its position on the NrK diagram, which is conceptually similar to the UVJ diagram,⁸ with an sSFR that is continuously increasing as a function of the distance to the dividing line. According to Arnouts et al. (2013), using a stricter UVJ selection should bias our sample toward galaxies with a higher sSFR, hence, at fixed mass, with a higher SFR, which is what we observe for the most massive bin. In this context, the fact that the gas mass does not change substantially is particularly interesting, and is another hint that the mechanism responsible for the downfall, whatever it is, is mostly impacting the SFE, and not the gas supply.

⁸By using rest-frame wavelengths that are further apart, this diagram has a larger dynamic range and will separate quiescent and star-forming galaxies more clearly than the UVJ diagram. The downside is that measuring the rest-frame K band is particularly difficult at high redshifts, while the near-UV is hardly accessible at low redshift.

References

- R.G. Abraham, N.R. Tanvir, B.X. Santiago et al., MNRAS **279**, L47 (1996)
- L.E. Abramson, D.D. Kelson, A. Dressler et al., ApJ **785**, L36 (2014)
- C. Allende Prieto, D.L. Lambert, M. Asplund, ApJ **556**, L63 (2001)
- S. Arnouts, E. Le Floc'h, J. Chevallard et al., A&A **558**, 67 (2013)
- I.K. Baldry, S.P. Driver, J. Loveday et al., MNRAS **421**, 621 (2012)
- M. Barden, B. Häufner, C.Y. Peng, D.H. McIntosh, Y. Guo, MNRAS **422**, 449 (2012)
- M. Barden, H. Rix, R.S. Somerville et al., ApJ **635**, 959 (2005)
- P.S. Behroozi, J. Silk, ApJ **799**, 32 (2015)
- P.S. Behroozi, R.H. Wechsler, C. Conroy, ApJ **770**, 57 (2013)
- E.F. Bell, D.H. McIntosh, M. Barden et al., ApJ **600**, L11 (2004)
- E.F. Bell, D.H. McIntosh, N. Katz, M.D. Weinberg, ApJs **149**, 289 (2003)
- E. Bertin, S. Arnouts, A&As **117**, 393 (1996)
- M. Béthermin, E. Daddi, G. Magdis et al., A&A **573**, A113 (2015)
- M. Béthermin, H. Dole, A. Beelen, H. Aussel, A&A **512**, 78 (2010)
- Y. Birnboim, A. Dekel, MNRAS **345**, 349 (2003)
- A.F.L. Bluck, J.T. Mendel, S.L. Ellison et al., MNRAS **441**, 599 (2014)
- A. Boselli, L. Cortese, M. Boquien et al., A&A **564**, A66 (2014)
- A. Boselli, S. Eales, L. Cortese et al., PASP **122**, 261 (2010)
- R.J. Bouwens, G.D. Illingworth, P.A. Oesch et al., ApJ **754**, 83 (2012)
- R.G. Bower, A.J. Benson, R. Malbon et al., MNRAS **370**, 645 (2006)
- G.B. Brammer, P.G. van Dokkum, P. Coppi, ApJ **686**, 1503 (2008)
- J. Brinchmann, S. Charlot, S.D.M. White et al., MNRAS **351**, 1151 (2004)
- V.A. Bruce, J.S. Dunlop, M. Cirasuolo et al., MNRAS **427**, 1666 (2012)
- V.A. Bruce, J.S. Dunlop, R.J. McLure et al., MNRAS **444**, 1001 (2014)
- G. Bruzual, S. Charlot, MNRAS **344**, 1000 (2003)
- D. Calzetti, L. Armus, R.C. Bohlin et al., ApJ **533**, 682 (2000)
- G. Chabrier, PASP **115**, 763 (2003)
- R. Chary, D. Elbaz, ApJ **556**, 562 (2001)
- L. Ciesla, M. Boquien, A. Boselli et al., A&A **565**, A128 (2014)
- L. Ciesla, A. Boselli, M.W.L. Smith et al., A&A **543**, A161 (2012)
- C.J. Conselice, ApJs **147**, 1 (2003)
- L. Cortese, J. Fritz, S. Bianchi et al., MNRAS **440**, 942 (2014)
- D.J. Croton, V. Springel, S.D.M. White et al., MNRAS **365**, 11 (2006)
- E. Daddi, F. Bournaud, F. Walter et al., ApJ **713**, 686 (2010)
- E. Daddi, A. Cimatti, A. Renzini et al., ApJ **617**, 746 (2004)
- E. Daddi, H. Dannerbauer, D. Stern et al., ApJ **694**, 1517 (2009)
- E. Daddi, M. Dickinson, G. Morrison et al., ApJ **670**, 156 (2007)
- T. Dahlen, B. Mobasher, S.M. Faber et al., ApJ **775**, 93 (2013)
- R.S. de Jong, A&A **313**, 377 (1996)
- A. Dekel, Y. Birnboim, MNRAS **383**, 119 (2008)
- A. Dekel, A. Zolotov, D. Tweed et al., MNRAS **435**, 999 (2013)
- J.L. Donley, A.M. Koekemoer, M. Brusa et al., ApJ **748**, 142 (2012)
- B.T. Draine, A. Li, ApJ **657**, 810 (2007)
- D. Elbaz, E. Daddi, D. Le Borgne et al., A&A **468**, 33 (2007)
- D. Elbaz, M. Dickinson, H.S. Hwang et al., A&A **533**, 119 (2011)
- H.C. Ferguson, M. Dickinson, M. Giavalisco et al., ApJ **600**, L107 (2004)
- J. Franco, D.P. Cox, PASP **98**, 1076 (1986)
- M. Fumagalli, I. Labb  , S.G. Patel et al., ApJ **796**, 35 (2014)
- J.M. Gabor, R. Dav  , MNRAS **427**, 1816 (2012)
- A. Galametz, A. Grazian, A. Fontana et al., ApJs **206**, 10 (2013)
- F. Galliano, S. Hony, J. Bernard et al., A&A **536**, A88 (2011)

- G. Gavazzi, G. Consolandi, E. Viscardi et al., *A&A* **576**, A16 (2015)
V. González, R. Bouwens, G. Illingworth et al., *ApJ* **781**, 34 (2014)
N. Grevesse, A.J. Sauval, *Space Sci. Rev.* **85**, 161 (1998)
N.A. Grogin, D.D. Kocevski, S.M. Faber et al., *ApJs* **197**, 35 (2011)
K. Guo, X.Z. Zheng, H. Fu, *ApJ* **778**, 23 (2013)
Y. Guo, H.C. Ferguson, E.F. Bell et al., *ApJ* **800**, 39 (2015)
B. Häussler, D.H. McIntosh, M. Barden et al., *ApJs* **172**, 615 (2007)
S. Heinis, V. Buat, M. Béthermin et al., *MNRAS* **429**, 1113 (2013)
P. Hennebelle, G. Chabrier, *ApJ* **684**, 395 (2008)
T.M. Hughes, L. Cortese, A. Boselli, G. Gavazzi, J.I. Davies, *A&A* **550**, A115 (2013)
O. Ilbert, S. Arnouts, E. Le Floc'h et al., *A&A* **579**, A2 (2015)
A. Karim, E. Schinnerer, A. Martínez-Sansigre et al., *ApJ* **730**, 61 (2011)
R.C. Kennicutt Jr., *ARA&A* **36**, 189 (1998)
L.J. Kewley, M.A. Dopita, *ApJs* **142**, 35 (2002)
L.J. Kewley, S.L. Ellison, *ApJ* **681**, 1183 (2008)
A.M. Koekemoer, S.M. Faber, H.C. Ferguson et al., *ApJs* **197**, 36 (2011)
M. Kriek, P.G. van Dokkum, I. Labb   et al., *ApJ* **700**, 221 (2009)
P. Lang, S. Wuyts, R.S. Somerville et al., *ApJ* **788**, 11 (2014)
N. Lee, D.B. Sanders, C.M. Casey et al., *ApJ* **801**, 80 (2015)
J. Leja, P.G. van Dokkum, M. Franx, K.E. Whitaker, *ApJ* **798**, 115 (2015)
A.K. Leroy, A. Bolatto, K. Gordon et al., *ApJ* **737**, 12 (2011)
J.M. Lotz, J. Primack, P. Madau, *AJ* **128**, 163 (2004)
G.E. Magdis, E. Daddi, M. B  thermin et al., *ApJ* **760**, 6 (2012)
G.E. Magdis, E. Daddi, D. Elbaz et al., *ApJ* **740**, L15 (2011)
G.E. Magdis, D. Elbaz, E. Daddi et al., *ApJ* **714**, 1740 (2010)
B. Magnelli, D. Elbaz, R.R. Chary et al., *A&A* **496**, 57 (2009)
B. Magnelli, D. Lutz, A. Saintonge et al., *A&A* **561**, 86 (2014)
B. Magnelli, D. Lutz, P. Santini et al., *A&A* **539**, A155 (2012a)
B. Magnelli, P. Popesso, S. Berta et al., *A&A* **553**, 132 (2013)
B. Magnelli, A. Saintonge, D. Lutz et al., *A&A* **548**, A22 (2012b)
C. Mancini, A. Renzini, E. Daddi et al., *MNRAS* **450**, 763 (2015)
F. Mannucci, G. Cresci, R. Maiolino, A. Marconi, A. Gnerucci, *MNRAS* **408**, 2115 (2010)
M. Martig, F. Bournaud, R. Teyssier, A. Dekel, *ApJ* **707**, 250 (2009)
D.H. McIntosh, E.F. Bell, H. Rix et al., *ApJ* **632**, 191 (2005)
A. Muzzin, D. Marchesini, M. Stefanon et al., *ApJ* **777**, 18 (2013)
A.B. Newman, R.S. Ellis, K. Bundy, T. Treu, *ApJ* **746**, 162 (2012)
K.G. Noeske, B.J. Weiner, S.M. Faber et al., *ApJ* **660**, L43 (2007)
S. Noll, D. Burgarella, E. Giovannoli et al., *A&A* **507**, 1793 (2009)
M. Pannella, C.L. Carilli, E. Daddi et al., *ApJ* **698**, L116 (2009a)
M. Pannella, D. Elbaz, E. Daddi et al., *ApJ* **807**, 141 (2015)
M. Pannella, A. Gabasch, Y. Goranova et al., *ApJ* **701**, 787 (2009b)
M. Pannella, U. Hopp, R.P. Saglia et al., *ApJ* **639**, L1 (2006)
C.Y. Peng, L.C. Ho, C.D. Impey, H. Rix, *AJ* **124**, 266 (2002)
Y.-J. Peng, S.J. Lilly, K. Kova   et al., *ApJ* **721**, 193 (2010)
M. Pettini, B.E.J. Pagel, *MNRAS* **348**, L59 (2004)
S. Ravindranath, H.C. Ferguson, C. Conselice et al., *ApJ* **604**, L9 (2004)
A. R  my-Ruyer, S.C. Madden, F. Galliano et al., *A&A* **563**, A31 (2014)
A. Renzini, Y.-J. Peng, *ApJ* **801**, L29 (2015)
G. Rodighiero, E. Daddi, I. Baronchelli et al., *ApJ* **739**, L40 (2011)
F. Salmi, E. Daddi, D. Elbaz et al., *ApJ* **754**, L14 (2012)
B. Salmon, C. Papovich, S.L. Finkelstein et al., *ApJ* **799**, 183 (2015)
E.E. Salpeter, *ApJ* **121**, 161 (1955)
P. Santini, H.C. Ferguson, A. Fontana et al., *ApJ* **801**, 97 (2015)

- P. Santini, R. Maiolino, B. Magnelli et al., A&A **562**, 30 (2014)
M.T. Sargent, E. Daddi, M. Béthermin et al., ApJ **793**, 19 (2014)
C. Schreiber, D. Elbaz, M. Pannella et al., A&A **589**, A35 (2016)
C. Schreiber, M. Pannella, D. Elbaz et al., A&A **575**, A74 (2015)
N. Scoville, H. Aussel, K. Sheth et al., ApJ **783**, 84 (2014)
J. Silk, G.A. Mamon, Res. Astron. Astrophys. **12**, 917 (2012)
J. Silk, M.J. Rees, A&A **331**, L1 (1998)
L. Simard, D.C. Koo, S.M. Faber et al., ApJ **519**, 563 (1999)
L. Simard, C.N.A. Willmer, N.P. Vogt et al., ApJs **142**, 1 (2002)
R.E. Skelton, K.E. Whitaker, I.G. Momcheva et al., ApJs **214**, 24 (2014)
J.S. Speagle, C.L. Steinhardt, P.L. Capak, J.D. Silverman, ApJs **214**, 15 (2014)
D.P. Stark, R.S. Ellis, A. Bunker et al., ApJ **697**, 1493 (2009)
D.P. Stark, M.A. Schenker, R. Ellis et al., ApJ **763**, 129 (2013)
C.L. Steinhardt, J.S. Speagle, P. Capak et al., ApJ **791**, L25 (2014)
A. Stockton, E. McGrath, G. Canalizo, M. Iye, T. Maihara, ApJ **672**, 146 (2008)
C.M.S. Straatman, I. Labb , L.R. Spitler et al., ApJ **783**, L14 (2014)
A.W. Strong, J.B.G.M. Bloemen, T.M. Dame et al., A&A **207**, 1 (1988)
A.R. Tomczak, R.F. Quadri, K.H. Tran et al., ApJ **783**, 85 (2014)
A. van der Wel, E.F. Bell, B. H ussler et al., ApJs **203**, 24 (2012)
K.E. Whitaker, M. Franx, J. Leja et al., ApJ **795**, 104 (2014)
K.E. Whitaker, P.G. van Dokkum, G. Brammer, M. Franx, ApJ **754**, L29 (2012)
R.J. Williams, R.F. Quadri, M. Franx, P. van Dokkum, I. Labb , ApJ **691**, 1879 (2009)
S. Wuyts, N.M. F rster Schreiber, A. van der Wel et al., ApJ **742**, 96 (2011)
H.J. Zahid, L.J. Kewley, F. Bresolin, ApJ **730**, 137 (2011)
V. Zubko, E. Dwek, R.G. Arendt, ApJs **152**, 211 (2004)

Chapter 6

Reaching the Distant Universe with ALMA

6.1 Introduction

In my first published paper (Schreiber et al. 2015, see also Chap. 2) we were able to measure FIR-based star formation rates for a large sample of galaxies, thanks to the deep *Herschel* surveys that were observed during the lifetime of the satellite. This allowed us to put new constraints on the properties (and existence) of the Main Sequence of star-forming galaxies, from $z = 4$ to the present day. However, as can be seen from Fig. 2.13, most of our results at $z = 4$ are based on extrapolations of a *single* measurement, obtained by stacking the most massive galaxies ($M_* > 3 \times 10^{11} M_\odot$). In fact, we were able to probe only a tenth of the total SFR density at these epochs: having reached the limits of what *Herschel* alone can provide, learning more about the $z \geq 4$ Universe calls for more powerful tools.

To this end, we have proposed with Roger Leiton, Maurilio Pannella and David Elbaz (together with other collaborators¹ from the United States and Chile) an ALMA program to specifically target $z = 4$ galaxies that were too faint to be unambiguously detected by *Herschel*. Our goal with this proposal was to spend no more than 2 min of observing time on each galaxy to detect the dust continuum emission at an observed wavelength of $870 \mu\text{m}$ (band 7), i.e., targeting the rest-frame $170 \mu\text{m}$ which can be related to the star formation rate of the galaxy with an error of about 50 % (because the dust temperature is unknown and cannot be constrained with a single FIR photometric measurement). Even with this modest integration time, ALMA can detect $z = 4$ galaxies with SFRs of the order of $150 M_\odot/\text{yr}$, i.e., ten times lower than what *Herschel* was able to reach.

With such a data set, one could precisely constrain the normalization and the dispersion of the Main Sequence at $z = 4$, being affected neither by the uncertain clustering corrections that plague stacked *Herschel* measurements, nor by the poorly constrained dust-corrections that need to be applied to UV-based SFRs. Current

¹Neil Nagar, Mark Dickinson, Ezequiel Treister, Gustavo Orellana Gonzalez, Carolina Finlez, Sabrina Cales, Paula Calderon, and Yun-Kyeong Sheen.

cosmological models of gas accretion (e.g., Davé et al. 2011) predict that this normalization should rise continuously with redshift. On the other hand, observations of the $z > 4$ Universe tend instead to show a normalization that saturates at the $z \sim 4$ value (Stark et al. 2009; González et al. 2010; Bouwens et al. 2011), although recent studies Stark et al. (2013), Salmon et al. (2015) argue for a rise after correcting for the contamination of emission lines to the stellar mass estimates of $z > 5$ galaxies. Still, all these determinations are based on the UV light alone, using dust-correction recipes that were established in the Local Universe (Meurer 1999; Calzetti et al. 2000). Without a direct measurement of the dust content of these high-redshift galaxies, there is indeed no better choice available, but our ALMA survey could provide valuable constraints on the relation between the observed UV spectrum and the dust extinction.

Our proposal was accepted on April 9th 2014, and we received all the data on February 17th 2015. In this chapter, I describe the way our sample was built (Sect. 6.2), give an overview of the observations (Sect. 6.3.2), describe how I reduced the images (Sect. 6.4) and then measured the fluxes (Sect. 6.5), and finally introduce some preliminary results (Sect. 6.6), some of which were recently submitted for publication (Schreiber et al. 2016). I also report the detection of several other galaxies that happened to be in the field of view of our observations, but that were not part of our proposed sample. These are mostly massive galaxies at $z = 1$ to 3 (Sect. 6.7), except for two peculiar objects that are not detected in any *HST* image, and only show up in the *Spitzer* IRAC bands (Sect. 6.9). These two galaxies are among the brightest ALMA detections in our data, and a first determination of their redshift, based on the available photometry, would place them at $z > 4$.

6.2 Sample Selection

Our main sample consists of 113 galaxies found within the three CANDELS fields observable by ALMA, i.e., GOODS–South, UDS and COSMOS. These galaxies were selected from the catalogs introduced in Chap. 2 for having a photometric (or spectroscopic) redshift within $3.5 < z < 5.0$ and a stellar mass larger than $5 \times 10^{10} M_{\odot}$, as derived respectively by EAzY and FAST. We decided to observe all the galaxies satisfying these criteria, regardless of their *UVJ* classification (see Sect. 2.2.6), since at these redshifts the rest-frame *J* band falls in the IRAC channel 3, which is too shallow (at least in the UDS and COSMOS fields) to detect these distant objects. Furthermore, the number of massive quiescent galaxies in this sample is expected to be low: extrapolating the trend from $z < 4$ yields a quiescent fraction of about 20 % (see however Straatman et al. 2014, who report a substantially larger fraction of 35 %).

Using the stacked SED from Chap. 2 (Sect. 2.3.2), we predicted the $870 \mu\text{m}$ flux for each galaxy, assuming that they are all star-forming and that their SFR (and therefore, with a simple conversion factor, their L_{IR}) is following the Main Sequence from Eq. 2.9 with a scatter of 0.3 dex (see Sect. 2.4.4). The mean predicted flux was

~2 mJy, and we aimed for an RMS of 0.2 mJy so as to detect 80 % of the sample. To optimize the integration time, groups of objects that were close enough were combined into a single pointing.

Among the three CANDELS fields, GOODS–South was found to contain significantly fewer $z = 4$ galaxies. This could be explained either by cosmic variance, or from the fact that the GOODS–South photometry is of higher quality, and therefore that a significant fraction of the $z = 4$ galaxies in the other fields are spurious. Indeed, with this program we are pushing photometric redshifts to their limits, in a regime (massive and dusty distant galaxies) which is poorly sampled by spectroscopic campaigns to date. For this reason, we have proposed the spectroscopic follow-up of these $z = 4$ galaxies to confirm their redshift with the KMOS instrument on the VLT. This 20h program was recently approved, on July 2nd 2015. Coming back to the ALMA program, we decided to compensate for this lack of $z = 4$ massive galaxies in GOODS–South by adding 13 spectroscopically confirmed galaxies of lower stellar mass. The chances of detecting these objects in the continuum are slim, but these observations will nevertheless provide useful upper limits on their dust content. Also, by tuning the receptor frequency, we allow for the possibility to detect the [CII] emission line for some of them. These galaxies are in the “secondary sample”, and at the time of writing this thesis, their data have not yet been analyzed.

6.3 Description of the Observations and Data

6.3.1 Notes on Interferometric Imaging

Since ALMA is an interferometric facility, the angular resolution of an observation does not depend on the seeing, but on the configuration of the 36 antennae during the data acquisition: the most compact configuration (where the antennae are the closest to one another, i.e., at most 160 m apart) corresponds to the worst angular resolution of about 2'', while the most extended configuration (with a maximum distance of 10 km) allows reaching angular scales as small as 0.02'', i.e., less than a pixel of the best *HST* images (see, e.g., the impressive observations of a lensed galaxy in ALMA Partnership et al. 2015). This value is inversely proportional to the longest *baseline* (i.e., the longest distance between two antennae within the configuration), and this is due to the fact that an interferometer provides observations in the Fourier domain, which is also called the “ (u, v) plane” (see, e.g., Wilson et al. 2009). Instead of observing directly the intensity of the light in a pixel corresponding to a given position of the sky, like any conventional imager, the measured signal in each pair of antennae is combined and yields the light intensity of a given *scale* defined by the physical distance between the two antennae. To recover the emission at all scales in a given region of the sky, one has to cover the *whole* (u, v) plane, i.e., observe with all possible combinations of antennae distances and relative positions. Of course, this is infeasible, as it would necessitate either an infinite number of antennae or an infinite

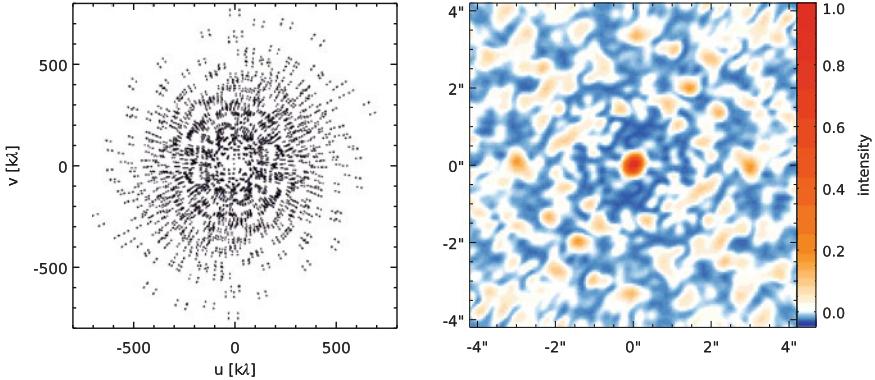


Fig. 6.1 *Left* Coverage in the (u, v) plane of one of the galaxies we observed in the GOODS-South field. One can see that this plane is only sparsely sampled, and that due to the short integration time, we are missing some information about the object. Indeed, using longer integration times would have increased the (u, v) coverage without moving the antennae thanks to the rotation of the Earth around itself. *Right* Beam (or PSF) of the corresponding image. This beam shows a number of negative features (blue) and secondary lobes (orange), which are the result of the sparse sampling of the (u, v) plane. These features make it hard to interpret a raw imaging of the (u, v) visibilities. It is possible, to some extent, to *deconvolve* this so-called “dirty beam” from the raw image, for example using the CLEAN algorithm, to improve the visual quality of the image and allow direct flux measurements (Color figure online)

integration time with varying antenna configurations. Therefore, any interferometric observation only covers a fraction of the (u, v) plane (see, e.g., Fig. 6.1, left), and consequently not all the angular scales of the target object are properly recovered. As written above, the longest baseline defines the angular resolution of the observations, but in fact this sparse sampling also implies the existence of a *largest recoverable scale*, which is defined instead by the shortest baseline. For example, in the most extended configuration (in which the shortest distance between two antennae is 250 m) one cannot recover structures that are more than $0.4''$ wide (versus $8''$ in the most compact configuration). This means that if one observes an extended object with a too sharp resolution, part of the flux will be irretrievably lost, or “resolved out”, and the total intensity of the source will be underestimated.

Even if the flux is not resolved out, using a too high angular resolution at fixed integration time can have an impact on the final signal to noise ratio (S/N) of the measurement. Indeed, if the galaxy is spatially resolved, the signal will be split among the various observable scales. Consequently, compared to a coarser resolution observation, a smaller fraction of the data is used to constrain the emission of the galaxy on its main angular scale, and the signal to noise ratio of the total flux will be lower. This is why, when preparing an ALMA proposal, one requests a given sensitivity *at a given angular scale*. If the observations have been obtained with a better angular resolution than that requested (which happens regularly due to the observatory’s

practical constraints²), achieving the requested S/N will require longer integration time, and one can later apply some filtering on the received data to remove (or down-weight) the high resolution measurements. This last step is called “tapering”, and is conceptually equivalent to picking the right aperture size to measure a flux, avoiding too large apertures that would add more noise than signal.

As can be seen from Fig. 6.1 (right), the point spread function (PSF) reconstructed from a limited (u, v) coverage is complex, and contains both positive and negative features. In fact, the integral of this PSF is zero, reflecting the fact that interferometric images have no background. This PSF is called the “dirty beam”. The most common technique to get rid of these sidelobes is the CLEAN algorithm (Högbom 1974). Briefly, the algorithm locates the brightest peak in the image and subtracts the dirty beam, rescaled to a fraction of the measured peak flux (e.g., 10 %). The procedure is repeated until the brightest peak falls below a given flux threshold (e.g., half the requested RMS). The resulting image is called the “residual map”, and is dominated by noise. Then, each subtracted peak is added back to the map, replacing the dirty beam by a Gaussian ellipsoid which is usually fit to the core of the dirty beam, so as to preserve the angular resolution. This ellipsoid is called the “restoring beam”, and the resulting image is called the “clean image”. This procedure has been used extensively in radio astronomy, and is also the standard imaging technique for ALMA.

One last point I will describe here is the existence of the *primary beam*, which is essentially the interferometric equivalent of the field of view of a conventional telescope. This primary beam is independent of the antenna configuration, and is in fact the PSF of each individual antenna. As such, it is a roughly Gaussian profile whose width scales with observed wavelength. In our case, the FWHM is about $18''$. This means in particular that the output level of a source of fixed flux is not uniform within a given pointing: it is maximal in the center, and drops by a factor of ten for sky positions that are $15''$ away from it. In a typical data reduction run, images are produced as seen by the antennae, i.e., the noise level is constant across the whole image (which is good to find the detections), but the absolute flux of each pixel is attenuated by the primary beam (which is bad for flux measurements). Therefore, one has to apply a “primary-beam correction” to the image (i.e., simply divide the raw image by the profile of the primary beam): the pixel values are then directly measuring the intrinsic flux, but the noise level becomes highly non-uniform.

6.3.2 General Properties of Our Data

Our only aim was to measure the integrated flux of each galaxy, and therefore we only needed a resolution good enough that the emission can be unambiguously attributed

²Moving a 12 m antenna across several kilometers takes some amount of time, as the transporters can only reach speeds of 12 km/h. For this reason, it is sometimes more efficient to schedule a low-resolution observation in an extended configuration and spend a bit more time on-source, rather than changing the configuration.

to each target. The high resolution *HST* images show that the typical angular size of these objects in the rest-frame UV is about $0.3''$, although some are substantially more extended. It is not obvious that the FIR sizes are in any way correlated to the UV sizes, as both usually come from totally different regions of the galaxy. Still, the measured UV sizes can be used as lower limits, and to be safe we requested that the target sensitivity be achieved on a spatial scale of $0.7''$, i.e., using almost the most compact configuration.

Our proposed sample was observed for a total of 5.9 h of telescope time, including overheads, for an average observing time of 1 min and 19 s per object. All our targets were observed, but the quality of the data varies from one field to another. I will therefore describe each field separately, and provide a summary in Table 6.1.

In the following, the angular resolution is determined by fitting a Gaussian profile to the core of the dirty beam (where the amplitude is at least half that of the peak) using GALFIT (Peng et al. 2002). A comparison of these beams in the three different fields is shown in Fig. 6.2. The RMS is measured individually for each pointing using the median absolute deviation of the non-primary-beam-corrected image.

GOODS-South

This field was observed in July 2014, with a total telescope time of 1 h and 10 min, and an on-source integration time of 1 min and 15 s per target. All targets were observed in one scheduling block, and were visited twice: first with a 50 s pointing, and about 40 min later for an additional 25 s. It turned out that the array was in a relatively extended configuration when the observations were undertaken, so that the achieved resolution is $0.36'' \times 0.31''$ with a uniform RMS of 0.13 mJy/beam. Although the noise is about twice lower than the one we requested, most galaxies will be resolved at that resolution. To optimize the detection rate, I therefore imaged the data using a tapering of $0.7''$ so as to recover the requested angular resolution. By testing different values, I found that $0.7''$ was providing the best compromise between signal-to-noise ratio and peak flux. After this step, the RMS increases to

Table 6.1 Summary of the observations in the three fields

Field and data ^a	Targets ^b (main)	Angular resolution	RMS (mJy/beam)	Detections	
				main ^c	others ^d
GS (high res.)	27	$0.36'' \times 0.31''$	0.13	6	
GS (tapered)	27	$0.69'' \times 0.69''$	0.22	9	7
UDS (SB1)	42	$1.48'' \times 0.69''$	0.23	9	4
UDS (SB2, high res.)	42	$0.43'' \times 0.34''$	0.25	2	
UDS (SB2, tapered)	42	$0.59'' \times 0.53''$	0.26	4	
UDS (merged+tapered)	42	$0.92'' \times 0.73''$	0.21	12	4
COSMOS	44	$1.14'' \times 0.70''$	0.15	13	8

^aThe data sets we use for science are highlighted in boldface

^bNumber of targets that belong to the primary sample

^cNumber of targets from the primary sample that are detected at more than 3σ significance

^dNumber of additional galaxies detected in the field of view at more than 5σ significance

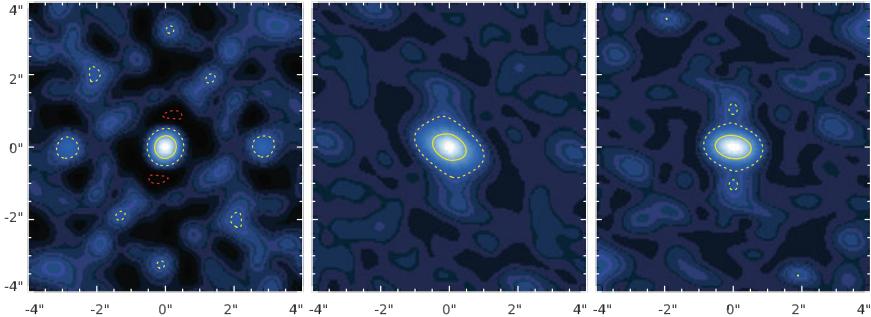


Fig. 6.2 Zoom on the core of the dirty beams of GOODS–South (left, after tapering), UDS (middle) and COSMOS (right). All three images are shown with the same intensity scale, and correspond to a $8'' \times 8''$ area. The contours indicates the region where the amplitude of the beam is larger than half of the peak value (yellow solid line) and a tenth (yellow dotted line), and negative by more than a tenth of the peak value (red dotted line). This diversity illustrates how the shape of the beam can vary depending on the coverage of the (u, v) plane: the GOODS–South beam has the strongest sidelobes and negative features because of the tapering, while UDS has the weakest variations thanks to the fact that the field was observed twice and at different elevations (Color figure online)

0.22 mJy/beam (min: 0.17 mJy, max: 0.31 mJy), and three more objects are detected (a detailed description of the detections is provided below).

The calibration was based on the quasars J0334-401, J0348-2749 and J0522-3627.

UDS

The UDS was observed twice: 1 h and 23 min in June 2014, and 1 h and 20 min in December 2014, for a total of 2 min and 31 s of on source integration per target. Both runs have used the same observing strategy as for GOODS–South, i.e., visiting each target twice in a single scheduling block.

The first observing run was taken in an extended configuration, with a resolution of $0.43'' \times 0.34''$, and an RMS of 0.25mJy/beam. This angular resolution is similar to that of GOODS–South, but the achieved RMS is substantially higher. It is likely that the observers on site judged that this run was not of good enough, and decided to schedule another one later.

The second observing run was acquired in a compact configuration, so that the angular resolution is substantially coarser and reaches $1.48'' \times 0.69''$, for an RMS of 0.23 mJy/beam (min: 0.20 mJy, max: 0.25 mJy).

On its own, this last run satisfies our requested criteria. However, to optimize the sensitivity, in the following I combine the data from both runs (this step is described in more detail in the next section). The resulting resolution is $0.58'' \times 0.43''$, i.e., still substantially smaller than that requested. Also, the core of the beam is not Gaussian and these numbers do not really make justice to its real profile. I therefore taper the scales below $0.5''$ to recover a more uniform beam that is $0.92'' \times 0.73''$, i.e., similar to that of the second observing run but with an axis ratio closer to unity. The resulting RMS is 0.21 mJy/beam (min: 0.19 mJy, max: 0.23 mJy). This additional data allows the detection of several more objects above the 3σ level.

The calibration for the first run was based on the quasars J0215-0222, J0241-0815 and J2258-279, while that for the second run was based on the quasars J0108+0135 and J0217+0144, as well as Neptune.

COSMOS

The COSMOS field was observed for 1 h and 45 min, with 1 min and 25 s of on-source integration. Here the sources were observed for 43 s twice, each time within a separate scheduling block. The angular resolution reaches $1.14'' \times 0.70''$ for a uniform RMS of 0.15 mJy/beam (min: 0.14 mJy, max: 0.16 mJy), i.e., 25 % smaller than requested.

For the first scheduling block, the calibration was based on the quasars J1010-0200, J1058+0133 and J1058+015. For the second scheduling block, the last quasar was replaced by Callisto.

6.4 Data Reduction

ALMA data are received pre-reduced by an astronomer on site, which is very convenient to get a quick overview of the achieved sensitivity and detection rate. However, the quality of these pre-reductions turned out to be mixed, from ok to quite poor. In the following sections I describe the modifications I made to the pre-reduction in each field.

GOODS-South

In *GOODS-South*, the person in charge of the reduction noticed that the resolution was higher than requested, and therefore used tapering on one of the targets to measure the effective RMS at the $0.7''$ resolution. However, he/she did not re-reduce the other targets, which were shipped at the highest resolution. Since the reduction script is also provided with the data, I could easily do it myself. The high-resolution images are not used in the following analysis.

UDS

The pre-reduced data in UDS were of good quality. However, as written in the previous section, the UDS field was observed twice at very different period of the year, resulting in two scheduling blocks (SBs) of about one hour each. The pre-reduction used data from the last SB only, disregarding all the data from the first SB. Therefore, I reduced by myself the data from this first SB, running the automatic pipeline. After several hours of computations, the pipeline delivered a calibrated “Measurement Set” (MS) that is ready to be used for imaging.

For testing purposes, I chose one of the bright source from the second SB and tried to image the corresponding pointing using data from the first SB only. The resulting image was of poor quality, with a very high RMS of about 1 mJy and pronounced striped patterns, so the source was hardly visible. Within `casa`, I used `plotms` to manually browse through the visibility data, looking for anomalous features. I noticed a small group of visibilities with amplitudes systematically higher by a factor of 200 compared to the rest of the data. I flagged these outliers and relaunched the imaging,

only to notice that the image quality did not improve much. Inspecting the visibilities again, I could not find any other striking feature.

Using another angle of approach, I tried imaging this same source only using data from a reduced frequency range, in particular imaging each of the four sidebands individually. I realized that the issue came from the second sideband, and that the other three were behaving properly. Disabling this faulty sideband, i.e., trashing 25 % of the data, the final imaging quality was good, and the RMS went down to 0.24 mJy/beam. I then tried to refine further the frequency selection within the faulty sideband to see if part of the data could be used, but it appears that the whole sideband is corrupted. This could be an issue of calibration, and should be investigated further.

I then merged the two scheduling blocks together to form a single measurement set. I first used the `split` procedure to discard the second sideband from the first scheduling block, then the procedure `concat` to create the new merged measurement set. Since the first data set was observed at relatively high angular resolution, I also used tapering to remove the scales below 0.5''. To do so, I used the same script as in *GOODS-South*, and imaged all the targets.

COSMOS

Finally, in COSMOS the provided reduction was extremely poor. There was a typo in the imaging script that made all the reduced images invalid (all pixels where “not a number”). The person in charge of the reduction did not notice this, and reported zero detections. On top of this, only one of the two scheduling blocks was used in the imaging process, leading to sub-optimal sensitivity. Therefore I merged the two scheduling blocks (using the casa procedure `concat`), fixed the typo in the script, and relaunched the imaging process.

6.5 Flux Measurement and Detection Rate

If a target is point-like, its flux can be read directly from the pixel value of the corresponding peak in the primary-beam-corrected clean image. This measured flux is called the “peak flux”. This image-based measurement is the simplest way to estimate the flux of an object, and it has the nice advantage that it allows blind detections of all the sources in the field of view, simply by locating all the pixels of the map that lie above a chosen confidence threshold. For this preliminary analysis, I used SExtractor (Bertin and Arnouts 1996) to identify all the pixels above a 3σ threshold, then cross matched the resulting “detection” list to our target list. Detections which were more than 1.5'' away from any target were rejected as being either likely spurious ($\text{SNR} < 5$) or real but unrelated to our targets ($\text{SNR} > 5$), while those between 0.5 and 1.5'' were kept only if their SNR was higher than 4. The resulting number of detections are summarized for each field in Table 6.1.

Overall, we detect 34 of our primary targets with this method, with a fairly low detection rate of 30 %. Considering only the *UVJ* star-forming galaxies, the detection rate is slightly higher and reaches 39 %, although 9 of our detections are classified as

UVJ quiescent galaxies. The measured peak fluxes are found to be on average about a factor of ~ 2.7 (median: 1.8) below our predictions.

The main drawback of using peak fluxes is that they will underestimate the total flux of extended objects. Analyzing the *HST* images, we found that the typical angular size of our targets was about $0.3''$. At the resolution achieved in our observations, these sources should be barely resolved, and the peak flux should be a good first approximation. On the other hand, we cannot exclude that the extent of these galaxies is larger in the FIR domain, because dust is usually not physically correlated with the UV bright regions that *HST* reveals. Some of our targets are also clearly extended in the *HST* images, and we expect the peak flux to fall short significantly for these objects. This could explain part of the disagreement with the prediction, and therefore we need a more refined flux measurement that takes into account this possible spatial extension.

With non-interferometric data, the simplest way to measure the flux of an extended structure is to use aperture photometry. If the aperture is large enough, the total flux of the object is recovered by summing up the pixel values within a chosen aperture (which is usually circular). If the aperture is too small (which is often the case if one wants to avoid excessive contamination from neighboring sources), one can extrapolate the total flux of the object by assuming some profile or simply using the growth curve of the PSF. With interferometric data, this is probably not be the best approach. As written in the introduction, the integral of the dirty beam is zero, and therefore the flux of any source measured within an infinite aperture will also be zero. Now, this is true only if the flux is measured on the dirty image. If the cleaned image is used, the result will depend on the details of how the image was cleaned: down to which flux threshold, with which technique, etc. For example, a typical threshold to stop cleaning is twice the image RMS, since choosing a lower threshold may prevent the cleaning algorithm from converging. All the flux that falls below this threshold will not be cleaned, and cannot be measured by aperture photometry. Hence, while it is possible to measure a flux within an aperture directly from the map, it is not trivial to determine the proper aperture correction that takes care of both the flux that falls out of the chosen aperture *and* the flux that was not cleaned.

To prevent these issues, the most reliable way to measure the flux of an extended object is to perform a profile fitting directly in the (u, v) plane, without using the reduced imaging at all. In particular, the resulting fluxes do not depend on the CLEAN algorithm, and tapering is not necessary since the model knows its own extents and handles the various measured scales correctly.³ The `casa` pipeline provides this functionality through the `uvmodelfit` procedure, which is essentially a crippled version of GALFIT that understands interferometric data. Only one profile can be fitted at a time, and a limited number of models are provided: a point source, elliptical Gaussian profiles, and elliptical disks. More evolved procedure have been published (e.g., `uvmultifit`, Martí-Vidal et al. 2014), but these were not considered for this

³To come back to the analogy with aperture photometry: this is equivalent to using a PSF-convolved model to measure the flux of an object. In this case, no aperture is necessary since the fit automatically weights pixels depending on the expected flux of the model.

preliminary analysis. One can refer to Martí-Vidal et al. (2014) for a comparison of existing alternatives, and their respective capabilities.

For each source detected with the first method, I used `casa` and `uvmodelfit` to fit a Gaussian profile to each detection, which is the most physically plausible model available (it should be noted that a Gaussian profile is equivalent to a Sérsic profile of index $n = 1/2$). I used as starting position the location derived from the peak flux, an angular size of $0.5''$, a total flux of 0.8 mJy , and an axis ratio of 1. Together with the position angle, all these parameters were left free to vary in the fit, and a total of six fitting iterations were performed for each source (the total flux is usually stable after the fourth or fifth iteration). The resulting fluxes are found to be larger than the peak fluxes in 90 % of the case, with a median increase of 37 % (31, 37 and 47 % in GS, UDS and COSMOS respectively) and a flux that is more than doubled for 8 galaxies. For three objects the integrated flux is actually smaller by 10 to 20 %, but these differences are comparable to the RMS of the image. This suggests that most our galaxies are substantially resolved in the FIR, even at $0.7''$ resolution: the median measured size is $0.3''$ (min: $0''$, max: $1.6''$), which is consistent with the size estimate from the *HST* images (although there is little correlation on a galaxy to galaxy basis).

The `uvmodelfit` procedure returns uncertainties on the derived parameters, in particular on the total flux measurement. These are found to be systematically larger than the RMS of the map by a factor of 2 on average, suggesting that they are not severely underestimated, but simulations have to be made to assess the reliability of these uncertainties. This can be done by creating visibility data for mock galaxies of varying shapes, e.g., with the `casa` simulator `simobserve`, and using `uvmodelfit` to recover the input flux. This is still work in progress.

Using these improved flux measurements, the tension with the prediction is reduced. The measured fluxes are still on average a factor of 1.9 lower, although the median factor is now consistent with unity (1.03).

To confirm the accuracy of these measurements, I used our target “0-23751” from the GOODS–South field. This galaxy has already been observed by ALMA within the ALESS program (PI: I. Smail, Cycle 0). It was targeted as one of the brightest sub-mm source in the Extended Chandra Deep Field South (ECDFS), which contains GOODS–South. The quality of the ALESS data for this object (“ALESS-13”) is poorer than our observations: the RMS is 0.42 mJy/beam , and the beam is much larger ($1.36'' \times 1.15''$). Still, this independent measurement can be used as a consistency check.

In Hodge et al. (2013), the total flux of this source as measured by ALMA is reported as $8.0 \pm 0.6\text{ mJy}$, and is consistent with the single-dish LABOCA flux of $8.8 \pm 1.2\text{ mJy}$. To measure this total flux, Hodge et al. (2013) used the `casa` procedure `imfit` which fits a galaxy profile on the reconstructed clean image. Here, using the (u, v) plane fitting, I measure $7.7 \pm 0.4\text{ mJy}$. The tension between these measurements is less than 1σ , and I therefore consider them as compatible.

I list all the measured fluxes for the galaxies in our main sample in Table 6.2, and for the other galaxies within the field of view in Table 6.3.

Table 6.2 Galaxies from our main samples that were detected with a significance higher than 3σ on the ALMA maps. The “name” column contains the ID of each object: the first number is the identifier of the field (0: GOODS–South, 1: UDS, 2: COSMOS), and the second number is the CANDELS ID

Name	z_{phot}	$\log_{10}(M_*)$	UVJ	RMS	Speak	$S_{\text{peak}}^{\text{err}}$	$S_{u,v}$	$S_{u,v}^{\text{err}}$	$R_{u,v}$	$R_{u,v}^{\text{err}}$
		M_{\odot}		mJy	mJy	mJy	mJy	mJy	”	”
0-3973	3.70	11.19	Q	0.22	0.73	0.22	0.77	0.24	0.00	1.00
0-4356	4.61	11.40	SF	0.20	2.69	0.20	2.50	0.25	0.11	0.06
0-4936	3.43	11.07	SF	0.20	0.60	0.20	1.32	0.46	0.63	0.24
0-5375	4.33	11.06	SF	0.18	0.92	0.18	1.37	0.27	0.11	0.12
0-5652	4.45	11.12	SF	0.22	0.84	0.22	1.11	0.34	0.23	0.14
0-6374	4.54	10.80	SF	0.25	0.82	0.25	2.62	1.06	1.57	0.54
0-12407	4.40	11.42	Q	0.28	2.21	0.28	2.02	0.36	0.32	0.08
0-13375	4.76	10.12	SF	0.21	0.78	0.21	1.25	0.62	0.52	0.28
0-16822	4.53	10.94	SF	0.30	2.26	0.30	1.85	0.33	0.27	0.09
0-23751	4.48	11.34	Q	0.25	6.51	0.27	7.72	0.39	0.30	0.02
1-2720	4.55	11.83	SF	0.21	2.64	0.21	3.39	0.34	0.42	0.10
1-4319	4.55	11.92	Q	0.22	1.47	0.22	3.29	0.55	0.70	0.17
1-6218	4.00	11.39	SF	0.21	1.57	0.21	1.86	0.27	0.00	9.08
1-11201	3.76	11.52	SF	0.21	0.85	0.21	1.74	0.48	0.59	0.21
1-13854	3.71	10.83	SF	0.23	0.96	0.23	1.19	0.37	0.32	0.29
1-16843	3.95	11.15	SF	0.22	2.05	0.22	2.89	0.40	0.39	0.13
1-16932	3.68	11.14	SF	0.22	1.33	0.22	1.69	0.32	0.17	0.15
1-23809	3.81	10.72	SF	0.20	0.62	0.20	0.95	0.37	0.32	0.36
1-25382	3.61	11.10	SF	0.21	0.96	0.21	1.65	0.47	0.57	0.27
1-30625	4.16	11.35	Q	0.20	1.52	0.20	2.16	0.32	0.29	0.13
1-35143	4.26	11.48	SF	0.19	0.62	0.19	0.79	0.29	0.22	0.34
1-35579	4.36	10.94	Q	0.19	0.81	0.20	1.25	0.36	0.57	0.28
2-3662	3.51	10.95	Q	0.15	0.73	0.15	1.14	0.27	0.51	0.33
2-14723	3.93	11.16	SF	0.15	4.02	0.15	4.72	0.23	0.24	0.09
2-15925	4.20	10.93	SF	0.15	0.85	0.15	1.13	0.26	0.34	0.22
2-16517	4.30	11.81	SF	0.16	1.03	0.16	1.64	0.28	0.59	0.20
2-16676	3.84	11.68	Q	0.15	1.14	0.15	1.52	0.25	0.23	0.22
2-17145	4.11	11.48	SF	0.15	0.99	0.15	1.32	0.26	0.31	0.20
2-19794	3.65	10.94	SF	0.15	0.68	0.15	1.93	0.41	1.27	0.33
2-20877	4.61	12.13	SF	0.15	2.17	0.15	3.19	0.29	0.63	0.13
2-27853	4.58	11.86	Q	0.15	5.14	0.15	10.95	0.36	0.74	0.04
2-31823	3.77	11.02	SF	0.15	0.47	0.15	1.72	0.47	1.52	0.48
2-33803	3.83	11.22	SF	0.16	3.13	0.16	3.55	0.25	0.23	0.08
2-34209	3.63	11.02	SF	0.15	0.80	0.15	0.98	0.23	0.00	1.00
2-38011	4.23	10.81	SF	0.15	0.68	0.15	1.59	0.38	0.78	0.24

Table 6.3 Other detections in the field of view of our observations. These galaxies are not part of the $z = 4$ sample. The “name” column gives either the CANDELS ID of the most likely counterpart (where the ID is defined as in Table 6.2), or a new identifier if no counterpart was found in the CANDELS catalog at less than 1'', in which case the source is considered as spurious unless otherwise specified

Name	d_{xmatch}	z_{phot}	$\log(M_*)$	RMS	S_{peak}	$S_{\text{peak}}^{\text{err}}$	$S_{u,v}$	$S_{u,v}^{\text{err}}$	$R_{u,v}$	$R_{u,v}^{\text{err}}$
			M_{\odot}	mJy	mJy	mJy	mJy	mJy	''	''
0-6811	0.30	0.55	10.30	0.25	5.25	0.91	12.46	4.51	1.40	0.42
AZ4-G2				0.27	1.84	0.37	5.27	1.71	1.33	0.35
0-24965	0.29	2.17	10.12	0.29	2.98	0.56	6.47	1.93	0.75	0.20
AZ4-G4				0.24	10.44	2.08	11.90	4.05	0.58	0.21
AZ4-G5				0.29	3.49	0.66	3.46	0.93	0.35	0.13
AZ4-G6				0.17	1.62	0.25	1.69	0.43	0.18	0.12
AZ4-G7				0.20	4.73	0.89	6.76	1.93	0.50	0.16
AZ4-U1				0.21	1.41	0.27	1.37	0.36	0.00	1.00
AZ4-U2				0.22	1.73	0.32	2.08	0.50	0.37	0.20
AZ4-U3				0.22	2.23	0.43	2.14	0.55	0.00	1.00
AZ4-U4				0.21	1.72	0.33	2.03	0.52	0.23	0.33
180443	0.34 ^a	1.52	10.14	0.14	1.34	0.17	1.98	0.28	0.33	0.24
2-16460	0.28	2.21	10.98	0.16	1.08	0.18	1.63	0.33	0.40	0.18
2-21040	0.24	2.81	11.65	0.15	6.15	0.22	7.27	0.34	0.24	0.10
AZ4-C4	b	3.05	11.85	0.15	1.87	0.35	2.46	0.56	0.09	0.53
AZ4-C5	c	> 4	~ 11	0.16	3.03	0.17	3.92	0.27	0.37	0.12
AZ4-C6	c	> 4	~ 11	0.15	2.98	0.16	3.24	0.25	0.00	1.00
2-1981	0.03	3.10	11.62	0.15	3.63	0.18	4.79	0.31	0.41	0.06
2-2213	0.21	1.95	11.03	0.15	1.41	0.21	1.86	0.36	0.32	0.18

^aThis source is associated to a galaxy that actually falls out of the HST-WFC3 coverage. It is detected in the shallower K_s -band catalog of Ilbert et al. (2013), and its ID corresponds to that given in this catalog

^bThis source has a clear counterpart in both the HST and *Spitzer* IRAC images. It was not included in the CANDELS catalog because it falls close to a bright star (see Sect. 6.8). Its stellar mass is probably overestimated owing to the presence of a dusty AGN

^cThese objects only have counterparts in the *Spitzer* IRAC images (see Sect. 6.9)

6.6 The $z = 4$ Main Sequence

6.6.1 Calibration of the SFR

In Chap. 3, I developed a new template library and gave prescriptions for the evolution of the dust temperature with cosmic time. By extrapolating the trend observed at lower redshifts, this prescription suggests that the average dust temperature at $z = 4$ should be around 36 K. It turns out that the best-fit Chary and Elbaz (2001) SED of the $z = 4$ stacked *Herschel* photometry from Schreiber et al. (2015) was substantially colder compared to this prescription, with a temperature of about 31 K (lower than that

measured at $z = 3$). Obviously, measuring a dust temperature at $z = 4$ with *Herschel* photometry only is daring, as the longest accessible wavelength, 500 μm , is actually at the peak of the SED, and no photometry constrains the Rayleigh–Jeans tail. In this situation, any uncertainty on the measured 500 μm flux translates into a larger uncertainty for the predicted 870 μm flux. For this stacked SED, a substantial part of this uncertainty comes from the clustering correction. The fact that the $z = 4$ best-fit SED is found to be colder than expected suggests that the prescription we adopted in Schreiber et al. (2015) was inadequate, and underestimated the amount of flux boosting caused by clustering. This uncertainty was in fact one of the main motivations for this ALMA survey.

Using the warmer SED (36 K), the stacked *Herschel* photometry yields an L_{IR} that is essentially identical (3 % difference), but a predicted 870 μm flux that is a factor of 20 % lower. Conversely, if this is really the average temperature of $z = 4$ galaxies, it means that each of our measured ALMA fluxes correspond to an L_{IR} that is 20 % higher than initially predicted.

Using the fluxes derived in the previous section, and the 36 K dust SED, I extrapolate the total L_{IR} from the observed 870 μm flux, disregarding any photometry in the *Herschel* bands for now. Using the prescription of Kennicutt (1998), I convert this L_{IR} into a star formation rate, neglecting the unobscured contribution from the UV. The stellar masses are taken directly from the catalogs introduced in Schreiber et al. (2015), and were computed by M. Pannella.

6.6.2 The SFR– M_* Relation

I show in Fig. 6.3 the resulting $z = 4$ Main Sequence, as seen by ALMA. Although this is still work in progress, we can already see at first glance that the galaxies in our sample follow roughly the relation we derived in Schreiber et al. (2015). However, before interpreting this observation any further, a number of concerns should be voiced.

First, we have in this sample some extremely massive galaxies, with $M_* \sim 6 \times 10^{11} \text{ M}_\odot$, and one (2-20877) that is as massive as 10^{12} M_\odot . Because these galaxies are faint and usually detected only in a handful of NIR broadbands, their redshift is not extremely well constrained, and there is generally a secondary peak of the redshift probability distribution around $z = 2$. If a massive and highly obscured $z = 2$ galaxy is mistakenly put at $z = 4$ because of this uncertainty, its stellar mass will reach unreasonable levels, as we observe here (see also Sect. 6.9). Therefore, there is a chance that some of our targets are actually at lower redshifts. For example, the most problematic case, 2-20877, has a $z_{\text{phot}} = 4.61$. In the 3DHST catalogs, this galaxy is placed at an even higher redshift, with $z_{\text{phot}} = 5.22$. However, the CANDELS redshift compilation instead gives a substantially closer solution with $z_{\text{phot}} = 2.86$, and a more reasonable stellar mass around $2 \times 10^{11} \text{ M}_\odot$. Therefore, instead of being more than a factor of two below the $z = 4$ Main Sequence, it would be on top of the

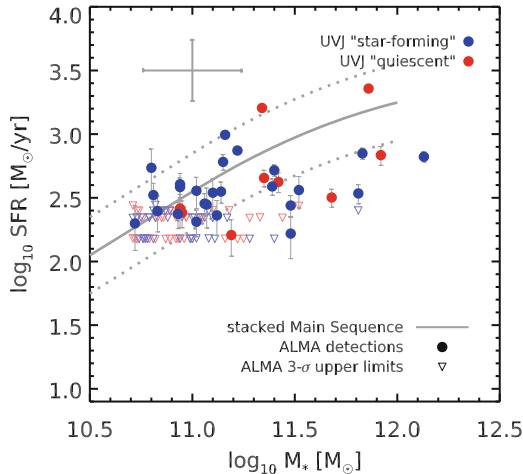


Fig. 6.3 The $z = 4$ Main Sequence, as seen by ALMA. SFRs are derived from the observed $870 \mu\text{m}$ continuum. Our detections are shown with *filled circles*, colored according to their respective *UVJ* classification (*blue* star-forming, *red* quiescent). For non-detections, I only show 3σ upper limits derived from the RMS of the map, scaled up by a factor of 1.6 to account for the typical spatial extent that is measured in the detections. The *gray line* in the background shows the $\text{SFR}-M_*$ relation from Schreiber et al. (2015), and the scatter measured at lower redshifts (0.3 dex) is shown with *dotted lines*. The large error bars in the top-left corner display the typical uncertainty on the stellar mass (*horizontal bar*) as estimated by comparing our masses to that of the 3DHST catalog for these same objects, and the SFR (*vertical bar*) assuming a T_{dust} scatter of 4 K (Color figure online)

$z = 2.8$ Main Sequence. The same argument may hold for the other massive galaxies that we find systematically below the Main Sequence.

The only way to unambiguously disentangle the two possibilities is to spectroscopically confirm the redshift. Usually, spectroscopic confirmation of $z \sim 4$ galaxies is obtained from the $\text{Ly}\alpha$ line ($\lambda_{\text{rest}} = 0.1216 \mu\text{m}$), which is redshifted into the optical domain, and is therefore “easily” accessible with reasonable integration times. However, being emitted in the rest-frame FUV, the intensity of this line is extremely sensitive to the presence of dust. For this reason, most spectroscopic detections at these redshifts consist of galaxies that are essentially dust-free (and, in virtue of the correlation between stellar mass and attenuation, of relatively low stellar mass). Apart from $\text{Ly}\alpha$, the other bright emission lines usually found in star-forming galaxies are the $\text{H}\alpha$ ($\lambda_{\text{rest}} = 0.6563 \mu\text{m}$), $\text{H}\beta$ ($\lambda_{\text{rest}} = 0.4341 \mu\text{m}$), $[\text{O II}]$ ($\lambda_{\text{rest}} = 0.3727 \mu\text{m}$) and $[\text{O III}]$ ($\lambda_{\text{rest}} = 0.5007 \mu\text{m}$) lines. At $z = 4$, these lines are shifted into the near-infrared. Observing in this wavelength domain is quite challenging, owing to the many emission lines emitted by our atmosphere (OH lines), but it is now routinely achieved with NIR spectrometers like KMOS (at the VLT) or MOSFIRE (at Keck). So far, these instruments have been used mostly to study galaxies at $z = 2$ to 3, measuring star formation rates and extinction from the $\text{H}\alpha$ and $\text{H}\beta$ lines. At $z > 2.5$

however, H_{α} is out of the reddest observable window (the K band), and the brightest observable lines are [O II] and [O III].

During my PhD, we have proposed a KMOS program using these lines to get the redshifts of massive dusty galaxies at $z = 2.8$ to 3.9 (PI: R. Leiton). This program was accepted, although with a moderate ranking, and half of our proposed targets have been observed. We received the data in early 2015, and the reduction is still in progress. During the last ESO call, we have also proposed a similar program to target specifically our ALMA detections, and are now waiting for the deliberation of the TAC. If we do obtain spectroscopic confirmation of our targets, not only will this allow us to clean our sample from low redshift contaminants, but it will also reduce significantly the uncertainty on the stellar mass estimates, which is currently of the order of 0.2 dex. However, it is clear that even then, the stellar mass will remain our dominant source of uncertainty (see Buat et al. 2014), since most of the photometry (including the first two *Spitzer* IRAC bands) probes the rest-frame UV-to-optical which is sensitive to dust extinction, but also because the star formation history of these high redshift galaxies is poorly understood.

The second caveat associated to this $z = 4$ Main Sequence is the uncertainty on the dust temperature of each individual galaxy. Assuming a scatter of 4 K, similar to what is observed at lower redshifts (see Chap. 3), our SFRs are about as uncertain as the stellar masses, i.e., about 0.2 dex. This uncertainty can be greatly improved by using the information from the shorter FIR wavelengths, as given by *Herschel*. Thanks to the high angular resolution of ALMA, we are now in a better position to properly extract the *Herschel* fluxes and decompose them into multiple counterparts. This is however quite time consuming, and is therefore still in progress. As a case study, I use such an approach to constrain the FIR SED of three ALMA detections in Sects. 6.8 and 6.9.

6.7 Other Galaxies in the Field of View

Within the field of view of our ALMA observations, I detected a total of 19 source at a significance of $>5\sigma$ that were not in our target list, 10 of which have a clear counterpart in the deep *HST* and/or *Spitzer* IRAC imaging, and 7 of which can be identified to a galaxy in the CANDELS catalogs. The remaining 9 objects with no detectable counterpart are considered to be likely spurious, and are not considered in the following.

As can be read from Table 6.3, for the 7 objects that were successfully cross-matched to the CANDELS catalogs, the average redshift is $\langle z \rangle = 2.04$, and the average stellar mass is $\langle M_* \rangle = 1.6 \times 10^{11} M_\odot$. This is consistent with the typically properties of sub-mm galaxies (e.g., Béthermin et al. 2015), and suggests that these are mostly real detections. However, the measured ALMA flux of AZ4-G1, 12 mJy, is well above what would be expected by extrapolating the *Herschel* photometry (about 0.3 mJy) and the pixel value of the LABOCA map (-1.5 ± 1 mJy). This galaxy is also at $z_{\text{spec}} = 0.515$, which is relatively low and unexpected for a sub-mm detection.

All these hints suggest that, although the ALMA emission coincides very well with the *HST* image, this source must be spurious. I performed a similar inspection of the other detections, and did not find any other such inconsistency.

This leaves a small sample of 6 ALMA detected galaxies at $1.5 < z < 3.1$. At these redshifts, the ALMA measurement is tracing the dust mass, and it can be used to estimate the gas mass of each of these galaxies (see Chap. 5).

Then, there remains 3 detections that have no counterpart in any catalog. The two brightest, AZ4-C5 and AZ4-C6, are detected only in *Spitzer* IRAC, and are discussed in Sect. 6.9. The third one, AZ4-C4, was not extracted in the CANDELS catalogs because of its proximity to a bright ($H = 14.3$) star, and is discussed in the next section.

6.8 A Massive $z = 3$ Galaxy Hidden Behind a Bright Star

AZ4-C4 is a 5.3σ ALMA detection that is located $9.8''$ away from the phase center. It has no counterpart in any known catalog because of its extreme proximity with a bright star. Furthermore, it has a clear detection in the radio with a peak flux of 0.157 mJy (Schinnerer et al. 2007), suggesting that it hosts an AGN. In this section, I present the method I used to measure the photometry of this object, the stellar population modeling that was used to estimate both its photometric redshift and stellar properties, and the interpretation of the *Herschel* and ALMA fluxes to derive its star formation activity.

Because of the neighboring star, standard blind source extraction techniques like SExtractor tend to fail and do not detect this object. The method I have chosen to solve this issue is to mask the galaxies from the image, fit a PSF at the position of the star, and subtract it from the image. Owing to the large dynamic range between the flux of the star and that of the galaxy (about 10 magnitudes), a proper characterization of the PSF is crucial here, since a small uncertainty on this PSF will translate in a large error for the fainter galaxy. Also, the PSF of most instruments is varying across the whole COSMOS field, and it is important to use a *local* PSF that has at least the same orientation to properly subtract the spikes.

I extract the photometry on the images observed in the following broadbands: CFHT-*U*, Subaru-*B*, *HST* F606W, F814W, F125W, F140W and F160W, Vista-*K_s* and *Spitzer* IRAC channels 1 and 2. The first step is to build the PSF. For each image, I build a sample of 30–50 bright stars (about 17–14 magnitudes), and extract a large cutout (10–30'', depending on the image) for each star. I then discard those that are either saturated, are close to another bright star, or have a different orientation than the star I want to subtract. Finally, I normalize each cutout by the value of the central pixel (i.e., the peak of the corresponding star), and combine them using median stacking to produce a high signal-to-noise PSF.

The second step is to subtract the PSF from the image. To do so, I first build a mask to filter out the pixels that contain flux from other sources, in particular from AZ4-C4, so that they do not influence the fit. I also mask the core of the star, so that

the fit operates mostly on the more extended features, which are the ones I want to subtract with most care. Using this mask and the PSF that was built earlier, I use GALFIT to find the best-fit flux and position of the star, and obtain a residual map where the contribution of the star was subtracted. To set the background level to zero, I identify empty regions on the map where the residuals of the star are low, and that are far away enough from AZ4-C4. I compute the median pixel value of these regions, identify this value as the background of the image, and subtract it from each pixel. The uncertainty of this background value is taken into account in the error budget later on.

By inspecting the residual images, which are reproduced in Fig. 6.4, I can see that there are in fact three galaxies behind this star. One is AZ4-C4, another is a compact galaxy that is even closer to the star, and the third is a very diffuse galaxy that is found substantially further. AZ4-C4 appears to have two components in the *H*-band image, one of them being more compact and brighter than the other. In the light of the radio detection, this could be another signpost of AGN activity. It is however almost undetected in any image shortward of the *J* band, indicating that it must be at relatively high redshifts. As for the other two galaxies, although they differ in morphology, their colors appear to be similar. They are both visible clearly in the *U*-band image, indicating that they are most likely low-redshift and dust-free young galaxies.

The last step is now to extract the photometry. For each of these galaxies, I draw circular apertures of varying diameters, which are chosen to encompass most of the flux of each object, and compute the sum of the pixel values that fall inside each aperture. Using the PSFs derived above, I correct this flux measurement for the flux outside of the aperture assuming that the sources are point-like (which they are not, in fact, and therefore the true flux could be slightly higher). The uncertainty on the flux of each object is estimated from the RMS in empty regions of the image, assuming that pixel values are not correlated (which is also wrong in most images, and the actual uncertainty should be slightly higher). To account for the errors in the PSF subtraction and the flux calibration, the signal-to-noise ratio of each measurement is limited to 10. The resulting photometry for AZ4-C4 is shown in Fig. 6.4.

With Tao Wang, we then use FAST (Kriek et al. 2009) to fit this data with Bruzual and Charlot (2003) templates that were built with an exponentially declining star formation history, the Salpeter (1955) IMF, allowing the age to vary between 0.1 and 10 Gyr, the attenuation to vary between $A_V = 0$ to 6, fixing the metallicity to the solar value, and scanning redshifts between $z = 0$ and 8. With these assumptions, AZ4-C4 is found to be at $z = 3.05^{+0.04}_{-0.45}$, with a tail extending toward $z = 2.5$. At the best-fit redshift, it is attributed a large stellar mass of $7 \times 10^{11} M_\odot$, but a relatively low $SFR = 22 M_\odot/\text{yr}$, i.e., a factor of ~ 40 below the Main Sequence. The best-fit stellar population is therefore relatively old, although the code also needs to put an extinction of $A_V = 0.7$ to reproduce the observed data. Owing to the lack of any strong break in their photometry, the redshift of the other two galaxies should be poorly constrained. The compact one is found at $z = 1.23$, while the extended one is found at $z = 0.83$, both with fairly low stellar masses of the order of $\sim 10^9 M_\odot$. These are not particularly interesting, and will not be discussed any further.

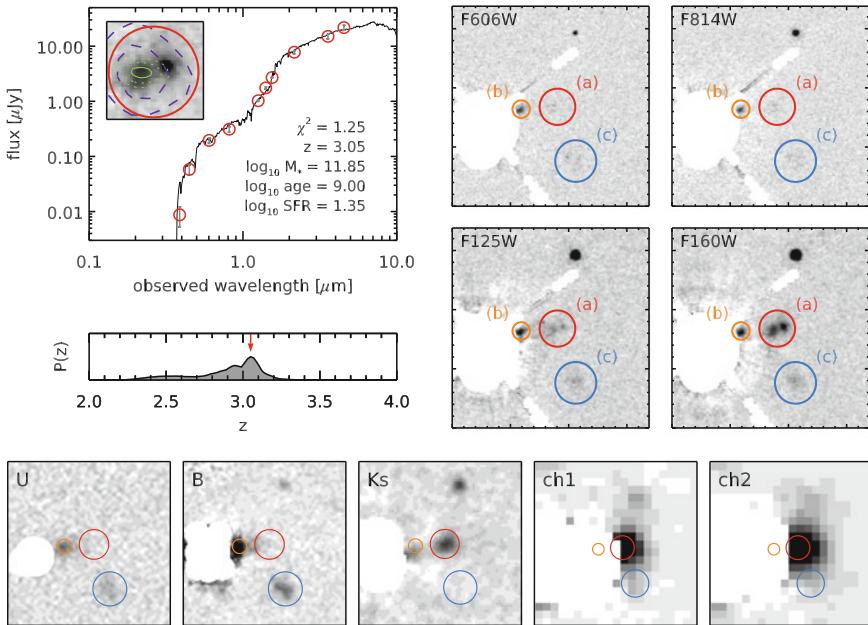


Fig. 6.4 Top-left UV to NIR photometry of AZ4-C4, as obtained by aperture photometry on the different images. The red circles show the measured fluxes, and the black solid line is the best-fit model from FAST. The corresponding best-fit parameters are listed in the bottom-right corner. The inset in the top-left corner shows a zoom-in on the galaxy, the chosen aperture (red circle), the ALMA contours (green, 3 to 4σ as dotted lines, 5σ as a solid line), and the radio contours (purple dashed lines, 5 and 8σ). Middle-left Redshift probability distribution, as inferred from the χ^2 of FAST. The smallest χ^2 solution is indicated with a red arrow. Right and bottom Optical to NIR imaging of AZ4-C4, after subtraction of the bright neighboring star. The regions with large residuals have been masked and appear in white color. The colored circles indicate galaxies that were not included in the CANDELS (or 3DHST) catalogs. The size of the circle gives the aperture that was used to measure the fluxes. (a), in red, is AZ4-C4. (b) and (c) have a very comparable photometry, both in terms of flux and shape, and are fitted with young blue stellar populations at $z \sim 1$, although the redshift is poorly constrained owing to the lack of a clear break in the SED. Both are clearly detected in the U band, indicating that they must be mostly dust-free galaxies at $z < 2$ (Color figure online)

I now come back to AZ4-C4 to see what can be learned from the ALMA detection. Assuming a redshift of $z = 3$ and the typical FIR SED of that redshift ($T_{\text{dust}} = 34$ K, see Chap. 3), the measured ALMA flux translates into an SFR = $(441 \pm 100) M_{\odot}/\text{yr}$, i.e., about a factor of two below the Main Sequence. This SED, which is reproduced in Fig. 6.5, seems to match the observed PACS photometry, although the signal-to-noise ratio of these measurements is quite low. On the other hand, it predicts a flux of about 9 mJy at the tip of the SED, which corresponds to SPIRE 350 μm . This value is above the 3σ detection limit of 7.2 mJy from Elbaz et al. (2011), but actually this limit only applies to galaxies that are “clean”, i.e., free from contamination from their neighbors. In this region of the SPIRE map, it

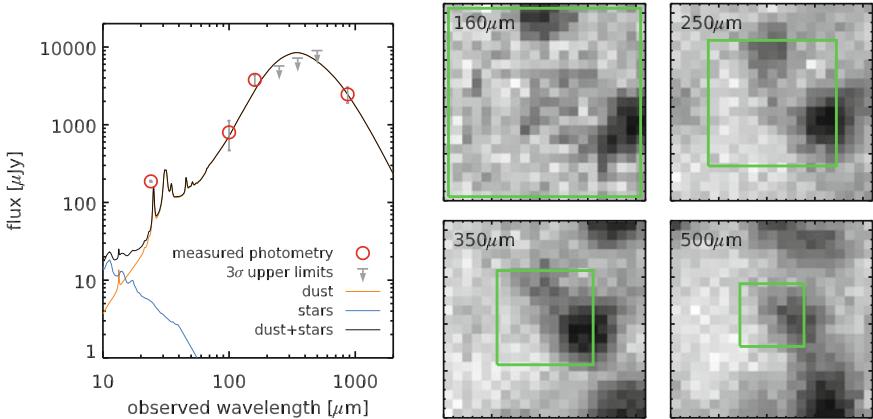


Fig. 6.5 *Left* MIR and FIR photometry of AZ4-C4. Detections are shown with *empty red circles*, while downward gray arrows list the 3σ upper limits, in the absence of a detection. Note that these upper limits in the SPIRE bands are only valid for “clean” sources, and are probably too low since AZ4-C4 lies in a crowded environment. The typical FIR SED of $z = 3$ galaxies is adjusted to the observed ALMA flux and reproduced with a *solid orange line*; the stellar continuum estimated from the UV-NIR photometry (Fig. 6.4) is shown with a *solid blue line*; and the sum of the two is shown with a *solid black line*. *Right* *Herschel* imaging of AZ4-C4. Here the contribution of the star is negligible, and I assume that all the flux at this position can be attributed to AZ4-C4. The green square shows a $50'' \times 50''$ region around AZ4-C4. There is a tentative detection in the PACS image, and possibly also in the SPIRE 250 μm map, but the other images are too confused to reliably extract the flux of this galaxy (Color figure online)

is clear that there are a number of such neighbors, and decomposing the observed flux proves difficult. To see if the chosen SED is indeed consistent with the available SPIRE data, I tried to subtract the extrapolated fluxes from the observed maps, and looked at the residuals, which do not reveal any hint of an over-subtraction. Hence this SED can be considered as compatible with the *Herschel* data.

However, the *Spitzer* MIPS detection appears in clear excess compared to this SED. Combined with the radio detection, this is probably another sign that this galaxy hosts a strong AGN. In fact, now that we know the redshift, we can also see what we can learn from the radio emission. It turns out that, if one was to assume that all the radio flux comes from star formation, and therefore that it obeys the radio-FIR correlation (as given in Pannella et al. 2015), the measured radio flux would correspond to an infrared luminosity that is *ten times* larger than that derived from ALMA. This is true even if we perturb the redshift within its allowed confidence interval, and therefore demonstrates the presence of an AGN. If I subtract the 24 μm flux predicted by the best-fit FIR SED, the residual flux is around 100 μJy . Using the relation of Lutz et al. (2004) that links the rest-frame 6 μm luminosity to the X-ray luminosity for AGNs, this flux would correspond to an intrinsic $L_X \sim 10^{44}$ cgs which is above (but close to) the detection limit of the Chandra 2Ms survey (6×10^{43} cgs at $z = 3$). The fact that the galaxy is not detected in the X-ray image suggests either

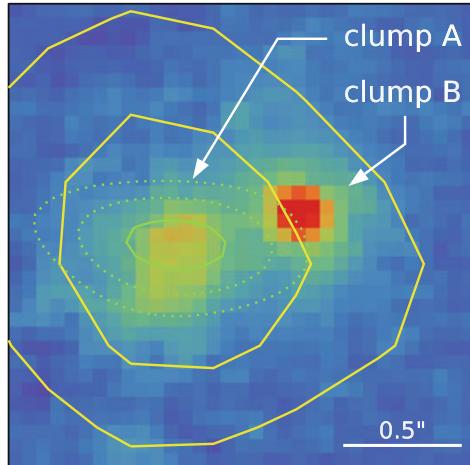


Fig. 6.6 Zoom in on AZ4-C4, after subtraction of the nearby star. The background false-color image is the *HST H*-band image (rest-frame $0.4 \mu\text{m}$), the yellow contours show the radio emission as measured by the VLA, and the green contours show the locus of the ALMA flux. Neither the radio nor the submm are resolved (the elongated shape of the ALMA profile is just caused by the ellipticity of the beam). This is the same figure as the inset in Fig. 6.4, with the contrast modified so as to better grasp the light distribution from the *HST* imaging. The two “clumps” that compose this galaxy are dubbed “clump A” and “clump B”, and are indicated with arrows. Although “clump B” appears much brighter, it is actually just more concentrated. The same amount of light is emitted by “clump A”, but on a more extended scale (Color figure online)

that the AGN is obscured (as is the case for most submm host galaxies, Alexander et al. 2005), or that the excess at $24 \mu\text{m}$ originates from another source (e.g., another close-by galaxy that I could not deblend).

In Fig. 6.6 I show a zoom-in on AZ4-C4 to better grasp the light distribution. The fact that both the radio and the ALMA emission originate from only one of the two “clumps” seen in the *HST H*-band image is intriguing. One possibility is that this is a merger of two galaxies, one being gas-rich and star-forming (clump A), and the other one being old and quiescent (clump B). Both clumps appear to have the same colors in the *HST* images, which would be consistent with this scenario if the star-forming galaxy is also strongly reddened by dust. Another (maybe more exotic) possibility is that the radio emission originates from a jet, emitted from an AGN that resides in clump B. This jet would turn out to be oriented toward clump A, where it compresses the gas and triggers a starburst (e.g., Elbaz et al. 2009). This hypothesis could be tested by resolving the radio emission, since the current beam of the radio observations in COSMOS is too large to allow any morphological analysis. For reference, at $z = 3$, the proper distance between the two clumps is 3.8 kpc.

6.9 Discovery of Two New High-Redshift Dusty Galaxies

In this section, I report the discovery of two unexpected 20σ ALMA sources close to some of our targets in the COSMOS field (AZ4-C5 and AZ4-C6). Surprisingly, these two galaxies have no counterpart in any catalog, and in fact there is little to no detectable emission in the deep optical-to-NIR broadband images, except in the first two *Spitzer* IRAC channels (3.6 and 4.5 μm) where both galaxies are clearly detected (see Fig. 6.7). Together with the other co-investigators, we dubbed these objects the first “dark ALMA galaxies”. They are interesting in many aspects, in particular for the fact that they could be the most distant massive and dusty star-forming galaxies ever detected. However, this claim can only stand if their redshift is spectroscopically determined. For this reason, I have proposed to use the “spectral scan” capabilities of ALMA to locate the [C II] emission line, which is the brightest line in the FIR. This proposal has been accepted and highly ranked, and will be observed sometime during the coming year. In the following, I review the interesting properties of these two objects, and try to constrain the redshift with the little amount of information currently available.

6.9.1 Optical to NIR Photometry

The UV to NIR fluxes of both objects are obtained in a similar fashion as for AZ4-C4 (see previous section), using aperture photometry. Here also, the IRAC images were first deconfused by fitting the bright neighbors using GALFIT, subtracting their best-fit profiles, and performing the aperture photometry on the residual map. For the other bands, the contamination from these neighbors is negligible, and the photometry is performed on the observed maps directly.

Unfortunately, none of these galaxies is covered in the *HST*-ACS images (except for the shallow F814W imaging that covers the whole COSMOS field), and the optical photometry is therefore purely ground based. On top of this, one of the two galaxies, AZ4-C6, is actually located close to the edge of the *HST*-WFC3 chip, where the noise is larger than usual. Therefore, to obtain the best possible constraints, I also perform the photometry on the ground-based UltraVISTA Y , J and H bands. Finally, to help deriving the photometric redshift, I add to the list the IRAC channels 3 and 4. While these images are quite shallow in COSMOS and do not provide very stringent constraints, every bit of information is useful when deriving the redshift of a galaxy that is only clearly detected in two broad bands.

In the end, the two galaxies (AZ4-C5 and AZ4-C6, respectively) are detected at 4.6 and 8.1σ in the IRAC 3.6 μm channel, and 9.9 and 10.9σ in the 4.5 μm channel. AZ4-C5 is also weakly detected at 3.2σ in the 8 μm channel. As for the shorter wavelength, AZ4-C5 turns out to be non-detected in all bands, while AZ4-C6 has in fact a 5.8σ detection in the K_s band and tentative 3σ detections in the *HST J* and H bands. The reason why it was not included in the CANDELS and K_s -band based

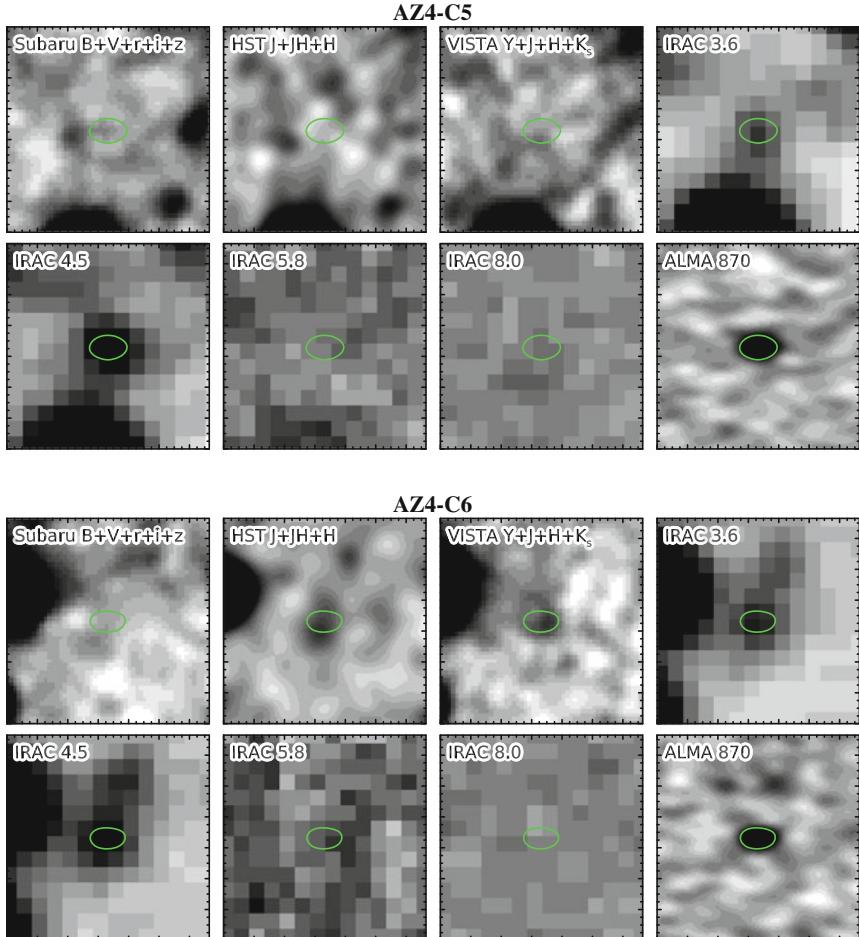


Fig. 6.7 Postage stamps of the two “dark ALMA galaxies” that we discovered in our ALMA survey. The first eight images show AZ4-C5, and the remaining ones show AZ4-C6. For each galaxy, the following images are displayed: • the sum of the Subaru images in the B , V , r , i and z bands, • the sum of the HST WFC3 images in the F125W, F140W and F160W bands, • the sum of the VISTA Y , J , H and K_s bands, • the four *Spitzer* IRAC bands, and • the ALMA image, non-primary-beam-corrected. The first three images are smoothed with a Gaussian kernel to reveal the extended and faint structures that would otherwise be undetectable. On each image, I show a green contour corresponding to the 1 mJy level of the ALMA emission (Color figure online)

catalogs of Muzzin et al. (2013) and Ilbert et al. (2013) is probably because it appears to be substantially extended, and therefore has a surface brightness low enough to fall below the detection threshold.

Near AZ4-C6 is another galaxy that is faint in the HST imaging, but also seen in IRAC and ALMA, albeit with a lower signal-to-noise ratio for the latter. This is our main target “2-38011”, which was attributed a $z_{\text{phot}} = 4.23$ and a flux at $870 \mu\text{m}$

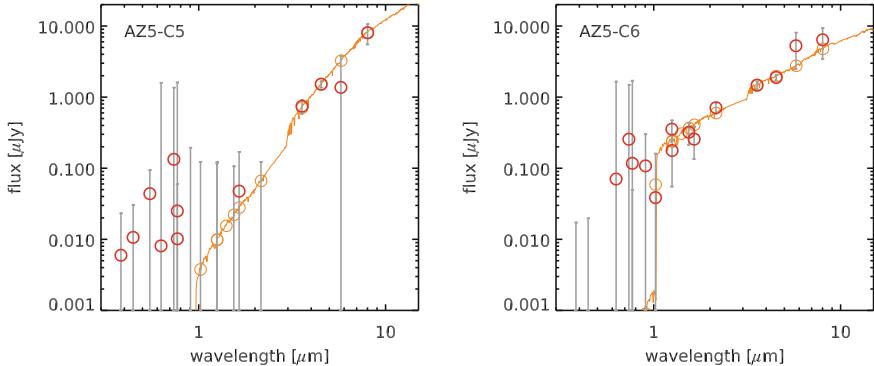


Fig. 6.8 *U* to *Spitzer* IRAC photometry of the two dropouts (*left* AZ4-C5, *right* AZ4-C6). The measured fluxes are shown with *empty red circles*, with their associated error bars in *light gray*. To guide the eye, the optimal best-fit SED from FAST is shown with a *solid orange line* (at $z = 6.9$ and $z = 7.45$, respectively, although the precise redshift is very uncertain), and the model fluxes are shown with *empty orange circles* (Color figure online)

that is half that of AZ4-C6 (with a substantial spatial extension of $0.78''$). A third source is tentatively detected in the ALMA image, to the left of 2-38011, but it has no counterpart in any other band.

The resulting SEDs are shown in Fig. 6.8.

6.9.2 MIR to FIR Photometry

While there is a clear signal in the SPIRE bands at the position of both galaxies, the number of other possible counterparts for this emission is large, and therefore the FIR fluxes cannot be accurately associated to our two dropouts. This is especially true for AZ4-C6, since our primary target (2-38011) is detected on this image: we therefore have a $z = 4$ galaxy very close to another galaxy of unknown redshift, making it impossible to properly decompose the observed flux. The simplest thing to do would be to extract the SPIRE fluxes assuming that they are produced by a single source, and use the resulting values as upper limits. However, these constraints are not very stringent, in particular for AZ4-C6 where $S_{500} < 40$ mJy. For the present work however, we do not need to bother with the SPIRE measurements, as the fluxes in both MIPS and PACS (essentially non-detections) already bring significant pieces of information and can be measured straightforwardly.

The situation of AZ4-C5 is a bit more complex, owing to the presence of a nearby $z = 0.46$ bright star-forming galaxy. This galaxy is so bright that one can see the secondary lobes of the MIPS and PACS PSFs. If these are not properly reproduced in the model PSF, part of the flux in the sidelobes will be attributed to the surrounding galaxies. Therefore, I have built a custom PSF from the *Herschel* observations of

Vesta,⁴ resampling it to the pixel size of our maps, and rotating it to match the actual orientation of the satellite when this image was obtained.

6.9.3 A First Estimate of Their Physical Properties

Back in April, when we proposed the ALMA spectral scan, I did a first quick estimation of the redshift and stellar mass of these two objects. I extracted the IRAC fluxes using PSF-fitting, and assumed non-detection in the other bands at the 5σ level. I derived a first redshift estimation by looking for an observed SED in the CANDELS catalogs that, properly redshifted, would match this photometry. This told me that the redshift could be either around $1.5 < z < 2.5$, or at $z > 4$ with a peak around $z = 7$. At the time, we claimed that the $z = 2$ solution was ruled out from the fact that the galaxies are not detected by *Spitzer* MIPS or *Herschel*. Indeed, a galaxy at this redshift with a $870 \mu\text{m}$ flux of 3 mJy would be very bright in the mid- to far-IR, substantially above the detection limit in all *Spitzer* and *Herschel* bands. In this section, I review these estimations and claims with a more careful and quantitative data analysis.

Using FAST, I fitted the updated UV to NIR photometry, similarly to AZ4-C4, and taking into account the non-detections. Since I allow extreme attenuations ($0 < A_V < 6$), the code can actually find a suitable fit at all $z > 1.5$. The resulting $P(z)$ and best-fit parameters are shown in Fig. 6.9.

For AZ4-C5, all the solutions below $z = 4$ require $A_V > 4$ and an age larger than 1 Gyr. This is a very peculiar and unlikely combination. At $z > 4$, the needed attenuation is still fairly high, and never drops below $A_V = 3$, while the age falls below 1 Gyr only beyond $z > 5$. With such a large amount of extinction, the stellar mass is in fact very large: the best fit M_* already reaches $10^{11} M_\odot$ at $z = 2.5$, and $10^{12} M_\odot$ at $z = 6$. It is therefore possible that this galaxy hosts a buried AGN that makes the IRAC colors redder (Donley et al. 2012) and forces FAST into using unrealistically large attenuations. If this is true, then obtaining the redshift from this photometry is simply hopeless. One argument against this hypothesis is that, at $z \geq 5$, the ratio of the UV to IR luminosity is extremely low (see later in Sect. 6.9.5), and would be even lower if the contribution of a putative AGN is removed from the optical SED. Another possibility is that the IRAC photometry is inaccurate, although I tend not to favor this explanation since the measured color is perfectly consistent with the non-detection in the K_s band, and the tentative detection at $8 \mu\text{m}$. Lastly, it could also be that the $4.5 \mu\text{m}$ flux is contaminated by strong emission lines, as shown, e.g., in Labb   et al. (2013). $\text{H}_\alpha + [\text{N II}]$, the brightest line, would imply $5.1 < z < 6.6$, but then the $3.6 \mu\text{m}$ band should also be contaminated by $[\text{O III}]$. To reach the red IRAC color we observe ($[3.5] - [4.5] = 1$), one would have to consider $[\text{O III}] + \text{H}_\beta$ at $6.97 < z < 9$. However this interpretation would create a tension with the measured $8 \mu\text{m}$ flux, which is fairly high. Furthermore, since the dropouts are strong submm

⁴<https://nhscsci.ipac.caltech.edu/pacs/data/PSFs/vesta20091109/>.

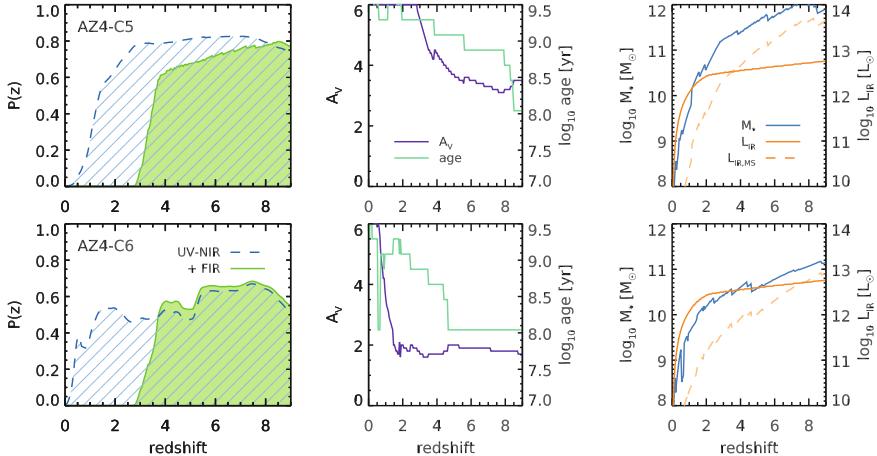


Fig. 6.9 Derived physical quantities of both dropouts (top AZ4-C5, bottom AZ4-C6). Left Redshift probability distribution inferred from the reduced χ^2 of the best-fit SED, either using only the U to IRAC photometry (blue dashed line) or all the photometry including the FIR (solid green line). Middle Best-fit attenuation (A_V , purple solid line) and light-weighted stellar age (green solid line) as a function of the redshift. Right Best-fit stellar mass (M_* , blue solid line) and total infrared luminosity (L_{IR} , orange solid line) as a function of redshift. I also show the expected L_{IR} based on the measured stellar mass, assuming that the galaxy lies exactly on the Main Sequence and that most of the SFR is obscured (orange dashed line) (Color figure online)

emitters, one can reasonably expect that they are substantially obscured, and therefore that emission lines should be strongly attenuated. Similarly, one can also interpret the non-detection in the K_s band as the signature of the Balmer break, in which case the resulting constraints on the redshift would be $5.5 < z < 9$. But here again, the Balmer break is probably not a significant spectral feature in dust-rich galaxies. *In fine*, several hints point toward $z > 5$, but the precise redshift remains elusive.

Although AZ4-C6 has a more complete photometry, the absence of any break in the measurements also leads to a poorly constrained redshift probability distribution. However, here the fit parameters are more reasonable. The attenuation stays constant at $A_V = 2$ all the way from $z = 2$ to 9, and rises rapidly at $z < 2$ to reach $A_V = 4$ at $z = 1$. On the other hand, the best-fit age is found around 1 Gyr at $z < 3.5$, and then rapidly drops to 100 Myr at $z > 4.5$. Except at $z < 2$, none of these values are particularly intriguing. The stellar mass also behaves reasonably, being equal to $10^{10} M_\odot$ at $z = 1.5$ and $10^{11} M_\odot$ at $z = 7.5$. In the end, the available photometry does not give very strong constraints on the redshift, although here also $z > 4.5$ would be preferred.

It turns out that the most stringent constraints on the redshift are provided by the combination of the strong detection in ALMA together with the weak (or non-) detections in the *Spitzer* MIPS and *Herschel* PACS images. To derive these constraints, I assume that the dropouts do not have atypical dust properties, and therefore that the effective temperature of their dust SED is close to the average value at a given redshift

(Chap. 3). This can be wrong and influence the best-fit redshift in both directions, since an SED that is well fitted by a given T_{dust} at a given redshift can be equally well described by a colder T_{dust} at a lower redshift, or a warmer T_{dust} at higher redshift.⁵ Since these two dropouts are among our brightest ALMA detections, they are more likely to be starbursting systems, which are known to show warmer T_{dust} than the average (Elbaz et al. 2011; Magnelli 2014; Béthermin et al. 2015). Therefore the redshift constraints I derive here could be biased toward lower redshifts, but I prefer to remain conservative and assume standard T_{dust} values. Then, at each redshift in the grid created by FAST, I normalize the dust SED to the observed ALMA flux, and combine it with the best-fit stellar template from FAST to build a full SED from the UV to the FIR. The contribution of the stellar component to the MIR and FIR fluxes is usually negligible, except at high redshifts where it starts to be the dominant source of the observed 24 μm . I finally measure the χ^2_{dust} by comparing the model SED against the observed MIPS and PACS fluxes (or non-detections), and add it to the χ^2_{stellar} originally produced by FAST using the UV to NIR photometry. Finally, I compute the redshift probability distribution with $\exp(-\chi^2_{\text{dust}} - \chi^2_{\text{stellar}})$.

The result is overplotted in Fig. 6.9. In both cases, the constraints from the dust emission exclude all the solutions at $z < 3.5$ and tend to favor $z > 5$. This is also illustrated in Fig. 6.10 where I show the resulting fit at three different redshifts: $z = 2$ where the fit is definitely poor, $z = 4$ where the fit becomes good, and $z = 7.5$ to show that it remains acceptable even at extremely high redshifts.

Therefore, with this more detailed and precise analysis, I confirm that the two dropouts are most likely to be massive galaxies at $z > 4$, with a preference for $z > 5$. The precise redshift remains to be determined though, reinforcing the need for the spectroscopic scan I describe in the next section.

6.9.4 Measuring the Redshift with ALMA

Since both objects are very faint in the observed optical and NIR, and also because they probably host large amounts of dust, getting a spectroscopic redshift from the usual optical emission lines (Ly_α , H_α , [O II], etc.) would require unreasonable integration times. Our best chance to get a redshift is to target instead one of the many luminous lines in the far-IR.

The most convenient way of determining FIR redshifts is the so-called “CO-ladder”. These lines are produced by rotational de-excitation of the CO molecule, e.g., CO(5-4) for the transition from the rotational quantum level $J = 5$ to $J = 4$, and the flux of each line relative to the CO(1-0) line is determined by the Spectral Line Energy Distribution (SLED). Depending on a number of factors, in particular the strength of the current star formation and also the geometry of the molecular

⁵ T_{dust} is roughly measured from the position of the peak of the FIR SED in the rest frame. In fact, what we actually measure as observers is $T_{\text{dust}}^{\text{obs}} = T_{\text{dust}}/(1+z)$, hence the degeneracy between z and T_{dust} .

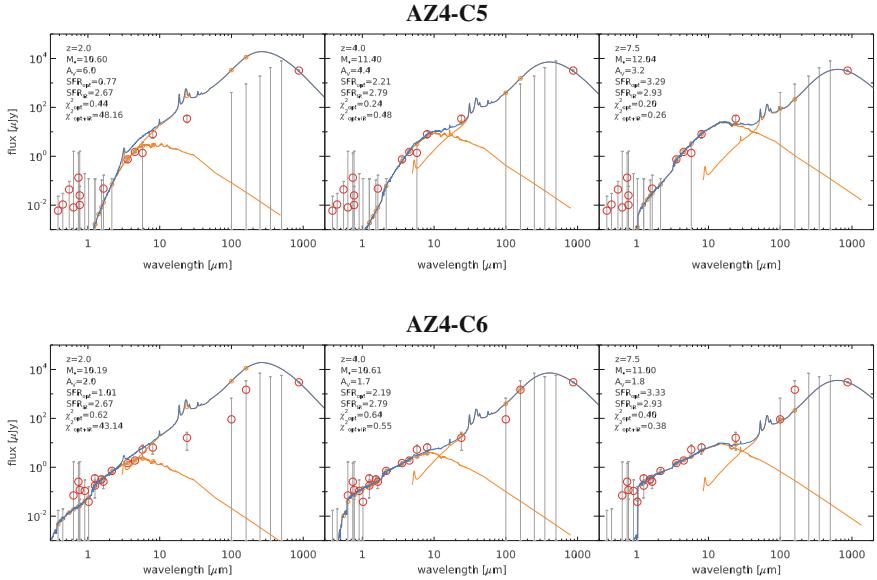


Fig. 6.10 Panchromatic SED of the two dropouts (top AZ4-C5, bottom AZ4-C6). Each panel shows the best-fit model SED at different redshifts, $z = 2$, $z = 4$ and $z = 7.5$ from left to right. On all figures, measured fluxes are shown with empty red circles, the stellar and dust SEDs are shown individually with two solid orange lines, while the combined SED is shown with a solid blue line. The corresponding model fluxes are shown with orange small empty circles (Color figure online)

clouds, this SLED will vary greatly from one galaxy to another (see, e.g., the review of Carilli and Walter 2013). In fact, high- J CO lines luminosities are well correlated with the total infrared luminosity, and are therefore good tracers of star formation. Low- J CO lines, on the other hand, and in particular the CO(1-0) line, are good tracers of the molecular gas mass (e.g., Daddi et al. 2010). The greatest advantage of these lines is that they are relatively close to one another, and it is therefore easy to have at least two such lines covered within a given submm band (see Table 6.4), thereby unambiguously establishing the redshift.

There are tools available to prepare such scans with ALMA⁶ that can automatically prepare the spectral setup to cover a given redshift range and guess the line fluxes to estimate the necessary integration time. Alternatively, one can estimate the lines fluxes using published scaling relations (e.g., Daddi et al. 2015), knowing, e.g., the L_{IR} , and prepare the spectral setup manually in the ALMA Observing Tool (AOT).

For our dropouts, the optimal strategy for a CO spectral scan is to observe in band 3, and target the CO(5-4) and CO(6-5) lines, covering $3.2 < z < 5.9$. However, with M. Pannella, we estimated the line flux to be of the order of 1 mJy, with a relatively large uncertainty owing to the unknown CO SLED. Reaching a 5σ significance for a line of that flux would require about 6 h of ALMA time (including calibration,

⁶<https://github.com/aconley/ALMAzsearch>.

Table 6.4 ALMA passbands available for observing in Cycle 3 (omitting band 8 and 9, which have poor atmospheric transmission)

Band 7	Band 6	Band 4	Band 3
811 – 1091 μm	1091 – 1422 μm	1840 – 2400 μm	2521 – 3571 μm

which is the most time consuming part of a spectral scan), which we judged was too expensive and unlikely to be observed. In fact, so far, CO spectral scans of $z > 4$ galaxies have only been obtained for lensed galaxies (e.g., from the SPT survey, Weiβ et al. 2013).

Instead, I decided to go for another strategy and target the [C II]-158 μm line in both band 7 and band 6. This line is the brightest in the FIR, and is well correlated with the total IR luminosity, with a scatter of a factor of two (Stacey et al. 1991; Helou et al. 2001). However, there is a known saturation of this correlation (Graciá-Carpio et al. 2011), and starburst galaxies tend to show a [C II] deficit compared to Main Sequence galaxies of similar L_{IR} (Díaz-Santos et al. 2013), and this has to be taken into account. Although only a single line would be observed, the redshift determination would still be unambiguous because the chosen sensitivity will not permit the detection of any other line in these bands. In fact, within band 7 and 6, the only other observable bright lines would be the high-J CO ladder at $z < 2.5$, which is a domain I already ruled out with the FIR photometry.

After discussion with the other co-investigators (D. Elbaz, M. Pannella, T. Wang and R. Leiton), and based on the first crude photometric redshift estimation, we decided to scan the redshift range $5.3 < z < 6.8$. To do so, I configured 8 spectral setups, divided into two contiguous spectral scans: one at the end of band 7 (covering $5.3 < z < 5.9$), and one at the beginning of band 6 (covering $5.9 < z < 6.8$). This is illustrated in Fig. 6.11. I thought of adding an additional spectral scan at the other end of the band 7 to cover the range $4.3 < z < 4.7$ (the redshift window between $z = 4.7$ and 5.3 is hardly accessible with [C II] because of a substantial drop of atmospheric transmission within the band 7, as shown in Fig. 6.11). However that would have requested too much observing time (about 6 h). Also, at the time of writing the proposal, the redshift probability distribution was more peaked toward higher redshifts than what I derived in the previous section. If the proposal gets accepted, I may reconsider the chosen spectral setups and maybe move the band 6 scan back into the band 7 to cover $z \sim 4.5$.

To estimate the necessary integration time, I based the estimation of the flux on the recent observations of $z = 5$ –6 Lyman Break Galaxies from Capak et al. (2015), in particular on their brightest galaxy, HZ10. It is found at $z = 5.657$ with a continuum flux at 870 μm of 1.3 mJy, i.e., about a third of the flux of our dropouts. Its [C II] line flux is 1.57 Jy km/s with a line width of 127 km/s, corresponding to a peak flux of about 7 mJy. I then assume that both HZ10 and our dropouts are in the regime where the [C II] line flux is saturated (which appears to be the case at least for HZ10, from Fig. 4 in Capak et al. 2015). Therefore, although our dropouts have three times larger L_{IR} , I conservatively assume that they will have roughly the same

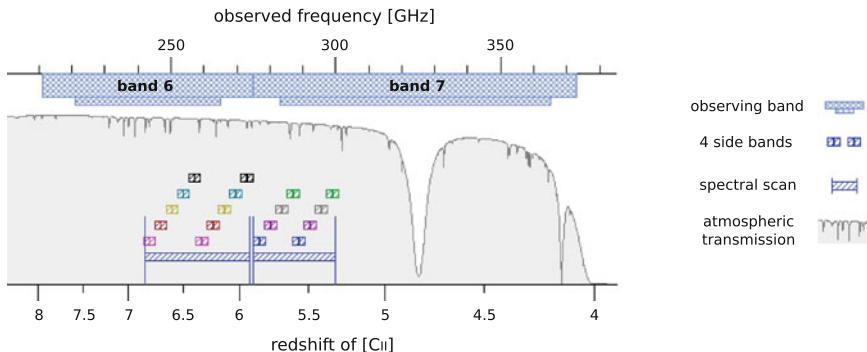


Fig. 6.11 Observing setup for the ALMA [CII] spectral scan. This figure shows the two spectral scans that I proposed during the last ALMA call for proposal. The frequency range covered by each spectral scan is indicated with a striped *dark blue region*. Within each scan, this figure shows the individual spectral setups that compose the scan (the *small colored squares*), each setup containing four passbands that are observed simultaneously. The atmospheric transmission curve is shown in the background with arbitrary units (Color figure online)

[CII] flux as HZ10, with some error margin to take into account the dispersion of the [CII]– L_{IR} correlation. I therefore settled for a sensitivity of 1 mJy per spectral element with a resolution of 32.6 km/s (or 31.3 MHz, just enough to get about 3 spectral element over the FWHM of the line). The total integration time needed to achieve this sensitivity in both scans is 3 h, including overheads. The proposal has been accepted and ranked “A” (i.e., within the top 5 % of all submitted proposals).

6.9.5 Potential Scientific Outcome

For these proposed ALMA observations, I requested a sharper angular resolution than what we got in our Cycle 2 data. Indeed, a by-product of the two spectral scans is that we will reach extreme signal-to-noise ratios on the continuum ($S/N \sim 70$), which will enable precise size measurements should the beam size be small enough. Note however that this is a dangerous game, since the [CII] profile is known to often be substantially more extended than the galaxy itself. To prevent complications in the [CII] flux measurements, we should therefore avoid observing these galaxies with a too high angular resolution. We thus decided to settle on an angular resolution of 0.5'', i.e., a factor 1.5 better than the resolution achieved for the continuum measurement, where both galaxies were found to be unresolved (or barely resolved). By measuring the size of the star-forming region, we will derive the SFR surface density. Knowing the redshift, a natural and quick follow up with either ALMA or PdBI will deliver the CO luminosity, which will in turn give us access to the molecular gas mass. Having access to both this quantity and the SFR surface density will provide a unique measurement of the star formation efficiency in a very high redshift galaxy.

Lastly, the faint optical to NIR fluxes of both objects translate into a strong lower limit on the InfraRed eXcess, $\text{IRX} = \log_{10}(L_{\text{IR}}/L_{\text{UV}})$. AZ4-C6 is found to have $\text{IRX} > 3$ at all $z < 5$, and $\text{IRX} > 2$ at $z > 5$, while AZ4-C5 has $\text{IRX} > 5$ at $z < 4.5$ down to a minimum of $\text{IRX} = 3$ at $z = 9$. These values are much higher than what was recently reported for the $z = 5$ LBGs of Capak et al. (2015), who found at most $\text{IRX} = 0.5$ and argued for a strong change of the dust content at these redshifts. If our dropouts are indeed confirmed to be at $z > 5$, they will complement this latter sample and show that there exists dusty galaxies with ISM properties similar to $z = 2\text{--}4$ star-forming galaxies, even up to $z \sim 6$. This is expected especially among the most massive systems, which are known at $z < 4$ to be the most dusty (e.g., Pannella et al. 2009; Buat et al. 2012; Heinis et al. 2014), and less likely to be selected in LBG samples (Wang et al. in prep.).

References

- D.M. Alexander, I. Smail, F.E. Bauer et al., *Nature* **434**, 738 (2005)
 ALMA Partnership, C. Vlahakis, T.R. Hunter et al., *ApJ* **808**, L4 (2015)
 E. Bertin, S. Arnouts, *A&AS* **117**, 393 (1996)
 M. Béthermin, C. De Breuck, M. Sargent, E. Daddi, *A&A* **576**, L9 (2015)
 R.J. Bouwens, G.D. Illingworth, P.A. Oesch et al., *ApJ* **737**, 90 (2011)
 G. Bruzual, S. Charlot, *MNRAS* **344**, 1000 (2003)
 V. Buat, S. Heinis, M. Boquien et al., *A&A* **561**, A39 (2014)
 V. Buat, S. Noll, D. Burgarella et al., *A&A* **545**, A141 (2012)
 D. Calzetti, L. Armus, R.C. Bohlin et al., *ApJ* **533**, 682 (2000)
 P.L. Capak, C. Carilli, G. Jones et al., *Nature* **522**, 455 (2015)
 C.L. Carilli, F. Walter, *ARA&A* **51**, 105 (2013)
 R. Chary, D. Elbaz, *ApJ* **556**, 562 (2001)
 E. Daddi, F. Bournaud, F. Walter et al., *ApJ* **713**, 686 (2010)
 E. Daddi, H. Dannerbauer, D. Liu et al., *A&A* **577**, A46 (2015)
 R. Davé, K. Finlator, B.D. Oppenheimer, *MNRAS* **416**, 1354 (2011)
 T. Díaz-Santos, L. Armus, V. Charmandaris et al., *ApJ* **774**, 68 (2013)
 J.L. Donley, A.M. Koekemoer, M. Brusa et al., *ApJ* **748**, 142 (2012)
 D. Elbaz, M. Dickinson, H.S. Hwang et al., *A&A* **533**, 119 (2011)
 D. Elbaz, K. Jahnke, E. Pantin, D. Le Borgne, G. Letawe, *A&A* **507**, 1359 (2009)
 V. González, I. Labb  , R.J. Bouwens et al., *ApJ* **713**, 115 (2010)
 J. Graci  -Carpio, E. Sturm, S. Hailey-Dunsheath et al., *ApJ* **728**, L7 (2011)
 S. Heinis, V. Buat, M. B  thermin et al., *MNRAS* **437**, 1268 (2014)
 G. Helou, S. Malhotra, D.J. Hollenbach, D.A. Dale, A. Contursi, *ApJ* **548**, L73 (2001)
 J.A. Hodge, A. Karim, I. Smail et al., *ApJ* **768**, 91 (2013)
 J.A. H  gbom, *A&A* **15**, 417 (1974)
 O. Ilbert, H.J. McCracken, O. Le F  vre et al., *A&A* **556**, 55 (2013)
 R.C. Kennicutt Jr., *ARA&A* **36**, 189 (1998)
 M. Kriek, P.G. van Dokkum, I. Labb   et al., *ApJ* **700**, 221 (2009)
 I. Labb  , P.A. Oesch, R.J. Bouwens et al., *ApJ* **777**, L19 (2013)
 D. Lutz, R. Maiolino, H.W.W. Spoon, A.F.M. Moorwood, *A&A* **418**, 465 (2004)
 B. Magnelli, D. Lutz, A. Saintonge et al., *A&A* **561**, 86 (2014)
 I. Mart  -Vidal, W.H.T. Vlemmings, S. Muller, S. Casey, *A&A* **563**, A136 (2014)
 G.R. Meurer, T.M. Heckman, D. Calzetti, *ApJ* **521**, 64 (1999)

- A. Muzzin, D. Marchesini, M. Stefanon et al., *ApJS* **206**, 8 (2013)
M. Pannella, C.L. Carilli, E. Daddi et al., *ApJ* **698**, L116 (2009)
M. Pannella, D. Elbaz, E. Daddi et al., *ApJ* **807**, 141 (2015)
C.Y. Peng, L.C. Ho, C.D. Impey, H. Rix, *AJ* **124**, 266 (2002)
B. Salmon, C. Papovich, S.L. Finkelstein et al., *ApJ* **799**, 183 (2015)
E.E. Salpeter, *ApJ* **121**, 161 (1955)
E. Schinnerer, V. Smolčić, C.L. Carilli et al., *ApJS* **172**, 46 (2007)
C. Schreiber, M. Pannella, D. Elbaz et al., *A&A* **575**, A74 (2015)
C. Schreiber, M. Pannella, R. Leiton et al. (2016). ArXiv e-prints, [arXiv:1606.06252](https://arxiv.org/abs/1606.06252)
G.J. Stacey, N. Geis, R. Genzel et al., *ApJ* **373**, 423 (1991)
D.P. Stark, R.S. Ellis, A. Bunker et al., *ApJ* **697**, 1493 (2009)
D.P. Stark, M.A. Schenker, R. Ellis et al., *ApJ* **763**, 129 (2013)
C.M.S. Straatman, I. Labb  , L.R. Spitler et al., *ApJ* **783**, L14 (2014)
A. Wei  , C. De Breuck, D.P. Marrone et al., *ApJ* **767**, 88 (2013)
T.L. Wilson, K. Rohlfs, S. H  ttemeister, Tools of Radio Astronomy (2009)

Chapter 7

Conclusions and Perspectives

In this thesis, I presented a variety of results related to the study of galaxy evolution. Using the best estimates available today of stellar masses and unbiased FIR-based star formation rates, I revisit the correlation between these two quantities, the *Main Sequence* of star-forming galaxies, and extend it from $z = 2$ to $z = 3.5$ (Schreiber et al. 2015). Then, using preliminary data from our ALMA survey, I extend further the study of this relation up to $z \sim 4$, in a regime that could not be probed by the *Herschel* satellite. At all $z < 3$ and at least for $M_* > 3 \times 10^9 M_\odot$, I measured the dispersion of SFRs around this relation and found it to be relatively small, reinforcing the idea that the majority of the star formation in the Universe is happening through a steady, secular process that has yet to be clearly identified.

Motivated by the high-quality IR SEDs that I obtained in this first work, I designed a new library of model SEDs to provide a fine description of the dust continuum average temperature and PAH mass fraction, and calibrated the evolution of both quantities using the observed *Herschel* stacked SEDs. The resulting library will be published together with the following analysis in a paper to be submitted. I found that the average dust temperature in these Main Sequence galaxies is going down with cosmic time, confirming already published trends. I also report for the first time an evolution of the PAH mass fraction that exceeds the expected trend from the varying metallicity. Both observations suggest that the ISM properties of Main Sequence galaxies were actually evolving through time, possibly because distant galaxies were forming their stars in more compact regions, with higher gas fractions.

I then used this new library to interpret the SED of $z = 1$ galaxies of varying stellar mass, and measure their dust content. Using this measurement to estimate the gas content (via the metallicity), I showed that Main Sequence galaxies at this redshift and $M_* > 5 \times 10^{10} M_\odot$ are evolving with significantly lower star formation efficiencies compared to galaxies of lower stellar mass. This decrease of SFE actually goes together with a reduced slope of the Main Sequence, and becomes more pronounced toward $z = 0$. This suggests that massive star-forming galaxies are in a state of *global decline* starting from $z = 1.5$, experiencing a *slow downfall* of their star formation

activity. Here as well, the precise mechanism that drives this evolution is elusive, although several candidates are known and discussed. This work will be submitted to *A&A* in August.

Finally, at higher redshifts ($z > 5$), I reported the discovery of two massive and dusty star-forming galaxies in the early Universe that are seen by ALMA but missed by *Hubble* because of their extreme distance and dust content. The mere existence of these objects, if confirmed to be at the right redshift, would add more variety to the known $z = 5\text{--}6$ Universe and show that dust is still a key ingredient to properly characterize star-forming galaxies beyond the Lyman Break population. I will wait for the result of the spectroscopic confirmation I proposed before publishing this analysis.

All of these results beg for further investigation. In my opinion, there is still much progress to be made on the observational side to provide the last bricks necessary to reconstruct the evolution of galaxies from their birth to their death.

Star formation rates based on the dust emission are scarce, and while *Herschel* is now resting in peace somewhere in the solar system, we have yet to extract all the information present in the fantastic images it produced. I believe some of this information can be recovered today by optimizing our interpretation of the images (see, e.g., Sect. A.2 where I describe a method to go further than traditional flux catalogs), but it is also clear to me that a larger fraction will be unlocked by follow-up MIR-to-radio observations with better spatial resolution and depth, either from ALMA, the JVLA, or *JWST*-MIRI. I illustrate this point in Fig. 7.1, where I show a simulated patch of the sky as observed by current and future facilities. These simulated images are built from a catalog produced by EGG, a code that I developed and describe in Chap. 4, and that I will present in a paper to be submitted in the coming months. In this figure, look in particular at the two rightmost images that show the comparison between what we have today with *Herschel* and what ALMA will eventually provide after sufficient time investment.¹ While the observed wavelength is obviously different, ALMA can provide accurate “anchors” in the Raleigh–Jeans tail of the FIR SED. This is a priceless piece of information: not only does it allow direct measurement of dust (and gas) masses which, as shown in Chap. 5, are key quantities to study the way galaxies consume their gas content, but it can also be used together with MIR data (from either *Spitzer* or, soon, *JWST*-MIRI) to better interpret the *Herschel* images by pin-pointing which galaxies are the most likely counterpart of a given *Herschel* blob.

Star formation rates are only half of the story though. As written in Chap. 6, at high redshifts ($z > 3$) the dominant source of uncertainty is probably the stellar mass, owing to the poor coverage of the rest-frame near-IR which is redshifted into the *Spitzer* IRAC bands. While the most massive galaxies are indeed seen on these images, low mass galaxies are too faint to be detected. This is also illustrated in Fig. 7.1 (middle panel), where I compare the *Spitzer* IRAC image against a typical

¹Observing this region down to the depth I choose here would not require more than a few tens of minutes of integration time, but ALMA has a small field of view (about 20'' diameter) that make it much less efficient at achieving large area surveys.

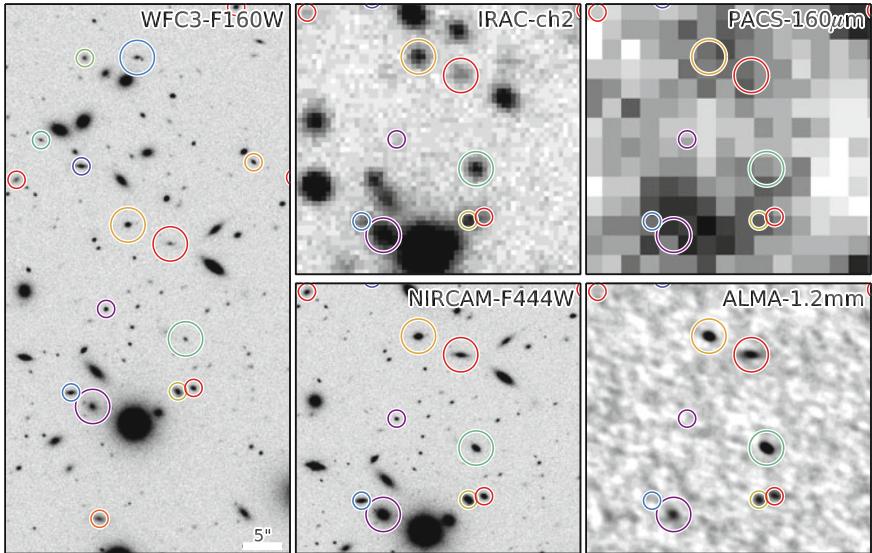


Fig. 7.1 Simulated images created from a catalog produced by EGG (Chap. 4). On all images, the *colored circles* show $z > 2$ galaxies, *smaller circles* being for galaxies less massive than $3 \times 10^{10} M_{\odot}$. The color is chosen randomly for display purposes and has no specific meaning. *Left Hubble WFC3-F160W* (*H* band, $1.6 \mu\text{m}$) at the depth of GOODS-South. Image created with *SkyMaker*. *Top, middle Spitzer IRAC ch2* ($4.5 \mu\text{m}$) at the depth of GOODS-South. Image created with *SkyMaker*. *Top, right Herschel PACS 160 μm* at the depth of GOODS-South. *Bottom, middle James Webb NIRCAM-F444W* ($4.4 \mu\text{m}$) with a similar exposure time as the *Hubble* image (depth estimate according to the *JWST* website), an image quality that will be obtained soon after the launch of the satellite (if only in limited regions of the sky). Image created with *SkyMaker*. *Bottom, right ALMA 1.2 mm* at a depth three times better than that achieved (or proposed) by current surveys. The image was created with *SkyMaker* and smeared by the observed ALMA beam from our ALMA survey in UDS (Chap. 6). I then added the noise manually by convolving a random Gaussian field with the corresponding dirty beam

JWST-NIRCAM image (the depth of which, a 5σ limit of 26.5 mag in $1''$ aperture, is estimated from the *JWST* website² assuming an integration time similar to that of *Hubble* in GOODS-South). The gain in precision, both in terms of depths and confusion, is substantial. I show here the NIRCAM-F444W image, but *JWST* is in fact equipped with a large array of filters that will provide good wavelength sampling of the NIR up to very high redshifts. Furthermore, the spectroscopic capabilities of the NIRSPEC instrument will allow for the first time the detection of H_{α} at $z > 2.5$. Not only will this eliminate the uncertainty linked to the photometric redshift and provide independent measures of the SFR, this will also provide direct quantification of the strength and equivalent width of these emission lines, which can heavily contaminate the broadband photometry at high redshifts (typically at $z > 5$), hence bias the stellar mass estimates.

²<http://www.stsci.edu/jwst/science/sensitivity>.

Both ALMA and *JWST* can therefore provide (directly or indirectly) better measurements of the SFR and M_* of galaxies at higher redshift than what *Herschel* and *Hubble* allowed me to do in this thesis, but also at intermediate redshift (say, $z = 1$) and lower stellar masses ($M_* < 5 \times 10^9$). This will allow us to explore further and more thoroughly the properties of the Main Sequence. At low redshifts, one important question that I could not address in a definite way in Chap. 2 is the connection between feedback and the scatter of the Main Sequence. Indeed, numerical simulations predict that supernova-driven feedback should have a large impact on low-mass galaxies (owing to their shallow gravitational potential) and create a substantial scatter in their SFR. However, this signature can only be observed in galaxies of very low stellar masses, which are difficult to observe today. *JWST* (and to some extent, ALMA) will change that.

Yet, it is true, measurements of this SFR– M_* correlation are abundant in the literature, and while it proved to be a useful tool to study the global history of star formation in the Universe, deeper answers are now to be found elsewhere, by studying more than just these two parameters that are the SFR and the stellar mass. For example, a natural follow-up of the work I presented in Chap. 5 is to study the variation of gas content and star formation efficiency not only *along* but also *within* the Main Sequence. Why is a galaxy found above the MS? We know that strong starbursts do have different star formation efficiencies (e.g., Daddi et al. 2010; Genzel et al. 2010; Béthermin et al. 2015), but what about *weak* starbursts? Are they just the interpolation between the Main Sequence and the strong starburst regime? Or is there really a bimodal process that only triggers strong starbursts? Is the answer varying with stellar mass, so that low-mass starburst are fundamentally different from their high-mass equivalent? And what about galaxies that are actually slightly *below* the Main Sequence? We have provided some hints in Chap. 5 that the SFE is the varying factor for massive galaxies, but what about less massive galaxies?

Nevertheless, to address these questions, it remains important to have a solid reference point, hence the need to properly calibrate the normalization and scatter of the Main Sequence, as well as using robust SFR and M_* measurements.

Lastly, a key question that cannot be addressed directly by observing the Main Sequence is to figure out how galaxies *leave* this sequence and die. The detection of massive quiescent galaxies has been reported as far back as $z = 4$ (e.g., Straatman et al. 2014), and the process through which these galaxies *quench* is not clearly identified. Many suspects are known, but the killer is still on the run.

One way to address this issue is to study the quiescent galaxies and look for clues of their recent past, e.g., by measuring their metallicity, morphology and gas content, and try to link these observations with the known profile of the suspects. For example, in our $z = 4$ ALMA survey, we targeted all the massive galaxies we could find, regardless of whether they are star-forming or quiescent. A good fraction ended up being non-detected. While it is likely that some of these are just not at the right redshift, there is also a good chance that we did observe some genuinely quiescent $z = 4$ objects. From there, there are several paths that one could follow. One could use ALMA to provide upper limits on the dust (and gas) content to see if these galaxies were quenched either by blowing out the gas or by preventing

its fragmentation. Alternatively, one could use *JWST*-NIRCAM to obtain detailed morphologies that even *Hubble*-WFC3 could not provide since these galaxies are often extremely compact. The interesting properties of $z = 4$ quiescent galaxies is that they must have quenched soon before they are observed, simply because of the age of the Universe (1.5 Gyr) and the time it takes to build a massive galaxy (1 Gyr to reach $M_* = 2 \times 10^{11} M_\odot$ with an exponentially rising star formation history of $\tau = 500$ Myr). Studying these objects can bring additional information that was lost in present-day elliptical, after several billion years of merging and subsequent relaxation.

References

- M. Béthermin, E. Daddi, G. Magdis et al., A & A **573**, A113 (2015)
E. Daddi, D. Elbaz, F. Walter et al., ApJ **714**, L118 (2010)
R. Genzel, L.J. Tacconi, J. Gracia-Carpio et al., MNRAS **407**, 2091 (2010)
C. Schreiber, M. Pannella, D. Elbaz et al., A & A **575**, A74 (2015)
C.M.S. Straatman, I. Labb , L.R. Spitler et al., ApJ **783**, L14 (2014)

Appendix A

phy++: A C++ Library for Numerical Analysis

A.1 Introduction

A.1.1 A Brief Overview

phy++ is a set of library and tools written in C++ that I developed during my PhD. The goal is to provide user-friendly vector data manipulation, as offered in interpreted languages like IDL¹, its open source clone GDL², or python & numpy³, but with the added benefit of C++: increased robustness, and optimal speed.

The library can be split into two components: the *core* library and the *support* library. The core library introduces the *vector* type, which is at the heart of *phy++*, while the support library provides functions and other tools to manipulate these vectors and do some common tasks, ranging from low level mathematics and programming (sorting, integrating, binning, ...) to higher level astrophysics-related tasks (such as cross-matching, stacking, SED fitting, ...). You can think of the core library as “the language” (the equivalent of IDL or python), and the support library as “the function library” (the equivalent of the `IDLASTRO`⁴, `numpy` or `astropy`⁵ libraries).

Below is an code sample written in *phy++* that illustrates its most basic functionalities.

¹<http://www.exelisvis.com/ProductsServices/IDL.aspx>.

²<http://gnudatalanguage.sourceforge.net/>.

³<http://www.numpy.org/>.

⁴<http://idlastro.gsfc.nasa.gov/>.

⁵<http://www.astropy.org/>.

```

vec2f img = fits::read("img.fits"); // read a FITS image
img -= median(img); // subtract the median of the whole image
float imax = max(img); // find the maximum of the image
vec1u ids = where(img > 0.5*imax); // find pixels at least half as bright
float sum = total(img[ids]); // compute the sum of these pixels
img[ids] = log(img[ids]/sum); // modify these pixels with a logarithm
fits::write("new.fits", img); // save the modified image to a FITS file

```

A.1.2 Why Write Something New?

The immediate goal of *phy++* is to provide a syntax as close as possible to that of IDL. IDL is an interpreted language that is widely used in the scientific community, in particular in astrophysics. Born in the late 1970s, this language provides intuitive manipulation of large arrays of data using vectorized operations: applying an operation on a given array does not require the user to write a loop to iterate over its elements and apply the operation. This leads to very concise code that is easy to write and read. Unfortunately, IDL suffers from a number of problems. I will start with the *political* and *ethical* problems.

- It is a proprietary, mostly⁶ closed-source program. This means that IDL is a black box and that people using it have no choice but to rely on the IDL developers for writing accurate code. While there is an extensive documentation, the algorithms used by the procedures are not always described. This is hardly acceptable for scientific code.
- IDL, like C++, combines several languages into one: a functional language and an object-oriented language. It also contains a huge support library providing many features (having used IDL for more than two years, I could not list them all). For this reason, and because it is proprietary, maintaining this language and adding new features costs a lot of money to its owner, Exelis. This money, in turn, is provided by science labs all around the world, who pay a yearly fee for a bunch of IDL licenses. This is totally fine in itself, but the fact is that most IDL users I have seen only make use of a small sub-set of IDL, one that has barely evolved in twenty years. In this context, the price that is paid is not justified.
- On top of that, the licensing model is that of *floating* licenses: only a fixed, maximum number of simultaneously running IDL instance is allowed in the whole lab. With the now common budget restrictions in research, labs typically buy fewer licenses than there are users. Even worse, it is often needed to run multiple instances of IDL on a single computer, e.g., when working on two projects simultaneously. This will consume two licenses, even though there is a single user. This leads to silly situations, typically when approaching specific deadlines (e.g., deadlines for requesting observing time on large telescopes) where everyone needs to use IDL at the same time, but there is not enough license available. Even worse, we have

⁶The procedures from the IDL library that are written in IDL language are actually open-source, but all the procedures written in native language are compiled and only the binary is provided.

seen cases in our lab of users being unable to run IDL on their new shiny computer because of incompatibility, not with IDL itself, but with the licensing software. Lastly, it should be noted that this licensing model relies on having network connection with a license server. This means that one cannot use IDL while traveling unless a proper SSH tunneling is in place.

These issues can be solved by switching to one of the free and open-source alternatives, like GDL. The downside is that these implementations are lacking behind IDL in terms of features, as some useful functions are still to be implemented. Worse, some functions cannot *legally* be implemented because they would violate IDL's copyright.

But that's only half of the story. Indeed, IDL and GDL also suffer from technical issues. I will list below the most important ones.

- Designed in the 1970s, IDL was born in an era where the available RAM was scarce, and that great care had to be taken to consume as few bytes of memory as possible. For this reason, the default integer type in IDL is a **short**, i.e., it occupies only two bytes in memory, while most languages (including some that are older than IDL itself) encode their integers on four bytes by default. The biggest issue with this choice is that the largest number one can store in a **short** is 32768. Being the default integer type, this creates quite a few surprises to the unexperienced user, and will fool even the expert from time to time.
- IDL is an interpreted language, meaning that the code you write is continuously read and interpreted by the IDL executable. While this is not an issue if you make good use of vectorization (the art of writing IDL code), performances are severely degraded once you write loops explicitly, because the content of the loop has to be *interpreted* and then *executed* on each iteration. And this is sometimes unavoidable.
- Like many interpreted languages, IDL is dynamically typed. This means that the type of a variable can change from one line to another, and that a variable containing a string can be assigned a number. While sometimes convenient, this comes at a cost: performance. Most IDL programs I have seen do not use this feature, yet they have to pay for it anyway.
- But worse than dynamic typing, and this is my main concern, variables in IDL are not *declared* before they are used. This means that if you do a typo in the name of one of your variables, chances are that the code will still run. Indeed, IDL cannot know that this was not intended, and will think that you want to create or modify a new variable. It will then do its best to carry on, and the result will be unpredictable. This, together with the fact that variables are almost not *scoped* (i.e., a variable created inside a **for** loop is still valid outside of the loop) makes it very easy to write confusing and buggy code. The most frightening part is that, in a good fraction of the cases, the output will be meaningful, and you can go on with your calculation never realizing that something went wrong. And publish that.

Avoiding the aforementioned issues is possible, but it requires coding with a fair amount of rigorousness and self discipline. My limited experience with astronomers

taught me that these are not particularly common character traits in the field, probably because we are all self taught programmers, but also because most of the code we write never goes out of our own computer and therefore does not get the chance to be reviewed or corrected by someone else. My conclusion is that, when it comes to checking the validity of a code, as much work as possible has to be done by the *language* itself (or its compiler), e.g., by being designed so that some errors cannot even be made, and that most of the remaining ones are identified *before* running the program and reported to the programmer so that he/she can fix them.

Switching to more modern interpreted languages like `python` or `Julia`⁷ would solve a few of these issues, in particular the first one. But the other items on this list are unfortunately inherent to most interpreted languages.⁸ To avoid these traps, the only solution today is to use *statically typed*, compiled languages, like C++.

Now, there are already some libraries in C++ that are addressing the topic of vector data manipulation. One can cite `Eigen`⁹ or the more recent `blaze-lib`.¹⁰ These are powerful libraries that have inspired *phy++* in some way, but their issue is that they are more oriented toward algebra, meaning that they have vectors and matrices, but no data type for arrays of higher dimensions (i.e., tensors¹¹).

Therefore, seeing that a gap had to be filled, *phy++* was created.

A.1.3 Why C++?

There are many different compiled languages that offer similar or better performances than C++. In particular, the most famous ones are Fortran and C. C is impractical to use because it has not been developed with user-friendliness in mind, and no mechanism exist to improve that. This is a system language, and it does that perfectly, but not much more. Fortran is known as the fastest of all, and it is particularly well suited for numerical analysis. While few languages are harder to read than Fortran 77, things have become much better since Fortran 90 (which is not used as often as it should be). However, Fortran is relatively bad at doing anything else than numerical analysis, which is annoying the moment you want to do something that is a bit off the tracks. C++ on the other hand, with all its disadvantages, is probably the best fit thanks to its almost unlimited capacity for adaptation. And it also happens to be the language I am most familiar with.

Since the beginning, C++ has always been good at performances, first because it is a language that compiles directly into assembler instructions, but also thanks to its philosophy: “you only pay for what you ask for”. But its main disadvantage is its *complexity*: it contains almost the whole C language, plus all the layers that

⁷<http://julialang.org/>.

⁸The best counter example is probably Java.

⁹<http://eigen.tuxfamily.org>.

¹⁰<http://code.google.com/p/blaze-lib/>.

¹¹Eigen actually has a tensor module, but it is unsupported.

were added on top of it, one year after another, starting from classes, exceptions, then templates. The end result is that it is a challenging task to master all the aspects of this language.

But the good news is: you do not have to master all of C++, and for your sanity you probably should not. Indeed, there are a number of *sub-languages* made out of a subset of C++ that are completely self-sufficient, i.e. you can use them to write any program. In other words, there are multiple, very different ways of writing the same program in C++. Typically, modern programs only use a small fraction of the whole language, e.g., leaving aside most of what was inherited from C (raw arrays, raw pointers, explicit memory management, etc.). A special class of such sub-languages are those that are tailored specifically to address a given task, as opposed to being open to any purpose. These are called *domain-specific languages* (DSL), and only require learning a few of C++’s rules and concepts, plus the rules introduced by the sub-language itself. The *phy++* library is an example of such domain-specific languages, its domain being vector data manipulation.

In short, although C++ is a very complex language, it is only necessary to learn a fraction of it to be able to use *phy++* correctly. Of course, the more one knows about C++, the more one will be able to take advantage of all the features of *phy++* in an optimal way.

A.1.4 Documentation

In this thesis, I do not include the library’s full documentation. I figured this would be pointless for one major reason: the library, although fairly mature, is still being conceived. New functions and features are added on a regular basis. Therefore, the documentation is still very much unstable, and if I was to include it here, it would become obsolete several months after the publication of this manuscript. Because it currently consists of more than a hundred pages, I realized this would be a waste of time and resources.

If you are interested, you can of course read the current, updated and full¹² documentation online. It is available either in a web-oriented format¹³ or as a compiled PDF document.¹⁴ I give in Fig. A.1 a screenshot and description of the web interface.

¹²Actually, at the time of writing this sentence, only half of the functions are documented.

¹³http://cscreib.github.io/phypp/doc/category_support_01_intro.html.

¹⁴<http://github.com/cscreib/phypp/raw/master/doc/latex/phypp.pdf>.

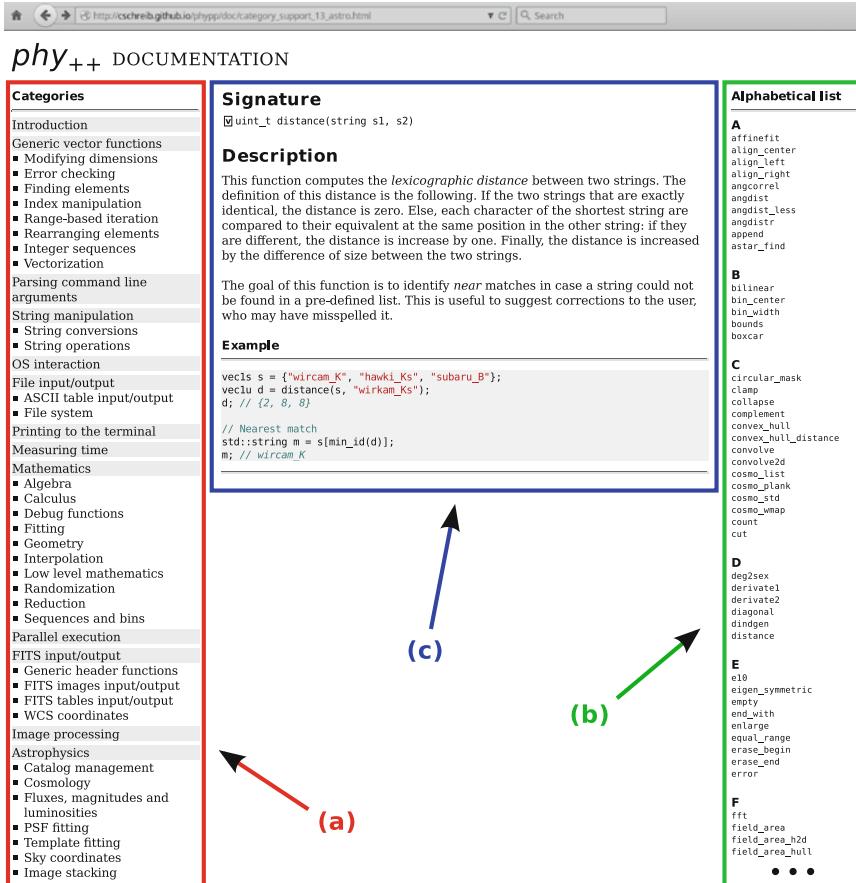


Fig. A.1 Example web page in the online documentation of the *phy++* library (http://cschreib.github.io/phypp/doc/category_support_01_intro.html). Three main areas are highlighted on this screenshot: **a** the *category* menu, where the functions of the library are grouped by themes and sub-themes to ease the discovery of new functions; **b** the *alphabetical* menu, which lists all the functions of the library by alphabetical order to allow quick access to the documentation of a known function; and **c** the central panel where the documentation is displayed, giving the *signature* of the function (i.e., what arguments it expects), a short descriptive text, and a code sample to illustrate the usage of the function

A.2 Application: **pixfit** and **gfit**

Using the *phy++* library, I have written most of the important codes involved in this thesis, for example the EGG tool that I introduce in Chap. 4. In this section I describe two other codes that I have written at the end of my PhD.

Most of the galaxies that we detected with ALMA (see Chap. 6) should be relatively bright in the *Herschel* SPIRE images. However, because of the poor angular resolution, interpreting these images is challenging. To obtain more precise flux estimations, I developed two programs, `pixfit` and `gfit`. These are still in the process of being tested, and I did not have time to reach a stable solution at the time of writing this manuscript. Still, I hope to be able to publish the codes in the near future. In the following, I give a brief description of the philosophy behind this novel approach, and postpone a more detailed assessment of the performances and robustness to a future work.

Conventional tools used to extract FIR fluxes (like FASTPHOT, Béthermin et al., 2010) perform point-source fitting at various pre-determined positions of the image simultaneously using linear algebra, assuming that the noise of the image is Gaussian. If there is no strong overlap between two extracted objects (or, alternatively, if the positions of the emitting sources are known perfectly), the resulting fluxes and error estimates have been shown to be reliable (see, e.g., Wang et al. in prep.). However, extracting fluxes in the highly confused SPIRE maps remains a challenge, since most objects are blended. In Wang et al. (in prep.), the situation is improved by bringing additional prior information on the expected fluxes of the faintest galaxies, but this comes at a price: the output flux catalog becomes model dependent. Even then, the number of SPIRE 500 μm sources extracted in a typical *Herschel* deep field does not exceed a hundred, compared to the thousands of MIPS 24 μm detections that we know are contributing, to some extent, to the observed 500 μm emission.

The approach that I chose with these new tools is to think of the flux catalog as only an intermediate product in the chain of data analysis: what we have in input is an observed map, and what we want in output is a catalog of SFR, L_{IR} , or M_{dust} . In fact, the flux catalog is only a translation of the observed map into a format that is easier to manage, but the issue is that this translation, as I argue above, is not unique. In most cases, we do not know what fraction of a given 500 μm flux should be attributed to this or that galaxy, and building a flux catalog requires making assumptions (e.g., “the brightest galaxy at 24 μm will be the brightest at 500 μm ”).

However, if we give up on the idea of building a conventional flux catalog, where each galaxy has either its own flux or no flux at all, one can get rid of these assumptions. For example, the idea behind `pixfit` and `gfit` is the following: for galaxies that are too close to one another on a given image (e.g., the SPIRE 500 μm map), I give up measuring their individual fluxes, and combine them into a single “flux group”, for which I can measure the total flux accurately (e.g., with aperture photometry after subtracting the neighboring sources). In this case, “too close” can be defined arbitrarily, for example by choosing a given fraction of the width of the PSF,

or a fraction of pixels on the rasterized image.¹⁵ The measured flux is then stored into a separate list, and each galaxy that belongs to the group is linked to this measurement. This first task of extracting the fluxes and making the flux groups is performed by `pixfix` on each FIR image independently. In particular, this means that two galaxies can be grouped in one image, where the angular resolution is poor, but not in another, where the resolution is sharper. This is made in a fully automatic way, by just specifying in input a list of prior positions, and defining the distance threshold below which two sources must be grouped. An example is show in Fig. A.2.

The output of this procedure is very similar to a conventional flux catalog, since each galaxy *can* have its individual flux extracted from each image, provided that it was not grouped with any other galaxy. If this is not the case, then for each band there is an additional column that indicates the ID of the flux group that contains the flux of this galaxy, and a second catalog is built to store these flux groups. It contains four columns: the group ID, a reference to the image this group was extracted from, the extracted flux and the associated uncertainty.

The next step is to properly interpret this data. Standard SED fitting codes assume that one has access to individual flux measurements in all bands, and these codes do not know how to deal with the flux groups I introduced above. Some particular codes can be given upper limits in case of a non-detection, but treating these in a statistically correct way is not trivial, and requires non-linear fitting algorithms. Indeed, while the likelihood associated to a measurement is a Gaussian, that associated to an upper limit is an error function. Therefore, the contribution of an upper limit to the χ^2 is:

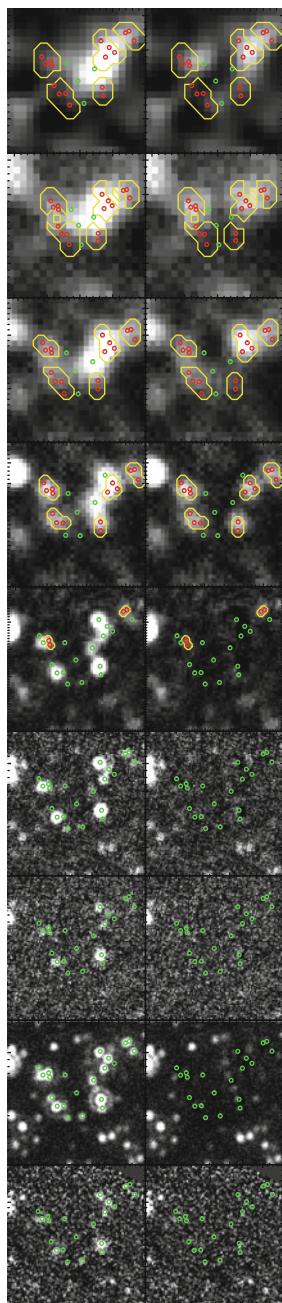
$$\chi^2 = -\frac{1}{2} \log \left[\frac{1}{2} \left(1 + \operatorname{erf} \left(\frac{\text{limit} - \text{model}}{\text{error} \sqrt{2}} \right) \right) \right], \quad (\text{A.1})$$

where *limit* is the estimated upper limit, *model* is the attempt at modeling the corresponding flux, and *error* is the uncertainty on the upper limit.¹⁶ If a galaxy is grouped in an image, the flux of the corresponding group can be used as an upper limit. As written at the beginning of this section, not only is this suboptimal, but this approach is also incorrect since each galaxy will be fitted independently. Indeed, while the upper limit will ensure that no individual model goes above the flux present on the map, there is no constrain on the *sum* of all the model fluxes: if the measured flux on

¹⁵ Actually a similar approach is used in the extraction code of (Magnelli et al., 2009), where sources that are distant by less than a pixel are not fitted individually. The main difference with the approach I introduce here is that only one of their galaxies is kept in the prior list and arbitrarily “wins” all the observed flux.

¹⁶This expression is numerically unstable for large deviations above the upper limit. Setting $d \equiv (\text{limit} - \text{model})/\text{error}$, then for $d < -3$, this formula can be approximated with good accuracy by $d^2 + 2 \log(-2 d \sqrt{\pi}/2.0)$. Note the similarity with the regular formula for a Gaussian weight, which is just d^2 .

Fig. A.2 Example application of *pixfit* in GOODS-South. In the top row are the observed images. From *left to right*: *Spitzer* 16 and 24 μm , *Herschel* PACS 70, 100 and 160 μm , *Herschel* SPIRE 250, 350 and 500 μm , and LABOCA 870 μm . Each postage stamp covers the same region of the sky. The *bottom row* are the same image after subtracting the galaxies that have individual flux measurements, leaving only the fluxes of the groups. Each *open circle*, whether *green* or *red*, is a prior position used to extract the fluxes. *Green circles* are galaxies that have an individual flux measurement, while *red circles* show galaxies that were grouped with their neighbors for being too closely packed. A *yellow contour* indicates the extent of the corresponding flux group, and the area that is used to perform aperture photometry



the map is 20 mJy, and we use this value as an upper limit for two galaxies that lie in this region, then each galaxy can reach 20 mJy individually, for a combined flux of 40 mJy that will clearly overshoot what is observed.

That is where the `gfit` tool comes in. This program understands the catalogs produced by `pixfit`, and can perform SED fitting of multiple galaxies simultaneously. In particular, if two galaxies have some of their flux grouped, the program will model these fluxes individually, sum them up, and compare the result to the measured flux of the group in the χ^2 , like any regular measurement. The fit can then be made using linear algebra, and is therefore very fast.

This main feature of performing simultaneous SED fitting is a double-edged sword though. The major downside is that if I have 100 templates in my SED library (e.g., corresponding to different values of T_{dust}), finding the optimal χ^2 requires testing each and every possible combination of templates for all the galaxies in the group, and each additional galaxy increases the computation time by a factor of 100. Obviously, this means that the problem can become computationally prohibitive. To avoid this, I first sample the parameter space of the library with a coarse grid, say of only 10 templates. I locate the combination of SEDs that produces the best χ^2 , and refine the grid around this region with 10 more templates. With this approach, the accuracy on the best-fit parameters is unchanged, but the complexity drops from 100^N to 2×10^N . Without a super computer, this can still be too much if the prior density is too large. In practice though, I never had to fit more than 6 galaxies simultaneously in a given group, although I have only applied this method to a handful of cases. This problem can also be tackled with more sophisticated algorithms for global minimization, but I have not investigated this path any further.

At present, both tools are written and are feature complete. I have tested them on some of our ALMA detections, trying to better constrain their SEDs. The results seemed reasonable, but these tools really have to be tested on simulated images before any output can be trusted. I will do this later, when time permits.

Below is a excerpt from the code of `pixfit`, to illustrate how the *phy++* library looks like in a “real world” situation.

```

// [...]

// First build the masks of each group.
// We want to make sure that the same pixels are not counted twice in different
// groups, so we have to exclude the regions where two (or more) groups overlap.
for (uint_t i : range(group_cat.ra)) {
    if (group_cat.fit[i]) {
        // This is a 'group_fit'
        // We don't need to care about it any more.
        continue;
    }

    // Locate the sources that are part of this group
    vec1u id = where(old_cat.group_aper_id == group_cat.id[i]);
    phypp_check(!id.empty(), "aper group ", group_cat.id[i], " is empty...");

    // Extract just what we need from the whole map
    uint_t xmi = max(0, floor(min(tx[id]) - group_aper_size));
    uint_t xma = min(img.dims[1]-1, ceil(max(tx[id]) + group_aper_size));
    uint_t ymi = max(0, floor(min(ty[id]) - group_aper_size));
    uint_t yma = min(img.dims[0]-1, ceil(max(ty[id]) + group_aper_size));

    vec2i tgrp = grp_map(ymi--yma, xmi--xma);

    // Convert coordinates to the local map
    tix[id] -= xmi;
    tiy[id] -= ymi;

    // Build the aperture mask
    vec2b mask(tgrp.dims);
    for (uint_t j : id) {
        // Create the aperture for this source
        vec2b taper = translate(aper, tdy[j], tdx[j]) > 0.5;
        vec1u idi, idp;
        subregion(
            mask,
            {tiy[j]-hsize, tix[j]-hsize, tiy[j]+hsize, tix[j]+hsize},
            idi, idp
        );

        // Add it to the mask
        mask[idi] = mask[idi] || taper[idp];
    }

    // Flag pixels that already belong to another group
    vec1u ido = where(tgrp != 0 && mask);
    tgrp[ido] = -1;
    // Remove these pixels from the mask
    mask[ido] = false;
    if (count(mask) == 0) {
        warning("group ", group_cat.id[i], " has empty mask");
    }

    // Set pixels of this group
    tgrp[where(mask)] = group_cat.id[i];
    // Save back in the whole map
    grp_map(ymi--yma, xmi--xma) = tgrp;
}

// [...]

```

References

- M. Béthermin, H. Dole, A. Beelen, H. Aussel, A&A **512**, 78 (2010)
B. Magnelli, D. Elbaz, R.R. Chary et al., A&A **496**, 57 (2009)

Index

A

Active galactic nucleus, 20, 38, 41
Aperture correction, 33
Atmospheric transmission, 15

B

Balmer break, 14
Baseline (interferometric), 169
Birth rate parameter, 6
Black body, 14
Black hole (supermassive), 20
Blue cloud, 24
Bulge (galactic), 21, 24

C

Cold dust, 39
Cold flow (gas), 31
Color diagram, 34, 40
Cosmic infrared background, 16, 47

D

Depletion timescale, 5
Diffraction, 15
Disk (galactic), 5, 10, 11, 24
Dust absorption, 13, 38
Dust grain, 13
Dwarf galaxy, 4

E

Elliptical galaxy, 21
Extinction curve (dust), 17

F

Feedback (AGN), 20, 70
Feedback (stellar), 19, 70

G

Gas fraction, 25

I

Infall (cosmological), 5, 20, 31
Initial mass function, 38, 39
Intergalactic medium, 5
Interstellar medium, 13, 31

L

Luminosity function, 30
Luminous infrared galaxy, 30
Lyman break, 14

M

Main Sequence (galaxies), 6, 23, 30
Median absolute deviation, 51
Merger, 70
Merger (major), 5, 30
Metallicity, 17

N

Neutron star, 61

P

Point spread function, 33
Polycyclic aromatic hydrocarbon, 30
Primary beam, 171

Q

- Quasi-stellar object, 20
- Quenching (AGN), 20
- Quenching (gravitational), 23
- Quenching (halo), 23
- Quenching (morphological), 21
- Quenching (stellar), 19
- Quiescent galaxy, 21, 39

R

- Radiation pressure, 19
- Radiative-mode (AGN), 20
- Radio jet, 20
- Radio-mode (AGN), 20
- Red cloud, 24
- Redshift (photometric), 37
- Redshift (spectroscopic), 37

S

- Scatter stacking, 51
- Schechter function, 44
- Source confusion, 15
- Specific star formation rate, 6
- Spectral energy distribution, 31
- Spectral energy distribution fitting, 38
- Spectral slope, 16, 24

Spiral arm, 5

- Stacking (image), 18, 45
- Star formation efficiency, 25
- Star formation history, 4, 30, 38
- Star formation rate, 4, 38
- Starburst (galaxy), 10, 11, 31, 42
- Starburstiness, 66
- Stellar lifetime, 13, 61
- Stellar mass, 5, 37
- Stellar-mass completeness, 33, 42
- Stellar remnants, 61
- Stellar winds, 19

T

- Tapering, 171
- Thermal radiation, 13, 38
- Tidal tails, 8
- Toomre criterion, 8, 21

U

- (u, v) plane, 169

W

- White dwarf, 61