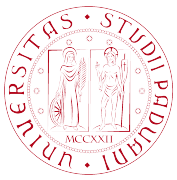

UNIVERSITÀ DEGLI STUDI DI PADOVA

UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia "Galileo Galilei"

SCUOLA DI DOTTORATO DI RICERCA IN ASTRONOMIA
XXIV ciclo
Sede Amministrativa: Università degli Studi di Padova

Cosmological Surveys with Spitzer & Herschel Space Observatories: Luminosity Functions and Cosmological Evolution

Dottoranda : Lucia Marchetti

Supervisore
Ch. mo Prof. Alberto Franceschini
Dip.to di Fisica e Astronomia
Università di Padova

Co-Supervisore
Dr. Mattia Vaccari
Dip.to di Fisica e Astronomia
Università di Padova

Direttore della Scuola
Ch. mo Prof. Giampaolo Piotto
Dip.to di Fisica e Astronomia
Università di Padova

Contents

List of Figures

List of Tables

Sommario

Le lunghezze d'onda infrarosse contengono delle informazioni fondamentali per descrivere l'origine delle galassie, dei nuclei galattici attivi e, più in generale, la storia evolutiva della formazione stellare, della produzione dei metalli e dell'accrescimento gravitazionale. In sostanza quindi, rappresentano un punto di vista del tutto complementare a quello investigato dalle più classiche *survey* nell'ottico.

Tuttavia, i vari processi fisici che si susseguono nella formazione galattica lasciano la loro impronta sulla forma dello spettro di emissione globalmente a tutte le lunghezze d'onda, ciascuno più o meno dominante nei diversi regimi di frequenza. Per questo motivo solo tramite dettagliati studi multi banda sulle proprietà delle galassie siamo in grado di capire approfonditamente la nature di questo tipo di sorgenti.

In questo contesto, un grosso passo avanti è stato compiuto nell'ultima decade, grazie all'avvento di nuovi telescopi e di grandi *survey* osservative condotte a tutte le lunghezze d'onda, consentendo quindi, per la prima volta, un reale studio completo, dai raggi X al Radio, delle SED di un significativo numero di galassie. In particolare, gli osservatori spaziali Spitzer ed Herschel stanno dando un grosso contributo in questo senso, permettendo, per la prima volta, osservazioni nel medio-, lontano- infrarosso e nel regime sub-millimetrico campionando diverse lunghezze d'onda.

In questo lavoro sfruttiamo diverse *survey* multibanda condotte in maniera congiunta da Spitzer e Herschel con il fine di studiare le funzioni di luminosità nelle bande del medio-, lontano-infrarosso e del sub-millimetrico e determinare poi l'evoluzione delle sorgenti infrarosse.

Più nel dettaglio abbiamo basato la nostra analisi sui dati raccolti nelle *survey* sia profonde che su grandi aree, e quindi meno profonde, condotte da Spitzer, e nella *Herschel Multi-tiered Extragalactic Survey* (HerMes, Oliver et al. (2012)) condotta da Herschel. Queste campagne osservative includono le osservazioni più profonde e più vaste in area mai condotte a queste lunghezze d'onda, analizzando il per lo più inesplorato intervallo di lunghezze d'onda infrarosse fino a 500 μm in aree di cielo che vanno dai 0.01 gradi^2 ai 10 gradi^2. Inoltre, nella maggior parte delle aree da loro coperte, è stato possibile sfruttare un considerevole numero di campagne osservative sia fotometriche che spettroscopiche consentendo di raccogliere informazioni anche nelle restanti bande dello spettro elettromagnetico dal lontano ultravioletto al vicino infrarosso (per esempio, GALEX nel lontano e vicino UV, SDSS, INTWFS, CFHTLS e altre nell'ottico, 2MASS e UKIDSS nel vicino IR) ed essere così in grado di caratterizzare dettagliatamente le proprietà fisiche delle sorgenti rilevate.

Una parte del presente lavoro è stata dedicata alla costruzione e alla validazione del database denominato Spitzer-Selected Multi-Wavelength Wide-Area Data Fusion. In questo database abbiamo combinato i dati nel medio- e lontanto-IR raccolti da Spitzer nelle campagne

osservative *Spitzer Wide-area InfraRed Extragalactic* (SWIRE, Lonsdale et al. (2003)) in sei aree di cielo, *Spitzer Deep-Wide Field Survey* (SDWFS) nell'area di Bootes, *Spitzer Extragalactic First Look Survey* (XFLS) nell'area di XFLS, assieme ai dati nell'UV, nell'Ottico e nel vicino IR, campionando circa 70 gradi^2 in totale. Abbiamo sfruttato al massimo l'ampiezza e la ricchezza dei dati raccolti da SDSS e INTWS nell'ottico e da 2MASS e UKIDSS nel vicino-IR, così come altre osservazioni ottiche ottenute all'interno dei progetti SWIRE, SWDFS e XFLS. Abbiamo così aggiunto a questa già ricca quantità di dati tutte le informazioni spettroscopiche disponibili da SDSS e NED.

Il Spitzer Data Fusion rappresenta un ideale punto di partenza per studi sulle proprietà statistiche delle galassie, come lo studio dettagliato delle distribuzioni di energia spettrale (SED) per la stima dei redshift fotometrici, delle masse e del tasso di formazione stellare (SFR). Estato usato per validare le osservazioni condotte da Herschel in SDP all'interno del consorzio HerMES e per produrre la prima versione pubblica dei cataloghi SPIRE. Estato pensato per essere compatibile con le piattaforme rese disponibili dal VO e sarà reso pubblico assieme ad una collezione di altre quantità ancillari come le curve di trasmissione e gli indici di conversione tra magnitudini Vega ed AB per tutti i filtri disponibili.

Lo studio delle funzioni di luminosità è spesso reso più complicato a causa, da un lato, delle difficoltà di condurre osservazioni su grandi aree fino a bassi livelli di flusso, dall'altro delle difficoltà di associare una corretta stima del redshift alle sorgenti man mano osservate. In questo lavoro abbiamo perciò assemblato le osservazioni più profonde e su grandi aree di Spitzer ed Herschel con lo scopo di selezionare un campione di galassie infrarosse nel modo più completo e affidabile, così come la migliore collezione collezione di dati ancillari UV/Ottico/NIR atti ad identificare le controparti di queste sorgenti e determinarne una stima corretta del redshift. Grazie alle osservazioni di Spitzer ed Herschel siamo infatti in grado, per la prima volta, di campionare la luminosità bolometrica infrarossa integrata tra 8 e 1000 μm delle sorgenti ad ogni epoca cosmica e quindi consentire di studiare nel dettaglio i processi di formazione stellare oscurati dalla polvere del mezzo intergalattico in funzione del tempo cosmico.

Ad ogni modo, anche avendo a disposizione la migliore collezione di dati, studiare le funzioni di luminosità delle galassie rimane complicato e ricco di aspetti controversi, a causa della presenza di effetti di selezione sulle osservazioni dovute, per esempio, al limite di sensibilità degli strumenti che rendono per lo più incompleti i campioni di galassie selezionati e introducono quindi dei *bias* nella stima corretta delle funzioni di luminosità. È quindi chiaro che solo tramite un confronto diretto delle stime ottenute utilizzando diversi metodi statistici sugli stessi campioni di dati che possiamo essere sicuri di quantificare l'impatto di questi effetti di selezione sui nostri risultati.

Grazie alla Spitzer Data Fusion, siamo in grado di descrivere le funzioni di luminosità locali ($0 < z \lesssim 0.5$) delle sorgenti selezionate nei campioni di grande area raccolti dai fotometri MIPS e SPIRE, rispettivamente operanti su Spitzer ed Herschel. Abbiamo sfruttato ampiamente le informazioni multi banda raccolte nel database per analizzare le SED delle sorgenti rilevate da questi strumenti e quindi stimare, nel sistema di riferimento a riposo, le luminosità monocromatiche sia nelle bande a MIPS 24/70/260 μm che in quelle a SPIRE 250/350/500 μm e la luminosità infrarossa bolometrica integrata tra 8 e 1000 μm per ciascuna sorgente. Successivamente abbiamo utilizzato svariato metodi statistici per stimare la funzione di luminosità locale di campioni completi di galassie selezionate in flusso per ciascuna di queste bande: il classico metodo $1/V_{max}di$Schmidt (1968) e una sua incarnazione leggermente modificata detta $1/V_{est}$ introdotta da Page & Carrera (2000); un metodo parametrico di *maximum*

likelihood basato sulla statistica Bayesiana descritto da Kelly et al. (2008) ed infine un metodo semi-parametrico introdotto da Schafer (2007).

Abbiamo poi applicato gli stessi metodi ad un campione di dati raccolti dai due fotometri PACS e SPIRE, di Herschel, distribuiti dal consorzio HerMES e centrati nell'area della survey *Cosmic Evolution Survey* (COSMOS), osservazioni in cui la sensibilità di PACS e SPIRE a il catalogo di redshift fotometrici di COSMOS ci consentono di studiare l'evoluzione delle sorgenti di Herschel fino a $z \sim 5$. Nello specifico abbiamo studiato, in funzione del redshift, le funzioni di luminosità monocromatiche, la funzione di luminosità bolometrica infrarossa integrata tra 8 e 1000 μm e la densità del tasso di formazione stellare cosmica delle sorgenti rilevate da SPIRE a 250 μm. Data per assodata la presenza di incertezze nelle stime sopra citate, abbiamo applicato due metodi indipendenti per misurare il tasso evolutivo delle sorgenti, ottenendo risultati consistenti tra loro e quindi rendendo ancora più solide le nostre stime. I nostri risultati pongono dei vincoli stringenti sulle predizioni modellistiche dei modelli semi-analitici che descrivono la formazione della galassie sulla base di principi primi in accordo con lo scenario cosmologico standard, e dai nostri confronti emerge un sostanziale disaccordo tra i risultati osservativi da noi ottenuti e quelli predetti dai modelli nel quantificare il tasso evolutivo delle sorgenti rilevate da Herschel, soprattutto ad alto redshift.

Summary

Infrared wavelengths contain a substantial amount of information about the origin of galaxies and active galactic nuclei and about the evolutionary history of star formation, metal production and gravitational accretion. They present a widely complementary view with respect to more classical surveys in the optical.

However, the different physical processes occurring in galaxies all leave their imprint on the global and detailed shape of the spectrum, each dominating at different wavelengths. Thus only detailed analysis of the multi-wavelength properties of a galaxy should, in principle, allow us to fully understand the nature of that object.

In this context, a major development in the last decade has been the advent of new observing facilities and large surveys at all wavelengths of the spectrum, enabling astronomers for the first time to observe the full Spectral Energy Distributions (SEDs) of significant samples of galaxies over a large redshift range at wavelengths from the X-rays to the Radio. In particular Spitzer and Herschel have been giving a great contribution, allowing, for the first time, the sampling of the MIR, FIR and SMM part of the spectrum at multiple wavelengths.

In this work we exploited a number of multi-wavelength extragalactic surveys in fields jointly observed by Spitzer and Herschel to determine the IR galaxy luminosity function at MIR, FIR and SMM wavelengths and study the evolution of IR sources with cosmic time.

We used data in areas jointly covered by Spitzer Deep & Wide Extragalactic Surveys and the *Herschel Multi-tiered Extragalactic Survey* (HerMes, Oliver et al. (2012)). These surveys include the deepest and widest observations ever carried out at these wavelengths, probing the almost unexplored IR wavelength range up to 500 μm in fields of sizes from 0.01 deg^2 to 10 deg^2. Besides, over the better part of their covered areas, it was possible to exploit numerous photometric and spectroscopic surveys in order to completely cover the wavelength range from the FUV to the NIR (e.g. GALEX in the FUV/NUV, SDSS, INTWFS, CFHTLS and others in the Optical, 2MASS and UKIDSS in the NIR) and thus characterize the physical properties of detected sources in great detail.

Part of the present work has been devoted to the construction and validation of the so called Spitzer-Selected Multi-Wavelength Wide-Area Data Fusion. In this database we combined Spitzer mid- and far-infrared data from the *Spitzer Wide-area InfraRed Extragalactic* (SWIRE, Lonsdale et al. (2003)) survey in six fields, the *Spitzer Deep-Wide Field Survey* (SDWFS) in the Bootes field, the Spitzer Extragalactic First Look Survey (XFLS) in the XFLS field, with data at UV, Optical and NIR wavelengths, covering about 70 deg^2 in total. We fully exploited the power of the SDSS and INTWFS databases, as well as complementary near-infrared data from 2MASS and UKIDSS and further optical imaging obtained by the SWIRE, SDWFS and XFLS teams. We then added to this rich dataset all the spectroscopic information available from SDSS and NED.

The Spitzer Data Fusion represents an ideal starting point to perform statistical studies on galaxies like detailed SED fitting analysis to estimate photometric redshifts, masses as well as the Star Formation Rates (SFRs).It has been used to validate the Herschel SDP observations within the HerMES consortium team and to produce the first release of the official SPIRE catalogues. It was developed to be VO-compliant and will be made publicly available, together with a collection of ancillary information such as the transmission curves and Vega-to-AB conversion factors of all spectral filters.

The determination of the Galaxy Luminosity Function is often hampered by the difficulties of covering a wide area down to faint fluxes on one hand and determining counterparts and their redshifts for detected sources in a complete and reliable manner on the other. In this work we have thus assembled and exploited the deepest and widest Spitzer and Herschel extragalactic surveys to select IR galaxy samples in a complete and reliable manner, and the best UV/Optical/NIR ancillary data to identify them and estimate their redshifts. Thanks to Spitzer and Herschel observations we are now able, for the first time, to reliably sample the IR bolometric luminosity of sources at virtually any epoch and thus provide important insights into dust obscured star formation activity across Cosmic Time. Even with the best dataset, however, accurately constructing the Luminosity Function remains a tricky pursuit, since the presence of observational selection effects due to e.g. detection thresholds in apparent magnitude, colour, surface brightness or some combination thereof can make any given galaxy survey incomplete and thus introduce biases in the Luminosity Function estimates. Only a comparison of results coming from different Luminosity Function estimators applied on the same samples can ensure we can assess the impact of these biases in a robust manner.

Armed with the Spitzer Data Fusion, we were able to describe the $0 < z \lesssim 0.5$ Local Luminosity Function of sources selected in wide fields by the Spitzer MIPS and Herschel SPIRE imagers. We fully exploited the multi-wavelength information collected within the Spitzer Data Fusion to perform a complete SED fitting analysis of MIPS and SPIRE sources and thus estimate the monochromatic rest-frame luminosities at both MIPS 24/70/260 μm and SPIRE 250/350/500 μm as well as the IR bolometric luminosity integrating the SED between 8 and 1000 μm. We then implemented a number of different statistical estimators to evaluate the local luminosity functions of flux-limited samples in these bands: the classical $1/V_{max}$ estimator Schmidt (1968) and the modified $1/V_{est}$ version by Page & Carrera (2000); a parametric maximum likelihood technique (ML) based on a bayesian approach as described in Kelly et al. (2008) and finally a *semi-parametric* approach introduced by Schafer (2007).

We have then applied the same tools to the PACS/SPIRE dataset provided by the PEP/HerMES teams in the *Cosmic Evolution Survey* (COSMOS) field, where PACS/SPIRE sensitivity and the COSMOS photometric redshift catalog allowed us to probe the evolution of Herschel sources up to $z \sim 5$. Namely, we have studied the redshift-dependent FIR/SMM Monochromatic and IR Bolometric Luminosity and the inferred Cosmic Star Formation Rate Density of SPIRE 250 μm sources. Given the uncertainties at play, two independent methods to measure the evolutionary rates of Herschel sources were applied, returning consistent results and thus strengthening our conclusions. Our results put strong constraints on modeling predictions, and their comparison with semi-analytical models for the formation of galaxies into a self-consistent cosmological context starting from first principles readily shows substantial disagreements in the observed/predicted evolutionary rates of Herschel sources at moderate-to-high redshifts.

Introduction

The determination of the star formation history of the Universe has been a major thrust to observational cosmology and galaxy evolution studies for decades. Understanding how the structures that we observe in the Local Universe came to be is in fact a fundamental step toward understanding how our Universe has formed and how it evolves along cosmological timescales.

For all the successes of ground-based astronomy, however, the Earth's atmosphere has always been a formidable barrier to astronomers. The efforts of casual star gazers are often frustrated by clouds, and for professional astronomers the atmosphere is a serious handicap because, even on a clear night, the air above our heads blocks most of the radiation which arrives from space and distorts what little information does reach the ground. It was thus only following the X-ray pioneering studies of 1960s and 1970s and with the establishment of a coordinated long-term program of space astronomy missions, mainly by NASA and ESA, that the richness of the full electromagnetic spectrum has been made available to astronomers. Thanks to an ever more sophisticated technology, we have been able to develop a new generation of space telescopes which opened new windows on the sky and revolutionized the way in which mankind views the Universe.

One of the first demonstrations of what was in store for space astronomers at largely unexplored wavelengths came after the launch in 1983 of NASA's Infrared Astronomic Satellite (IRAS). The all-sky survey carried out by the satellite provided a first census of the Universe in a number of infrared bands and discovered a new type of sources which were exceedingly bright at these wavelengths. These objects, mostly connected with enhanced star formation activity (or starburst) episodes, had infrared luminosities which could typically be in excess of $L \geq 10^{12} L_{\odot}$, and were shown to evolve very rapidly over the limited cosmological volumes probed by IRAS deeper observations.

After the great success of IRAS, in the 1990s NASA's Cosmic Background Explorer (COBE) and ESA's Infrared Space Observatory (ISO) combined their observational capabilities to shed light on the higher-redshift infrared Universe. In particular, COBE was built with the aim to measure the cosmic infrared background. In so doing, it discovered that the cosmic background light from galaxies is dominated by two peaks: the first peak contains the power output from the ultraviolet and the optical, dominated by the emission from hot young stars, while the second peak arises at infrared to sub-millimeter wavelengths. This was an extremely important discovery as it implied that around half of all the star formation activity in the Universe, characterized by energetic phenomena radiating in the ultraviolet band, is obscured by dust and thus absorbed and later re-emitted, mostly in the infrared range. The observed intensity of such a background light also implied that the strong evolution observed by IRAS in the local Universe continued up to higher redshifts. This was soon confirmed

by ISO deeper surveys, which probed individual sources up to $z \sim 1$ and started to clarify the physical properties of the sources making up the cosmic infrared background. On the whole, this meant that, if we really wanted to understand the first stages of galaxy formation, when a very rich interstellar medium was present, we had to observe more carefully at these wavelengths.

Some fifteen years after COBE's launch, we are leaving a period of uttermost interest in infrared astronomy. NASA's Spitzer Space Telescope, launched in 2003 and whose operational life has been extended at least until 2015, has given us the possibility to study the Universe at MIR and FIR wavelengths up to $z \sim 5 - 6$. Spitzer observations have greatly expanded our knowledge of most astrophysical phenomena, from the evolution of distant galaxies and AGNs to the composition and effects of the interstellar medium.

The latest great space observatory, ESA's Herschel, is also optimized for long-wavelength observations, and extends our observing capabilities into the far-infrared and sub-millimeter. Its instrumentation is optimized for studying the formation and the evolution of stars and galaxies. In particular the Herschel Multi-tiered Extragalactic Survey (HerMES) and the PACS Extragalactic Probe (PEP) Key Programs provide us an impressive amount of data to disentangle issues related with dust-obscured star-formation in the early Universe.

In this context of ever deeper surveys at most wavelengths, it is even more difficult and important to reliably measure galaxy properties in the Local Universe: difficult because the very possibility to carry out extremely deep observations leads to most observing time being spent on the deepest pencil-beam surveys rather than shallower wider-area ones, and important because the increasingly detailed knowledge of the high-redshift Universe needs similarly well-defined local benchmarks to trace the formation and evolution of galaxies across cosmic time in great detail. Perhaps more importantly, in the era of multi-wavelength surveys and virtual observatories, shallow wide-area surveys with large data rates are likely to profit the most from the paradigm shift caused in astronomical research by the easy access to a number of otherwise separate databases for science exploitation.

This work capitalizes on the above trends by exploiting Spitzer and Herschel major survey projects and putting them in a multi-wavelength context. Spitzer catalogs are combined with shorter-wavelength photometry and spectroscopy and are then used to enhance the source extraction on Herschel maps and to better detail the SEDs of Herschel sources. The accurate SED fitting process of complete Herschel maps is used to determine the evolution of the Luminosity Function at FIR/SMM wavelengths and thus place stronger constraints on models for the formation and evolution of infrared galaxies.

Chapter 1

Galaxy Formation & Evolution Scenario

For several decades cosmologists have considered the determination of the density parameter and the expansion rate of the Universe their prime task. Now that this goal has seemingly largely been achieved (see Sec. 2.1) the most pressing issues in cosmology is turned to be understand how the Universe evolved from a very primitive initial state into what we are observing around us today - galaxies of different morphologies, the large scale structure of their distribution, clusters of galaxies, and active galaxies. We seek to study the formation of stars and of metals, and also the processes that reionized the intergalactic medium.

The boundary conditions for studying these processes are now very well defined. A few years ago, the cosmological parameters in models of galaxy evolution, for instance, could vary freely because they had not been determined sufficiently well at that time. Today, a successful model needs to come up with predictions compatible with observations, but using the parameters of the standard model. There is little freedom left in designing such models. In other words, all the ingredients are prepared and what we need is to find the fundamental agreement between models and observations in order to complete the galaxy formation and evolution scenario.

The fundamental driver of progress in cosmology is then through observations. The advent of large galaxy surveys, either wide spectroscopic surveys probing the nearby Universe (e.g., SDSS) or narrower surveys using photometric redshifts and often in the IR domain (e.g., with Spitzer and Herschel) to probe distant galaxies in the optical and IR domains, has led to formidable progress in understanding galaxy formation. Nevertheless, it is difficult to link the galaxies we see at high redshift with the ones we see in local Universe, and one is prone to Malmquist bias, as well as aperture and other selection effects.

In our work we try to give our contribution to link these two window on galaxies, by using a big multi-wavelength collection of data and especially exploiting the most promising FIR data at the moment collected by Herschel both in local and in high redshift Universe. In this Chapter, as an introduction to our original results presented along this work, we try to give a panoramic view on the current status of galaxy formation and evolution scenario, stressing out the importance of a multi-wavelength approach to address the properties of galaxies at any epoch.

1.1 Galaxies as Observed at Different Epochs: A Multi-Wavelength Point of View

Galaxies represent the fundamental constituent of our Universe. They are gravitationally bound systems composed of a large number of stars, gas, dust, and other form of matter. They span a large range in dimension, luminosity, mass and show different characteristics according to the epoch (and thus distance) at which we observe them.

1.1.1 Local Galaxies & Morphological Classification

Historically, optical photometry was the method used to observe galaxies. Thus, the morphological classification defined by Hubble in 1926 (Hubble 1926) is still the best-known today. This classification is still true and well suited to describe the behaviour of the "normal" local galaxies. According to this classification, mainly based on the visual inspections of these objects, the galaxies are divided into *Ellipticals*, *Spirals*, *Irregulars* and *S0* or *Lenticulars* (see Fig. 1.1).

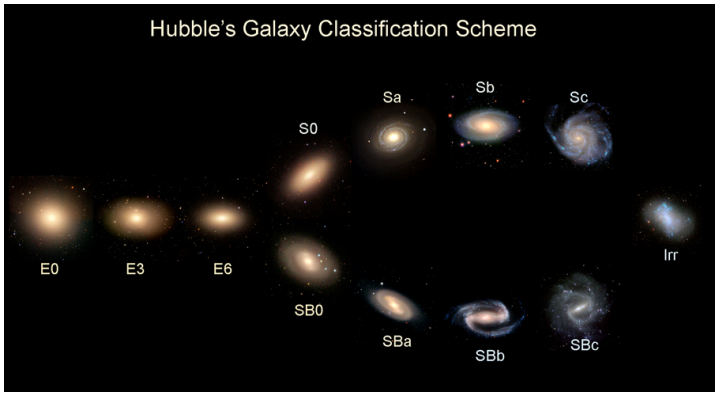

Figure 1.1: Hubble morphological classification of galaxies.

They are usually divided into two basic types: early-type galaxies including ellipticals and lenticulars, and those of the late-types including spirals and irregulars.

The ellipticals, which account for only around 10% of the observed bright galaxies, are elliptical in shape and have no discernible spiral structure. They are usually red in colour, due to the type of stars they are composed of, they have very little dust and show no sign of active star formation. Ellipticals show no significant rotational motions and their shape is thought to be sustained by the anisotropic 'hot' motions of the stars within them. These galaxies mainly host old stars characterized by a high metallicity, mostly red giants. These, however, dominate the light but not the mass of these galaxies which is mostly in little stars of the main sequence. Ellipticals occur preferentially in dense regions, e.g. inside clusters of galaxies. The distribution of stellar masses of elliptical galaxies is very broad, extending from 10^5 to $10^{12} M_\odot$.

Lenticular, or S0, galaxies were added later by Hubble to bridge the gap between normal spirals and ellipticals. Around 20% of galaxies we see present this morphology. They are more elongated than elliptical galaxies but have neither bars nor spiral structure. They present, thus, intermediate characteristics between spirals and ellipticals in term of their morphology,

(thick disc and prominent bulge), but are very similar to the second ones in term of their stellar population.

Hubble's division into normal (S) and barred (SB) spirals depends on whether the prominent spiral arms emerge directly from the nucleus, or originate at the ends of a luminous bar projecting symmetrically through the nucleus. Spirals often contain copious amounts of dust, and the spiral arms in particular show evidence of ongoing star formation (i.e. lots of young supergiant stars), giving the arms a blue colour. The nucleus or bulge of a spiral galaxy resembles an elliptical galaxy in morphology, luminosity profile and colour. Many spirals also demonstrate some kind of 'activity' (non-thermal emission processes connected to star formation).

The subdivision of S and SB into a, b or c depends on how tightly the spiral arms are wound up. Spirals show ordered rotational motion which can be used to estimate their masses. Irregular galaxies have no apparent structure and no rotational symmetry. They are relatively rare, are often faint and small and are very hard to see. These galaxies can be usually divided into two groups: type I irregulars, displaying a rotation pattern similar to the spiral one, and those of type II showing an irregular kinematics. Less luminous galaxies are often called dwarfs and irregular galaxies are generally dwarfs.

There are other types of objects which do not fit into the Hubble scheme. The "normal" galaxies described by Hubble present emission mainly by stars; therefore, the spectral energy distribution (defined in Chap. 5) of the radiation from such galaxies is in principle a superimposition of the spectra of their stellar population. The class of objects named as active galactic nuclei (AGN) generate significant fraction of their luminosity from gravitational energy that is released in the infall of matter onto a supermassive black hole and thus represents a different type of objects with comparison to those in the Hubble scheme.

Another exception is represents by the starburst galaxies, characterised by star formation rate level of $\sim 100 M_{\odot}$/yr in contrast with the typical value of $< 10 M_{\odot}$/yr of the normal galaxies. If many young stars are formed we would expect these starburst galaxies to radiate strongly in the blue or in the UV part of the spectrum, corresponding to the maximum of the Planck function for the most massive and most luminous stars. This expectation is not fully met though: star formation takes place in the interior of dense molecular clouds which often also contain large amounts of dust. If the major part of star formation is hidden from our direct view by layers of absorbing dust, these galaxies will not be very prominent in blue light. However, the strong radiation from the young, luminous stars heats the dust; the absorbed stellar light is then emitted in the form of thermal dust emission in the infrared and submillimeter regions of the electromagnetic spectrum of these galaxies can thus be extremely luminous in the IR. They are called ultra-luminous infrared galaxies (ULIRGs).

A Multi-Wavelength Approach

Clearly, all the physical process regulating the matter cycle and the perturbations induced by the environment must be considered and understood simultaneously to get the complete scenario of galaxy properties. A coherent and complete understanding of galaxy evolution through cosmic time requires a simultaneous analysis of all the different galaxy components. This can be achieved only through a multi-wavelength analysis.

In the last 25 years the advent on new technologies allowed astronomers to extend to large angular scales the observations on the Universe with a continuous coverage, from the

X-rays to centimeter Radio. The Local Universe has been completely or partially map by all sky or shallow extended surveys such as ROSAT (Voges et al. 1999) in the X-Rays, GALEX (Martin et al. 2005a) in the Ultraviolet, SDSS (York et al. 2000) in the optical, 2MASS (Skrutskie et al. 2006) in the near-infrared, Spitzer/SWIRE (Lonsdale et al. 2003) in the infrared, (now) Herschel/HerMES in the Far-infrared/Sub-millimeter (Oliver et al. 2012) and NVSS (Condon et al. 1998) in the Radio continuum. The nearby Universe has also been observed in spectroscopic mode.

Thus our knowledge of the nearby objects as well as the process that took place in them is very high at most all the λ regime. We know that stars emit as black body in the ultraviolet to the near-infrared spectral domain, with a different contribution depending on their mean age. Massive young stars dominate the short wavelength range while evolved system mainly emit in the near-infrared. The gaseous component, either atomic or molecular can be observed using some specific emission lines, HI line at 21 cm and via carbon monoxide lines emitted in the millimeter domain respectively. The dust component is heated by the interstellar radiation field and emits in the mid- and far- infrared and sub-millimeter domain following different emission processes depending on wether or not it is in thermal equilibrium with the radiation. The millimeter radio and centimeter emission is due to the free-free emission of unbound and massive stars, and by synchrotron emission due to relativistic electrons accelerated in supernovae remnants spinning in week magnetic fields. Both event are tightly related to recent star formation events and have a relative weights that changes with λ, where the free-free emissions is being important at short wavelength while synchrotron dominating at long wavelengths. The accretion phenomenon on a compact source in binary systems, or the emission of the hot gas component permeating massive elliptical galaxies or forming the diffuse interstellar medium of gas-rich systems, can be observed in the X-rays (see Tab. 1.1 as a summary).

1.1.2 High-Redshift Galaxies

The morphological and photometric features of galaxies show time evolution.

From a theoretical point of view there are known expected effects of this evolution. The best known effect is the evolution in colour called Butcher-Oemler for which, in clusters at $z > 0.1$ a higher fraction of blue galaxies is observed, compared to local clusters, interpreted as evidence of the evolution of galactic population.

A further evolution effect, theoretical and not observed, is the so called "passive evolution" on the basis of which red galaxies (ellipticals) appear brighter and bluer at higher redshift. This effect is due to the fact that the farther the distance the younger the star population is, therefore bluer.

The simplest way to measure galaxy type distribution with time is to count them in function of flux. B- band counts, for example, sensitive to young stars, show an excess of galaxies. First it was thought that these might be proto-galaxies at $z \sim 3$, but later observations confirmed that indeed they were forming dwarf galaxies at $z \sim 1$. If one observes the evolution of the luminosity function for different spectral types, at $z \sim 1$, one can note that, while more luminous galaxies evolve nearly passively, the less luminous ones seem to increase in blue and decrease in red.

Observations conducted by deep surveys like HDF Williams et al. (1996), GOODS (Dickinson et al. 2004) or COSMOS (Scoville et al. 2007), only to mention some, confirm that up to

Domain λ	**X-ray** 1-100 Å	**UV** 912-3500 Å	**Visible** 3500-7500 Å	**NIR** 0.75-5 μm	**MIR** 5-20 μm	**FIR-Submm** 20 μm - 1 mm	**Radio** 1mm-1m
Continuum							
Emission	Black Body Thermal bremsstrhalung	Black Body	Black Body	Black Body	Thermal emission	Modified Black body	Synchrotron Free-Free
Emitting source	Accretion disk in binary systems	Young stars	Intermediate age stars	Old stars	PAH, hot dust grains [a]	Cold dust grains	Relativistic electrons in weak magnetic fields
	Hot gas						HII Regions
Main Emission Lines							
Emission lines		Atomic hydrogen metals	Atomic hydrogen metals	Atomic & molecular hydrogen, molecules	PAH	[CII],CO,molecules	HI(21cm)
Origin		HII regions	HII regions	HII regions	PDR[b]	Gian molecular clouds	Diffuse ISM[c]
Absorbtion lines		Hydrogen, metals	Hydrogen			HI(21 cm)	
Origin		Stellar atmosphere, ISM, IGM[d]	Stellar atmosphere	Stellar atmosphere		Diffuse ISM	

Table 1.1: Emitting sources and emission processes at different wavelengths (Boselli 2011).

[a] Policyclic Aromatic Hydrocarbons, see Chap. 3, Sec. 3.2
[b] Photodissociation Region
[c] Interstellar Medium
[d] Intergalactic Medium

$z \sim 1$ bright galaxy evolution seems almost passive, but at $z \sim 3$ galaxies appear essentially different. In particular at redshift $z \sim 3$ galaxies do not follow Hubble sequence and they appear irregular and fragmentary. This last effect might be due, in part at least, to the fact that high redshift galaxies are composed mostly by very young stars of O B type emitting in the UV. Local galaxies too, if observed in the UV, show a fragmentary structure because in this band star formation areas are pointed out.

From an observational point of view, galaxy evolution has been studied in detail and is currently quite precisely known between redshift 0 and approximately 1, mostly based on deep optical imaging and spectroscopy.

One important aspect of these investigations concerns the fraction of optical-UV radiant energy emitted by young stellar populations, which is absorbed by dust in the ISM medium and re-emitted at long wavelengths. It has been shown that, while being modest for evolved galaxies like early-type spirals, and negligible in E/S0 galaxies, this fractional energy may become strong or even dominant during active phases of star formation, galaxy assembly, and nuclear activity (Kormendy & Sanders (1992); Franceschini et al. (1994); Sanders & Mirabel (1996)).

Mid- and far-infrared observations of the high-redshift universe are then needed for precise quantification of the dust-extinction effect: without a direct observation of the dust-reradiated part of the young stellar population emission, a reliable estimate of the star formation rate is difficult (e.g. Poggianti & Wu (2000)).

In the redshift interval from local up to $z \sim 1$, relatively deep IR observations have been obtained with early space infrared missions like IRAS and ISO (see Chap. 3). Based on the fair sensitivity and moderate spectral coverage of the latter, the galaxy and AGN evolution in the redshift interval $0 < z < 1$ has been mapped (Franceschini et al. (2001); Chary & Elbaz (2001); Rowan-Robinson (2001), Rowan-Robinson (2009); Elbaz et al. (2002); Lagache et al. (2004); Pozzi et al. (2004)), but there was insufficient sensitivity at higher redshifts for any significant constraints to be set there.

Multi-wavelength studies on galactic scales might still be prohibitive at high redshifts, but the avilable data allows us to reconstruct the spectral energy distributions (SED) from the X-Ray to centimeter radio regimes. SED and optical and near-IR spectra are now useful tools since they can be reproduced by models and compared to observations in samples of objects at different redshift, thus providing us with one of the most powerful tools for constraining galaxy evolution.

For a long time, both theoretical expectations (Franceschini et al. (1994); Blain & Longair (1993); see also Kormendy & Sanders (1992)) and preliminary observational results by millimetric and sub-millimetric surveys with SCUBA (see Chap. 3) and IRAM (Greve et al. (2005); Bertoldi et al. (2007)) have produced evidence that major events of star formation in the universe happened at high redshifts, $z > 1$, and are likely responsible for the origin or assembly of the most massive galaxies today. Such events would be accompanied by major energy release, but, because of heavy dust extinction, would be detectable only through far-infrared observations. Sub-millimetric surveys indeed have identified sources that are extremely luminous at long wavelength, but almost undetectable in the optical (e.g. Chapman et al. (2005); Mobasher et al. (2005); Rodighiero et al. (2007)).

A dramatic improvement in our knowledge of the infrared emissivity of source populations at redshifts $z > 1$, corresponding to the the era of peak star formation and AGN activity, has been made possible with the deep sky exploration by the Spitzer Space Telescope (SST,

Werner et al. (2004), see Chap. 3). The Multiband Imaging Photometer (MIPS, Rieke et al. (2004)) on SST combined sensitivity and spectral coverage, particularly with the $\lambda_{\rm eff} = 24$ μm band, adequate for identifying and characterizing substantial numbers of far-IR galaxies up to $z \simeq 3$ (Le Floc'h et al. (2005); Dole et al. (2006b); Caputi et al. (2007); Lonsdale et al. (2004); Rowan-Robinson et al. (2005)). The MIPS instrument also includes photometric imagers at $\lambda_{\rm eff} = 70$ and 160 μm (although the latter was limited by source confusion). A further crucial contribution by the SST mission was the capability offered by the Infrared Spectrometer (IRS, Houck et al. (2004)) to obtain high-quality far-IR spectra of high-redshift sources for the first time.

An historical difficulty in identifying and interpreting luminous high-redshift sources, i.e. the difficult association of the optical counterparts to the IR-mm sources due to the poor spatial resolution and the large IR error-boxes, is also largely overcome by observations with the Spitzer Infrared Array Camera (IRAC, Fazio et al. (2004)). IRAC deep multiwavelength imaging at 3 to 10 μm allows continuous band-merging and association of far-IR and optical/near-IR source catalogues in this critical spectral transition region (e.g. Egami et al. (2004)).

The multiwavelength observations by Spitzer have fully confirmed previous expectations of numerous populations of luminous galaxies and active nuclei at high redshifts. Sources at z¿1 and up to ∼3 are routinely detected with the sensitive, high-spatial resolution 24 μm channel of Spitzer/MIPS (see Cha[. 3).

MIPS surveys at longer wavelengths (70 and 160 μm) have produced more limited samples of distant sources (Frayer et al. (2006a), Frayer et al. (2006b), Frayer et al. (2009)), because of the much longer integration times required and the lower mapping speed at faint flux levels. These faint far-IR Spitzer sources are also more difficult to identify, because of the poor Spitzer resolution at such long wavelengths.

More recent results obtained exploiting deep MIPS 24 μm data are shown in Fig. 1.2, top panel, by Rodighiero et al. (2010). In this plot the star formation rate activity shows a quick increase from $z = 0$ to $z \sim 1$ where we can observe a wide peak, it remains nearly constant up to $z \sim 3$ to decrease, then, at higher redshifts. A complete different scenario is however found by those study based only UV selected data, like Hopkins 2006, reported in the low panel of Fig. 1.2. These studies find a peack of the SFR at redshift ~ 2 and a rapid decrease at higher redshift.

It is important to stress that both these solutions come from strong over-extrapolation on data, especially at high redshift, and are thus affect by great uncertainties. To really understand what happens at high redshift regime we would need multi-wavelength studies and more deep and precise observations in the FIR bands, in order to directly sample SFR activity in a wide range of redshift. Such a kind of studies are now possible thanks to the Herschel Space Observatory (see Chap. 3) which is able for the first time to really collect data on star forming galaxies at redshift up to ~ 4.

In this work we exploit Herschel data collected within HerMES survey, as well as a wide collection of multi-wavelength data both in the local and in the high redshift Universe to study the evolution of the luminosity function of galaxies and thus give a contribution to solve the scenario of galaxy formation and evolution across cosmic time (see Chap. 4 and Sec. 3.8.3 for a description of the data and Chap. 7 for a discussion of the results).

The already published works based on the very first data collection udertaken with Herschel have already proven the great potential of this innovative instrument especially with the HerMES survey (Oliver et al. 2012).

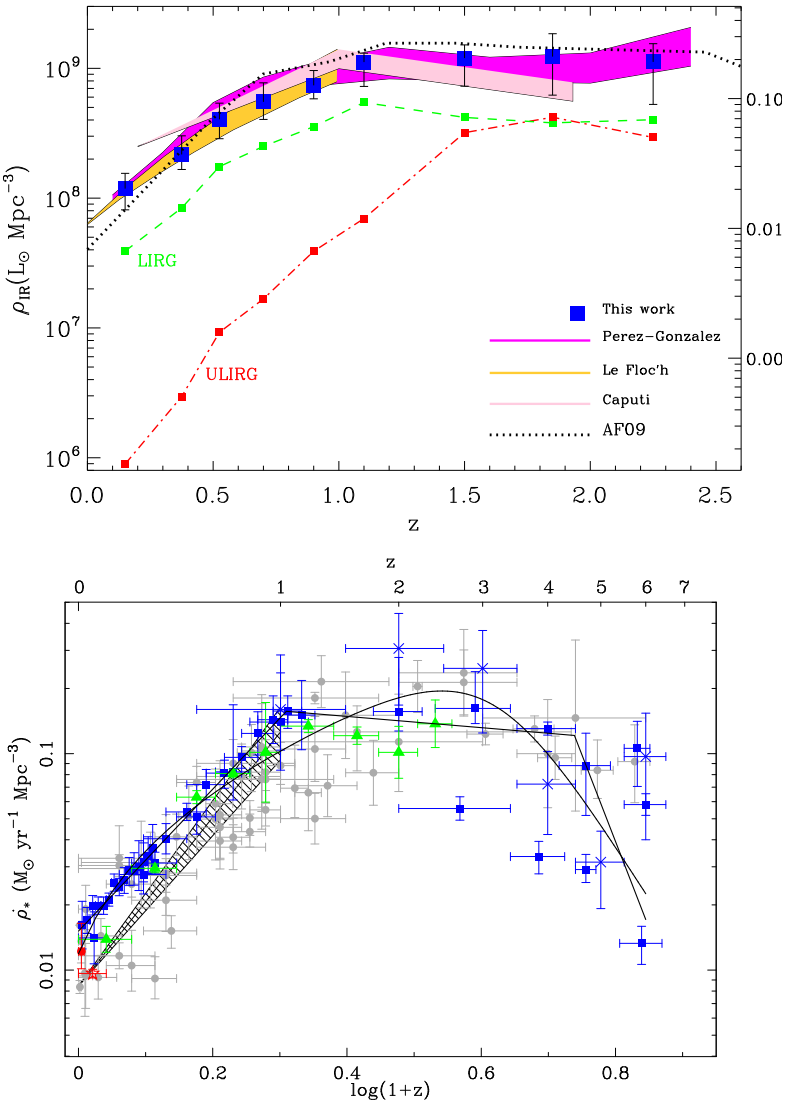

Figure 1.2: Comparison the Evolution of the star formation rate density as from IR and UV luminosity function estimates. **Top**: Evolution of the comoving bolometric IR luminosity density with redshift Rodighiero et al. (2010). The right-hand axis contains the conversion to star-formation rate based on a scaling law by Kennicutt (1998). See the paper for a detailed description of the plot; **Bottom**: Evolution of SFR density with redshift (Hopkins & Beacom 2006) where the grey point fitted by the filled black line peaking at $z \sim 2$ are UV selected data by Hopkins (2004).

HerMES is detecting over 100,000 sources using two instrument, PACS and SPIRE sampling 70/100/160 and 250/350/500 μm respectively (see Sec. 3.8.3. The FIR colours of the Herschel sources help addressing the question of how much of the energy production comes from accretion (AGN) and how much from star formation. First results on an SDSS sample of AGN (Hatziminaoglou et al. 2010) find that one third are detected by SPIRE, with the long wavelength colours indistinguishable from star forming galaxies. Modelling of the full SED required the combined contribution of both AGN and starburst components, with the former

dominating the emission at the MIR wavelengths and the latter contributing mostly to the FIR wavelengths. This suggests that SPIRE detects the star formation in AGN, with little contamination from any dusty torus, offering high hopes for disentangling nuclear and star formation activity.

The wealth of data in these fields mean we can explore the FIR properties of many known samples. HerMES first results on Lyman break galaxies have already shown that we can detect U-band dropout sources with stacking (Rigopoulou et al. 2010) and FUV drop-out sources individually (Burgarella et al. 2011).

HerMES, thanks to the database introduced in this work in Chap. 4, give to the community a complete data set to understand the global relationship between FIR and optical galaxies, the effect of dust attenuation in optical/UV populations, and phenomena in individual galaxies. First results comparing HerMES and GALEX (Buat et al. 2010) confirm that total infrared luminosity accounts for 90 per cent of the total star formation rate, though this reduces to 70 per cent when considering the lower star formation rate systems ($\dot{M}_* < 1\mathrm{M}_\odot\mathrm{yr}^{-1}$).

These ancillary data can also be used to investigate the detailed properties of the FIR galaxies, e.g. their morphology. One study has explored galaxies with morphological classifications at $2 < z < 3$ and shows that the mean SFR for the spheroidal galaxies is about a factor of three lower than for the disk like galaxies (Cava et al. 2010).

It is then clear that Herschel observations are giving us a perfect collection of data to really push beyond our knowledge on galaxy formation and evolution scenario.

1.2 Cosmic Background Radiation

The cosmic microwave background (CMB) is a remnant of the early hot phase of the Universe, namely thermal radiation from the time before recombination. CMB contains a great deal of information about our Universe and represent a strong constraint to the emissivity of the single galaxies we can ever observe or deduced by models. Therefore, one might ask whether background radiation also exists in other wavebands, which then might be of similar value for cosmology. In Chap. 3 we will focus to the discovery of the IR background and it's key role to stress the importance of IR studies on galaxies, here we want to report a summary of the evidence of the background light at other λs in order to complete the census of such an important constraint in the galaxy evolution studies.

The neutrino background that should be present as a relic from the early epochs of the Universe, in the form of a thermal distribution of all three neutrino families with $T \sim 1.9$ K, is likely to remain undiscovered for quite some time due to the very small cross-section of these low-energy neutrinos.

Indeed, apparently isotropic radiation has been found in wavelength domains other than the microwave regime. Following the terminology of the CMB, these are called background radiation as well. However, the name should not imply that it is a background radiation of cosmological origin, in the same sense as the CMB. From the thermal cosmic history, no optical or X-ray radiation is expected from the early phases of the Universe. Hence, for a long time it was unknown what the origin of these different background radiations may be.

At first, the early X-ray satellites discovered a background in the X-ray regime (cosmic X-ray background, CXB). Later, the COBE satellite (see Chap. 3) detected an apparently isotropic radiation component in the FIR, the cosmic infrared background (CIB).

In the 1970s, the first X-ray satellites discovered not only a number of extragalactic X-ray sources (such as AGNs and clusters of galaxies), but also an apparent isotropic radiation component, the CXB. Its spectrum is a very hard (i.e., flat) power law, cut off at an energy above ~ 40 keV. Initially, the origin of this radiation was unknown, since its spectral shape was different from the spectra of sources that were known at that time. For example, it was not possible to obtain this spectrum by a superposition of the spectra of know AGNs.

ROSAT, with its substantially improved angular resolution compared to earlier satellites (such as the Einstein observatory), conducted source counts at much lower fluxes, based on some very deep images. From this, it was shown that at least 80% of the CXB in the energy range between 0.5 keV and 2 keV is emitted by discrete sources, of which the majority are AGNs. Hence it is natural to assume that the total CXB at these low X-ray energies originates from discrete sources, and observations by XMM-Newton seem to confirm this.

It was supposed for a long time that the CXB is, at higher energies, produced by a hot intergalactic gas at temperatures of $k_B T \sim 30$ keV.

This model was excluded, however, by the precise measurement of the thermal spectrum of the CMB by COBE, showing that the CMB has a perfect blackbody spectrum. If a postulated hot intergalactic gas were able to produce the CXB, it would cause significant deviations of the CMB from the Planck spectrum, namely by the inverse Compton effect. Thus, the COBE results clearly ruled out this possibility.

Thanks to Herschel we are now able to push even more constraint on the background analysis, being able to really measure the contribution in time of the FIR emitting sources.

1.3 Theoretical Tools for Galaxy Formation & Evolution Studies

The study of the galaxy formation and evolution needs, as we already pointed out, both observations and models able to describe the nature of the physical processes detectable at different wavelengths. The physical processes acting on baryons to create structure like stars or galaxies are complex and require sophisticated models able to describe the interplay of many different physical processes acting at the same time. There are two approaches for modeling these processes that are the bases of galaxy formation and evolution: the Hydrodynamical simulations and the so-called semi-analytical models. The two techniques are based on different and simplifying assumptions and they both are useful for studying different aspects of galaxy formation and evolution. In the following sections we give a very brief description of both these approaches. At the end of this work in Chap. 7 we will present some preliminary comparison between our observational results and some recent semi-analytical models predictions.

1.3.1 Hydrodynamical Simulations

Describing the formation and evolution of galaxies by means of hydrodynamical simulation is very difficult and some-how impossible in great details: the physical processes that determine the formation of stars in galaxies occur on very small length-scales, whereas the evolution of structures, which defines, e.g., the merger rate, happens on cosmological scales. Hence it is impossible to treat both scales together in a single simulation. Furthermore, the physical laws determining the behavior of gas (hydrodynamical processes such as shock fronts and friction;

radiation processes) are too complicated to be modeled in a detailed simulation, except in those which are confined to a single galaxy. In addition, many of the gas processes are not understood sufficiently well to compute their effects from basic physical laws.

SPH (Smoothed Particle Hydrodynamics) is a Lagrangian numerical method following the motion of a set of gas particles. Differently from the semi-analytical models, SPH does not impose any restriction on the halo geometry and solves directly the equations for the evolution of both dark matter and gas particles leading to accurate representation of resolved objects. By solving the hydro-dynamic and thermo-dynamic equations we can compute the thermal energy, velocity and position of each particle at any given time. Estimates of gas properties at each particle position are derived by smoothing over the nearest neighboring particles. Given these gas properties they can be used to compute the forces which act on each particle to update their thermal energies, velocities and positions. However, due to the smoothing, the description of the regions where the gas properties vary discontinuously (or at a scale much smaller than the smoothing scale) is poor.

The limitation of this method lies in the high computational cost which does not allow the detailed analysis of the parameter space of the galaxy formation process. However limited resolution restricts the accessible dynamical range and the computational requirements of cosmological simulations limit the parameter space explorations.

1.3.2 Semi-Analytical Models

To overcome the intrinsic limitation of describing all the physical processes with a quantitative physical model we can parametrize the functional behavior of those processes. The semi-analytical models assume that dark matter haloes are spherically symmetric and that infalling gas is shock-heated to the virial temperature of the halo and then settles with a universal density profile. Given the halo merger trees (see Fig. 1.3), several assumptions in the treatment of gas are made in order to obtain simple, analytic solutions to the complex hydrodynamical processes at the bases of galaxy formation.

In these models the dynamics of the gas is strongly coupled with the evolution of dark matter halos. Processes such as star formation rate are modeled by means of simple and physically motivated recipes. The set of free parameters is fixed by comparing models to observations. One of the usual objection made to these models is that they seem to contain too many free parameters, but it is essential to underline that the choice of such parameters is physical and not statistical and it is established simply because of our lack of understanding of the physics of galaxy formation. It is worthwhile to stress that the number of constraints on parameters is larger than the number of parameter used.

Anyway, these models are computationally cheaper and can reach high mass resolutions in the halo merger trees. In the CDM model [1], massive halos are formed by the merging of halos of lower mass. An extension to the Press-Schechter theory [2] allows us to compute the statistical properties of these merger processes of halos. By means of these, it is then possible to generate a statistical ensemble of merger histories for any halo of mass M today. Each individual halo is then represented by a merger tree. Alternatively, such merger trees can

[1] Cold Dark Matter (CDM) model

[2] The Press-Schecter model describe the spherical collapse for a non-linear evolution of a density perturbation in the Universe. Despite being simplistic, it represents the fundamental principles of gravitational collapse and yields approximate relations.

also be extracted from numerical simulations of structure formation, by following the mass assemble history of individual halos. The statistical properties of halos of mass M at redshift z are then obtained by analyzing the ensemble of merger trees. Each individual merger tree specifies the merger processes that have led to the formation of a particular halo.

The merger processes can be divided into *major* and *minor* merger according to the mass ratio of the merging halos. Major mergers refers to the formation of a spheroidal stellar population where the ratio of the two halos component is, e.g., 1:3, in this case, the disk populations of the two merging galaxies are dynamically heated to commonly form an elliptical galaxy; minor mergers refers to the case in which the masses of the two components in a merger are very different, the gas of the smaller component will basically be accreted onto the more massive halo, where it can cool again and form new stars. By this process, a new disk population may form. In this model, a spiral galaxy is created by forming a bulge in a 'major merger' at earlier times, with the disk of stars and gas being formed later in minor mergers and by the accretion of gas. Hence the bulge of a spiral is, in this picture, nothing but a small elliptical galaxy, which is also suggested by the very similar characteristics of bulges and ellipticals, including the fact that both types of object seem to follow the same relation between black hole mass and velocity dispersion.

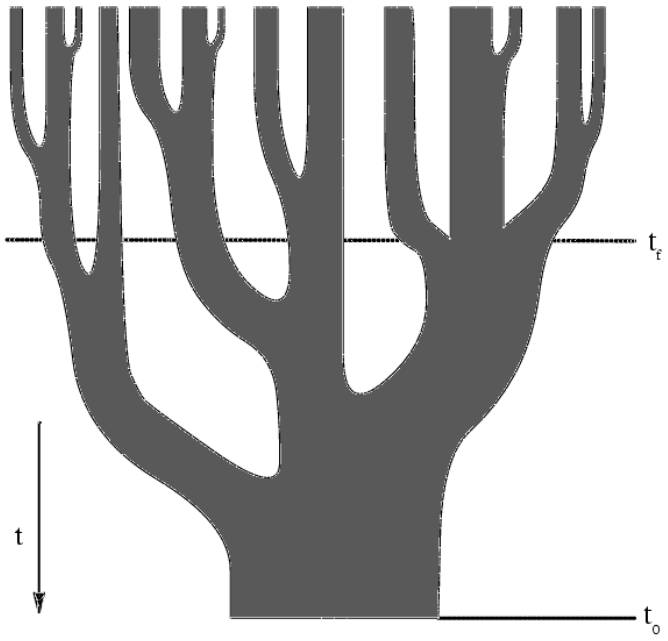

Figure 1.3: Typical merger tree as expected in a hierarchical CDM model of structure formation. A massive halo at the present time t_0 has formed by mergers of numerous halos of lower mass, as indicated in the figure. One defines the time of halo formation as the time t_f at which one of the sub-halos had reached half the mass of the current halo.

Semi-analytical models can then be described by means of a number of properties they consider in their framework. Beyond the merger trees already mentioned they can consider also the scenario in which in a halo do not undergo any merger process at a given time, thus gas can cool (*cooling process*) and the cooling rate is determined by the chemical composition and the density of the gas. If the density is sufficiently high and cooling is efficient, gas can be

transformed into stars and the star formation is parametrized by a factor of proportionality between the star-formation rate and the rate at which gas cools.

Besides they include, the *feedback* process that happens when shortly after the formation of stars, the more massive of them will explode in the form of supernovae. This will re-heat the gas, since the radiation from the SN explosions and, in particular, the kinetic energy of the expanding shell, transfers energy to the gas. By this heating process, the amount of gas that can efficiently cool is reduced; this reduction increases with the star-formation rate. This leads to a self-regulation of star formation, which prevents all the gas in a halo from being transformed into stars.

These are only some of the properties that these models can take into account to describe the complex process of galaxy formation, we refer the reader to more specific literature for a more detailed description since this is beyond the intention of this work.

The free parameters in semi-analytic models, such as the star-formation efficiency or the fraction of energy from SNe that is transferred into the gas, need to be fixed by comparison with some key observational results. For example, one requires that the models reproduce the correct normalization of the Tully-Fisher relation [3] and that the number counts of galaxies match those observed. Although these models are too simplistic to trace the processes of galaxy evolution in detail, they are highly successful in describing the basic aspects of the galaxy population, and they are continually being refined.

For these reason, studies such as the one we propose in this work, that collect a huge amount of multi-wavelength information at different redshift, are of fundamental importance in studying the statistical properties of galaxy evolution, to better constrain the parameter of semi-analitical models and thus complete the scenario of how, when and where the galaxies form.

Hierarchical models *vs* Cosmic downsizing

Another important rule covered by the observations within the context of the galaxy evolution is try to solve the contradictions between two cosmological scenarios both existing in the modern cosmology.

Semi-analytical models describe a hierarchical model of structure formation where that smaller-mass objects are formed first, with more massive systems forming later in the cosmic evolution. There is ample evidence for this to be the case; e.g., galaxies are in place early in the cosmic history, whereas clusters are abundant only at redshifts $z > 1$.

However, looking more closely into the issue, apparent contradictions are discovered from observation of particular phenomena. For example, the most massive galaxies in the local Universe, the massive ellipticals, contain the oldest population of stars, although their formation should have occurred later than those of less massive galaxies. In turn, most of the star formation in the local Universe seems to be associated with low- or intermediate-mass galaxies, whereas the most massive ones are passively evolving.

The phenomenon that massive galaxies form their stars in the high-redshift Universe, whereas most of the current star formation occurs in galaxies of lower mass, has been termed 'downsizing'. This downsizing can be studied in more detail using redshift surveys of galaxies

[3] Tully and Fisher in 1977 observed that galaxy luminosity and rotational velocity at a given scale length linearly correlate in a log-log plane. The slope of this relation depends on the band in which one observes.

like the one present in this work. Within this context our study assume a more great importance in trying to give a contribution in understanding which is the best formation scenario that is able to describe our observational results.

Chapter 2

Statistical Studies of Galaxy Evolution

One of the fundamental problems of modern cosmology is to explain the formation of structures in the Universe, and of galaxies in particular, that are the main baryonic component of the cosmological large-scale structure. Understanding galaxy formation and evolution is however a challenging problem whose solution requires a concerted approach combining observational, theoretical semi-analytical and numerical work. The main problem connected with these studies is that we are able to see the Universe at different epochs, due to the finite speed of light, but we are unable to track the evolution of an individual galaxy or galaxies. Thus, to study evolution, we must compare the statistical properties of samples of galaxies seen at different epochs. For this reason, starting from as early as the 1930s, astronomers have tried to quantify the statistical nature of evolution and large-scale structure of galaxies by studying their distribution as a function of flux (i.e. their source counts) or, when distances are available, as a function of luminosity (i.e. their luminosity function).

Clearly, these studies need large complete samples of objects with accurately known flux densities or apparent magnitudes. The simplest way of obtaining these samples is to count the number of objects brighter than different flux density limits and compare these with the expectations of standard world models. The analysis can then be greatly refined when redshifts and spectroscopic evidence are available for large source samples (Longair 2008).

The most direct way of constraining the evolutionary history of galaxies and trying to resolve the discrepancies would be to derive the redshift-dependent luminosity and mass functions from local/wide and high-z/deep, unbiassed surveys.

In our work we give our contribution in this scenario by investigating the properties of the luminosity functions both in the local and in the high-z Universe using the best collection of data now available in the IR/SMM regime taken within Spitzer and Herschel cosmological surveys. In this Chapter we introduce the concept of *luminosity functions* the problems connected in its evaluations and how we can use them to extrapolate information on evolution of galaxy.

2.1 Cosmological Definitions

Any cosmological study on galaxy formation and evolution assumes a typical cosmological framework to describe the observational quantities like distances, volumes, luminosity and number densities. In this Section, we want to explain the fundamental tools needed to investigate this quantities into a cosmological context where the general relativity describes the geometry of the space by metrics which determine the separation between points. In particular we describe here the definition we use in our analysis, like cosmological constants and k-corrections definition.

2.1.1 Cosmological Models and Distance Measures in Cosmology

One of the critical problems when studying the Universe over large scales is to understand how we can describe its geometry and how we can measure it. In an expanding universe we have to pay particular attention because the distances between comoving objects are constantly changing, and Earth-bound observers look back in time as they look out in distance. For these reason there are many ways to specify the distance between two points. The unifying aspect is that all distance measures somehow measure the separation between events on radial null trajectories, i.e. trajectories of photons which terminate at the observer (Hogg 1999).

The fundamental parameter that describes the motion in an expanding universe is the *Hubble Constant* H_0 . This parameter is the constant of proportionality between recession speed v and distance d in the expanding universe,$v = H_0\, d$, its value changes with time and H_0 refers to the present epoch. If we assume that the Universe is, looking at large enough scales, homogeneous and isotropic, one gets a simplification of the metric of the Universe. It may be described by the line element $\mathrm{d}s$ of the Robertson-Walker metric in polar coordinates:

$$\mathrm{d}s^2 = \mathrm{d}t^2 - a(t)^2 \left[\frac{\mathrm{d}r^2}{1-kr^2} + r^2(\mathrm{d}\theta^2 + \sin^2\theta\, \mathrm{d}\,\phi^2)\right] \tag{2.1}$$

where $a(t)$ is the scale factor which depends on time, as to describe an expanding universe. Since $a(t)$ has the dimension of length, r is dimensionless and k, which is the curvature of space, may be set to $k = 0$, $+1$ or -1 by choosing an appropriate scaling of r. Evaluating for today ($t = t_0$), one can introduce the following constants:

$$\begin{aligned} a_0 &= a(t_0) \quad & H_0 &= \frac{\dot{a}(t_0)}{a(t_0)} \\ \rho_0 &= \rho(t_0) \quad & q_0 &= -\frac{\ddot{a}(t_0)a(t_0)}{\dot{a}(t_0)^2} \end{aligned}$$

where q_0 is the deceleration parameter. If we estimate $H_0 = 100h$ km s^{-1}Mpc^{-1}, where h is a dimensionless number parametrizing our ignorance, than the inverse of the Hubble constant is the *Hubble Time* t_{H}:

$$t_H \equiv \frac{1}{H_0} = 9.78 \times 10^9\, h^{-1}\ \mathrm{yr} = 3.09 \times 10^{17}\, h^{-1}\ \mathrm{s} \tag{2.2}$$

and the *Hubble Distance*, D_H is:

$$D_H \equiv \frac{c}{H_0} = 3000\,h^{-1}\ \text{Mpc} = 9.26 \times 10^{25}\,h^{-1}\ \text{m} \tag{2.3}$$

These quantities set the scale of the Universe. In this context the mass density ρ of the Universe and the value of the cosmological constant Λ would affect the time evolution of the metric, that would be, without the simplification explained in Eq. 2.1, a metric described by the Einstein field equations. However we can consider the mass density parameters as purely kinematic parameters, and so in the form:

$$\Omega_{\rm M} \equiv \frac{8\pi\,G\,\rho_0}{3\,H_0^2} = \frac{\rho_0}{\rho_{cri,0}} \tag{2.4}$$

$$\Omega_\Lambda \equiv \frac{\Lambda\,c^2}{3\,H_0^2} \tag{2.5}$$

$$\Omega_k = 1 - \Omega_M - \Omega_\Lambda \tag{2.6}$$

(Peebles 1993) where

$$\rho_{cri,0} = \frac{3H_0^2}{8\pi G} \tag{2.7}$$

is the *critical density* or the average density of matter in the Universe today that would be needed exactly to halt, at some point in the future, the cosmic expansion. The Universe that has precisely the critical density is said to be $flat$ or *Euclidean*. Ω_k measures the "curvature of space"'. These parameters completely determine the geometry of the Universe if it is homogeneous, isotropic, and matter-dominated.

The baryonic matter in the Universe we can see directly (as stars or illuminated by stars) or indirectly (from absorption or by projections) accounts for at most 10% of Ω_M. The dynamical properties of galaxies and galaxy clusters and the effects of gravitational lensing also suggest that the majority of mass cannot be seen. It is still unclear what those remaining 90% consist of, but as this matter seems to only interact gravitationally with baryonic matter, it is referred to as *dark matter*.

The current theory says that our Universe appears geometrically flat and under expansion. In this scenario in our estimations we refer to the typical value of $\Omega_\Lambda = 0.72$, $\Omega_M = 0.28$ and $H_0 \sim 72$ km s^{-1}Mpc^{-1} with a negative parameter of deceleration q_0 and the curvature index $k = 0$. Different models, as the *Einstein de Sitter* Universe, assume different values for these parameters.

One of the most common ways to describe large astronomical distances is by the concept of redshift:

$$z \equiv (\lambda_0 - \lambda_{em})/\lambda_{em} \tag{2.8}$$

Observations show, that the more distant an object is, the more its light is shifted to longer wavelengths. Absorption and emission lines of atoms are commonly used to measure a redshift. The redshift effect may be explained in two ways. One may interpret it as the Doppler effect that light affects due to the movement of the source relative to the observer. That motion then arises from the expansion of the Universe. The other way to look at it is a

general expansion of space which "stretches" the light wave (cosmological redshift). In special relativity, redshift is related to radial velocity v by:

$$1 + z = \sqrt{\frac{1 + v/c}{1 - v/c}} \tag{2.9}$$

where c is the speed of light. The difference between an object's measured redshift z_{obs} and its cosmological redshift z_{cosm} is due to its (radial) peculiar velocity v_{pec}; i.e. the cosmological redshift could be defined as that part of the redshift due solely to the expansion of the Universe, or Hubble flow. The peculiar velocity is related to the redshift difference by:

$$v_{pec} = c\frac{(z_{obs} - z_{cosm})}{(1 + z)} \tag{2.10}$$

where is assumed $v_{pec} \ll c$. If we consider small v/c and small distance d, in the expanding Universe, the velocity is linearly proportional to the distance and all the distance measures, angular diameter distance, luminosity distance, etc, converge; we can write:

$$z \approx \frac{v}{c} = \frac{d}{D_H} \tag{2.11}$$

where D_H is the Hubble distance. The cosmological redshift is directly related to the scale factor $a(t)$, or the "size" of the Universe. For an object at redshift z

$$1 + z = \frac{a(t_0)}{a(t_e)} \tag{2.12}$$

where $a(t_0)$ is the size of the Universe at the time the light from the object is observed, and $a(t_e)$ is the size at the time it was emitted. Redshift is independent of cosmology, but it does not correspond to a distance one could measure with a ruler. To obtain a distance measure in a proper length scale, one needs to take spacetime into account.

2.1.2 Radial Comoving Distance

As light needs time to get from an object to the observer, one can define a distance that may be measured between the observer and the object with a ruler at the time the light was emitted, the *proper distance*. The *radial* (or line-of-sight) *comoving distance* is the *proper distance* at the time the light was emitted times the ratio of scale factors now to then $(1 + z)$. One can then define the function (Peebles 1993):

$$E(z) \equiv (\Omega_M(1 + z)^3 + \Omega_k(1 + z)^2 + \Omega_\Lambda)^{\frac{1}{2}} \tag{2.13}$$

$E(z)$ is proportional to the time derivative of the logarithm of the scale factor $\dot{a}(t)/a(t))$. For this reason, $H(z) = H_0E(z)$ is the Hubble constant as measured by a hypothetical astronomer working at redshift z. Since $\mathrm{d}z = \mathrm{d}a$, $\mathrm{d}z/E(z)$ is proportional to the time-of-flight of a photon traveling across the redshift interval $\mathrm{d}z$, divided by the scale factor at that time. The total line-of-sight comoving distance is then given by integrating these contributions:

$$D_C = D_H \int_0^z \frac{\mathrm{d}z'}{E(z')} \tag{2.14}$$

where D_H is the Hubble distance defined in Eq. 2.3. The line-of-sight comoving distance between two nearby events (i.e. close in redshift or distance) is the distance which we would measure locally between the events today if those two points were locked into the Hubble flow.

2.1.3 Transverse Comoving Distance

The *transverse comoving distance* is a quantity used to get the comoving distance perpendicular to the line of sight. In other words it is used to get the comoving distance that two objects at the same redshift have which are separated on the sky by an angle $\delta\theta$. In an Euclidean geometry this distance would be $D_S = D_C \delta\,\theta$, but this no longer holds true in a curved spacetime where the transverse comoving distance D_M is defined as $D_S = D_M\,\delta\theta$ which holds true in any spacetime. Using Robertson-Walker metric we get:

$$D_M = \begin{cases} D_H \frac{1}{\sqrt{\Omega_k}} \sinh[\sqrt{\Omega}_k D_c/D_H] & \text{for } \Omega_k > 0 \\ D_C & \text{for } \Omega_k = 0 \\ D_H \frac{1}{\sqrt{|\Omega_k|}} \sin[\sqrt{|}\Omega_k|D_C/D_H] & \Omega_k < 0 \end{cases} \tag{2.15}$$

where the trigonometric functions sinh and sin account for what is called " the curvature of space". For $\Omega_\Lambda = 0$, there is an analytic solution to the equations

$$D_M = D_H \frac{2[2 - \Omega_M(1-z) - (2-\Omega_M)\sqrt{1+\Omega_M z}]}{\Omega_M^2(1+z)} \tag{2.16}$$

2.1.4 Luminosity Distance & k-correction

Once a suitable co-moving distance scale for the specfied cosmology has been described, the measurement of general galaxy properties, such as their flux or flux density, S, and distance, can then be described in terms of their intrinsic luminosity, L, and redshift, z. In Euclidean geometry the flux of an object is emitted isotropically, over a sphere of radius r and is related to the objects luminosity by:

$$S = \frac{L}{4\pi r^2} \tag{2.17}$$

Dealing with the problems of cosmological measures we need to define the *Luminosity distance* D_L, so that the Eq. 2.17 holds true with $r = D_L$ for any spacetime. The luminosity distance is given as (Weinberg 1972):

$$D_L(z) = (1+z) D_M(z) \tag{2.18}$$

or

$$D_L(z) = (1+z) D_C(z) \tag{2.19}$$

in the case of our Universe as explained in Sec. 2.1.1. Since the speed of light in a vacuum is believed to be constant, wavelength and frequency are interchangeable, so that the luminosity density can be written as either L_λ or L_ν and they are related by:

$$L_\nu \mathrm{d}\nu = L_\lambda \mathrm{d}\lambda \tag{2.20}$$

giving

$$L_\nu = \frac{c}{\nu^2} L_\lambda \quad or \quad \nu L_\nu = \lambda L_\lambda \tag{2.21}$$

The *Spectral Energy Distribution* (SED) of a galaxy with flux S, defined by $S = \int S_\nu \mathrm{d}\nu$ may then be described by νS_ν, expressed in terms of the luminosity distance, D_L.

If νL_ν is not constant, due to the redshift the observed flux varies. Taking those effects of redshift into account, the following relations can be found:

$$\nu_0 S_{\nu_0} = \frac{\nu_e L_{\nu_e}}{4\pi D_L^2(z)} \qquad \nu_0 = \frac{\nu_e}{(1+z)}$$

$$\begin{aligned} \Rightarrow S_{\nu_0} &= \frac{\nu_e}{\nu_0} \frac{L_{\nu_e}}{4\pi D_L^2(z)} \\ &= (1+z) \frac{L\nu_e}{4\pi D_L^2(z)} \end{aligned} \tag{2.22}$$

where ν_0 is the observed frequency, and ν_e is the frequency the light was emitted at. What we really want to do is to relate the light that we observe at a specific wavelength to the emitted light in the object's rest-frame:

$$S_{\nu_0} \mathrm{d}\nu_0 = \frac{L_\nu[\nu_0(1+z)](1+z)\mathrm{d}\nu_0}{4\pi D_L^2} \tag{2.23}$$

Thus we find:

$$S_{\nu_0} = \frac{L_\nu[\nu_0(1+z)](1+z)}{4\pi D_L^2} \tag{2.24}$$

The k-correction is defined exactly to "correct" the fact that sources observed at different redshifts are, in general, compared with standards or each other at different rest-frame wavelengths. This can be defined as the ratio of the observed luminosity to the rest-frame luminosity that the galaxy emitted, which is equivalent to that coming from the rest-frame SED at those two frequencies:

$$k(z) = \frac{\int R(\nu) L_{\nu_0} \mathrm{d}\nu}{\int R(\nu) L_{\nu_e} \mathrm{d}\nu} \tag{2.25}$$

where $R(\nu)$ is the responsivity of the photometric band. Rest-frame flux/luminosity densities, measured at a given wavelength/frequency, may thus be determined for a galaxy using the k-correction, determined from a galaxy's SED (provided one knows the galaxy's rest-frame SED, redshift and observed flux/luminosity density).

2.1.5 Number Counts

Having explained problems connected with studying objects at differente epochs and position, now we can explain the simple method of counting objects in order to obtain the required complete sample of sources. We can distinguish between *integral counts* and *differential counts*:

- integral counts: $N(>S_\nu)$ or the number of sources on a unit sky's area with $S > S_\nu$;
- differential counts: $N(S_\nu) = \mathrm{d}N(S)/\mathrm{d}S$.

A simple method to obtain integral counts is to calculate:

$$N(>S_\nu) = \sum_i \frac{1}{\Omega(S_i)} \tag{2.26}$$

where the sum is over all sources with $S_i > S_\nu$ and $\Omega(S)$ is the area of sampling (or the area of the sky where a source with any flux S_ν is detectable). We can associate to these counts a Poisson uncertainty.

Suppose now that sources are uniformly distributed in Euclidean space, or, similarly, the sample is complete and homogeneous. Consider first sources with luminosities in the range L and $L + \mathrm{d}L$. In a survey to a limiting distance r, given by the inverse square law, $r = (L/4\pi S)^{1/2}$. The number of sources within distance r in the solid angle Ω is:

$$\begin{aligned} N(>S_\nu)\mathrm{d}L &= \frac{\Omega}{3} r^3 N(L)\,\mathrm{d}(L) \\ &= \frac{\Omega}{3}\left(\frac{L}{4\pi S}\right)^{3/2} N(L)\,\mathrm{d}(L) \end{aligned} \tag{2.27}$$

Integrating over the luminosity function of the sources,

$$N(>S_\nu) = \frac{\Omega}{3(4\pi)^{3/2}} S^{-3/2} \int L^{3/2} N(L)\,\mathrm{d}L \tag{2.28}$$

that become $N(>S_\nu) \propto S^{-3/2}$ if all sources have the same luminosity. This result is known as *integral Euclidean source counts* for any class of extragalactic object.

It will be noted that integral counts here described have the disadvantage that the number of sources counted to different flux densities are not independent since bright objects contribute to the counts at all lower flux densities. In other words in this method the flux-density data bins are not independent. For this reason it is statistically preferable to work in terms of *differential source counts* in which the sources counted in each flux density bin are independent. In this case we can obtain the counts as:

$$\mathrm{d}\,N(S) = N(S)\,\mathrm{d}S \propto S^{-5/2}\,\mathrm{d}S \tag{2.29}$$

It is useful to compare the observed counts and what we can predict using different world models. If we consider the world model with $\Omega_\Lambda = 0$ and the others described by $\Omega_0 + \Omega_\Lambda = 1$ with $\Omega_0 = 1,\ 0.6,\ 0.3$ and finite value of Ω_Λ we obtain that the predicted differential counts depart rapidly from the Euclidean expectation even at small redshifts. The major result is that the slope of the differential and integral source counts must be smaller than the Euclidean

prediction described in Eq. 2.28: if $N(\geq) \propto S^{\beta}$ then $\beta < 1.5$. This value can be used to interpret the observational results with the aim to find the model to describe the evolution of our Universe.

2.2 The Luminosity Function

The luminosity function (LF) is a fundamental statistical measure of galaxies' properties (Schmidt 1968), describing number density of galaxies as a function of their luminosity. The LF is an important tool for understanding the distribution of galaxies or quasars over cosmological time and so provide a census of their evolution. The galaxy luminosity function probes several aspects of the galaxy population, like:

- the evolution of stellar populations and star formation histories, e.g. Faber et al. (2007);
- the dependence of galaxy properties on environment, e.g. Croton et al. (2005);

As a consequence, studies of the redshift-dependent luminosity functions place relevant constraints on models of structure formation and galaxy evolution, e.g. Bower et al. (2006).

By definition, the LF, denoted by $\Phi(L)$ (in units of Mpc^{-3} per unit logaritmic interval of luminosity), describes the relative number of galaxies of different luminosities by counting them in a representative volume of the Universe which then measures the co-moving number density of galaxies per unit luminosity, L, such that

$$dN = \Phi(L)dLdV. \tag{2.30}$$

where dN is the observed number of galaxies within a luminosity range [L, L+dL] and volume dV. When working in luminosities it is common practice to apply log intervals of L. The quantity $\Phi(L)$ can be normalized such that

$$\int_0^{\infty} \Phi(L)dL = \rho, \tag{2.31}$$

where ρ is the number of objects per unit volume V, and thus $\Phi(L)dL$ gives the number density of objects within a given luminosity range. In general, the density function can be defined by $\rho(\mathbf{x})$, where $\mathbf{x}$ represents the 3D spatial cartesian co-ordinates such that the total number N of objects in the volume V is

$$N = \int_V \rho(\mathbf{x})\, d\mathbf{x} \tag{2.32}$$

However, it is common to compute ρ from the measured redshifts z and angular co-ordinates. Thus, for a given sample within a respective minimum and maximum redshift range $z_{\min}$ and $z_{\max}$ and solid angle Ω at a distance r it is possible to compute,

$$N = \int_{z_{\min}}^{z_{\max}} \rho(z)\frac{dV'}{dz}dz, \tag{2.33}$$

where $\rho(z)$ is now the density as a function of redshift and $dV' \equiv \Omega r^2 dr$ is a solid angle-integrated differential volume element (Johnston 2011).

The various techniques and estimators used to get the LF are numerous and most relatively easy to understand. We can ideally start doing a selection of galaxies with flux greater then an assigned value ($S_{\nu_{lim}}$, where ν stands for the mean frequency of the pass-band) in a particular spectral range (either in X-rays, optical, IR, radio). This selected sample is defined as "flux-limited" and includes only galaxies with flux $S_\nu > S_{\nu_{lim}}$ (another possible way is to get a "volume-limited" sample, but this is much more difficult and uncertain).

Evolution of source populations can be seen by a statistical change in the luminosities, or number densities or by a change in the shape of the luminosity function. Thus every class of cosmic sources will have a LF-shape that is specific to that class. When a survey sample is observed many different classes of galaxies will be detected and so the resulting LF will be the sum of all contributions from morphologically different objects. In the same way we can express the luminosity function in terms of number counts :

$$N(S_{\nu_{lim}}) = \int_0^\infty \int_0^{z_{max}:S(L,z)>S_{\nu_{lim}}} \phi(L,z)\,\mathrm{d}L\,\mathrm{d}z \tag{2.34}$$

or equivalently

$$N(S_{\nu_{lim}}) = \int_0^\infty \int_{L_{min}:S>S_{min}}^\infty \phi(L,z)\,\mathrm{d}L\,\mathrm{d}z \tag{2.35}$$

Now the problem is to examine the distribution of sources of the selected sample as well as to deal with the intrinsic incompleteness of a given galaxy survey. In the literature there is a plenty of examples about how scientists try to solve these issues in different cases and in the next sections we try to give an overview of such studies, then focusing on the particular methods implemented in our work.

2.3 Luminosity Function Estimators

From as early as the 1930s, astronomers have tried to quantify the statistical nature of the evolution and large-scale structure of galaxies by studying their luminosity distribution as a function of redshift - known as the galaxy luminosity function (LF). Accurately constructing the LF remains a popular and yet tricky pursuit in modern observational cosmology where the presence of observational selection effects due to e.g. detection thresholds in apparent magnitude, colour, surface brightness or some combination thereof can make any given galaxy survey *incomplete* and thus introduce biases into the LF.

Numerous sophisticated statistical approaches have been devised to tackle these issues; all have advantages, but not one is perfect. A comparison between results coming from different statistical procedure applied on the same sample of data can give a more unbiased LF estimate. This is what we try to do in our own work and the results will be presented in Chap. 6, 7. In this section we present the more generally used methods in LF estimates, some of them are actually exploit in our work: the $1/V_{max}$ estimator technique (Schmidt 1968); the parametric maximum likelihood estimator method (Sandage et al. 1979) with a modified version adapted to the bayesian statistics (Kelly et al. 2008); the stepwise maximum likelihood estimator technique (Efstathiou et al. 1988; Saunders et al. 1990), also known as the non-parametric maximum likelihood technique; and finally some more generalized methods that have emerged

within the last few years those propose a more rigorous statistical framework within which to determine the LF compared to some of the more traditional methods (e.g., Schafer (2007)). A brief description of each can be found below and a summary of advantages and disadvantages of each method is reported in Tab. 2.1.

2.3.1 The V/V_{max} Test and $1/V_{max}$ Estimator

One of the main difficulties in constructing an accurate LF from flux-limited survey samples is the problem of completeness and non-uniform distribution of sources in space. In addition we are hindered observationally by the notorious *Malmquist bias* effect. This means that we are biased to observe intrinsically brighter objects at higher redshifts and observe only the fainter objects over smaller nearby volumes.

Classically, when pioneering galaxy surveys were shallow, it was common practice to exclude clustered regions like the Virgo cluster and members of the Local Group to try and avoid biases in the shape and in the parameters of the LF. The *classical* method, as coined by Felten (1977), represents the first rudimentary binned approach to determining the LF by assuming that the distribution of sources within the sampled volume is spatially homogeneous, i.e. with no strong large-scale clustering. From this starting point we count the number of galaxies N within a volume V such that

$$\Phi \equiv \frac{N}{V} \tag{2.36}$$

The volume, $V(M)$, is calculated for the maximum distance that each galaxy with an absolute magnitude, M_i, could have and still remain in the sample.

Whilst this method is relatively straightforward to apply, its basic assumption about sample homogeneity is well understood to be a handicap. A straightforward development from the *classical*'s is the V/V_{max} test, a method to investigate the uniformity of the objects in space, firstly described by Kafka (1967) but more formally applied by Schmidt (1968). In order to apply this test we have to select a sample of objects that is considered complete within defined flux density or apparent magnitude limits and for which we know the distance, or redshift, for all the objects. The essence of the test is to ask whether or not the distribution of objects is statistically uniform in the accessible region of space defined by the observational selection criteria. Once selected the flux density limited sample, for each object having flux density S ($S > S_{lim}$), we may calculate the ratio V/V_{max}, where V is the volume of space enclosed by the redshift z of the object ($V = (\Omega/3)D^3$, where Ω is the solid angle of sampling) and V_{max} is the max volume within the object would still be included in the complete sample bringing its value of flux to that of the flux limit (in other words, fixed L_ν, we let D to increase until we get $S = S_{lim}$ or $D = D_{lim}$; a schematic construction of the test is shown in Fig. 2.1).

The volume V_{max} corresponds to the redshift z_{max} at which a source of intrinsic luminosity L would be observed to have flux density S_{lim}. We thus obtain:

$$V/V_{max} = \frac{\int_0^r D^2 \, \mathrm{d}r}{\int_0^{r_{max}} D^2 \, \mathrm{d}r} \tag{2.37}$$

where D is the distance measured and r and r_{max} are the comoving radial distance coordinates corresponding to z and z_{max} respectively. The volumes used in the test are comoving-coordinate volumes at the present epoch. Now, if we suppose that the distribution of galaxies

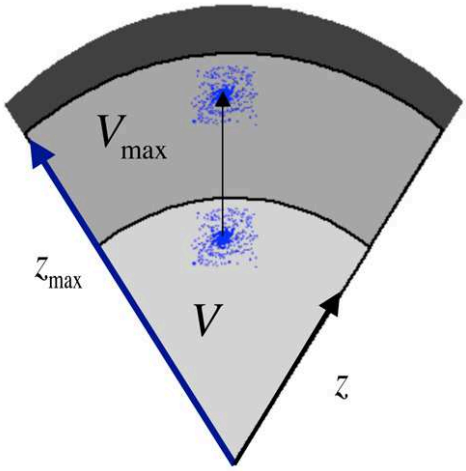

Figure 2.1: The construction of the V/V_{max} test. The basic construction of the method considers the radio between the volume V, in which a galaxy is observed to the volume V_{max}, the maximum volume the said galaxy could occupy and still be observed.

in space is uniform, the mean value of V/V_{max} is then:

$$\langle V/V_{max} \rangle = \frac{\int_0^{r_0} \left(\frac{V}{V_{max}} \right) D^2 \, \mathrm{d}r}{\int_0^{r_0} D^2 \, \mathrm{d}r} = \frac{1}{2} \tag{2.38}$$

Moreover, if the source distribution is uniform, the value of V/V_{max} should be uniformly distributed between 0 and 1, which is true if they have constant comoving number densities. Since the sample used is usually limited, it is common to evaluate V/V_{max} and compare this with the mean value 0.5 for an uniform distribution. If, however, the mean value of this ratio is different from 0.5, this is considered to be either a signature of incompleteness and/or an indication of evolution. A deviation from 0.5 means that there must be some gradient in the density (either positive or negative) outwards. Larger values than 0.5 mean that the density increases outwards and would imply a density evolution where galaxies were more numerous in the past; smaller values than 0.5 mean that the density decreases outward and would imply that galaxies were less numerous in the past. For a uniform distribution between 0 and 1, the standard deviation is:

$$\sigma_0 = 1/\sqrt{12} = 0.288 \tag{2.39}$$

When the number of sources N is sufficiently large, the central limit theorem may be used and so the probability distribution of $\langle V/V_{max} \rangle$ approaches a Gaussian distribution with standard deviation $N^{-1/2}$ of the original distribution, that is:

$$\sigma = \frac{\sigma_0}{N^{\frac{1}{2}}} \tag{2.40}$$

The procedure can be further extended to consider the distribution of objects within a particular redshift bin, what is referred to as a *banded* V/V_{max} test. These calculations can be carried out for all objects in the sample within given redshift limits. Correspondingly, luminosity functions as a function of cosmic epoch can be derived directly from the samples.

Schmidt (1968) introduced the intuitive and powerful $1/V_{max}$ estimator for LF evaluation as a simple extension of the V/V_{max} test. The quantities $V_{max}(L_i)$ are the maximum volumes within which the ith object in the catalog could lie, and still be in the catalog. V_{max} thus depends on the survey limits, the distribution of the objects in space, and the way in which detectability depends on distance. In the simplest case, a uniform distribution in space is assumed. Given the $V_{max}(L_i)$, an estimate of the LF is:

$$\Phi(B_{j-1} < L \leqslant B_j) = \sum_{B_{j-1}<L\leqslant B_j} \frac{1}{V_{max}(L_i)} \tag{2.41}$$

in which its value is computed in bins of luminosity, bounded by the B_j. Usually express in the differential form as

$$\phi_{1/V_{\rm max}}(L, z) = \frac{1}{\Delta L} \sum_{i=1}^{N} \frac{1}{V_{{\rm max},i}}, \tag{2.42}$$

where N is the number of objects within some volume-luminosity region. Errors in the LF can be evaluated using a Poisson statistical distribution:

$$\sigma_{\phi(L)} = \sqrt{\sum_{B_{j-1}<L\leqslant B_j} \frac{1}{(V_{max}(L_i))^2}} \tag{2.43}$$

This method implies binning of the luminosity data, a non-parametric technique, and as such does not need to assume an analytic form. It does however contain the underlying assumption that galaxies have a uniform distribution in space that can be tested with the already explained V/V_{max} test, but that still remains difficult to assess when there are evidence of incompleteness.

Nevertheless, the $V_{\rm max}$ estimators have evolved, been improved and refined over the years to accommodate the many different types of survey that have steadily grown in size and complexity. One of these is the one implemented in Page & Carrera (2000), the so called V_{est}, which we also used in our present work. They improved the method to take into account systematic errors in the $V_{\rm max}$ test introduced for objects close to the flux limit of a survey. As they point out, for evolutionary studies of galaxies the traditional approach, as extended by Avni et al. (1980) and Eales (1993), is very common but can distort the apparent evolution of extragalactic populations. Through the use of Monte Carlo simulations, with a sample of 10,000 objects and simulating an un-evolving two-power law model X-ray LF, Page & Carrera (2000) compare the $1/V_{\rm max}$ estimate of the differential LF given by Eq. 2.42 to their improved binned approximation of the $\phi_{\rm est}$, which assumes that ϕ does not change significantly over the luminosity and redshift intervals ΔL and Δz, respectively, and is defined as

$$\phi_{\rm est} = \frac{N}{\int_{L_{\rm min}}^{L_{\rm max}} \int_{z_{\rm min}}^{z_{\rm max}(L)} (dV)/(dz) dz dL}, \tag{2.44}$$

where N is the number of objects within some volume-luminosity region.

A graphical representation of the difference between the volume-luminosity test in the classical $1/V_{max}$ estimate and that modified by Page & Carrera (2000) is reported in Fig. 2.2.

The shaded region in the plot (a) is the region of the volume-luminosity plane in the interval ΔL Δz that has been surveyed. This shaded area corresponds to the double integral

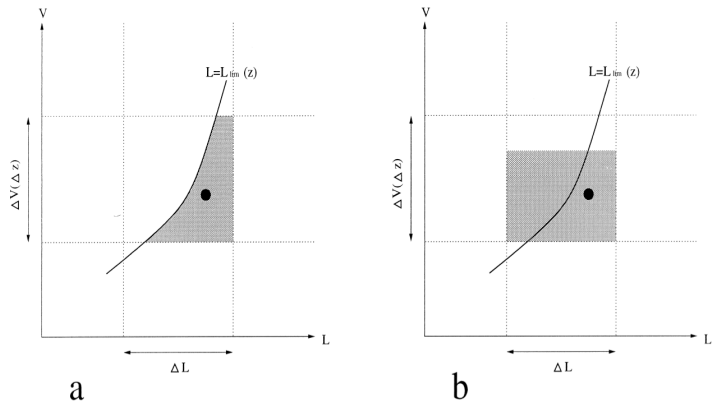

Figure 2.2: Volume-luminosity space 'available' to an object (the black dot) in (a) ϕ_{est} (Page & Carrera 2000) and (b) $\phi_{1/V_{max}}$ (Schmidt 1968). The line $LL_{lim}(z)$ is the minimum detectable luminosity of an object of redshift z and is determined by the flux limit of the survey

at the denominator in Eq. 2.44. The shaded region in Fig. 2.2 (b) has an area equal to $[\Delta L V_{max}(i)]$ for object i, represented by the black spot. This area is clearly not the same as that corresponding to the really surveyed region (the shaded region in Fig. 2.2 (a)) and hence there is no guarantee that $\phi_{1/V_{max}}$ will give a good estimate of ϕ. However it is known that $1/V_{max}$ estimate is a good evaluation of ϕ when the entire redshift-luminosity interval ΔL Δz corresponds to objects brighter than the flux limit of the survey (i.e. in Fig. 2.2 the curve $L = L_{lim}(z)$ passes completely above the ΔL Δz region) or when ΔL is very small. The first situation generally applies for luminosity functions of objects which are much brighter than the flux limits. Hence for luminosity functions in most of the redshift intervals, the two methods will produce the same results, particularly for the highest luminosity bins of any given redshift shell. At the lowest luminosities in each redshift shell, where objects are close to the survey flux limit and the shaded portion of volume-luminosity space in Fig. 2.2 is much smaller than the rectangular ΔL Δz region, the two methods produce the most different results.

2.3.2 Parametric Maximum Likelihood Estimator

The maximum likelihood estimator (MLE) is a bery old concept of statistical analysis, but only with Fisher (1912) we see the first formal derivation. In terms of its application within the context of observational cosmology, it was Sandage et al. (1979), hereafter STY, who were the first to see it as a powerful approach to estimate galaxy LFs. This is a parametric technique which therefore assumes an analytical form for the LF and thus eliminates the need for the binning of data (as usually required by most non-parametric methods, e.g Sec.2.3.1). The equally popular non-parametric counterpart of the MLE, the step-wise maximum likelihood (SWML), is discussed below.

The derivation of the MLE begins by considering x, a continuous random variable, that is described by a probability distribution function (PDF) given by

$$f(x;\theta), \tag{2.45}$$

where θ represent the parameter we wish to estimate. As we shall see, in practice θ represents

more than one parameter of the LF, or a parameter vector. If x represents our observed data then the likelihood function, $\mathcal{L}$, can be written as

$$f(x_1, x_2, ..., x_N|\theta_1, \theta_2, ..., \theta_k) = \mathcal{L} = \prod_{i=1}^{N} f(x_i; \theta_1, \theta_2, ..., \theta_k) \tag{2.46}$$

where x_i are N independent observations. It is often the case that the likelihood function is expressed in terms of the logarithmic likelihood such that

$$\ln\{\mathcal{L}\} = \sum_{i=1}^{N} \ln f(x_i; \theta_1,\theta_2, ..., \theta_k) \tag{2.47}$$

Constraining the $\theta_1, \theta_2, ..., \theta_k$ parameters is, in principle, a straightforward matter of maximizing the likelihood function $\mathcal{L}(\theta)$ or $\ln\{\mathcal{L}(\theta)\}$ such that

$$\frac{\partial(\mathcal{L} \text{ or } \ln\{\mathcal{L}\})}{\partial\theta_j} = 0, \quad j = 1, 2, ..., k \tag{2.48}$$

In the context of estimating the parameters of the LF we consider a galaxy at redshift z for which we can define the cumulative luminosity function (CLF) and thus determine the probability that the galaxy will have an absolute magnitude brighter than M as

$$p(M|z) = \frac{\int_{-\infty}^{M} \phi(M')\rho(z)f(M')dM'}{\int_{-\infty}^{\infty} \phi(M')\rho(z)f(M')dM'}, \tag{2.49}$$

where $\rho(z)$ is the density function for the redshift distribution, $f(M')$ is the completeness function which for a 100% complete survey would be

$$f(M') = \begin{cases} 1, & M_{\text{lim}}^{\text{bright}} \leq M' \leq M_{\text{lim}}^{\text{faint}} \\ \\ 0, & \text{otherwise.} \end{cases} \tag{2.50}$$

It follows that the probability density for detected galaxies is given by the partial derivative of $p(M, z)$ with respect to M,

$$p(M_i, z_i) = \frac{\partial p(M, z)}{\partial M} = \frac{\phi(M_i)}{\int_{M_{\text{faint}(z_i)}}^{M_{\text{bright}(z_i)}} \phi(M')dM'} \tag{2.51}$$

Note that the density functions have cancelled, thus rendering the technique insensitive to density inhomogeneities. Finally, the likelihood is maximized to give

$$\mathcal{L} = \prod_{i=1}^{N} p(M_i, z_i). \tag{2.52}$$

Most commonly a Schechter function is assumed where the parameters that we wish to estimate are α and M_*. These two parameters are defined by

$$\Phi(M) = 0.4\ln(10)\phi^* \frac{\left(10^{0.4(M^*-M)}\right)^{(\alpha+1)}}{e^{10^{0.4(M^*-M)}}}. \tag{2.53}$$

that is the Schechter function in terms of absolute magnitude (in Eq. 2.62 is reported the Schechter function in terms of absolute luminosities connected to this by the relation $\frac{L}{L_*} = 10^{-0.4(M-M_*)}$). MLE can be also used applying other analytic functions as illustrated in Sec. 2.3.3.

Although the MLE method has become more popular than other traditional non-parametric methods, there are aspects not to be overlooked. This approach typically assumes a restrictive and somewhat ad hoc parametric form and has not been used to give an estimate of the normalization parameter ϕ^* of the LF that consequently has to be estimated by independent means. Some examples of this independent procedure are reported in Davis & Huchra (1982) and later Blanton et al. (2003). Furthermore confidence intervals based on the errors derived from the various procedures are typically derived by assuming that the uncertainties on the LF parameters have a Gaussian distribution. While this is valid for very large sample sizes, it is not necessarily a good approximation for finite and small samples. This is particularly problematic if one is employing the best-fit results to extrapolate the LF beyond the bounds of the selection function. Motivated by this issues, Kelly et al. (2008) have developed a *Bayesian* method for estimating the LF. They derive the likelihood function of the LF by relating the observed data to the true LF, assuming some parametric form, and derive the posterior probability distribution of the LF parameters, given the observed data and describe a Markov chain Monte Carlo (MCMC) algorithm for obtaining random draws from the posterior distribution. These random draws allow one to estimate the posterior distribution for the LF, as well as any quantities derived from it. The MCMC method therefore allows a straightforward method of calculating uncertainties on the parameters of the LF. Because the Bayesian approach is valid for any sample size, one is therefore able to place reliable constraints on the LF and related quantities even below the survey flux limits. We report in the following Sec. 2.3.3 details on the Kelly et al. (2008) approach, which turns out to be a powerful tool in the LF computation as reported by Patel et al. (2012).

2.3.3 Bayesian Approach to the Parametric Maximum Likelihood Estimator

In maximum likelihood analysis, one is interested in finding the estimate that maximizes the likelihood function of the data. For a given statistical model, parameterized by θ, the likelihood function, $p(x|\theta)$, is the probability of observing the data, denoted by x, as a function of the parameters θ. In Bayesian analysis, one attempts to estimate the probability distribution of the model parameters, θ, given the observed data x. The probability distribution of θ given x is related to the likelihood function as

$$p(\theta|x) \propto p(x|\theta)p(\theta) \tag{2.54}$$

where $p(x|\theta)$ is the likelihood function of the data, and the term $p(\theta)$ is the prior probability distribution of θ; the result, $p(\theta, x)$, is called the posterior distribution. The prior distribution,

$p(\theta)$, should convey information known prior to the analysis. In general, the prior distribution should be constructed to ensure that the posterior distribution integrates to 1, but does not have a significant effect on the posterior. In particular, the posterior distribution should not be sensitive to the choice of prior distribution, unless the prior distribution is constructed with the purpose of placing constraints on the posterior distribution that are not conveyed by the data. The contribution of the prior to $p(\theta|x)$ becomes negligible as the sample size becomes large.

From a practical standpoint, the primary difference between the maximum likelihood approach and the Bayesian approach is that the maximum likelihood approach is concerned with calculating a point estimate of θ, while the Bayesian approach is concerned with mapping out the distribution of θ in the parameter space. The maximum likelihood approach uses an estimate of the sampling distribution of $\hat{\theta}$ to place constraints on the true value of θ. In contrast, the Bayesian approach directly calculates the probability distribution of θ, given the observed data, to place constraints on the true value of θ.

In terms of LF evaluation, we can easily understand that LF estimate is related to the probability density of (L, z)

$$p(L, z) = \frac{1}{N}\phi(L, z)\frac{dV}{dz} \tag{2.55}$$

where N is the total number of sources in the observable universe and is given by the integral of ϕ over L and $V(z)$; $p(L, z)\mathrm{d}L\mathrm{d}z$ is the probability of finding a source in the range $L, L+\mathrm{d}L$ and $z, z+\mathrm{d}z$. Eq. 2.55 separates the LF into its shape, given by $p(L, z)$, and its normalization, given by N. Once we have an estimate of $p(L, z)$, we can easily convert this to an estimate of $\phi(L, z)$ using Eq. 2.55.

In general it is easier to work with the probability distribution of L and z instead of directly with the LF because $p(L, z)$ is more directly related to the likelihood function. $\phi(L, z)$ can be described,as we have seen, by a parametric form with parameter θ, then we can derive the likelihood function for the observed data. As we already said, the likelihood function is the probability of observing one's data, given the assumed model. The presence of flux limits and various other selection effects can make this difficult, as the observed data likelihood function is not simply given by Eq. 2.55. In this case, the set of luminosities and redshifts observed by a survey gives a biased estimate of the true underlying distribution, since only those sources with L above the flux limit at a given z are detected. In order to derive the observed data likelihood function, it is necessary to take the survey's selection method into account. This is done by first deriving the joint likelihood function of both the observed and unobserved data, and then integrating out the unobserved data. Then, the probability $p(L, z)$ becomes :

$$p(L, z|\theta) = \frac{\phi(L, z|\theta)p(selected|L, z)}{\lambda}\frac{dV}{dz} \tag{2.56}$$

where $p(selected|L, z)$ stand for the probability connected with the selection functors of the survey and λ is the expected number of sources and is determined by:

$$\lambda = \iint \phi(L, z|\theta)p(selected|L, z)dlogL\frac{dV}{dz}dz \tag{2.57}$$

where the integrals are taken over all possible values of redshifts and luminosities.

Because the data points are independent, the likelihood function for all N sources in the universe is

$$p(L, z|\theta) = \prod_{i=1}^{N} p(L_i, z_i|\theta) \tag{2.58}$$

In reality, we do not know the luminosities and redshifts for all N sources, nor do we know the value of N, as our survey only covers a fraction of the sky and is subject to a selection function. As a result, our survey only contains n sources. Because of this, the selection process must also be included in the probability model, and the total number of sources, N, is an additional parameter that needs to be estimated. Then the likelihood becomes:

$$p(d|\theta) = p(N, \{L_i, z_i\}|\theta) = p(N|\theta)p(\{L_i, z_i\}|\theta) \tag{2.59}$$

Where $p(N|\theta)$ is the probability of observing N objects and $p(\{L_i, z_i\}|\theta)$ is the likelihood of observing a set of L_i and z_i both given the model LF. Is it possible to assume that the number of sources detected follows a Poisson distribution (Patel et al. 2012), where the number of detectable sources is given by λ. Then, the term $p(N, \{L_i, z_i\}|\theta)$ could be written as the product of individual source likelihood function since each data point is independent.

$$p(N|\theta)p(\{L_i, z_i\}|\theta) = \frac{\lambda^N e^{-\lambda}}{N!} \prod_{i=1}^{N} \Phi(L, z|\{\theta\})p(selected|L, z)\frac{dV}{dz} \tag{2.60}$$

Then we can use the likelihood function for the LF to perform bayesian inference by combining it with a prior probability distribution, $p(\theta)$ to compute the posterior probability distribution, $p(\theta|d_i)$ given by the Bayes' theorem :

$$p(\theta|d_i) = \frac{p(\{d_i\}|\{\theta\})p(\{\theta\})}{\int p(\{d_i\}|\{\theta\})p(\{\theta\})\mathrm{d}\theta} \tag{2.61}$$

where the denominator is the bayesian evidence and is determined by integrating the likelihood over the prior parameter space. This last is needed to normalize the posterior distribution.

Calculating the Bayesian evidence is computationally expensive since it involves integration over n-dimensions for an n parameter LF model. Therefore, Markov chain Monte Carlo (MCMC) methods, used to examine the posterior probability, perform a random walk through the parameter space to obtain random samples from the posterior distribution. MCMC gives as a result the maximum of the likelihood (see App. A for more details), but an algorithm is needed to investigate in practice the region around the maximum. Kelly et al. (2008) proposed to use the Metropolis-Hastings algorithm (MHA; Metropolis et al. (1953), Hastings (1970)) in which a proposed distribution is used to guide the variation of the parameters. The algorithm uses a proposal distribution which depends on the current state to generate a new proposal sample. The algorithm needs to be tuned according to the results and the number of iterations as well as the parameter step sizes change from a case to another. Once we obtain the posterior distribution, we have the best solution for each of the parameters describing the LF model that we have chosen at the beginning.

Parameterising the Luminosity Function

Using the popular STY79 method, as well the one based on the bayesian statistics introduced by Kelly et al. (2008), implies to assume a parametric form of the LF from the beginning

of the process able to describe the observed data. This choice is not strainghtforward and during years the LF models chosen by scientist varied.

The most common of these models is

- **The Schechter Function** named after the reference work by P. Schechter (Schechter (1976), see also Press & Schechter (1974)) who originally derived it in the form of the mass function during their studies of structure formation and evolution. It is typically written in the form given by

$$\Phi(L)dL = \phi^* \left(\frac{L}{L_*}\right)^{\alpha} \exp\left(\frac{-L}{L_*}\right) \frac{dL}{L_*}, \tag{2.62}$$

  where, ϕ^* is a normalization factor defining the overall density of galaxies, usually quoted in units of $h^3\mathrm{Mpc}^{-3}$, and L_* is the characteristic luminosity. The quantity α defines the faint-end slope of the LF and is typically negative, implying relatively large numbers of galaxies with faint luminosities. LF studies of the local Universe ($z \lesssim 0.2$) have estimated this parameter being close to $\alpha \sim -1.0$, formally implying an infinite number of faint galaxies.

Although the Schechter form of the LF has been very successful as a generic fit to a wide variety of survey data, it has also been shown that galaxy surveys sampled, for example, in the IR have yielded LFs that do not seem to be fitted by a standard Schechter form. If the optical/near-IR LLFs display the classical "Schechter" exponential convergence at high luminosities, LLFs for galaxies selected at longer wavelengths show flatter and flatter slopes. This flattening is progressive with λ and it is connected to the increasing with λ of the incidence of the starburst contribution to the luminosity (Franceschini 2000).

Here we report some of the other analytic forms that have been used to describe the general shape of the luminosity function:

- **The Log Gaussian Function**: This function is introduced by (Saunders et al. 1990) to fit IRAS (see Chap. 3) IR LF and is defined as

$$\Phi(L) = \phi_* \left(\frac{L}{L_*}\right)^{1-\alpha} \exp\left[-\frac{1}{2\sigma^2}\log^2\left(1+\frac{L}{L_*}\right)\right] \tag{2.63}$$

  Usually this function is called **modified Shechter funtion** since the form is very similar to the Schechter one defined in Eq. 2.62. Another similar function has been introduced more recently by Serjeant & Harrison (2005) to predict the luminosity functions at a number of IR/Submm λs from 70 to 850 μm, and is defined as

$$\Phi(L) = \phi_* \ln(10) \left(\frac{L}{L_*}\right)^{1-\alpha} \exp\left(-\frac{L}{L_*}\right) \tag{2.64}$$

- **The Cumulative Two Power Law Function**: The cumulative two power law function tries to represent the LF by two power laws in the form of (Yahil et al. 1991)

$$\Psi(L) = \phi_* \left(\frac{L}{L_*}\right)^{\alpha} \exp\left(1+\frac{L}{L_*}\right)^{-\beta} \tag{2.65}$$

- **The Double Power Law Function**: The double power law function represent the LF by two power laws in the form of Rush & Malkan (1993) and Franceschini et al. (2001):

$$\Psi(L) = \phi_* L^{1-\alpha} \left(1 + \frac{L}{L_*\beta}\right)^{-\beta} \tag{2.66}$$

More recently Serjeant & Harrison (2005), together with the previous Eq. 2.64 a double power law given by:

$$\Psi(L) = \frac{\phi_*}{\left(\frac{L}{L_*\beta}\right)^{\beta} + \left(\frac{L}{L_*\beta}\right)^{\gamma}} \tag{2.67}$$

2.3.4 Non-Parametric Step-Wise Maximum Likelihood Estimator

The step-wise maximum likelihood method (SWML) was introduced by Efstathiou et al. (1988) and represents the non-parametric version of the STY79 method. The technique does not depend on an analytical form for $\phi(L)$ (or $\phi(M)$ in units of absolute magnitude). Instead the LF is effectively parameterized as a series of N_p step functions allowing to define the following initial setup:

$$\Phi(L) = \phi_k, \qquad L_k - \frac{\Delta L}{2} < L < L_k + \frac{\Delta L}{2}, \tag{2.68}$$
$$\text{where, } k = 1, ..., N_p$$

The differential LF can then be expressed as:

$$\phi(L) = \sum_{i=1}^{N} \phi_i W(L_i - L), \tag{2.69}$$

where $W(x)$ represents two window functions,

$$W(x) \equiv \begin{cases} 1, & -\frac{\Delta L}{2} \leq x \leq \frac{\Delta L}{2}, \\ 0, & \text{otherwise.} \end{cases} \tag{2.70}$$

Therefore, it can be shown that the expression for the step-wise likelihood is given by

$$\mathcal{L}(\{\phi_i\}_{i=1,...,I} \,|\, \{L_k\}_{l=1,...,K}) = \prod_{k=1}^{N_{\rm obs}} \frac{\sum_{l=1}^{K} W(L_l - L_k)\phi_l}{\sum_{l=1}^{K} \phi_l H(L_{\rm lim}(z_k) - L_l)\Delta L} \tag{2.71}$$

which is more commonly written and easily computed in logarithmic form:

$$\ln\mathcal{L} = \sum_{i=1}^{N} W(L_i - L_k) \ln\phi_k - \sum_{i=1}^{N} \ln\left\{\sum_{j=1}^{N_p} \phi_j \Delta L H[L_i = L_{min}(z_i)]\right\} + const \tag{2.72}$$

where the Heaviside function $H(x)$ reads

$$H(x) = \begin{cases} 0, & x \leq -\Delta L/2, \\ (x/\Delta L + 1/2), & -\Delta L/2 \leq x \leq \Delta L/2, \\ 1, & x \geq \Delta L/2. \end{cases} \tag{2.73}$$

It is clear that, since the likelihood function involves ratios of ϕ_k, we need to define some sort of constraints to fix the otherwise arbitrary normalization. Efstathiou et al. (1988) defines the following constraints:

$$g = \sum_k \phi_k (L_k/L_f)^\beta \Delta L - 1 = 0 \tag{2.74}$$

where L_f is a fiducial luminosity and β is a constant. The constraint is introduced into the likelihood equation using a Lagrangian multiplier λ, thus we maximize $\ln\mathcal{L}' = \ln\mathcal{L} + \lambda g(\phi_k)$ with respect to the ϕ_k and λ.

The errors on the LF parameters are assumed to be asymptotically normally distributed giving a covariance matrix:

$$\mathrm{cov}(\phi_k) = [\mathbf{I}(\phi_k)]^{-1}, \tag{2.75}$$

where $\mathbf{I}(\phi_k)$ is the information matrix (see Eadie et al. 1971):

$$I(\phi_k) = -\left[\begin{array}{cc} \partial^2 \ln\mathcal{L}/\partial\phi_i\partial\phi_j + (\partial g/\partial\phi_i)(\partial g/\partial\phi_i) & \partial g/\partial\phi_j \\ \partial g/\partial\phi_i & 0 \end{array} \right]_{\phi=\phi_k} \tag{2.76}$$

Strauss & Willick (1995) noted two drawbacks of the method. The first concerns discretization of the LF using step functions. A bias is introduced into the selection function due to having discontinuous first derivatives. The authors instead suggest interpolating through the steps and then calculating the selection function to reduce this bias. The second drawback is the sensitivity of the LF to the choice of bin size. Koranyi & Strauss (1997) showed by example simulation that if the total number of bins is too small, this can dramatically underestimate the faint-end slope of the LF.

Heyl & Hernquist (1997) extended the use of the SWML method by generalizing it in a similar way as Avni et al. (1980) did for $V/V_{\max}$ by combining various surveys with different magnitude limits, coherently. Moreover, this extension also provided an absolute normalization and was used to probe density evolution in the LF by spectral type.

Springel & White (1998) also explored evolution and provided a variation of the method that instead of employing simple step functions to model the desired functional form, models the selection function as a series of linked piecewise power laws. This method provides accurate information on the shapes of the selection function and the LF and has a number of computational advantages. For example, it does not require iterative solutions and provides error estimates easily.

2.3.5 *Semi-Parametric* Estimators

Despite the continuing popularity of the classical statistical approaches as $1/V_{max}$ or MLE, there has been renewed interest in the exploration of more innovative statistical approaches, motivated by the intent to overcome the biases of the precedent methods. One of these new techniques can be considered a mixture of the parametric and non-parametric estimators and so called a $semi-parametric$ estimator. We report here the major characteristics of the $semi-parametric$ method developed by Schafer (2007) which we implemented in our own analysis.

Schafer (2007) introduced the *semi-parametric* method in order to estimate the luminosity functions given redshift and absolute magnitude (luminosity) measurements from an irregularly truncated sample of objects (i.e. a flux-limited sample), which actually represents a

bivariate density estimation problem. The problem connected with truncated data is illustrated with real data in Fig. 2.3 taken from Schafer (2007). In a truncated or flux-limited sample, only objects with apparent magnitude within some range are observable. When this bound on apparent magnitude is transformed into a bound on absolute magnitude, or absolute luminosity, the truncation bounds take an irregular shape, and may vary with redshift; beside, k-correction further complicates this boundary.

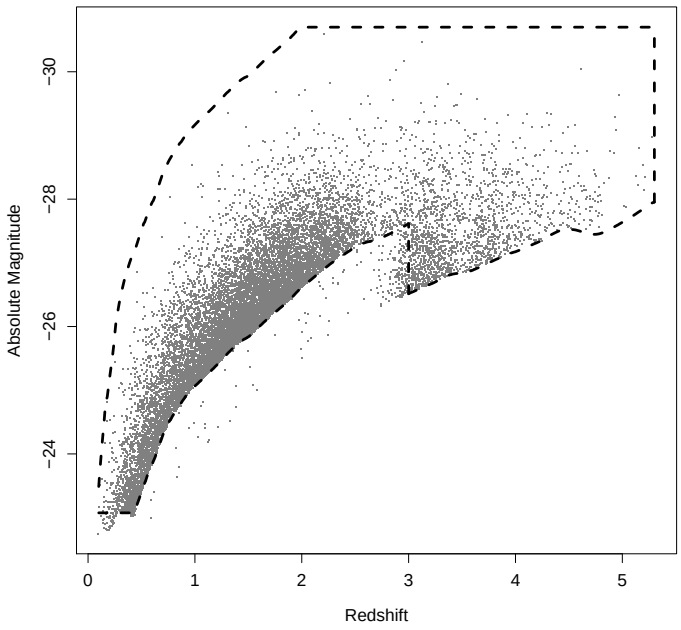

Figure 2.3: Distribution of a truncated sample of objects in the Absolute Magnitude - z plane (the same would be in the plane Absolute luminosity - z), taken from Schafer (2007). After applying all the cut in the sample selection based on flux limit of the survey (original selection is explained in Richards et al. (2006)), the objects used to evaluate the LF are the one within the irregular dashed region. It is clear that we need to pay particular attention while dealing with such distribution of objects by a statistical point of view.

The advantages of this methods are summarized as: (1) it does not assume a strict parametric form for the bivariate density (differently to the parametric MLE); (2) it does not assume an independence between redshift and absolute magnitude; (3) it does not require to split the data into arbitrary bins (contrary to the non-parametric MLE); (4) it naturally incorporates a varying selection function. This is accomplished by writing the bivariate density $\phi(z, L)$ as:

$$\log\phi(z, L) = f(z) + g(L) + h(z, L, \theta) \tag{2.77}$$

where $h(z, L, \theta)$ assumes a parametric form and is introduced to model the dependence between the redshift z , the absolute luminosity L and the real valued parameter θ. The function f and g are estimated in a completely free-form way. The advantage of this method is that we can estimate the evolution of the LF without assuming a strict parametric form for the bivariate density, while a parametrized form is taken to model only the relation between the redshift

and the absolute luminosity.

Besides, Schafer (2007) used a different concept of likelihood called *local* likelihood (Loader 1996; Hjort & Jones 1996) which differs to the *global* likelihood usually considered by the other statistical methods. Very briefly, the difference between these likelihoods can be summarized as follow. Be $X = (X_1, X_2, ..., X_N)$ observations of independent, identically distributed random variables from a distribution with density f_0. The classical *global* likelihood estimate chooses a single estimate from a class F of candidates for f_0 and is maximized as:

$$\sum_{j=1}^{n} \log f(X_j) - \left\{ n \left[\int f(x) \mathrm{d}x - 1 \right] \right\} \tag{2.78}$$

The *local* likelihood, instead of seeking the single member of the class to be the estimate of f_0, approximate $f_0(x)$ for x near u, yielding the *local* estimate $\hat{f}_u$. It is not important that $\hat{f}_u(x)$ is not a good estimate of $f_0(x)$ for x far from u, since many such local estimates will be found and then smoothed together (see Fig. 2.4 for a graphical illustration); then the final likelihood estimator $\hat{f}_L L$ will be, by smoothing together the local estimates:

$$\hat{f}_{LL}(x) \equiv \frac{\left[\sum\limits_{u \in G} K^*(x, u, \lambda) \hat{f}_u(x) \right]}{\sum\limits_{u \in G} K^*(x, u, \lambda)} \tag{2.79}$$

where G forms a grid $u \in G$ of equally spaced values of a gaussian density with mean 0 and variance 1; the term $K^*(x, u, \lambda)$ is a kernel function defined as such that:

$$\sum_{u \in G} K^*(x, u, \lambda) = 1 \quad \forall x \tag{2.80}$$

If G is sufficiently large, the amount of smoothing is completely dominated by the parameter λ of the kernel function. This kind of likelihood has the advantage that tuning the value of λ it is possible to adjust the amount of smoothness in the estimator.

Coming to the bivariate density estimation, the local density likelihood is incorporated into the case for flux-limited survey data by including the dependence between the redshift, z, and absolute luminosity, L. The simplest first order approximation to $\mathbf{h}$ is made from Eq. 2.77 by making the following assumption of a linear proportionality on both z and L, such that $h(z, L, \theta) = \theta z L$.

After an extensive derivation, a global criterion for the likelihood is found to be given by:

$$\begin{aligned}
\mathcal{L}^*(\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{z}, \boldsymbol{L}, \theta) \equiv \sum_{j=1}^{n} w_j \Bigg(& \sum_{u \in G} K^*(z_j, u, \lambda) a_u(z_j) \\
& + \sum_{u \in \mathcal{G}} K^*(L_j, v, \lambda) \boldsymbol{b_v}(L_j) + \boldsymbol{h}(z_j, L_j, \theta) \\
& - \int_{\mathcal{A}} \Bigg\{ \exp(\boldsymbol{h}(z, L, \theta)) \left[\sum_{u \in G} K^*(L, v, \lambda) \exp(\boldsymbol{b_v}(L)) \right] \\
& \times \left[\sum_{u \in G} K^*(L, v, \lambda) \exp(\boldsymbol{a_v}(L)) \right] dL\, dz \Bigg\} \Bigg),
\end{aligned} \tag{2.81}$$

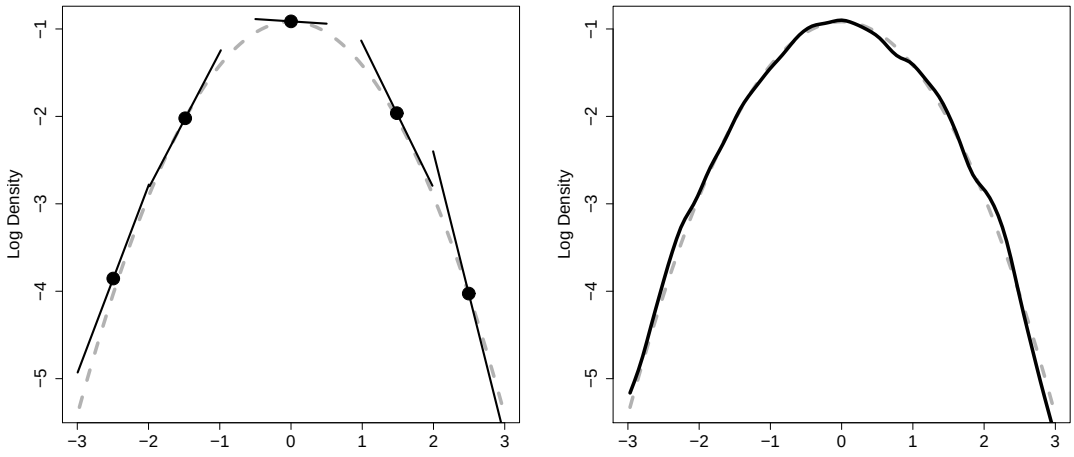

Figure 2.4: Illustration of local likelihood density estimation, taken from Schafer (2007). The dashed curves in both panel are the logarithm of the Gaussian density with mean 0 and variance 1 (f_0). **Left** : the local linear estimates $\hat{f}_u$ of the density for each u. **Right**: $\hat{f}_L L$ estimate obtained by smoothing together the local estimates shown in the left panel.

where $\mathbf{a_u}(z)$ and $\mathbf{b_v}(L)$ are degree p polynomials which form part of the smoothing term of K^* for local estimates. $\mathcal{A}$ defines the region of the (z, L) plane outside which the data are truncated. The quantity w_j is a weighting to take incompleteness into account and is defined as the inverse of the selection function.

It is important to notice that this method assumes *complete* data-set in the un-truncated region and it is necessary to pay attention to this issue when applying the method to samples that suffer for some kind of incompleteness. We report in Fig. 2.5 the results obtained by Schafer (2007) when applying this method to evaluate the LF evolution for a quasar sample selected by Richards et al. (2006). For more details about LF evolution and its evaluation see Sec. 2.4

2.4 Luminosity Functions & Cosmological Evolution

An intrinsic problem when studying galaxy evolution is our inability to trace the individual evolutionary process of a galaxy through cosmological time. Since current theories of galaxy formation and evolution appear to be far from converging on a convincing and accepted scheme, studies of the statistical properties of populations of galaxies at different epochs are still very much needed to constrain them. From a purely practical point of view what we can try to investigate is how the distribution of galaxies varies in luminosity and density across cosmic time and thus investigate the luminosity function as a function of redshift.

When talking about evolution, we refer in practice to two major trends:

- *Pure Density evolution (PDE).* This means to study the mean number density as a function of redshift. The PDE scenario assumes that galaxies were more numerous in

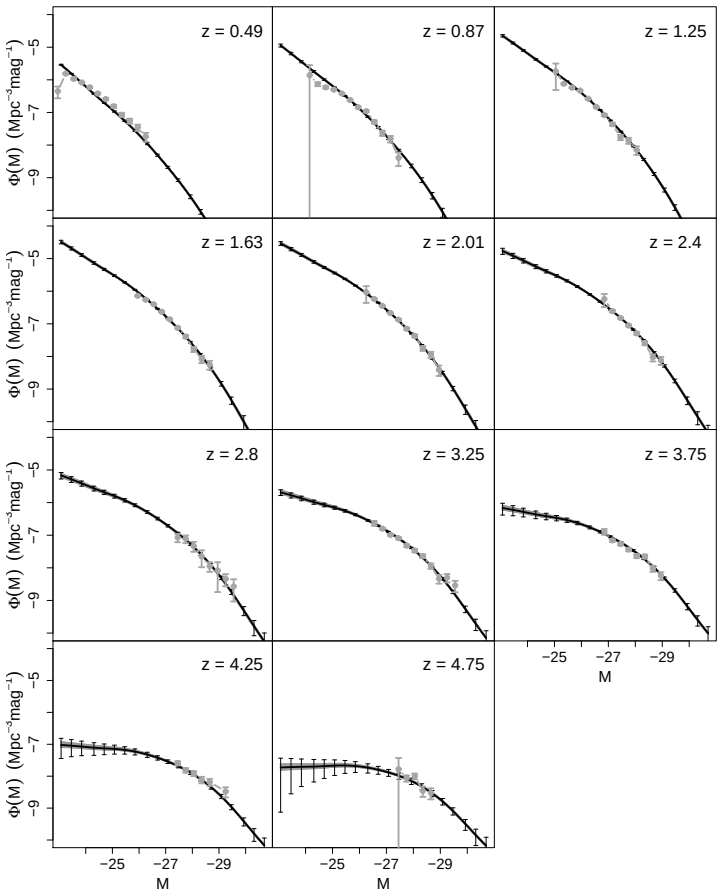

Figure 2.5: Estimates of the LF at different redshift using the *semi-parametric* method by Schafer (2007). For details about the plot see Schafer (2007).

the past, while keeping roughly the same luminosity distribution (e.g. due to merging). A positive PDE means that the galaxies were more numerous in the past.

- *Pure Luminosity evolution (PLE).* In this case the comoving density of the source population keeps constant in time, while it is the luminosity to change coherently with time. A positive PLE means that the galaxies were more luminous in the past.

More likely, of course, the total evolution of galaxies should be considered as the combined effect of these two scenario.

There are different approaches applied to constrain evolution. The one more commonly used and that we applied to our study is the method that adopts a strict parametric form based on some physical assumptions regarding our current understanding of luminosity and/or number (density) evolution. The evolutionary scenarios previously described are actually parametrized as follow:

$$\phi_*(z) = \phi_*(0)(1+z)^{\alpha_D} \tag{2.82}$$

$$L_*(z) = L_*(0)(1+z)^{\alpha_L} \tag{2.83}$$

α_D and α_L are, respectively, the density and luminosity evolution parameter, while L_* and ϕ_* are the parameters previously seen describing the analytic models to the LFs (Sec. 2.3.3). The PLE is sometimes reported in units of absolute magnitudes: $M_*(z) = M_*(0) - E(z) = M_*(0) - \alpha_L 2.5\log(1+z)$. Fig. 2.6 reports how the evolution factors would change the shape of the LF when it is described by a Schechter function, Eq. 2.62. In the figure α_D and α_L are γ and β respectively. The panel indicates the various components that make up the shape of the LF i.e. the power-law slope at the faint end (the steepness of which is governed by α); M_* (or L_* in units of absolute luminosity), which characterizes the turnover from the faint end to the bright-end; and the exponential cut-off which constrains the bright end of the LF. Panel (b) then shows how pure luminosity evolution (PLE) affects the shape of the LF by incorporating the PLE model into the Schechter function for a range of β values given by [1.5, 1.0, 0.5, -0.5, -1.0 , -1.5] at a fixed redshift of $z = 0.2$. These are indicated, respectively, from left to right on the plot by the dotted lines. The red solid line shows the universal LF from panel (a). As one would expect, the LF is affected only in its position in magnitude with no shift in number density. Panel (c) now shows the case where only pure density evolution (PDE) is present by adopting a range of γ given by [6.0, 4.0, 2.0 and -2.0, -4.0, -6.0] as, respectively, indicated by the dotted lines from top to bottom in the plot. It is clear that the shape of the LF is now dominated by shifts in the ϕ direction.

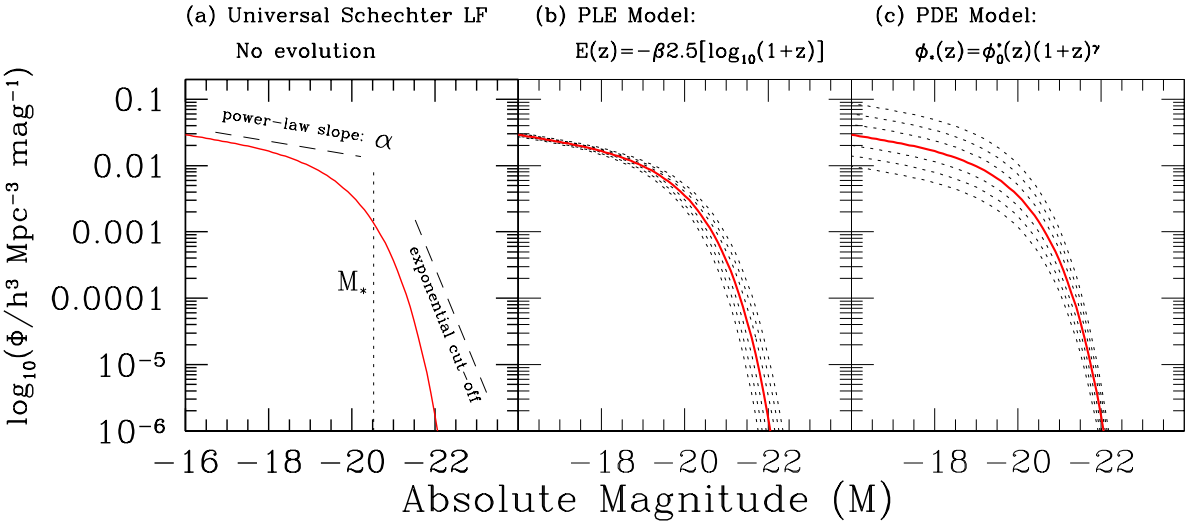

Figure 2.6: The characteristic shape of the Schechter luminosity function with evolution scenarios (Johnston 2011). The LF has been modeled using constrained LF parameter values based on the 2dFGRS Norberg et al. (2002) results. In panel (a) this implies: $\alpha = -1.21$, $M_* = -19.61$ and $\phi_* = 1.61 \times 10^{-2} h$ Mpc^{-3}. The steepness of the faint-end slope is determined by the α parameter and the characteristic magnitude, M_*, indicates the 'knee' of the LF. Panel (b) shows the effect of introducing a pure luminosity evolution (PLE) model. In terms of magnitudes this model is shown as $E(z)$ and is dependent on the evolutionary parameter β. The red line shows the LF from panel (a) and the dotted lines (from left to right on the plot) represents the same LF for a range of $\beta = 1.5$, 1.0, 0.5, -0.5, -1.0 and -1.5, at a redshift of $z = 0.2$. In panel (c) a pure density evolution (PDE) model is alternatively introduced in to the LF. As with panel (b), the dotted lines (from top to bottom on the plot) represent LFs for the range of $\gamma = 6.0$, 4.0, 2.0 and -2.0, -4.0 and -6.0 at the same redshift.

These parametrisations have been long debated in the past and continue to be a critical point in the interpretation of evolution. The proponents of density evolution imagined that

the shape of the luminosity function changed only slightly over time, but that amplitude fell steadily from $z \simeq 2.5$ until the present; the advocates of luminosity evolution claimed that the apparent increase in numbers was due to a shape-invariant luminosity function sliding up from low luminosity to high as we look back toward $z = 2.5$. Now that the data have improved, it is apparent that neither camp was altogether correct. In some respects, luminosity evolution has turned out to be closer to the truth, but it is still not an adequate description of the actual evolution history. It is now clear that both the shape of the luminosity function and its normalization changed over the history of the Universe. Obviously, if the real luminosity function had turned out to be a power-law of fixed shape, both evolution modes could have been correct.

The discussion on which parametrization best fits the observed evolution does not simply concern which model better fits the data, but also what physical processes caused the evolution that we observe. Speaking about the AGN evolution, if pure density evolution had been correct, we can imagine that the reason for the evolution was that at $z = 2.5$ many galaxies were active, and as time went on they drops from substantial luminosity to none in a very short time. Contrarily, if pure luminosity evolution were correct, we can imagine an analogous picture, in which some fraction of galaxies became active at redshifts of a few, but their characteristic luminosity declined smoothly from $z = 2.5$ to the present. This is to stress that concluding in favour of density or luminosity evolution based on confrontation with statistical data is informative on the physical processes driving the evolution.

An alternative approach to the above parametric one is termed as a *free-form* technique, requiring no strict parametric assumption for the evolutionary model. A more recent example of this method is the one by Dye & Eales (2010) who extends a previous idea of Peacock & Gull (1981) by introducing a reconstruction technique that discretizes the functions of z and luminosity and employs an adaptive grid in $(L - z)$ space to constrain the evolution.

Concluding the discussion about the evolution we have to mention that another parameter frequently used is the *luminosity density* of galaxies, generally indicated as $\rho_{\rm L}$, that is the integral of the luminosity function down to a limiting value L_{min} of luminosity (Boselli 2011). The limit corresponds to the observational limit or the extrapolated value of the survey when the parametric or non-parametric LF are used respectively. This quantity is particularly important since it can be easily transformed into a *Comoving Star Formation Rate Density* (SFRD) whenever the observed bands are directly related to the emission of the young stellar components, or in *stellar mass densities* when the adopted filters are sensitive to the emission of low-mass stars dominating the stellar mass in galaxies. For wavebands dominated by AGN emission (like the hard X-rays), the luminosity density can be traced back to a rate of accretion onto massive nuclear black-holes. In the non-parametric case, $\rho_{\rm L}$ is given by

$$\rho_{\rm L} = \sum_{i=1}^{N_b} L_i \phi_i \mathrm{d}L_i \tag{2.84}$$

where N_b is the number of bins of luminosity of range dL_i, L_i is the mean luminosity of each bin and ϕ_i is the estimated value of the luminosity function within the bin i. In the case of a parametric representation of the luminosity function, the luminosity density is given by the relation

$$\rho_{\rm L} = \int_{L_{min}}^{\infty} L\phi(L)\mathrm{d}L \tag{2.85}$$

where $\phi(L)$ is described by a particular functional form such as those reported in Sec. 2.3.3. It is otherwise possible to fit the non-parametric LF results with a functional form and then perform the integral reported in Eq. 2.85.

Luminosity density can be converted into physical densities by using the relation explained in Sec. 5.4. Measured for different ranges of redshift, luminosity functions can thus be used to trace the variation of some important physical parameters per unit of comoving volume through cosmic time. Among these, the most widely determined is the variation of the star formation rate density $\rho_{\rm SFR}$ (SFRD) with look-back time, also referred to as the population's average star formation history (SFH). As we discuss in Chap. 5, the star formation rate of galaxies can be obtained by combining multi-λ data necessary, for instance, for appropriate corrections for dust extinction. At the same time, multi-λ data can provide independent estimates of the star formation rate density at a given redshift, and make the determination of $\rho_{\rm SFR}$ possible at some redshift where other indicators are not accessible for technical reasons (e.g. sensitivity of the detectors, etc). In Chap. 3 we report some of the efforts trying to reconstruct the SFRD of the Universe by using the IR emission of galaxies. As we will discuss, such early attempts were not conclusive due to their observational limitation. Our present analysis, developed in the following chapters (Chap. 6 and 7), will completely overcome such limitations and provide rather conclusive results on the subject.

Estimator	Type	Author	Advantages	Disadvantages
V/V_{max} & $1/V_{max}$	Non-parametric	Schmidt (1968)	Easy to apply No analytic form of LF assumed Shape and normalization of LF provided Test of evolution provided Test of completeness provided	Sensitive to density fluctuation
Parametric MLE	Parametric	Sandage et al. (1979)	No Binning of the data required Not sensitive to density fluctuation	Analytic form of LF required Cancels out density fluctuations No normalization of LF provided
Bayesian Parametric MLE	Parametric	Kelly et al. (2008)	No Binning of the data required Normalization of LF provided	Cancels out density fluctuations Priors for LF paramaters required
Non-Parametric MLE	Non-Parametric	Efstathiou et al. (1988)	No parametrization of the LF required Goodness of fit provided Independent of clustering effects	Binning of data required No normalization of LF provided
Semi-Parametric MLE	Semi-Parametric	Schafer (2007)	No parametrization of the LF required	Complete data-set required

Table 2.1: Summary of Advantages & Disadvantages of LF estimators. As we have mentioned in Chap. 2, no one of the LF estimators ever constructed is perfect or free of biases and each of them has advantages and disadvantages that guide the user to choose one or another according to the specifics of each study. In this table we try to summarize the major aspects of the methods described in this work.

Chapter 3

Extragalactic Surveys with Infrared Observatories

Observations at infrared wavelengths, loosely defined as the wide spectral interval from 1 to 1000 μm, are essential to study diffuse media in galaxies, including all kinds of atomic, ionic and molecular gases and dust grains. Hence they are particularly suited to investigate the early phases in galaxy evolution, when a very rich ISM (*Interstellar Medium*) is present in the forming systems.

In the local Universe, infrared emission is known to constitute about one third of the total light emitted from galaxies (Soifer & Neugebauer 1991). However, as can be inferred from the CIRB (*Cosmic Infrared Background*), this contribution increases significantly with redshift, accounting for half of the total energy produced by extragalactic sources, and is due to a corresponding increase in dust-obscured star formation and/or accretion activity (Dole et al. 2006a; Puget et al. 1996; Hauser 2001). For these reasons, if we are interested in investigating the early stages of the Universe and its later evolution we have to study this peculiar band, which is best done from outside the Earth's atmosphere.

With the launch of the Infrared Astronomical Satellite (IRAS) in January 1983, a number of studies of faint sources provided evidence for the existence of galaxy populations with enormous infrared luminosities, while being unnoticeable in the optical-UV. Later, the more powerful Infrared Space Observatory (ISO), and now the Spitzer Space Telescope and the Herschel Space Observatory have greatly improved our knowledge in this important range of wavelengths, and offered clear evidence for cosmological evolution for the faint IR sources.

In this Chapter we introduce the infrared spectral region and outline the properties of galaxies as viewed in this large wavelength band. We then present a short history of infrared astronomy and finally describe the Spitzer and Herschel missions, instruments and extragalactic surveys.

3.1 The Infrared Spectral Region

The infrared (IR) spectral region roughly spans the wavelength range between 1 μm and 300 μm. It can be roughly subdivided into the near-IR (NIR, 1-3 μm), the mid-IR (MIR, 3-30 μm) and the far-IR (FIR, 30-300 μm) regions, with longer wavelengths referred to as the sub-millimetre (sub-mm, 300 μm-1 mm) and millimetre (mm) regions. Fig. (3.1) shows the

transmission of Earth's atmosphere at NIR and MIR wavelengths. Water vapor and carbon dioxide are especially important absorbers throughout the IR, with water vapor making the atmosphere completely opaque over large wavelength ranges from almost all ground-based observing sites. Other absorbers, such as ozone, cause more limited (in wavelength) trouble, but can be important for specific scientific programmes. In addition, the atmosphere emits molecular lines and a very broad thermal continuum, which at each wavelength is proportional to the product of the emissivity and the Planck function evaluated at the temperature of the atmosphere.

The thermal emission from the telescope also contributes significantly to the thermal continuum emission at wavelengths longer than a few μm. For these reasons, when observing in these bands, it is necessary to observe from outside the terrestrial atmosphere and pay attention to the noise due to the structure of the telescope that can decrease, for example, using an efficient cooling system.

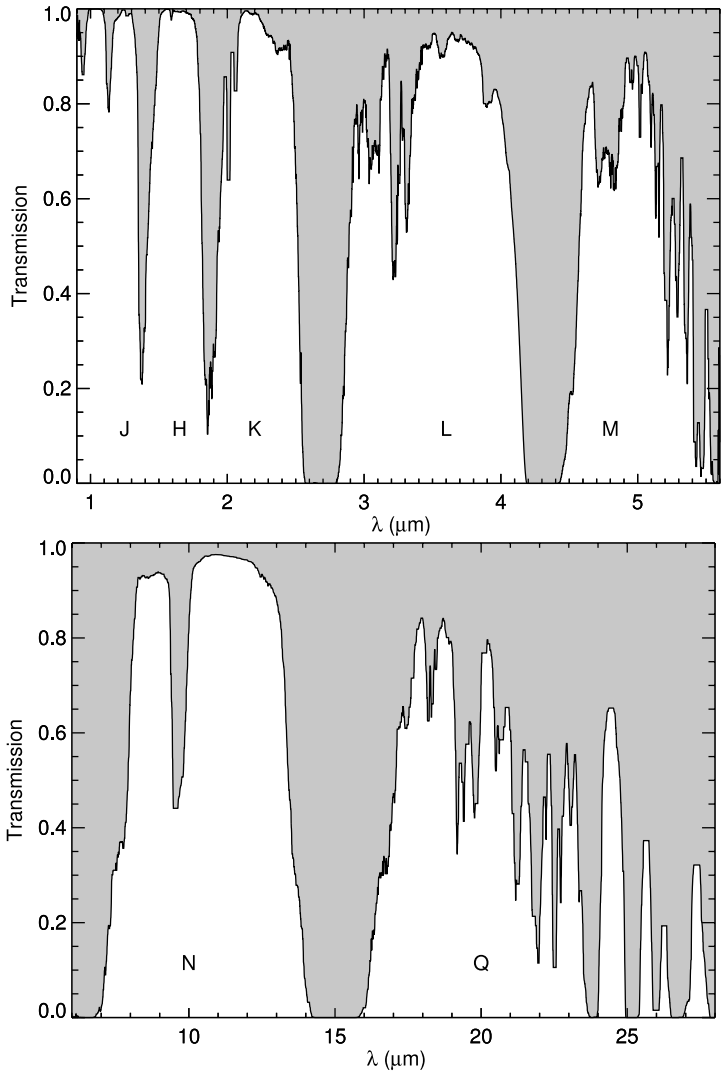

Figure 3.1: Atmospheric transmission in the near- and mid-infrared spectral region. Filter names in each spectral window are shown. Longwards of 30 μm the athmosphere is completely opaque up to 200 μm.

3.2 The Origin of Infrared Emission & Infrared Galaxies

Observations at long wavelengths, in the wide interval from 1 to 1000 μm, are essential to study diffuse media in galaxies, including all kinds of atomic, ionic and molecular gases and dust grains. Hence they are particularly suited to investigate the early phases in galaxy evolution, when a very rich ISM is present in the forming system.

The ISM amounts to $\sim 10\%$ of the baryonic mass of a local late-type system (Franceschini 2000), and can be considered as a mix of gas and dust. Small amorphous silicate and carbonaceous grains (i.e. dust particles) are very common in the interstellar medium of galaxies and amount to $\sim 1\%$ of the ISM mass.

The production of dust grains is a complex process. It is believed that condensation nuclei for dust grains mostly form in dense regions of the ISM, which are better shielded from UV photons. Thus they can be formed during the process leading to the birth of a star when a solar nebula is produced, silicate grains are formed and then blown away by the wind of the T-Tauri phase. Besides, graphite grains and silicate grains can be formed in the cold atmospheres of evolved stars, the former from carbon stars, the latter from OH-HI stars. Finally, another typical process leading to grain production could be a Supernova explosion. In this case, direct evidence of dust production came from the case-study of SN 1987a (CO and SiO molecules found in the ejecta), the dark spots observed in the synchrotron nebula of Crab, the IR mapping by ISO of Cas-A which resolved clumpy emission associated with the fast moving knots (Lagage et al. 1996).

Dust particles interact with photons emitted by astrophysical sources by absorbing, scattering, and polarizing the light (the combined effect of absorption and scattering takes the name of extinction). They also emit photons at wavelengths typically much greater than those of the absorbed photons. In particular, having low heat capacities, dust grains are transiently heated by absorbing UV or optical photons and hence can obscure regions of star-formation at these wavelengths. The dust then re-radiates this energy from MIR to sub-mm wavelengths, so optically obscured star-forming regions may be seen by observing at these latter wavelengths.

The interstellar IR emission spectrum is incredibly rich and shows a great wealth of fine details. In the mid-1970s, ground-based and airborne studies revealed relatively broad emission features in mid-infrared (mid-IR) spectra of bright HII regions, planetary nebulae (PNe), and reflection nebulae. The opening up of the IR sky by the Infrared Space Observatory (ISO) and the Spitzer Space Observatory has revealed the incredible richness of the mid-IR spectrum . It is dominated by major emission features at 3.3, 6.2, 7.7, 8.6, 11.2, 12.7, and 16.4 μm. In addition, there are weaker features at 3.4, 3.5, 5.25, 5.75, 6.0, 6.9, 7.5, 10.5, 11.0, 13.5, 14.2, 17.4, and 18.9 μm. These features are perched on broad emission plateaus from 3.2 to 3.6, 6 to 9, 11 to 14, and 15 to 19 μm. Moreover, many of the well-known features shift in peak position, vary in width, and/or show substructure, revealing a sensitivity to the local physical conditions. The IR emission features dominate the mid-IR spectra of almost all objects with associated dust and gas and illuminated by UV photons (Tielens 2008).

Large polycyclic aromatic hydrocarbon or PAHs molecules carry the infrared IR emission features that dominate the spectra of most galactic and extra-galactic sources. PAHs compose a family of hydrocarbon molecules with C atoms arranged in a honeycomb structure of fused six-membered, aromatic rings with peripheral H atoms. The C atoms are bonded to each other or to an H atom (for peripheral C atoms) through three sigma bonds, resulting in a planar structure. The fourth electron of each C atom is in a p orbital sticking out of the plane.

The p electrons on adjacent C atoms overlap to form π bonds and delocalized electron clouds above and below the plane. This conjugation is the origin of the high stability of these types of species. The well-known IR emission features (see Fig. 3.2) are characteristic of the stretching and bending vibrations of aromatic hydrocarbon materials (Allamandola et al. 1989). The 3-μm region is characteristic of CH stretching modes, and the 3.3-μm band results from the CH stretching mode in aromatic species. Pure CC stretching modes generally fall between 6.1 and 6.5 μm, vibrations involving combinations of CC stretching and CH in-plane bending modes lie slightly longward (6.5-8.5 μm), and CH in-plane wagging modes give rise to bands in the 8.3-8.9 μm range. The 11-15 μm range is characteristic of CH out-of- plane bending modes. Longward of 15 μm, emission bands reflect in-plane and out-of-plane ring bending motions of the carbon skeleton; hence these modes are more molecule specific, particular to the longer wavelengths.

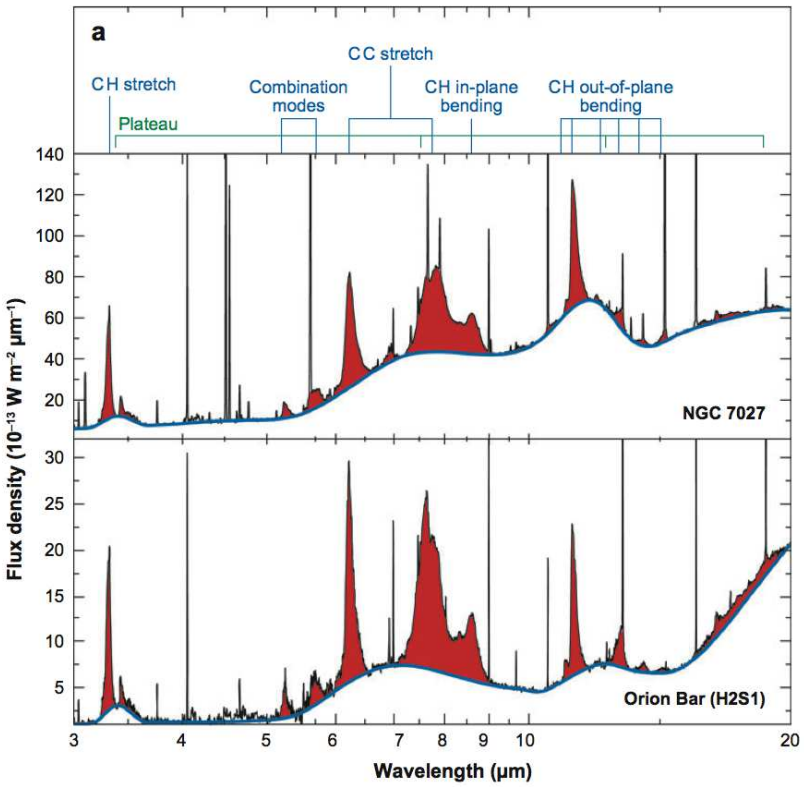

Figure 3.2: The mid-infrared spectra of the photodissociation region in the Orion Bar and in the planetary nebulae NGC 7027. These spectra are dominated by a rich set of emission features labeled with the vibrational modes of polycyclic aromatic hydrocarbon molecules (PAHs) at the top. Images from Tielens (2008).

Thanks to ever more sophisticated and high sensitive spectrometer we are now able to conduct detailed analysis of the molecular structure of the PAH. Of particular interest in the interpretation of the mid-IR emission of galaxies is the fractional abundance of various side groups of the emitting PAHs listed in Tab. 3.1.

Functional group	$\lambda(\mu\text{m})$	**Fraction** (relative to aromatic H)
Aromatic H	3.3	1
Aliphatic CH_3	3.40	0.02
Aliphatic deuterium CH_2D	4.65	0.02
Hydroxyl (OH)	2.77	< 0.002
Amine (NH_2)	2.88 - 2.95	< 0.01
Aldehydic (HCO)	5.9	0.006 (?)
Nitrile (C≡N)	4.48	< 0.01 (?)
N substituted in the ring structure (fraction relative to C)	6.2	0.04 (?)
Acetylenic (C≡CH)	3.03	< 0.003

Table 3.1: Table of the overwhelming aromatic nature of the unidentified infrared bands (Tielens 2008).

Interpreting the SEDs of infrared galaxies is then more complicated than in the other bands and this is clearly visible if we look at a typical spectral energy distribution of a luminous far-infrared galaxy (see Fig. 3.3). The observed SED of a galaxy depends on the properties of the stellar populations (e.g. ages and metallicities), the dust model (e.g. composition, size distributions), and the relative geometry of the two. The full range of possible combinations of all these is such that, fully accounting for all possible processes is today not possible (Sajina et al. 2006). As we are observing a system really unhinged by the presence of absorption the spectrum can be divided into different components: at the longest wavelengths, the sub-mm spectrum can be approximated by a black-body spectrum with optically thin emissivity varying as $\nu^{1.5}$; at shorter wavelengths there is a warm power-law component associated with dust grain of smaller size with decreasing wavelength; in the range between 5 and 20 μm the spectrum is dominated by the intense emission of polycyclic aromatic hydrocarbon molecules, or PAHs, which can account for up to 20% of the total infrared luminosity; finally at the shortest wavelengths, the stellar population of the galaxy is dominant, but it suffers the extinction as $e^{-\tau_\nu}$, where τ_ν describes the attenuation of the spectral energy distribution as a function of wavelength.

Since star formation activity happens in dusty astrophysical environments, the astrophysical processes at the origin of star formation need to be studied in the infrared. Thus, in order to understand how galaxies form and evolve and how the mechanisms that power them evolve with cosmic time, we must study their infrared emission in all possible detail.

Over the last 30 years or so, thanks to the deployment of mostly space-based observatories and detectors, we have been able to carry out extensive infrared extragalactic surveys. Here we give a panoramic view of these instruments and their surveys explaining how our knowledge has been improved since the advent of the first observations in this wavelength range, in particular with respect to number counts and luminosity functions and their evolution with redshift.

3.3 IRAS and the Discovery of Infrared Galaxies

The InfraRed Astronomical Satellite (IRAS Soifer et al. 1987) provided the first deep view of the entire IR sky. Its "all sky survey" was designed to search for objects with IR emissions at 12, 25, 60 and 100 μm. IRAS operated between January and November 1983 and

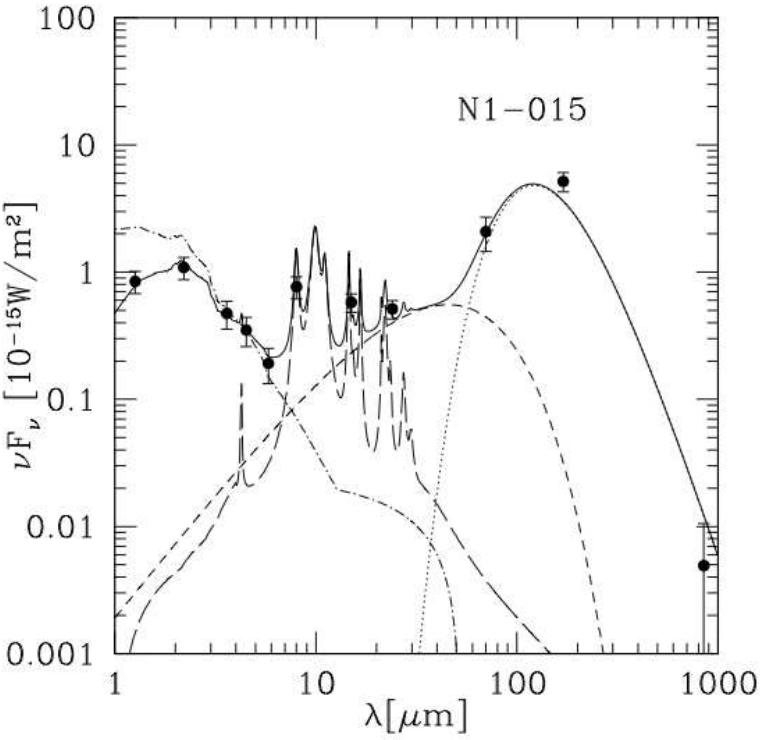

Figure 3.3: An example of an IR SED model for the luminous infrared galaxy N1-015. This includes a greybody (dotted line), a warm power-law (short-dash), PAH emission (long-dash), and unextincted stellar emission (dot-dash) with $e^{-\tau_\nu}$ extinction applied. The thick solid line is the total and the filled circles are the observed broad-band flux densities. Images from Sajina et al. (2006).

surveyed 96% of the sky, with completeness limits of 0.5 Jy at 12, 25 and 60 μm and 1.5 Jy at 100 μm. Its most sensitive band was at 60 μm managing and detected approximately 25,000 IR-emitting galaxies. IRAS detected IR emission from galaxies, primarily from spirals, but also from quasars (QSOs), Seyfert galaxies and early type galaxies. Moreover it discovered a new class of galaxies which radiate most of their energy in the infrared, many of them dusty starburst galaxies. The most luminous of these infrared galaxies (named "ultra-luminous infrared galaxies": ULIRGs) have QSO-like bolometric luminosities (LIRGs: $L \geq 10^{11} L_\odot$, ULIRGs:$L \geq 10^{12} L_\odot$, Sanders & Mirabel (1996). These galaxies appear often to be characterized by interactions or merging processes. Actually, IRAS gave us our first real census of this "new" type of galaxy emitting the bulk of their energy at infrared wavelengths. IRAS looked at galaxies with FIR luminosities of $10^6 L_\odot \leq L_{FIR} \leq 10^{13} L_\odot$, covering a redshift range from zero out to above $z = 0.4$.

The main catalogs produced on the basis of the IRAS survey were the Point Source Catalogue (PSC), Bidelman (1988) (later developed further through spectroscopic studies into the Point Source redshift Catalogue (PSCz), Saunders et al. (2000)) and the Faint Source Catalogue (FSC), Moshir et al. (1992), which managed to get down to a 60μm flux limit of $\sim 200 - 300$ mJy.

The IRAS survey, despite its moderate sensitivity, provided an important tool for the first statistical studies of deep and homogeneous FIR-selected galaxy sample, allowing a systematic

statistical investigation of star-formation and AGN activity in mostly local sources. Based on the Point Source Catalog and the more recent Faint Source Catalog extracted from the IRAS data, several authors have also looked at evidences for evolution in subsamples reaching various depths. At the time of IRAS the situation was somewhat unclear, although evolution seemed to be detected at 60 μm in most cases. Using PSC data, with solid identifications and radial velocities, Saunders et al. (1990) claimed a very strong evolution effect in their sample, while Fisher et al. (1992) found theirs compatible with no evolution (the latter has however a somewhat lower median redshift). Using the deeper FSC data, Lonsdale et al. (1990) found evidence for strong evolution from source number counts. Oliver et al. (1992) reported also evolution in their spectroscopic subsample, though at a slightly milder rate. Deeper studies (e.g. Hacking et al. (1987); Franceschini et al. (1988)) made on much smaller specific areas (a few tens of square degrees) suggest a large excess of faint detections. In order to clarify the problem and to complete these studies by providing the "missing link" between extensive, shallow catalogs ($f_\nu(60\mu\text{m}) > 0.2$Jy) and small, very deep studies ($f_\nu(60\mu\text{m}) <\sim$ 100mJy),Bertin et al. (1997) have undertaken an identification programme of IRAS sources at 60 μm on some selected areas in the sky, focusing themselves especially on the faintest flux domain achievable with the Faint Source Survey (FSS) "plates": $f_\nu(60\mu\text{m}) \approx 120$mJy with a signal-to-noise ratio of 4. In this work Bertin et al. (1997) confirmed the evolution found by previous studies, estimating strong evolution in FIR luminosity $\propto (1+z)^{3.2\pm0.2\pm0.3}$ or in comoving density $\propto (1+z)^{6.0\pm0.5\pm0.7}$. Besides, they found that sources in excess are genuine and are generally associated with faint, relatively red, optical counterparts which is interpretable as being M_* massive starbursting galaxies at redshifts $>\sim 0.1$.

A pioneering study of the photometric properties of distant IRAS galaxies byFranceschini et al. (1994) suggested that the starbursts events explaining the large bolometric luminosities are considerably hidden by dust at UV/visible wavelengths, which was indicated to explain the lack of optically luminous galaxies in optical surveys at $z \geq 0.5$.

We summarize here some of these historical results. In Fig. 3.4 is shown the luminosity function of a variety of classes of extragalactic sources detected by IRAS. In Fig. 3.5, Hacking et al. (1987) showed that the counts below 100 mJy from the IRAS survey are higher than expected if no evolution has taken place out to a redshift of approximately 0.2. Starburst galaxies cannot account for the reported 100 μm background without extreme evolution at high redshifts. The normalised differential counts showed that there were more faint IRAS galaxies than expected, in the same sense as the counts of radio sources, X-ray sources and quasar, although the range of redshift sampled was very much smaller than those of the extragalactic radio sources and quasar.

In Fig. 3.6 we report the more recent results on IRAS number counts by Oliver et al. (1992) and Bertin et al. (1997).

3.4 COBE and the Cosmic Infrared Background (CIRB)

The Cosmic Background Explorer (COBE, (Boggess 1992)), was a NASA space satellite launched November 18, 1989. It carried three instruments on board, the Diffuse Infrared Background Experiment (DIRBE) to search for the cosmic infrared background radiation, a Differential Microwave Radiometer (DMR) to map the cosmic radiation sensitively, and a Far Infrared Absolute Spectrophotometer (FIRAS) to compare the spectrum of the cosmic

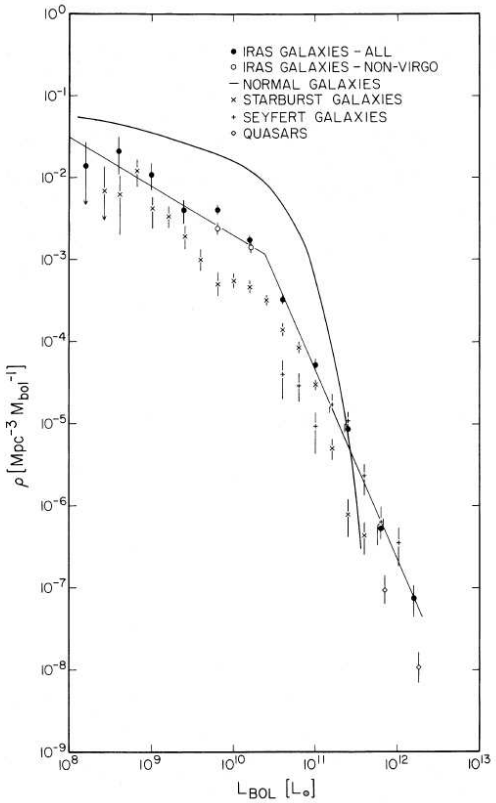

Figure 3.4: The 60 μm IRAS luminosity function, normalized to $H_0 = 75$ km s^{-1} Mpc^{-1} and plotted in units of bolometric luminosity. Filled and open circles represent far-infrared luminosity function derived for the IRAS bright galaxy sample, including and excluding the Virgo cluster respectively. Solid curve represents analytical fit to normal galaxy luminosity function taken from Schechter (1976). Crosses represent optically selected starburst galaxies, and plus signs represent the optically selected Seyfert galaxies, both taken from Huchra (1977). Open diamonds represent optically selected quasars taken from Schmidt & Green (1983). Straight lines represent best fit of two power laws to bright galaxy luminosity function excluding Virgo galaxies. Images from Soifer et al. (1987).

microwave background radiation with a precise calibration blackbody. The main duties of this instrument were determine precise measurements on the Cosmic Microwave Background (CMB) and to search for a diffuse Cosmic Infrared Background (CIRB), first predicted by Partridge & Peebles (1967) and thus to contribute to the studies of the evolution of infrared galaxies.

The extra-galactic background light, I, can be defined using the flux, $S(L, z)$, emitted from all the galaxies, dN, (i.e. the number counts) over the whole sky by:

$$I = \int S(L, z) \frac{\mathrm{d}N}{\mathrm{d}\Omega \mathrm{d}S} \mathrm{d}S \tag{3.1}$$

at a specific observation band of frequency, ν. In terms of the luminosity function, this may be written as:

$$I_\nu = \int S_\nu(L, z) \phi(L, z) \mathrm{d}logL \frac{\mathrm{d}V}{\mathrm{d}\Omega} \tag{3.2}$$

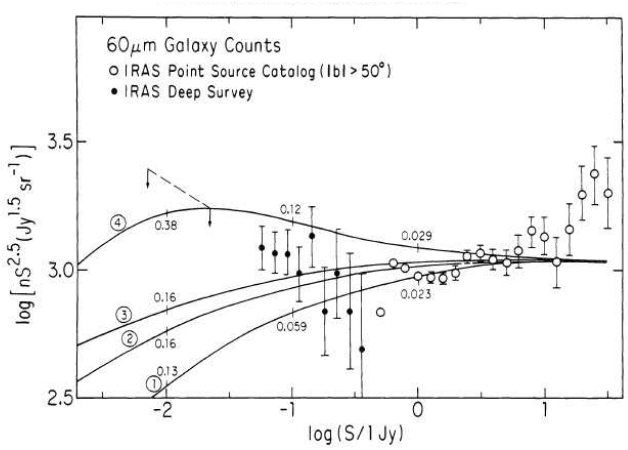

Figure 3.5: Predicted and observed 60 μm normalized differential source counts $nS^{2.5}(\mathrm{Jy}^{1.5}\mathrm{sr}^{-1})$. The solid curves show the model source counts as function of 60 μm flux density S. Tick marks indicate median model redshifts at different flux densities. The circled number refer to the four models described in Tab. 1 of Hacking et al. (1987). The open symbols mark source counts from the IRAS Point Source Catalog (1985) with galactic latitudes $|b| \geq 50$ degrees. Filled circle represent counts from deep survey. The image is from Hacking et al. (1987).

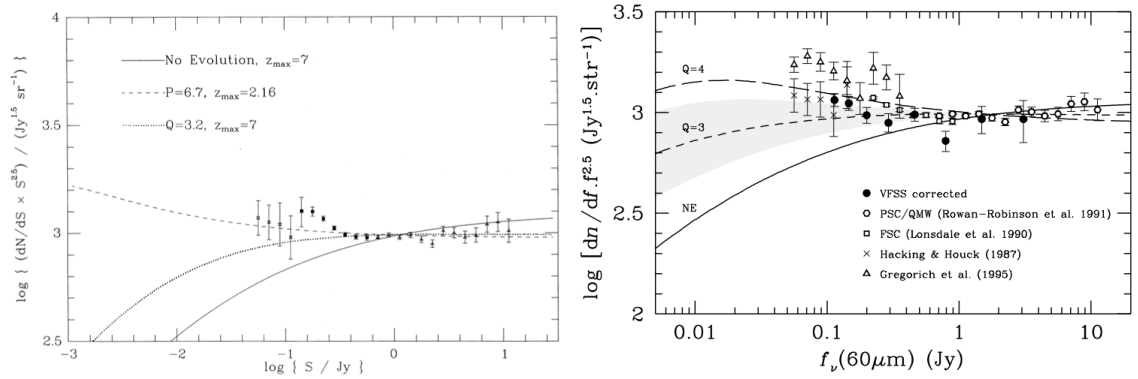

Figure 3.6: Predicted and observed 60 μm normalized differential source counts $nS^{2.5}(\mathrm{Jy}^{1.5}\mathrm{sr}^{-1})$. **Left**: Oliver et al. (1992). Triangles are data from Rowan-Robinson et al. (1990); solid squares are Saunders et al. (1990) and the crosses are Hacking et al. (1987). The theoretical curves are for no evolution, pure luminosity and density evolution, respectively defined as: the former, $\frac{L(t)}{L(t_0)} = f(z)$ with an exponential dependence on time which for an $\Omega = 1$ universe takes the form $f(z) = exp\frac{2}{3}Q[1-(1+z)^{-3/2}]$ with $Q = 3.2 \pm 1.0$; the last as $\frac{\Phi(t)}{\Phi(t_0)} = g(z)$ with a power-law dependence on $(1+z)$, $g(z) = (1+z)^P$ with $P = 6.7 \pm 2.3$. The redshift cut-offs used here have a negligible effect on the counts at these flux levels. **Right**: Bertin et al. (1997). Corrected VFSS number counts compared with previous studies and simple no-evolution (NE) and $L_* \propto (1+z)^Q$ models at 60μm.The correction was done by subtracting to each bin ($f_\nu(60\mu\mathrm{m}) < 0.3$ Jy) the bias measured in Monte-Carlo simulations, with the Q = 3.5 model. This bias depends weakly on the exact value of Q; for $3 \leq Q \geq 4$, the typical uncertainty is $\approx 5\%$, and has been included in the error bars. Note that the other counts have not been corrected for the Eddington bias, and might be overestimated in their faintest bins. The grey area indicates approximately the 1? confidence domain for number counts in the context of a$L_* \propto (1+z)^Q$ pure luminosity evolution model, drawn from the background fluctuation analysis.

Observations of the integrated cosmic background provide us with integral constraints on the number counts. Integrating the counts over flux shows us how much of the background

light, in a particular waveband, is resolved into individual sources. A comparison of the integral of number count distributions at different wavelengths, as in eq. 3.1, with the observationally determined extragalactic background help us to differentiate between separate models of evolution of cosmic sources in that band.

The determination of the cosmic background is extremely important because it is a measure of the total output of all galaxies over the history of the Universe. COBE showed that the cosmic background light from galaxies is dominated by two peaks: one peak in the optical/near-IR dominated by the emission from hot young stars and aging stars; the second peak coming from emission in the infrared and sub-mm regime corresponding to the so called CIRB. The integrated extragalactic background light in the far-infrared and sub-millimeter region of the spectrum is approximately equal to, or slightly larger than, the integrated background light in the optical and UV part of the spectrum (see Fig. 3.7 and Fig. 3.8).

The existence of such a background radiation was predicted some years before discovery by Franceschini et al. (1994) based on an analysis of deep IRAS and sub-mJy radio galaxy counts and on the assumption that the main phases of star formation in galaxies are heavily extinguished by dust absorption. The detection of the CIRB at the level found by COBE implied that the most extinguished models among those considered by Franceschini et al. could well explain this intensity.

The original detection of the CIRB was performed by Puget et al. (1996) and was later confirmed with independent analyses by various other groups using FIRAS (Fixsen et al. 1998), as well as data from the DIRBE experiment in three broad-band channels at λ =240, 140 and 100 μm (Hauser et al. 1998). These studies have confirmed that it is possible to understand the CIRB as the product of the light radiated by dust over a Hubble time primarily powered by star formation (the accretion around black-hole contribution is $< 20\%$). To appreciate the relevance of this discovery consider that extragalactic backgrounds at other wavelengths contain only modest contributions by distant galaxies: the radio background is clearly dominated by radio-loud AGN; the microwave background includes photons generated at $z \sim 1500$; the X-ray and γ-ray backgrounds are dominated by distant quasars and AGN. On the other hand, the DIRBE experiment on board COBE has brought to the first detection of the integrated emission by distant galaxies in the form of an isotropic signal at far-IR and sub-mm wavelengths. Then the CIRB's intensity exceeding the optical background suggests that galaxies in the past should have been much more active in the far-IR than in the optical, and very luminous in an absolute sense. For this reason if we want to develop a complete understanding of galaxy formation, this background light must be resolved into sources, whose properties need to be characterized.

COBE proved that the CIRB is very intense and peaks at 100-200 μm, and therefore (since galaxy dust emission is the only reasonable mechanism to produce it) that the strong evolution seen in the local universe by IRAS must continue up to high redshift and that the integrated infrared emission of galaxies is a fundamental portion of the overall cosmic energy budget.

3.5 ISO & SCUBA and the High-Redshift Infrared Universe

After the IRAS' success, plans were made by ESA to develop a new space mission, the Infrared Space Observatory (ISO, Kessler et al. 1996), which was the first cryogenic space mission operated as an observatory. With ISO the aim was to observe the singles sources that contribute to the total CIRB detected by COBE. While IRAS only sampled the 10 to

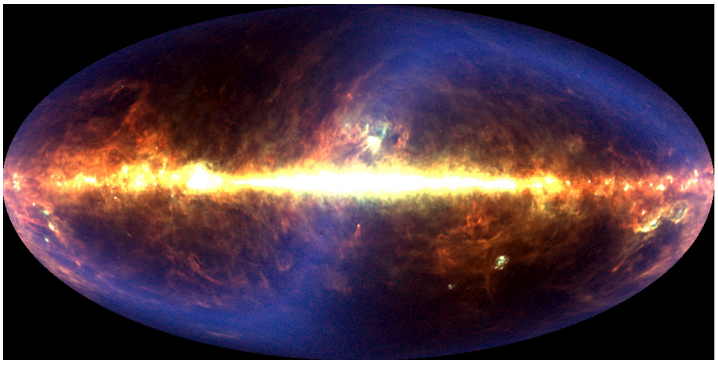

Figure 3.7: Projection of the entire infrared sky created from years of observations by the robot spacecraft COBE. Our Solar System is evidenced most prominently by the S-shaped blue sash called zodiacal light, created by small pieces of rock and dust orbiting between the Sun and Jupiter. Our Galaxy is evidenced by the bright band of light that crossed the middle of the image, created mostly by dust that laces the disk of our Milky Way. Close inspection of the image reveals that the background is not completely dark, indicating that our Universe itself provides a diffuse glow, created by the dust left over from the star formation throughout the Universe.[2]

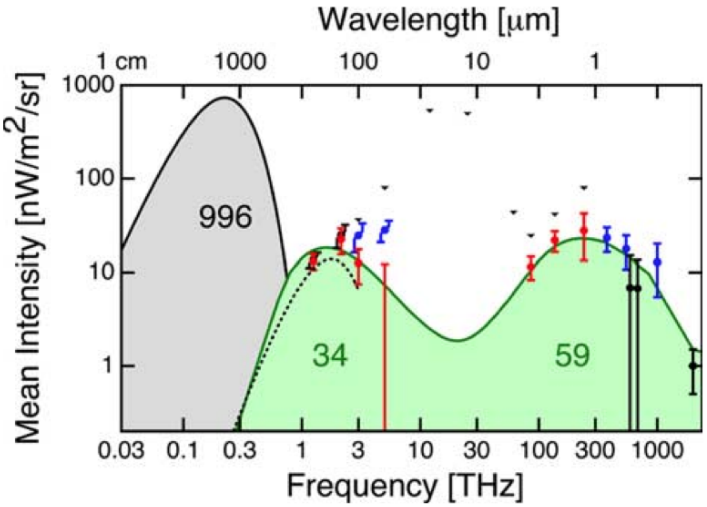

Figure 3.8: **Left**: The cosmic infrared and optical background from foreground subtracted total intensity measurements only (Wright 2004). The numbers in the bumps indicate the integrated intensity in nW/ m^2/sr within each bump. (For interpretation of the references to colour in this figure legend, the reader is referred to the original paper by Wright (2004).

100 μm wavelength domain, producing a survey sample of 25,000, mostly local, galaxies, ISO broadened out the observational wavelength range to cover the whole 3.5 μm to 180μm interval, with its main bands at 6.7, 12, 15, 60, 90, 135, and 175 μm. ISO was launched in November 1995, had a 2.5-year lifetime, improved on the IRAS angular resolution (down to 1.5-90″, a hundred times finer) and was about 1000 times more sensitive in the mid-IR. The great success of ISO has been to overcome the problems implied by the limiting effects of the Earth's atmosphere and of thermal emission from the telescopes and instruments when observing in the IR bands, by putting a cooled telescope into space (the whole observatory was put inside an enormous Dewar). ISO provided us with the ability to truly explore the

Spectral Energy Distributions (SEDs) of the dusty, infrared emitters; it provided us with a first real chance to ascertain the nature of their central engines, the physical conditions within dusty galaxies, and their evolution with cosmic time up to $z \leq 1$.

The satellite was composed of an imaging photopolarimeter (ISOPHOT), a NIR/MIR camera (ISOCAM), a short wavelength spectrometer (SWS) that covers 2.4-45 μm and a long-wavelength spectrometer (LWS) covering 43-197μm (Genzel & Cesarsky 2000).

The largest single Open Time project conducted by ISO has been the European Large Area ISO Survey (ELAIS) mapping an area of 12 deg^2 at 15 μm with ISOCAM and at 90 μm with ISOPHOT (Oliver et al. 2000). Other important ISO surveys were the deep ones performed in the HDFN and HDFS areas (Aussel et al. 1999; Oliver et al. 2002).

With ISO's improved resolution and sensitivity over IRAS, the limiting flux densities were pushed down to about 0.1 mJy at 6.7 μm and 0.5 mJy at 15 μm (Oliver et al. 2002). Probing early star-formation hidden by dust, the ISO satellite permitted us to study starburst (SB) galaxies at higher redshifts than previously possible with IRAS. By producing more detailed investigations into star formation regions, it allowed us to start to distinguish between starbursts (SBs) and active galactic nuclei (AGNs) as the central engine driving ultra luminous IR galaxy (ULIRG) emission (Genzel et al. 1998).

ISO confirmed earlier IRAS indications of an excess of faint IR sources and showed that the mid-infrared source counts require strong cosmological evolution. In Fig. (3.9) we report the 15μm differential source counts. From >10 mJy down to 0.4 mJy the counts increase rapidly with a slope in the integrated counts of $\beta = -3$. This is significantly steeper than expected in an Euclidean model without luminosity/density evolution ($\beta = -2.5$). Below 0.4 mJy down to the faintest flux densities sampled the slope is flatter ($\beta = -1.6$) and the counts appear to converge. At the $\sim$ 0.4 mJy peak, the observed source counts are an order of magnitude above non-evolution models, obtained by extrapolating the local IRAS 12 μm luminosity function (Rush & Malkan 1993).

The 15 μm integral number counts agree with model predictions over more than five orders of magnitude in flux density (Pearson & Rowan-Robinson (1996), Guiderdoni et al. (1997), Franceschini et al. (2001), Roche & Eales (1999)). Common to all these models is the assumption of strong luminosity and/or density evolution of dusty star formation in bright normal spirals and starburst galaxies, with varying contributions from AGNs.

Using the data acquired in the longer wavelength bands, at FIR wavelengths (100-240 μm from IRAS and ISO), presented a clearer picture of the sources being detected there. The interest in this new exiting regime of λ were accreting more and more with time, but for years such kind of studies were completely precluded to ground-based astronomy that suffered for the lack of the necessary technology. Even from superb sites the atmospheric transparency is generally poor, and even in the transmission 'windows' the high background photon power and associated noise fluctuations from the atmosphere limit the observing sensitivity. However, in terms of the three fundamental factors that govern the impact of a particular waveband - the collecting area of the telescope, the sensitivity of the detector, and the availability of imaging systems - submillimetre astronomy has finally come of age. The 15-m James Clerk Maxwell Telescope (JCMT) is situated at a high, dry site - Mauna Kea in Hawaii. With detectors cooled to well below 1 K, sky-background noise sensitivity levels are achievable in all the sub-millimetre atmospheric windows. Array detectors were the obvious next step, and the extremely ambitious SCUBA project was approved by the JCMT Board in 1987.

SCUBA, the Sub-millimetre Common-User Bolometer Array (Holland et al. 1999), was

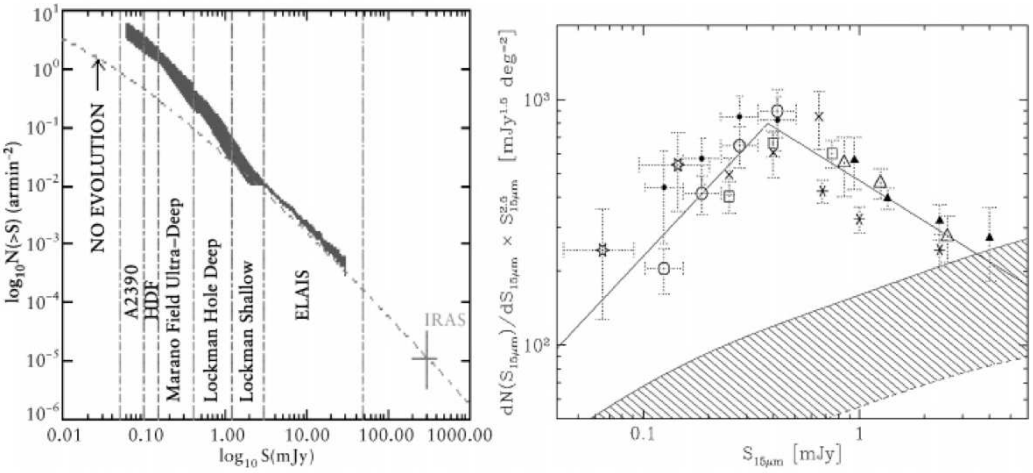

Figure 3.9: **Left**: Summary of integrated 15 μm source counts from the different ISO surveys (Elbaz et al. 1999), normalizing downwards by a factor 1.5 the counts of Serjeant et al. (2000)), compared to no evolution models (continuous) matching the IRAS counts (the Rush & Malkan (1993) counts were renormalized downwards by a factor of 2). **Right**: Differential 15 μm counts with the shaded area marking the counts predicted with non-evolution models (from Elbaz (2000)). The counts are normalized to a Euclidean distribution of non-evolving sources which would have a slope of index -2.5 in such a Universe. Data points: A2390 (open stars, Altieri et al. (1999)), HDFN (open circles, Aussel et al. (1999)), HDFS (filled circles, Oliver et al. (2002)), Marano (open squares, crosses, stars, Elbaz et al. (1999)), Lockman hole (open and filled triangles, Elbaz (2000)). Image from Genzel & Cesarsky (2000).

designed and constructed at the Royal Observatory, Edinburgh in collaboration with Queen Mary, University of London. It was delivered to the JCMT in 1996 and was fully operational by 1997. SCUBA had two arrays of bolometric detectors. The long-wave array had 37 detectors operating in the 750 and 850 μm atmospheric transmission windows, while the short-wave array 91 detectors for observations at 350 and 450 μm. Both arrays had approximately the same field-of-view with a diameter of 2.3', and could be used simultaneously by means of a beam-splitter. There was also three pixels available for photometry in the transmission windows at 1.1, 1.35 and 2.0 mm, and these was located around the edge of the long-wave array. The detectors was cooled to approximately 100 mK to limit background noise from heat emitted by the instrument.

The main results from SCUBA surveys were to show that the present day theoretical model on galaxy formation based upon optically selected samples significantly underestimated the observed surface density of sub-millimeter sources: this demonstrates that a significant fraction ($> 80\%$) of the star-formation activity in the high-redshift universe may have been missed in previous optical studies (Smail et al. (1997)). SCUBA has identified a population of high-redshift dusty starburst galaxies which contribute a significant fraction of the extragalactic background at 850 μm Hughes et al. (1998) and showed a density of dust-enshrouded star-formation at $z > 2$ which is at least a factor of ~ 5 greater than that deduced from Lyman limit systems. SCUBA found that the majority of the sub-mm background is composed of sources with 850 μm fluxes between $\sim 0.8 - 10$ mJy.

ISO & SCUBA investigated the sources of the IR and sub-mm background at the two wavelength extremes of the CIRB, the mid-IR and the millimeter. Having resolved the CIRB background into discrete sources, people started to discern the nature of the source popula-

tions. ISO 15 μm sources account for 80% of the CIRB and have a median redshift of $z \sim 0.7$ (Elbaz et al. 2002). Chary & Elbaz (2001) divided the ISO galaxies into ULIRGs, LIRGs, starburst, spirals and AGN populations, deducing that LIRGs would only be visible out to $z \sim 1.2$. By extrapolating their results to 140 μm they found that ULIRGs and LIRGs were the major contributing populations to the CIRB (at 67%), but that the ISO galaxies make little contribution to the millimeter (mm) background, thus implying that the mm and sub-mm background must be composed of sources that are further away. Redshifts for SCUBA 850 μm galaxies above 3 mJy (contributing 30% of the CIRB at that wavelength) have a median around $z \sim 2.2$ (Chapman et al. 2005). To go significntly fainter, we needed to wait for Spitzer results, described in details in the next section, which provided counts of more than 3 orders of magnitude deeper than ISO & SCUBA, and surveying much larger samples than previously available.

3.6 The Spitzer Space Telescope

The launch of the Spitzer Space Telescope (Werner et al. 2004) opened a new era for infrared astronomy. Its vast leap in sensitivity over previous observatories, both in space and on ground, has enabled virtually every aspect of observational astronomy to gain very significant improvement. The satellite was launched in August 2003 and was inserted into an Earth-trailing heliocentric orbit. Spitzer is constrained to point no closer than 80 deg toward and no further than 120 deg from the Sun, but even with these constraints 35% of the sky is visible at any time, and the entire sky is visible every six months.

Spitzer is the final element in NASA's Great Observatories Programme, and an important scientific and technical cornerstone of the astronomical search for the Origins Programme. It incorporates an 85 cm diameter telescope primary mirror (see Tab. 3.2 for more details), cooled to as low as 5.5 K, and three scientific instruments providing imaging and spectroscopy at wavelengths from 3.6 to 160 μm. The overall configuration of the flight system is shown in Fig. (3.10). The spacecraft and solar panel were provided by Lockheed Martin. The telescope, cryostat, and associated shields and shells make up the Cryogenic Telescope Assembly (CTA), built by Ball Aerospace, who also built the Multiband Imaging Photometer for Spitzer (MIPS) and Infrared Spectrograph (IRS) instruments. The third instrument, the Infrared Array Camera (IRAC), was built at NASA's Goddard Space Flight Center. The Spitzer flight hardware consists of a spacecraft that operates at roughly room temperature and the CTA that is cooled by a combination of superfluid liquid helium, helium boil-off gas, and radiative cooling and operates at much lower temperatures.

3.6.1 Spitzer Imaging Instruments

Spitzer has two NIR and MIR imaging instruments, IRAC and MIPS, and an IR spectrograph, IRS (Infrared Spectrograph) providing a multi-wavelength coverage of the IR Universe.

- IRAC: InfraRed Array Camera (Fazio et al. 2004). It is a four-channel camera that provides simultaneous $5.2' \times 5.2'$ images at 3.6, 4.5, 5.8 & 8 μm. Thanks to the low background available in space, it can reach faint limiting fluxes in modest exposure times. In fact, IRAC can reach similar depths (19th magnitude) as the 10 m W. M. Keck Telescope with the NIRC2 camera at 3.6 μm in only one-hundredth of the exposure time

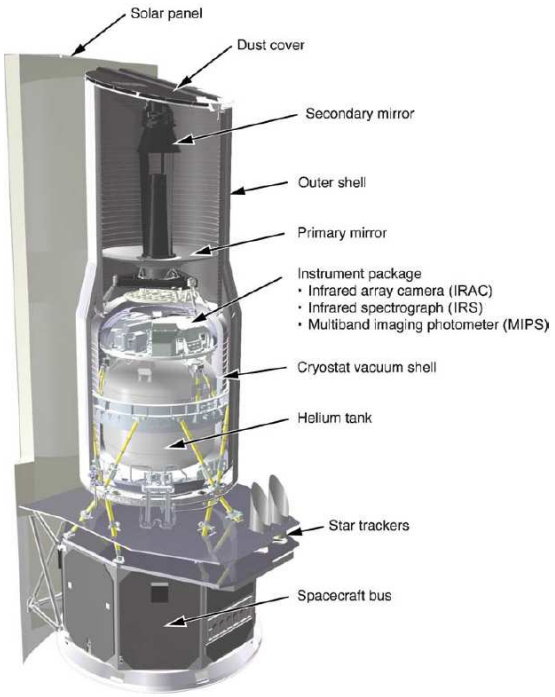

Figure 3.10: The Spitzer Space Telescope. The outer shell forms the boundary of the Cryogenic Telescope Assembly (CTA), which incorporates the telescope, the cryostat, the helium tank, and the three instruments, the Multiband Imaging Photometer, the Infrared Spectrograph, and the Infrared Array Camera. Principal Investigator-led teams provided the instruments, while Ball Aerospace provided the remainder of the CTA. Lockheed Martin provided the solar panel and spacecraft bus. The observatory is approximately 4.5 m tall and 2.1 m in diameter; the mass at launch was 861 kg. The dust cover atop the CTA was jettisoned 5 days after launch. Image from Soifer et al. (2008).

Telescope	
Aperture (diameter)	85 cm
Orbit	Solar (Earth-trailing)
Cryogenic Lifetime	5.5 years (est.); 5.7 years actual
Wavelength Coverage (passband centers)	3.6 - 160 μm (imaging)
Diffraction Limit	5.5 μm
Image Size	1.5" at 6.5 μm
Field of View (of imaging arrays)	$\sim 5' \times 5'$ (each band); at 160 μm a single field of view is $0.053' \times 5.33'$
Telescope Minimum Temperature	5.6 K (cryo); 27.5 K (post-cryo)
Maximum Tracking Rate	1.0"/ sec

Table 3.2: Spitzer telescope main characteristics, from the Spitzer telescope handbook [3].

and does so with a far larger field of view: $5' \times 5'$ versus $40'' \times 40''$. The IRAC bands are only weakly affected by dust extinction and are well suited to probing stellar mass at fairly high redshifts and therefore to detect elliptical/spheroidal galaxies, Surace et al.

(2005).

- MIPS: Multiband Imaging Photometer for Spitzer (Rieke et al. 2004). It produces imaging and photometry in three broad spectral bands, centered nominally at 24, 70 and 160 μm, and low-resolution spectroscopy between 55 and 95 μm. The instrument contains three separate detector arrays each of which resolves the telescope Airy disk with pixels of size $\lambda/2D$ or smaller. While the IRAC bands probe the stellar mass, the MIPS bands are sensitive to on-going star formation and cooler highly dusty regions. The high sensitivity at 24 μm of the MIPS instrument allows infrared-luminous galaxies to be traced to high redshifts. Sensitivity to obscured starbursts at $z \sim 2$ is enhanced due to redshifting of the 7.7 μm feature into the 24 μm band.

- IRS: Infrared Spectrograph (Houck et al. 2004). Consists of four separate modules (Short- Low, Short-High, Long-Low, Long-High) which provided low (R $\sim$ 60 - 130) and moderate (R $\sim$ 600) resolution spectroscopic capabilities from 5.2 to 38 μm. Spectra could be obtained in either staring or mapping mode. In addition, the IRS provided imaging in two filters (13-18 and 18-26 μm) and onboard software to autonomously identify point sources and accurately place them (by offsetting the telescope) in any of the IRS slits.

See Tab. 3.3 for a summary of the instrumentation.

λ [μm]	λ/Dλ	FoV	Pixel	Sensitivity [μJy]
IRAC: InfraRed Array Camera				
3.6	4.7	$5.21' \times 5.21'$	1.2	1.6 (3.4)
4.5	4.4	$5.18' \times 5.18'$	1.2	3.1 (4.3)
5.8	4.0	$5.21' \times 5.21'$	1.2	20.8 (21)
8.0	2.8	$5.21' \times 5.21'$	1.2	26.9 (27)
MIPS: Multiband Imaging Photometer for Spitzer				
24	5	$5.4' \times 5.4'$	2.55	110
70	4	$2.7' \times 1.4'$	5.20	14.4 mJy
55-95	15-25	$0.32' \times 3.8'$	10.1	57, 100, 307 mJy (60, 70, 90 μm)
160	5	$0.5' \times 5.33'$	16×18	29 (40) mJy
IRS: Infrared Spectrograph				
5.2-14.7	64-128	$3.7'' \times 57''$	1.8	250
13-18.5	$\sim$3	$54'' \times 80''$	1.8	116 (80)
9.9-19.5	$\sim$600	$4.7'' \times 11.3''$	2.3	$1.2 \times 10 - 18$ W/m^2
14.3-35.1	64-128	$10.6'' \times 168''$	5.1	1500
18.9-37.0	$\sim$600	$11.1'' \times 22.3''$	4.5	$2 \times 10 - 18$ W/m^2

Table 3.3: Spitzer science instrument main characteristics, from the Spitzer telescope handbook [5].NB.The sensitivities given are for point sources, and are only representative; IRAC sensititivity is given for intermediate background - the first number in each case with without confusion, and the second number (in parentheses) includes confusion; IRS sensitivity is given for low background at high ecliptic latitude (note that for IRS, sensitivity is a strong function of wavelength); MIPS sensitivity is given for low background; 70 observations can be confusion limited; Because of a bad readout at one end of the slit, spectral coverage for 4 columns in MIPS SED is reduced to about 65-95 μm; 160 μm is often confusion limited, 29 mJy refers to no confusion and 40mJy refers to the estimated confusion limit.

Some beautiful examples of the Spitzer photometric imaging capabilities are reported in Fig. 3.12 and Fig. 3.13.

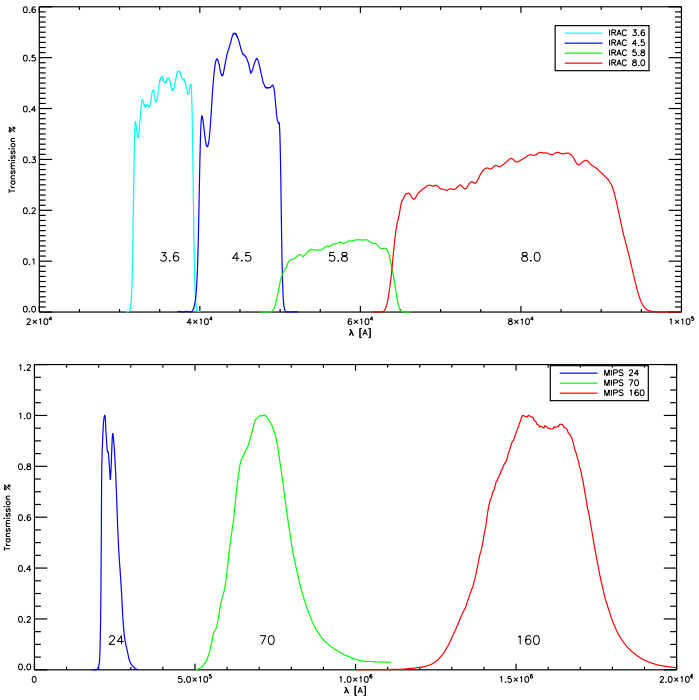

Figure 3.11: IRAC & MIPS filter transmission. **Top**: IRAC filters transmission; **Bottom**: MIPS filters transmission.

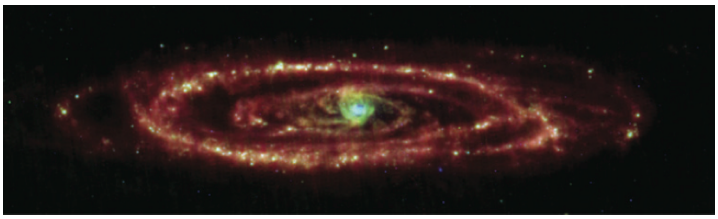

Figure 3.12: The Andromeda Nebula, Messier 31 as imaged by Spitzer at 24 (blue), 70 (green), and 160 (red)μm.The excess of blue/green colours in the nuclear region signals warm dust in circumstellar shells in the bulge. Images from Soifer et al. (2008).

3.6.2 The Spitzer Legacy Science Programme

Astronomers have fully exploited Spitzer's orders-of-magnitude gain in imaging and spectroscopic to address a wide variety of topics in extragalactic research (Soifer et al. 2008). Spitzer provided the scientific community with the most powerful tool yet available for astronomical explorations between 3.6 and 160μm, thanks to its three cryogenically-cooled science instruments. Compared to the 15 μm band (probed by ISOCAM and to a lesser extent by Spitzer's IRS peak-up imaging camera), whose sensitivity is enhanced for sources in the redshift range 0.5-1.5 due to the prominent polycyclic aromatic hydro-carbon (PAH) features, the

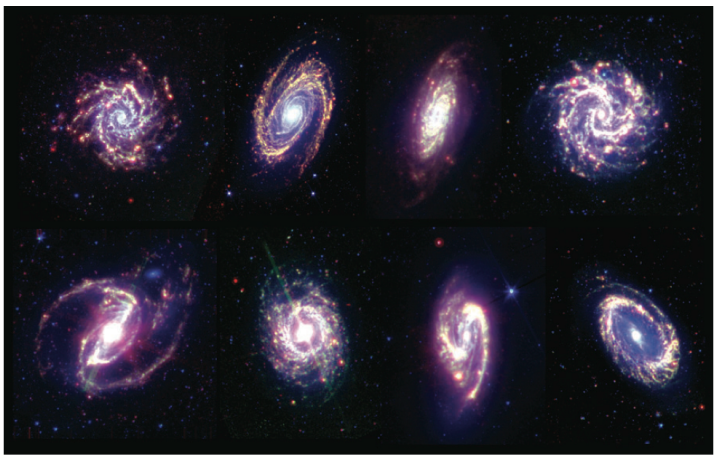

Figure 3.13: Selected SINGS spiral galaxies as imaged by Spitzer at 3.6 (blue), 8 (green), and 24 (red) μm illustrate the wide variation of star-formation activity configurations as compared to the stellar distribution in spiral galaxies. The galaxies are, clockwise from upper left-hand corner, NGC628 SAc, NGC3031 = M81 SAab, NGC3521 SABbc, NGC3184 SABcd, NGC4725 SABab, NGC3627 SABb, NGC3351 SBb, and NGC1097 SBb. Images from Soifer et al. (2008).

Spitzer 24μm band makes it possible to explore the Universe well up to $z = 2.5$. Coupled with better sensitivity and increased efficiency in large areal coverage, the 24μm observations with Spitzer dramatically improved the quality and statistics of number counts in the mid-infrared (Marleau et al. 2004). As the infrared member of NASA's family of Great Observatories, Spitzer has been used very successfully in multi-spectral studies with its companion observatories, the Chandra X-Ray Observatory and the Hubble Space Telescope (HST). Spitzer has also observed objects currently accessible only in the infrared, most notably detecting radiation from extrasolar planets for the first time (Werner et al. 2006). Much of the extragalactic work results from Spitzer surveys of areas of the sky ranging in extent from tens of square arcminutes to tens of square degrees. Most of these surveys cover areas that are well studied with many other observatories, thereby producing a unique multi-wavelength view of the universe. The most popular within these "Cosmic Windows" were targeted by the competitive Spitzer Legacy Science Programme coordinated by the Spitzer Science Telescope and were carried out during the Spitzer Cold Mission (2003-2009). These projects are:[6]

- SWIRE: The Spitzer Wide-Area Infrared Extragalactic Survey (Cycle 0 Legacy), observed ~ 50 deg^2 broken down into six separate regions with IRAC and MIPS.
- FLS: The First Look Survey (extragalactic component) (Cycle 0 DDT) observed 4 deg^2 with IRAC and MIPS.
- SDWFS: The Spitzer Deep Wide-Field Survey (Cycle 4) observed about 8.5 deg^2 in the NOAO Deep Wide-Field Survey Bootes area with IRAC, following major GTO programmes using IRAC and MIPS.
- S-COSMOS: The Spitzer Cosmic Evolution Survey (Cycle 2 & 3) observed the 2 deg^2 Cosmic Evolution Survey field with IRAC and MIPS.

[6] `http://irsa.ipac.caltech.edu/data/SPITZER/docs/spitzermission/observingprograms/legacy/`

- SPUDS: (Cycle 4), The Spitzer Ultra Deep Survey extended the 1 deg^2 field UKIDSS Ultra Deep Survey with an IRAC and MIPS deep imaging survey.
- SIMPLE: The Spitzer IRAC/MUSYC Public Legacy in E-CDFS (Cycle 2), images 0.25 deg^2 around the Chandra Deep Field South with IRAC.
- GOODS: The Great Observatories Origins Deep Survey (Cycle 0 Legacy) observed 320 $'^2$ centered on the Hubble Deep Field North and the Chandra Deep Field South fields with IRAC and MIPS (at 24 μm). Spitzer's deepest observations have been made over portions of the GOODS field. A number of GO programmes and The Far-Infrared Deep Extragalactic Legacy Survey (Cycle 3) in part extends this imaging to 70 μm.
- SINGS/LVLS: The Spitzer Infrared Nearby Galaxies Survey (Cycle 0 Legacy) carried out an imaging and spectroscopic study of 75 nearby galaxies with IRAC, MIPS and IRS. The Local Volume Legacy Survey (Cycle 4) extends IRAC and MIPS imaging to a volume complete sample of 258 galaxies within 11 Mpc.

The surveys (apart from SINGS/LVLS which targeted individual objects) are listed in approximate order of increasing depth of the Spitzer observations. Other major imaging surveys have been conducted as part of Guaranteed Time Observer (GTO) and General Observer (GO) programmes. Following the exhaustion of the onboard cryogen, Spitzer's longer wavelength channels could not operate in a useful manner anymore, but its IRAC 3.6 and 4.5 μm channels could continue operations with an observing efficiency and sensitivity very close to those they offered during the Cold mission phase. The Spitzer Warm mission has therefore focused on the so called Spitzer Exploration Science Programme, whose observations require more than 500 hours of Spitzer observing time, providing an opportunity for large-scale investigations not possible during the cryogenic mission. The Spitzer Extended Deep Survey (SEDS) of a few very small fields observed by HST and The Spitzer Extragalactic Representative Volume Survey (SERVS) of about 1/3 of the SWIRE footprint are examples of Exploratory Science projects which have allowed to better sample the IRAC depth-area parameter space, as shown in Fig. 3.14. [7]

The shared aim of this ambitious extragalactic observing programme was to develop a clearer understanding of galaxy evolution processes in passive and actively star forming galaxies. As well as mapping AGN, starbursts, spirals and spheroids out to high redshifts, it aimes at studing the influence of environment on all scales, the history of star formation, the assembly of stellar mass and the role of gravitational accretion processes in AGN. While ISO's sensitivity was such that very few sources were detected beyond $z \sim 1$, Spitzer detects sources up to the highest redshifts and gives us the possibility of a more detailed investigation into the mechanisms regulating galaxy evolution (see Fig. 3.15).

Deep Spitzer observations have enabled the direct measurement of the star formation in galaxies up to $z \sim 3$. The 24 μm luminosities of galaxies gauge their bolometric IR luminosities, and hence obscured star formation rate, which represents ∼75% of the total star formation in the Universe. The results show a dependence of star-formation activity with redshift, with a peak at $z \sim 1$, an approximately constant rate to $z \sim 2.5$, and possibly a decrease in the SFR per volume at higher redshift. LIRGS (galaxies with $10^{11} L_{\odot} < L_{IR} < 10^{12} L_{\odot}$) dominate the comoving star-formation rate density for $z < 1.5$. ULIRGs at $z \sim 2$ represent ∼40% of the

[7] `http://irsa.ipac.caltech.edu/data/SPITZER/docs/spitzermission/observingprograms/es/`

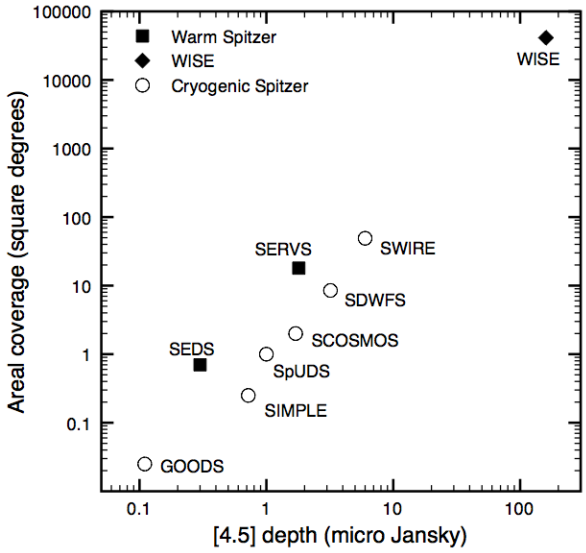

Figure 3.14: Area versus depth for SERVS compared to other surveys at wavelengths of $\sim 4.5\mu$m (Mauduit et al. 2012). For consistency, the depth shown is the 5σ limiting flux for point sources, excluding confusion noise, calculated from the Spitzer performance estimate tool [9] in each case. The surveys are: GOODS, the Spitzer IRAC/MUSYC Public Legacy in E- CDFS (SIMPLE) survey (Spitzer programme identifier [PID] 20708), the Spitzer Ultra Deep Survey (SpUDS, PID 40021, P.I. J.S. Dunlop), S-COSMOS, the Spitzer Deep Wide-Field Survey (SDWFS, Ashby et al. (2009)), SWIRE and the Spitzer Extragalactic Deep Survey (SEDS, PIDs 60022, 61040, 61041, 61042, 61043 P.I. G. Fazio). The point for the Wide- Field Infrared Explorer (WISE) is the band 2 depth from Mainzer et al. (2005).

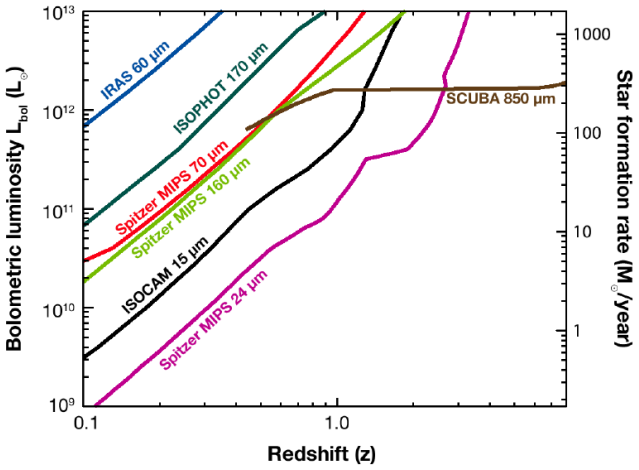

Figure 3.15: Sensitivity to the bolometric luminosity and star-formation rate, assuming star forming galaxies of various infrared and sub-millimeter experiments. Detections of at least 10 sources in the surveys can be expected in the areas above the curves. We assumed the scenario of a typical deep survey (when available). IRAS 60 μm ($S > 1$ Jy, all sky); ISOCAM 15 μm ($S > 250$ Jy, 2 deg^2); ISOPHOT 170 μm ($S > 180$ mJy, 5 deg^2); Spitzer/MIPS 24 μm ($S > 80$ μJy, 5 deg^2); Spitzer/MIPS 70 μm ($S > 25$ mJy, 5 deg^2); Spitzer/MIPS 160 μm ($S > 50$ mJy, 5 deg^2); SCUBA 850 μm ($S > 1$ mJy, 1 deg^2). Images from Lagache et al. (2005).

massive ($> 10^{11} M_{\odot}$) galaxies at that epoch and are undergoing their final mass accumulation. The specific star-formation rate for massive galaxies ($\dot{M}/M_*$) increased by more than an order of magnitude from $z \sim 0.8$ to $z \sim 2$. In Fig. (3.16), the plot IR Luminosity Density vs z shows the star formation rate in terms of redshift. In fact if we assume the relation between SFR and the luminosity of newly formed stars, $\dot{M} = K\,L$ (where typically, $K = 1.7 \times 10^{10}$, $\dot{M}$ is in solar masses per year and L is in solar bolometric luminosity units, Kennicutt 1998) we can associate to an increasing observed IR luminosity, an increasing SFR.

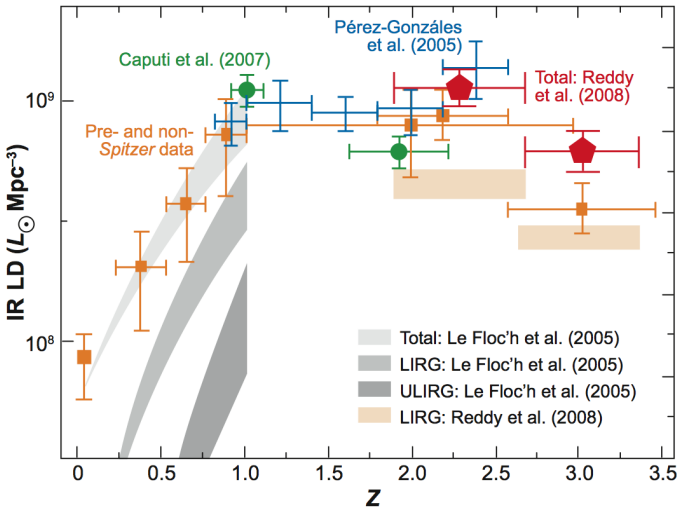

Figure 3.16: The infrared luminosity density (IR LD) as a function of redshift z to $z = 3.5$. The orange points represent pre- and non-Spitzer data. The Spitzer-based data include results from Reddy et al. (2008), Caputi et al. (2007), Pérez-González et al. (2005a), and Le Floc'h et al. (2005). The bands indicate the total, LIRG, and ULIRG contribution to the total IR LD at $z < 1$, as determined by Le Floc'h et al. (2005). Image from Soifer et al. (2008).

Like previously ISO, and consistently with the observed increase of the average star formation rate with redshift, Spitzer shows faint counts far in excess of the predictions of no-evolution extrapolations. At all FIR wavelengths, particularly at 24 μm, where counts reach more than 3 orders of magnitude deeper than previous missions, and 70 μm, where the counts go much deeper than ISO, Spitzer's extensive surveys provide much larger samples than previously available.

Frayer et al. (2006b) presents number counts at 70 μm in the GOODS-N field, Fig. (3.17). They show that the differential source counts turn over at $8-10$ mJy; below this flux level, the counts fall steeply with a slope comparable to that reported at 24 μm. Down to their limiting flux of $\sim$ 1.2 mJy, their number counts are fit tolerably well by the Lagache et al. (2004) models which fits apparently well also the number counts at 24 μm as described by Shupe et al. (2008). At the bright end, the results are consistent with those of Dole et al. (2004), who report a number density of 240 galaxies/deg^2 brighter than $\sim$ 15 mJy, which exceeds the Euclidean extrapolation by a factor of $\sim$ 10. Dole et al. (2004) evaluates number counts at 160 μm, extending them down to 50 mJy, where $N(> 50\text{mJy}) \sim 290$ galaxies/deg^2 and the differential counts exceed the Euclidean extrapolation by $\sim$ 5 times. Above $\sim$ 200 mJy, the counts agree with previous results from the ISO Far-Infrared Background (FIRBACK) survey at 170 μm. They conclude that most of the sources responsible both for the 160 μm counts

and for the 70 μm counts at ~ 15 mJy lie at $z \sim 0.7$, with a tail to $z \sim 2$.

Models which allow evolution of only the more luminous sources, including the possibility of discontinuous changes in the evolution rate, or large changes to source SEDs, have less difficulty in fitting the counts. By contrast, models which use smoothly changing evolution rates, such as those of Rowan-Robinson (2001), have more difficulty.

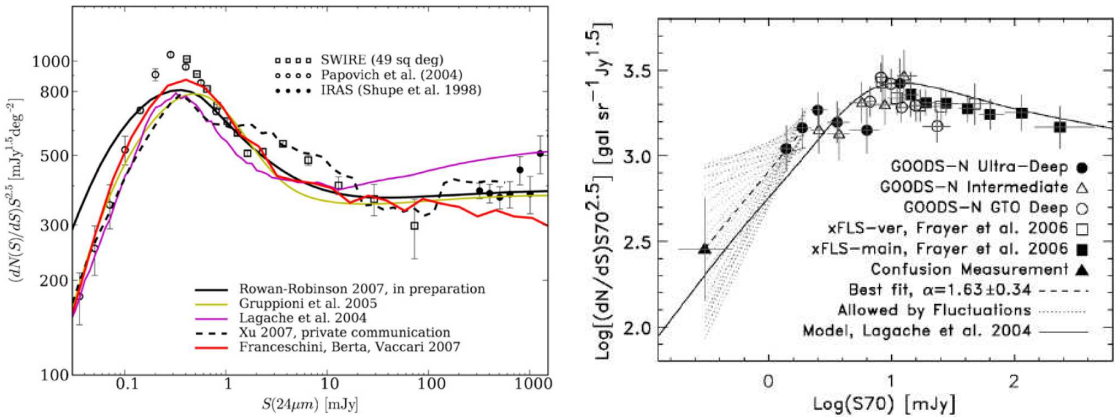

Figure 3.17: **Left**: Plot of normalized number counts at 24 μm with several models overlaid. Image from Shupe et al. (2008). **Right**: The 70 μm counts from Spitzer surveys. Image from Frayer et al. (2006b).

The resolution of the CIRB into individual sources has been a central scientific goal for Spitzer. By integrating the counts in each Spitzer bandS, it is possible to derive a lower limit to the background; this limit can be increased by extrapolating the counts to zero flux. In no case does this estimate disagree with direct measurements or upper limits on the background (Hauser & Dwek 2001).

At 160 μm the point sources detected by Spitzer fall far short of completely resolving the CIRB. Dole et al. (2006b) addressed the shortfalls by exploiting the smaller beam and substantially higher sensitivity of Spitzer at 24 μm (compared to the longer wavelengths) to enhance the depth at the longer wavelengths via a stacking analysis. Their results are included in Tab. 3.4, which gives the contribution to the CIRB at 24, 70, and 160 μm from galaxies individually detected at 24 μm and from fainter 24 μm sources based on an extrapolation of the Spitzer data. The table also includes the Dole et al. best estimate of the total CIRB.

Wavelength	**24 μm**	**70 μm**	**160 μm**
Background due to sources $> 60\mu$Jy	2.16 ± 0.26	5.9 ± 0.9	10.7 ± 1.6
Background due to sources $< 60\mu$Jy	0.54	1.2	2.6
Total from 24 μm sources	2.7	7.1 ± 1.0	1.35 ± 1.7
Cest estimate of total background	2.7	6.4	15.4

Table 3.4: Contribution of 24 μm galaxies to the CIRB (Dole et al. 2006b). Note: All brightness are in units of nW m^{2}str^{-1}.

We report in Fig. (3.19) a graph detailing which observations have contributed to the extragalactic background measurement and its interpretation in terms of resolved sources. This figure shows with more detail that the optical and infrared cosmic backgrounds are well separated and that the power in the infrared is comparable to that in the optical. In contrast,

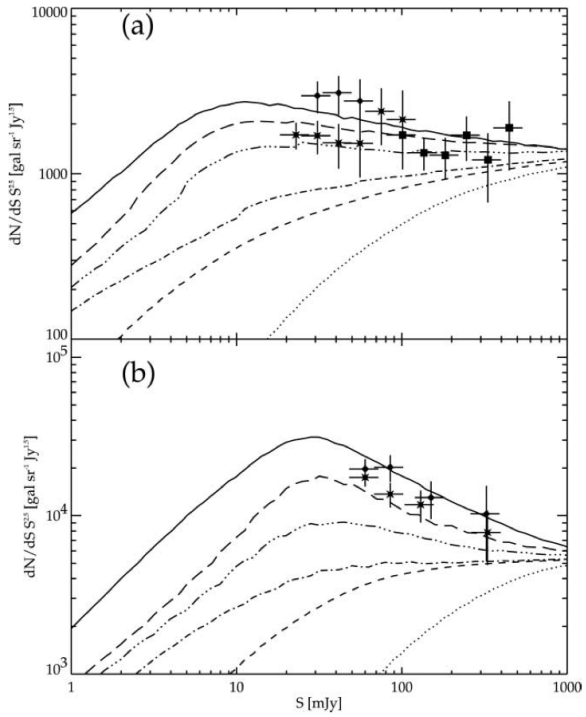

Figure 3.18: Cumulative contributions to the differential source counts at 70 and 160 μm of galaxies, as a function of redshift, form the model of Lagache et al. (2004): (a) 70 μm; (b) 160 μm. Galaxies contributing to the counts at redshifts 0.1 (dotted line), 0.3 (dashed line), 0.7 (dot-dashed line), 0.9 (triple-dot-dashed line), and 1.1 (long-dashed line). The total contribution is the upper solid line. The image is from Dole et al. (2004), see the original paper for the explanation of the symbols.

we know that locally, the infrared output of galaxies is only one third of the optical output. This implies that infrared galaxies grow more luminous with increasing z faster than optical galaxies do. This fact is also illustrated in Fig. 3.20, based on the predictions of the Lagache et al. (2004) model.

3.7 Forthcoming Initiatives in Infrared Astronomy

We have summarized in the previous sections the main results obtained during the last 15 years in the infrared extragalactic astronomy. In May 2009 the Herschel Space Observatory was launched and now and for the following years we have the opportunity to exploit its exciting data (the early Herschel results was published in an Astronomy and Astrophysics special issue, Walmsley et al. (2010)). I will describe Herschel Observatory and Mission in the following section, but to emphasize the importance of IR and sub-mm astronomy, it must be said that the next few years have other missions already in the pipeline, in particular ALMA & JWST.

The *Atacama Large Millimeter/sub-millimeter Array* (ALMA), one of the largest ground-based astronomy projects of the next decade, is a major new facility for world astronomy. ALMA will be comprised of a giant array of 12-m antennas, with baselines up to 16 km and state-of-the-art receivers that cover all the atmospheric windows up to 1 THz. An additional,

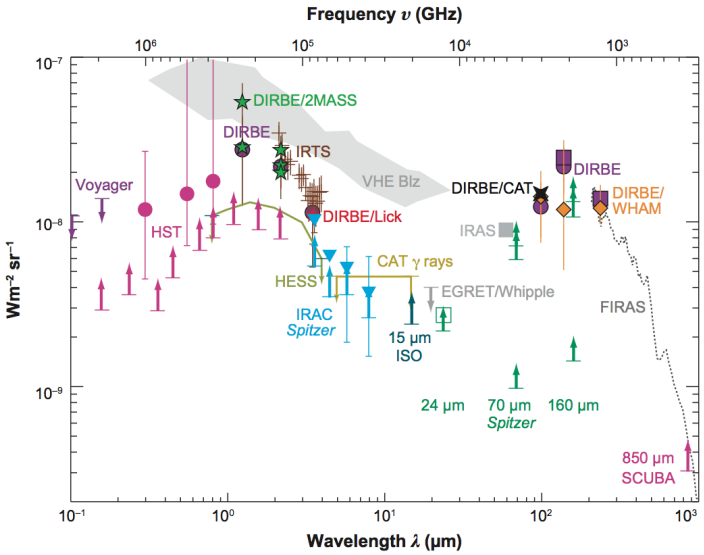

Figure 3.19: The extragalactic background light spectral energy distribution from 0.1 to 1000 μm, adapted from Dole et al. (2006b). At Spitzer wavelengths from $24 - 160\,\mu$m, the green arrows are lower limits based on directly observed sources. The higher green arrows at 70 and 160 μm are based on stacking numerous 24 μm detections, as discussed in Dole et al. (2006b) whereas the highest green arrows at these wavelengths reflect the estimated contribution of all 24 μm sources, including those beyond the sensitivity limit of the stacking analysis. The green open square is the estimated 24 μm background when these faint sources are included. Note that the corrected 70 and 160 μm background estimates from the MIPS data are essentially equal to the backgrounds estimated at these wavelengths by other techniques. See Dole et al. (2006b) for identification of the other datasets and background estimates included in this figure. Image from Soifer et al. (2008).

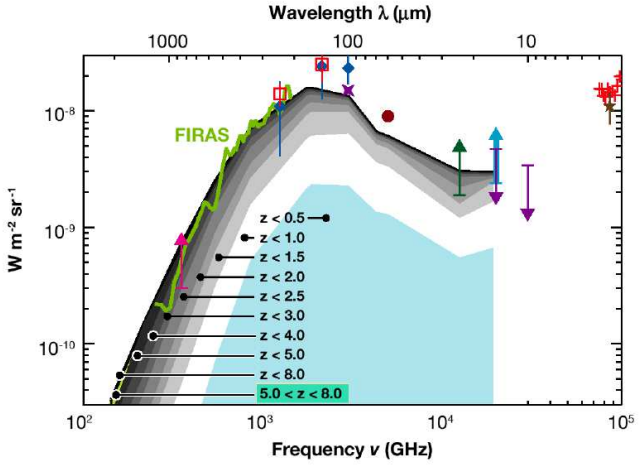

Figure 3.20: Cumulative contribution to the CIRB of galaxies at various redshifts from 0.5 to 8, from the model of Lagache et al. (2004). The reference of the image's data are reported in Lagache et al. (2005).

compact array of 7-m and 12-m antennas will greatly enhance ALMA's ability to image extended sources. Construction of ALMA started in 2003 and will be completed in 2013. The ALMA project is an international collaboration between Europe, East Asia and North Amer-

ica in cooperation with the Republic of Chile. ALMA is located on the Chajnantor plain of the Chilean Andes in the District of San Pedro de Atacama, 5000 m above sea level. ALMA will enable transformational research into the physics of the cold Universe, regions that are optically dark but shine brightly in the millimeter portion of the electromagnetic spectrum. Providing astronomers a new window on celestial origins, ALMA will probe the first stars and galaxies, and directly image the formation of planets. ALMA will operate at wavelengths of 0.3 to 9.6 millimeters, where the Earth's atmosphere above a high, dry site is largely transparent, and will provide astronomers unprecedented sensitivity and resolution. ALMA will be a complete astronomical imaging and spectroscopic instrument for the millimeter/sub-millimetre.

The *James Webb Space Telescope* (JWST), with a 50-Kelvin telescope and ∼60 times the collecting area of Spitzer, will far surpass the sensitivity and angular resolution of Spitzer in the mid-IR when it is launched in the next decade. JWST will enable more detailed studies of the distant universe revealed by Spitzer. Its imaging and spectroscopy at $\lambda < 25\mu$m will provide powerful probes of sources that Spitzer can detect only in deep imaging.

The relevance of IR and sub-mm astronomy has grown in recent years but the years ahead will be even more packed of discoveries thanks to all facilities that at this time and in the years to come will explore the sky at these wavelengths.

3.8 The Herschel Space Observatory

Herschel is the fourth 'cornerstone' mission in the ESA science programme [10]. With a 3.5 m Cassegrain telescope it is the largest space telescope ever launched. It is performing photometry and spectroscopy in approximately the 55-671 μm range, bridging the gap between earlier infrared space missions and ground-based facilities.
We can easly understand the importance of Herschel looking at the Fig. 3.19: the SPIRE and PACS wavelengths range (70-500 μm) cover the CIRB peak and the Herschel sensitivity and spatial resolution will be crucial to resolve it in galaxies.

Herschel has been designed to observe the 'cool universe'; it is observing the structure formation in the early universe, resolving the far infrared cosmic background, revealing the cosmologically evolving AGN/starburst symbiosis and galaxy evolution at the epochs when most stars in the universe were formed, unveiling the physics and chemistry of the interstellar medium and its molecular clouds, the wombs of the stars, and unravelling the mechanisms governing the formation and evolution of stars and their planetary systems, including our own solar system, putting it into context. In short, Herschel is opening a new window to study how the universe has evolved to become the universe we see today, and how our star, the Sun, our planet, the Earth, and we ourselves fit in. Herschel is operated as an observatory facility. It is available to the worldwide scientific community, roughly two thirds of the observing time is 'open time' allocated through standard competitive calls for observing proposals.

Herschel was successfully launched on 14 May 2009. However, it was conceived almost 30 years earlier as the Far InfraRed and Submillimetre Space Telescope (FIRST), which was formally proposed to ESA in November 1982. It was incorporated in the ESA 'Horizon 2000' long-term plan for implementation and in November 1993 ESA decided that FIRST would be implemented as the fourth 'cornerstone' mission.

[10] `http://herschel.esac.esa.int/`, `http://herschel.cf.ac.uk/`

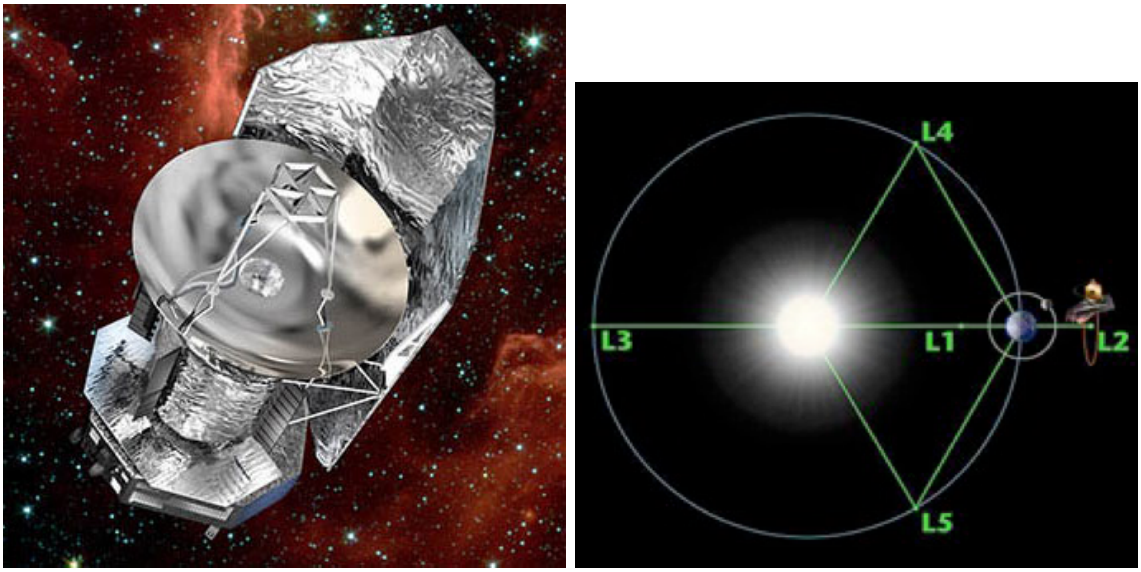

Figure 3.21: The Herschel Space Observatory and its orbit.

The mission, approved in 1993 (a spacecraft design with a 3 m telescope passively cooled to 160 K, two science instruments), was changed in the light of the experience gained from the ISO mission: the current mission is based upon a spacecraft employing a superfluid helium cryostat reusing ISO technology, autonomously operating in a orbit around the 2nd Lagrangian point (L2) in the Sun-Earth/Moon system (see Fig. 3.21).

The prime science objectives of Herschel are intimately connected to the physics of the processes taking place in the inter-stellar medium (ISM) in the widest sense. Near and far in both space and time, going from solar system objects and the relics of the formation of the Sun and the Solar System, to star formation in and feedback by evolved stars to the ISM, to the star-formation history of the universe, galaxy evolution and cosmology.

3.8.1 The Herschel Telescope & and Focal-Plane Instruments

The Herschel telescope (Doyle et al. 2009) was constructed to be as large as possible, still be compatible with the size of the launcher (Ariane 5) and no in-flight deployable structures, low mass, and as cold as possible with passive cooling.

The optical design is that of a classical Cassegrain telescope with a 3.5 m physical diameter primary and an 'undersized' secondary, yielding an effective primary diameter of 3.28 m. The primary mirror has been made out of 12 segments, the secondary is a single piece. Similar to ISO, a fully passive design was adopted.

In the Tab. 3.5 the main characteristic of the telescope are summarized.

The science payload consists of three instruments, provided by consortia of institutes led by PIs:

- PACS: The Photodetector Array Camera and Spectrometer, Poglitsch et al. (2010), PI: A. Poglitsch, Max-Planck-Institut für extraterrestrische Physik (MPE), Garching. PACS is a camera and low to medium resolution spectrometer for wavelengths in the

Telescope	
Primary physical/effective diameter	3.5 / 3.28 m
Secondary diameter	30.8 cm
System/primary f-number	8.70 / 0.5
Wave front error best-focus (centre/edge)	4.8 / 5.5 μm
Angular resolution	$\sim 7" \times (\lambda_{\rm obs}/100\ \mu{\rm m})$
Operational temperature	~85 K

Table 3.5: Herschel telescope main characteristics.

range 55-210 μm. It employs four detector arrays, two bolometer arrays and two Ge:Ga photoconductor arrays. The bolometer arrays are dedicated for wideband photometry, while the photoconductor arrays are to be employed exclusively for spectroscopy with a resolution of a few thousand. PACS can be operated either as an imaging photometer, or as an integral field line spectrometer.

- SPIRE: The Spectral and Photometric Imaging REceiver, Griffin et al. (2010), PI: M. J. Griffin, Cardiff University. SPIRE is a camera and low to medium resolution spectrometer complementing PACS for wavelengths in the range 194-672 μm. It comprises an imaging photometer and a Fourier Transform Spectrometer (FTS), both of which use bolometer detector arrays. There are a total of five arrays, three dedicated for photometry and two for spectroscopy.

- HIFI: The Heterodyne Instrument for the Far Infrared, de Graauw et al. (2010), PI: T. de Graauw, in late 2008 succeeded by F. Helmich, SRON Netherlands Institute for Space Research, Groningen. HIFI is a very high resolution heterodyne spectrometer covering the 490-1250 GHz and 1410-1910 GHz bands. It utilises low noise detection using superconductor-insulator-superconductor (SIS) and hot electron bolometer (HEB) mixers, together with acousto-optical and autocorrelation spectrometers. HIFI is not an imaging instrument, it observes a single pixel on the sky at a time.

The three instruments enabling Herschel to offer broad band photometric imaging capability in six bands with centre wavelengths of 70, 100, 160, 250, 350, and 500 μm, imaging spectroscopy over the entire Herschel wavelength coverage, and very high resolution spectroscopy over much of this range. A number of observing modes are provided, including point source photometry, small, and large area photometric imaging, and the observation of a single spectral line, or one or more spectral ranges, in either a single position or in various mapping modes.

A summary of the most important instruments characteristics are shown in Tab. 3.6 (Pilbratt et al. 2010). In Fig. 3.22 the PACS & SPIRE filter transmissions are reported.

Since the beginning of the observations, Herschel got results beyond expectation showing really good optical performances as illustrated in Fig. 3.23 reported in Pilbratt et al. (2010).

3.8.2 The Herschel Mission and its GT and OT Science Programmes

After the launch and the early check-out phase, the Herschel inflight mission phases proceeded as follows:

HIFI	**Heterodyne spectrometer**
Wavelength coverage	157−212 & 240−625 μm
Field-of-view (FOV)	single pixel on sky
Detectors	5×2 SIS & 2×2 HEB mixers
Spectrometers	auto-correlator & acousto-optical
Spectral resolving power	typically 10^6
PACS	**2-band imaging photometer**
Wavelength coverage	60−85 or 85−130, 130−210 μm
Field-of-view (FOV)	$0.5F\lambda$ sampled $1.75' \times 3.5'$
Detectors	64×32 & 32×16 pixel bol. arrays
Diffraction limited beams	5.5", 7.7", 12" respectively
PACS	**integral field spectrometer**
Wavelength coverage	55–210 μm
Field-of-view (FOV)	(5×5 pixel) 47" $\times$ 47"
Detectors	two 25×16 pixel Ge:Ga arrays
Spectral resolving power	1000−4000
SPIRE	**3-band imaging photometer**
Wavelength bands ($\lambda/\Delta\lambda \sim 3$)	250, 350, 500 μm
Field-of-view (FOV)	$2F\lambda$ sampled 4' x8'
Detectors	139, 88 & 43 pixel NTD bol. arrays
Diffraction limited beams	18", 25", 36" respectively
SPIRE	**imaging fourier transf. spectrometer**
Wavelength coverage	194−324 & 316−671 μm
Field-of-view (FOV)	$2F\lambda$ sampled circular 2.6'
Detectors	37 & 19 pixel NTD bol. arrays
Spectral resolving power	370−1300 (high) / 20−60 (low)

Table 3.6: Herschel science instrument main characteristics. Acronyms relating to the detectors: superconductor-insulator-superconductor (SIS), hot electron bolometer (HEB), gallium-doped germanium (Ge:Ga), and neutron transmutation doped (NTD).

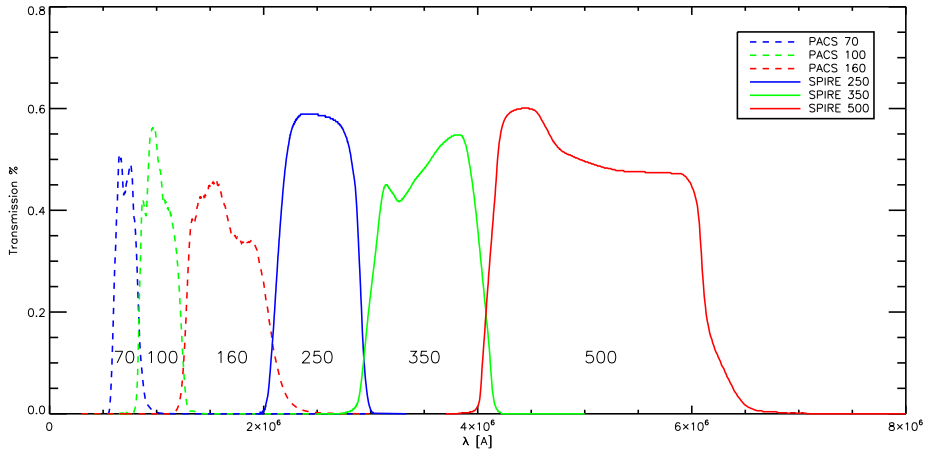

Figure 3.22: PACS & SPIRE filter transmissions

- Commissioning Phase (CoP) : Complete check-out of spacecraft functions and performance. Switch-on and functional verification of the instruments. Cryo-cover opening and thermal stabilization. Nominal duration of two months.
- Performance Verification Phase (PVP) : instrument performance, calibration, focal plane geometry, and pointing determination. Testing, optimizing, verifying, and releasing the various instrument/observing modes for use. Nominal duration of three months.
- Science Demonstration Phase (SDP): use of the released observing modes to execute selected observations from the approved Key Programme (KP, see later in the text). Assessing, optimizing, and releasing the KP observations for execution. Nominal duration of six weeks.
- Routine Science Phase (RSP) : execute released observations employing released observing modes. Perform engineering and routine calibration observations to optimize the observing modes and the quality of the data obtained. The RSP will continue until Helium exhaustion, and will be followed by archive phases.

The first three phases have been successfully concluded for SPIRE and PACS and the first results were published in a special *Astronomy and Astrophysics* issue (Walmsley et al. 2010). Some problem have been found in HIFI. In particular, on observational day 81 (August 2, 2009), telemetry from Herschel showed that HIFI had entered an anomalous state because of a single event upset, caused by a cosmic particle that damaged the voltage supply for the instrument's control electronics. With HIFI unavailable, the PVP activities had to be replanned, making more use of PACS and SPIRE and a new phases called HIFI Priority Science Programme (PSP) was performed to test the HIFI functionality. Now all work fine and the first results were published in a special *Astronomy and Astrophysics* issue (Walmsley et al. 2010).

Herschel is an observatory, and is thus open to the astronomical community, but a large chunk of observing time has been allocated prior to launch to a number of key science programmes. Some are allocated to the instrument teams, so that they get data back from the

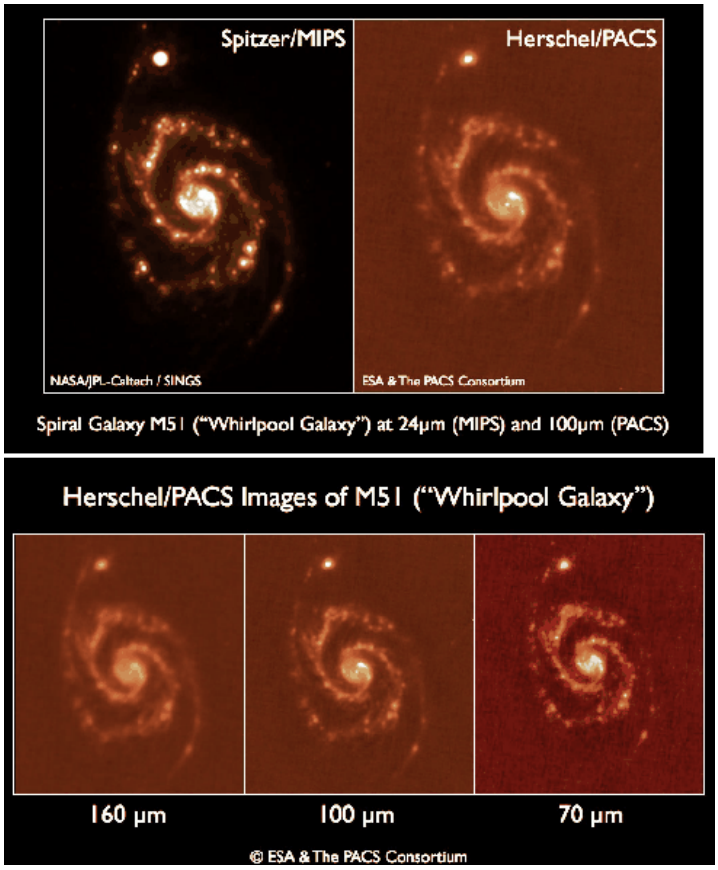

Figure 3.23: **Top**: Spitzer/MIPS and Herschel/PACS M51 images at 24 μm and 100 μm, respectively. Because aperture sizes and wavelengths have similar ratios and Spitzer has perfect optical performance, it can be concluded that Herschel has good optical performance at 100 μm. **Bottom**: Herschel/PACS images at 160, 100, and 70 μm, demonstrating the im-provement of the angular resolution with shorter wavelengths (Pilbratt et al. 2010).

mission they have worked so hard to build. Others are allocate to the general community, and the telescope time is bid for in the same way that other observatories work. The Heschel observing time is then shared between guaranteed and open time (GT and OT).

The basic rules were defined in the Science Management Plan (SMP) as part of the AO for the science payload in 1997. In the nominal mission ∼20 000 hours are available for science, 32% is GT (mainly owned by the PIs consortia). The remainder is OT, which is allocated to the general community (including the GT holders) on the basis of AO for observing time. In each AO cycle the GT is allocated first, followed by the OT. All observations made in the first year of the RSP will have proprietary times of 12 months, while for all observations made later, the proprietary time will be 6 months, with a simple 'bridging scheme' so that no observation will become public before observations that were executed earlier become public as well. A small amount of the open time can be allocated as discretionary time. All proposals are assessed by the HOTAC, and all observing data are archived and will be available to the entire community after the proprietary time has passed.

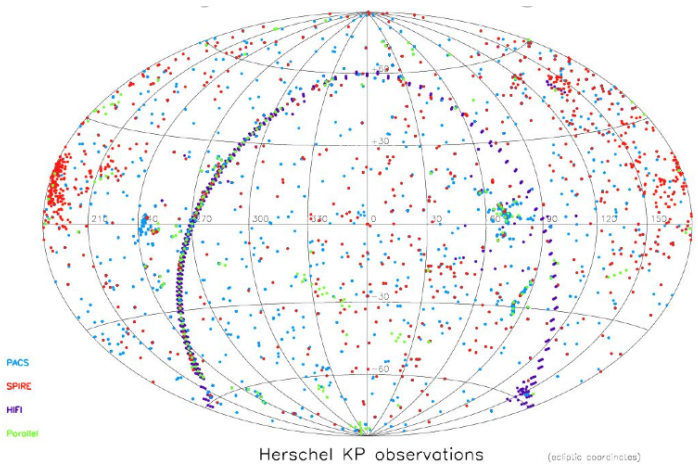

Figure 3.24: Key Programmes observations plotted in ecliptic coordinates, with colour-coding denoting the instrument. "Parallel" indicates performing 5-band photometric imaging using SPIRE and PACS simultaneously.

Table 3.7: Herschel Approved KP proposals by science area (Pilbratt et al. 2010)

Science area	KP GT		KP OT		KP TOTAL	
	#	h	#	h	#	h
Solar System	1	293.7	1	372.7	2	666.4
ISM/SF	10	2337.5	10	2113.2	20	4450.7
Stars	2	544.6	0	0	2	544.6
Gal - AGN	5	983.7	8	1930.3	13	2914.0
Cosmology	3	1719.4	2	962.6	5	2682.0
Total	21	5878.9	21	5378.8	42	11257.7

Also considering that Herschel would not have the benefit of an all sky survey for much of its wavelength coverage, the Herschel Science Team (the team, including the Mission Scientist, G. Pilbratt, and instruments PIs, ruling the scientific exploitation of the observatory) required that KPs in the form of large spatial and spectral surveys were to be selected and executed early in the mission, so that the results could be followed up by Herschel itself. An initial AO limited to KP observing proposals therefore was issued. This process took place from February 2007 to February 2008. There were 21 GT and 62 OT proposals, out of which by coincidence also 21 were awarded observing time (Tab. 3.7 and Fig. 3.24).

Herschel is carrying out surveys of unprecedented size and depth, vastly improving the state of observations in the under-explored FIR and Sub-mm wavebands. It can observe a huge range of objects, from comets in the Solar System to galaxies in the distant reaches of the Universe. The science areas are split into the six broad categories: *Solar System*, *Planet Formation and Planetary system*, *Star and their environments*, *Interstellar Medium and Star Formation*, *Galaxies*, *Cosmology*. Altogether 3 calls for proposals have been issued,

and the projects selected are many in each mentioned category. Our present work is centered on extragalactic and cosmological observations, for this reason we report here only the key programme extragalactic surveys within Herschel GT and OT. More detailed descriptions of the whole Herschel observing programme are at `http://herschel.cf.ac.uk/mission/key-programmes/`.

- HerMES: the Herschel Multi-tiered Extragalactic Survey (GT KP, Oliver et al. (2012), in prep. [11]). HerMES is an astronomical project to study the evolution of galaxies in the distant Universe enabled now by the the sub-mm bands sampled by Herschel. The light from most of these galaxies will have taken more than 10 billion years to reach us, which means we will see them as they were 3 or 4 billion years after the big bang. It is the biggest project on Herschel, the observations are done with both SPIRE and PACS instruments, surveying $\sim 70\mathrm{deg}^2$ from $20' \times 20'$ to $3.6^\circ \times 3.6^\circ$ (for a total of 900 hours observing time) and 12 clusters.

- The Dusty Young Universe (GT KP, [12]). Observations of some of the most distant quasars in the far-infrared shows that a substantial amount of dust was created in the first billion years after the Big Bang. However, the wavelengths observed to date haven't been able see the bulk of the emission, which is now possible using SPIRE and PACS. By observing over 100 of the most distant quasars it is possible to analyse their properties. In addition, PACS spectroscopy of four of the brightest quasars will allow the different origins of the far-infrared emission to be disentangled.

- PEP: PACS Evolutionary Probe (GT KP, Lutz et al. (2011) [13]. It is aimed to study the rest-frame far-infrared emission of galaxies in the past 11 billion years (up to redshifts of about 3) as a function of environment. The survey will shed new light on the constituents of the cosmic IR background and their nature, as well as on the co-evolution of AGN and starbursts. The observation are PACS only but PEP is coordinated with SPIRE observations of the same fields in the HerMES programme. It surveys 2.7 deg^2 from $10' \times 10'$ to $85' \times 85'$ (550 hours of observing time) and 10 clusters.

- Herschel-ATLAS (OT KP, Eales et al. (2010a) [14]. The Herschel ATLAS is the Herschel survey covering the largest total area, it will survey 550 deg^2 of sky in five far-infrared and sub-mm bands, i.e. one eightieth of the sky. There are five main science programmes, the principal one being a survey of dust emissions in 50,000 nearby galaxies. This follows on from the highly successful Sloan Digital Sky Survey, an optical survey that surveyed hundreds of thousands of galaxies in the nearby universe. However, Sloan missed the objects that are hidden by dust and didn't tell us anything about the gas in the galaxies out of which the stars form. H-ATLAS will do both. The observations are done using both SPIRE and PACS in parallel mode.

- GOODS: Great Observatories Origins Deep Survey (OT KP, Elbaz et al. (2011)). The GOODS fields are small regions of the sky which have been observed very deeply in all

[11] `http://hermes.sussex.ac.uk/`
[12] `http://www.mpia.de/Public/Forschung/Instrumentation/DustyYoungUniverse/index.html`
[13] `http://www.mpe.mpg.de/ir/Research/PEP/index.php`
[14] `http://www.h-atlas.org/`

available wavebands from X-ray to radio by many of the greatest space observatories, including Hubble, Chandra and Spitzer. By observing a sky area totalling about 300 square arcminutes with PACS, the project aims at detecting some of the faintest objects, and galaxies in the earliest stages of the Universe. The observations are done with both SPIRE and PACS instruments: PACS very deep imaging of the GOODS fields (330 hours); SPIRE deep imaging of the GOODS-N Field (30 hours).

Most of the results presented in this work on the Herschel LF is done based on the data taken within the HerMES projects. For this reason in the next section we give some more details about this survey and its results.

3.8.3 The Herschel Multi-tiered Extragalactic Survey (HerMES) Project

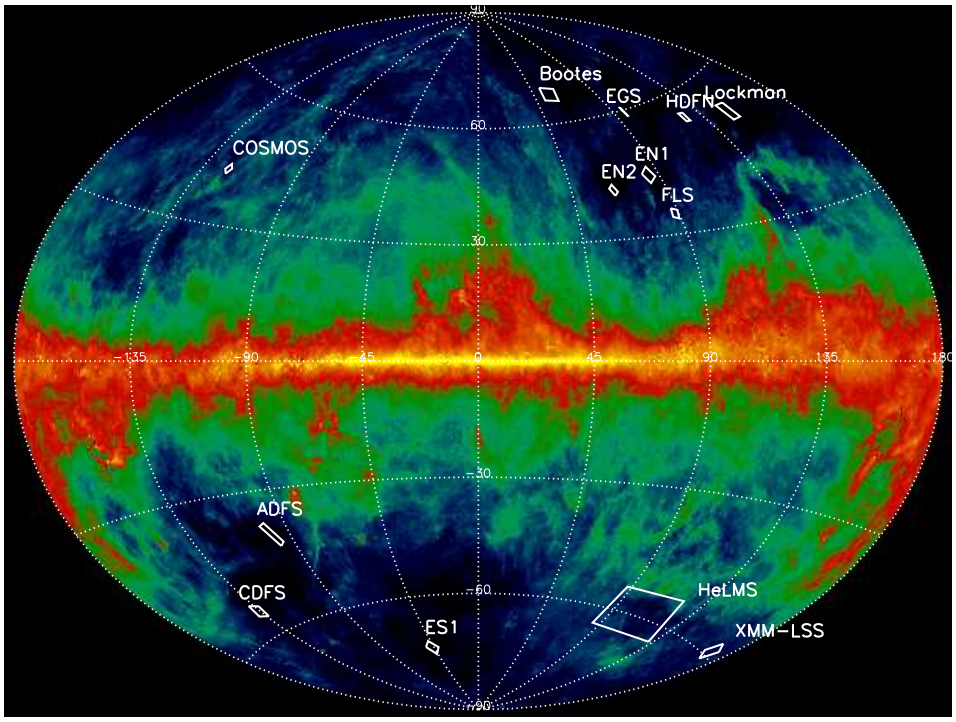

Figure 3.25: HerMES fields overlaid on IRAS/COBE sky maps by Schlegel et al. (1998) in Galactic coordinates. Most fields were chosen so as to be the largest fields of their size characterized by very-low FIR background and far from bright IRAS sources, whereas a few were chosen due to the availability of a large body of multi-wavelength data at non-IR wavelengths and/or to their being observable from both hemispheres.

The Herschel Multi-tiered Extragalactic Survey, HerMES, is a legacy program designed to map a set of nested fields totaling ∼380 deg^2. Fields range in size from 0.01 to ∼20 deg^2, using Herschel-SPIRE (at 250, 350 and 500 μm), and Herschel-PACS (at 100 and 160 μm), with a wider component of 270 deg^2 with SPIRE alone (Oliver et al. 2012). The goal of this programme is to provide a legacy survey of star forming galaxies over the wavelengths at which the galaxy SEDs and the CIRB background peak (see Fig. 3.26)

The survey is defined by Astronomical Observing Requests (AORs). Details of the observing modes can be found in the Herschel observer's manuals (especially the SPIRE Observers

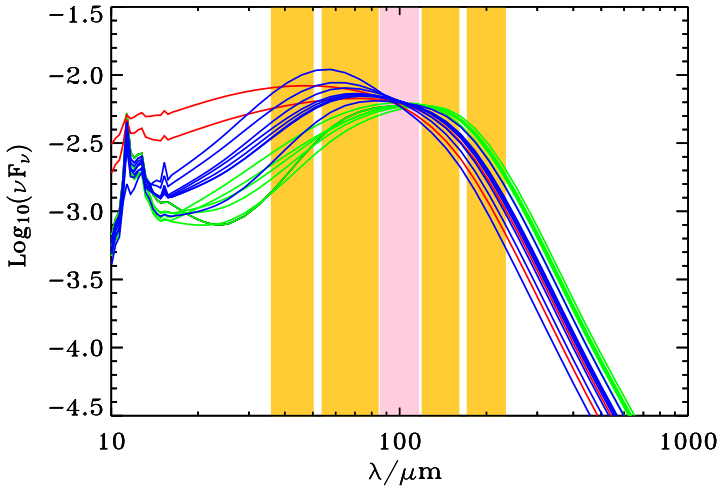

Figure 3.26: SED coverage with Herschel/HerMES SPIRE/PACS bands, Oliver et al. (2012). Model spiral (green), star-burst galaxy (blue) and AGN (red) Spectral Energy Distributions (SEDs) normalised to the same $L_{\rm FIR}$ and plotted in their rest-frame with the *Herschel*-PACS and *Herschel*-SPIRE bands at $\lambda = 100, 160, 250, 350$ and 500μm plotted at $\lambda/(1+z)$ for a galaxy at $z = 1.5$. Note that the *Herschel*-SPIRE band at 250μm measures a similar flux density for all and so is a reasonable proxy for the $L_{\rm FIR}$ for these templates.

Manual, v2.2, HERSCHEL-DOC-0798 [15]). The point source sensitivities of the different observing modes have been quoted using official mission values and are reported in Tab. 3.8.

Mode	**Wavelength [μm]**				
	100	160	250	350	500
SPIRE scan (nominal, 30"/s)			64	53	76
SPIRE scan (fast, 60"/s)			91	75	108
PACS scan (medium, 20"/s)	42	80			
Parallel mode (nominal, 20"/s)	71	135	37	30	44

Table 3.8: HerMES point source sensitivities (5-σ sensitivities [mJy $\sqrt{\rm N_{scan}}$]) for a single scan in various observing modes.

The combination of PACS and SPIRE bands can extensively probe the peak of the far infrared spectral energy distributions of star forming galaxies and thus measure the infrared luminosity, $L_{\rm IR}$, with great accuracy. For this reason the primary criterion in the construction of the observation is to sample the $(L_{\rm IR}, z)$ plane of star-forming galaxies uniformly and with sufficient statistics up to a redshift of $0 < z \leq 3$ at least. By applying the Lagache et al. (2003) models it has been possible to evaluate the area needed to reach the required resolution in luminosity and redshift to best sample the $(L_{\rm IR}, z)$ plane (see Fig. 3.27). Each tier thus probes a given $(L_{\rm IR}, z)$ region bounded by the areal constraint and the flux limit. An optimized sampling of the wavelength space is achieved by combining HerMES with the PACS Evolutionary Probe survey.

[15] `http://herschel.esac.esa.int/Documentation.shtml`

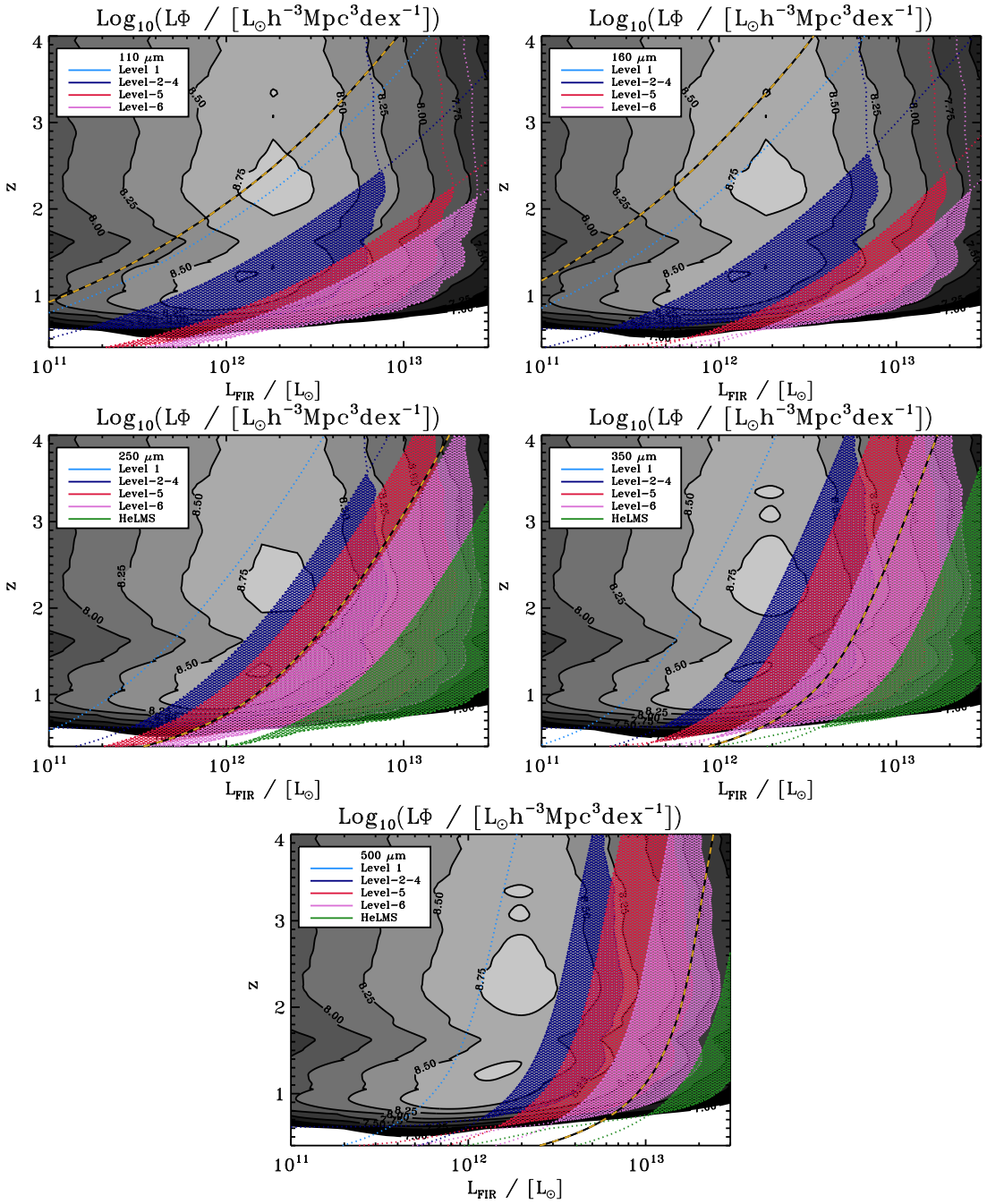

Figure 3.27: Far infrared luminosity density in $\log(L_\odot h^{-3}\mathrm{Mpc}^3\mathrm{dex}^{-1})$ (grey-scale and contour diagram) as a function of far infrared luminosity (x-axis) and redshift (y-axies) - from the model of Lagache et al. (2003) (Oliver et al. 2012). The power of different survey elements to probe this space are indicated by overlays. Each panel shows survey elements at different wavelengths; reading left-to-right from the top they are 100, 160, 250, 350 and 500 μm. Surveys are deemed to properly sample the space if they can detect galaxies of these FIR luminosities at the 5-σ instrumental noise level and with more than 75 galaxies in bins of $\Delta \log L\, \Delta z = 0.1$. These two constraints are marked with dotted lines and are hatched. The different survey levels are shown with: Levels 2-4 - blue; Level 5 - red; Level 6 - magenta and HeLMS - green. Level-1 (cyan) does not have enough volume to satisfy the number of galaxies criterion and so only the instrumental noise limit is shown. The 5σ confusion noise levels (after 5σ clipping) from Berta et al. (2011) [at 100 and 160μm] and Nguyen et al. (2010) [at 250, 350 and 500μm] with yellow/black lines. Note the bimodal peaks at $z \sim 1$ and $z \sim 2.5$

HerMES was thus designed to comprise a number of tiers of different depths and areas (see Tab. 3.12 and Tab. 3.13 here reported fromOliver et al. 2012 for the complete census of fields; for a summary of the Levels depths respectively a summary of levels depths and expected number of detected sources by mean of models see Tab. 3.9).

Levels	Area [deg^2]	$5\sigma_{250}$ [mJy]	$N_{\rm Val}$ [10^3]	$N_{\rm Glenn}$ [10^3]	$N_{\rm cat}$ [10^3]
PACS Ul.	0.012				
Level 1	0.15	4	2.2	2.0 ± 0.1	—
Levels 2-4	6.0	10	17	22.4 ± 0.9	
Level 5	37	15	53	73.6 ± 2.3	52
Level 6	52	26	20	28.1 ± 0.6	30
H-ATLAS	570	45	76	90.6 ± 2.9	115
Level 7 (HeLMS)	270	64	130		24

Table 3.9: Projected SPIRE survey results for the 250μm band. This table simplifies the survey giving approximate instrumental noises in 4 tiers (L1 includes GOODS-N). The 5σ confusion noise from Nguyen et al. (2010) is 29 mJy, approximately the Level 6 depth. Numbers of 250 μm sources are estimated from: a count model [NVal]Valiante et al. (2009); HerMES $P(D)$ analysis [$N_{\rm Glenn}$] Glenn et al. (2010) and from raw number counts in fields that are obtained at these depths, extracted as described in [$N_{\rm cat}$] Smith et al. (2012)

.

HerMES samples the higher luminosity objects, which are bright but rare, in the wide shallow tiers, and the lower luminosity galaxies, which are faint but very numerous and marginally confused, in the deep narrow-field tiers. The total time allocated for HerMES is 909.3 hours. This comes from the Guaranteed Time awarded to the SPIRE instrument team (850 hr), from one of the Herschel Mission Scientists (M. Harwit, 10 hr) and from members of the Herschel Science Centre (B. Altieri, L. Conversi, M. Sanchez Portal and I. Valtchanov, 40hr). ESA also effectively contributed 9.3 hours as we agreed for our Abell 2218 observations in the Science Demonstration Phase to be made public immediately and so were not charged for these observations.

Confusion is a serious issue for Herschel and SPIRE, in particular, and is an important driver in deciding about the survey depths; making a maximum use of the full spectrum of ancillary data, it is possible to limit the confusion problem at the moment of source identification. In order to pursue multi-wavelength analyses, the selected fields (Fig. 3.8.3) are among the most intensively observed at all wavelengths. These incude: radio (VLA, WRST, GMRT, ATCA); sub-mm (SCUBA, Bolocam, AzTEC, MAMBO); mid and far infrared (Spitzer, ISO, AKARI); near-infrared (UKIRT, VISTA); optical (HST, Subaru SuprimeCAM, CFHT MegaCAM, KPNO MOSAIC1, CTIO MOSAIC2, INT WFC); UV (GALEX) and X-ray (XMM-Newton, Chandra). Extensive redshift and/or photometric redshift surveys are either available or underway for most of these fields. However, all these data were available in separate and patchy databases and therefore it has been necessary to collect them in a unique and homogeneous dataset which is the Spitzer-Selected Multi-Wavelength Wide-Area Data Fusion described in the following chapter (the Spitzer Data Fusion represnts also the guide to the XID source extraction Roseboom et al. (2010) explained hereafter).

An additional consideration was that the contamination from Galactic (*cirrus*) emission should be minimal. The defining criterion was coverage at mid/far infrared wavelengths not

accessible to Herschel, or where Herschel is relatively inefficient due to its warm primary mirror. Specifically, we required Spitzer MIPS coverage at 24 and 70 μm. At the time of design the one exception to this was the AKARI Deep Field South, which did not have Spitzer coverage but did have coverage at 65, 90, 140, and 160 μm from AKARI. However, this field has since been observed by Spitzer MIPS (Scott et al. 2010). The HeLMS field, which was added in 2011 for studying large-scale structure and the bright end of the number counts, does not have ancillary data from Spitzer. However, being located on the SDSS Stripe 82 region, HeLMS does have ancillary coverage from many other facilities. The cumulative area of all major Herschel-SPIRE extra-galactic Key Program surveys as a function of instrumental noise and for the HerMES fields is shown in Fig. 3.28. As an extra point, since we have already introduced that before, the Spitzer-SERVS and VISTA-VIDEO surveys were approved after HerMES and designed with reference to HerMES. So, almost all the SERVS and VIDEO fields were included in HerMES Level 5 observations. However, the SERVS and VIDEO field in ELAIS S1 was not quite within the original HerMES planned observations, which were only at Level 6. HerMES thus included additional deeper observations covering the SERVS/VIDEO field.

It is striking to compare the Herschel-SPIRE sub-millimetre surveys with previous sub-millimetre surveys. To do this it has been explored the sensitivity of surveys to a canonical galaxy with a modified blackbody spectral energy distribution with emissivity, $\beta = 1.5$, and temperature $T = 35\,\mathrm{K}$. These are shown in Fig. 3.28.

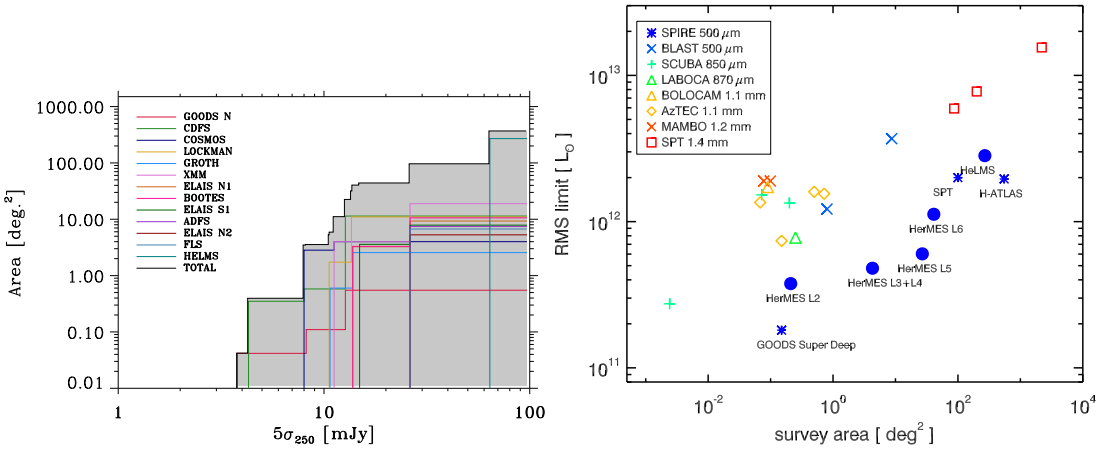

Figure 3.28: **Right**: Cumulative area against 5-σ instrumental noise level at 250μm for the HerMES blank-field surveys with SPIRE (Oliver et al. 2012). The colour-coding breaks this down into individual survey fields. **Left**: Luminosity limit verses redshift for submm surveys to date. The luminosity limit was calculated assuming a modified blackbody of 35K at $z = 2$ (references for the points are reported in Oliver et al. (2012)

The first HerMES observation were carried out on 12th September 2009. This was the first half of HerMES SPIRE observations of Abell 2218, as part of the Herschel Science Demonstration Phase (SDP) that, for HerMES, has been designed to exercise most of the observing modes that were to be used in the full survey, and the SPIRE observations, are described in Oliver et al. (2010). This includes the observations of GOODS-N (Fig. 3.29). The SDP observations finished on 25th October 2009; AORs are available under the proposal ID SDP_soliver_3. The program is now essentially completed, except for the HeLMS wide-area extension that was decided late in the mission being carried out as part of the

Routine Phase (proposal ID KPGT_soliver_1). The current ESA schedule is on `herschel.esac.esa.int/observing/ScheduleReport.html` and the observing log can be followed on `herschel.esac.esa.int/observing/LogReport.html`.

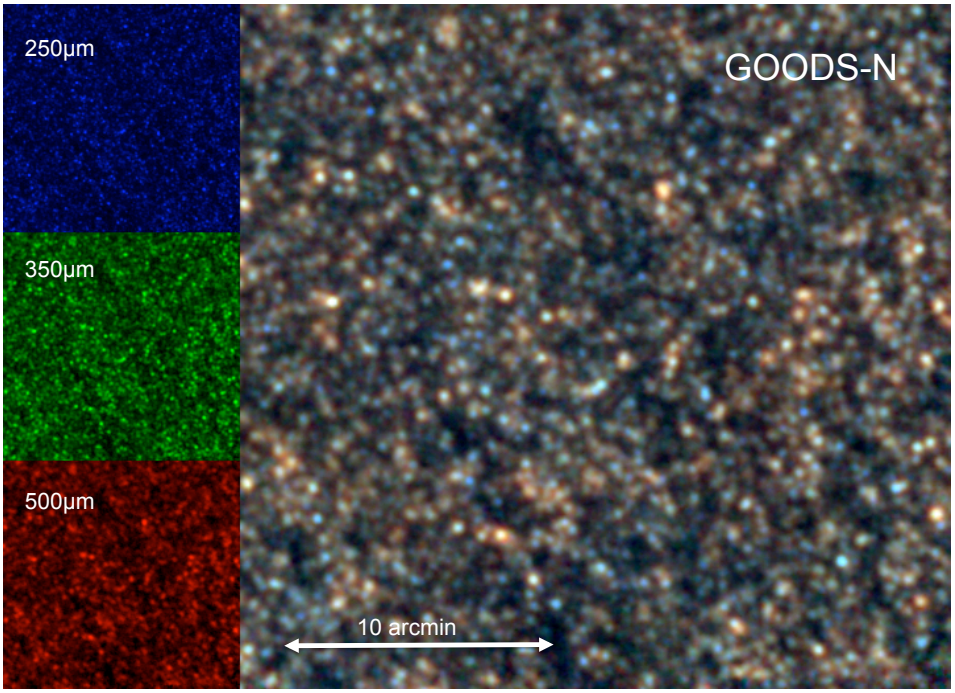

Figure 3.29: Three colour Herschel-SPIRE image of the GOODS-North region. This is a subset of our GOODS-N observation. The left-most panels show the single band images of the cluster, while the central panel shows the three colour image.

The data products are summarized in Tab. 3.10. The Herschel source catalogues from SPIRE and PACS data (SCAT, Smith et al. (2012) and PCAT, Aussel *et al.*, in prep. respectively) consist of the usual independent lists where sources are selected from data at one wavelength without reference to any other. In addition these products will include fluxes estimated for sources from other catalogues (including sources from public Spitzer catalogues). Maps from SPIRE and PACS data (SMAP Levenson et al. (2010) and PMAP respectively) are suitable for extended source analysis, fluctuation analysis etc. The SPCAT product include all Herschel bands (SPIRE+PACS). Upper limits are listed for sources detected in some Herschel bands but not others. The XID products are SPIRE products created by using the original method described in Roseboom et al. (2010) and involve associations with a variety of large homogenous catalogues, including, but not necessarily limited to, public Spitzer catalogues. These are lists with photometry at the positions of known 24μm galaxies and thus are approximately 90 per cent complete (see Sec. 3.8.4 for more details on the method and its products; the multi-wavelength photometry associated to these SPIRE extractions comes from the Spitzer Data Fusion described in Chap. 4 in the shared areas).

Name	Description	Minimum Parameters
SCAT	SPIRE Source Catalogues	Positions, Fluxes, errors, SNRs, etc.
SMAP	SPIRE Maps	Maps of flux, noise and coverage
PCAT	PACS Source Catalogues	Positions, Fluxes, errors, SNRs, etc.
PMAP	PACS maps	Maps of flux, noise and coverage
SPCAT	SPIRE/PACS band-merged catalogues	Positions, Fluxes, errors, SNRs, etc.
CLUS	Catalogues & Maps for Clusters	As above for maps and catalogues
XID	Cross identifications with selected homogenous catalogues at other wavelengths.	Fluxes, errors, SNRs, positions, positional offsets

Table 3.10: Deliverable Data Products.

The first data release (DR1)

The first data release (DR1) has been published on the 3^{rd} of April 2012 and is published through the Herschel Database in Marseille, HeDaM [16]. All the consortium worked hardly to make it possible within the scheduled time-line and part of the projects described in this thesis have been devoted to this issue. This release included Herschel SPIRE sky maps and object catalogues. The maps were made using 250, 350 and 500 μm filters. These sub-millimeter wavebands were not used at all before the Herschel Mission. The maps cover $\sim$74 deg^2 of the sky, i.e. a volume of 6.6e8 (Mpc)^3 at $z < 1.5$ (and many of the galaxies that we see are expected to be at $z > 1.5$), greater than the SDSS which maps a volume of 3.5e8 (Mpc)^3 for $z < 0.17$. The maps range in depth but are mostly at or below the SPIRE confusion limit and so provide a very high quality view of the sub-millimeter sky, limited primarily by the diameter of the Herschel mirror. The catalogues extracted from these maps include over 50,000 entries, representing over 17,000 galaxies. Extensive simulations have demonstrated that these catalogues are very high quality and $\sim$ 90% of the point-like galaxies having very reliable positions and flux measurements. This data release follows two early data releases (July 2010 and September 2011) which were limited to the brightest catalogued sources over a smaller range of fields. The data releasing process will end with the end of the mission and the Second Data Release (DR2).

The SPIRE data used to derive the LFs described in Chap. 6 and Chap. 7 are based on the XID catalogues. The method has been developed during the SDP and improved along with the development of the mission. A first description of the method has been published in Roseboom et al. (2010) and an updated version of the method will be published in Roseboom *et al.*, in prep.; the data products we use in our work come from this last incarnation of the method that has been improved along with new data a new results acquired within HerMES. In the following Section (Sec. 3.8.4) we present a summary of the main characteristics of the method.

[16] `http://hedam.oamp.fr/HerMES/`

3.8.4 HerMES XID Method & Data Products

It is very difficult to measure accurate flux densities for sources in astronomical images dominated by confusion noise. Great advances in the sensitivity of instruments at long wavelengths (FIR/SUBMM/Radio) has meant that the blended signal from numerous, unresolved, faint sources now form a non-negligible fraction of the observed telescope background. Hence confusion noise, i.e. fluctuations in this background, is now the dominant source of noise in deep imaging. Confusion acts to increase the positional uncertainty of sources dramatically (e.g. Hogg (2001)), making cross-identifications with other wavelengths problematic; besides, correlations between the confusing background and sources above the confusion limit result in, at best, flux boosting of detected sources above the confusion limit and, at worst, complex blends of correlated confusion noise, resulting in spurious sources (Scheuer (1957), Condon (1974)).

In the recent history we can find two main ways to source extraction in this complicated conditions: traditional source detection methods, combined with probabilistic approaches for flux boosting and source identification have been used to good effect on sub-mm surveys performed with SCUBA (e.g. Ivison et al. (2007)); a more statistical approach, that choose to ignore individual sources and look at the aggregate properties of sources via either stacking (e.g. Dole et al. (2006b)), or the map statistics themselves via the pixel intensity distribution, the so-called $P(D)$ (e.g. Patanchon et al. (2009)).

Both approaches have advantages and disadvantages. Working with individual sources allows the true variation of sub-mm galaxy properties and their correlations with other observables to be properly investigated. However, finding multi-wavelength identifications for individual sub-mm sources is usually difficult, and generally reliable identifications can be found for only a fraction of sources (Roseboom et al. (2009)). Statistical approaches have the advantage of using all the available data, and hence provide greater precision in the parameters of interest. However, interpretation of these statistically-derived quantities is sometimes complicated, and highly dependent on the choice of parameterisation.

The Method

The XID method is based on a combination of the methods previously described. By using a linear inversion technique to fit for the flux density of all known sources simultaneously, the ability to work on individual sources is retained, while the information in the map itself can be used to distinguish the contributions from each source. This approach has been used by Scott et al. (2002) to fit the flux densities of SCUBA 850 μm sources in the 8 mJy survey; Magnelli et al. (2009) to fit the Spitzer 24μm flux density of IRAC detected sources in GOODS-N; and also by Béthermin et al. (2010) to fit the BLAST for 24μm detected sources in the Extended Chandra Deep Field South (ECDFS). The key to this approach is its simplicity; the only assumptions are that all sources are unresolved by the telescope, and that the positions of all sources are known. If these assumptions hold, then in the limit of infinite signal-to-noise ratio in the image the resulting flux density measurements would be perfect, irrespective of source density.

The method performs cross-identification and source photometry in map-space so as to minimize source-blending effects. It make use of a combination of linear inversion and model selection techniques to produce reliable cross-identification catalogues based on Spitzer MIPS

24 μm source positions.

One of the key features of HerMES is that all of the planned survey fields contain existing *Spitzer* data from a range of legacy surveys. More importantly the tiered nature of HerMES is well matched to the variable quality of the Spitzer data, in particular the MIPS 24 μm observations. This is highlighted by comparing the S_{250} and S_{24} sensitivities. Using a compilation of pre-Herschel empirical models (e.g. Fernandez-Conde et al. (2008); Le Borgne et al. (2009); Franceschini et al. (2010)) Roseboom et al. (2010) estimate that 0.4–24 per cent of $S_{250} > 1$ mJy sources have $S_{250}/S_{24} > 100$, with the majority of these (up to 70 per cent) lying in the range $1.2 < z < 1.6$ where the 24 μm band is coincident with the 10 μm silicate feature present in strong absorption in typical starburst galaxies.

It is also clear from existing measurements of the cosmic IR background (CIRB) from BLAST (Devlin et al. 2009) that sources already detected at 24 μm with Spitzer are the dominant contributor at these wavelengths. In particular, Pascale et al. (2009) shows that greater than 90 per cent of the CIRB at BLAST/SPIRE wavelengths can be accounted for by 24 μm sources with $S_{24} \geq 100$ μJy. Hence we can be confident that using the 24 μm source lists as a model for the positions of sources in the SPIRE maps is appropriate. It is also worth considering that in the deepest fields (i.e. GOODS-N) the source density of 24 μm sources is $\sim$ 24,000 deg^{-2}, or $\sim$ 2 SPIRE 250 μm beam elements per source. Thus even recovering the SPIRE fluxes for the detected 24 μm sources, involves going significantly beyond the confusion limit.

It should be noted that this algorithm has been developed in parallel with the other data reduction techniques (i.e. Smith et al. (2012); Levenson et al. (2010)) for use in the first SDP science papers from HerMES. Thus while this approach has proven to give the best performance under testing, it is clear that several aspects could be easily improved. To investigate possible systematics in XID photometry, it has been compared to the generated using a combination of source detection and extraction, via Sussextractor and p-statistic methods to match the resulting source lists with existing 24 μm catalogues.

Sussextractor source lists are provided for each SPIRE band by SCAT (Smith et al. 2012). These source lists contain all SPIRE sources detected in the maps at a significance of greater than 3σ. The monochromatic SPIRE source lists are then matched to the same 24 μm catalogues used as an input to XID algorithm. The matching is performed by finding potential counterparts within a search radius of 10 arcsec, 14 arcsec, and 20 arcsec for the 250, 350 and 500 μm bands respectively. For each of these potential IDs has been calculated the p-statistic. The uncertainty of the SPIRE position is calculated using Equation B8 of Ivison et al. (2007). All IDs with $p < 0.1$ are considered. A complete sample is constructed by taking the best ID with $p < 0.1$ for each SPIRE source. Alternatively a 'clean' sample is constructed by taking only those cases where the separation is less than 0.6×FWHM, and there is only one potential ID with $p < 0.1$.

Fig. 3.30 compares the flux density estimates for sources in the LH-SWIRE, LH-North fields and GOODS-N from the XID catalogues and the SCAT+p-stat listings. Only those sources which are in common and are found at greater than 5σ in both catalogues are presented. The FLS field is omitted for clarity. While there is a large scatter between the two estimates for all sources, a good agreement can be seen for the 'clean' ones. The bulk of the sources which are discrepant between the two catalogues can be found above the one-to-one line in Fig. 3.30, i.e. $S_{\rm XID} < S_{\rm SCAT}$. This is a natural consequence of the XID algorithm considering all known sources simultaneously, and thus deblending confused cases into their individual 24

μm detected components.

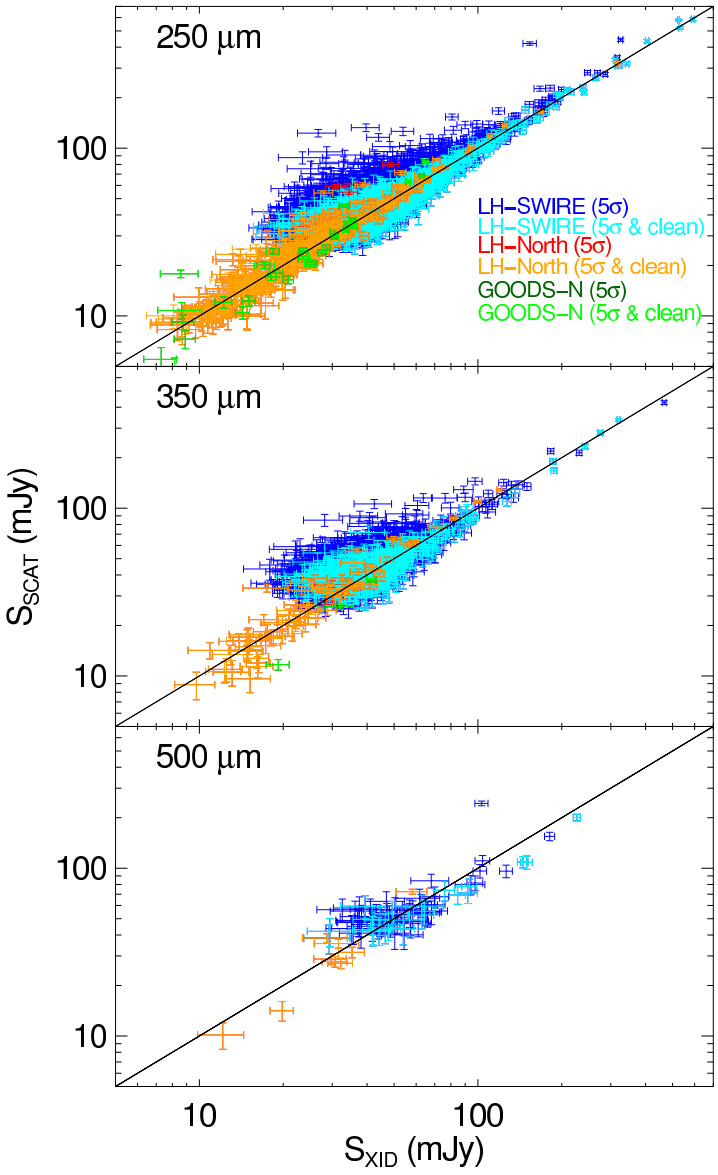

Figure 3.30: Comparison of flux densities from XID catalogue to those from the SCAT SussExtractor-derived source catalogues in the Lockman Hole SWIRE, Lockman Hole North and GOODS-N fields. SCAT sources are matched to the 24 μm sources via the p-statistic. Sources present in both catalogues at 5σ are presented, as well as a "clean" sample where $p < 0.1$, seperation$< 0.6\times$FWHM$_{\rm SPIRE}$, and there are no alternative IDs with $p < 0.1$. XID fluxes are also required to have $\rho < 0.8$ and $\chi^2 < 5$, as described in the text. Good agreement can be seen between the XID and SCAT flux densities for "clean" sources. This suggests that any discrepencies between SCAT and XID are solely due to issues with source blending.The figure is from Roseboom et al. (2010).

In order to achieve consistency between the 3 SPIRE bands the method only carry out the model selection stage of the algorithm for the 250 μm band. An alternative approach, where all three SPIRE bands are treated independently, was initially considered, but found to give poor results. In particular the increase in beam size from 250 to 350 to 500 μm results in

a decreased ability to deblend at long wavelengths and a preference to retain fewer sources. This naturally leads to inconsistencies between the measurements in the different bands. Thus it was decided to use the 250 μm results to determine which sources were indeed present at the SPIRE bands. It is possible to use the pre-Herschel mock catalogues of Fernandez-Conde et al. (2008) to estimate the number of 350 and 500 μm sources missed by this requirement. Assuming a uniform sensitivity across the SPIRE bands, and the depth of the deepest field considered here (GOODS-N; 4 mJy), the number of sources the additional incompleteness due to requiring a detection at 250 μm is only 0.5 per cent. However it is clear that this estimate is highly dependent on the range of SEDs used in the Fernandez-Conde et al. models.

While the input source list is defined by the 24 μm flux density limits, the XID method use source positions from Spitzer IRAC 3.6 μm imaging where there is deep co-incident data and previous associations between the two data sets have been made. This occurs in all of our fields, with the exception of LH-North, and the wider area of HDF-N. The IRAC positional accuracy is typically ~ 0.2 arcsec (as opposed to ~ 1 arcsec for 24 μm) and hence using these eliminates any error in the flux density solutions introduced by astrometric errors.

The Data Products

The resulting HerMES XID catalogues contain the complete input 24 μm source catalogue, as well as any previously associated data sets at other wavelengths (see Chap. 4), as well as the best estimate of the SPIRE flux density for each 24 μm source passing our input selection criteria.

In addition to the flux density and error in each band for each input 24 μm source, the SPIRE component of the XID catalogues contain a number of extra columns describing diagnostics of the fitting process and local source confusion. These extra measures include:

- Maximum absolute value of the Pearson correlation coefficient calculated on the covariance matrix of the flux density solution (hereafter refered to as ρ).
- χ^2 of the source solution in the neighbougrhood of source (7 pixel radius).
- The background level estimated in the fitting.
- The number of sources in the segment containing this source.
- The ID number of the segment.
- The PRF-smoothed flux density at the position of this source, ignoring contributions from neighbouring sources and the background.
- The number of 24 μm sources within a radius of the FWHM with greater than 50 per cent of the flux density of this source.
- The 'purity' of the SPIRE flux density, based on the ratio of this source's 24 μm flux density to the 24 μm flux density smoothed with the SPIRE PRF at this position (see Brisbin et al. (2010)).

The reason for including these extra columns is to enable samples of varying quality to be extracted from the XID catalogues based on differing scientific requirements. From an

early assessment of the XID algorithm performance the recommended quality cuts for typical science applications were:

- $S_\lambda > 5 \times \Delta S_\lambda$;
- $\rho_\lambda < 0.8$;
- $\chi^2_\lambda < 5$.

Given the more detailed analysis presented below, these cuts have proven to return very reliable samples, although possibly at the expense of completeness. Hence they represent fairly conservative guidelines for the use of the XID catalogues. In the 2012 updated version of the XID catalogues the refined extraction technique and the increased quality of the observations have brought to a change in this criterion of recommended quality cuts to:

- $S_\lambda > 5 \times \Delta S_\lambda$;
- $\chi^2_\lambda < 5$.

where ΔS_λ is the quantity referring to the Signal to Noise Ratio (SNR) at a given selection λ including confusion, thus named "Total" SNR_λ or SNRT_λ. These are again conservative guidelines that could be changed according to each specific requirements.

To quantify the effectiveness of these new techniques they consider simulated SPIRE images. They consider two simulated cases: a 'deep' map, where $\sigma_{\mathrm{conf}} \gg \sigma_{\mathrm{inst}}$; and a 'shallow' map where $\sigma_{\mathrm{inst}} \geq \sigma_{\mathrm{conf}}$. In each case they simulate a $2.2^o \times 2.2^o$ patch of sky in all three SPIRE bands, taking the mock catalogues of Fernandez-Conde et al. (2008) (FC08) as an input. While many mock catalogues exist at these wavelengths, the FC08 mocks were found to give the best match to the observed confusion noise and source colours in real SPIRE data.

Simulated maps are produced from the positions and flux densities quoted in the mock catalogues by first making noise-free maps in each band, using the known SPIRE PRF parameters. Secondly, Gaussian noise and a flat background are added.

The performance of source extraction and cross-identification methods are typically characterised by two metrics: the completeness, i.e. the fraction of sources recovered at a given flux density; and the reliability or mis-ID rate. While the notion of completeness translates well to the methods presented here, reliability is not an intuitively useful quantity when performing XIDs in the map-space. We know (or assume) that all of our 24 μm sources are reliable; the aim is solely to accurately measure their flux densities at other wavelengths. Thus the second metric by which XID methods is judged is flux density accuracy. While it is useful to assess the completeness and flux density accuracy of the method on totally artificial maps, it is also possible to calculate these metrics for the real data by injection of mock sources into the observed maps. This has the advantage of reproducing the true noise properties of the data, as well as highlighting the confusion noise in the presence of angular clustering.

As maps are already heavily affected by confusion, it is possible to inject one source at a time into the map, and then run the XID source extraction algorithm, taking the input position of the mock source and the neighbouring 24 μm sources into account. Using SDP data, they inject mock sources with flux densities in the range 3–200 mJy at random positions. To maintain consistency with the properties of the real 24 μm input catalogues, test positions within 3 arcsec of an existing 24 μm source are also excluded, as was done with the fully

articifial simulations. As a result the total number of test positions is $\sim$ 3000–5000 per field, with 300–500 per test flux density. Fig. 3.31 shows the completeness and flux accuracy determined by this method.

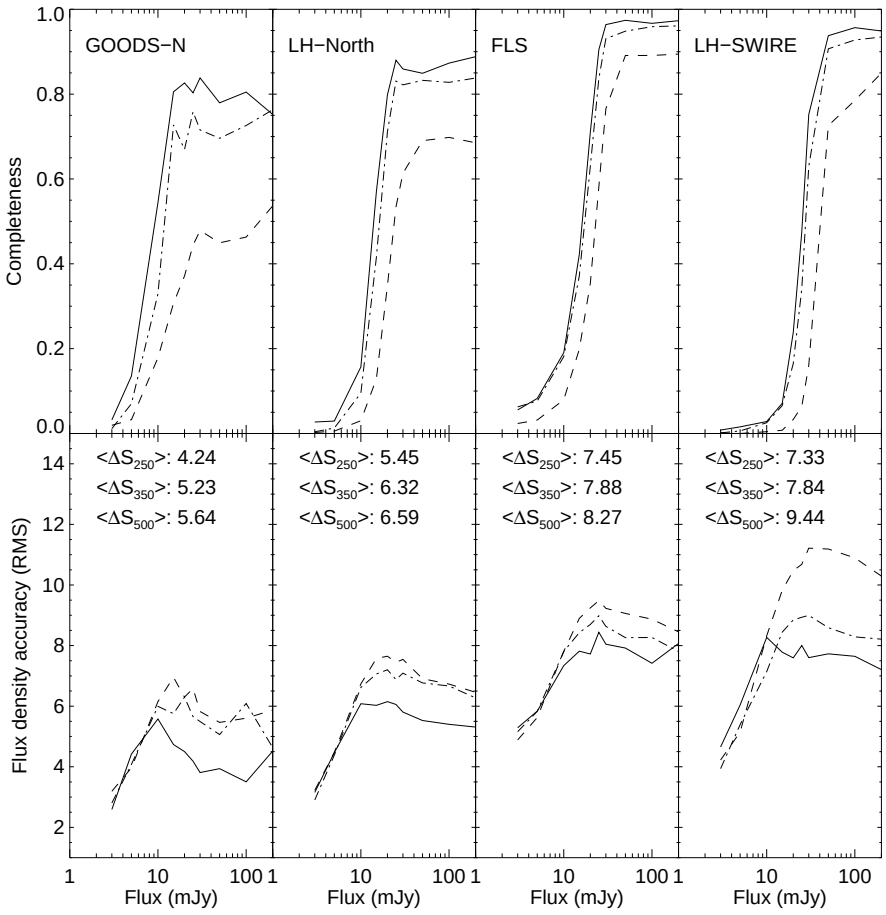

Figure 3.31: **Top**: completeness determined by injection of mock sources into SPIRE observed maps and defined as the ratio of the number of sources recovered at $> 5\sigma$ and $\rho < 0.8$ to the number of input positions. **Bottom**: flux density accuracy determined by injection of mock sources into SPIRE observed maps and defined as the RMS of the input-output flux density. In calculating the recovered flux density accuracy all input positions recovered with $\rho < 0.8$ are considered. Sources are injected one at a time so as to avoid increasing the source confusion. In each panel the results for the 250 μm (solid line), 350 μm (dot-dashed line) and 500 μm (dashed line) bands are shown. Mean flux density error for each band is shown in the top right corner of each of the lower panels. The figure is from Roseboom et al. (2010).

It can be seen that the completeness never reaches 100 per cent in any field. The values rise sharply from faint flux densities and then plateau at a quasi-constant value above a certain flux density level. This is due to the effect of the $\rho < 0.8$ criteria. Somewhat counter-intuitively, this is a bigger problem in the fields with deeper SPIRE/MIPS 24 μm data. The reason for this is simple; the input source density is much higher in the deep fields and, as there is no prior assumption on the SPIRE flux density, this affects all flux densities equally. If the $\rho < 0.8$ criteria is removed the residual $\sim$ 20–50 per cent incompleteness in the deep fields is recovered, but at the expense of flux density accuracy. For sources with $\rho < 0.8$ in the

GOODS-N field the 1σ flux density error is 4.24, 5.23 and 5.64 mJy for the 250, 350 and 500 μm bands, respectively. For sources with $\rho > 0.8$ the comparable values are 6.3, 5.9, and 6.9 mJy, an increase of $\sim$ 10–50 per cent.

As a final cross-check of the completeness estimates it has been compare the raw differential number density of sources found in both the XID and SCAT+p-stat catalogues to the best estimates of the source densities from (Oliver et al. 2010). Fig. 3.32 shows the differential number density of sources in our XID and SCAT+p-stat catalogues, in the LH-SWIRE, FLS, and LH-North fields. GOODS-N observations are excluded as the number of sources detected is too small for this comparison to be useful. Encouragingly, at bright flux densities (i.e. > 50 mJy), both the XID and SCAT+p-stat catalogues show reasonable agreement with Oliver et al. (2010), although cosmic variance introduces a large scatter at the highest flux densities. Both the XID and SCAT+p-stat are seen to be incomplete at faint flux densities, although in each band the XID catalogue is significantly more complete at flux densities $\sim$ 20–30 mJy. Taking the Oliver et al. result to represent the total number of sources, Tab. 3.11 quotes the XID and SCAT+p-stat catalogue 50 per cent completeness levels. These values are in good overall agreement with the completeness estimates found via simulations and source injection.

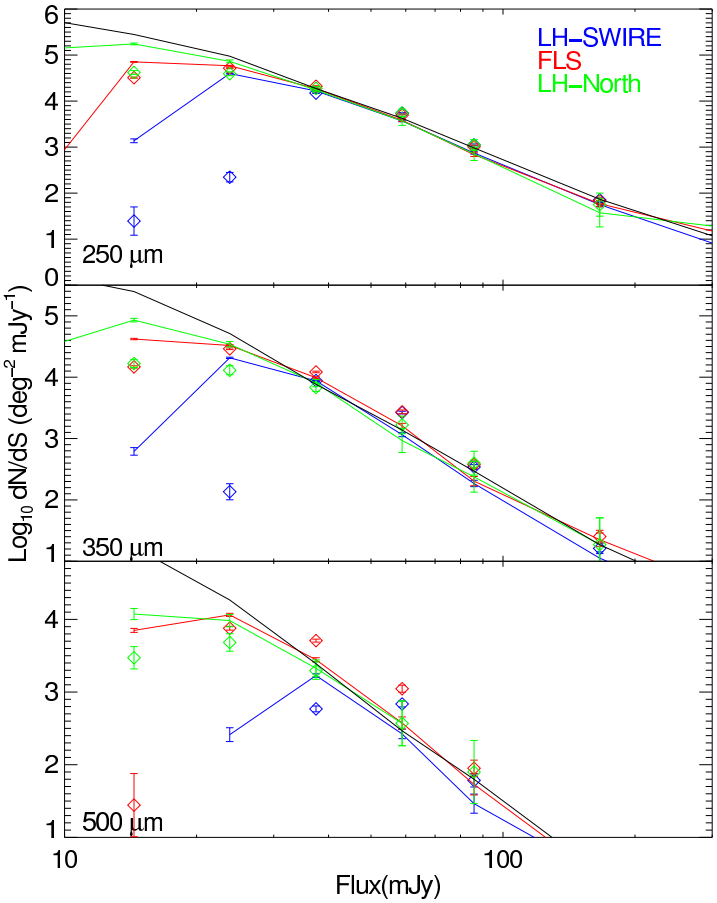

Figure 3.32: Differential number density of sources in the XID catalogues and SCAT v3 release catalogues. The black line is the current best estimate of true source density from Oliver et al. (2010). The solid lines are XID, open symbols are SCAT in all panels. The figure is from Roseboom et al. (2010).

Using the 24 μm prior input list can introduce an additional incompletness which is strongly dependant on the relative depth of the existing 24 μm data to SPIRE data. From the combination of deep SPIRE and Spitzer 24 μm observations in GOODS-N they estimate an incompleteness due to the 24 μm limit in the other fields of $\sim$ 20 per cent at 250 μm, increasing to $\sim$ 40 per cent at 500 μm. However this incompleteness is dominated by the faintest SPIRE sources (i.e. less than 30–40 mJy), and we can be confident that the XID catalogues are complete at bright fluxes.

	S_{250}[50 per cent] (mJy)			S_{350}[50 per cent] (mJy)			S_{500}[50 per cent] (mJy)		
	Source Injection		Counts	Source Injection		Counts	Source Injection		Counts
Field	XID	XID	SCAT+p-stat	XID	XID	SCAT+p-stat	XID	XID	SCAT+p-stat
GOODS-N	9.5	–	–	12.1	–	–	n/a	–	–
LH-North	14.1	13.3	26.7	16.4	23.8	30.5	24.1	23.6	31.5
FLS	16.4	21.6	23.3	17.5	21.6	23.3	23.2	22.9	25.5
LH-SWIRE	25.4	27.	36.4	27.8	26.7	35.6	42.	36.4	44.4

Table 3.11: Completeness estimates (50 per cent) for XID and SCAT+p-stat catalogues for real observations of SDP fields. Completeness is estimated via both injection of sources into the map, and by comparing the number density of sources in the resulting catalogues to the best estimate of the true source density from Oliver et al. (2010).

Set	Level	Target	Mode	$N_{\rm AOR}$	T [hr]	$N_{\rm rep}$	$N_{\rm samp}$	l_1 [']	l_2 [']	θ [deg]	$\Omega_{\rm nom}$ [deg^2]	$\Omega_{\rm max}$ [deg^2]	$\Omega_{\rm good}$ [deg^2]	DR
1	CD	Abell 2218	Sp. Nom.	2	9.29	100	1118	4	4	84		0.14	0.10	SDP
2	CD	Abell 1689	Sp. Nom.	8	1.97	48	235	4	4	18		0.11	0.08	
3	CD	MS0451.6-0305	Sp. Nom.	8	1.97	48	235	4	4	5		0.11	0.08	DR1
4	CS	RXJ13475-1145	Sp. Nom.	8	1.97	48	234	4	4	17		0.11	0.08	
5	CS	Abell 1835	Sp. Nom.	8	1.97	48	236	4	4	16		0.11	0.08	
6	CS	Abell 2390	Sp. Nom.	8	1.97	48	235	4	4	81		0.11	0.08	
7	CS	Abell 2219	Sp. Nom.	8	1.97	48	234	4	4	66		0.11	0.08	DR1
8	CS	Abell 370	Sp. Nom.	8	1.97	48	233	4	4	70		0.11	0.08	
9	CS	MS1358+62	Sp. Nom.	8	1.97	48	235	4	4	76		0.11	0.08	
10	CS	Cl0024+16	Sp. Nom.	8	1.97	48	235	4	4	61		0.11	0.08	
11	CH	MS1054.4-0321	Sp. Nom.	8	2.18	16	131	15	10	22		0.24	0.16	
12	CH	RXJ0152.7-1357	Sp. Nom.	8	2.18	16	131	15	10	165		0.24	0.16	
13	L1	GOODS-S	Sp. Nom.	76	20.22	76	730	20	20	14		0.51	0.35	
22	L2	COSMOS	Sp. Nom.	24	50.13	24	388	85	85	70		3.49	2.82	
14	L2	GOODS-N	Sp. Nom.	1	13.51	30	416	30	30	42		0.64	0.55	SDP
15	L2	ECDFS	Sp. Nom.	19	8.78	19	232	30	30	44		0.79	0.58	DR1
17	L3	Groth Strip	Sp. Nom.	7	3.54	7	85	67	10	130		0.82	0.60	DR1
18	L3	Lockman-East ROSAT	Sp. Nom.	7	3.2	7	87	30	30	77		0.77	0.57	
18B	L3	Lockman-East Spitzer	Sp. Nom.	4	4.53	4	32	80	40	149		1.78	1.40	
19	L3	Lockman-North	Sp. Nom.	1	3.91	7	104	35	35	1		0.74	0.65	SDP
23	L4	UDS	Sp. Nom.	7	10.54	7	110	66	66	20		2.46	2.02	
24	L4	VVDS	Sp. Nom.	7	10.39	7	110	66	66	21		2.46	2.02	
22B	L5	COSMOS HerMES	Sp. Nom.	8	25.20	8	128	110	110	70		5.04	4.38	
27	L5	CDFS SWIRE	Sp. Fast	10	41.72	20	81	190	150	99		12.18	11.39	
28	L5	Lockman SWIRE	Sp. Fast	2	13.51	2	16	218	218	2		18.2	17.37	SDP
28B	L5	Lockman SWIRE	Sp. Fast	8	41.26	8	58	220	180	50		15.26	7.63	
42	L7	HeLMS	Sp. Fast	11(10)	103.4	2		1560	750	15	270			
20	L3	Lockman-North	PACS	12	13.96	11		30	30	42	0.25			SDP
20B	L3	Lockman-North	PACS	20	20.89	20		30	30	42	0.25			
21	L3	UDS HerMES	PACS	25	25.93	14		30	30	0	0.25			
25	L4	UDS	PACS	12	40.19	7		57	57	0	0.9			
29	L5	EGS HerMES	Parallel	7	22.68	7	93	150	40	131		3.50	2.67	
30	L5	Bootes HerMES	Parallel	5	20.33	5	70	80	80	0		4.21	3.25	DR1
31	L5	ELAIS N1 HerMES	Parallel	5	20.82	5	72	95	95	38		3.74	3.25	DR1
32	L5	XMM VIDEO1	Parallel	4	13.44	4	65	106	75	107		3.20	2.72	
32B	L5	XMM VIDEO2	Parallel	4	8.88	4	53	106	44	107		2.12	1.74	
32C	L5	XMM VIDEO3	Parallel	4	13.44	4	53	106	75	107		3.19	2.73	
33	L5	CDFS SWIRE	Parallel	4	50.42	4	57	204	170	101		11.87	10.89	
34	L5	Lockman SWIRE	Parallel	4(2)	71.22	4		215	215	154		17.86	16.08	
39B	L5	ELAIS S1 VIDEO	Parallel	4	17.72	4	56	138	80	87		4.42	3.72	
35	L6	ELAIS N1 SWIRE	Parallel	2	28.0	2	28	207	192	55		13.37	12.28	
36	L6	XMM-LSS SWIRE	Parallel	6	45.58	2	29	180	180	82		21.62	18.87	DR1
37	L6	Bootes NDWFS	Parallel	4	27.99	2	30	243	80	145		11.3	10.57	DR1
38	L6	ADFS	Parallel	2	18.11	2	28	190	122	80		8.58	7.47	DR1
39	L6	ELAIS S1 SWIRE	Parallel	2	17.9	2	28	140	81	91		8.63	7.86	
40	L6	FLS	Parallel	2	17.1	2	29	160	138	5		7.31	6.71	SDP
41	L6	ELAIS N2 SWIRE	Parallel	2	17.1	2	26	177	119	147		9.06	7.80	

Table 3.12: Summary of the HerMES observations. The full set of Astronomical Observation Requests (AORs) are available through ESA's Herschel Archive. We have grouped $N_{\rm AOR}$ observations of the same field at the same level made with the same mode and areal size into a 'set' (the number of AORs still to be scheduled after 2011 Dec 21 is indicated in parentheses). The first five columns give: the set identification number; the design level; the target name, the Herschel observing mode and the number of AORs in the set. T is the time used or allocated for this set. $N_{\rm rep}$ is the total number of repeats of the observing mode in the set. All our SPIRE nominal (30" s^{-1}) and fast mode (60" s^{-1}) (Sp. Nom. and Sp. Fast) observations include a scan in the nominal and orthogonal direction, so 1 repeat is 2 scans). For SPIRE observations that have been executed $N_{\rm samp}$ is the median number of bolometer samples per pixel in the 250 μmmap (6" $\times$ 6" pixels). The error per pixel in our SPIRE maps as processed by the standard HIPE pipeline are $\sigma^2_{250} = \sigma^2_0/N_{\rm samp}$ with $\sigma^2_0 = 896 \pm 11, 1554 \pm 27$ and ~ 1440 mJy^2 beam^{-2} for Parallel, Sp. Nom. and Sp. Fast modes respectively. l_1, l_2 are sides of a rectangle with near homogenous coverage. θ is the roll angle with short-axis of that rectangle measured East of North. For SPIRE observations that have been executed $\Omega_{\rm max}$ is the total area of pixels with any 250 μm coverage and $\Omega_{\rm good}$ is the area of pixels where the number of bolometer samples per pixel in the 250 μm map is greater than $N_{\rm samp}/2$. For PACS fields or unobserved SPIRE fields $\Omega_{\rm nom}$ gives the nominal area of region. The final column indicates which observations are included in our data releases; observations marked SDP were released in our Early Data Release, observations marked SDP or DR1 have been released through `http://hedam.oamp.fr` as part of HerMES DR1. Set numbers #16 and #26 were removed from the programme.

Fields	Area		Cummulative	Observations		5-σ noise level (for band in μm)			
	Nominal	Extra		PACS	SPIRE	110	160	250	350	500
	[deg^2]					[mJy]				
Abell 2218	0.0050	0.0050	0.1	P	1	4.1	7.9	6.4	5.3	7.6
Abell 1689	0.0050	0.0050	0.11	P	2	3.6	6.9	9.2	7.7	11.0
8 Targets	0.04	0.04	0.15	P	3-10	5.7	10.9	9.2	7.7	11.0
2 Targets	0.03	0.03	0.18	P	11-12			13.9	11.6	16.7
Various	0.18	0.18	0.36	E	E	6.1	11.7	14.2	11.9	17.1
GOODS-N	0.042	0.042	0.04	G,P	G,14	2.2	4.1	3.8	3.1	4.5
GOODS-S	0.11	0.087	0.13	G,P,33	13,15,27,33	2.1	2.9	4.3	3.6	5.2
GOODS-S	0.012	0.012	0.14	G,P,33	13,15,27,33	1.1	2.1	4.6	3.8	5.5
GOODS-S	0.018	0.0060	0.15	G,P,33	13,15,27,33	1.6	3.0	4.6	3.8	5.5
GOODS-S	0.023	0.0060	0.15	G,P,33	13,15,27,33	2.0	3.8	4.6	3.8	5.5
COSMOS	2.0	2.0	2.15	P	22,22B	7.7	14.7	8.0	6.6	9.5
ECDFS	0.25	0.14	2.29	P,33	15,27,33	7.6	14.5	8.0	6.6	9.6
GOODS-N	0.25	0.208	2.5	P	14	4.7	8.9	8.2	6.8	9.9
Lockman-East	0.25	0.25	2.75	P	18,18B,28B,34,28	6.5	12.3	9.6	7.9	11.5
Lockman-North	0.25	0.25	3.0	20,20B,34	19,28B,34,28	7.4	14.1	10.6	8.8	12.7
Groth Strip	0.25	0.25	3.25	P,29	17,29	7.1	13.6	10.7	8.9	12.8
UDS HerMES	0.25	0.25	3.5	21,25,32,36	23,25,32,36	6.8	12.9	11.2	9.3	13.4
UDS	0.7	0.7	4.2	25,32,36	25,32,36	11.2	21.4	11.2	9.3	13.4
VVDS	2.0	2.0	6.2	25,32C,36	25,32C,36	28.8	54.9	11.2	9.3	13.4
CDFS SWIRE	11.4	11.1	17.3	33	27,33	31.5	60.2	12.7	10.5	15.2
Lockman SWIRE	16.1	15.6	32.9	34	28,28B	35.3	67.3	13.6	11.2	16.2
EGS HerMES	2.7	2.5	35.4	29	29	26.6	50.8	13.8	11.3	16.4
Böotes HerMES	3.3	3.3	38.6	30,37	30,37	26.6	50.8	13.8	11.3	16.4
ELAIS N1 HerMES	3.3	3.3	41.9	31,35	31,35	26.6	50.8	13.8	11.3	16.4
ELAIS S1 VIDEO	3.7	3.7	45.6	39B,39	39B,39	28.8	54.9	14.9	12.2	17.8
XMM-LSS VIDEO	7.7	5.0	50.6	32,32B,32C,36	32,32B,32C,36	28.8	54.9	14.9	12.2	17.8
COSMOS Hermes	4.4	2.4	53.0		22B			15.9	13.3	19.1
ELAIS N2 SWIRE	7.9	7.9	60.9	41	41	49.9	95.1	25.8	21.2	30.8
FLS	6.7	6.7	67.6	40	40	49.9	95.1	25.8	21.2	30.8
ADFS	7.5	7.5	75.1	38	38	49.9	95.1	25.8	21.2	30.8
ELAIS S1 SWIRE	7.9	4.2	79.2	39	39	49.9	95.1	25.8	21.2	30.8
ELAIS N1 SWIRE	12.3	9.1	88.3	35	35	49.9	95.1	25.8	21.2	30.8
Böotes NDWFS	10.6	7.3	95.6	37	37	49.9	95.1	25.8	21.2	30.8
XMM-LSS SWIRE	18.9	15.0	110.6	36	36	49.9	95.1	25.8	21.2	30.8
Various	570.0	570.0	681.0	A	A	86.3	164.0	25.8	21.2	30.8
SPT	100.0	100.0	781.0		S			44.5	37.1	53.0
HeLMS	270.0	270.0	1051.0		42			45.3	37.5	54.1
								64.0	53.0	76.5

Table 3.13: HerMES survey with sensitivities in the context of other survey programmes being undertaken by Herschel. The "observations" columns refer to the AOR set numbers of Tab. 3.12 for HerMES or for other Key Programmes we use: "E" for Egami cluster programme, "G" for GOODS-H, "P" for PEP, "A" for H-ATLAS and "S" for SPT (see Tab. 6 in Oliver et al. (2012) for a summary of the Herschel programs). The sensitivities are estimated consistently using HSPOT v5.1.1. These are single pixel sensitivities and ignore the benefits of matched filters, particularly for unconfused fields, e.g. H-ATLAS quote empirical 5-σ sensitivities of 105, 140, 32, 36, 45 mJy for the five wavelengths so the sensitivities in this Table should be scaled by 1.22, 0.85, 0.72, 0.97, 0.85 to obtain a consistent comparison with H-ATLAS. The sensitivity of HerMES observations have been calculated including data from shallower tiers as described in the text. Other surveys are treated independently. Cluster observations are listed before blank fields. The fields are ordered in increasing 250 μm flux limit then right ascension. The area is defined by the PACS observations for Levels 1-4 (above the second horizontal line), otherwise we use $\Omega_{\rm good}$ from Tab. 3.12 or $\Omega_{\rm nom}$ for HeLMS. We tabulate three areas: the nominal area for each field; the 'doughnut' area which excludes any deeper sub-fields within; and the cumulative area of all fields higher in the table. The 5-σ confusion noise (after 5σ cut) from Nguyen et al. (2010) is 24.0, 27.5, 30.5 mJy (at 250, 350 and 500 μm), approximately the Level 6 depth. GOODS-S also has PACS data not listed here at 70 μm over 0.11 deg^2 to a 5-σ depth of 1.9 mJy.

Chapter 4

Spitzer Data Fusion : A Spitzer-Selected Multi-Wavelength Catalogue

Notwithstanding the recent progress of astronomical archives, databases and virtual observatory tools, multi-wavelength catalogues readily useable for interpretative work are difficult to come by for a variety of reasons, generally related with the different resolution and sensitivity achievable at different wavelengths and required to study galaxies at different redshifts.

In this chapter we present Version 1.0 of the Spitzer Data Fusion combining datasets extending from the FUV to the FIR and covering the most popular multi-wavelength extragalactic survey wide-area fields, while in the next chapters we will discuss its exploitation for SED template fitting and for the determination of the luminosity function of sources detected in Spitzer's MIPS and Herschel's SPIRE bands.

4.1 Spitzer Data

In the course of its 2003-2009 cold mission the Spitzer mission (Werner et al. 2004) has revolutionized galaxy formation and evolution studies, providing nearly-optical-quality imaging over wide-areas at MIR wavelengths in its IRAC (Fazio et al. 2004) 3.6, 4.5, 5.8, 8.0 μm and MIPS (Rieke et al. 2004) 24 μm channels as well as lower-resolution imaging in its MIPS 70 and 160 μm channels. This wealth of wide-area MIR and FIR imaging data, combined with a number of imaging surveys at FUV to NIR wavelengths, has allowed the astronomical community to greatly improve modeling constraints on the SM and SFR of galaxies up to high redshift. in producing the Spitzer Data Fusion we attempted to provide an homogeneous database of FUV to FIR photometry for Spitzer-selected sources covering the most popular wide-area extragalactic fields observed by Spitzer.

4.1.1 SWIRE Fields

The Spitzer Wide-area Infrared Extragalactic (SWIRE) survey (Lonsdale et al. 2003) covers 50 deg^2 spread over six fields selected for their low levels of obscuration due to dust. Here

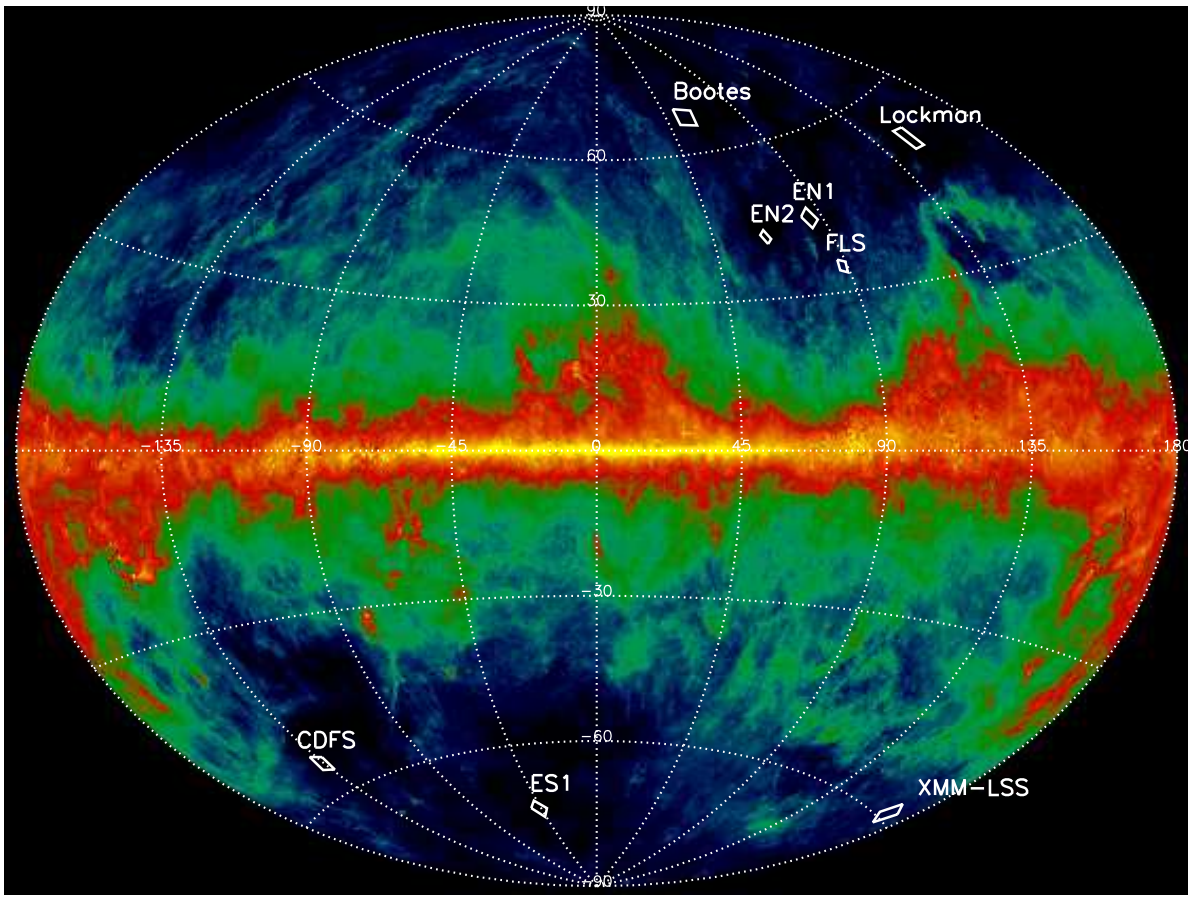

Figure 4.1: Spitzer Data Fusion fields overlaid on IRAS/COBE sky maps by Schlegel et al. (1998) in Galactic coordinates. Spitzer Data Fusion fields are a subset of HerMES fields in Fig. 3.8.3.

we briefly describe the properties of the SWIRE Final Data Release (FDR), which are better detailed by Surace et al. (in prep).

IRAC and MIPS 24 μm Catalogues

Sources were detected and fluxes extracted from IRAC mosaics using the SExtractor software (Bertin & Arnouts 1996). Aperture fluxes were extracted within five separate apertures (ap1/ap2/ap3/ap4/ap5), which for IRAC range from 1 to 6 times the IRAC beam FWHM. Colour-Magnitude diagrams were constructed for various types of objects, in particular main-sequence stars. It was found that the scatter in these diagrams were minimized through the use of ap2, corresponding to a 3.8 arcsec aperture diameter and roughly twice the IRAC beam FWHM. We used these IRAC ap2 fluxes to produce our database.

Map Making, source extraction and flux measurement on MIPS 24 μm mosaics were performed using MOPEX/APEX by Makovoz and Marleau 2005[1]. MOPEX/APEX APEX uses PRF fitting to carry out point source extraction but also flags extended sources and computes

[1] `http://ssc.spitzer.caltech.edu/postbcd/mopex.html`

fluxes within several apertures. We used these MIPS 24 μm PRF fitting fluxes to produce our database.

IRAC four single-band catalogues and MIPS 24 μm catalogue were combined into a 5-band catalogue using SSC's `bandmerge` software (see `http://ssc.spitzer.caltech.edu/postbcd/bandmerge.html`]). Briefly, `bandmerge` computes for each band-to-band pairing a χ^2 value based on positional uncertainties, then checks the linkages for consistency, falling back to second- or third-choice links if necessary. The final output includes only the best final linkages, so that a source from a given band appears only once in the merged list. This catalogue contains only sources lying within the region which has full coverage in all four IRAC bands.

For more information about the IRAC and MIPS 24 μm data processing see Surace et al. (2005) and Shupe et al. (2008). Note however how the SWIRE FDR uses deeper MIPS 24 μm catalogues than discussed by Shupe et al. (2008), obtained using APEX rather than SExtractor. In the course of this work, however, we discovered (Shupe & Vaccari, priv comm, see also Starikova et al. (2012)) how bright and/or extended nearby MIPS-24 sources are occasionally broken down in multiple components or lost altogether by this new procedure. We thus devised a updated MIPS 24 μm source extraction procedure which uses SExtractor and is both more reliable for bright and/or extended sources and deeper for fainter and/or point-like sources and thus more suitable for Herschel Source Extraction. Measured fluxes were calibrated following the latest MIPS 24 μm nominal calibration (Engelbracht et al. 2007) and colour-corrected so as to place them on a constant $\nu\, S_\nu$ scale which is appropriate for a wide range of Galaxy and AGN SEDs.

MIPS 70 and 160 μm Catalogues

MIPS 70 and 160 μm maps and catalogues were produced using MOPEX/APEX as done for MIPS 24 μm observations. Flux calibration was carried out following the latest MIPS instrumental calibration by Gordon et al. (2007) and Stansberry et al. (2007), and fluxes were colour-corrected as as to put them (like their 24 μm counterparts) on a constant $\nu\, S_\nu$ scale.

4.1.2 Bootes Field

The Bootes field was defined by the NOAO Deep Wide Field Survey (NDWFS) optical imaging covering ~ 9 deg^2. The SDWFS (PI Daniel Stern, Spitzer PID 40839) and MAGES (PI Buell Jannuzi, Spitzer PID 50148) Spitzer programmes later provided IRAC and MIPS data respectively over the whole NDWFS field. In this work, SDWFS DR 1.1 IRAC public images by Ashby et al. (2009) were re-extracted using an updated SWIRE/SERVS SExtractor pipeline, while MAGES MIPS data were re-reduced following the best practices employed in the reduction of similar data within the SWIRE fields at 24 μm (Shupe & Vaccari, priv comm) as well as at 70 and 160 μm (Seymour & Tugwell, priv comm).

4.1.3 XFLS Field

The Extragalactic First Look Survey (XFLS) field is defined by the Spitzer First Look Survey (FLS, PI Tom Soifer, Spitzer PID 26), which provided IRAC and MIPS data over $\sim$ 4 deg^2. In this work, as done for the Bootes field, XFLS IRAC public images by Lacy et al. (2005) were re-extracted using an updated SWIRE/SERVS SExtractor pipeline, while XFLS

MIPS data were re-reduced following the best practices employed in the reduction of similar data within the SWIRE fields at 24 μm (Shupe & Vaccari, priv comm) as well as at 70 and 160 μm (Seymour & Tugwell, priv comm).

4.2 Ancillary Datasets

A number of ancillary datasets, including photometry from the FUV to the NIR and optical spectroscopy, were cross-correlated with the Spitzer catalogues described in the previous Section. Here we briefly review their properties in order of increasing wavelength.

4.2.1 GALEX Photometry

The Galaxy Evolution EXplorer (GALEX) mission (Martin et al. 2005a) was conceived to carry out a wide-ranging set of imaging and spectroscopic surveys. Five imaging surveys are underway in a FUV band (1350–1750 Å) and a NUV band (1750–2800 Å) with 5" resolution (FWHM) and 1" astrometry, and a cosmic UV background map. Three overlapping slitless grism spectroscopic additionally performed surveys over the 1350–2800 Å band with $R \sim 100$, resulting in greater than 100,000 galaxies with redshifts ($0 < z < 2$), extinction, and SFR.

The GALEX Deep Imaging Survey (DIS) was conceived to provide UV photometric observations for the most popular fields for galaxy evolution studies, and has surveyed 80 deg^2 down to $m_{AB} \sim 25$, including all of our fields. Version 1.0 of the Spitzer Data Fusion uses the GALEX GR4/5 data release, while the latest, and arguably last, GALEX data release, or GALEX GR6, will be incorporated in a future version.

4.2.2 SDSS Photometry and Spectroscopy

The Sloan Digital Sky Survey (SDSS, York et al. (2000)) is an optical of survey mapping about one-quarter of the entire sky in detail, determining the positions and absolute brightnesses of hundreds of millions of celestial objects. The SDSS telescope (Gunn et al. 2006) saw first light in May 1998, and entered routine operations in April 2000. It is located at Apache Point Observatory (APO) in Sunspot, New Mexico, it has a 3 deg diameter field of view, and the imaging uses a drift-scanning camera with 30 2048×2048 CCDs at the focal plane which image the sky in five broad filters covering the range from 3000 Å to 10000 Å. The imaging is carried out on moonless and cloudless nights of good seeing, and the resulting images are calibrated photometrically (Tucker et al. 2006) to a series of photometric standards around the sky. The five filters in the imaging array of the camera, u, g, r, i, and z, have design effective wavelengths of 3550, 4770, 6230, 7620, and 9130 Å, respectively (Fukugita et al. 1996), while the close-to-final values are 3560, 4680, 6180, 7500, and 8870 Å. The 95% completeness limits of the images are $u, g, r, i, z = 22.0, 22.2, 22.2, 21.3, 20.5$ (Abazajian et al. 2004), although these values depend as expected on seeing and sky brightness. To meet stringent requirements in both signal-to-noise ratio and photometric calibration, imaging data are declared to be of survey quality only if the night is determined to be photometric, with a zero-point uncertainty below 1%, and if the seeing is better than 1.5 arcsec. The imaging data saturate at about 13, 14, 14, 14, and 12 mag for point sources.

Photometric calibration in SDSS has been carried out in two parallel approaches. The first uses an auxiliary 20-inch photometric telescope (PT) at the site, which continuously

surveys a series of US Naval Observatory (USNO) standard stars which are used to define the SDSS *ugriz* photometric system (Smith et al. 2002). Transformations between the *ugriz* and native SDSS 2.5-meter *ugriz* photometric systems (*AB* magnitude) and zero-points for stars in patches surveyed by the 2.5-m telescope are determined with these data (Tucker et al. 2006). These secondary patches are spaced roughly every 15 deg along the imaging stripes. This approach has allowed the SDSS photometry to reach its goals of calibration errors with an rms of 2% in g, r, and i, and 3% in u and z (Ivezić et al. 2004), as measured from repeat scans (see the discussion in Ivezić et al. (2007)). This calibration process has been used in earlier SDSS data releases but is not ideal for several reasons. A second approach, developed with DR6 and called ubercalibration (Padmanabhan et al. 2008), does not use information from the PT to calibrate individual runs, but rather uses the overlaps between the 2.5-m imaging runs to tie the photometric zero-points of individual runs together and measure the 2.5-m flatfields, and to determine the extinction coefficients on each night. This is the default calibration adopted in the latest SDSS data releases. The results of the improved photometric calibration are detections with uncertainties of roughly 1% in g, r, i, and z, and 2% in u, substantially better than the uncertainties in previous data releases.

After astrometric calibration (Pier et al. 2003) the properties of detected objects in the five filters are measured in detail (Lupton et al. 2002). Subsets of these objects are selected for spectroscopy, including galaxies (Strauss 2002), quasars (Richards et al. 2002), and stars. The spectroscopic targets are assigned to a series of plates containing 640 objects each (Blanton et al. 2003), and spectra are measured using a pair of double spectrographs, each covering the wavelength range 3800−9200 Å with a resolution $\lambda/\Delta\lambda$ which varies from 1850 to 2200. These spectra are calibrated in wavelength and flux, and classifications and redshifts, as well as spectral types for stars, are determined by a series of software pipelines (SubbaRao et al. 2002).

The SDSS is the first very-large-area survey to use electronic light detectors, so the images it produces are substantially more sensitive and accurate than earlier very-wide-area surveys, which relied on photographic plates. The results of the SDSS are electronically available to the scientific community and the general public, both as images and as precise catalogues of all objects discovered. The SDSS photometric and spectroscopic data are merged into a sophisticated database and are made available both through an object-oriented database (the Catalogue Archive Server, CAS) and as flat data files (the Data Archive Server, DAS) in progressive data releases.

The SDSS provides 5-band *ugriz* 95% complete photometry down to $r \sim 22.2$ and complete spectroscopy down to $r \sim 17.77$ selected on the basis of the photometric input catalogue. For normal galaxies, this effectively limits the range of measured spectroscopic redshifts to $z < 0.15$ and the range of accurately measured photometric redshifts at $z < 0.5$, and is thus very well suited for local studies.

The SDSS Data Release 7 (DR7, Abazajian et al. (2009) marks the completion of the original goals of the SDSS. It includes 11,663 deg^2 of imaging data. The catalogue contains five-band photometry for 357 million distinct objects. The survey has complete spectroscopy over 9380 deg^2. There are over 1.6 million spectra in total, including 930,000 galaxies, 120,000 quasars, and 460,000 stars. The astrometry has all been recalibrated with the second version of the USNO CCD Astrograph Catalogue, reducing the rms statistical errors at the bright end to 45 mas per coordinate. A systematic error in bright galaxy photometry due to poor sky determination was better quantified and judged to be a less severe problem than previously

reported for the majority of galaxies. Finally, a series of improvements to the spectroscopic reductions were carried out, including better flat fielding and improved wavelength calibration at the blue end, better processing of objects with extremely strong narrow emission lines, and an improved determination of stellar metallicities

The SDSS has recently entered a new phase, referred to as SDSS-III (Eisenstein et al. 2011), which between other things will greatly improve the spectroscopic completeness for Luminous Red Galaxies up to $z \simeq 0.7$ observational cosmology and galaxy formation and evolution. The first SDSS-III data release, or SDSS DR8 (Aihara et al. 2011), contains the final SDSS imaging products, which have been used to define the later spectroscopic sample. Several improved photometric redshift estimates are now provided as part of the data release (Csabai et al. in prep, `http://www.sdss3.org/dr8/algorithms/photo-z.php`), and the now completed PRIMUS spectroscopic redshift database (Coil et al. 2011) has allowed (Sheldon et al. in prep, `http://arxiv.org/abs/1109.5192`) to produce reliable photometric redshift PDFs.

Version 1.0 of the Spitzer Data Fusion uses the SDSS DR7, while the SDSS DR8 and future versions will be used in a future version.

Since SDSS photometry has been reduced and calibrated in a very careful and specific way with respect to other optical surveys used in the Spitzer Data Fusion we pay particularly attention to select the best quality of photometry for our objects and to create a dedicated photometric flag to guide the user in the exploitation of SDSS data ingested in our catalogue. We firstly selected only objects classified as `primary` detections, i.e. the SDSS catalogue entries that are the best version, in terms of photometric quality, of a given astronomical source on the sky; we thus discard those classified as `secondary` detections, i.e. non-primary objects, and those called `family` detections, i.e. deblended objects.

We then create a flag in order to identify the objects with magnitudes within the reliability limits given in the documentation. As we already said, SDSS use a particular photometric system based on the concept of `luptitudes` that substitute the more classical concept of magnitude as explained in Lupton et al. (1999). This is an original photometric measurement which expresses magnitude as inverse hyperbolic sine (or "ashin") magnitudes. The transformation from linear flux measurements to asinh magnitudes is designed to be virtually identical to the standard astronomical magnitude at high signal-to-noise ratio, but to behave reasonably at low signal-to-noise ratio and even at negative values of flux, where the logarithm in the Pogson [2] magnitude fails. This allows us to measure a flux even in the absence of a formal detection. The asinh magnitudes are characterized by a softening parameter b, the typical 1-sigma noise of the sky in a PSF aperture in 1" seeing. The relation between detected flux f and $asinh$ magnitude m is

$$m = -2.5/\ln(10) * [asinh((f/f_0)/(2b)) + \ln(b)]. \quad (4.1)$$

where f_0 is given by the classical zero point of the magnitude scale, i.e., f_0 is the flux of an object with conventional magnitude of zero. The quantity b is measured relative to f_0, and thus is dimensionless. In our catalogue we report the already translated values of

[2]The Pogson magnitude is the standard astronomical magnitude system, where one increment in magnitude is an increase in brightness by the fifth root of 100. A star of 1st magnitude is therefore 100 times as bright as a star of 6th magnitude. That is, for two objects $M1 - M2 = -2.5log(F1/F2)$ where $M1$ and $M2$ are the magnitudes of two objects, and $F1$ and $F2$ are their luminous fluxes.

magnitudes following the Eq. 4.1 (that is the default of the released magnitudes in the official SDSS catalogues), but we add a flag based on the original luptitudes system to characterise the objects of the catalogue. We based our flag on intervals in magnitude limits and f_0 (zero flux objects) values reported in Tab 4.1 (from [3]). The table also lists the flux corresponding to $10f_0$, above which the asinh magnitude and the traditional logarithmic magnitude differ by less than 1% in flux.

Band	b	**Zero-Flux Magnitude** $[m(f/f_0 = 0)]$	$m(f/f0 = 10b)$
u	1.4×10^{-10}	24.63	22.12
g	0.9×10^{-10}	25.11	22.60
r	1.2×10^{-10}	24.80	22.29
i	1.8×10^{-10}	24.36	21.85
z	7.4×10^{-10}	22.83	20.32

Table 4.1: Reference number for SDSS `luptitudes` vs SDSS magnitudes.

The so called "mag(nitude) flag" assumed the following values (we defined a flag for each band):

- 1 : $\text{mag} \leq \text{mag}_{\text{lim}}$
- 2 : $\text{mag}_{\text{lim}} \leq \text{mag} \leq \text{mag}$ $(f/f_0 = 0)$: object fainter than the magnitude limit
- 3 : $\text{mag}(f/f_0 = 0) \leq \text{mag}$: object not detected

We also followed the recommendation by the SDSS collaboration in order to select "clear" photometric objects as explained at `http://cas.sdss.org/astro/en/help/docs/realquery.asp#flags`. To be on the safe side, being as conservative as we can, we do not discard objects that are not identified as "clear" objects leaving them in the catalogues and creating a flag called "reliability flag" that is 1 when the photometry is clear and 0 when an object has unclear photometry (this flag is single and it refers to the photometry from all the bands). SDSS stores in its database many other photometric flags; while downloading the catalogue we paid attention to download the more useful of them along with the data in each band to allow users to do the most disparate studies.

SDSS Spectroscopic and Photometric Redshifts

We use the SDSS database not only for the photometry of our objects but also for their spectroscopic redshifts and photometric redshifts. All the spectroscopic information for SDSS spectroscopic objects are included in the `SpecObjAll` SDSS catalogue. Like for the photometry solution described above we select only objects defined as `sciencePrimary` or, in other words, which have spectra that are considered good enough for science discarding all the objects that do not satisfy this criterion.

Photometric redshifts for the SDSS photometric objects computed using the template fitting technique by Csabai et al. (2003) were first made available as part of the SDSS Early Data Release (EDR) and Data Release 1 (DR1). No photometric redshifts were then made available for DR2 through DR4. Starting with DR5, however, there are two versions of

[3] `http://www.sdss.org/dr7/algorithms/edr.tb21.html`

photometric redshifts in the SDSS databases, in the `Photoz` and `Photoz2` tables respectively, computed with different algorithms.

The `Photoz` table contains photometric redshift based on the template fitting method applied on SDSS EDR by Csabai et al. (2003). The template fitting approach simply compares the expected colours of a galaxy (derived from template spectral energy distributions) with those observed for an individual galaxy. Since a representative set of photometrically calibrated spectra in the full wavelength range of the filters is not easy to obtain, they have used the empirical templates of Coleman et al. (1980) extended with spectral synthesis models. These templates were adjusted to fit the calibrations (see Budavári et al. 2000). `Photoz` photometric redshifts were computed for each SDSS detected object, leading to a low quality of redshifts for fainter objects. Csabai et al. (2003) e.g. claim that photometric redshift collected in `Photoz` present some problems in the range of z near 0 and 1 (note that both `Photoz` and `Photoz2` must be smaller than 1 by design) for galaxies with $r > 21$. This is shown in Fig. 16 in the Csabai paper. With DR7, the SDSS team have made substantial improvements to the `Photoz` photometric redshift code, using a hybrid method combining the template-fitting approach of Csabai et al. (2003) and an empirical calibration using objects with both observed colours and spectroscopic redshifts. The estimation method uses a k-d tree (following Csabai et al. (2007) to search in the ubercalibrated $u-g, g-r, r-i, i-z$ colour space for the 100 nearest neighbors of every object in the estimation set (i.e., the galaxies for which we want to estimate redshift) and then estimates redshift by fitting a local hyperplane to these points, after rejecting outliers. If an object lies outside the bounding box of the 100 nearest neighbours in colour space, the photometric redshift is less reliable, and the object is flagged accordingly. They use template fitting to estimate the k-correction, distance modulus, absolute magnitudes, rest frame colours, and spectral type. They search for the best match of the measured colours and the synthetic colours calculated from repaired (Budavári et al. 2000) empirical template spectra at the redshift given by the local nearest neighbor fit. The rms error of the redshift estimation for the reference set used to check the reliability of the photo-z against a measured spectroscopic redshift decreases from 0.044 in DR6 to 0.025 in DR7 with this improved algorithm. Iteratively removing the outliers beyond 3 σ gives rms errors of 0.028 and 0.020 for the old and new methods, respectively. In addition, the reliability of the quoted errors is much higher.

The photometric redshifts in the `Photoz2` table are a second set of photometric redshifts for galaxy objects, calculated using a Neural Network (NN) technique to calculate redshifts and the nearest neighbor error (NNE) method to estimate redshift errors (Oyaizu et al. 2008). `Photoz2` photometric redshifts were only computed for SDSS sources which were reliably detected at slightly brighter levels than required for `Photoz` and have not changed since DR6 (so do not use the ubercalibrated magnitudes). However, SDSS team provide a value-added catalogue containing the redshift probability distribution for each galaxy, p(z), calculated using the weights method presented in Cunha et al. (2009). The p(z) for each galaxy in the catalogue is the weighted distribution of the spectroscopic redshifts of the 100 nearest training-set galaxies in the space of dereddened model colours and r magnitude. Cunha et al. (2009) showed that summing the p(z) for a sample of galaxies yields a better estimation of their true redshift distribution than that of the individual photometric redshifts. Mandelbaum et al. (2008) found that this gives significantly smaller photometric lensing calibration bias than the use of a single photometric redshift estimate for each galaxy (Abazajian et al. 2009).

4.2.3 Deep Optical Photometry

Deep optical imaging was obtained over the last decade covering most of our fields by the SWIRE, CFHTLS, NOAO/Bootes and SSC/XFLS teams. Multi-band *ugriz* imaging is thus available from the INT WFC in LH, EN1, EN2 and XFLS (McMahon et al. 2001; Babbedge 2004; González-Solares et al. 2011), from the KPNO MOSAIC1 in LH (González-Solares et al. 2011), from the CTIO MOSAIC2 in CDFS (Siana et al. priv comm) and from the CFHT MegaCam in XMM (CFHTLS W1 T0006, `http://terapix.iap.fr/rubrique.php?id_rubrique=259`). BVR (Berta et al. 2006) and Iz (Berta et al. 2008) imaging is available from the WFI and VIMOS camera respectively in ES1. BRI and JHK_s imaging is available from the KPNO MOSAIC1 and NEWFIRM cameras respectively in Bootes (NDWFS DR3, Jannuzi et al. (2004) and IBIS DR2, Gonzalez et al. (2010). Single-band CFHT u (González-Solares et al. 2011) and KPNO MOSAIC1 R (Fadda et al. 2004) imaging is also available in XFLS. The vast majority of optical data available in LH, EN1, EN2 and XFLS was re-reduced by González-Solares et al. (2011) using the Cambridge Astronomical Survey Unit (CASU) pipeline (Irwin & Lewis 2001), ensuring an homogeneous astrometric and photometric calibration against SDSS and 2MASS. An area of 8 deg^2 split between ES1 and CDFS will also be covered in *ugri* by the ongoing VST-VOICE survey (PI Covone & Vaccari). Observations have started in the 2011/2012 observing season and will continue for 4 more years.

4.2.4 2MASS Photometry

The Two Micron All Sky Survey (2MASS, Skrutskie et al. (2006)) project was designed to close the gap between our current technical capability and our knowledge of the near-infrared sky and thus provide direct answers to immediate questions on the large-scale structure of the Milky Way and the Local Universe. Its survey has uniformly scanned the entire sky in three NIR bands to detect and characterize point sources brighter than about 1 mJy in each band, with SNR greater than 10, using a pixel size of 2.0 arcsec. This survey use two highly-automated 1.3-m telescopes, one at Mount Hopkins, Arizona, and one at CTIO, Chile. Each telescope is equipped with a three-channel camera, each channel consisting of a 256×256 array of HgCdTe detectors manufactured by Rockwell International Science Center, now Rockwell Scientific. These arrays have 40 μm pixel pitch and are sensitive between wavelengths of 0.8 and 2.5 μm. Two dichroic mirrors within each camera permit simultaneous imaging of an $8.5' \times 8.5'$ field at a pixel scale of 2 arcsec/pixel in the three 2MASS wave bands. These bands, defined primarily by interference filters, as well as by detector quantum efficiency, transmission/reflection losses at the system's optical surfaces, and the atmosphere, largely correspond to the classical J, H, and K bands defined by Johnson (1962), with the exception that the 2MASS 'K-short (K_s) filter excludes wavelengths longward of 2.31 μm to reduce thermal background and airglow and includes wavelengths as short as 2.00 μm to maximize bandwidth. It is capable of observing the sky simultaneously at J (1.24 μm), H (1.66 μm), and K_s (2.16 μm). The 7.8 s of integration time accumulated for each point on the sky and strict quality control yielded a 10 σ point-source detection level of better than 15.8, 15.1, and 14.3 magnitudes at the J, H, and K_s bands, respectively, for virtually the entire sky (Skrutskie et al. 2006). The 2MASS All-Sky Data Release includes 4.1 million compressed FITS images covering the entire sky, 471 million source extractions in a Point Source Catalogue, and 1.6 million objects identified as extended in an Extended Source Catalogue. 2MASS has timely

released data products to the astronomical community. The All-Sky Data Release took place in March 2003 (this releases covers 99.998% and supersedes all earlier incremental data releases. Version 1.0 of the Spitzer Data Fusion uses the 2MASS All-Sky Point Source Catalogue (PSC) and Extended Source Catalogue (XSC) by Cutri et al. (2003) which can be accessed through IPAC's IRSA/GATOR Service [4] and both fully cover all of our fields. The PSC is recommended for the great majority of unresolved sources, while XSC is recommended for the few sources which show up as extended at 2MASS depth and resolution.

4.2.5 UKIDSS Photometry

The United Kingdom Infrared Deep Sky Survey (UKIDSS) is a suite of five public surveys which began in May 2005 and is now seeking funding for continuation beyond its nominal completion at the end of 2012. These surveys all use the Wide Field Camera (WFCAM) on the United Kingdom Infrared Telescope (UKIRT), which has four Rockwell Hawaii-II 20482 PACE arrays, with a pixel scale of 0.4 arcsec, giving a solid angle of 0.21 deg^2 per exposure. UKIDSS is a set of five survey of complementary combinations of depth and area, employing the wavelength range $0.83-2.37$ μm in up to five filters $ZYJHK$, and extending over both high and low Galactic latitude regions of the sky. The data, as the SWIRE optical data by González-Solares et al. (2011), are reduced and calibrated at the Cambridge Astronomical Survey Unit (CASU) using a dedicated software pipeline (Irwin & Lewis 2001) and are then transferred to the WFCAM Science Archive (hereafter the WSA, Hambly et al. (2008)) in Edinburgh where are ingested into a sophisticated SQL database and detections in the different passbands are merged. The UKIDSS data are released by the WFCAM Science Archive at intervals of six to twelve months.[5] There are three surveys targeting Extragalactic Fields. The Large Area Survey (LAS) is a wide, relatively shallow survey that will cover 4000 deg^2 from within the footprint of the SDSS, in the four bands $YJHK$. The depth, $K = 18.4$, is some 3 mag deeper than 2MASS. The Deep Extragalactic Survey (DXS), is of intermediate depth, $K = 21.0$, and covers 35 deg^2 in J and K within SIWRE fields. The deepest survey is the Ultra Deep Survey (UDS), which covers 0.78 deg^2 in JHK to a depth $K = 23.0$ within XMM. Then there are two surveys targeting Galactic fields. The Galactic Plane Survey (GPS) covers some 1800 deg^2, defined by the sections of the Galactic latitude band -5 deg $< b < +5$ deg that are contained within the Declination limits -15 deg $< \delta < +60$deg. This region is imaged in JHK to a depth $K = 19.0$. Finally, the Galactic Clusters Survey images 11 stellar open clusters and star formation associations, covering 1400 deg^2, in all five bands $ZYJHK$, to a depth $K = 18.7$.

Version 1.0 of the Spitzer Data Fusion uses the DXS bandmerged catalogue from the UKIDSS DR9Plus Data Release, covering portions of XMM, LH and EN1 in J and K and overlapping with SWIRE over 15 deg^2.[6]. An area of 12 deg^2 split between ES1, XMM and CDFS will also be covered in $ZYJHK_s$ by the ongoing VISTA-VIDEO survey (PI Jarvis), and the ESO Science Archive and the VISTA Science Archive have started distributing the first preliminary data products.

As for the SDSS photometry, we apply specific solutions in downloading UKIDSS data from the online database. First of all we try to avoid bad photometry and/or duplicate objects by selecting only objects identified as "unique" sources in the UKIDSS catalogue. Because

[4] `http://irsa.ipac.caltech.edu/applications/Gator/`

[5] `http://surveys.roe.ac.uk/wsa/`

[6] `http://surveys.roe.ac.uk/wsa/dr9plus_release.html`

of the spacing of the detectors in WFCAM, and the restrictions on guide star brightness, there will always be overlap regions between adjacent frame sets. Source merging is done on a set-by-set basis; hence after source merging there are usually a small number of duplicate sources in the table. A process known as seaming takes place after source merging is complete, whereby duplicates are identified and flagged. The flagging attribute is `priOrSec`, and the meaning of the flag is quite simple: if a source is not found to be duplicated in overlap regions, then `priOrSec`=0; if a source is duplicated, then `priOrSec` will be set to the frameSetID of the source that should be considered the best one to use out of the set of duplicates. Presently, the choice of which is best is made on the basis of proximity to the optical axis of the camera, the assumption being that this will give the best quality image in general. So, if a particular source has a non-zero `priOrSec` that is set to it's own value of `frameSetID`, then this indicates that there is a duplicate elsewhere in the table, but this is the one that should be selected as the best (i.e. this is the primary source). On the other hand, if a source has a non-zero value of `priOrSec` that is set a different `frameSetID` than that of the source in question, then this indicates that this source should be considered as a secondary duplicate of a source who's primary is actually to be found in the frame set pointed to by that value of `frameSetID`. Thus in our Spitzer Data Fusion we selected only objects with `priOrSec`=0 or `priOrSec`=`frameSetID`.

Besides, accordingly to the documented recommendation on photometry reliability, we associate a "reliability flag" to the UKIDSS data collected in our database. This flag is based on the UKIDSS flag `ppErrBits` (Post-processing error quality bit flags). `ppErrBits` tells if a source has problematic detection like: is deblended, is detected in a bad pixels in default aperture, is located in a poor flat field region, is close to saturated, etc. (see `ppErrBits` on line documentation at `http://surveys.roe.ac.uk/wsa/www/gloss_p.html#dxssource_pperrbits` for more information). Since this information is stored in a bit flag, this could be a cryptic information for an non-expert user; for this reason we add a "reliability flag" which is a simple Boolean flag whose value is 1 for "good" detections and 0 for "bad" detections, i.e. where `ppErrBits` is not 0. If our "reliability flag" is 0 than the user knows that the detection could have some problems and that it would be preferable to discard the source from the sample or to check which specific problem may affect the source, as indicated by the `ppErrBits` flag available for each UKIDSS band.

Finally, the magnitude limits of UKIDSS depend on the aperture corrections associated to each aperture magnitude, which are different between different bands and between different apertures. Derived detector zero-points, i.e. the magnitudes at which an object gives a total (corrected) flux of 1 count/s, are stored within the UKIDSS database and are appropriate for generating magnitudes in the natural detector+filter system based on Vega. These zero-points have been derived from a robust average of all photometric standards observed on any particular set of frames, but for different airmass or values of the extinction, that could change from an observation to another, the zero-point needs to be estimated for the correct values of these quantities. Once corrected, the zero-points can be used to estimate the magnitudes and magnitude limits of each band and aperture following the instruction stored in the `photZPcat` at [7]. To help the non-expert user, we have already stored in our Spitzer Data Fusion the magnitude limits of the UKIDSS reference aperture magnitudes.

[7] `http://surveys.roe.ac.uk/wsa/www/gloss_p.html`

4.2.6 SWIRE Photometric Redshifts

Accurate photometric redshifts for more than a million galaxies over ~ 35 deg^2 with good multi-band optical coverage within the SWIRE fields were computed by Rowan-Robinson et al. (2008) using a preliminary version of the Spitzer and optical catalogues used in the Spitzer Data Fusion. Since the present work substantially extends the multi-wavelength coverage available in Spitzer fields and thus increases the number of Spitzer sources for which photometric redshifts can be reliably computed, an updated photometric redshift was produced, incorporating improvements suggested by recent work. Version 1.0 of the Spitzer Data Fusion uses the latest version of the SWIRE photometric redshift catalogue at the time of writing, or `zcatrev12ff2`, which is available at `http://astro.ic.ac.uk/~mrr/swirephotzcat/zcatrev12ff2.dat.gz` and described by Rowan-Robinson et al. (in prep). The SWIRE photometric redshift catalogue provides the largest accurate photometric redshift catalogue for SWIRE sources at $z > 0.5$.

4.2.7 NED Spectroscopic Redshifts

The NASA Extragalactic Database (NED) at `http://nedwww.ipac.caltech.edu/` was searched for spectroscopic redshifts, so as to complement the SDSS spectroscopic database. A bulk download script was implemented selecting all reliable spectroscopic redshifts within the footprints of our fields, and the result was merged with recently released public data which have not appeared on NED as yet and with proprietary spectroscopic redshifts available to the SWIRE and/or HerMES consortia.

4.3 Spitzer Data Fusion Construction

The Spitzer Data Fusion is constructed by combining Spitzer IRAC and MIPS as well as ancillary catalogues according to the following positional association procedure.

4.3.1 Spitzer Bandmerging

The Spitzer Data Fusion is a multi-wavelength catalogue of Spitzer/IRAC-selected sources. As detailed above, the SWIRE (PI Carol Lonsdale, Spitzer PIDs 142, 181, 182, 183, 184, 185) Final Data Release (Surace et al. in prep), our own Bootes catalogues based on SDWFS (PI Daniel Stern, Spitzer PID 40839) public images by Ashby et al. (2009) and our own XFLS catalogues based on XFLS (PI Tom Soifer, Spitzer PID 26) public images by Lacy et al. (2005) provide the primary IRAC selection in the different fields.

For SWIRE fields, IRAC and MIPS 24 μm single-band source lists are combined using SSC's `bandmerge` software, as described in Sec. 4.1.1, mostly limiting reliable associations to angular separations smaller than about 2.5 arcsec. In the Bootes and XFLS fields, sources from our own IRAC 3.6 and 4.5 μm catalogues are bandmerged into a two-band catalogue finding their nearest neighbour within a 1.0 arcsec search radius, and IRAC 5.8 and 8.0 μm and MIPS 24 μm are associated to the resulting IRAC 3.6/4.5 μm catalogue using a 1.5 arcsec (5.8 and 8.0 μm) and 2.5 arcsec (24 μm) search radius respectively.

In all fields, MIPS 70 and 160 μm catalogues are matched against MIPS 24 μm positions using a nearest-neighbour association with a search radius of 10 and 20 arcsec (determined

by the PSF FWHM at the two wavelengths and at 24 μm), following Rowan-Robinson et al. (2008) and Vaccari et al. (in prep) where a 24 μm detection is required to guarantee the reliability of 70 and 160 μm detections but also to more accurately pinpoint their positions. Experience has shown that these search radii maximize the completeness and reliability of the association process.

The reliability R of each MIPS 24 μm vs IRAC, MIPS 70 μm vs MIPS 24 μm and MIPS 160 μm vs MIPS 24 μm association is based on the actual positional separation d and source density σ of the denser 'reference' catalogue as follows : $R = 1 - \pi d^2 \sigma$).

4.3.2 Astrometric Registration

Both the Spitzer-selected catalogue and the ancillary catalogues are then astrometrically registered against 2MASS, which provides the best combination of astrometric accuracy and source density across the whole sky. 2MASS is also the astrometric reference frame adopted by the Spitzer Science Center for Spitzer operations and the Spitzer Heritage Archive, and its use should thus also minimize systematic astrometric offsets to be found between Spitzer-detected sources and their 2MASS counterparts. The median (RA,Dec) positional offsets against 2MASS are computed for high-significance detections and subtracted for each ancillary catalogue in each field, so that the average offset against 2MASS is then very close to zero. The positional median absolute deviation (MAD) for all sources after astrometric registration provides an indication of the level of astrometric accuracy of each ancillary catalogue at the faint end (see Table4.2), while astrometric accuracy for bright non-saturated objects is substantially (a factor of two or more) better.

4.3.3 Multi-Wavelength Bandmerging

Following astrometric registration, ancillary catalogues are matched against Spitzer/IRAC positions using a simple nearest-neighbour association and an homogeneous search radius of 2.5 arcsec. While more refined multi-wavelength cross-correlation techniques following e.g. Sutherland & Saunders (1992) or Roseboom et al. (2009) may be useful in deeper Spitzer fields, the astrometric accuracies and source densities of individual catalogues under consideration are mostly well-matched to reliably provide a straightforward nearest-neighbour association. The number of matches obtained with each ancillary catalogue is reported in Tab. 4.3. For each matched source in each ancillary catalogue the reliability of the match is defined based on the angular separation and on the source density of the ancillary catalogue

4.4 Public Release of the Spitzer Data Fusion

The Spitzer Data Fusion aims to be the most complete publicly available multi-wavelength catalogue of Spitzer-selected sources over the wide-area "Cosmic Windows" surveyed by Spitzer and now Herschel. This is by definition an ongoing work as we improve the data reduction of Spitzer data (including new data from Spitzer Warm surveys such as SERVS, Mauduit et al. (2012)) ad continue the collation of previously proprietary multi-wavelength data. Here we describe the general properties of its Version 1.0.

The Spitzer Data Fusion provides ample opportunities for galaxy formation and evolution studies. Of particular importance is its application to the Herschel Multi-Tiered Extragalactic

Field	IRAC	IRAC	MIPS	MIPS	GALEX	GALEX	SDSS	SDSS	OPT-DEEP	OPT-DEEP	UKIDSS	UKIDSS
Name	RA	Dec	RA	Dec	RA	Dec	RA	Dec	RA	Dec	RA	Dec
ES1	0.183	0.160	0.860	0.817	0.636	0.619	NA	NA	NA	NA	NA	NA
XMM	0.186	0.165	0.688	0.614	0.623	0.619	NA	NA	NA	NA	0.165	0.155
CDFS	0.141	0.144	0.866	0.838	0.612	0.614	NA	NA	NA	NA	NA	NA
LH	0.164	0.149	0.631	0.665	0.590	0.587	0.200	0.192	0.223	0.192	0.155	0.149
EN1	0.141	0.160	0.681	0.705	0.602	0.614	0.170	0.176	0.175	0.171	0.136	0.139
EN2	0.161	0.181	1.011	0.982	0.611	0.608	0.194	0.187	0.204	0.192	NA	NA
Bootes	0.180	0.181	0.622	0.598	0.669	0.667	0.197	0.204	NA	NA	NA	NA
XFLS	0.147	0.187	0.709	0.673	0.534	0.560	0.192	0.213	0.194	0.213	NA	NA

Table 4.2: Astrometric registration of the Spitzer Data FusionThe positional median absolute deviation (MAD) after the astrometric registration against 2MASS provides an indication of the level of astrometric accuracy of each ancillary catalogue.

Field Name	IRAC 3.6/4.5 μm	MIPS 24 μm	MIPS 70 μm	MIPS 160 μm	GALEX FUV/NUV	SDSS *ugriz*	OPT-DEEP misc	2MASS *J*/*H*/*K*	UKIDSS *J*/*K*	Area deg^2
ES1	390231	61236	2246	961	85039	NA	146537	10904	NA	7.0
XMM	498027	69629	3823	1702	104344	NA	327024	14794	151565	8.5
CDFS	462638	97002	4096	1813	101705	NA	177745	12952	NA	7.5
LH	660682	110516	5548	2417	158981	217005	432490	17139	226838	11.0
EN1	575524	102406	4652	2133	116180	210571	363949	21210	334955	9.5
EN2	272412	59378	2331	970	63774	103460	173880	11443	NA	4.5
Bootes	677522	41969	4325	2825	159218	228757	592136	7007	NA	8.5
XFLS	107720	16712	2252	322	29208	62437	82576	11682	NA	4.5

Table 4.3: Number of sources contained in the Spitzer Data Fusion at different wavelengths. The main catalog selection (and thus the total number of sources) is at IRAC 3.6/4.5 *mu*m, and the other channels are matched against this sample.

Survey (Oliver et al. 2012). Version 1.0 was extensively since the Herschel Science Demonstration Phase work by the HerMES consortium, e.g. to improve the Herschel source extraction process (Roseboom et al. 2010) and carry out the SED fitting of Herschel sources (e.g. Eales et al. (2010b), Rowan-Robinson et al. (2010), Vaccari et al. (2010)).

Version 1.0 of the Spitzer Data Fusion will be made publicly available at `http://www.astro.unipd.it/background/df/` concurrently with the SWIRE Final Data Release as VO-compliant fits files which can be easily read by Fortran using NASA's `cfitsio` library or by IDL using NASA's `astrolib` library as well as by Starlink's `stilts/topcat` applications, so as to integrate easily with public astronomical science archives hosted e.g. at IRSA (`http://irsa.ipac.caltech.edu/`), MAST (`http://archive.stsci.edu/`) and Vizier (`http://vizier.u-strasbg.fr/viz-bin/VizieR`). For each field, Version 1.0 provides a "main" catalogue file and a number (one for each telescope used in ancillary observations) of "ancillary" catalogue files. The "main" catalogue reports the position and one flux measurement per each wavelength but only a limited number of other quantities and flags. Each "ancillary" catalogue contains several aperture and total flux measurements as well as a comprehensive set of photometric quantities and flags, allowing the expert user to make the most of ancillary observations. The "main" and "ancillary" catalogues have got the same number of entries following the same order, but can also be matched through the Spitzer running index which appears as the first column in both the "main" and all the "ancillary" catalogues. The rows contained in Version 1.0 of the "main" catalogue are illustrated in Sec. 4.5. Future releases will be fully documented and will include detailed coverage levels at all wavelengths, flux limits for non-detections, homogeneous star-galaxy flags, improved estimates of the association reliability and estimates of the completeness and reliability as a function of coverage at all wavelengths. This will provide detailed survey selection functions and thus enable a wider range of statistical studies.

Spitzer fluxes are expressed in μJy (IRAC and MIPS 24 μm) or mJy (MIPS 70 and 160 μm), where UV/Optical/NIR fluxes are expressed in magnitudes. Magnitudes are mostly in the *AB* system but occasionally in the Vega system when public ancillary data are thus provided, and in such cases Vega-to-AB conversion factors are provided.

4.5 Detailed Contents of the Spitzer Data Fusion "Main" Catalogue

As we said, the Spitzer Data Fusion Version 1.0 provides a "main" catalogue file and a number of "ancillary" catalogues for each field. Here we list the quantities included in the "main" catalogue, i.e. a subset which should allow a quick inspection of the multi-wavelength database and serve as reference values for object indices, coordinates and fluxes at most wavelengths of interest.

1. Index IRAC
2. RA IRAC
3. Dec IRAC
4. Index MIPS-24

5. RA MIPS-24
6. Dec MIPS-24
7. Index MIPS-70
8. RA MIPS-70
9. Dec MIPS-70
10. Index MIPS-160
11. RA MIPS-160
12. Dec MIPS-160
13. Index UV (GALEX)
14. RA UV (GALEX)
15. Dec UV (GALEX)
16. Index Optical (SDSS)
17. RA Optical (SDSS)
18. Dec Optical (SDSS)
19. Index Optical (Extra)[8]
20. RA Optical (Extra)
21. Dec Optical (Extra)
22. Index NIR (2MASS - Point Source Catalogue, PSC)
23. RA NIR (2MASS - Point Source Catalogue, PSC)
24. Dec NIR (2MASS - Point Source Catalogue, PSC)
25. Index NIR (UKIDSS - Deep Extragalactic survey, DXS)
26. RA NIR (UKIDSS - Deep Extragalactic survey, DXS)
27. Dec NIR (UKIDSS - Deep Extragalactic survey, DXS)
28. FUV AB mag (GALEX)
29. NUV AB mag (GALEX)
30. FUV AB mag error (GALEX)
31. NUV AB mag error (GALEX)

[8]"Extra" for Optical quantities means that they come from ancillary optical survey, different from SDSS and that could be different from field to field, as explained in Sec. 4.2.3.

32. u AB mag (SDSS)
33. g AB mag (SDSS)
34. r AB mag (SDSS)
35. i AB mag (SDSS)
36. z AB mag (SDSS)
37. u AB mag error (SDSS)
38. g AB mag error (SDSS)
39. r AB mag error (SDSS)
40. i AB mag error (SDSS)
41. z AB mag error (SDSS)
42. u AB mag (Extra, as before)
43. g AB mag (Extra)
44. r AB mag (Extra)
45. i AB mag (Extra)
46. z AB mag (Extra)
47. u AB mag error (Extra, as before)
48. g AB mag error (Extra)
49. r AB mag error (Extra)
50. i AB mag error (Extra)
51. z AB mag error (Extra)
52. J Vega mag (2MASS - PSC)
53. H Vega mag (2MASS - PSC)
54. K Vega mag (2MASS - PSC)
55. J Vega mag error (2MASS - PSC)
56. H Vega mag error (2MASS - PSC)
57. K Vega mag error (2MASS - PSC)
58. J Vega mag (UKIDSS - DXS)
59. H Vega mag (UKIDSS - DXS)

60. K Vega mag (UKIDSS - DXS)
61. J Vega mag error (UKIDSS - DXS)
62. H Vega mag error (UKIDSS - DXS)
63. K Vega mag error (UKIDSS - DXS)
64. IRAC 3.6 micron flux [μJy]
65. IRAC 4.5 micron flux [μJy]
66. IRAC 5.8 micron flux [μJy]
67. IRAC 8.0 micron flux [μJy]
68. IRAC 3.6 micron flux error [μJy]
69. IRAC 4.5 micron flux error [μJy]
70. IRAC 5.8 micron flux error [μJy]
71. IRAC 8.0 micron flux error [μJy]
72. IRAC 3.6 micron SNR
73. IRAC 4.5 micron SNR
74. IRAC 5.8 micron SNR
75. IRAC 8.0 micron SNR
76. MIPS 24 micron flux [μJy]
77. MIPS 24 micron flux error [μJy]
78. MIPS 24 micron SNR
79. MIPS 24 micron / IRAC Distance [arcsec]
80. MIPS 24 micron Reliability
81. MIPS 70 micron flux [mJy]
82. MIPS 70 micron flux error [mJy]
83. MIPS 70 micron SNR
84. MIPS 70 micron / 24 micron Distance [arcsec]
85. MIPS 70 micron Reliability
86. MIPS 160 micron flux [mJy]
87. MIPS 160 micron flux error [mJy]

88. MIPS 160 micron SNR
89. MIPS 160 micron / 24 micron Distance [arcsec]
90. MIPS 160 micron Reliability
91. Redshift - master redshift
92. Redshift error - master redshift error (where available)
93. Redshift flag: 6 = NED spec-z & Spitzer detection; 5 = NED spec-z & SDSS detection; 4 = SDSS spec-z; 3 = SDSS photoz2 "z2cc2"; 1 = SDSS phot-z "z1"
94. SDSS Photometric Redshift z1
95. SDSS Photometric Redshift z2d1
96. SDSS Photometric Redshift z2cc2
97. SDSS Spectroscopic Redshift
98. NED Spectroscopic Redshift
99. SDSS Photometric Redshift Error z1
100. SDSS Photometric Redshift Error z2d1
101. SDSS Photometric Redshift Error z2cc2
102. SDSS Spectroscopic Redshift Error
103. NED Spectroscopic Redshift Error
104. SDSS stype Flag [9]
105. SDSS zconf Flag [10]
106. SDSS zstatus Flag [11]
107. SDSS specclass Flag [12]
108. Spitzer Catalogue Reliability Flag : 0 = duplicates, tile intersection and edge of fields (UNRELIABLE); 1 = full coverage and moderately reliable; 2 = full coverage and fully reliable

An illustration of the ample wavelength coverage provided by the Spitzer Data Fusion (along with Herschel data) is shown in Fig. 4.2, which relies on an extensive database of instrumental transmissions assembled for the Spitzer Data Fusion and included in App. B.

[9]SDSS photometric flag, see `http://cas.sdss.org/DR7/en/help/browser/browser.asp` for details.
[10]SDSS spectroscopic flag, see `http://cas.sdss.org/DR7/en/help/browser/browser.asp` for details.
[11]SDSS spectroscopic flag, see `http://cas.sdss.org/DR7/en/help/browser/browser.asp` for details.
[12]SDSS spectroscopic flag, see `http://cas.sdss.org/DR7/en/help/browser/browser.asp` for details.

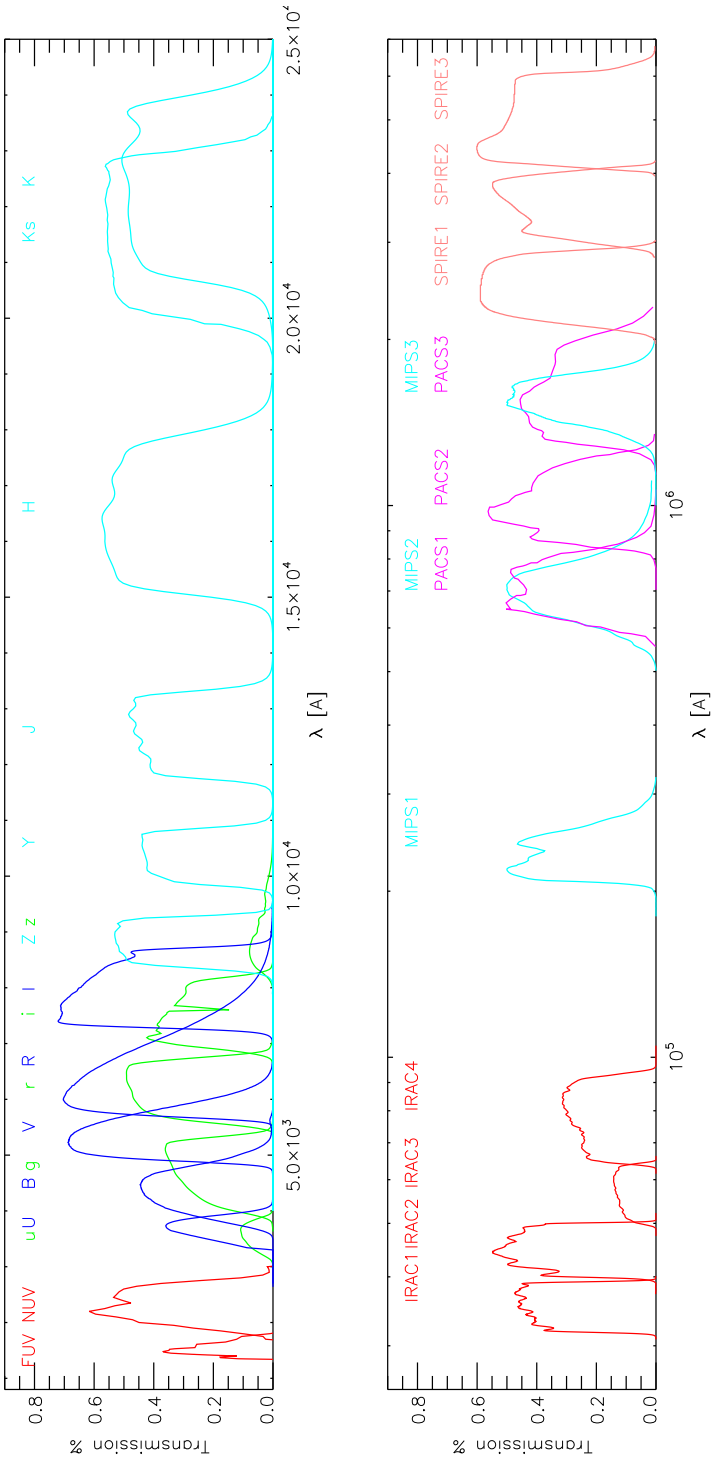

Figure 4.2: Instrumental Transmission of 'Optical' and 'Infrared' Filters.

4.6 Star-Galaxy Separation

Some form of discrimination between stars (or star-like objects) and galaxies is included within most of the catalogues we combined into the Spitzer Data FusionGiven the heterogeneous nature of most flags, however, but also the wealth of multi-wavelength data available within our fields, in our work we decided to derive an independent assessment of the star-galaxy nature of Spitzer sources based on a number of criteria used in the literature. This allows to exploit colour criteria in star-galaxy discrimination in a more homogeneous manner by making full use of the photometry available at different depth levels over different portions of the field. A star-galaxy binary flag is provided in the catalogue based on a combination of several criteria, namely

- Vaccari et al. (2005), Fig. 18 based on r-band and 24 μm flux (0001 = 1 = first bit)
- Shupe et al. (2008), Fig. 2 based on K-band and 24 μm flux (0010 = 2 = second bit)
- Vaccari et al. (2005), Fig. 19 based on r-band, K-band and 24 μm flux (100 = 4 = third bit)
- Rowan-Robinson et al. (2005), Fig. 1 and 2 based on r-band, i-band and 3.6 μm flux (1000 = 8 = fourth bit)

A source detected in all of the above bands and satisfying all of the four criteria will thus be flagged as $1+2+4+8 = 15$, and similarly for sources satisfying only a few of these criteria.

4.7 IRAC Source Counts

While an accurate characterization of the completeness of the Spitzer Data Fusion at different wavelengths and as a function of a number of parameters (also known as the survey selection function) is beyond the scope of this Thesis, due to its Spitzer selection the completeness at Spitzer wavelengths is particularly important and therefore deserves some further attention.

Detailed completeness estimates require extensive simulations to be carried out in all fields, but a rough assessment of catalogue completeness can be obtained by comparing measured source counts against published ones. This is illustrated in Fig. 4.4 for Spitzer IRAC bands. The departure of measured source counts in any given field from published counts (corrected for incompleteness) provides a rough measure of the level of completeness achieved in the field. These rough completeness estimates are in reasonable agreement with the results of detailed simulations carried out in some sub-field at a few flux levels, and can thus be confidently used as a guide when assessing (in)completeness levels. In particular, Fig. 4.4 shows how the Bootes field is deeper and the XFLS is shallower than all SWIRE fields, as expected due to different exposure times.

4.8 MIPS 24 μm Source Counts

While IRAC images are deeper, the completeness of our MIPS 24 μm catalogs is particularly important since a deep MIPS 24 μm catalog allows us not only to reliably associate

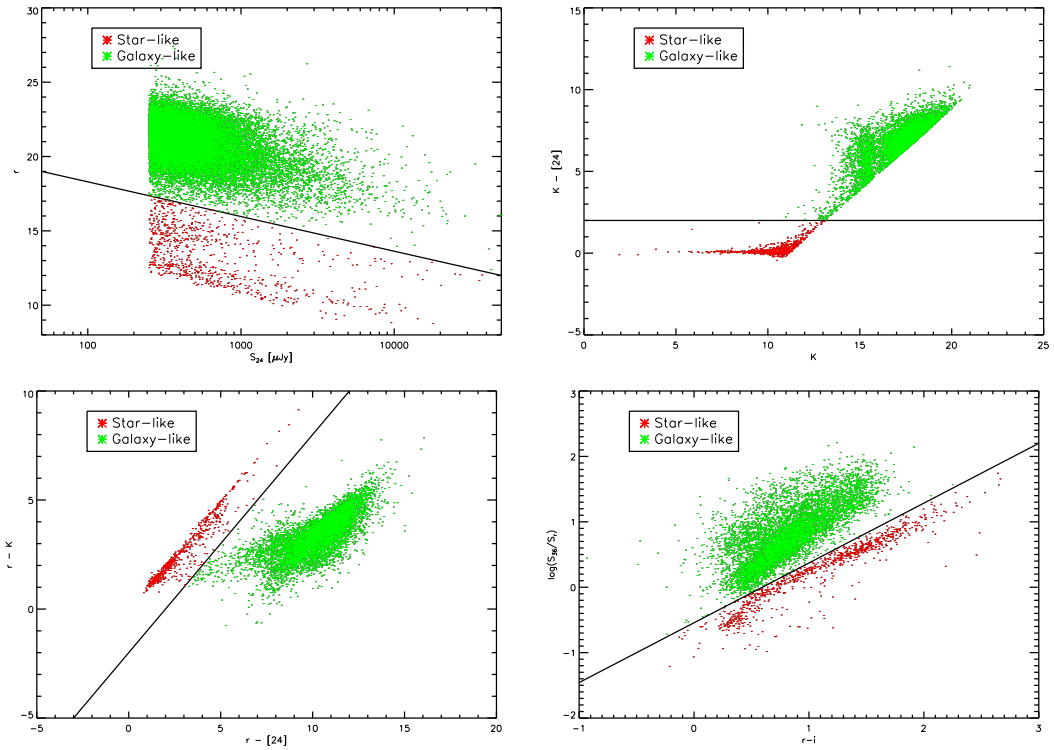

Figure 4.3: Star-Galaxy Separation using multi-band imaging. Top Left : the r vs S_{24} colour-colour plot employed (albeit using ISO 15 μm rather than MIPS 24 μm by Vaccari et al. (2005)). Top Right : the $K-[24]$ vs K colour-magnitude plot employed by Shupe et al. (2008). Bottom Left : the $r-K$ vs $r-[24]$ colour-colour plot employed (albeit using ISO 15 μm rather than MIPS 24 μm by Vaccari et al. (2005)). Bottom Right : the $r-[3.6]$ vs $r-i$ colour-colour plot employed by Rowan-Robinson et al. (2008).

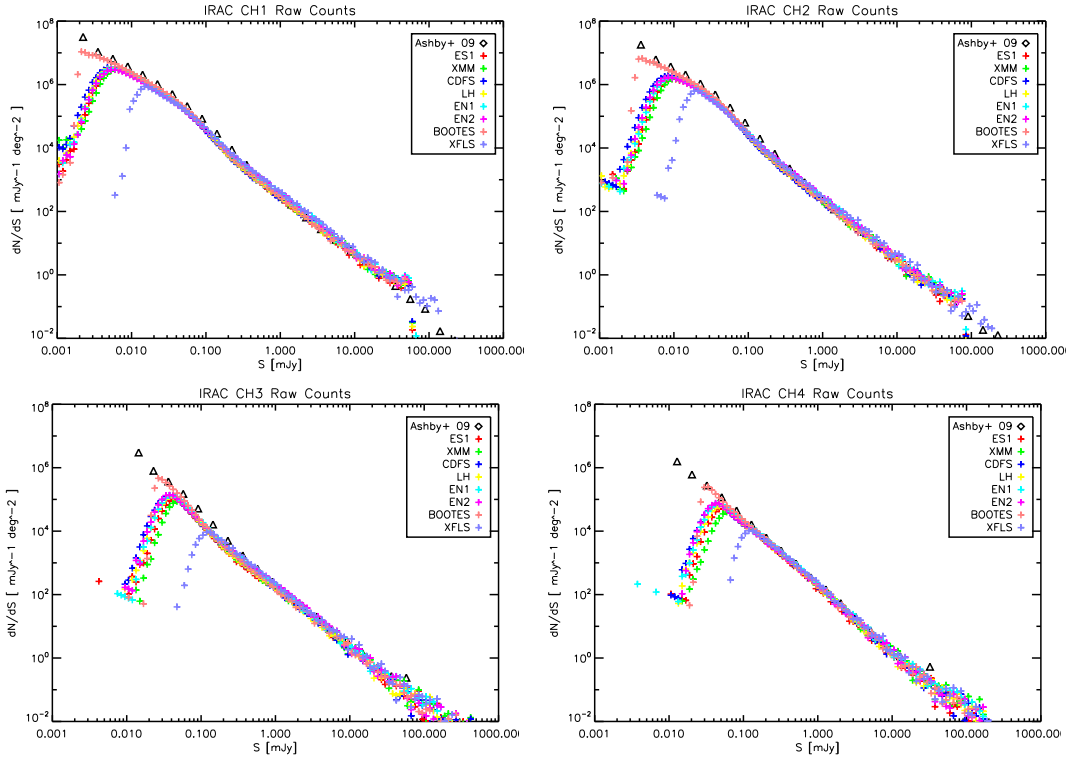

Figure 4.4: Spitzer Data Fusion IRAC Completeness : IRAC 4-band Raw Source Counts against Ashby et al. (2009).

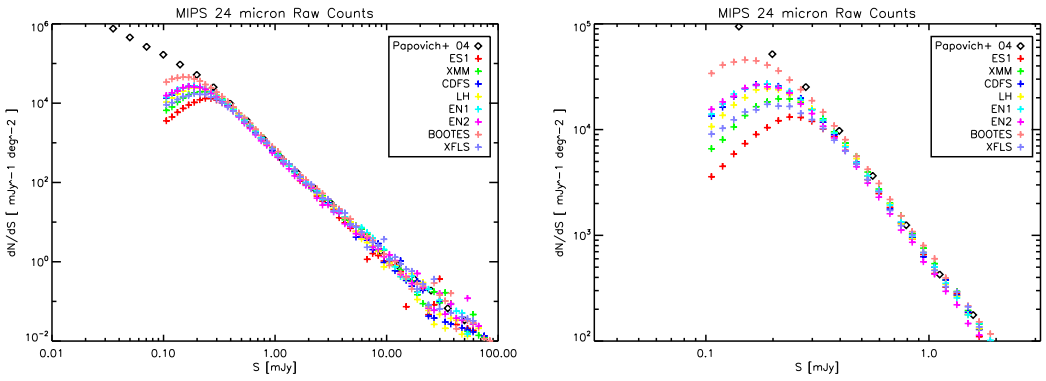

Figure 4.5: Spitzer Data Fusion MIPS 24 μm Completeness : MIPS 24 μm Raw Source Counts against Papovich et al. (2004).

MIPS 70 and 160 μm sources with their IRAC counterparts but also, most importantly, to more efficiently detect sources in Herschel maps. Fig. 4.5 illustrates how MIPS 24 μm raw source counts compare with published source counts (corrected for incompleteness). As for IRAC, this provides an indication of the Spitzer Data Fusion MIPS 24 μm completeness levels which is in reasonable agreement with the results of detailed simulations carried out in some sub-field at a few flux levels. The Bootes field is thus deeper than most SWIRE fields due to its longer exposure time, whereas ES1 (shorter exposure time), XFLS (shorter exposure time) and XMM (higher IR background levels) are all shallower than most SWIRE fields.

4.9 The SWIRE-SDSS & The Spitzer-SDSS Samples

Three SWIRE fields (LH, EN1, EN2) as well as Bootes and XFLS fully overlap with the SDSS DR7 photometric coverage, although SDSS spectroscopy is only available for part of the EN1 field. Considering only the area fully covered by all Spitzer bands, the SWIRE-SDSS and Spitzer-SDSS fields jointly cover 23 and 35 deg^2 respectively.

4.9.1 Optical Magnitude and Redshift Distribution

SDSS optical flux measurements are complete and reliable down to about $r_{AB} \sim 22.2$, and at these flux limits SDSS photometric redshifts are complete and reliable up to $z \sim 0.5$, whereas SDSS spectroscopic redshifts are available for all sources down to $r \sim 17.7$, where the vast majority of sources is at $z < 0.15$.

Fig. 4.6 and Fig. 4.7 shows the Optical Magnitude and Redshift Distribution of some MIPS-selected sub-samples, such as those used in Chap. 6 for the determination of the MIPS and SPIRE local luminosity function. Please note that, while unreliable photometric redshifts (`Photoz`) are available for fainter sources, as shown in the Figures, in our study of the MIPS and SPIRE local luminosity function in we only use the reliable photometric redshifts (`Photoz2`) for sources brighter than $r_{AB} = 22.0$ to be on the safe side.

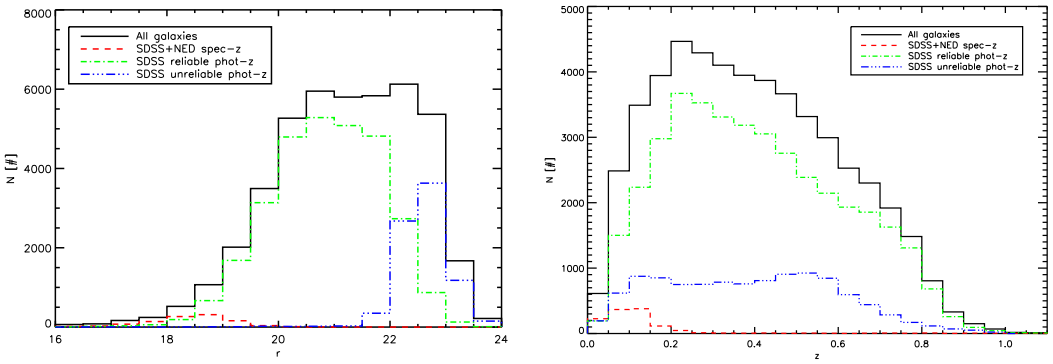

Figure 4.6: Optical Magnitude and Redshift Distribution of SWIRE-SDSS 24 μm Full Sample. Left : Magnitude Distribution of 24 μm full sample. Right : Redshift Distribution of 24 μm full sample. The cut-off observed at $r > 23$ and $z > 0.8$ is due to the SDSS magnitude limit and to the boundary conditions imposed on SDSS photometric redshifts. SDSS reliable photometric redshifts (`Photoz2`) are available for the bulk of SDSS sources brighter than ~ 22 while unreliable photometric redshifts (`Photoz`) are also available for fainter sources.

4.9.2 IRAC Colours

In the early phases of the Spitzer mission, Lacy et al. (2004) and Stern et al. (2005) first employed IRAC 4-band colours as a diagnostic tool, and particularly for the identification of AGNs. More recently, Sacchi et al. (2009) and Trichas et al. (2009) have expanded the scheme to take into account indications from MIPS 24 micron and X-ray observations respectively. Fig. 4.8 illustrates how these diagnostics can be applied to the Spitzer Data Fusion to select AGNs in a way that is in very good agreement with SDSS spectral classification.

4.9.3 SWIRE vs ELAIS (Spitzer vs ISO)

The European Large-Area ISO Survey (Oliver et al. 2000; Rowan-Robinson et al. 2004) was the largest ISO Open Time Key Project. Aimed at a first characterization of the extragalactic sky at MIR and FIR wavelengths over large areas, ELAIS was in many ways a pathfinder to the SWIRE survey. More importantly, ELAIS did provide (albeit at bright fluxes only, limited by ISOCAM sensitivity) 15 μm flux measurements which can be combined with Spitzer's IRAC and MIPS measurements to produce useful IR-based diagnostic diagrams.

Fig. 4.9 e.g. shows a colour-colour plot based on IRAC 8.0 μm and MIPS 24 μm fluxes measured by SWIRE and ISO 15 μm fluxes measured by ELAIS (Vaccari et al. 2005). The discrimination between stars and galaxies much tighter than based on optical and IRAC 3.6 μm colours, and is confirmed by the independent assessment (based on optical imaging) from the original ELAIS catalogue.

4.9.4 FIR Colours

FIR colours of extragalactic sources measured in the MIPS bands can hint at their properties based on their FIR spectral index, in a tradition dating back to IRAS, and early Spitzer studies such as those by Frayer et al. (2006a) focusing on the XFLS field thus relied on such colours and on the sparse availability of spectroscopic redshifts. Fig. 4.10 and Fig. 4.11

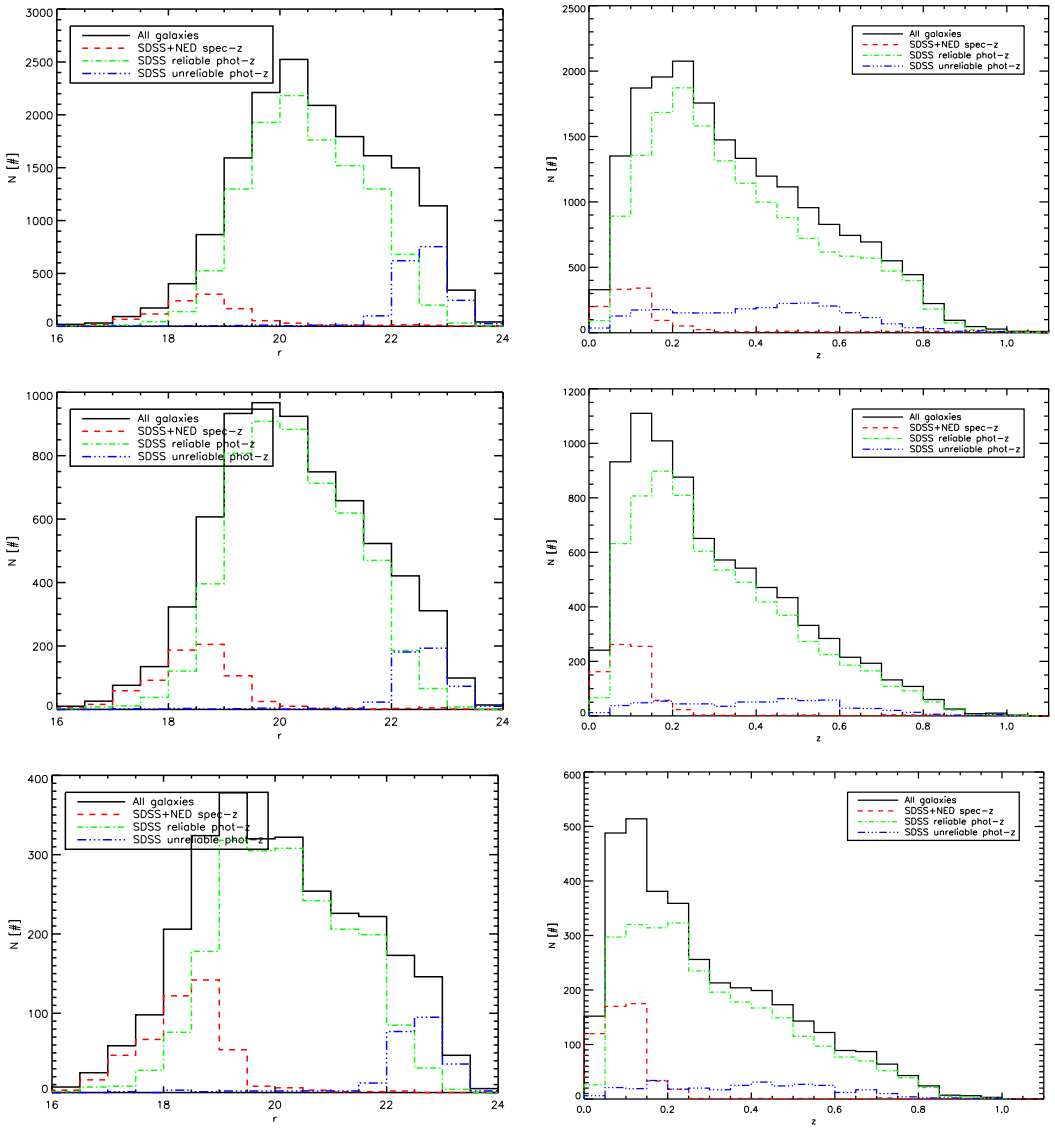

Figure 4.7: Optical Magnitude and Redshift Distribution of SWIRE-SDSS 24/70/160 μm Bright Samples. Left : Magnitude Distribution of Bright Samples. Right : Redshift Distribution of Bright Samples. Top : $S_{24} > 250$ μJy. Middle : $S_{70} > 10$ mJy. Bottom : $S_{160} > 50$ mJy. MIPS 24/70/160 μm Bright Samples are used in Chap. 6 to determine the MIPS Local Luminosity Function. SDSS reliable photometric redshifts (`Photoz2`) are available for the bulk of SDSS sources brighter than $r \sim 22$ while unreliable photometric redshifts (`Photoz`) are also available for fainter sources.

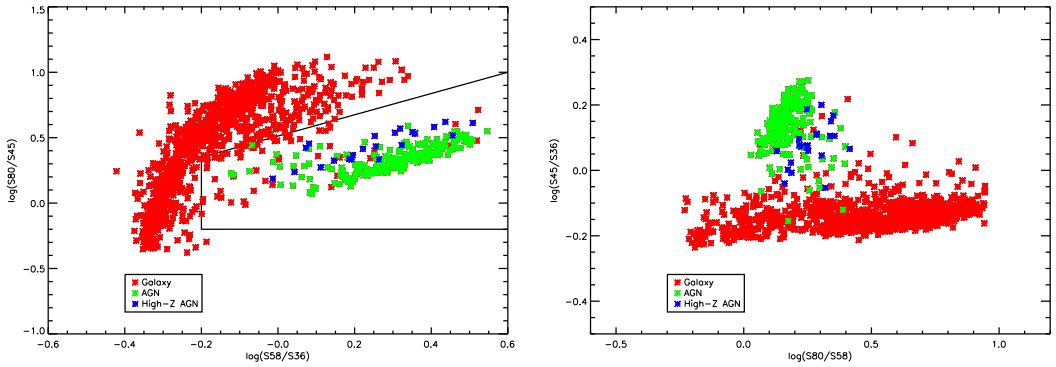

Figure 4.8: IRAC Colours for SWIRE-SDSS Spectroscopic Sample. Left : Lacy et al. (2004)'s Plot. Right : Stern et al. (2005)'s Plot. Label with SDSS Spectroscopic Classification (TBD).

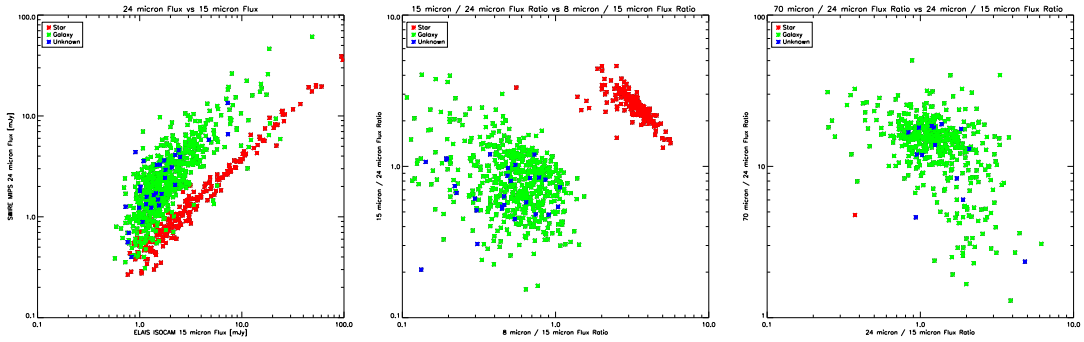

Figure 4.9: ELAIS-SWIRE-SDSS Joint Sample. Classification is from the ELAIS FA catalogue by Vaccari et al. (2005)

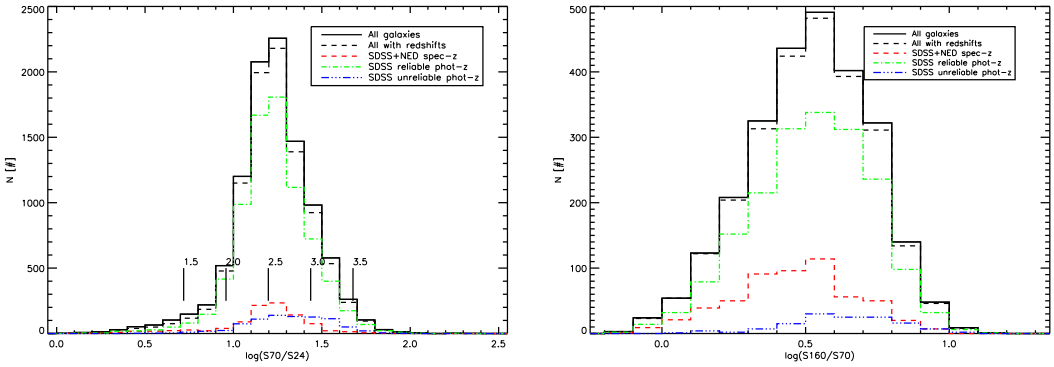

Figure 4.10: FIR Colours of SWIRE-SDSS MIPS Sample. Left : S_{70}/S_{24} flux ratio distribution. Left : S_{160}/S_{24} flux ratio distribution.

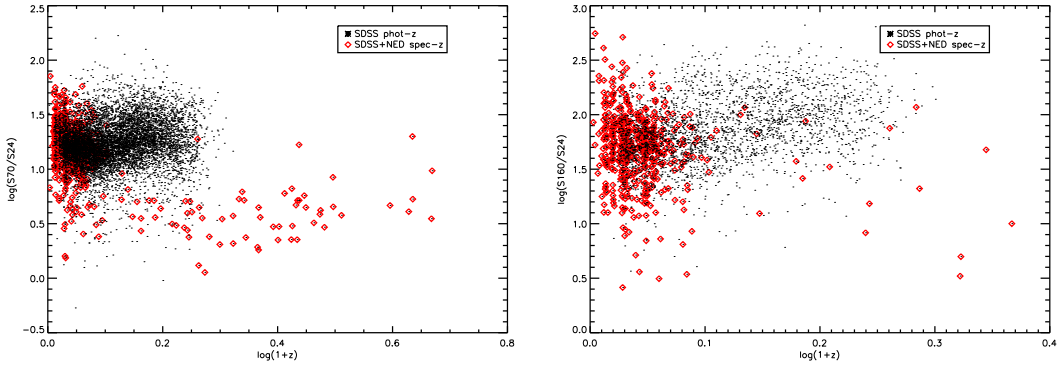

Figure 4.11: FIR Colours of SWIRE-SDSS MIPS Sample. Left : S_{70}/S_{24} flux ratio as a function of redshift. Right : Left : S_{160}/S_{24} flux ratio as a function of redshift. Spectroscopic redshifts measured in the $0 < z < 0.15$ range are mostly from SDSS while higher spectroscopic redshifts are mostly from NED.

shows the overall distribution of the S_{70}/S_{24} and S_{160}/S_{24} flux ratios and their trend with spectroscopic redshifts, confirming the average values and the trends seen by Frayer et al. (2006a).

4.9.5 FIR Luminosities

Using SDSS photometric and spectroscopic redshifts, we can compute the luminosity distances of our sources and thus their *observed-frame* monochromatic luminosities, as shown in Fig. 4.12. Thanks to the wide wavelength coverage provided by the Spitzer Data Fusion, however, we can carry out a SED fitting analysis for all sources (as done in Chap. 5) to determine the k-correction and thus obtain their *rest-frame* luminosities which can in turn be used to estimate the evolution of their luminosity function (as done in Chap. 6 and Chap. 7).

4.10 Applications of the Spitzer Data Fusion

The Spitzer Data Fusion was developed within the SWIRE and HerMES collaborations to meet the need for a well-documented multi-wavelength FUV-to-FIR catalogue of Spitzer-

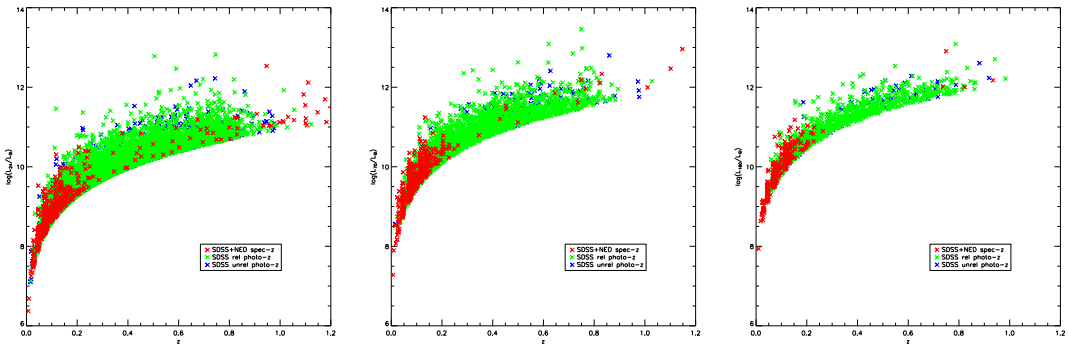

Figure 4.12: FIR Luminosity of SWIRE-SDSS MIPS Sources. Left/Centre/Right : $L_{24}/L_{70}/L_{160}$ as a function of redshift.

selected sources. It has thus filled a gap preventing the effective and timely exploitation of Spitzer and Herschel cosmological surveys, and in so doing ihas produced a number of useful applications beyond the main scope of this Thesis. These range from the possibility to carry out a panchromatic FUV-to-FIR SED fitting analysis of a very large sample of Spitzer sources (see Chap. 5), including those detected by the Spitzer Exploratory Reference Volume Survey (PI : Lacy & Farrah) by Mauduit et al. (2012), from the optimal astrometric registration of Herschel map-making by Levenson et al. (2010) to the optimization of Herschel source extraction by Roseboom et al. (2010), from the determination of the Local Luminosity Function (see Chap. 6) to the study of its evolution up to high redshift (see Chap. 7), thus fully exploiting Spitzer and Herschel data as well as ancillary data in a homogeneous manner.

Chapter 5

Using SED Fitting to Disentangle the Cosmic Star Formation History

Galaxies emit electromagnetic radiation over the full wavelength range and the distribution of their emitted energy over wavelength is called their Spectral Energy Distribution (SED). SEDs are our primary source of information about the properties of unresolved galaxies. Indeed, the different physical processes occurring in galaxies all leave their imprint on the global and detailed shape of the spectrum, each dominating at different wavelengths. Detailed analysis of the SED of a galaxy should therefore, in principle, allow us to fully understand the properties of that galaxy. SED fitting is thus the attempt to analyze a galaxy SED and to derive one or several physical properties simultaneously from fitting models to an observed SED.

Substantial progress in many areas that affect SED fitting has recently been made. A major development in the last decade has been the advent of new observing facilities and large surveys at all wavelengths of the spectrum, enabling astronomers for the first time to observe the full SEDs of significant samples of galaxies over a large redshift range at wavelengths from the X-rays to the radio. In particular Herschel has been giving a great contribution in this sense, allowing, for the first time, the sampling of the FIR and SMM part of the spectrum at multiple wavelengths. At the same time, tools and models have been created that aim to extract the complex information imprinted in the SEDs (Walcher et al. 2011) in an increasingly efficient and reliable manner.

In this Chapter we introduce the concept of SED fitting analysis and how this technique can be used to disentangle the SFH of galaxies. In particular we explain how we exploit the Spitzer Data Fusion described in Chap. 4 to perform the SED fitting analysis on the samples we then use to evaluate the MIPS and SPIRE luminosity functions in Chap. 6 and Chap. 7: keeping the (known) redshift of any given source fixed, we can determine the k-correction and monochromatic rest-frame luminosities (defined in Sec. 2.1.4) at every wavelength of interest on the basis of the best-fit SEDs, the integrated rest-frame IR bolometric (8-1000 μm) luminosity and thus estimate the SFR of galaxies using various SFR tracers (as explained in Sec. 5.4). To check the reliability of our results we use two different template-based SED fitting codes, `Hyperz` (Bolzonella et al. 2000) and `LePhare` (Arnouts et al. 1999), detail their use and compare their results in our applications.

5.1 SED Fitting Recipes

Fitting the SEDs of galaxies is a well-developed technique that has matured significantly in the last decade. Model predictions (hereafter often simply models) and fitting procedures (hereafter often simply codes) have improved significantly over this time, attempting to keep up with the vastly increased volume and quality of available data on the SEDs of galaxies. SED fitting can be used effectively to derive not only a redshift estimate but also a range of physical properties of galaxies, such as stellar masses, star formation rates, dust masses and metallicities, with care taken not to over-interpret the available data. In general the major ingredients of any SED fitting recipe are the SED models able to describe galaxy emission from the ultraviolet to the infrared, starting from galaxy formation to the production and re-processing of the radiation from the assembled galaxies. In the following Sections we discuss the main ingredients of a typical SED fitting analysis performed using template-based SED fitting codes.

5.1.1 Synthetic *vs* Observed SEDs

In general, performing a template-based SED fitting procedure means to compare the observed SEDs of galaxies to those which would be returned from a set of reference spectra, using the same observing conditions and photometric system. Then the result for a given object corresponds to the best fit of its observed SED by the set of synthetic spectra, in general through a standard χ^2 minimization procedure.

Thus, in order to perform the SED fitting analysis, apart from the complete set of filters and detectors transmission (that we already described in the previous Chapter), we need a reliable set of synthetic SEDs, or "templates" to be compared with our observed SEDs. These template SEDs can be either theoretical, i.e. based on numerical simulations, or phenomenological, i.e.based on photometric observations of ""typical" sources.

Constructing a reliable set of **synthetic SEDs** is far from easy and a detailed description of this process is beyond the scope of this Thesis; for this reason we limit the discussion to the main ideas behind the theory. Galaxies emit across the electromagnetic spectrum. Excluding those galaxies dominated by an accreting supermassive black hole at their nucleus (AGN), the ultraviolet to infrared spectra of all galaxies arises from stellar light, either directly or reprocessed by the gas and dust of the surrounding interstellar medium (ISM). Thus the UV-to-IR spectral energy distribution or SED contains a large amount of information about the stars of a galaxy, such as the stellar mass to light ratio, and the surrounding ISM, such as the total dust mass. However, to extract such information, models are necessary in order to connect physical properties of the galaxy with the observed SED. The starting point should be the stellar spectrophotometric model, then the transfer of the radiation of these stars in a galaxy through the ISM, and finally how to connect these with the larger picture of galaxy formation and evolution. Once modeled all these processes and created a set of synthetic templates that should represent a number of "true" galaxies with different physical properties, it is possible to perform the SED fitting procedure by comparing these synthetic SEDs to photometric observations. In this way the result of the SED fitting analysis gives us not only the best fit SED for each source, allowing us to classify sources according to their morphologies, e.g. as Ellipticals, Spirals, etc, but also all physical information contained in the model that has constructed the synthetic SED itself, like its Stellar Mass and its Star Formation History.

A major development in the last decade has been the advent of new observing facilities and large surveys at all wavelengths of the spectrum, enabling astronomers for the first time to observe the full SEDs of galaxies from the FUV to the FIR, from the local universe out to redshifts beyond 6. Indeed, the "true SED" of a galaxy as defined in the models considers the sum of all photons emitted from inside the volume defining the galaxy. To make the **observed SED** of a galaxy, however, this "true SED" is then filtered through the spectral response curves of the instruments and is redistributed spatially over the point spread function (PSF). Additionally, the measurement process not only adds noise, but also makes it necessary to combine flux measurements from different instruments.

The observed SED can thus be identified as a series of wavelengths and associated fluxes; this is only a simplification of the fact that the measurement process convolves the true SED with a spectral response curve, yielding a transmitted flux at an effective wavelength. In photometry, the response curve is much broader than a typical Gaussian curve like in spectroscopy and therefore needs to be represented with more care, i.e. tabulated as a response function. The response function in turn depends on the detector quantum efficiency, the instrument transmission and the filter in use. For this reason we spent so much time constructing, along with the Spitzer Data Fusion, a complete list of filters carefully described and ready to be ingested in a SED fitting code (see App. B). Another important issue when dealing with multi-wavelength photometric data collection is whether what is measured in each band is actually physically the same; this is what we refer to as "aperture matching problem", and is discussed in Sec. 5.1.3.

The construction of multi-wavelength SEDs is thus a complex and rich subject. Since the main focus of our work is the description of the FIR properties of our galaxies from a statistical point of view to enable the study of the evolution of their Luminosity Function, we are not particularly interested in the detailed description of the full physical properties of each detected source as returned by their rest-frame UV/Optical/NIR emission, and instead we fix our attention on being able to reproduce the rest-frame emission of galaxies at longer (MIR/FIR/SMM) rest-frame wavelengths, where galaxy emission often peaks and is closely associated with star formation activity. For this reason we did not construct a new sample of theoretical SEDs to fit our data, but we based our analysis on existing sets of SED phenomenological templates defined and adopted by previous IR studies (see Sec. 5.2.1).

5.1.2 Reddening and Dust Extinction

Studies of high-redshift galaxies and star formation obscured by dust have shown the importance of extinction due to dust in the "high-redshift" universe. Therefore, another effect that must be taken into account is dust extinction, or reddening, produced inside the galaxies themselves.

Interstellar dust has been a field of constant inquiry since it was first realized that an obscuring material existed between the stars and a large body of research exists on the composition, shape and distribution of dust. Most of our understanding of interstellar dust has come locally, from observations within our own Galaxy and the Magellanic clouds, along with theoretical and experimental laboratory work. It is generally accepted that the grains are of three different kinds; graphitic/amorphous carbon grains, amorphous silicate grains, and polycyclic aromatic hydrocarbons (PAHs), which may or may not be an extension of the carbonaceous grains (see Sec. 3.2 and Tab. 3.1). The former two were found to reproduce

the observed extinction along different lines of sight within our galaxy (Mathis et al. 1977), while the latter were added to explain unidentified emission bands in the MIR (Leger & Puget 1984). Other forms of dust have been suggested, such as SiC (Treffers & Cohen 1974), and ice is expected to form on grains in the coldest environments such as deep in molecular clouds, but generally only the previous forms are considered in the SED modeling of galaxies.

The effect of dust opacity is very important at the short wavelengths of the SED and strongly decreases with the increasing of wavelength, and in the FIR only dust emission needs to be considered. These effects on the optical-UV light are often described by two parameters: the reddening and the total obscuration. Reddening is the wavelength dependence of dust effects, including features, and takes account of the fact that shorter wavelength photons are more readily scattered and absorbed by dust. This is often parametrized by the colour excess $E(B-V)$ or the Balmer decrement H_α/H_β. The total obscuration is a measure of the total light absorbed or scattered out of our line-of-sight by dust either bolometrically or in a single band and can be considered the normalisation of the reddening. This is generally parametrized as $A(V)$. For relative measures correcting only for reddening is sufficient, however for absolute quantities the total obscuration must also be taken into account. This is especially important when the reddening is close to flat, i.e. only small effects by dust are visible on the spectrum.

For individual stars in the Milky Way, the Large and Small Magellanic clouds, extinction laws have been measured (Cardelli et al. 1989). When considering a galaxy as a whole, it must be taken into account that stars reside at different optical depths, depending on whether they lie on the side of the galaxy facing the observer or averted from the observer, and that the stellar light can be scattered into the observer's line-of-sight as well as out of it. Additionally, stellar populations of different age will have different extinction optical depths, and this extinction might have a different wavelength dependence.

These issues lead to the concept of *attenuation*, where the complexity of the actual star-gas geometry is wrapped into a single attenuation law, now not applied individually to each star in the galaxy, but applied to the full spectrum of the galaxy. Using an attenuation law, the dust obscuration of stellar light is expressed through a screen approximation, as if the dust was lying between us and the stellar population of the galaxy, with a wavelength-dependent reddening law (a_λ). The total amount of attenuation then depends only upon the thickness of the screen ($\Delta\tau$),

$$I(\lambda)_{obs} = I_{star}(\lambda)e^{-a_\lambda \Delta\tau} \tag{5.1}$$

The attenuation law was derived empirically for starburst galaxies by Calzetti et al. (1994) and Calzetti (1997), who fit the law with a simple polynomial as a function of $1/\lambda$. They found a law much greyer than the extinction laws of the Milky Way and LMC demonstrating the effects of geometry and mixing compared to simple extinction. Generally a simple power-law , $a_\lambda \propto \lambda^{-0.7}$, is able to reproduce the observed effective attenuation in galaxies (Charlot & Fall 2000). However, a simple attenuation law cannot account for differential geometries and star formation histories within and between galaxies. This can be seen with the higher optical depths observed for nebular emission lines relative to the underlying stellar continuum, indicating that the stars and gas that give rise to the lines and to the continuum see different amounts of dust (Calzetti et al. (1994); Calzetti (1997)) through the galaxy.

In the literature we can find many different reddening laws studied by different authors each describing a particular environment (galaxy, clusters, etc.). When performing the SED

fitting procedure it is important to select the reddening law that better describes the data. For this reason every SED fitting code stores into libraries many different reddening laws between which it is possible to choose the most suitable to reproduce the observed photometry. This is particularly important when we try to reproduce the complete shape of the SED from the UV to the IR and, in other words, the energy balance of the emitted and absorbed light in a galaxy. The FIR part of the SEDs of galaxies provide insight into the energetic processes taking place in the sources. Emission beyond a few microns from quiescent or star-forming galaxies is dominated by dust reradiating the stellar energy absorbed at UV/Optical wavelengths. Thus, a fully self-consistent description of the radiative processes in the IR regime needs to account for both the dust composition and the nature of the heating sources.

5.1.3 Aperture Matching

As we already mentioned, an important issue when assembling a multi-wavelength observed SED and performing a multi-wavelength SED fitting procedure is to make sure that the fluxes we measure at different wavelengths refer, at least within the measurement errors, from the galaxy and the same sky area. This is not trivial as public catalogues often measure fluxes in different ways and the instruments they rely upon for their measurements often have intrinsically very different spatial resolutions. This is of particular importance when every ancillary catalogue contains many flux measurements for each band, as is the case for our Spitzer Data Fusion, in which case it is easy to mismatch flux measurements in different bands.

Due to either the seeing of ground-based observations or the diffraction of the ground-based or space-based telescope itself, the flux from a point-like source is re-distributed over a point spread function (PSF) of a width that typically depends on the observation and on the wavelength. Moreover, galaxies are intrinsically extended and their morphology may depend on the wavelength at which they are observed. One of the main problems in the process of matching is the size of the PSF, which is typically narrowest at optical wavelengths, while UV and IR PSFs are broader (Walcher et al. 2011).

Another problem related to the resolution of the telescope are aperture biases. A rather simple manifestation of this is that objects that are further away will be seen as smaller on the sky. Therefore, in order to construct comparable samples at different redshifts, one must in principle adapt the size of the extracted aperture to the same physical size. A more complicated problem is the definition of the "total light" from an object. Indeed, the surface brightness profiles (SBP) of galaxies usually extend much beyond the observational surface brightness threshold. Different strategies have been developed to avoid these biases, such as either integrating over a full model for the SBP or simply using specific apertures to integrate the light only inside some physical radius, but each method has its own problems. One of these is that galaxies have different intrinsic morphology in different bands, thus complicating the application of consistent procedures, even when using data with comparable angular resolution.

Aperture matching between different bands is then crucial to the success of template fitting, and particular to the accuracy of photometric redshift estimates. For distant galaxies photometry in a 2 or 3 arcsec aperture will give the integrated light from the whole galaxy. For nearby galaxies photometry in the same aperture would be dominated by light from the central regions of the galaxy and might comprise only a few % of the integrated emitted light, but there are several options for estimating the integrated light. For example, optical and near

infrared catalogues generally provide Kron and Petrosian magnitudes, as originally defined by Kron (1980) and Petrosian (1976) (see also Graham & Driver (2005)). These integrated magnitudes are derived essentially by a curve of growth fitted to photometry derived in a series of apertures of different sizes. However in practice using integrated magnitude estimates for each photometric band to derive the integrated SED gives poor results for photometric redshifts. This is presumably because of the uncertainty introduced by the process of estimating the integrated magnitude, primarily because of different contributions of sky photon noise. A much more successful option is to start from photometry derived in a single small aperture in each band, and then apply an aperture correction derived in a single chosen band to all the bands; this is the approach followed by Rowan-Robinson et al. (2008) and González-Solares et al. (2011).

More in general, in template fitting methods such as `Hyperz` and `LePhare`, it is natural to try and use flux measurements providing a reliable estimate of the integrated emitted light, so that near and distant galaxies can be fitted in the same way. This also has the benefit that derived properties such as luminosity, star-formation rate, stellar mass and dust mass have a physical meaning for the galaxy as a whole. This is not always possible, however, since deep surveys are often hampered by confusion' making aperture fluxes increasingly more accurate than integrated fluxes with increasing depths, and a trade-off must then be established between these two needs.

Aperture matching is therefore very important, and particularly so in the UV/Optical/NIR where relative photometric errors can reach below 1 %, to obtain good photometric redshifts and physical properties, but much less so at longer wavelengths, where confusion and related flux uncertainties play the biggest role. In this work we're mostly interested in fitting the long-wavelength portion of the SEDs of galaxies with known (photometric or spectroscopic) redshifts, and thus we did not explore aperture matching in great detail. However in the Spitzer Data Fusion construction we tried to collect the largest number of flux estimates for each band in order to allow extensive testing and optimization of the aperture matching between different surveys. In this way the approaches adopted, between others, by the previous cited authors, can be further enabled by the Spitzer Data Fusion, as shown by preliminary work by Rowan-Robinson et al. (in prep).

5.2 SED Fitting Codes

The Spitzer Data Fusion together with the collection of ancillary information such as the transmission curves and Vega-AB conversion factors of all spectral filters, represents an excellent starting point to perform a panchromatic (FUV to FIR) SED fitting analysis on a large number of sources in Spitzer/Herschel joint wide fields. In this work we extensively exploit the information collected in the Spitzer Data Fusion to characterize the infrared properties of galaxies by fitting the SEDs of each source from the FUV to the FIR using two within the increasing number of template-based SED fitting codes which have been made publicly available: `Hyperz` (Bolzonella et al. 2000) and `LePhare` (Arnouts et al. 1999).

These codes both provide a complete framework to carry out a template-based SED fitting analysis, photometric redshift and stellar mass estimates, even if `LePhare` is much faster and flexible, as we will describe later. In our work we exploit three major samples, two in the low-redshift regime (see Chap. 6 for details) and one at high redshift (see Chap. 7 for details).

We started using `Hyperz` on the "low-z" sample and, after making sure the results of the two codes were in reasonable agreement when applied to these samples, we then proceeded to use `LePhare` on the "high-z" data due to the advantages it offered.

In this Section we describe the two SED fitting codes and how we set them up to perform the SED fitting on our samples.

5.2.1 `Hyperz` Photometric Redshift Code

`Hyperz` has been written by Bolzonella et al. (2000) and is a `Fortran 77` template-based SED fitting code; the code and the documentation are available at `http://webast.ast.obs-mip.fr/Hyperz/`, even though the much-improved latest version of the code can be obtained from the authors since the code is always under construction. The original version of the code has been written to evaluate the photometric redshift of galaxies working in the UV/Optical/NIR regime; then it has been modified to work also in other wavelength regimes and when the redshift of the sources is known. In this last case it can compute also stellar masses and other physical properties of galaxies fitting their full FUV-to-FIR photometry.

The public package of *Hyperz* provides the main code `Hyperz` and the modified version `Hyperzmass` and `Hyperzspec`. The first two codes fit in details the UV/Optical/FIR SED and give respectively the photometric redshift and the stellar masses of the sources along with other physical properties of the galaxy, while the third incarnation performs a simple χ^2 minimization procedure to the global observed SEDs which is particularly suitable to give the overall shape of the SEDs in a very straightforward way. Both `Hyperzmass` and `Hyperzspec` work keeping fixed the redshift of the sources. A weak point of the code is that there is no possibility to perform the analysis on photometric redshift, stellar masses and the fit to the whole SED in one go; `Hyperz`, `Hyperzmass` and `Hyperzspec` are all identified by the general name `Hyperz` since they use the same general framework, but they are actually three stand alone codes to be run separately (this is one of the reason why we then moved to the use use `LePhare` for more sophisticated studies).

`Hyperz` performs a standard χ^2 minimization procedure to find out the best fit of the photometric observed SEDs by the set of template spectra

$$\chi^2(z) = \sum_{i=1}^{N_{\rm filters}} \left[\frac{F_{{\rm obs},i} - b \times F_{{\rm temp},i}(z)}{\sigma_i} \right]^2 ,$$

where $F_{{\rm obs},i}$, $F_{{\rm temp},i}$ and σ_i are the observed and template fluxes and their uncertainty in filter i, respectively, and b is a normalisation constant.

It does not use any *Bayesian* marginalization to drive the code to the solution, even if Benítez (2000) demonstrated that introducing an *a priori* probability on the possible results would significantly improve the dispersion of $z_{\rm phot}$. The authors of `Hyperz` decided to not introduce such type of information, because the application of the *Bayesian* technique can introduce spurious effects in particular studies. However, they agrees that the method can be regarded with interest when the goal is some specific application or when one is dealing with poor data, in such a way that the introduction of hints allows to obtain useful results. Alternatively, the photometric redshift estimate can be safely improved introducing the Bayesian inference when prior information is not related to the photometric properties of sources. Examples of such priors that could be combined with the $z_{\rm phot}$ technique are the morphology

or the clues inferred from gravitational lensing modeling. In these cases, the user can easily introduce the interesting prior in the Fortran 77 code.

The package `Hyperz` provides a number of observed SEDs, a set of filters' transmissions, spectral synthesis models (e.g. Bruzual & Charlot (2003)) as well as 5 different reddening laws; the user can choose between these inputs already ingested in the procedure or add its own input list of models. This option is completely true for the template SEDs and the filters' transmissions; changing them along the process is really easy, since it means to simply add them in the list already implemented in the code (paying attention to maintain the same format of the file, e.g. `Hyperz` assumes that the transmission of the filters are all store in units of Energy, see Sec. 5.2.2 for more details on this point). It is not equally easy to change or add an original spectral synthesis model or reddening law. The code does not allow the user to add its own original inputs for these two quantities since this would mean change major part of the code itself. The reddening law are reported in Fig. 5.1 and are : (1) Allen (1976) for the Milky Way (MW); (2) Seaton (1979) fit by Fitzpatrick (1986) for the MW; (3) Fitzpatrick (1986) for the Large Magellanic Cloud (LMC); (4) Prevot et al. (1984) and Bouchet et al. (1985) for the Small Magellanic Cloud (SMC); (5) Calzetti et al. (2000) for starburst galaxies.

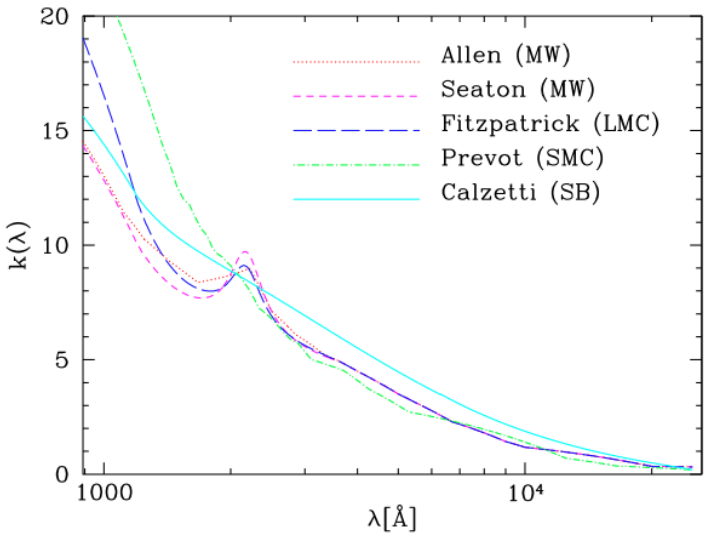

Figure 5.1: Extinction curves provided by `Hyperz`. The different laws are normalised to $k(B) - k(V) = 1$.

The code estimates, according to the input data, the input value A_V, corresponding to a screen model with $f_{obs}(\lambda) = f_{int}(\lambda)10^{-0.4A_\lambda}$ where f_{obs} and f_{int} are the observed and the intrinsic fluxes, respectively. The extinction at a wavelength λ is related to the colour excess $E(B-V)$ and to the reddening curve $k(\lambda)$ by

$$A_\lambda = k(\lambda)E(B-V) = \frac{k(\lambda)A_V}{R_V}\,,$$

with $R_V = 3.1$ except for the Small Magellanic Cloud ($R_V = 2.72 \pm 0.21$) and Calzetti's law ($R_V = 4.05 \pm 0.80$).

Once the code is install on the machine, the user has to set the *parameters* file to run the code. This setting file stores all the specific input read by the code (e.g., the photometric

catalogue, the template of SEDs, the filters list, the reddening law, etc..) and the required output. In the next section we describe how we set the *parameters* file for our purpose; for more details on the specific of the code we point to the on line documentation.

SED Fitting with `Hyperz`

In our work we exploit the `Hyperz` code (more specifically `Hyperzspec`) to fit the SEDs of our "low-z" samples of galaxies, keeping fixed the redshift of the sources, with a similar approach used by Bolzonella in her recent work. As SED templates we use the SWIRE templates by Polletta et al. (2007) and their slightly modified version by Gruppioni et al. (2010) listed in Tab. 5.1 and plotted in Fig. 5.2.

The SWIRE templates combines the SEDs of galaxies with those of AGNs. The library contains 20 templates including 1 elliptical, 7 spirals, 3 starbursts (SBs), 6 AGNs, and 3 composite (SBs+AGN) templates covering the wavelength range between 1000 Å and 1000 μm. The elliptical, spiral and starburst templates were generated with the GRASIL code (Silva et al. 1998). Templates of moderately luminous AGNs, representing Seyfert 1.8 and Seyfert 2 galaxies, were obtained by combining models, broad-band photometric data, and ISO-PHT-S spectra of a random sample of 28 Seyfert galaxies. Four additional AGN templates represent optically-selected QSOs with different values of infrared/optical flux ratios and one type 2 QSO. The QSO1 templates are derived by combining the SDSS quasar composite spectrum and rest-frame infrared data of a sample of SDSS/SWIRE quasars divided in three groups, all, and the 25% brightest and 25% weakest measurements per rest-frame wavelength bin. The type 2 QSO template (QSO2) represents the SED of the red quasar FIRST 013435.7-093102 (Gregg et al. 2002). The composite (AGN+SB) templates are empirical templates that well reproduce the SEDs of the following sources: the heavily obscured BAL QSO Mrk 231 (Berta 2005), the Seyfert 2 galaxy IRAS 19254-7245 (Berta et al. 2003), and the Seyfert 2 galaxy IRAS 22491-1808 (Berta 2005).

We add to this sample the ones slightly modified versions published during the Herschel PACS Science Demonstration Phase by Gruppioni et al. (2010). From those first PACS observation it was clear that in $\sim 10-12\%$ of cases the observed SEDs were very well reproduced by the Polletta et al. (2007) templates over the entire UV/Optical/NIR/IRAC range, while they were systematically underestimated in the MIPS/PACS range. In these cases, the PACS 100 and 160 μm data in flux density were always higher by up to a factor of ~ 4 than the template at the same $\lambda's$. This happens mainly for the Seyfert2/1.8 templates (for about 40% of the PEP sources fitted by these SEDs) and in less frequent cases also for the spiral ones. For these reason they modified the original SWIRE templates adding the followings to the starting sample: three new templates with a rest-frame 0.1-15 μm spectrum similar to the Seyfert2/1.8/Sdm of Polletta et al. (2007), but with a higher FIR bump, obtained by averaging together (in λ-bins) the observed rest-frame SEDs (normalised to Ks band) exhibiting an excess in the FIR and fitted by the same template; three new SB templates obtained by interpolating between the Sdm and the SB SED NGC6090 (to fill the large gap existing in the library between spirals and SBs).

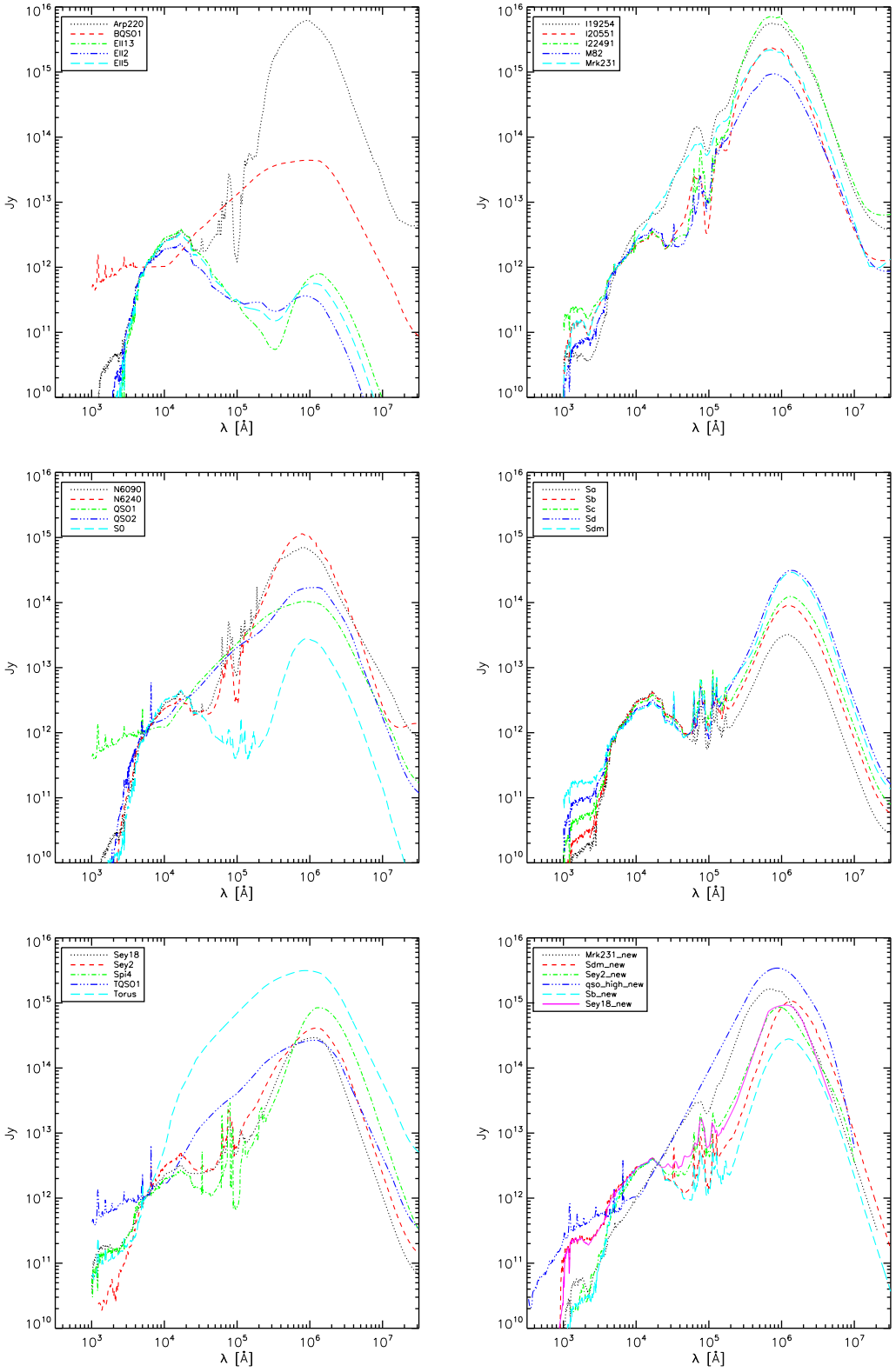

Figure 5.2: Templates used in the SED fitting procedure: panel 1 to 5, (Polletta et al. 2007), low right panel, (Gruppioni et al. 2010)

Index	SED CLASS	Reference
01	Ell13	Polletta+07
02	Ell5	Polletta+07
03	Ell2	Polletta+07
04	S0	Polletta+07
05	Sa	Polletta+07
06	Sb	Polletta+07
07	Sc	Polletta+07
08	Sd	Polletta+07
09	Sdm	Polletta+07
10	Spi4	Polletta+07
11	N6090	Polletta+07
12	M82	Polletta+07
13	Arp220	Polletta+07
14	I20551	Polletta+07
15	I22491	Polletta+07
16	N6240	Polletta+07
17	Sey2	Polletta+07
18	Sey18	Polletta+07
19	I19254	Polletta+07
20	QSO2	Polletta+07
21	Torus	Polletta+07
22	Mrk231	Polletta+07
23	QSO1	Polletta+07
24	BQSO1	Polletta+07
25	TQSO1	Polletta+07
26	Sb	Gruppioni+10
27	Sdm	Gruppioni+10
28	Sey2	Gruppioni+10
29	Sey18	Gruppioni+10
30	Mrk231	Gruppioni+10
31	qso_high	Gruppioni+10

Table 5.1: List of the SED used to perform SED fiitting analysis.

The templates are written as text files containing two columns: λ in Å and flux density in μJy. Since they were produced by different authors and they occasionally displayed a jagged appearance at the FIR and SMM wavelengths we are particularly interested in, we resampled them so as to have the same sampling in wavelength and smoothed them, without changing their overall begavior, in the FIR and SMM. We find that this set of templates well describes our sample of galaxies as explained below. As a further check we added to this set other templates like the ones published by Rowan-Robinson et al. (2010) or by Rieke et al. (2009) and we rerun the SED fitting code to see if different templates would have been preferred to the one in the Polletta et al. (2007)+Gruppioni et al. (2010) sample; this does not happen and the best-fit solutions are better, in terms of χ^2 of the fit, by using our first choice of templates.

As reddening law we choose the one by Calzetti et al. (2000) since it is more suitable for

SBs galaxies, and thus for IR-luminous galaxies like our sources, and we let the code explore the range of A_V between 0 to 4, to be as general as possible.

The code needs an input catalogue as follows: ID (a unique identifier for each source), $mag1$, $mag2$..., $magN$ (the source's observed AB apparent magnitudes in each band 1, 2,..N), $mag1_{err}$, $mag2_{err}$..., $magN_{err}$ (the source's magnitude error), Redshift. The error of the magnitudes are used by the code to explore the range of possible magnitudes that best fit the data. `Hyperz` does not allow the user to give different weights to different bands, while this could be important when the interest of the user is to fit a particular range of wavelengths with respect to others, especially when the data cover a wide range of wavelengths and the intrinsic photometric quality is inhomogeneous across different bands in a way that is not conveyed by the formal photometric errors. In such cases the user can modify the magnitude errors so that the codes indirectly gives an higher weight to a particular wavelength regime with respect to another one. Since our purpose is to well constrain the FIR/SMM peak of the SED even though the χ^2 is evaluated over the full wavelength range, we have no interest in a very good fit to the optical part of the spectrum and we want to avoid the case in which the code would prefer a certain best fit mostly based on the optical part of the spectrum since . We therefore set the photometric errors to be 0.5 magnitudes in all bands apart from two bands which best guide the fit in the IR regime (i.e. the 24 and 250 μm bands), where we set the errors to 0.1 magnitudes. Within such a scheme, we tried different value but noticed that the results did not change significantly and that the only important point was to fix to a sensible lower value the error associated to the key IR bands.

Besides, since the number of bands in the UV/Optical/NIR range are too many with respect to the ones in MIR/FIR/SMM range we decided to discard some of the bands in the UV/Optical/NIR domain in order to balance the fit and again prefer the solution that better reproduced the FIR peak.

The quality of the fits is very good for the majority of the objects in our SPIRE 250 μm "low-z" sample (see Fig.5.3 for some examples).

Once we have the best fit SED for each source we estimate the k-correction and the rest-frame monochromatic luminosity in each band. `Hyperz` does not give these quantities as output of the process so we have written an `IDL` procedure to evaluate them from the best fit SEDs of the sources. In particular we evaluate the IR bolometric luminosity integrating the best fit SED between 8 and 1000 μm, which can then used to evaluate the SFR using the Kennicutt (1998) and the SFR (UV+IR) as from Papovich et al. (2007) and Santini et al. (2009) (see Sec. 5.4 for details).

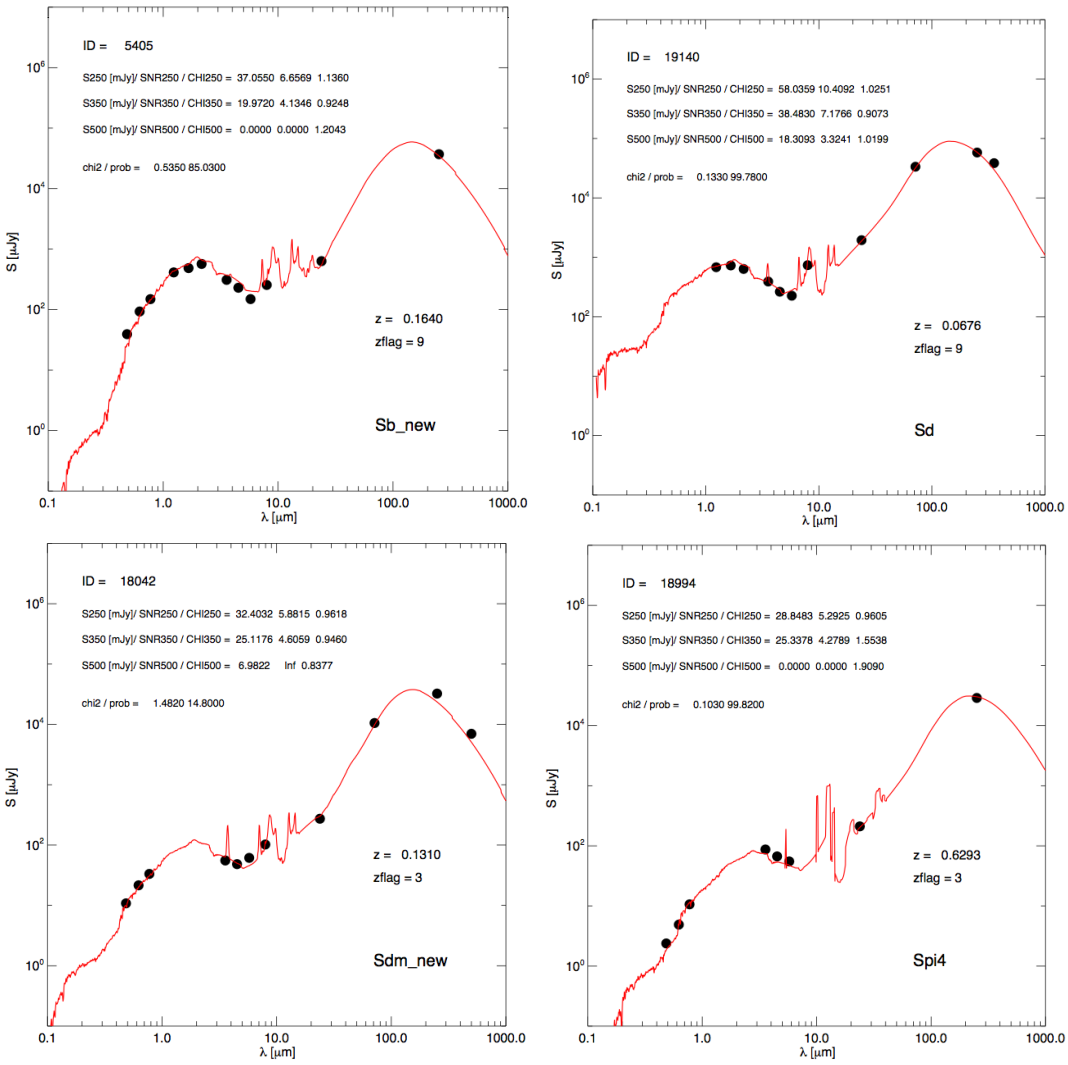

Figure 5.3: Typical `Hyperz` best fit results. The plot shows in red the best fit SED and in black the observed data points. The name of the SED template of best fit is on the right, the subscript *new* stands for the modified templates by Gruppioni et al. (2010). Above the SED template name is the redshift of the source and redshift flag: 3 stands for photometric, 9 stands for spectroscopic. At the top left one finds the ID of the source, the flux, the SNR and the χ^2 of the SPIRE extraction for the 250, 350 and 500 μm respectively, then the χ^2 and the probability of the fit given by `Hyperz`.

5.2.2 `LePhare` Photometric Analysis for Redshift Estimate

`LePhare` is a set of `fortran` programmes written by Arnouts et al. (1999). The main purpose of the code is to compute photometric redshifts using SED fitting applied to the Optical/NIR domain, but has been recently upgraded in order to fit the whole SED of galaxies, from the UV to the MIR/FIR/SMM. This package is composed of three standalone procedures which share the results of each other:

- A preliminary phase to select the SED models, the set of filters and to compute the template magnitudes, using stand-alone programmes. These allow to extract basic informations relative to the filters (λ_{mean} , AB-corrections, attenuation) and template SEDs (k-corrections and colour-colour diagrams as a function of redshift).

- The photometric redshift code based on a simple χ^2 minimization method

- A generator of realistic multi-colour catalogues taking into account observational effects.

This structure allows the code to be faster and more flexible than `Hyperz` so it is preferable when the number of bands and/or sources increases. To find the best fit it uses, like `Hyperz`, a χ^2 minimization, but, differently, it makes use of the *Bayesian* statistic. Additional constraints can be applied to the χ^2 fitting procedure with a simple option in the code, in this way the user can restrict the redshift, extinction ranges, the expected mass and absolute luminosity ranges. A prior on the redshift distribution, following a similar procedure than Benítez (2000), can be applied as e.g. done by Ilbert et al. (2006).

Another advantage of the code is that it allows one to fit separately the UV/Optical/NIR and the MIR/FIR/SMM data points in one go, using different templates and thus, for example, it can estimate at the same time the photometric redshift and stellar mass (fitting carefully the SED, with suitable phenomenological templates or synthetic models, up to the NIR band) and the fit to the FIR peak (fitting carefully the SED, with suitable phenomenological templates, from the MIR to the SMM) without loosing accuracy in the fitting at any wavelength.

`LePhare`, as the other template-based SED fitting code needs as input the photometric catalogue of sources, a list of template SEDs and a list of filters transmissions. Like `Hyperz`, it provides a list of templates SEDs, synthetic models, filters transmissions and extinction laws that the user can change or expand as desired. In particular, this code can be used even only to create simulated multi-colour catalogues simply based on a set of templates and filters. This peculiarity of the code has been exploited to construct well-defined statistical galaxy samples at faint magnitudes and use colour-selection or photometric redshift techniques and thus test the pre-selection of candidates for large spectroscopic surveys probing the high-redshift Universe. In fact, to achieve this goal, deep Optical/Infrared multi-band surveys covering areas of the order of a few degrees are required. Generating mock catalogues means to be able to address some critical issues, relative to the efficiency of a survey characterized by its observational conditions.

The way in which `LePhare` ingests the information contained in the filter transmissions is more sophisticated than `Hyperz`. For example, while `Hyperz` assumes that the transmission of the filters are all stored in units of Energy, `LePhare` gives the possibility to the users to select if the transmission of the filters are express in units of Energy or Photons. The transmissions (T_λ) are dimensionless (in %), however they refer either to a transmission in Energy or Photon which will slightly modify the magnitude estimates. The magnitude is :

$$mag(*) = -2.5log_{10}\frac{\int F_\lambda R_\lambda \mathrm{d}\lambda}{\int F_\lambda(Vega) R_\lambda \mathrm{d}\lambda} \tag{5.2}$$

If the transmission curve (T_λ) corresponds to energy then $R_\lambda = T_\lambda$, If the transmission curve (T_λ) corresponds to number of photons (N_ϕ) then $R_\lambda = \lambda T_\lambda$:

$$N_\phi = \frac{F_\lambda \mathrm{d}\lambda}{h\nu} = \frac{F_\lambda \lambda \mathrm{d}\lambda}{hc} \rightarrow -2.5 log_{10} \frac{\int F_\lambda(*)\lambda T_\lambda \mathrm{d}\lambda}{\int F_\lambda(Vega)\lambda T_\lambda \mathrm{d}\lambda} \rightarrow R_\lambda = \lambda T_\lambda \tag{5.3}$$

When `LePhare` builds the filter library, the filter shape given in input from the user is changed with respect to the original one as follows:

$$R_\lambda = T_\lambda \left(\frac{\lambda}{\langle \lambda \rangle} \right)^{tt} \tag{5.4}$$

where *tt* is the value of the *transmission* parameter (set to 0 for filters whose transmission is in Energy and to 1 for those in number of Photons) and $\langle \lambda \rangle$ is the mean wavelength of the filter.

The modification of filter's shape can be significant for long wavelength filters. Nevertheless it is often not the dominant source of errors with respect to other uncertainties relative to QE-CCD, telescope transmission, atmospheric extinction shape etc. These information are often hidden or totally absent in the documentation of each instruments for this reason even when the transmission parameter is set to the correct units, there is the possibility that the results would be incorrect for some percentage of uncertainties. When we construct our own compilation of filters (see App. B) we pay particular attention to collect all the possible information for all the filters we ingested in the list and to modify the shape of the filters according to the QE of CCDs or the telescope transmission when necessary. These modifications must be applied before to use the filters into a SED fitting procedures since there is no way to tell to the code how to modify the transmissions according to these elements. At any rate we test the problem in the procedure using filters with and without the correction for CCDs QE and did not find sensible differences in the results of the global SED fitting procedure.

Another filter issue that `LePhare` takes into account (which `Hyperz` does not implement) and that the user can set according to his need is the filter *calibration*. When the code has been modified to fit the whole range of wavelengths many tests have been done to be sure that the photometric accuracy of the fit at the longer wavelengths is equal to that in the Optical/NIR domain. In particular at long wavelengths the equivalent fluxes are taken as the monochromatic flux density calculated at the effective wavelength of the filter and for a reference spectrum that would result in the same energy received on the detector

$$F_\nu = \frac{\int F_\nu R_\nu \mathrm{d}\nu}{\int \frac{B_\nu}{B_{\nu_0}} R_\nu \mathrm{d}\nu} \tag{5.5}$$

where B_ν is the reference spectrum and ν_0 the effective frequency of the filter. In the code, the flux estimates are equivalent to consider $\frac{B_\nu}{B_{\nu_0}} = 1$ ($B_\nu = ctt$). Therefore there is a correction factor to account for with respect to the original flux estimated by the code. This correction is:

$$\langle F_\nu \rangle^{COR} = \langle F_\nu \rangle^{\texttt{LePhare}} \times \frac{\int R_\nu \mathrm{d}\nu}{\int \frac{B_\nu}{B_{\nu_0}} R_\nu \mathrm{d}\nu} = \langle F_\lambda \rangle^{\texttt{LePhare}} \times \frac{\int R_\lambda \frac{\mathrm{d}\lambda}{\lambda^2}}{\frac{1}{\lambda_0^2} \int \frac{B\lambda}{B_{\lambda_0}} R_\lambda \mathrm{d}\lambda} \tag{5.6}$$

where λ_0 is the effective wavelength defined as $\lambda_0 = \frac{\int R_\lambda B_\lambda \lambda \mathrm{d}\lambda}{\int R_\lambda B_\lambda \mathrm{d}\lambda}$. At long wavelengths, different conventions have been used for the reference spectrum. `LePhare` provides 6 different options to the users:

- 0 : the default that means $\frac{B_\nu}{B_{\nu_0}} = 1$ or $B_\nu = ctt$;
- 1 : the Spitzer/IRAC, ISO calibration $\nu B_\nu = ctt$;
- 2 : the SMM calibrations $B_\nu = \nu$;
- 3 : B_ν = Black Body at T = 10,000 K;
- 4 : A mixed calibration with ν_0 defined by $B_\nu = ctt$ and the flux estimated as B_ν = Black Body at T = 10,000 K, as adopted for the Spitzer/MIPS calibration;
- 5 : A mixed calibration with ν_0 defined by $B_\nu = ctt$ and the flux estimated as $B_\nu = \nu$, as adopted for the SCUBA calibration.

Like `Hyperz`, `LePhare` also provides many reddening laws and atmospheric extinctions curves; the Cardelli et al. (1989) law is hardcoded in the programmes and is the default law for the galactic extinction. Differently to `Hyperz`, using `LePhare` it is possible to easily add files of the extinction curve by following the format $(\lambda[\mathring{A}], k_\lambda)$. Besides, the user can choose to apply a different extinction law to each template SED used in the fit (while `Hyperz` let to use a unique extinction law to be applied on the whole set of templates).

Finally, `LePhare` gives in output also the integrated IR bolometric luminosity, the SFR estimates (based on the theory of the SFR tracers described in Sec. 5.4) and the k-corrections for each filters used to perform the SED fitting. Since we wrote our own `IDL` code to evaluate all these quantities from the `Hyperz` outputs, we use `LePhare` outputs to check our own estimates).

Here we have described the major properties of `LePhare`, stressing those peculiarities that make a difference with comparison to the simpler `Hyperz`. For more details on the code we suggest to see the online documentation at `http://www.cfht.hawaii.edu/~arnouts/LEPHARE/lephare.html`.

SED Fitting with `LePhare`

Thanks to its flexibility, `LePhare` is more suitable to perform the SED fitting on our "high-z" sample of galaxy. In this case, keeping fixed the redshift of the sources as for the previously described procedure with `Hyperz`, we find better results by fitting separately the UV/Optical/NIR and the MIR/FIR/SMM regimes, using two different sample of template SEDs, one for each wavelengths domain.

To fit the short wavelengths (up to 7 μm rest-frame) we use the 31 SEDs assembled by Ilbert et al. (2009) to compute the COSMOS photometric redshift (see Fig. 5.4). In this paper the goodness of the photometric redshifts, and so, of the SED fitting procedure, has been carefully calibrated using the zCOMOS spectroscopic data by Lilly et al. (2007), so we were confident that the recipe used by Ilbert et al. (2009) was the best solution also for our purposes. This SED sample is based on the Polletta et al. (2007) compilation, but they found that the blue observed colours of the spectroscopic sample were not fully reproduced by only

this sample. For this reason they adds to their list some additional templates generated by using Bruzual & Charlot (2003) (BC03) models with starburst ages ranging from 3 to 0.03 Gyr. In particular they extended the BC03 templates beyond 3 μm rest-frame using the Sdm template of Polletta et al. (2007). The final templates sample is composed by: 7 ellipticals, 12 spirals and 12 Starbursts.

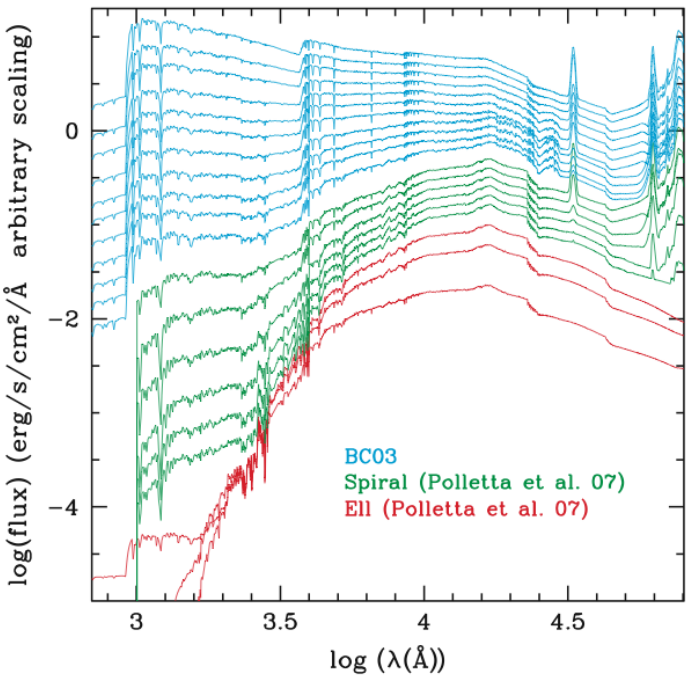

Figure 5.4: SED templates used to perform the SED fitting procedure with `LePhare` in the short wavelengths regime up to 5 μm. This set of templates is taken from Ilbert et al. (2009).

In order to limit the risk of catastrophic failures, they adopted the most suitable extinction curve depending on the SED template. For each galaxy of the zCOSMOS sample, they set the redshift to the spectroscopic redshift value. Then, they determined the best-fit-template and the appropriate colour excess $E(B-V)^{best}$. In this fit, they assumed an extinction curve $k(\lambda) = A(\lambda)/E(B-V)$. Since the extinction curves do not differ strongly at $\lambda > 3000$ Å, they fit the templates using only passbands with $\lambda > 3000(1+z)$ Å. With this procedure, the $E(B-V)$ best value does not depend significantly on the adopted extinction curve. The extinction curves differ strongly at $\lambda < 3000$ Å. Therefore, they used the rest-frame observed SEDs at $\lambda_{rest-frame} < 3000$ Å to discriminate between the different extinction curves. Summarizing their results, they found the best solution by adopting the Prevot et al. (1984) extinction curve for the templates redder than SB3, and Calzetti et al. (2000) for the templates bluer than SB3. They also allowed an additional bump at 2175 Å for the Calzetti extinction law if it produces a smaller χ^2. No reddening is allowed for galaxies redder than Sb. We apply the same approach to our SED fitting procedure.

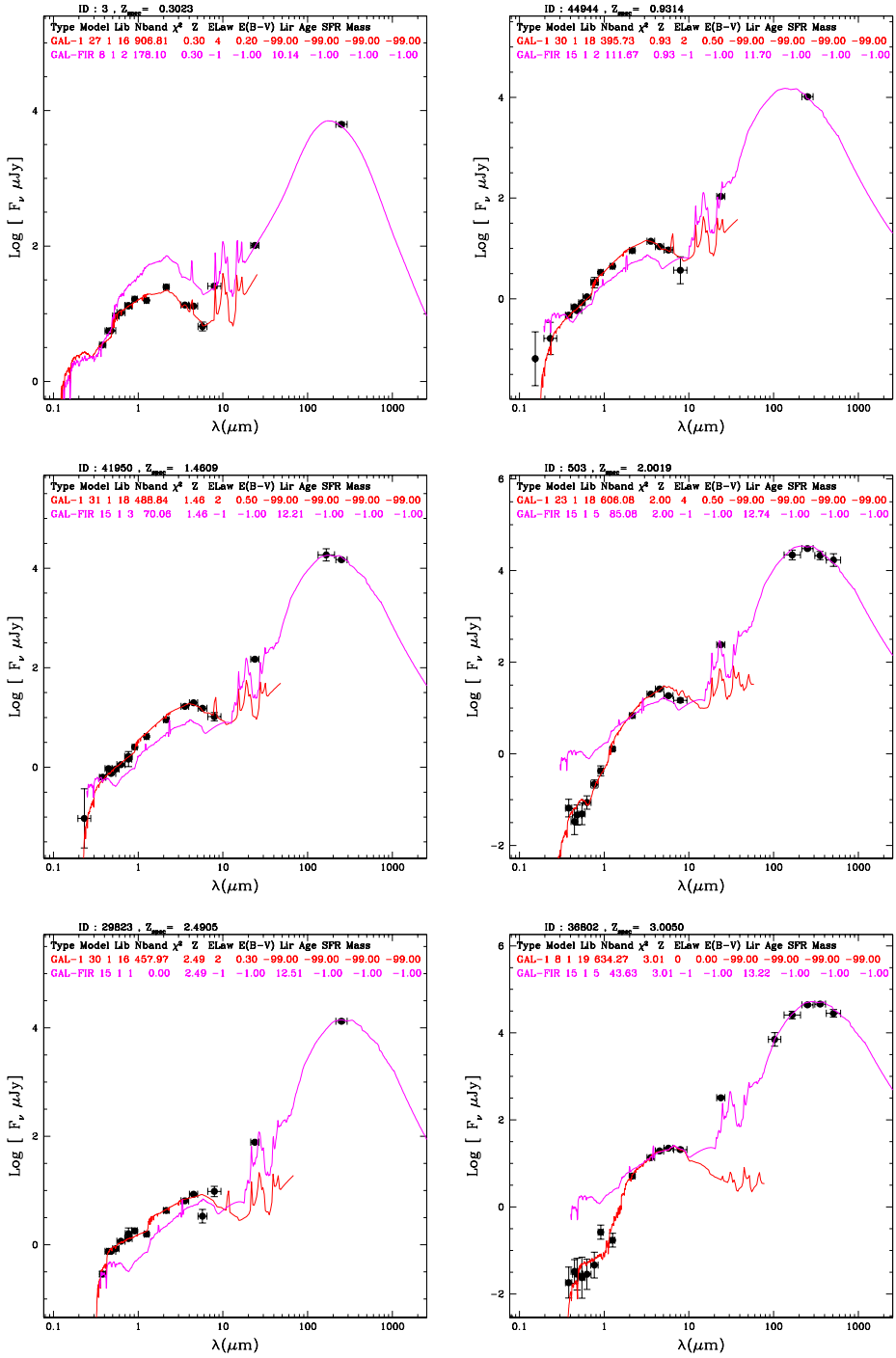

Figure 5.5: Typical `LePhare` best fit results for 6 sources in the COSMOS field. The UV/Optical/NIR best fit SED is reported in red, while the MIR/FIR/SMM best fit SED is reported in magenta.

To fit the longer wavelengths (from 7 μm rest-frame) we use the same templates used to fit the "low-z" sample of galaxies with `Hyperz`, Polletta et al. (2007) + Gruppioni et al. (2010).

Using 7 μm as the dividing line between the two wavelength regimes, the MIPS 24 μm band which is crucial to Herschel source extraction can be used to constrain the FIR peak together with (at least) the SPIRE 250 μm up to $z \sim 2.5$, while at higher redshifts PACS 100 and 160 micron detections (or upper limits) can be used instead.

As an additional benefit, using a double-fit approach we did not have to select *ad hoc* only a few bands to perform the fit in order to balance the number of bands between the short and the long wavelengths regimes; for this reason we perform the fit using all the bands available from the catalogues, which are e.g. 27 in the case of the PEP/HerMES COSMOS field used in Chap. 7 for the study of the evolution of the SMM luminosity function up to $z \sim 5$. Anyway, to ensure the goodness of the FIR/SMM fit where we do not always have data in all the bands, we reduce the error associated to MIPS 24 μm and SPIRE 250 μm fluxes to 0.1 magnitudes.

In Fig. 5.5 we report some examples of the fit we obtain in the COSMOS field using our SPIRE 250 μm selected sample. The SEDs of the sources are well constrained in all our redshift range ($0 < z < 5.0$).

5.3 Comparison Between Hyperz and LePhare SED Fitting Results

As we already mentioned in the text, the two codes, Hyperz and LePhare, which we used to perform the SED fitting analysis on our data, while conceptually similar have somewhat different capabilities and settings that render them differently applicable. In particular we stressed how LePhare is intrinsically more suitable for an "high-z" sample where the complexity of the data requires more sophisticated settings and more extensive testing (i.e. speed, in order to carry out such tests in a timely fashion) to get the best SED fitting results. On the other hand, performing SED fitting analysis of a "low-z" sample is in general easier and one can, at least in principle, use either code. To be on the safe side, in this Section we compare the SED fitting results obtained by using both codes on a subsample of our "low-z" SPIRE 250 μm sample described in Sec. 6.2.1, using only the Lockman Hole, which is a representative field in terms of areas (~ 11 deg^2) and number of sources within the "low-z" fields.

We fit the data using the two codes and the recipes previously described in this Chapter, which can be summarised as follows:

- Hyperz: single SED fitting procedure to fit from the FUV to the FIR regime. The input set of templates SED is the collection composed by Polletta et al. (2007)+Gruppioni et al. (2010); the reddening law is Calzetti et al. (2000). The photometric quantities and relative errors are treated as explained in Sec. 5.2.1.

- LePhare: double SED fitting procedure to fit separately the UV+Optical+NIR and the MIR+FIR+Submm regimes, using two different sample of template SEDs, one for each wavelengths domain, Ilbert et al. (2009) and Polletta et al. (2007)+Gruppioni et al. (2010) respectively. The reddening law and photometric quantities and relative errors are treated as explained in Sec. 5.2.2.

We also consider a further SED fitting analysis by setting up lePhare exactly in the same way we set up Hyperz, i.e. as follows:

- `LePhare`: single SED fitting procedure to fit from the FUV to the FIR regime. The input set of template SEDs is the collection composed by Polletta et al. (2007)+Gruppioni et al. (2010); the reddening law is Calzetti et al. (2000). The photometric quantities and relative errors are treated as explained in Sec. 5.2.1.

This last set-up allows us to actually check the intrinsic differences hardcoded in the two codes without biases coming from different input settings. A comparison of the best fit solution coming from the three different recipes is shown in Fig. 5.6.

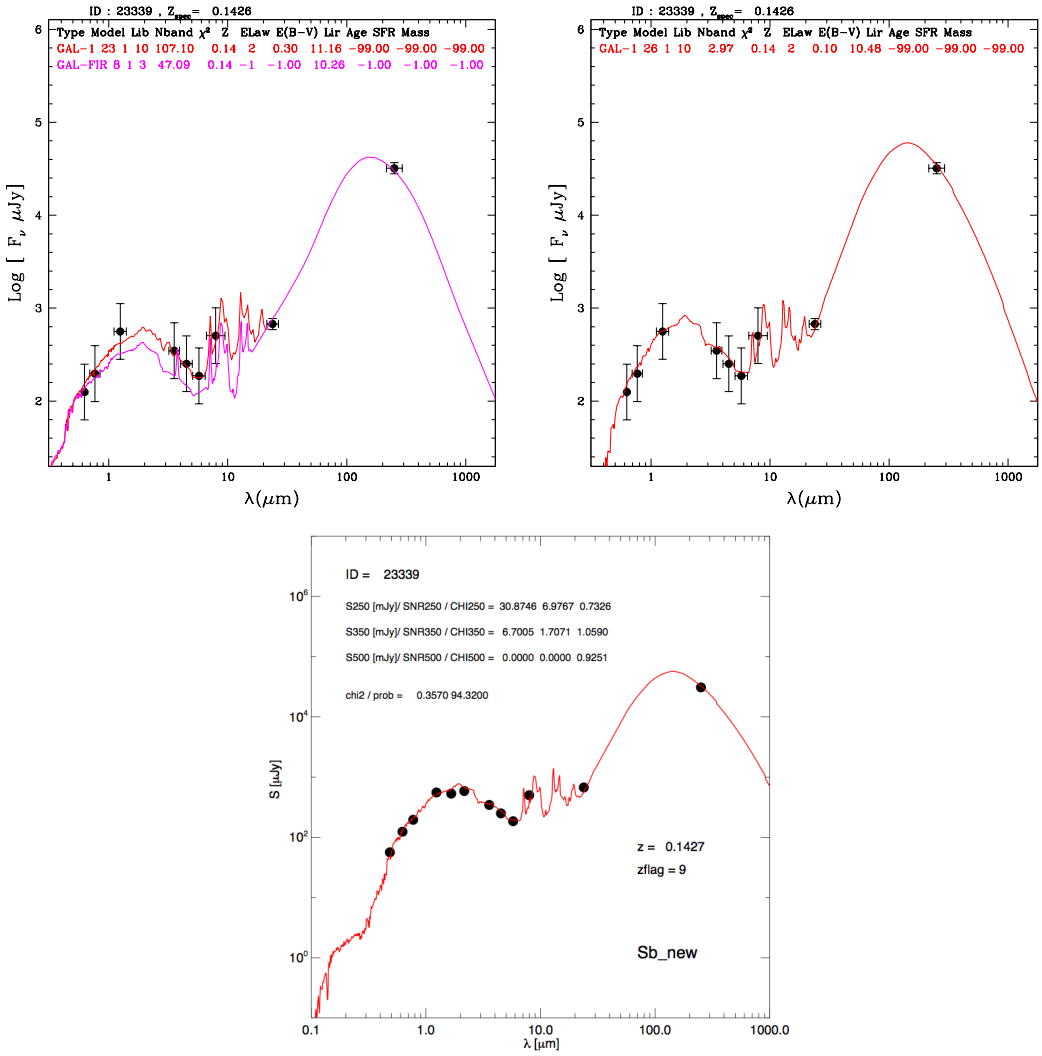

Figure 5.6: Comparison between results of three SED fitting recipes implemented on the same source of the Lockman Hole field. **Top left**: `LePhare` SED fitting procedure with separate fit to UV/Optical/NIR and the MIR/FIR/SMM implemented at the same time (the best fit SED are: one of the Sb templates within the Ilbert et al. (2009) list in the short wavelength regime and the Sd template within the list by Polletta et al. (2007) in the long wavelength regime) ; **Top right**: `LePhare` single SED fitting procedure *a' la* Hyperz; **Bottom**: `Hyperz` single SED fitting procedure from the UV to the SMM regime. The best fit SED for `Hyperz` and `LePhare` run with a single SED is the Sb template within the list by Gruppioni et al. (2010).

The plot shows the three best fit solutions for a single source of the sample fitted using the three SED fitting recipes separately in a successful case. More in general, we find that Hyperz and LePhare set to have the same input settings give us the same best SED solution for the majority of the sources in the "low-z" sample, while LePhare run using separate fit in the two wavelength domains gives more frequently a different best fit solution for the MIR/FIR/SMM, where this recipe can be compared with the other two, and increasingly so at higher redshifts. This can be easily understood since for the vast majority of sources in our sample, and increasingly so at higher redshifts, we only have a few MIR/FIR/SMM data points to be fitted (e.g. the 250 μm, from HerMES and 24 μm, from SWIRE), and in these cases the LePhare procedure of separating the fits to the two wavelength domains with two different sets of template SEDs has got more freedom to correctly reproduce the MIR/FIR/SMM observed SED without being mainly driven by the UV/OpticalNIR observed SED. In such a situation it is easy to understand that this last recipe will more often give us a different best fit solution. Whether this different best-fit solution is actually more correct will depend on a number of factors, which are difficult to quantify given the remaining uncertainties in the Infrared SEDs of high-redshift galaxies. What is clear, however, is that the difficulties in fitting the Optical and Infrared SEDs in a physically self-consistent manner makes the exploration of the two-fit solution with LePhare worthwhile for high-redshift galaxies, since it avoids the risk of weighting the fit heavily toward their optical properties.

Looking at the mismatches between the best-fit SEDs returned by the three different SED fitting recipes above from another point of view, Fig. 5.7 shows a comparison between rest-frame IR bolometric luminosities and rest-frame monochromatic SPIRE 250 μm luminosities for the LePhare two-fit and the LePhare one-fit solution against the Hyperz solution, respectively. From these plots it is clear that: a) at low redshifts the LePhare two-fit solution returns very similar results to the Hyperz solution in terms of best-fit SEDs and thus of luminosity estimates, since both procedures are reliable for the sample under consideration under the given settings; b) at higher redshifts the results of the LePhare two-fit solution diverge more often and more markedly (against Hyperz) than the results of the LePhare one-fit solution, as a combination of greater freedom in the SED fitting on the whole and increasing weight of the optical portion of the SED in the one-fit solution.

On the one hand, the comparison confirms the robustness of our SED fitting analysis of low-redshift sources but also the moderate degree of uncertainty of the SED fitting process when confronted with the sparsely sampled SEDs of higher-redshift sources. On the other hand, however, it confirms our decision to change our SED fitting analysis procedure moving to higher redshifts, where the intrinsic uncertainties of the photometric data and the complexity of the modeling itself require a quicker and more sophisticated SED fitting analysis which Hyperz does not allow.

In summary, the comparison confirms we can rather reliably use our Hyperz and LePhare SED fitting recipes for SPIRE samples at "low-z" and "high-z" respectively, as done in Chap. 6 and Chap. 7 respectively. Additionally, given the good agreement between the two recipes at "low-z", in Chap. 6 we also use MIPS samples to determine the local luminosity function in the three MIPS channels.

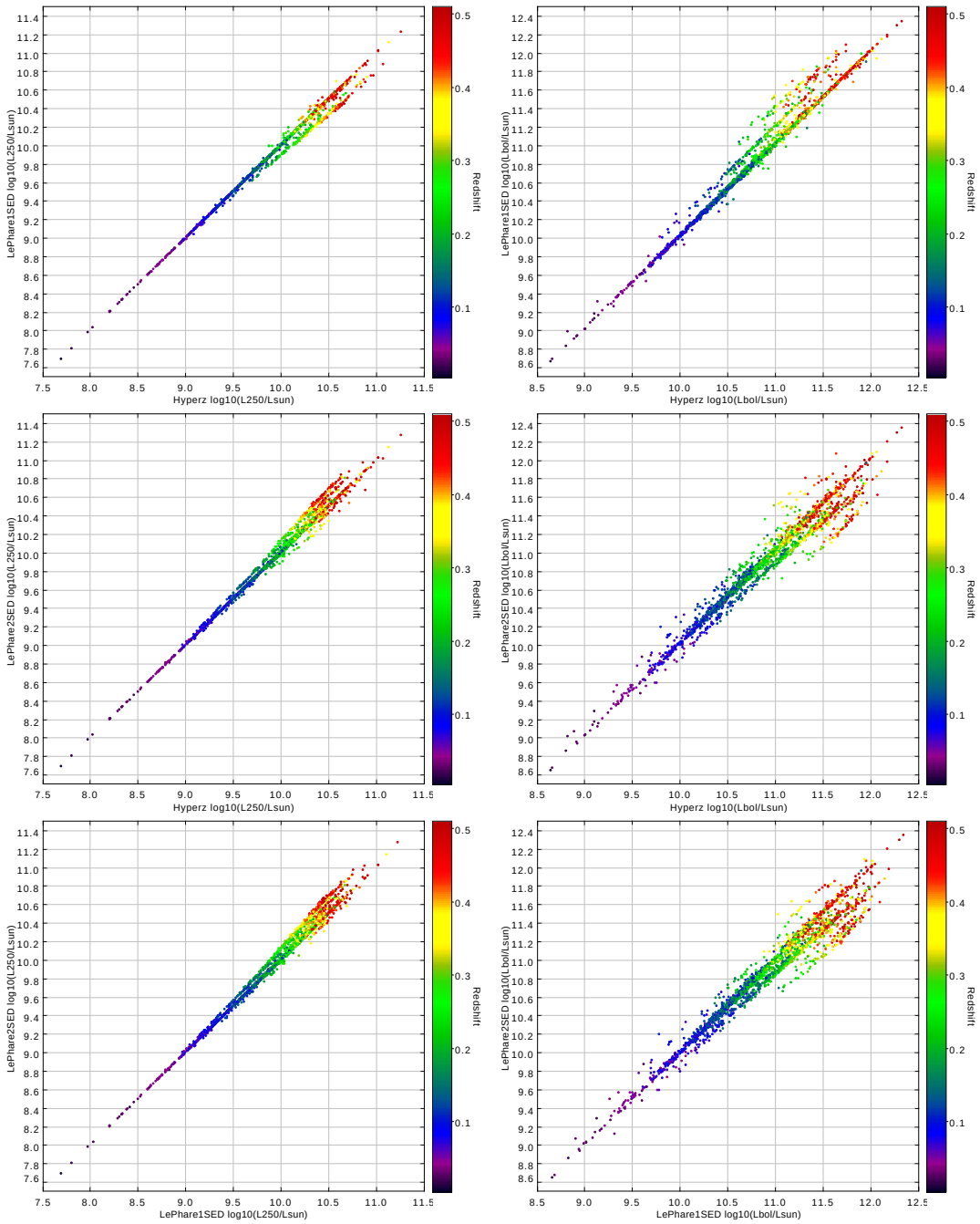

Figure 5.7: Comparison between the monochromatic SPIRE 250 μm luminosity and FIR bolometric luminosity estimates obtained with different SED fitting recipes. `Hyperz` with a single SED, `LePhare` with a single SED and `LePhare` with two SEDs (separately fit to Optical and IR) are considered.

5.4 Estimating Galaxy Star Formation Rates

When studying galaxy formation and evolution one has to bear in mind that these processes can be observationally constrained through the study of their present and past star formation activity and thus the first step is to understand how, where and when the stars come into being in the first place and what their distinctive observational signatures may be at different wavelengths.

From a simplistic point of view, the process of star formation can be described as a sequence of steps: firstly the gas component of the interstellar medium is distributed inhomogeneously within galaxies; then internal or external perturbations such as spiral density waves, gas compression (due to supernova explosions) or interactions with nearby companions can introduce turbulence in the gas, creating high density regions where gravity can break the hydrostatic equilibrium and make the gas collapse; gas collapses into giant molecular clouds and fragments into smaller systems which are at the origin of star clusters and where, in the densest regions, the temperature can increase sufficiently to start the hydrogen burning process and then a star is born.

The mass distribution of the star cluster in which stars are born is an important quantity to understand the whole process and is called Initial Mass Function (IMF). This function is expressed as ϕ(M) and is defined as the number density of stars with masses in the interval (M, M+dM) in a new generation of stars. If dN is the number of stars with masses in the range of mass indicated and formed during the time interval dt then

$$\mathrm{dN} = f(\mathrm{M}, t)\,\mathrm{d}t\,\mathrm{dM} = \psi(t)\,\phi(\mathrm{M})\,\mathrm{d}t\,\mathrm{dM} \tag{5.7}$$

where ϕ(M) is normalised as

$$\int_{\mathrm{M}_{low}}^{\mathrm{M}_{up}} = \mathrm{M}\,\phi(\mathrm{M})\,\mathrm{dM} = 1 \tag{5.8}$$

where M_{low} and M_{up} are the lower and upper mass cut-offs of the initial mass function, generally taken to be 0.1 and 100 $\mathrm{M}_{\odot}$ respectively, and $\psi(t)$ is the SFR or the total mass of stars formed per unit of time.

Generally the IMF is represented by a power-law

$$\phi(\mathrm{M}) \propto \mathrm{M}^{-\alpha} \tag{5.9}$$

Besides, it is common to express the IMF in logarithmic units and thus $\epsilon(M) = \frac{\mathrm{dN}}{\mathrm{dlog(M)}}$ and $\epsilon(\mathrm{M}) \propto \mathrm{M}^{-\alpha+1}$. We report in Tab. 5.2 some typical values for the IMF evaluated in the recent literature (for more details on each IMF estimates we refer the reader to the references).

In addition to the ones reported in Tab. 5.2 we have to mention other famous IMFs that are not described by a simple power law, but by a different more sophisticated functional form. These are:

- Chabrier (2001). This IMF is defined as:

$$\Phi(M) \propto \mathrm{M}^{-2.3}\exp\left[\left(-\frac{\mathrm{M}_C}{\mathrm{M}}\right)^{1/4}\right] \tag{5.10}$$

Reference	**Power law slope α**			
	$0.1 - 0.5\mathrm{M}_\odot$	$0.5 - 1\mathrm{M}_\odot$	$1 - 10\mathrm{M}_\odot$	$10 - 100\mathrm{M}_\odot$
Salpeter (1955)[1]	1.35	1.35	1.35	1.35
Kennicutt (1983)	0.4	0.4	1.5	1.5
Scalo (1986)	0.15	1.1	2.05	1.5
Scalo (1998)	0.2	0.2	1.7	1.3
Kroupa (1998)	0.5	1.2	1.7	1.7
Kroupa (2001)	-0.7	0.3	1.3	1.7
Franceschini et al. (2001) *Top-heavy 1*	1.0	1.0	1.0	1.0
Franceschini et al. (2001) *Top-heavy 2*	0.85	0.85	0.85	0.85

Table 5.2: Table of the most used power-law IMFs. Part of the numbers reported here are taken from the relative papers and some from Tab. 2 in Baldry & Glazebrook (2003)

- Larson (1998). This IMF is defined as:

$$\Phi(M) \propto \mathrm{M}^{-1.35}\exp\left(-\frac{\mathrm{M}_L}{\mathrm{M}}\right) \tag{5.11}$$

  This IMF recovers the Salpeter IMF at high masses, whereas at low masses the exponential cut-off determines the steep downfall after the peak mass of 0.25 $\mathrm{M}_\odot$.

The SFR, measured in $\mathrm{M}_\odot\ \mathrm{yr}^{-1}$, is generally estimated by observing the young stellar population; young and massive stars have relatively short lifetimes on the main sequence, thus their presence indicates recent star formation episodes in galaxies. Given $\mathrm{L}(\lambda)$ as the luminosity of the emission of the youngest stars, the SFR can be determined trough

$$\psi(t) = \mathrm{SFR} = k(\lambda)\mathrm{L}(\lambda) \tag{5.12}$$

where $k(\lambda)$ can be inferred from population synthesis models with several assumptions on the shape of the IMF and on the upper and lower mass cut-offs.

Within studies of the SFR in cosmological surveys, it is common to define the Specific Star Formation Rate (SSFR) as:

$$\mathrm{SSFR}[\mathrm{yr}^{-1}] = \frac{\mathrm{SFR}}{M_*} \tag{5.13}$$

where M_* is the stellar mass of the galaxy, which can be determined by the UV/Optical/NIR SED fitting of the galaxy using stellar population synthesis models, once assumed an IMF.

This is a widely used technique which can be blindly applied only if the IMF is universal, thus independent of morphological type, luminosity, redshift and if the star formation activity of the target galaxies has been constant for a time greater than or equal to the lifetime of the emitting stars. Stationarity in the star formation process is requested since, at a given wavelength, the number of newly formed emitting stars should equal the number of older emitting stars leaving the main sequence, so that the total emission of a galaxy is proportional to the number of emitting stars formed per year. This stationarity regime is reached only if stars formation is constant over a time scale comparable to, or larger than, the lifetime of the emitting stars.

The stationarity condition is quite restrictive since it is well known that the star formation activity of galaxies changes with time because of an inconstant supply of gas or because of possible interactions with the nearby environment. Therefore it is clear that modeling the SFR of a galaxy is a very complex topic and the number of potential star formation tracers is limited by a number of factors.

SED fitting codes that are able to fit the SED and give as an output the SFR for the fitted galaxies (like `LePhare` and `Hyperz`) use stellar population synthesis models to fit the extinction-corrected UV to NIR SED and then find out the time-dependent SFR that best reproduces the observed data. Once this relation is found, the present day SFR can be determined by simply measuring the SFR at $t = t_0$, where t_0 is the age of the observed galaxy. This procedure is subject to the uncertainty introduced by extinction corrections, but has the advantage that it does not need the stationarity conditions of SF and can be used for any kind of objects with a SFR which varies on short time scales. Its major limitation however is that it strongly depends on the assumed star formation history (SFH) that would heavily complicate the SFR estimates requiring a more careful analysis of the SED by a sophisticated *ad hoc* procedure.

Even if the restrictions on its use are numerous, the more common (and faster) approach uses Eq. 5.12 and evaluates $k(\lambda)$ using some recognised SFR tracer linked to readily available observational data. Here we report some of the most common such SFR tracers and, in particular, those we consider in our own work, focusing our attention on the star formation tracers based on the continuum emission (see Boselli (2011) for a more complete review).

The UV stellar continuum of galaxies ($\lambda \leq 2000$ Å), due to the emission of the young stellar populations, is often used to determine the present day star formation activity of late type galaxies. Given the relatively long lifetime of the emitting sources, UV luminosities can be transformed into SFRs, provided that the star formation activity of the observed galaxies is constant on time scales of a few 10^8 yr (the UV data should be corrected for dust extinction following the extinction laws already introduced). This dust-extinction-corrected star formation tracer is well tuned for normal, optically or UV selected galaxies, for low metallicity systems and any dust-poor objects in the nearby and far universe. It is unusable in highly obscured objects such as infrared selected galaxies or strong starbursts. Kennicutt (1998) proposes a calibration constant to transform UV luminosities into SFRs that is valid for: the spectral range between 1500 and 2800 Å, calibrated for a Salpeter IMF in the mass range 0.1 - 100 M$\odot$ and solar metallicity. The relation is

$$\mathrm{k(UV)} = 1.4 \times 10^{-28} \left[\frac{\mathrm{M_\odot yr^{-1}}}{\mathrm{erg} s^{-1} \mathrm{Hz}^{-1}} \right] \tag{5.14}$$

The validity of k(UV) over such a large of spectral range is due to the fact that the UV spectrum of star-forming galaxies is nearly flat if expressed in $S(\nu)$. Changing the IMF calibration law means changing the k(UV) value and alternative solutions following different IMF calibrations can be found in the literature, e.g. in Bicker & Fritze-v. Alvensleben (2005).

In dust-rich star-forming galaxies, the UV radiation emitted by the young stellar population is absorbed and re-emitted in the far-infrared domain. The total infrared luminosity and the monochromatic flux density are thus tightly related with the activity of star formation and can be transformed, under some conditions, into SFRs. In these objects, extinction-corrected UV or hydrogen recombination line fluxes are highly uncertain, given the large attenuation that can easily reach several magnitudes. Thus the FIR-luminosity has to be preferred for

quantifying the SFR for these obscured objects. Again, Kennicutt (1998) gives a calibration constant k(IR) for transforming the integrated rest-frame 8-1000 μm infrared luminosity (LIR) into SFR, again for a Salpeter IMF in the mass range 0.1-100 $M_{\odot}$ and solar metallicity.

$$k(IR) = 4.5 \times 10^{-44} \left[\frac{M_{\odot} yr^{-1}}{ergs^{-1} Hz^{-1}} \right] \tag{5.15}$$

Changing IMF would again change the value of k(IR) like in Bell (2003). The Kennicutt (1998) relation is valid only for highly obscured galaxies where most of the infrared emission is in the 10-120 μm spectral range, indicating a strong continuous star formation activity in the last 10^8 yr (like for the other indirect SF tracers the stationarity condition on the last $\sim 10^8$yr is required).

It should be stressed that this calibration cannot be applied to normal quiescent late-type galaxies in the nearby universe. Their IR emission not only comes from highly obscured star-forming regions, but from diffuse cirrus regions that are heated by the general interstellar radiation field.

Some authors have also proposed a calibration based on the combination of the UV and IR tracers. Based on the assumption that most of the photons originating in newly formed stars are absorbed and re-emitted by dust, the mid-IR emission is in principle the most sensitive tracer of the SFR. In addition, a small fraction of unabsorbed photons will be detected at UV wavelengths. Thus the combination of UV and IR luminosities should give the possibility to take into account the contribution to the SFR of those unabsorbed photons. An example of this estimator can be found in Papovich et al. (2007) and Bell et al. (2005) where the SFR is estimated as:

$$SFR[M_{\odot} yr^{-1}] = 1.8 \times 10^{-10} \times L_{UV+IR}/L_{\odot} \tag{5.16}$$

where

$$L_{UV+IR} = (2.2 \times L_{UV} + L_{IR}) \tag{5.17}$$

Fig. 5.8 compares the UV+IR and IR SFR estimates computed for our local SPIRE 250 μm sample in the Lockman Hole, showing how the UV contribution is mostly negligible for bright IR-selected sources.

A limit in using these relations is the need to determine the total infrared luminosity LIR. Thanks to Herschel we are now able to sample the IR peak of the SED in a more reliable way, and we are thus able to beat the large uncertainties in the k-correction affecting previous work. Before Herschel the need to overcome this issue brought several authors to use monochromatic infrared luminosities at relatively short wavelengths (from $\sim 15 - 25\mu$m). At these rest-frame wavelengths, dust is principally heated by the UV radiation emitted by the youngest stellar population, with a negligible contribution form the evolved populations, making the monochromatic infrared luminosity a direct tracer of the star formation activity. At 24 μm, the most widely used relations are proposed by Alonso-Herrero et al. (2006) calibrated on nearby luminous infrared galaxies and the one determined using the SINGS nearby galaxy sample by Calzetti et al. (2007), respectively:

$$SFR[M_{\odot} yr^{-1}] = 8.5 \times 10^{-38} L(24\mu m[ergs^{-1}])^{0.871} \tag{5.18}$$

and

$$\mathrm{SFR}[\mathrm{M}_{\odot}\mathrm{yr}^{-1}] = 8.5 \times 10^{-38}\mathrm{L}(24\mu\mathrm{m}[\mathrm{erg}s^{-1}])^{0.885} \tag{5.19}$$

Although proposed as alternative tracers, monochromatic infrared luminosities at wavelengths shorter than $\sim 15\mu$m should not be used to quantify the SFR in late-type galaxies since they are dominated by the emission of the PAHs., for the same reason the 24 μm calibrations should not be used for galaxies at $z \geq 0.9$ (where the rest-frame 24 μm start to be affected by the PAHs).

At higher wavelengths, the FIR-Radio correlation, which relates the total IR luminosity to the non-thermal 20 cm Radio continuum luminosity, is one of the tightest correlations observed in galaxies. This correlation is due to the fact that the relativistic electrons spinning weak magnetic fields is responsible for the synchrotron emission, are produced and accelerated in supernova remnants, which are related to the stellar population dominating the dust heating. Thus the radio continuum emission at 1400 MHz (20 cm) has often been proposed as an alternative SFR tracer for star-forming galaxies. This tracer has the great advantage to not be sensitive to the presence of dust and so can be used in highly obscured objects. Condon (1992) proposed the first attempt to calibrate the 20 cm radio luminosity of galaxies in SFR, and more recently Bell (2003) proposed a new calibration determined using the FIR-Radio correlation assuming a Salpeter IMF in the mass range 0.1-100 $\mathrm{M}_{\odot}$

$$\mathrm{SFR}[\mathrm{M}_{\odot}\mathrm{yr}^{-1}] = 5.52 \times 10^{-22}\mathrm{L}(1.4\mathrm{GHz})[\mathrm{WHz}^{-1}] \tag{5.20}$$

for $\mathrm{L}(1.4\mathrm{GHz}) > 6.4 \times 10^{21}[\mathrm{WHz}^{-1}]$, and

$$\mathrm{SFR}[\mathrm{M}_{\odot}\mathrm{yr}^{-1}] = \frac{5.52 \times 10^{-22}\mathrm{L}(1.4\mathrm{GHz})}{0.1 + 0.9(\mathrm{L}(1.4\mathrm{GHz})/6.4 \times 10^{21})^{0.3}} \tag{5.21}$$

for $3 \times 10^{19} \leq \mathrm{L}(1.4\mathrm{GHz}) \leq 6.4 \times 10^{21}[\mathrm{WHz}^{-1}]$.

The use of this calibration in the Radio should be considered with caution because of the possibly strong contamination of AGNs: although these objects follow the FIR-Radio correlation, the calibration of their IR luminosity in terms of SFR is highly uncertain because of their peculiar nature. Beside, the non-thermal Radio emission not only depends on the density of the relativistic electrons, n_e, but also on the intensity of the magnetic filed B, and thus uncertainties derive also from the shape of the magnetic field. Galaxy interactions might compress the magnetic field and thus increase the Radio continuum emission without affecting the electron density and so the SFR.

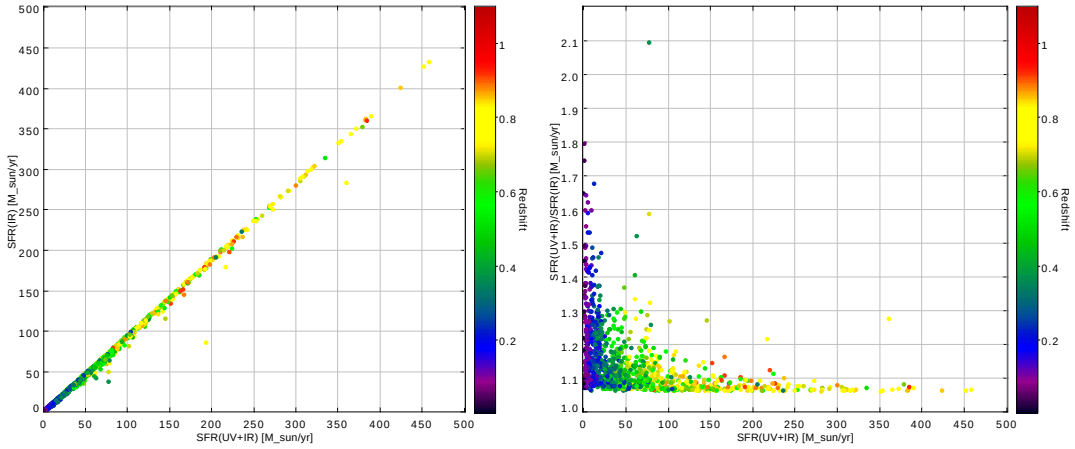

Figure 5.8: Comparison of the SFR estimates using UV+IR and IR luminosities in the COSMOS field according to Eq. 5.16 and Eq. 5.15 in the text. The SED fitting has been carried out on the sample used in Chap. 7 to study the evolution of the SPIRE Luminosity Function and using `LePhare` as described in Sec. 5.2.2. Since all the samples used in the following are IR-selected, it can be concluded that for our purposes the contribution to the SFR by the UV emission is very low with comparison to the one arising from IR emission.

Chapter 6

Spitzer & Herschel Local Luminosity Functions

The determination of the Galaxy Local Luminosity Function is often hampered by the difficulties of covering a wide area down to faint fluxes on one hand and determining counterparts and their redshifts for detected sources in a complete and reliable manner on the other.

In this Chapter, armed with the Spitzer Data Fusion, we describe the derivation of the Local Luminosity Function of sources selected in wide fields by the Spitzer MIPS and Herschel SPIRE imagers.

6.1 MIPS Local Luminosity Functions

As a first exploitation of the Spitzer Data Fusion we report here preliminary estimates of the MIPS local Local Luminosity Function (LLF) based on data obtained in the Lockman Hole (LH) field. Such a study has been useful to check the reliability of our Spitzer Data Fusion and of our SED fitting recipes in preparation for the estimate of the SPIRE LLF (see Sec. 6.2) and its evolution (see Chap. 7). This preliminary analysis will later be extended to all Spitzer-SDSS fields, as done for SPIRE.

The MIPS data as well all the ancillary catalogues we use to perform the SED fitting analysis have been already introduced in Chap. 4 along with the Spitzer Data Fusion construction process. In Sec. 6.1.1 we give some more details about the sample we used to estimate the MIPS Local Luminosity Functions.

6.1.1 MIPS 24/70/160 μm Sample in the Lockman Hole Field

The choice of the sample on which to base our Local Luminosity Function calculation was partly guided by the completeness of SWIRE single-band catalogues and partly by the completeness of the associated multi-wavelength (and redshift) information. The opportunity to increase the sample size where sample statistics was poorer, and in particular at 70 and 160 μm, also prompted us to dig deeper into these single-band catalogues. Adopted flux limits for the computation of the LLF are 0.25, 10 and 50 mJy at 24, 70 and 160 μm. At these flux levels, SWIRE IRAC imaging is expected to detect all MIPS sources at both 3.6 and 4.5 μm. We thus restrict our study of the IR local luminosity functions to sky areas jointly covered by

SWIRE at both 3.6 and 4.5 μm and require a joint IRAC detection at these two wavelengths. 70 and 160 μm sources must then have a reliable 24 μm association. Finally, a S/N cut is applied at $S/N > 5$ at 24 μm and at $S/N > 3$ at 70 and 160 μm and we make no use of SDSS unreliable photometric redshifts (associated to SDSS detected sources with $r_{AB} > 22.0$ as shown in Fig. 4.6).

The main source of sample incompleteness is likely to arise from the MIPS samples themselves. As discussed in Sec. 4.8, while an accurate estimate of catalogue completeness requires extensive simulations, a rough completeness estimate for the 24 μm component of the Spitzer Data Fusion can be obtained by comparing the raw source counts with published counts. A completeness correction was thus computed for the Lockman Hole field based on Fig. 4.5. Conversely, completeness of 70 and 160 μm catalogues was assessed by Vaccari et al. (in prep) for $S/N > 5$ sources using simulations. In order to extend our work to use $S/N > 3$ sources, we developed a Gaussian noise model adopting the source rms flux error obtained by Vaccari et al. (in prep) for each field and wavelength. This noise model is found to reproduce very well the completeness estimates obtained through simulations for $S/N > 5$ sources by Vaccari et al. (prep). While extending our sample to include $S/N > 3$ sources introduces additional uncertainty in terms of completeness correction and photometric accuracy, it is believed that the benefit of increasing the sample size (while maintaining a very high reliability thanks to the required detection at IRAC and MIPS 24 μm wavelengths) far outweighs the drawback caused by this additional uncertainty. The inverse of the completeness correction as a function of flux thus estimated was applied in the computation of the MIPS Local Luminosity Functions as a weighting factor for each source in our sample.

The requirement of a joint detection by IRAC in its two very sensitive channels at 3.6 and 4.5 μm, along with the moderate PSF FWHM of the MIPS 24 μm images, ensure the uttermost reliability of the 24 μm sample. On the other hand, while the reliability of 70 and 160 μm sources is ensured by a 24 μm detection, the larger beam size at these wavelengths and the moderate areal density of 24 μm sources increase the likelihood of chance associations. In Sec. 4.3.1 we accordingly defined a reliability for every 70 an 160 μm source on the basis of the distance of the 24 μm nearest neighbor and of the 24 μm source density. In the study of the MIPS Local Luminosity Functions, we restrict ourselves to sources whose reliability is higher than 0.95 and 0.75 at 70 and 160 μm respectively.

On the basis of these considerations we select three sample of sources at 24/70/160 μm respectively and we estimate the Local Luminosity Function for each band in different redshift bins. In Tab. 6.1 we report the number of sources in each band for the complete sample in the redshift range $0 < z \lesssim 0.5$, while in Fig. 6.1 we report its magnitude and redshift distribution.

	24 μm ($0 < z < 0.6$)	70 μm ($0 < z < 0.6$)	160 μm ($0 < z < 0.4$)	Area [deg^2]
LH	15696 (2521/13175)	3082 (1187/1895)	1304 (621/683)	$\sim$11.0

Table 6.1: Number of $0 < z \lesssim 0.5$ sources used for the determination of the MIPS Local Luminosity Function in the three MIPS channels at 24/70/160 μm. The number of sources with (spectroscopic/photometric) redshifts is indicated after the total number of sources.

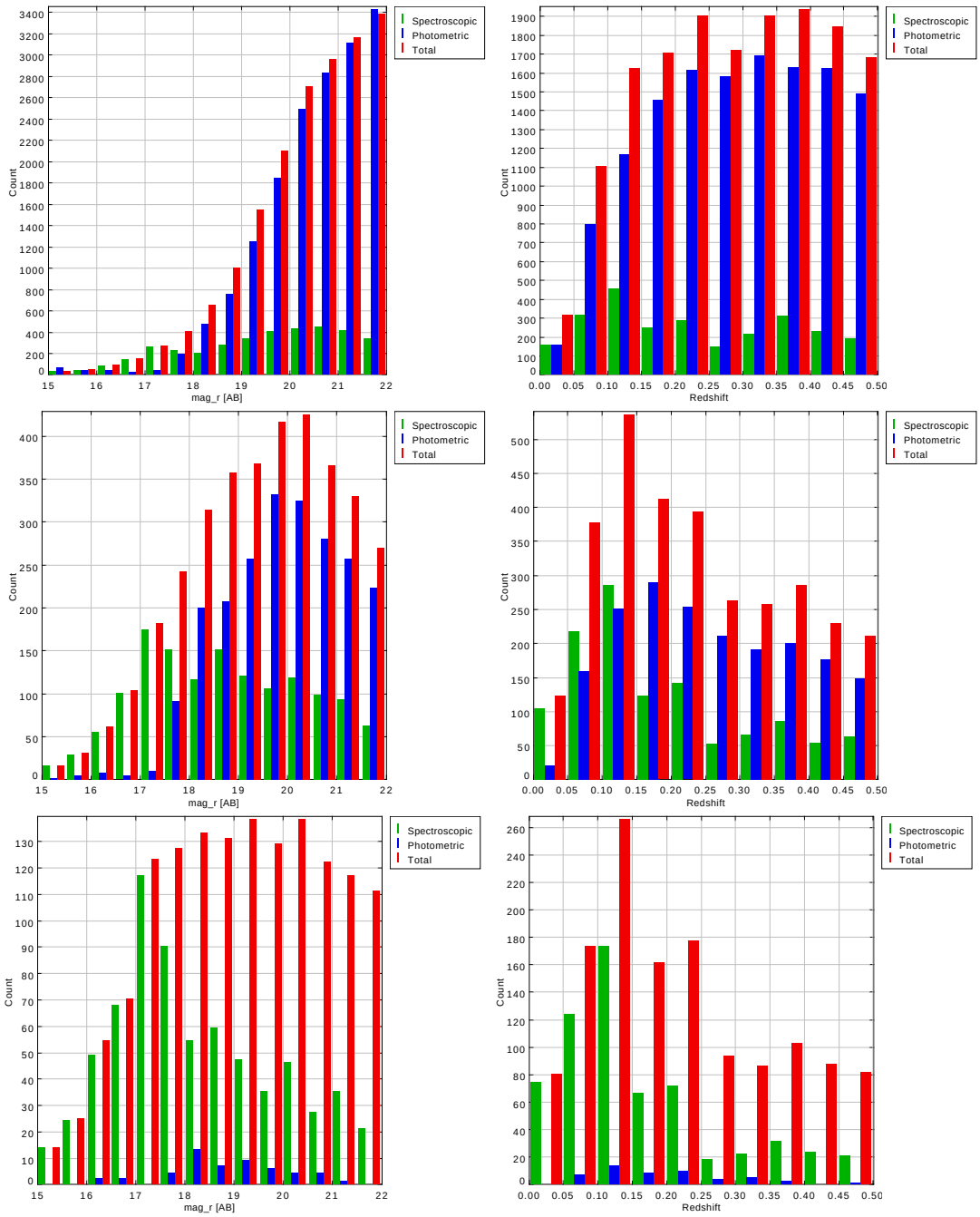

Figure 6.1: Optical Magnitude and Redshift Distribution of the MIPS 24/70/160 μm Bright Local Samples used in the computation of the MIPS Local Luminosity Functions. **Left** : Magnitude Distribution of MIPS Bright Local Samples. **Right** : Redshift Distribution of MIPS Bright Local Samples. **Top** : $S_{24} > 250$ μJy. **Middle** : $S_{70} > 10$ mJy. **Bottom** : $S_{160} > 50$ mJy.

6.1.2 MIPS Local Luminosity Function Estimates

We use the samples described in Sec. 6.1.1 along with data contained in the Spitzer Data Fusion to perform the multi-wavelength SED fitting analysis and thus determine the

observed rest-frame luminosities in each MIPS band. We set up the `LePhare` code a la' `Hyperz` performing a single fit from the UV to the IR, as described in Sec. 5.3. Following the discussion reported in the same Section, this procedure ensure we have a faster SED fitting recipe than using `Hyperz`, and enough constraints to ensure a reliable fit to the FIR/SMM peak where the lack of photometric data points can affect the result of the fit while using the `LePhare` procedure with two SEDs. To perform the fit we part of the photometric data available from the UV to MIPS wavelengths and set the photometric errors as described in Sec. 5.2.1. As template SEDs we use the SWIRE templates by Polletta et al. (2007) and the slightly modified versions by Gruppioni et al. (2010) described in Sec. 5.2.1. In Fig. 6.2 we report the resulting rest-frame luminosity distribution at each MIPS band as a function of redshift.

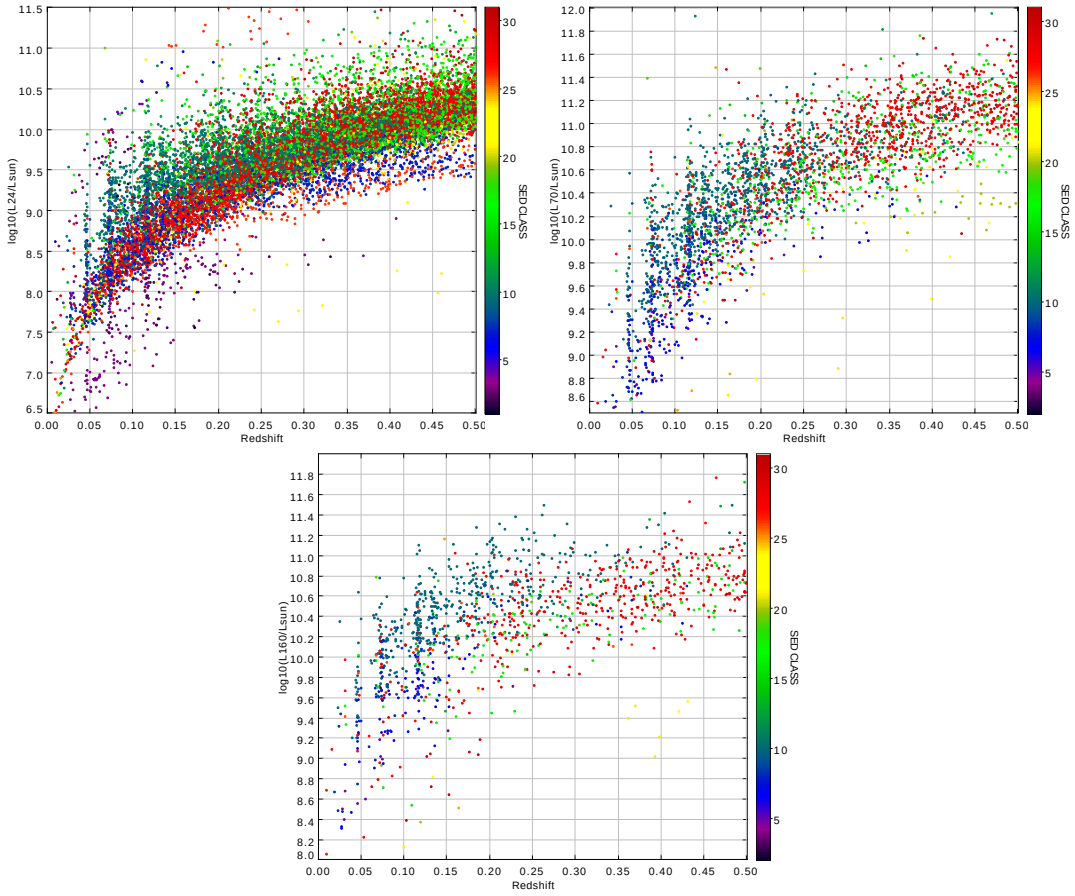

Figure 6.2: MIPS rest-frame luminosity distribution as a function of redshift in the Lockman Hole sample used for MIPS LLFs estimates. The plot is colour-coded by the best fit SED indices obtained as solution of the SED fitting procedure. The SED indices are listed in Tab. 5.1.

To evaluate the MIPS LLFs we use the $1/V_{max}$ technique (Schmidt 1968). The z_{max} and thus V_{max} estimates have been calculated combining the limits set by the two selection criteria introduced in Sec. 6.1.1: SDSS $r_{AB} > 22$ and $S_{24/70/160} > 250$ μJy / 10 mJy / 50 mJy respectively. Besides, since we estimate the $1/V_{max}$ in a number of redshift bins the V_{max} is

actually also limited by the z_{min} and z_{max} of each z-bin. Taking all of this into account the more general Eq. 2.42 for V_{max} becomes:

$$V_{max} = \frac{\Omega}{4\pi} \int_{z_{min}}^{z_{max}} \mathrm{d}z \frac{\mathrm{d}V}{\mathrm{d}z} \tag{6.1}$$

where z_{min} and z_{max} are the redshift boundaries resulting by taking into account both the redshift bin range and the selection factors:

$$z_{k,min} = z_{bin_k,min} \tag{6.2}$$
$$z_{k,max} = \min[z_{0,max}, .., z_{n,max}, z_{bin_k,max}] \tag{6.3}$$
$$\tag{6.4}$$

for all $0, ..., n$ selection factors and for each k redshift bin. For instance, in the case of the MIPS 24 μm luminosity function estimate in the z-bin $0.01 < z < 0.2$ these last equations become:

$$z_{0.01<z<0.2,min} = 0.01 \tag{6.5}$$
$$z_{0.01<z<0.2,max} = \min[z_{r_{AB},max}, z_{f24,max}, 0.2] \tag{6.6}$$
$$\tag{6.7}$$

where $z_{r_{AB},max}$ and $z_{f24,max}$ are the redshift at which a source in the sample reach the SDSS r_{AB} magnitude limit = 22 and the 24 μm flux limit = 250 μJy respectively; 0.01 and 0.2 are the minimum and the maximum of the redshfit bin.

In Tab. 6.2, 6.3 and 6.4 we report the values for the resulting LFs estimates in each band and redshift bins illustrated in Fig. 6.3, 6.4, 6.5. The error associate to each value of ϕ is estimate following the Poissonian statistics as shown in Eq. 2.43. We exclude from the caIculation the sources with $z < 0.01$ to avoid the range of redshift where even small redshift uncertainties may cause very large errors in luminosity function estimates. In these plots we compare our estimates to others reported in the literature. We report MIPS 160 μm LFs estimates for only two local bins with respect to others MIPS bands due to the very small number of 160 μm sources at higher redshifts.

MIPS 24 μm Local Luminosity Functions

$\log L$	$\log(\phi,\sigma)_{0.01<z<0.1}$	$\log(\phi,\sigma)_{0.1<z<0.2}$	$\log(\phi,\sigma)_{0.2<z<0.4}$	$\log(\phi,\sigma)_{0.4<z<0.6}$
8.00	-1.54 , -2.67	-2.25 , -2.59	-	-
8.25	-1.67 , -2.89	-1.60 , -2.35	-	-
8.50	-1.89 , -3.06	-1.68 , -2.81	-	-
8.75	-2.08 , -3.16	-1.88 , -3.24	-	-
9.00	-2.26 , -3.25	-2.13 , -3.55	-2.04 , -2.89	-
9.25	-2.36 , -3.30	-2.21 , -3.63	-2.17 , -3.55	-
9.50	-2.62 , -3.43	-2.37 , -3.71	-2.29 , -3.86	-
9.75	-2.85 , -3.54	-2.67 , -3.86	-2.43 , -4.08	-
10.00	-3.23 , -3.73	-3.06 , -4.05	-2.66 , -4.27	-2.53 , -3.94
10.25	-3.54 , -3.88	-3.53 , -4.29	-3.05 , -4.46	-2.74 , -4.41
10.50	-	-3.90 , -4.47	-3.54 , -4.71	-3.06 , -4.64
10.75	-	-4.05 , -4.55	-3.97 , -4.93	-3.57 , -4.90
11.00	-	-4.27 , -4.66	-4.43 , -5.16	-3.90 , -5.06
11.25	-	-4.74 , -4.89	-4.88 , -5.38	-4.45 , -5.34
11.50	-	-	-	-5.32 , -5.77
11.75	-	-	-	-5.62 , -5.92
12.00	-	-	-	-5.75 , -5.99

Table 6.2: MIPS 24 Local Luminosity Function estimates in four redshift bins within $0.01 < z < 0.6$ in the Lockman Hole field. L is expressed in unit of $L_\odot$ and LLF estimates and their errors are in [$\mathrm{Mpc}^{-3}\,\mathrm{dex}^{-1}$].

MIPS 70 μm Local Luminosity Functions

$\log L$	$\log(\phi,\sigma)_{0.01<z<0.1}$	$\log(\phi,\sigma)_{0.1<z<0.2}$	$\log(\phi,\sigma)_{0.2<z<0.4}$	$\log(\phi,\sigma)_{0.4<z<0.6}$
8.00	-1.99 , -2.02	-	-	-
8.25	-1.85 , -2.27	-	-	-
8.50	-1.76 , -2.34	-	-	-
8.75	-1.87 , -2.63	-	-	-
9.00	-1.94 , -2.80	-	-	-
9.25	-2.11 , -3.03	-	-	-
9.50	-2.26 , -3.23	-	-	-
9.75	-2.37 , -3.30	-2.20 , -3.24	-	-
10.00	-2.43 , -3.33	-2.39 , -3.56	-	-
10.25	-2.64 , -3.43	-2.59 , -3.80	-2.44 , -3.19	-
10.50	-3.28 , -3.75	-2.78 , -3.91	-2.74 , -3.94	-
10.75	-3.63 , -3.93	-3.17 , -4.11	-2.99 , -4.27	-
11.00	-	-3.89 , -4.47	-3.38 , -4.59	-2.97 , -3.67
11.25	-	-4.44 , -4.74	-3.75 , -4.82	-3.42 , -4.59
11.50	-	-	-4.60 , -5.24	-4.09 , -5.13
11.75	-	-	-5.35 , -5.62	-4.74 , -5.48
12.00	-	-	-	-5.62 , -5.92

Table 6.3: MIPS 70 Local Luminosity Function estimates in four redshift bins within $0.01 < z < 0.6$ in the Lockman Hole field. L is expressed in unit of $L_\odot$ and LLF estimates and their errors are in [$\mathrm{Mpc}^{-3}\,\mathrm{dex}^{-1}$].

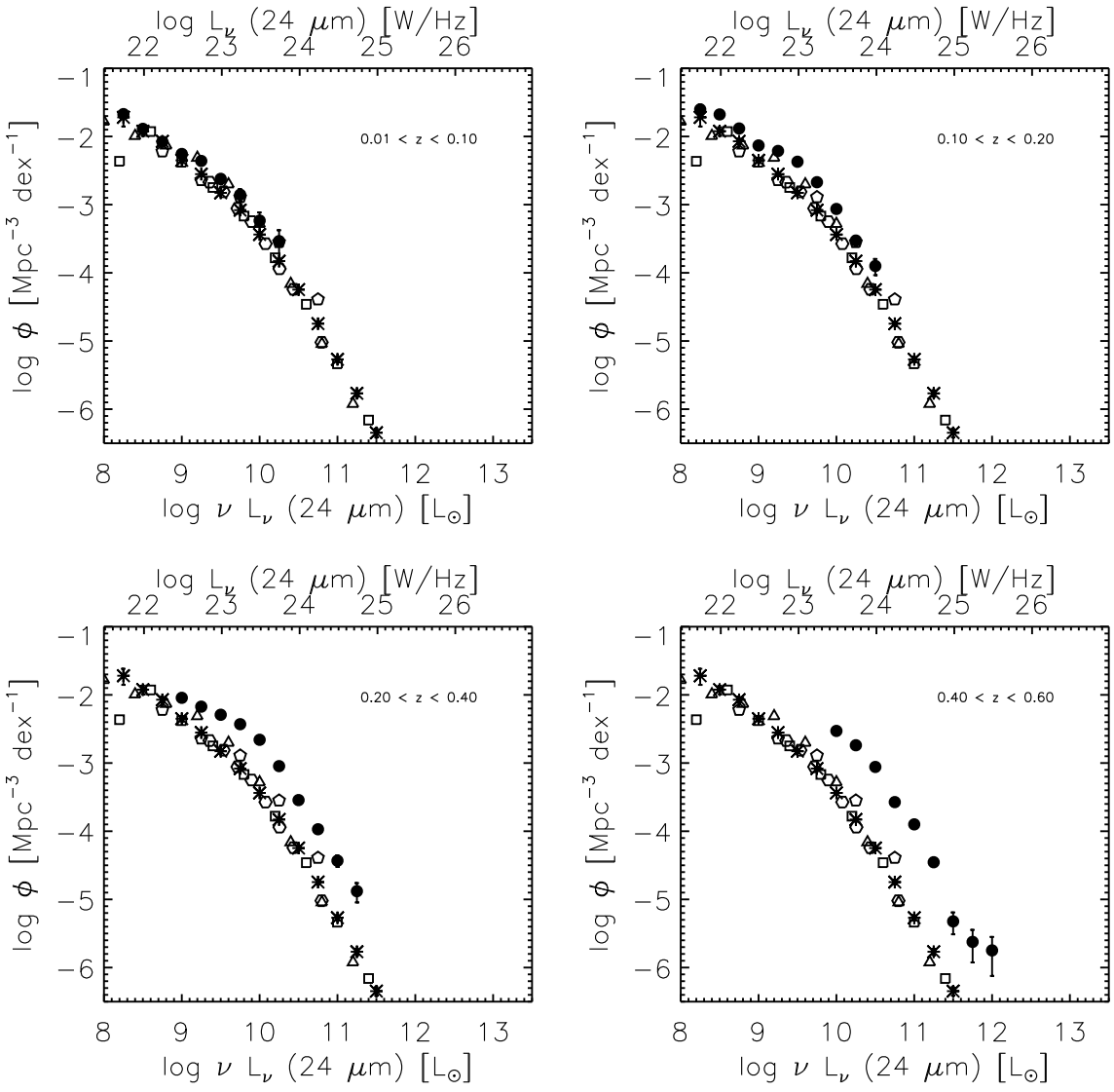

Figure 6.3: MIPS 24 μm Local Luminosity Function in four redshift bins within $0.01 < z < 0.6$ in the Lockman Hole field. Full black circle: our $1/V_{max}$ estimate; black empty squares:Marleau et al. (2007) MIPS 24 μm $1/V_{max}$ luminosity function estimate within $0 < z < 0.25$; black empty pentagons: Rodighiero et al. (2010) MIPS 24 μm $1/V_{max}$ luminosity function estimate within $0 < z < 0.25$; black asterisks: Babbedge et al. (2006) MIPS 24 μm $1/V_{max}$ luminosity function estimate within $0 < z < 0.25$.

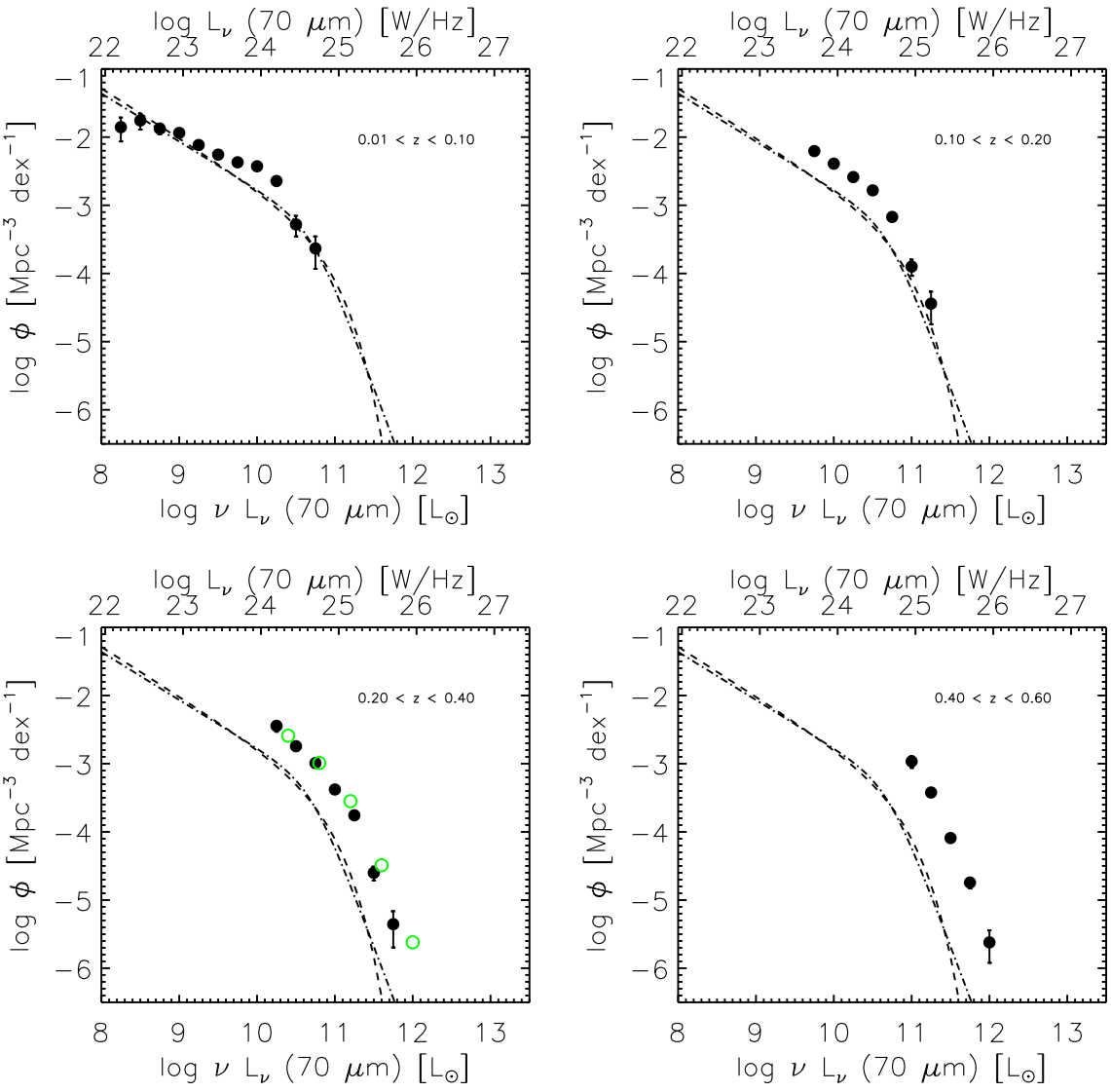

Figure 6.4: MIPS 70 μm Local Luminosiy Function in four redshift bins within $0.01 < z < 0.6$ in the Lockman Hole field. Black full circles: our $1/V_{max}$ estimate; green empty circles: Patel et al. (2012) $1/V_{max}$ MIPS 70 μm luminosity function estimate at the same reported redshift bin; black dot-dashed and dashed lines are local luminosity function prediction at 70 μm by Serjeant & Harrison (2005) estimated following Eq. 2.64 and Eq. 2.67 respectively.

MIPS 160 μm Local Luminosity Functions

$\log L$	$\log(\phi,\sigma)_{0.01<z<0.2}$	$\log(\phi,\sigma)_{0.2<z<0.4}$
8.25	-2.41 , -2.22	-
8.50	-2.19 , -1.79	-
8.75	-2.56 , -2.29	-
9.00	-2.84 , -2.30	-
9.25	-2.94 , -2.25	-
9.50	-3.20 , -2.38	-
9.75	-3.36 , -2.40	-
10.00	-3.62 , -2.59	-
10.25	-3.86 , -2.76	-2.65 , -3.38
10.50	-4.09 , -3.08	-3.06 , -4.07
10.75	-4.23 , -3.35	-3.48 , -4.51
11.00	-4.52 , -3.93	-3.99 , -4.91
11.25	-	-4.63 , -5.25
11.50	-	-5.58 , -5.73

Table 6.4: MIPS 160 Local Luminosity Function estimates in two redshift bins within $0.01 < z < 0.4$ in the Lockman Hole field. L is expressed in unit of $L_\odot$ and LLF estimates and their errors are in [$\mathrm{Mpc}^{-3}\,\mathrm{dex}^{-1}$].

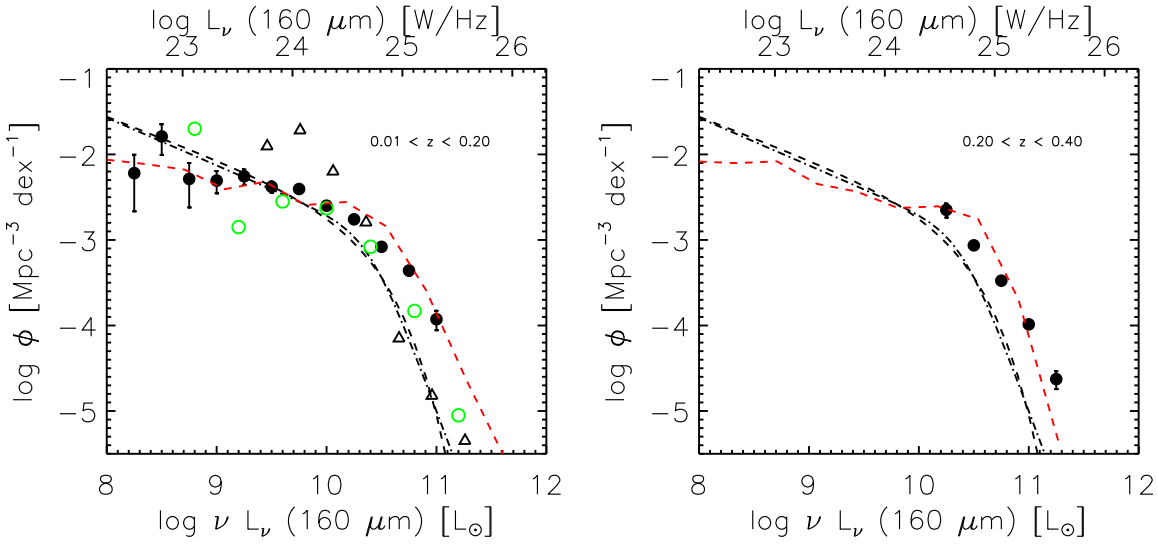

Figure 6.5: MIPS 160 μm Local Luminosity Function in two redshift bins within $0.01 < z < 0.4$ in the Lockman Hole field. Black full circles: our $1/V_{max}$ estimate; green empty circles: Patel et al. (2012) $1/V_{max}$ MIPS 160 μm luminosity function estimate at the same reported redshift bin; black empty triangles: Takeuchi et al. (2006) ISO 170 μm luminosity function estimate within $0 < z \lesssim 0.2$ converted to MIPS effective wavelengths using a spectral index of -1.0; dot-dashed and dashed lines are local luminosity function prediction at 160 μm by Serjeant & Harrison (2005) estimated following Eq. 2.64 and Eq. 2.67 respectively; Red dashed line: Fontanot et al. (2012) model prediction.

6.2 SPIRE Local Luminosity Functions

Thanks to Herschel FIR/SMM observations and its increased observational capabilities, with respect to previous IR instruments, we are now able to constrain in great detail the local shape of the IR LF and thus estimate a fundamental benchmark to then study the IR LF evolution using deeper observations. By using IR shallow and wide area observations, carried out by Herschel within HerMES, we are able to sample enough sky volume to accurately the luminosity/number distribution of low-z detected sources and thus to fix the best shape of the local LF. We can then (in Chap. 7) study the luminosity and density evolution of the LF (as described in Sec. 2.4) exploiting deep and small area surveys, which alone would not sample enough sky volume to describe the luminosity distribution of local sources and thus constrain the shape of the LF.

Preliminary results on the SPIRE LLF as been published in Vaccari et al. (2010) exploiting the first observations carried out within HerMES during Herschel's Science Demonstration Phase (SDP). In this Section we show how the increased statistics of more recent observations and the exploitation of more sophisticated statistical tools give us the possibility to put more rigid constraints on the local shape of the IR LLF, consolidating the preliminary results obtained during the SDP.

6.2.1 SPIRE 250 μm Sample in the HerMES Wide Fields

To estimate the SPIRE LLF we use the data of Herschel/HerMES L5/L6 observations (see Sec. 3.8.3 and 3.12 for more details on the observations) in 4 fields: Lockman Hole (LH), Extragalactic First Look Survey (XFLS), Bootes and ELAIS-N1 (EN1). These fields are covered by the Spitzer Data Fusion and the available ancillary data are thus described in Chap. 4. They are all part of both the Spitzer and SDSS survey programme and, crucially, they all enjoy a vast quantity of reasonably homogeneous ancillary observations and thus have uniform estimates of both photometric and spectroscopic redshifts by the SDSS plus some small spectroscopic surveys available from NED.

To estimate the SPIRE LLF we use the SPIRE flux estimates produced by the XID method explained in Roseboom et al. (in prep.), following Roseboom et al. (2010), illustrated in Sec. 3.8.4, applied on SCAT maps produced by Smith et al. (2012). SPIRE 250 μm channel is the more sensitive between SPIRE bands and thus we select the data on the basis of a SPIRE 250 μm reliability criterion (discussed Sec. 3.8.4) defined as: $\chi^2_{250} < 5$ and $SNRT_{250} > 4$. The SPIRE 250 catalogues of L5/L6 HerMES observations are uniformly highly complete and reliable down to 25/30/35 mJy at 250/350/500 μm respectively, as shown in Fig. 6.6 (left). In order to combine the data collected in these different fields (following Avni & Bahcall (1980) approach) we have to ensure a uniform completeness both in flux and in redshift coverage across fields, thus, due to some minor differences across the fields we decide to cut our sample down to 30 mJy. These minor differences are visible in Fig. 6.6 (righ) where we compare SPIRE 250 μm number counts estimated for the 4 fields and for the COSMOS deep field (described in Chap. 7). These discrepancies are consistent with the levels of cosmic variance predicted by theoretical models for fields of this size (Moster et al. 2011) as well as with the different depth of MIPS 24 μm observations available for these fields which were used to guide HerMES XID source extraction. In any case, we can notice that the differences are small and have major effects only at low fluxes, well below our selected flux limit. The

Field	250 μm	Area [deg^2]	Set
LH	2033 (925/1108)	11.0	34
XFLS	723 (425/298)	4.0	40
BOOTES	1738 (331/1407)	10.0	37
EN1	613 (111/502)	4.0	35
Total	5107 (1792/3315)	29.0	

Table 6.5: Number of $0 < z \lesssim 0.5$ sources used for the determination of SPIRE LLFs. The number of sources with (spectroscopic/photometric) redshifts is indicated after the total number of sources. The 250 μm sample is cut at $S250 > 30$ mJy according to the SPIRE 250 completeness. Set refers to Tab.3.12 and identifies the HerMES specific observing mode for each field.

greatest discrepancy is shown by XFLS where the SPIRE 250 μm detection reflect the known limitations of the MIPS 24 μm detections due to shorter exposure time with comparison to other SWIRE fields and Bootes (as described in Sec. 4.8).

Besides, since we only have SDSS photometric redshift estimates for some of our sources, we select them in order to keep only sources with SDSS $r_{AB} < 22.0$, thus brighter than the SDSS magnitude limit avoiding unreliable photometric redshifts. In Fig. 6.7 (low panels) we report SDSS r_{AB} and redshift histograms of our total sample composed by the 4 combined fields (LH+XFLS+Bootes+EN1) together, hereafter named "HerMES Wide Fields sample". In order to avoid effects of incompleteness in redshift, we limit our HerMES Wide Fields sample to $z \lesssim 0.5$, below the limit of reliability of SDSS redshift estimates. The number of sources of the HerMES Wide Fields sample is detailed in Tab. 6.5. Note we use the same SPIRE 250 μm selected sample to determine the LLF at 250/350/500 μm due to higher reliability and higher source density of the SPIRE 250 μm channel.

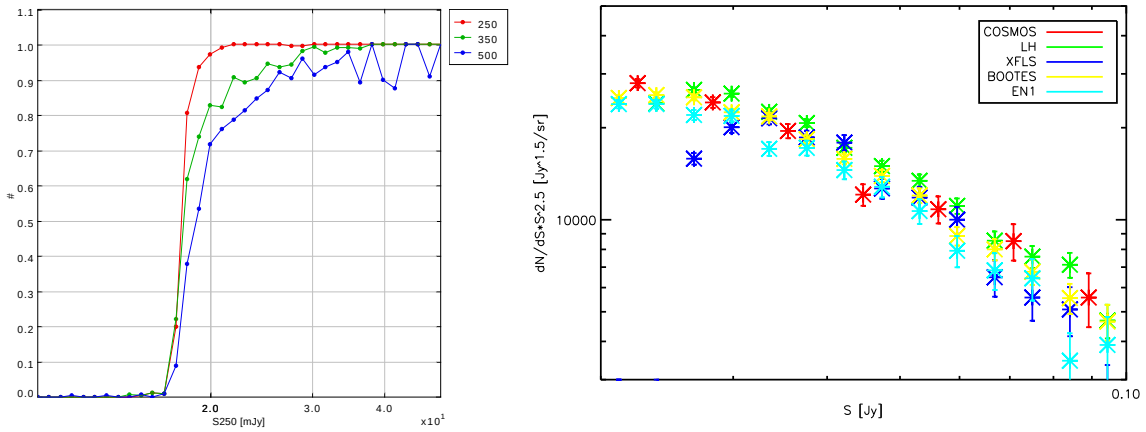

Figure 6.6: SPIRE 250 source counts and completeness based on XID catalogues by Roseboom et al (in prep.) for the HerMES Wide Fields sample used for the determination of SPIRE LLFs.

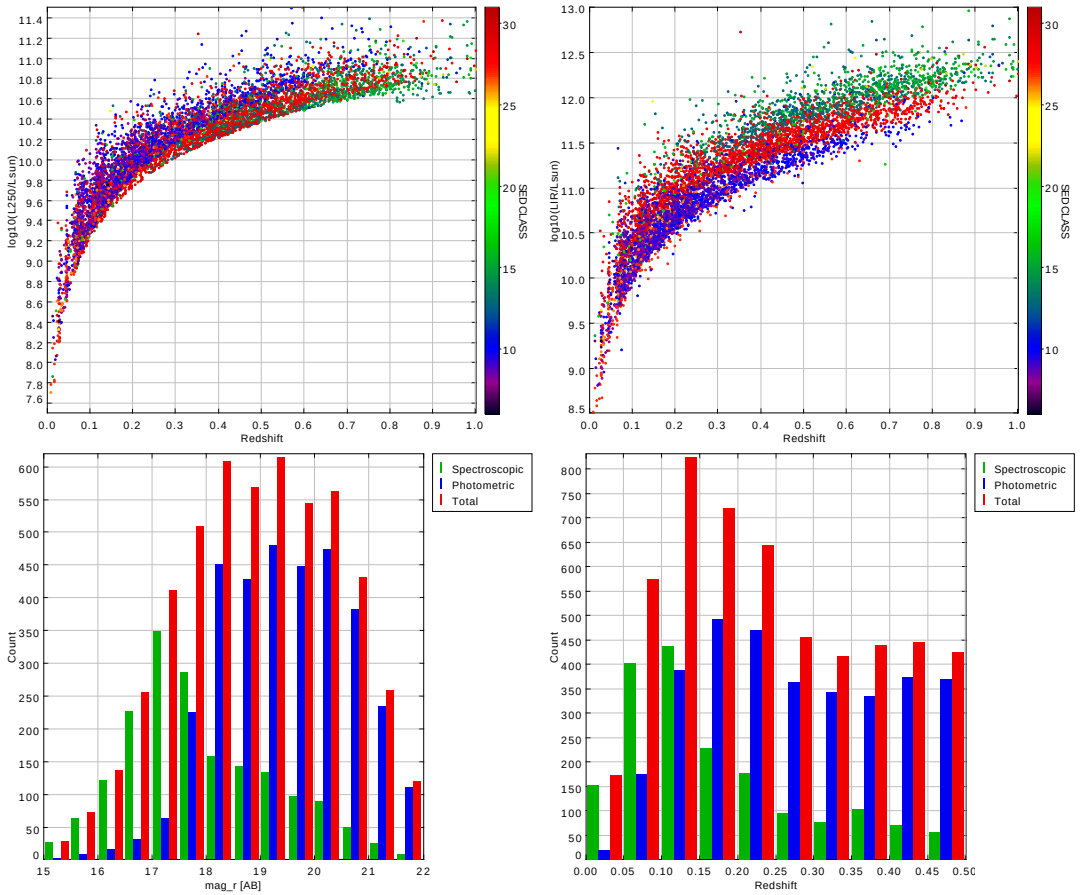

Figure 6.7: SPIRE 250 mum and IR Bolometric Luminosity vs redshift, SDSS r_{AB} and redshift histograms for the HerMES Wide Fields sample used for the determination of SPIRE LLFs. The $L - z$ plots are colour-coded according to the SED best fit class obtained by the SED fitting procedure following the list reported in Tab. 5.1. The histograms report the relative quantities for the photometric and spectroscopic samples in blue and in red respectively, against the total sample illustrated in red.

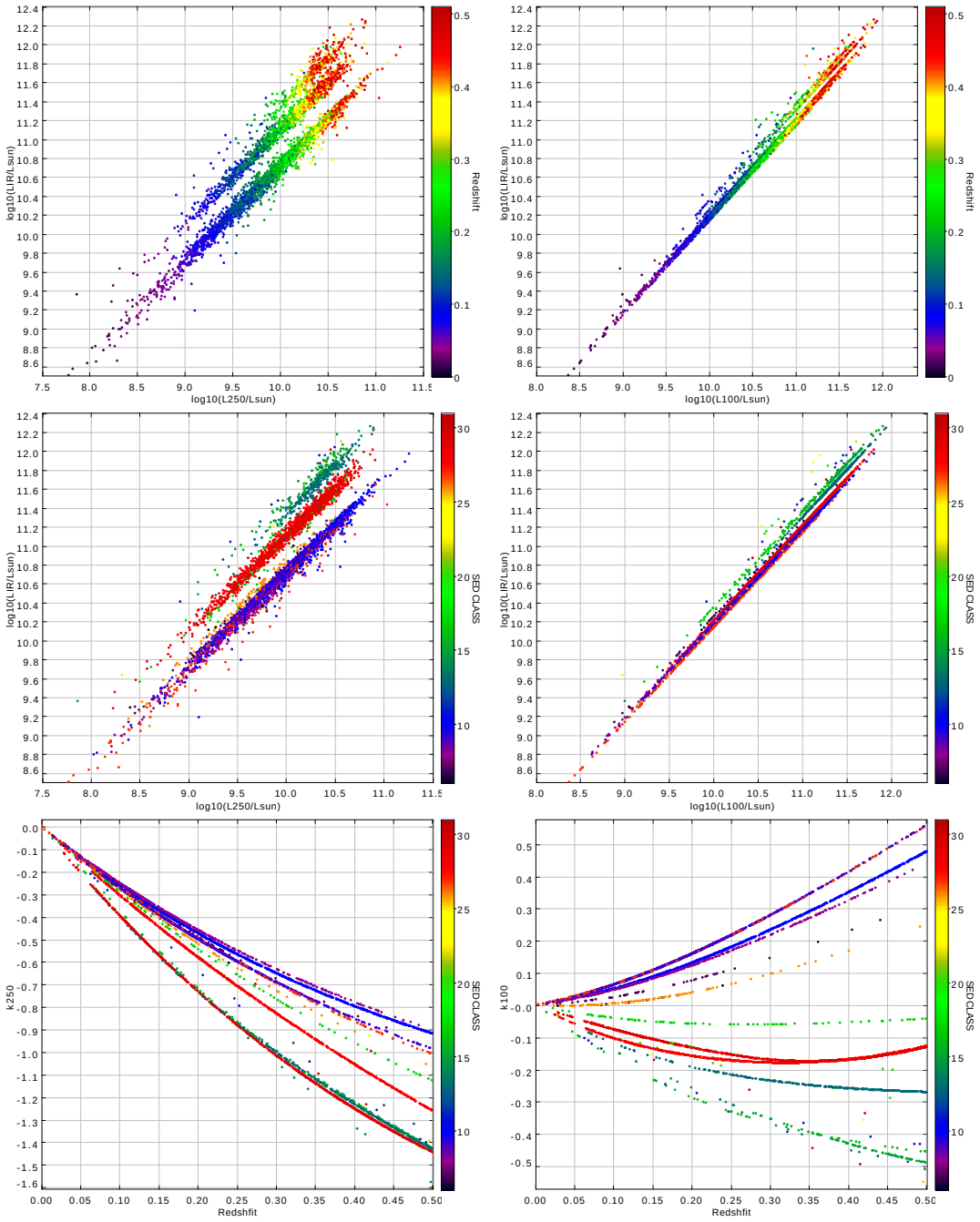

Figure 6.8: Correlation between rest frame SPIRE 250 μm/PACS 100 μm luminosities and IR bolometric luminosity; SPIRE 250 μm and PACS 100 μm k-corrections in function of redshift. **Top**: correlations between rest frame SPIRE 250 μm/PACS 100 μm luminosities and IR bolometric luminosity colour-coded according to the SED best fit class obtained by the SED fitting procedure following the list reported in Tab. 5.1. **Middle**: correlations between rest frame SPIRE 250 μm/PACS 100 μm luminosities and IR bolometric luminosity colour-coded as a function of redshift; **Bottom**: SPIRE 250 μm and PACS 100 μm k-corrections in function of redshift colour-coded according to the SED best fit class.

6.2.2 SPIRE Local Luminosity Function Estimates

Thanks to Spitzer Data Fusion we are able to perform the multi-wavelength SED fitting analysis of our HerMES Wide Fields sample and thus estimate the monochromatic and IR bolometric rest-frame luminosities and relative k-corrections. Following the discussion in Chap. 5 we perform the SED fitting analysis with `Hyperz` set as described in Sec. 5.2.1. To perform the fit we use only r_{AB}, i_{AB}, J_{2MASS}, IRAC-3.6, IRAC-4.5, IRAC-5.8, IRAC-8.0, MIPS-24, MIPS-70, MIPS-160, SPIRE-250 and SPIRE-350 magnitude, to balance the fit between the short and long wavelengths regimes, as already explained in Sec. 5.2.1. In particular, the SPIRE 350 μm and SPIRE 500 μ magnitudes to perform the SED fitting only if $SNRT_{350} > 5$ or $SNRT_{500} > 5$, respectively. As template SEDs we use the SWIRE templates by Polletta et al. (2007) and their slightly modified version by Gruppioni et al. (2010) listed in Tab. 5.1. In Fig. 6.7 (upper panels) we report the $L - z$ distribution for both L_{250} and LIR observed rest-frame luminosities obtained by the SED fitting procedure. Thanks to the multi-wavelength SED fitting procedure we are able also to investigate some classical relation extant between monochromatic res-frame luminosities at different λs. As an example we report in Fig. 6.8 a comparison between SPIRE 250 μm and PACS 100 μm monochromatic rest-frame luminosity respectively plotted against the IR bolometric luminosity. Historically, the 100 μm luminosity, as the one at MIPS 70 μm, has been considered a good indicators of the IR bolometric luminosity, due to a good correlation between the two (e.g., Patel et al. (2012) used the relation between MIPS 70 μm and the IR bolometric luminosity). In Fig. 6.8 we show that we confirm this trend by our SED fitting results while, on the contrary, the SPIRE 250 μm doesn't show a good correlation with the IR bolometric luminosity and thus it can not be used as an indicator of the total IR emission of the galaxy. This fact can be explained by looking into the very different behaviour of the k-corrections estimated at SPIRE 250 μm and PACS 100 μm respectively (lowest panels). The differences between these two are remarkable and this is reflected in the different behaviour of the resulting luminosities.

To estimate the LLFs in the SPIRE bands we exploit different LF estimators: the $1/V_{max}$ (Schmidt 1968) and the modified version $\phi_{\rm est}$ by Page & Carrera (2000), the bayesian parametric ML (Kelly et al. 2008; Patel et al. 2012) and the semi-parametric approach by Schafer (2007). All these method are explained in details in Chap. 2. Then we compare our results with precedent published results and to the $1/V_{max}$ estimates obtained using the COSMOS data in the same local redshift bins, as a connection to the results on the LF evolution of the following Chapter.

As for the MIPS LFs (following Eq. 6.1, 6.2 and 6.5) we estimate the z_{max}, and thus V_{max}, by combining all the selection factors that affects our local sample: $r_{AB} < 22$, MIPS 24 μm flux limit that guide the SPIRE 250 μm extraction, $S_{24} > 300$ μJy, and finally the flux limit in the SPIRE 250 μm band $S_{250} > 30$ mJy. We then estimate the LFs at SPIRE 250 μm as well as at 350 and 500 by using the SPIRE 250 μm selected sample extrapolating the luminosities at these other bands from the SED fitting results. The higher sensitivity of the SPIRE 250 μm channel with respect to 350/500 μm channels largely ensure that we do not miss sources detected only at these longer bands. Besides we estimate the IR bolometric luminosity functions using the integrated luminosity between 8 and 1000 μm.

As a summary, in Tab. 6.6 we report our $1/V_{max}$ luminosity function values for each bands and IR bolometric rest-frame luminosity per redshift bins. We also report the $\phi_{\rm est}$ values for the SPIRE 250 μm. We exclude from the caIculation the sources with $z < 0.02$ to avoid the

range of redshift where the photometric redshift uncertainties are higher and the results have higher associated uncertainties. The error associated with each value of ϕ and $\phi_{\rm est}$ is estimate following the Poissonian statistics as shown in Eq. 2.43.

$0 < z < 0.2$ **Luminosity Functions**

$\log L$	$\log(\phi,\sigma)_{250}$	$\log(\phi,\sigma)_{350}$	$\log(\phi,\sigma)_{500}$	$\log(\phi,\sigma)_{IR}$
8.00	-2.30, -2.45	-2.11 , -2.78	-2.22 , -3.29	-
8.25	-1.87, -2.46	-2.08 , -2.96	-2.24 , -3.47	-
8.50	-2.00, -2.74	-2.27 , -3.28	-2.40 , -3.74	-
8.75	-2.17, -3.03	-2.18 , -3.41	-2.63 , -4.03	-2.11 , -2.41
9.00	-2.19, -3.25	-2.38 , -3.70	-3.10 , -4.33	-1.97 , -2.55
9.25	-2.22, -3.45	-2.56 , -3.96	-3.78 , -4.67	-2.11 , -2.78
9.50	-2.38, -3.73	-2.98 , -4.27	-4.72 , -5.14	-2.12 , -2.89
9.75	-2.58, -4.01	-3.61 , -4.59	-	-2.16 , -3.04
10.00	-3.04 , -4.30	-4.453, -5.010	-	-2.30 , -3.44
10.25	-	-	-	-2.37 , -3.61
10.50	-	-	-	-2.50 , -3.79
10.75	-	-	-	-2.83 , -4.10
11.00	-	-	-	-3.28 , -4.39
11.25	-	-	-	-3.92 , -4.74
11.50	-	-	-	-4.56 , -5.06
11.75	-	-	-	-5.09 , -5.32
12.00	-	-	-	-5.56 , -5.56

$0.2 < z < 0.4$ **Luminosity Functions**

$\log L$	$\log(\phi,\sigma)_{250}$	$\log(\phi,\sigma)_{350}$	$\log(\phi,\sigma)_{500}$	$\log(\phi,\sigma)_{IR}$
8.50	-	-	-3.83 , -4.43	-
8.75	-	-	-3.15 , -4.07	-
9.00	-	-4.96 , -4.96	-2.83 , -4.12	-
9.25	-	-3.28 , -4.11	-3.31 , -4.67	-
9.50	-	-2.80 , -4.05	-3.91 , -5.11	-
9.75	-3.29 , -4.04	-3.17 , -4.56	-4.86 , -5.60	-
10.00	-2.77 , -4.10	-3.82 , -5.06	-5.56 , -5.95	-
10.25	-3.21 , -4.66	-4.75 , -5.54	-6.33 , -6.33	-5.27 , -5.27
10.50	-3.86 , -5.10	-5.49 , -5.91	-	-3.36 , -4.11
10.75	-4.84 , -5.59	-6.03 , -6.18	-	-2.95 , -4.16
11.00	-5.73 , -6.03	-	-	-3.11 , -4.24
11.25	-6.03 , -6.18	-	-	-3.38 , -4.61
11.50	-	-	-	-3.77 , -4.96
11.75	-	-	-	-4.38 , -5.36
12.00	-	-	-	-5.10 , -5.72
12.25	-	-	-	-6.33 , -6.33

Table 6.6: SPIRE 250/350/500 μm and IR bolometric rest-frame $1/V_{max}$ luminosity function values in the redshift ranges $0 < z < 0.2$ and $0.2 < z < 0.4$ using the HerMES Wide Fields sample; $\phi_{\rm est}$ values are reported for SPIRE 250 μm. L indicates $\nu\, L_\nu$ for the monochromatic LFs and L_{IR} indicates the integrated luminosity between 8 and 1000 μm for the IR bolometric rest-frame LF. L is expressed in unit of $L_\odot$ and LLF estimates and their errors are in [$\mathrm{Mpc}^{-3}\,\mathrm{dex}^{-1}$].

We perform the bayesian parametric ML (described in Sec. 2.3.3) by using the HerMES Wide Fields sample within the redshift bin $0.02 < z < 0.2$ and only for the SPIRE 250 μm

data, since this is our guiding band to other LF estimates.To find the best ML solution we performed a number of tests exploiting different functional forms shown in Sec. 2.3.3 and we then decide to report the best solution only for the modified Schecter fuction or log-gaussian functional form defined in Eq. 2.63 that is the choice made by different authors working in the infrared such as Saunders et al. (1990) and Gruppioni et al. (2010) with whom we want to compare our results. The detailed method is described in Sec. 2.3.3; here we only want to point out that, while Eq. 2.57 gives the expected number of objects in a sample composed by sources of the same morphological type and collected in a single field survey, in this case we have to change the equation to the following

$$\lambda = \sum_{SED} \sum_{fields} \iint \Phi(L, z|\theta) p(selected|L, z) d\log L \frac{\mathrm{d}V}{\mathrm{d}z} \mathrm{d}z \tag{6.8}$$

where we sum together the expected number of sources for each SED type, resulting from the SED fitting procedure, and surveys areas that describe our HerMES Wide Fields sample. In Tab. 6.7 we report the best solution for the best guess parameters of the log-gaussian functional form resulting by the MCMC procedure run on the HerMES Wide Fields sample with 5×10^5 iterations, while in Fig. 6.9 we report the histograms of the parameters produced by the MCMC procedures.

Parameter		σ
$\log(\mathrm{L}_*)$ $[\mathrm{L}_\odot]$	$9.71^{+0.05}_{-0.09}$	0.07
α	$1.22^{+0.04}_{-0.04}$	0.04
σ	$0.25^{+0.02}_{-0.02}$	0.02
ϕ_* $[\mathrm{Mpc}^{-3}\mathrm{dex}^{-1}]$	$-2.20^{+0.05}_{-0.04}$	0.04

Table 6.7: Best fitting parameters for the local SPIRE 250 μm determined using the parametric bayesian method. To obtain the fit we use the total HerMES Wide Fields sample within the redshift range $0 < z < 0.2$.

A summary of the results of our estimations is reported in the following figures. In Fig. 6.10, 6.11 we report the SPIRE 250 μm rest-frame LF estimated by using the $1/V_{max}$ and the parametric ML estimators using data in the sum of the fields in the HerMES Wide Field samlpe and for the single fields respectively. In Fig. 6.13 we report the SPIRE 250 μm rest-frame LF estimated by using the semi-parametric estimator described in Sec. 2.3.5 and the modified $1/V_{max}$ estimates following Page & Carrera (2000) described in Sec. 2.3.1. In Fig. 6.14, 6.15, 6.16 we report the SPIRE 350/500 μm and IR bolometric rest-frame LFs respectively. Finally in Fig. 6.17 we report the SPIRE 250 μm as well as 60/90 μm rest-frame LFs estimated in more local bins and compared with local predictions at these wavelengths given by different authors.

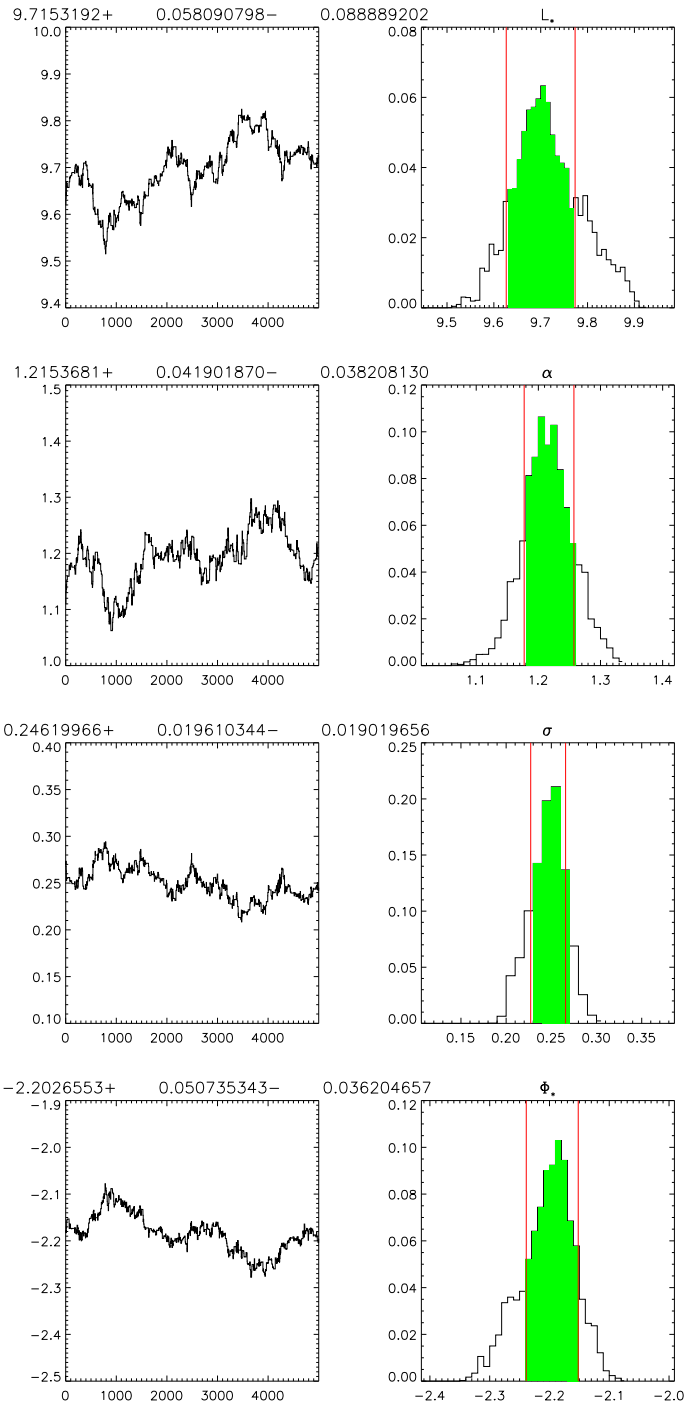

Figure 6.9: Series of values and histograms of the best fitting parameters for the SPIRE 250 μm local luminosity function within $0 < z < 0.2$ determined using the parametric bayesian method. In the series plot we report a tenth of the complete series coming from the MCMC procedure run with $5{\times}10^5$ steps.

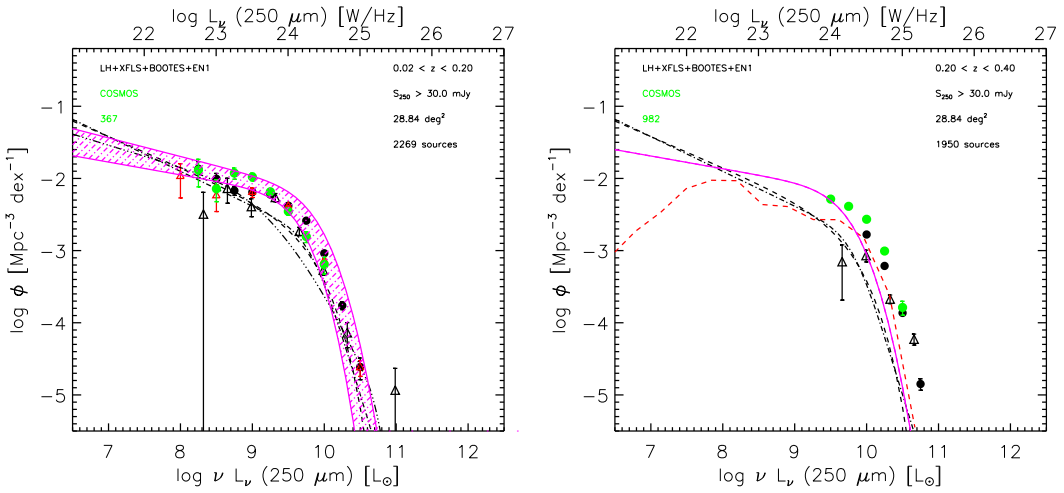

Figure 6.10: The SPIRE 250 μm local luminosity function estimates using the HerMES Wide Fields sample. Black filled circle: our $1/V_{max}$ estimates. Green filled circles: our $1/V_{max}$ results using COSMOS data (area 1.7 deg^2 and flux limited $f250 > 10$ mJy, see Chap. 7); red empty triangle: SPIRE 250 luminosity function estimate by Vaccari et al. (2010); black empty triangles: SPIRE 250 μm luminosity function by Eales et al. (2010b); red dashed line: Fontanot et al. (2012) model; black dashed dot dot dot line: Negrello et al. (2007) model. The magenta shaded region is the 1 σ best MCMC solution using the functional form by Saunders et al. (1990). The magenta line is the mean solution from the MCMC solution.

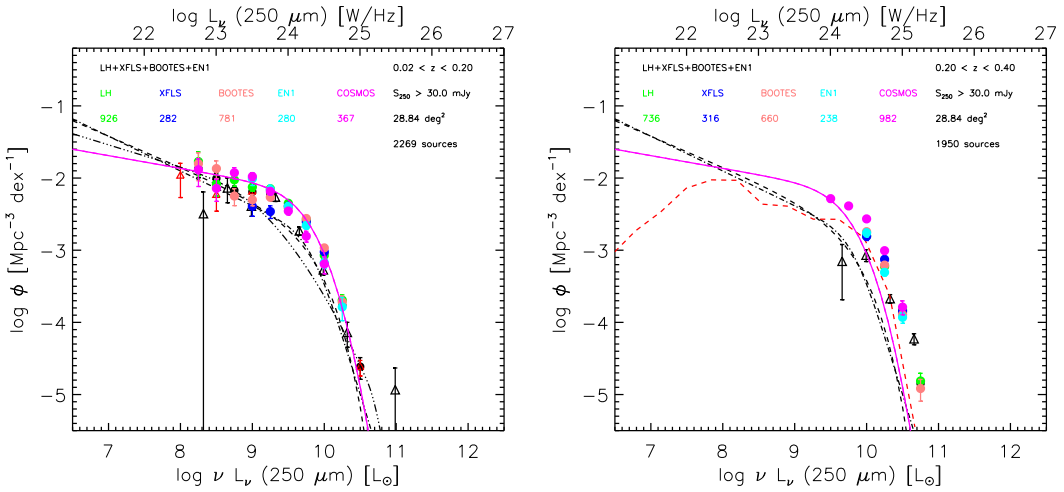

Figure 6.11: The SPIRE 250 μm local luminosity function estimates field to field in the HerMES Wide Fields sample. Colour-coded full circles: our $1/V_{max}$ results for each field; magenta full circles: our $1/V_{max}$ results using COSMOS data (area 1.7 deg^2 and flux limited $f250 > 10$ mJy, see Chap. 7); black empty triangles: SPIRE 250 μm luminosity function by Eales et al. (2010b); red dashed line: Fontanot et al. (2012) model; black dashed dot dot dot line: Negrello et al. (2007) model. The magenta line is the mean solution from the MCMC solution using data of the total HerMES Wide Fields sample.

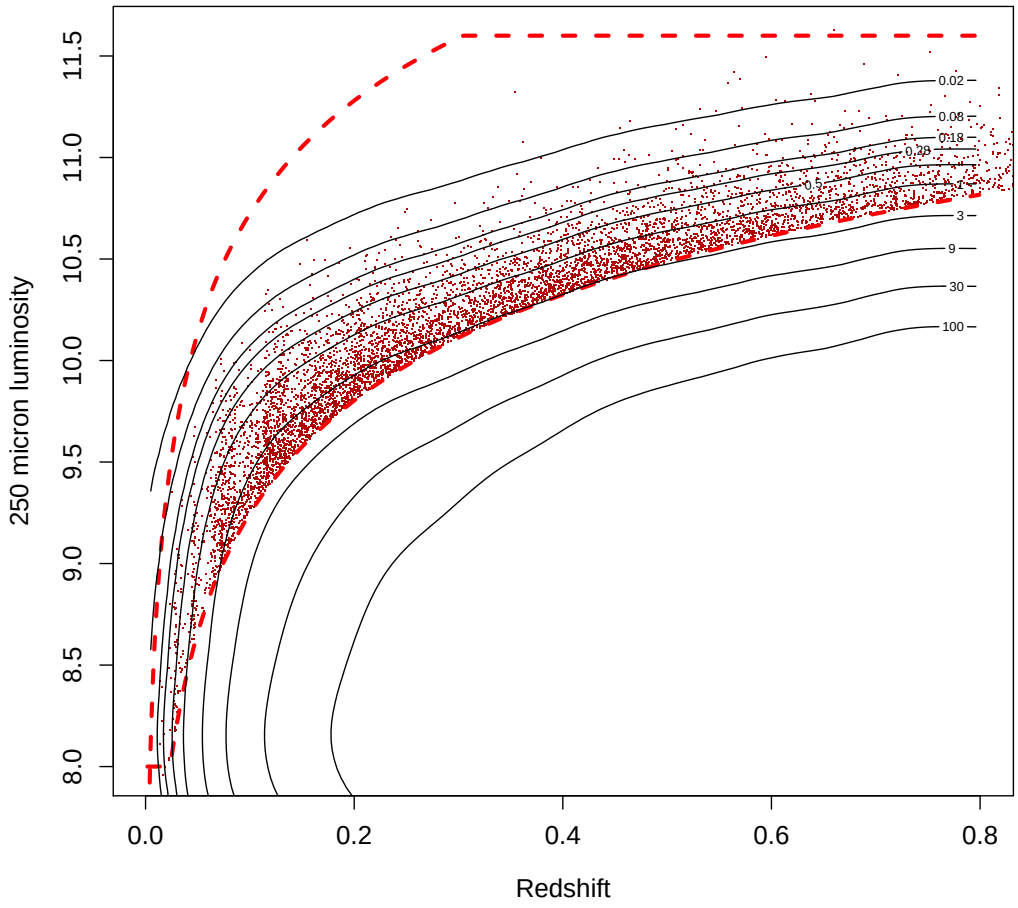

Figure 6.12: Distributions of SPIRE sources from the HerMES Wide Field sample in the 250 μm luminosity *vs* redshift plane as reconstructed by Schafer (2007) estimator. Red points: the data; the red dashed lines mark the flux limitations adopted in the application of the *semi-parametric* LF estimator by Schafer (2007); the continuous black lines are iso-density contours corresponding to the *semi-parametric* reconstrunctions of the source volume density as a function of luminosity and redshift.

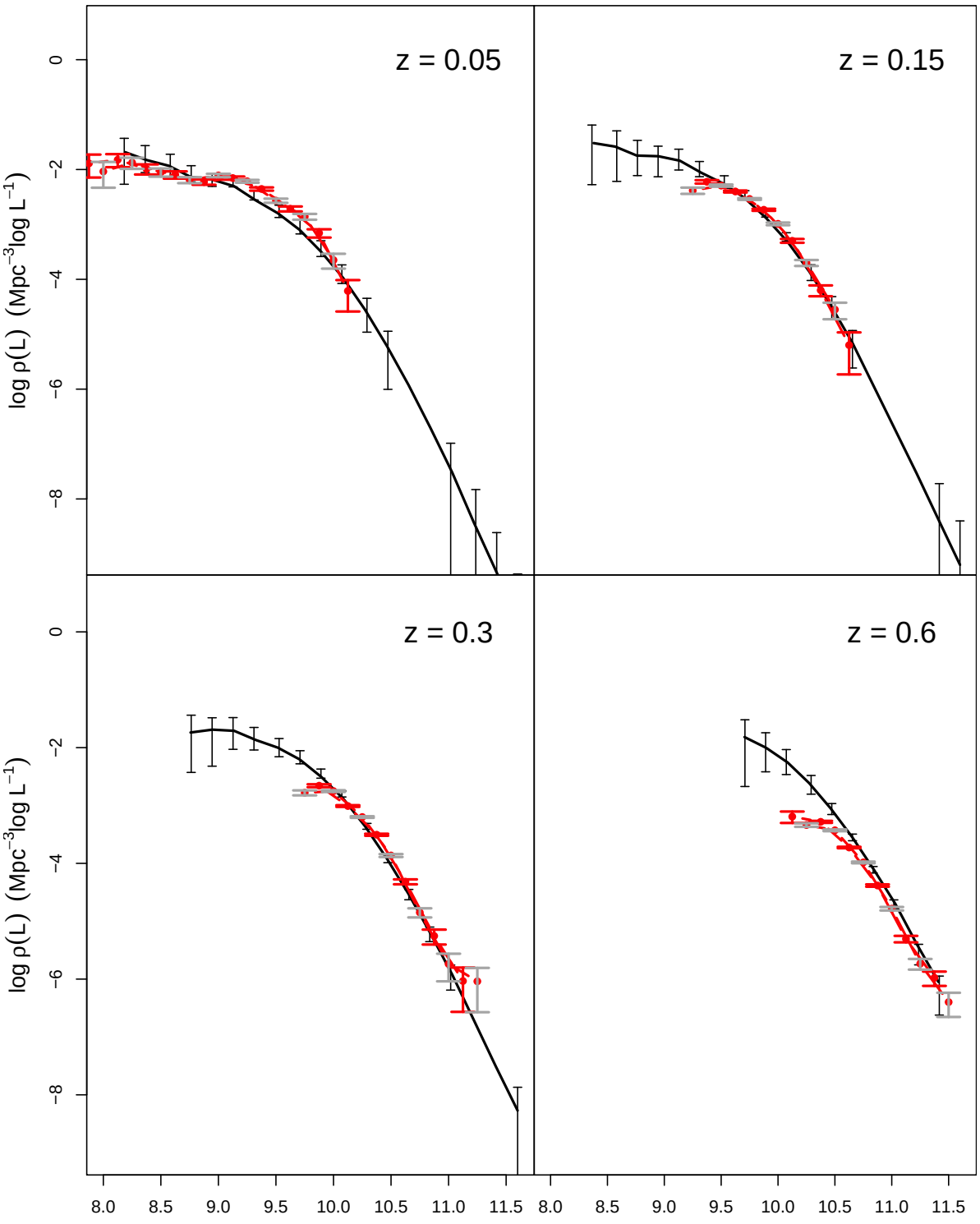

Figure 6.13: SPIRE 250 μm Local luminosity function estimated using the semi-parametric method by Schafer (2007) and the modified $1/V_{max}$ by Page & Carrera (2000). In grey our classic $1/V_{max}$ estimate, in red the estimate using Page & Carrera (2000) and in black the estimate using Schafer (2007).

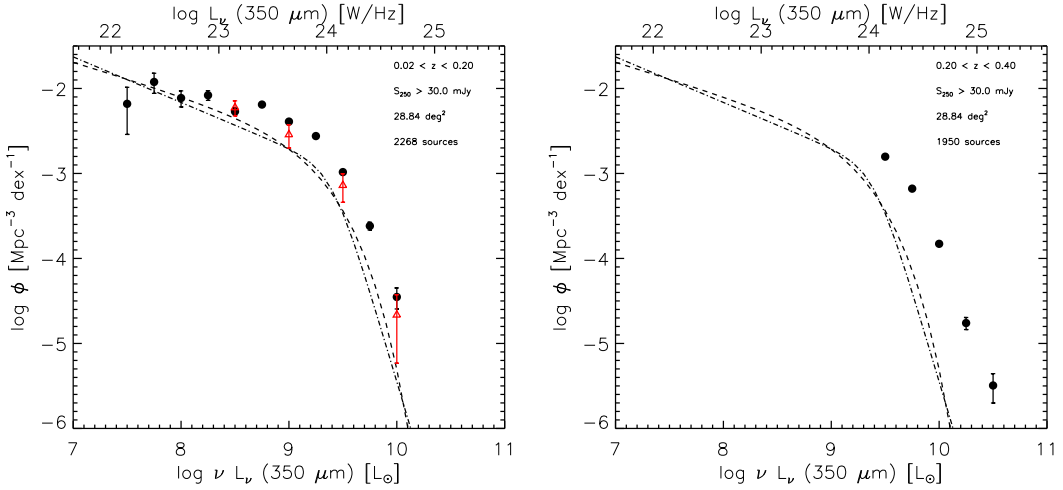

Figure 6.14: The SPIRE 500 local luminosity function estimates in the HerMES Wide Fields sample. Black full circles: our $1/V_{max}$; red empty triangles: SPIRE 500 μm local luminosity function estimate by Vaccari et al. (2010); black dot-dashed and dashed lines are local luminosity function prediction at 350 μm by Serjeant & Harrison (2005) estimated following Eq. 2.64 and Eq. 2.67 respectively.

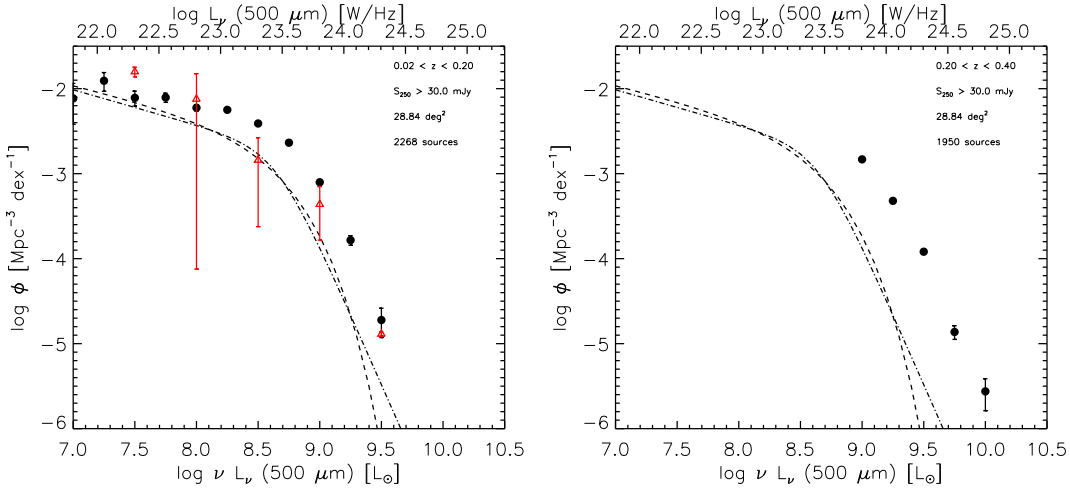

Figure 6.15: The SPIRE 500 μm local luminosity function estimates in the HerMES Wide Fields sample. Black full circles: our $1/V_{max}$ results; red empty triangles: SPIRE 500 μm local luminosity function estimate by Vaccari et al. (2010); black dot-dashed and dashed lines are local luminosity function prediction at 500 μm by Serjeant & Harrison (2005) estimated following Eq. 2.64 and Eq. 2.67 respectively.

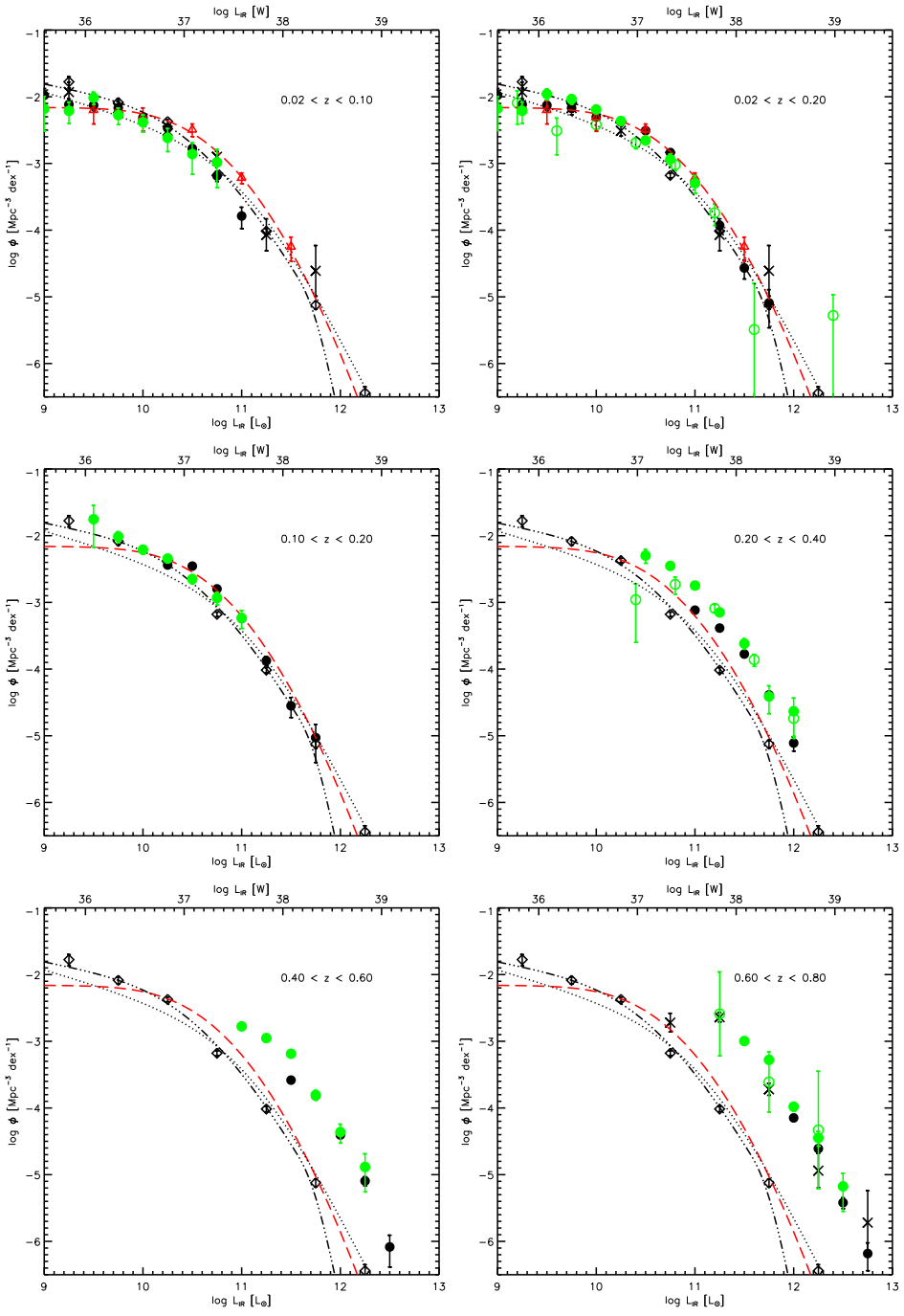

Figure 6.16: The IR bolometric rest-frame local luminosity functions in the HerMES Wide Fields sample. Black full circles: our $1/V_{max}$ results; green full circles: our $1/V_{max}$ results using COSMOS data (area 1.7 deg^2 and flux limited $f250 > 10$ mJy, see Chap. 7); green empty circles: the SWIRE IR bolometric rest-frame luminosity function by Patel et al. (2012) using a MIPS 70 and 160μm selected sample in LH and XMM-LSS; red empty triangles: IR bolometric rest-frame luminosity function estimate by Vaccari et al. (2010); red dashed line: χ^2 fit to the Vaccari et al. (2010) result using the functional form by Saunders et al. (1990); black empty diamonds: IRAS IR bolometric rest-frame luminosity function by Sanders (2003); black crosses: Rodighiero et al. (2010); black dashed dot dot dot line: Negrello et al. (2007) model; black dotted line: Valiante et al. (2009) model.

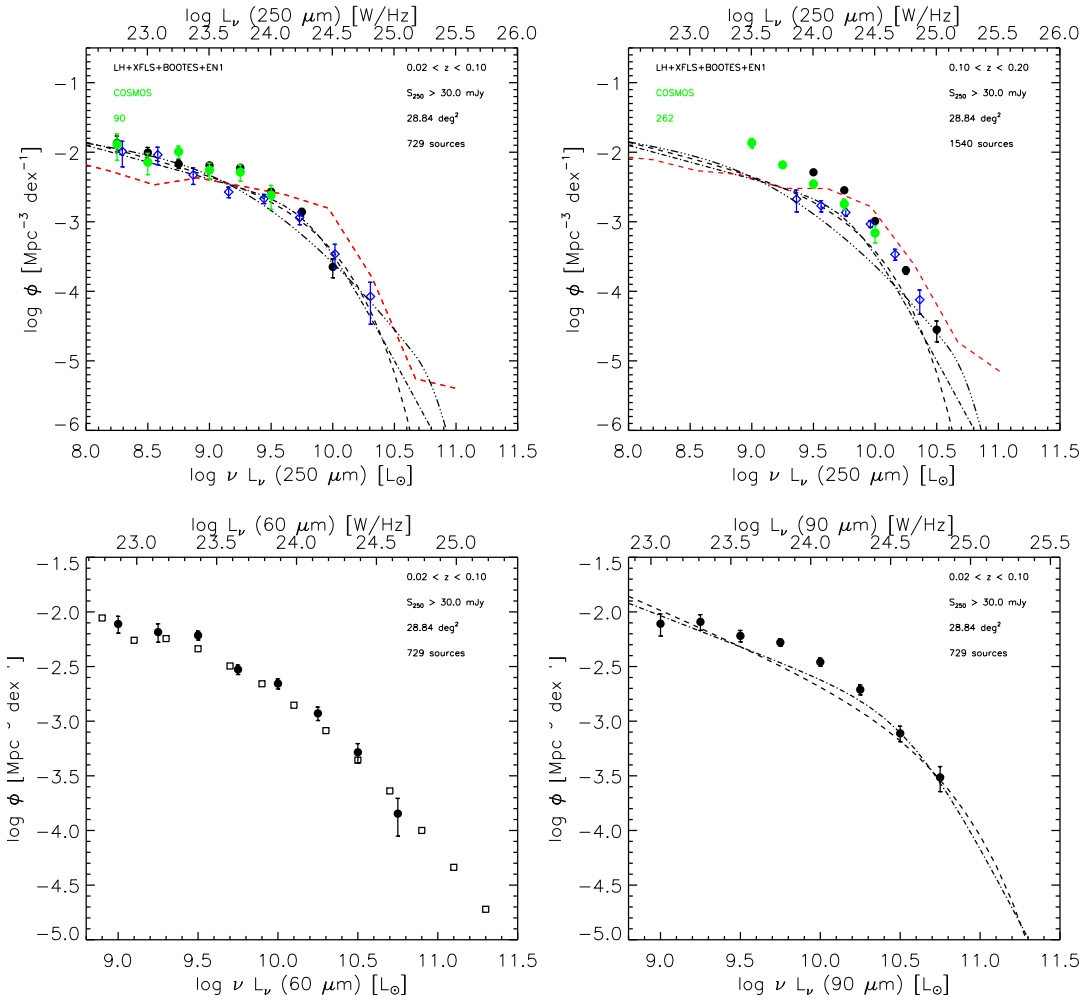

Figure 6.17: Very local bins for the SPIRE 250 μm luminosity functions and two extrapolation for the local luminosity functions at 60/90 μm using the HerMES Wide Fields sample. Black full circles: our $1/V_{max}$; green full circles: our $1/V_{max}$ results using COSMOS data (area 1.7 deg^2 and flux limited $f250 > 10$ mJy, see Chap. 7); empty blue diamonds: Dye et al. (2010) SPIRE 250 μm luminosity function estimates using the first ~ 14 deg^2 observed within the Herschel-ATLAS projects ; empty black squares: 60 μm local luminosity function estimate by Saunders et al. (1990); red dashed line: Fontanot et al. (2012) model; black dashed dot dot dot line: Negrello et al. (2007) model; black dot-dashed and dashed lines are local luminosity function prediction at 250/60/90 μm by Serjeant & Harrison (2005) estimated following Eq. 2.64 and Eq. 2.67 respectively.

6.3 Infrared Local Luminosity Density Estimates

Using some of the widest-area surveys carried out by Spitzer and Herschel, in this Chapter we studied the Local Luminosity Function of MIPS and SPIRE sources in unprecedented detail. Our LLFs at 24/70/160 μm (MIPS) and at 250/350/500 μm (SPIRE) put strong constraints on the Local Luminosity Density (LD) of the Universe throughout the MIR/FIR/SMM wavelength range. While the bulk of the discussion will be deferred to Chap. 7, here we stress a few points.

The MIPS 24/70/160 μm LLF was estimated in the Lockman Hole field and compared against previous measurements, showing a good agreement and improved constraints with respect to previous measurements. The possibility to extend the sample to cover an area at least three times as large as presented here will be pursued in the future and will lead to substantial improvements in our knowledge of the Local Luminosity Density at MIR/FIR wavelengths.

The SPIRE LLFs in different fields do not show any discrepancy beyond the small differences to be expected on the basis of cosmic variance. Our estimates mostly confirm and improve upon the SDP results published in Vaccari et al. (2010), thanks to our increased statistics. From these preliminary comparisons we notice that Fontanot et al. (2012) models seem to have some problems in reproducing the higher luminosity bins of our estimates at all considered redshift bins, but we are mostly in agreement with this model in the other luminosity bins.

As already discussed at the beginning of this Chapter and as we will stress in Chap. 7, the estimate of the local luminosity function in the SPIRE bands is of fundamental importance when studying the evolution of the SPIRE LFs at higher redshift. In practice, local luminosity function estimates guide the priors to be adopted when fitting the shape of the LF also at higher redshifts. Besides, thanks to the large volume sampled by shallow and wide areas surveys, these estimates allow us to calculate with very low level of uncertainties the SFRD in the local Universe. As explained in Sec. 2.4 by integrating the luminosity function in different redshift bins, whenever the observed bands are somewhat related to the emission of the young stellar components, we can evaluate an estimate of the SFR at those redshifts. In the case of the SPIRE local luminosity function we can easily use the IR bolometric luminosity as a tracer of SFR (as explained in Sec. 5.4) and thus the IR bolometric luminosity density as tracer of the SFRD.

We thus fit our $1/V_{max}$ local luminosity function estimates with a modified Schechter function described by Eq. 2.63. In this way we obtain the estimate of the local luminosity density (LLD) at each SPIRE bands and for the IR bolometric luminosity. The LLD values are reported in Tab. 6.8 and Tab.LLD.2. We then convert the estimate of the luminosity density into star formation rate density using the Kennicutt (1998) relation explained in Sec. 5.4 by using Eq. 5.12 and 5.15. The SFRD estimate we obtained from the IR bolometric luminosity density estimated at $0 < z < 0.1$ and $0 < z < 0.2$ are reported in Tab. 6.10 together with other SFRD estimates obtained by different authors using different SFR tracers. Our measurements are in good agreement with previous measurements at a variety of wavelengths, confirming we are not missing any substantial star formation activity in the local universe by adopting a SPIRE selection.

Band [μ**m**]	$\log(\mathrm{LLD}_{0<z<0.1})$	$\log(\mathrm{Err}_{+})$	$\log(\mathrm{Err}_{-})$
250	7.08	0.17	-0.18
350	6.58	0.17	-0.19
500	6.04	0.16	-0.17
IR	7.89	0.14	-0.15

Table 6.8: Local luminosity density estimates in the SPIRE 250/350/500 μm and for the IR bolometric luminosity using the local SPIRE sample within $0 < z < 0.1$. The values are reported as log(LLD) and log(errors) and are express in $\mathrm{L}_{\odot}$ Mpc^{-1}.

Band [μ**m**]	$\log(\mathrm{LLD}_{0<z<0.2})$	$\log(\mathrm{Err}_{+})$	$\log(\mathrm{Err}_{-})$
250	7.24	0.12	-0.12
350	6.75	0.12	-0.13
500	6.20	0.13	-0.13
IR	8.08	0.12	-0.12

Table 6.9: Local luminosity density estimates in the SPIRE 250/350/500 μm and for the IR bolometric luminosity using the local SPIRE sample within $0 < z < 0.2$. The values are reported as log(LLD) and log(errors) and are express in $\mathrm{L}_{\odot}$ Mpc^{-1}.

Reference	**SFR tracer**	$< z >$	SFRD (10^{-3} $M_{\odot}$ yr^{-1} Mpc^{-1})
Gallego et al. (2002)	[OII]	0.025	9.3 ± 3
Sullivan et al. (2000)	[OII]	0.15	23 ± 3
Hogg et al. (1998)	[OII]	0.20	11 ± 4
Gallego et al. (1995)	Hα	0.022	12 ± 5
Tresse & Maddox (1998)	Hα	0.2	25 ± 4
Sullivan et al. (2000)	Hα	0.15	14 ± 3
Pérez-González et al. (2003)	Hα	0.025	25 ± 4
Ly et al. (2007)	Hα	0.08	13 ± 4
Hanish et al. (2006)	Hα	0.01	16^{+2}_{-4}
Brinchmann et al. (2004)	Hα	0.15	29 ± 5
Dale et al. (2010)	Hα	0.16	10^{+6}_{-4}
Westra et al. (2010)	Hα	0.05	6 ± 2
Westra et al. (2010)	Hα	0.15	12 ± 3
Serjeant et al. (2002)	1.4 GHz	0.005	21 ± 5
Condon (1989)	1.4 GHz	0.005	21 ± 0.5
Sullivan et al. (2000)	FUV	0.150	39 ± 5
Martin et al. (2005b)	FUV+IR	0.02	21 ± 2
Bothwell et al. (2011)	FUV+IR	0.05	25 ± 1.6
Vaccari et al. (2010)	IR	0.1	22.3 ± 8.2
This work	IR	0.05	13.1 ± 3.6
This work	IR	0.1	20.5 ± 5.0

Table 6.10: Star formation rate density in the local Universe: literature results and from this work. This table is an updated version of the one reported in Bothwell et al. (2011).

Chapter 7

Evolution of Herschel Sources & Redshift-Dependent Luminosity Functions

The determination of the evolution of the Galaxy Luminosity Function is a challenging observational and statistical task, requiring a complete and reliable sample over a wide range of redshifts and luminosities, and thus requiring to image even relatively small fields down to very faint flux limits, with a correspondingly large investment in observing time.

The key physical quantity under investigation here is the IR bolometric luminosity, that Herschel allowed us to reliably measure for the first time for sources at virtually any epochs. This flux is then traceable to a measure of the rate at which primeval gas is transformed into stars, or is accreted into a nuclear black hole. Therefore, our data, in coordination with complementary information from the UV/Optical/NIR, will be needed to recover first reliable estimates of the history of generations of stellar populations in galaxies and that of gravitational nuclear accretion.

In this Chapter, we study the evolution of the FIR/SMM Luminosity Function and of the inferred Star Formation Rate Density using sources detected by the Herschel SPIRE imager in the COSMOS field, where the favourable k-correction in the SMM, the SPIRE sensitivity and the excellent ancillary data available in the field allow us to detect powerful starbursts up to a redshift of about 5.

7.1 PEP/HerMES 250 μm Sample in the COSMOS Field

The Cosmic Evolution Survey (COSMOS, Scoville et al. (2007)) is a project trying to bridge the gap in galaxy evolution studies between the deepest pencil-beam surveys such as the HUDF and GOODS, covering areas from tens to hundreds of arcmin^2, and shallower surveys such as SWIRE and SDSS covering tens to thousands of deg^2. The backbone of the project is a one-band HST WFC survey and a 30-band survey at UV/Optical/NIR wavelengths covering a 2 deg^2 area, complemented by extensive optical spectroscopy and accurate photometric redshifts as well as observations with virtually all ground-based and space-based large telescopes.

Due to its combination of size and depth, the COSMOS field is the field of choice for robust

galaxy evolutions studies needing large samples of faint objects. The PEP and HerMES projects have thus joined forces to observe the COSMOS field extensively with PACS and SPIRE, respectively, providing 5-band FIR/SMM photometry at 100/160/250/350/500 mum.

The PEP/HerMES sample was extracted from PACS and SPIRE maps using the techniques by Berta et al. (2011) and Roseboom et al. (2010) respectively, using the COSMOS MIPS 24 μm sample by Le Floc'h et al. (2009) reaching down to about 80 μjy as positional priors. The combination of MIPS, PACS and SPIRE depths is such that it allows us to be complete down to approximately 5/10/10/15/20 mJy at 100/160/250/350/500 μm (Berta et al. 2011; Roseboom et al. 2010). In particular, according to extensive simulations taking into account instrumental confusion and source clustering carried out by Roseboom et al. (2010) and detailed by Roseboom et al. (in prep), the SPIRE XID catalogue is highly complete and reliable down to about 10 mJy at 250 μm when cutting the sample at $SNRT_{250} > 3$ (including confusion) and $\chi^2_{250} < 5$. Since the SPIRE 250 μm channel is the most sensitive one of the PACS and SPIRE imagers, we decided to select our sample at this wavelength as specified above and to use PACS fluxes at 100 and 160 μm (when $SNR > 3$) and SPIRE fluxes at 350 μm and 500 μm (when $SNRT_{350,500} > 3$, and $\chi^2_{350,500} < 5$) to complete the FIR/SMM SED of our sources. While a $S250 > 10$ mJy sample may in principle be affected by source confusion at SPIRE wavelengths, early HerMES science work demonstrates how our source extraction method avoids most of the related issues. As a further check, we tried cutting the samples at brighter fluxes and higher $SNRT$ levels, i.e. at $S250 > 15$ mJy and $SNRT_{250,350,500} < 4$ or $S250 > 20$ mJy and $SNRT_{250,350,500} < 5$, and verified how this did not significantly affect our results.

In order to identify the optical counterparts to these MIPS/PACS/SPIRE sources, we rely on the latest Incarnation (Version 2.0) of the COSMOS photometric redshift catalogue by Ilbert et al. (in prep), which is an updated version of the one presented in Ilbert et al. (2009) and Ilbert et al. (2010)) and is based on the latest COSMOS multi-wavelength photometric catalog and compilation of spectroscopic redshifts by Capak et al. (in prep). We carry out a MIPS-24 vs IRAC nearest-neighbour matching using a 1.5 arcsec search radius whose results are shown in Fig. 7.1. The astrometric registration of the MIPS 24 μm sources (used to extract the PACS/SPIRE sources) and of the IRAC sources (used for the association with optical photometry) is excellent and with an estimated error well within the search radius. The optical identification and redshift completeness achieved through this process is of 95%, and does not depend on SPIRE flux, suggesting it does not depend on redshift either. In the following we thus make no further correction for optical photometry and/or photometric redshift incompleteness.

Simulated and achieved SPIRE 250 μm completeness levels are illustrated by Fig. 7.2, showing the results of simulations by Roseboom et al. (in prep) based on Roseboom et al. (2010) (Left), based on a realistic model of instrumental features and source properties, and the results of comparing observed counts as a function of redshift with the stacking analysis by Béthermin et al. (2012) (Right), estimating the counts as a function of redshift which. While simulations are bound to be optimistic to some extent, the stacking analysis, which was carried out in the same COSMOS field and used a slightly less complete version of the COSMOS photometric redshift catalogue, confirms that our sample is fundamentally complete down to 10 mJy at 250 μm.

The main properties of the sample which was thus selected are described in Fig. 7.3 and in Tab. 7.1. The redshift distribution is not particularly regular, showing some patterns at

both low and high redshifts. The former low-z patterns are likely to be due to actual cosmic structures spanning across the relatively small field, while the latter high-z patterns may be caused by biases in the photometric catalog and/or in the photometric redshift method. which however has been extensively tested and employs the best photometric dataset currently available in any field of a similar size.

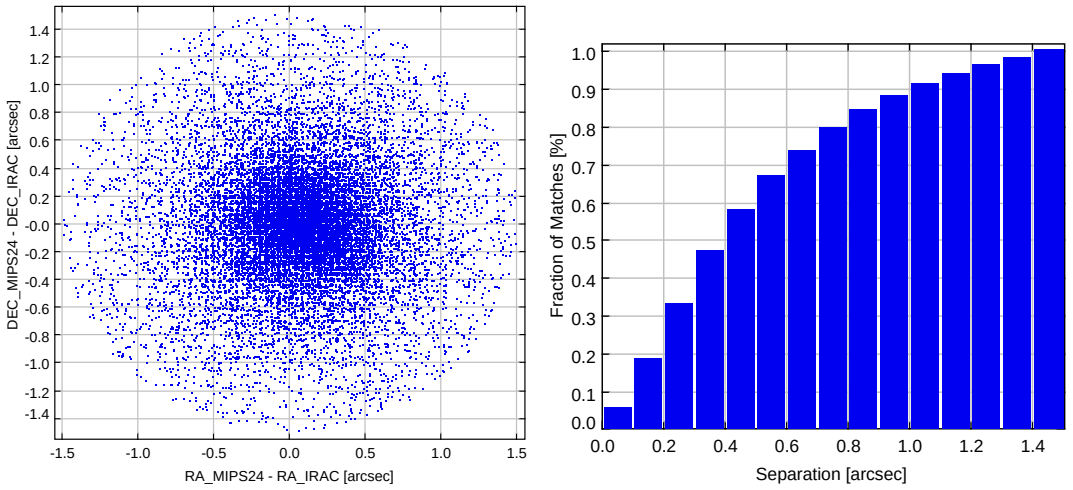

Figure 7.1: Optical Identification of SPIRE 250 μm Sources in the COSMOS Field. Left: MIPS μm vs IRAC Astrometric Offset. SPIRE sources are detected based on MIPS 24 μm priors and are associated to their optical counterparts using their IRAC coordinates. The limited (and symmetric astrometric offset between the two confirms the goodness of both astrometric solutions and thus of the matching process. Right: relative identification rate as a function of search radius. The 1.0 identification rate at 1.5 arcsec corresponds to our total 95% identification rate.

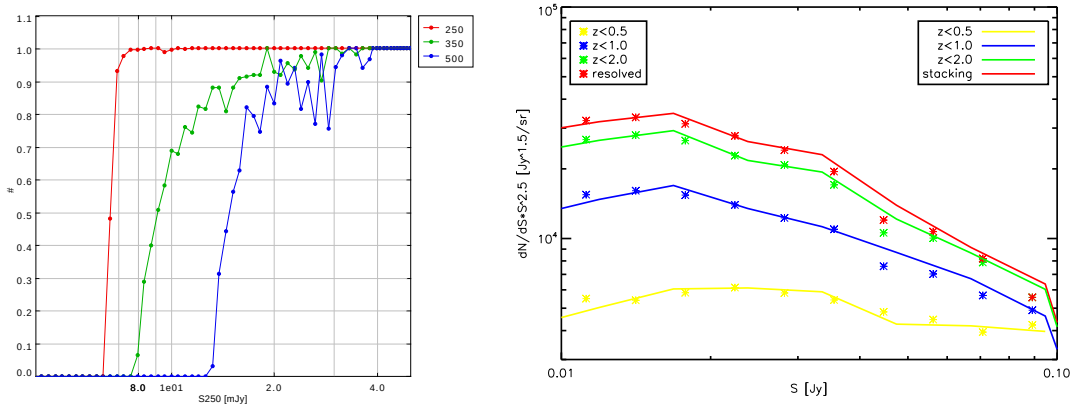

Figure 7.2: SPIRE Completeness of HerMES XID Catalogue in the COSMOS Field.]. Left: SPIRE SPIRE 250/350/500 μm completeness estimates of HerMES XID catalogue in the COSMOS field by Roseboom et al. (in prep) based on Roseboom et al. (2010) XID method. Right: HerMES XID catalogue measured 250 μm number counts as a function of redshift compared to 250 μm number counts estimates based on a stacking analysis by Béthermin et al. (2012). The two analysis confirm that our COSMOS SPIRE 250 μm sample is complete down to 10 mJy.

CH	P100	P160	S250	S350	S500	Area [deg^2]
COSMOS	3407 (36%)	3361 (35%)	9483 (100%)	6223 (65%)	1677 (18%)	1.7

Table 7.1: Number of sources used for the determination of the SPIRE Luminosity Function in the COSMOS field. The sample is selected at 250 micron and other numbers can thus be treated as detection rates.

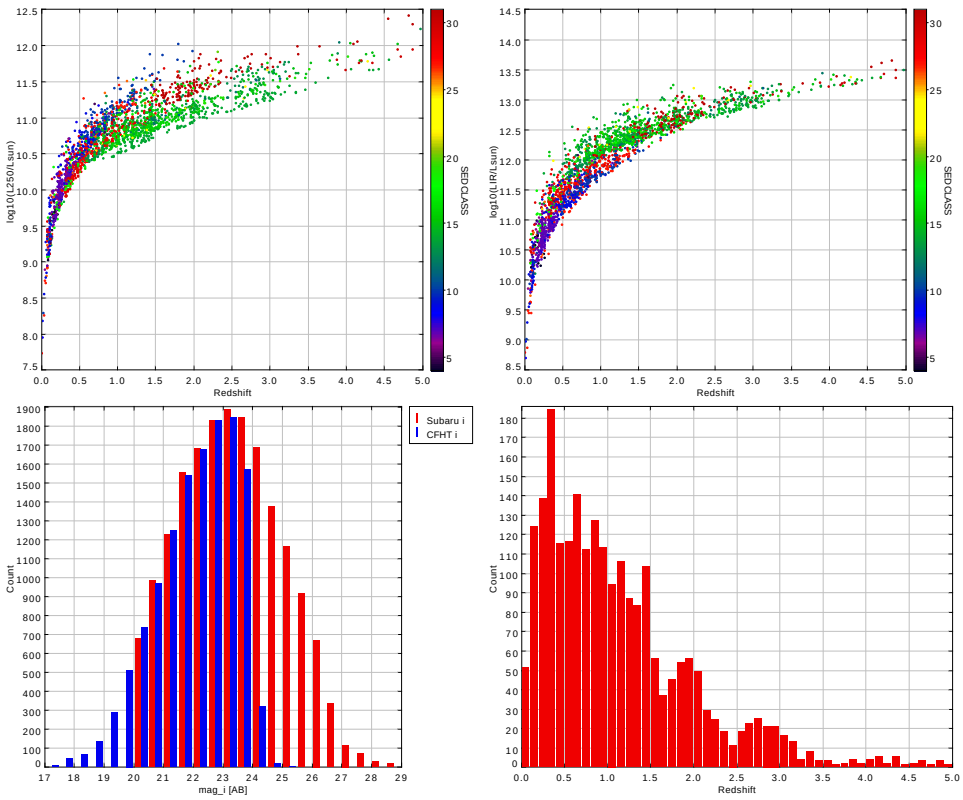

Figure 7.3: SPIRE 250 μm Sample General Properties in the COSMOS Field. Top Left: 250 μm Rest-Frame Luminosity as a function of redshift. Top Right: IR Bolometric Rest-Frame Luminosity as a function of redshift. Bottom Left: Observed i-band AB magnitude distribution based on shallower (Subaru) and deeper (CFHT) imaging. Bottom Right: Redshift Distribution based on COSMOS Photometric redshift Catalog.

7.2 $1/V_{max}$ Luminosity Function Estimates

Using the COSMOS Photometry Catalog and the COSMOS Photometric Redshift Catalog, we are able to perform the multi-wavelength SED fitting analysis of our SPIRE 250 μm sample and thus estimate the monochromatic and IR bolometric rest-frame luminosities and relative k-corrections. We assembled a 27-band photometric catalog reaching from the Optical to the PACS/SPIRE bands for each source and applied the two-fit `LePhare` SED fitting recipe described in Sec. 5.2.2. We divided the $L-z$ plane across the $0 < z < 5$ redshift range in intervals roughly according to sample sizes at different redshifts so as to sample the luminosity function in the $L-z$ in a statistically significant way at most luminosities and redshifts. Since our sample is effectively jointly selected at MIPS 24μm and SPIRE 250 μm and Tab. 7.1 shows

that sources are often not detected in the other PACS and SPIRE bands within our adopted limits, we tried using upper limits at PACS 100/160 μm and SPIRE 350/500 μm in the SED fitting, but concluded that this did not make much difference in the vast majority of cases while increasing the failure rate in the SED fitting, so we decided not to use upper limits in our analysis.

As done in Chap. 6, based on the best-fit SED at IR wavelengths determined by `LePhare` using the SWIRE templates by Polletta et al. (2007) and the slightly modified versions by Gruppioni et al. (2010) listed in Tab. 5.1, we determined monochromatic and IR bolometric luminosities for all sources along with their k-corrections. We then estimated the LFs at SPIRE 250 μm as well as at a number of shorter wavelengths such as 160, 90 and 60 μm and at IR bolometric (8-1000 μm) wavelengths based on the same SPIRE 250 μm sample and using the $1/V_{max}$ method. This wide variety of measurements ensures we can compare our results against a large number of previous estimates while breaking new ground in studying the LF of IR sources up to $z \sim 5$.

The z_{max} and thus V_{max} estimates were calculated combining the limits set by our two source detection criteria: $S_{24} > 80$ μJy and $S_{250} > 10$ mJy, following Eq. 6.2 and thus obtaining

$$z_{low<z<high,min} = z_{low} \tag{7.1}$$

$$z_{low<z<high,max} = \min[z_{f24,max}, z_{f250,max}, z_{high}] \tag{7.2}$$

$$\tag{7.3}$$

The comoving volume element is then integrated between the two redshift limits and over the observed sky area to obtain the V_{max} and weigh each source by a $1/V_{max}$ weight to compute the luminosity function estimates in each $L-z$ bin and at each rest-frame luminosity of interest.

While results for low-redshift sources are better discussed and illustrated in Chap. 6 and a detailed discussion of high-redshift results is deferred to later Sections of this Chapter, a comparison of our high-redshift results with previous work is illustrated in the following Fig. 7.4 shows the 250 μm LF compared against previous measurements at the same wavelength by Eales et al. (2010a) and Lapi et al. (2011) using HerMES and H-ATLAS data from Herschel SDP respectively. The agreement is rather good at all common redshifts, but our deep and medium-area sample allows us to reach up to higher redshifts while still having a reasonable statistics and redshift resolution, albeit losing the ability to probe the faint end of the luminosity function at the higher redshifts. Fig. 7.5 shows our results for the 160μm band, where observational results to compare against are currently not available. Fig. 7.6 and Fig. 7.7 compares our results with PEP SDP work at the same wavelengths by Gruppioni et al. (2010). While there seems to be a rather good agreement within errors throughout, we improve substantially on previous constraints based on very small areas, and particularly so since with increasing redshift our SPIRE 250 μm selection and our 350/500 μm flux measurements allow us to probe these rest-frame wavelengths more reliably than through a PACS selection. Finally, Fig. 7.8 shows the IR bolometric LF compared against a number of previous measurements, where our combination of depth and area allows us to break new ground at high redshift.

7.2.1 SPIRE 250 μm Luminosity Function

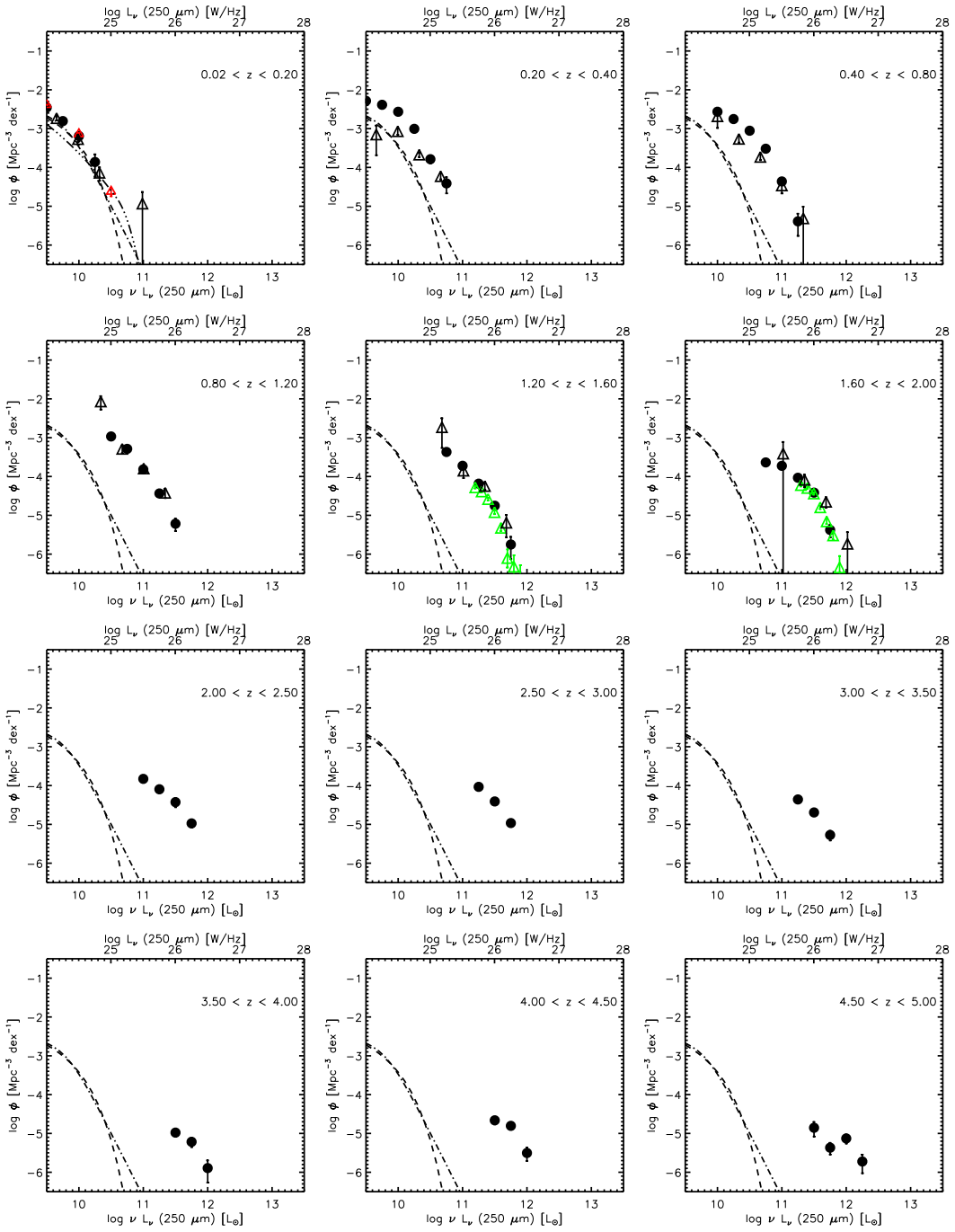

Figure 7.4: The SPIRE 250 μm bolometric luminosity functions using the $1/V_{max}$ technique compared to Eales et al. (2010b). Black filled circles: this work; black empty triangles: HerMES SDP 250 μm LF by Eales et al. (2010b); green empty triangles: H-ATLAS SDP 250 μm LF by Lapi et al. (2011); black empty triangles: HerMES SDP 250 μm LLF by Vaccari et al. (2010).

7.2.2 SPIRE 160/90/60 μm Luminosity Function

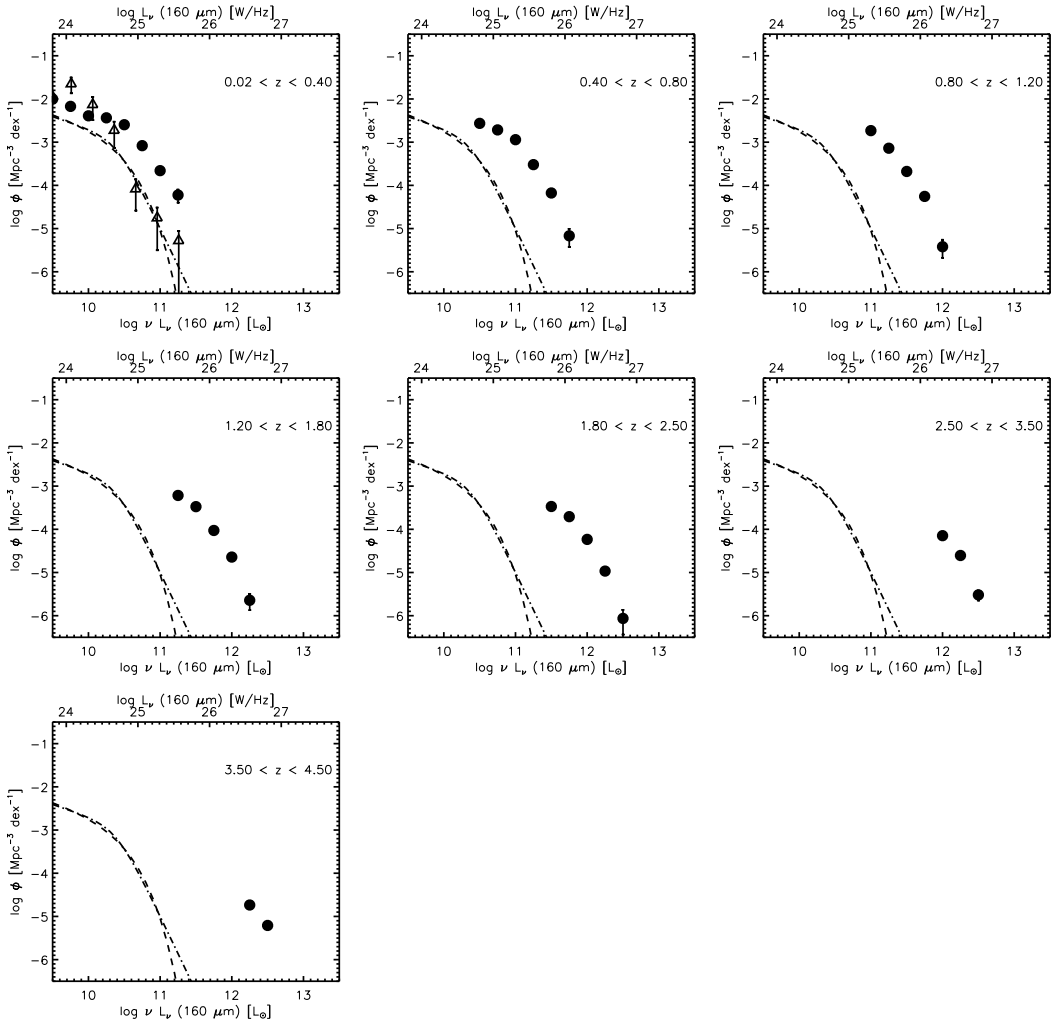

Figure 7.5: The SPIRE 160 μm luminosity functions using the $1/V_{max}$ technique. Black filled circles: this work; black empty triangles: FIRBACK 170 μm LLF by Takeuchi et al. (2006) converted to PACS effective wavelengths using a spectral index of -1.0.

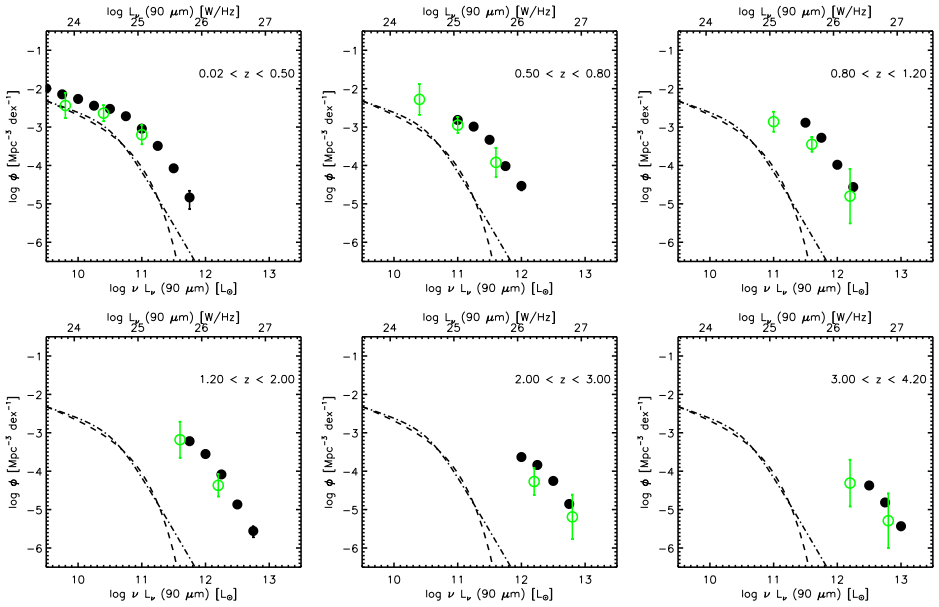

Figure 7.6: The SPIRE 90 μm luminosity functions using the $1/V_{max}$ technique. Black filled circles: this work; green empty circles: PACS SDP 90 μm LF by Gruppioni et al. (2010); black dot-dashed and dashed lines are local luminosity function prediction at 350 μm by Serjeant & Harrison (2005) estimated following Eq. 2.64 and Eq. 2.67 respectively.

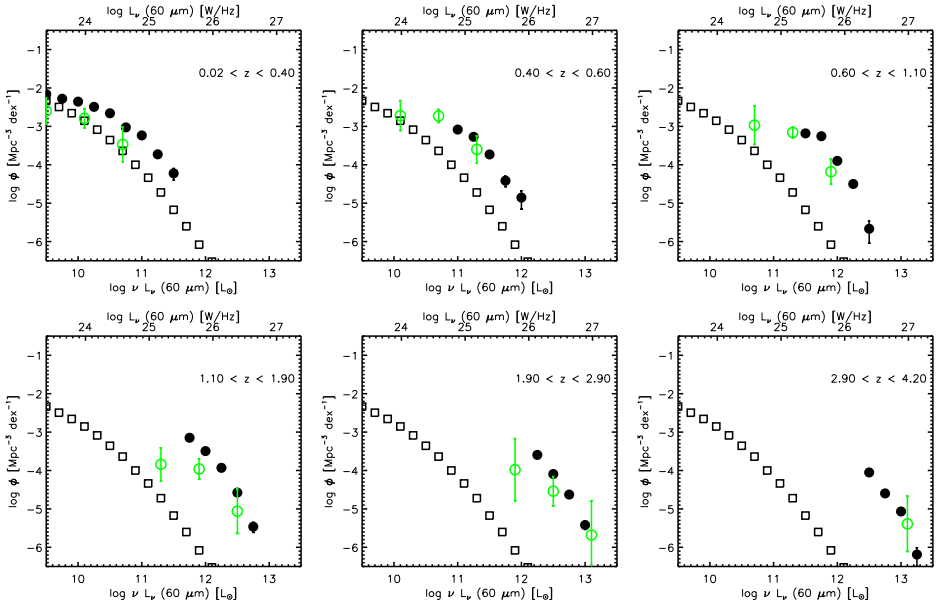

Figure 7.7: The SPIRE 60 μm luminosity functions using the $1/V_{max}$ technique. Black filled circles: this work; green empty circles: PACS SDP 60 μm LF by Gruppioni et al. (2010); black empty squares: 60 μm local luminosity function estimate by Saunders et al. (1990).

7.2.3 SPIRE IR Bolometric Luminosity Function

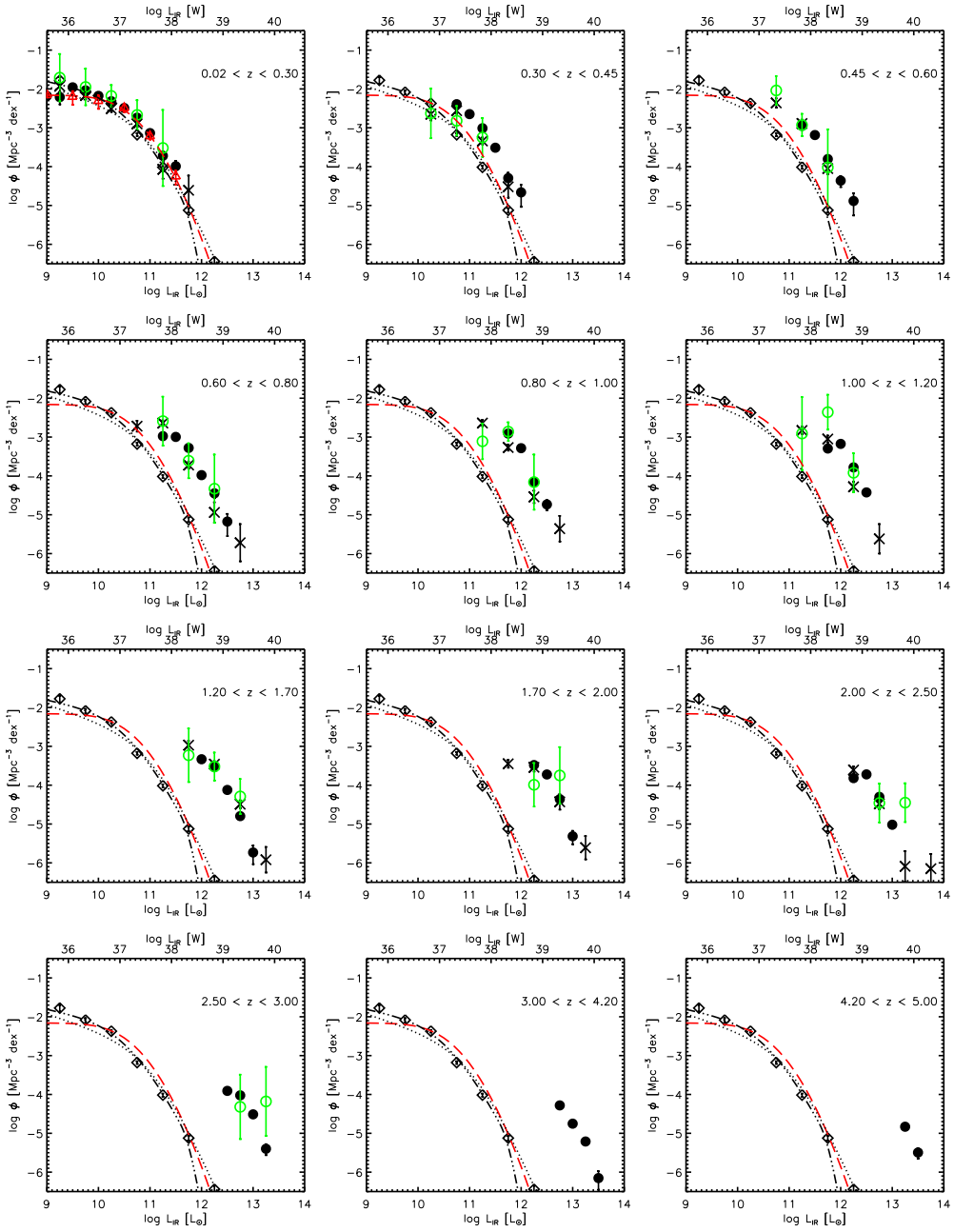

Figure 7.8: The SPIRE IR bolometric luminosity functions using the $1/V_{max}$ technique. Black filled circles: this work; black crosses; MIPS 24 μm selected IR Bolometric LF by Rodighiero et al. (2010); green empty circles: PACS 100/160 μm selected PEP SDP 60 μm LF by Gruppioni et al. (2010); red dashed line: IR Bolometric LLF best fit by Vaccari et al. (2010); Black dot-dot-dot-dashed line: IR Bolometric LLF model by Negrello et al. (2007); Black dotted line: IR Bolometric LLF model by Valiante et al. (2009) LLF model.

7.3 Results from Alternative Luminosity Function Estimators

The classical $1/V_{max}$ and Maximum Likelihood tests used for estimating LFs in the previous Sections are certainly relatively simple and offer complementary advantages, as previously discussed. However for completeness we have considered essential to test in addition these results and their robustness with independent methods.

7.3.1 Results From $1/V_{est}$ & A Modified Page & Carrera's Method

Page & Carrera (2000) proposed an extension of the classical $1/V_{max}$ LF estimator. This method, which the authors refer to as $1/V_{est}$ but is also known as ϕ_{est}, more carefully estimates the volume available for every sample source within the redshift-luminosity bin. As discussed in Chap. 2, the difference with $1/V_{max}$ is expected to become significant for redshift-luminosity bins affected by the survey flux-density limit (that typically happens at low-luminosities).

We report in Fig. 7.9 a comparison of the results of the $1/V_{max}$ (in grey symbols) and $1/V_{est}$ (red symbols) tests for the redshift-dependent LFs at $\lambda_{eff} = 250$ μm, based on the COSMOS sample, in 9 redshift bins from $z = 0.25$ to $z = 4.25$. As we see here, both estimators appear to suffer some incompleteness effects at low luminosities in all sampled redshift bins, manifested as sharp down-turns of the LFs at low luminosities. However, the Page & Carrera's method is less prone to the effect, thanks to a slightly better estimate of the available volume to the source. In any case, in all luminosity bins except the lowest ones, the agreement between the two methods is perfect.

Motivated by this slightly better performance of the $1/V_{est}$ estimator, we have proceeded to a modification of the original Page & Carrera (2000) test, allowing us to use it for estimating LFs in wavebands different from that of the primary selection (in our case 250 μm). With reference to Fig. 2.2, to every source i in any redshift-luminosity intervals we associate an available volume as appropriately given by $V_{est,i}$. We then re-distribute sources in bins of luminosity at the new wavelength and finally add up their weight as

$$\phi_{1/V_{est}}(L, z) = \frac{1}{\Delta L} \sum_{i=1}^{N} \frac{1}{V_{est,i}}, \tag{7.4}$$

where N is the number of sources falling in that bin. Errors on this are computed in the same way as for the $1/V_{max}$ method and thus as in Eq. 2.43.

7.3.2 Results from Schafer's *Semi-Parametric* Estimator

The *semi-parametric* method proposed by Schafer (2007) and already successfully applied by us in Chap. 6 in the context of the low-redshift wide-area sample analysis is applied and compared to $1/V_{max}$ results in Fig. 7.9. At high redshifts, however, the *semi-parametric* estimate appears to fail to provide a credible representation of the multi-wavelength luminosity-redshift planes and LFs. As we can see in the figure, the estimator has difficulties in reproducing sharp features in the LFs like turn-overs in luminosity and rapid changes as a function of redshift, as apparently indicated by the data. If we would assume that the estimated 250 μm LFs by the method provide a correct description of the real LFs, we would enormously exceed the observed number counts Oliver et al. (2010); Glenn et al. (2010) and the CIRB background measured intensity Lagache et al. (2005); Dole et al. (2006b).

One of the reasons of such difficulties of the *semi-parametric* method is illustrated in Fig. 7.10 reporting the distribution of sources from the COSMOS field as a function of the 250 μm rest-frame luminosity and redshift. The red dashed line contour marks the flux completeness, while the solid line contours indicate the reconstructed 2D distribution of the source density field according to the Schafer method. The top panel is based on the assumption that luminosities are computed using a single average far-IR SED, while in the bottom panel luminosities are computed with their own source-by-source k-correction based on the SED fitting as in Chap. 5. As we see, the two distributions are significantly different, with a spread obviously emerging in the latter case. Now, the Schafer's code can apparently only deal with the former distribution, but does not converge at all if we give it to digest the more correct latter distribution.

Another potential problem with this *semi-parametric* method is that it appears to strongly constrain the LFs well outside regions of the $L - z$ plane where we do really have data, an extrapolation that appears to us as largely unjustified.

Altogether, the *semi-parametric* method appears to work well within limited regions of the $L - z$ plane. In particular, it gives answers in excellent agreement with those of other estimators for the shallow low-redshift samples discussed in Chap. 6 and Fig. 6.13 for which the range covered in redshift is limited and where the spread induced by different k-corrections is small (see Fig. 6.12), because the k-correction itself is small. However the test becomes unreliable for more complex situations like those covered in the present Chapter.

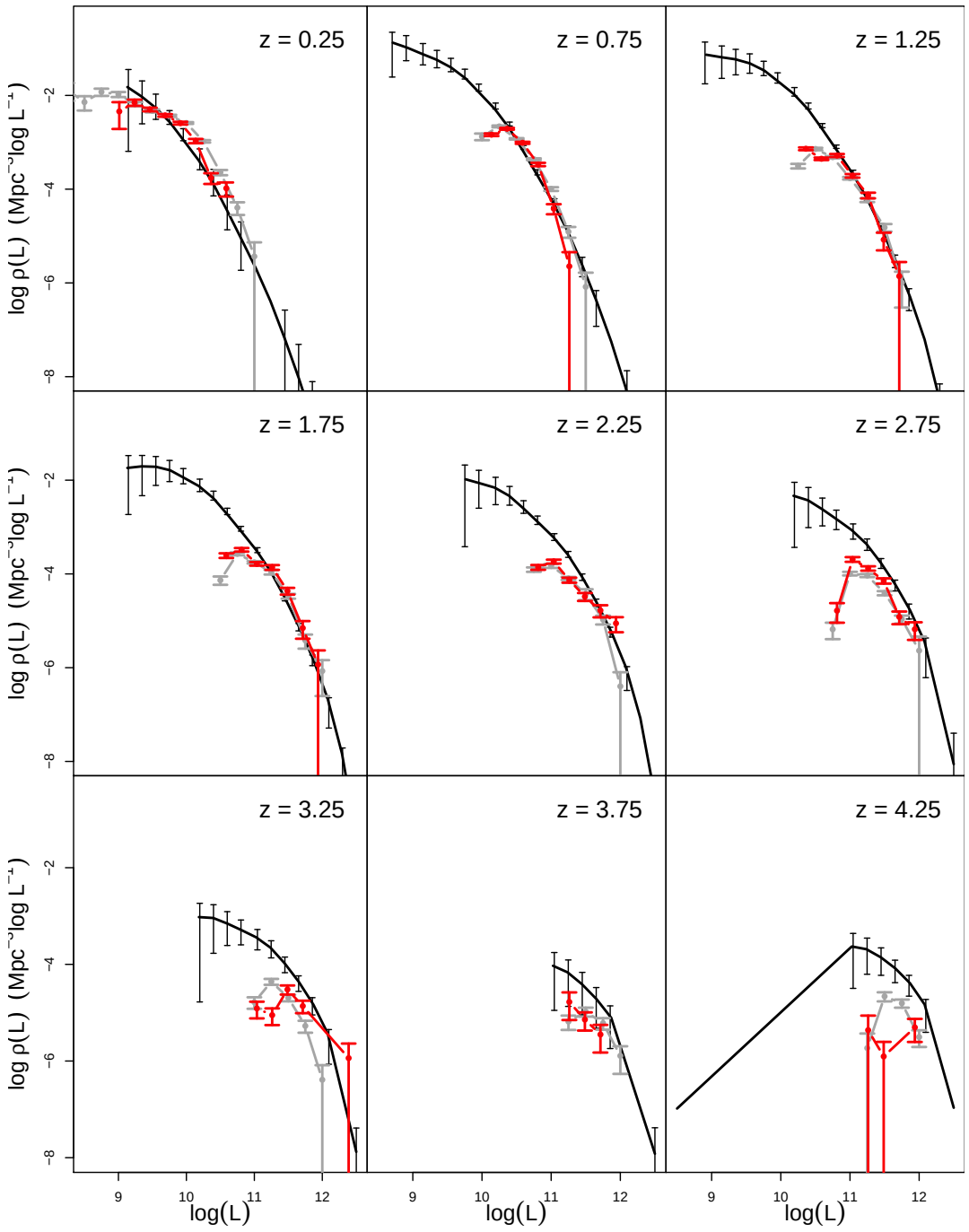

Figure 7.9: The SPIRE 250 μm luminosity function estimate using the Page & Carrera (2000) method (in red), the $1/V_{max}$ technique (in grey) and the Schafer (2007) method (in black).

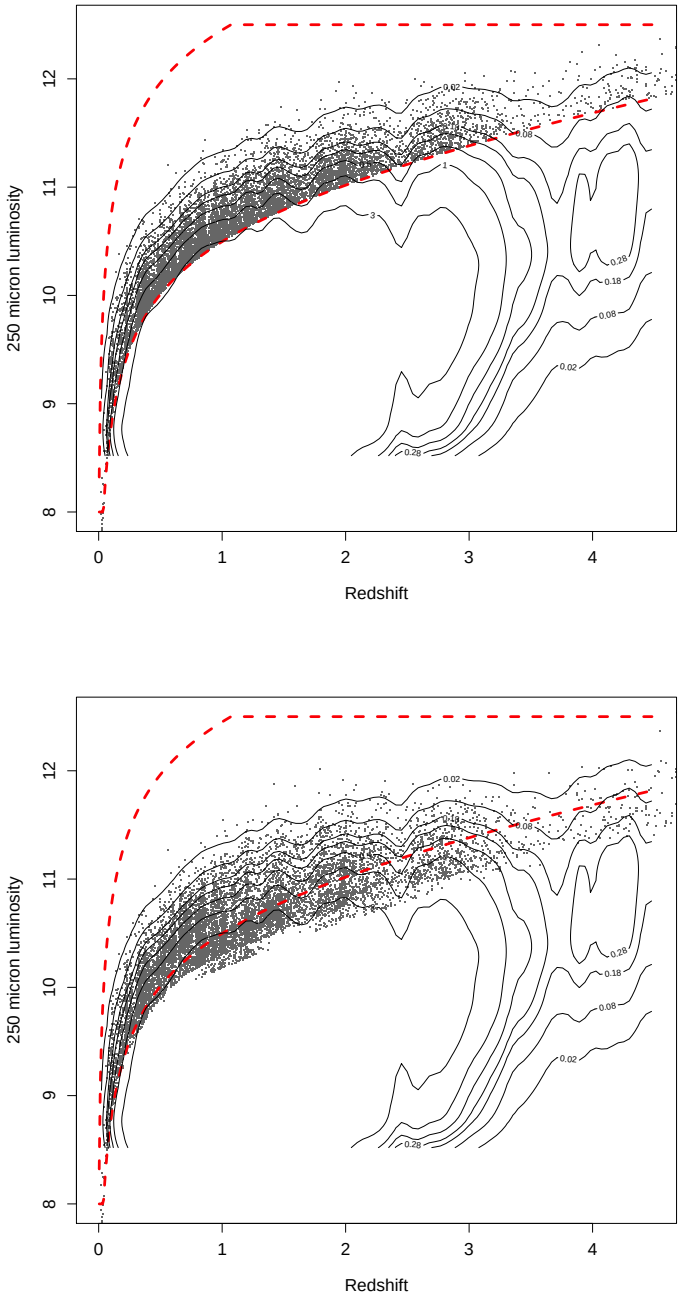

Figure 7.10: Distributions of SPIRE sources from the COSMO sample in the 250 μm luminosity *vs* redshift plane as reconstructed by Schafer (2007) estimator. Grey points: the data; the red dashed lines mark the flux limitations adopted in the application of the *semi-parametric* LF estimator by Schafer (2007); the continuous black lines are iso-density contours corresponding to the *semi-parametric* reconstrunctions of the source volume density as a function of luminosity and redshift. **Top**: the luminosities are calculated using an average IR SED for all sources; **Bottom**: the luminosities are calculated by using the k-correction based on the SED fitting procedure discussed in Chap. 5. Using the correct source-by-source best-fit SED as in the bottom panel, the Schafer (2007) code does not converge (contours here are those from the top-panel solution).

7.4 Evolution of Herschel Sources

The major motivation for this Thesis is not just the mere characterization of the statistical properties of the FIR and SMM source populations, but rather an exploration of how such properties evolve with cosmic time. To the depth of our present Herschel SPIRE and PACS surveys of the COSMOS field, this process of retrieving flash-backs in time is relatively straightforward, with some limitations. The latter are apparent in our Fig. 6.12 and Fig. 7.10, and in particular in the diagonal coverage of the data in the luminosity-redshift plane, the well known Malmquist effect of using flux-limited data. This implies that we can largely compare LFs in different ranges of luminosity at different epochs. Herschel data are of clear help here, however, in the moderate dependence of the luminosity lower boundary in Fig. 7.10 at increasing redshift, for $z > 1$, a modest increase because of the favourable k-corrections for high-z sources when observing in the SMM.

In conclusion, at least when considering the moderate-to-high luminosity regimes, our sampling of the LFs is guaranteed at least to $z \sim 4$ and cosmological evolution directly inferred from the observed LFs with minimal modeling.

7.4.1 Cosmological Evolution from V/V_{max} Analysis

A first classical method to check for evolution of a population of cosmic sources under investigation is to look for deviations from homogeneity in the sampled volumes as carried out with the V/V_{max} test.

We have applied this test on the shallow HerMES Wide fields sample and on the COSMOS field, separately, and report the V/V_{max} statistics as a function of source rest-frame 250 μm luminosity in Fig. 7.11 and Fig. 7.12. We see a coherent behaviour of the average $\langle V/V_{max} \rangle$ in luminosity bins, with a steady increase at increasing luminosity: while sources at $logL/L_{\odot} < 9.5$ are largely consistent with homogeneous distributions in space-time, an increasing fraction of the higher-luminosity ones are found at the boundaries of the sampled volume, which means they are numerous back in cosmic time but almost absent locally. The effect is obviously much more prominent in the COSMOS field than in the Wide fields one, due to the COSMOS much fainter flux limit. In both cases there is clear evidence for a strong luminosity dependence in the evolution of Herschel galaxies.

By its own nature, the V/V_{max} test provides us with quite an integrated measure of evolution, which is highly degenerate when we attempt to break it down in terms of changes of the average luminosity, or the average source comoving number density, or more radical deformations of the LF shapes. The V/V_{max} test, according to its description and analysis by Avni & Bahcall (1980), has very limited power to effectively constrain evolution patterns, limited in practice to testing density evolution, e.g. in bins of luminosity. For this reason, we have preferred to try mapping the Herschel sources as a function of cosmic epoch by looking directly at the evolution of their epoch-dependent luminosity functions, as discussed in the following.

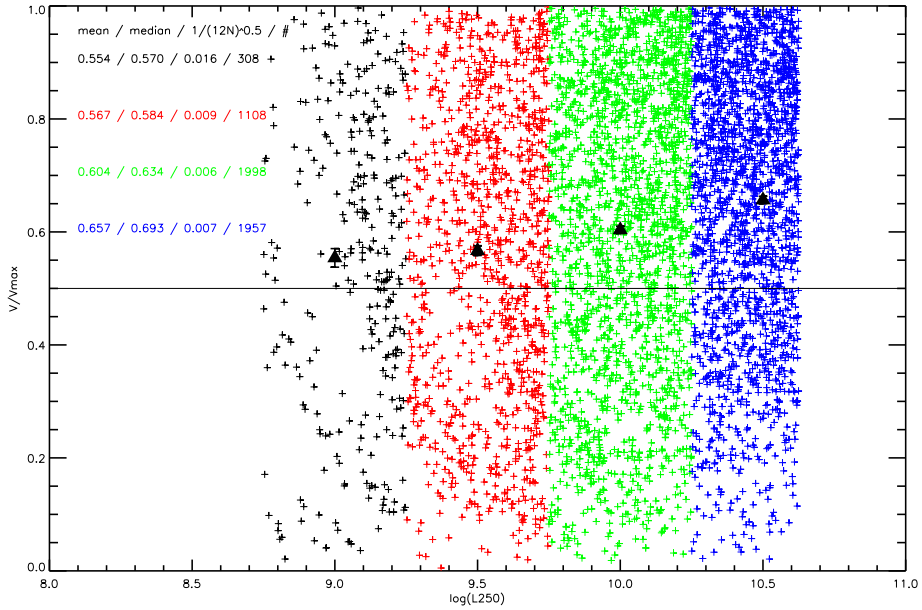

Figure 7.11: V/V_{max} for SPIRE Sources in the HerMES Wide Fields

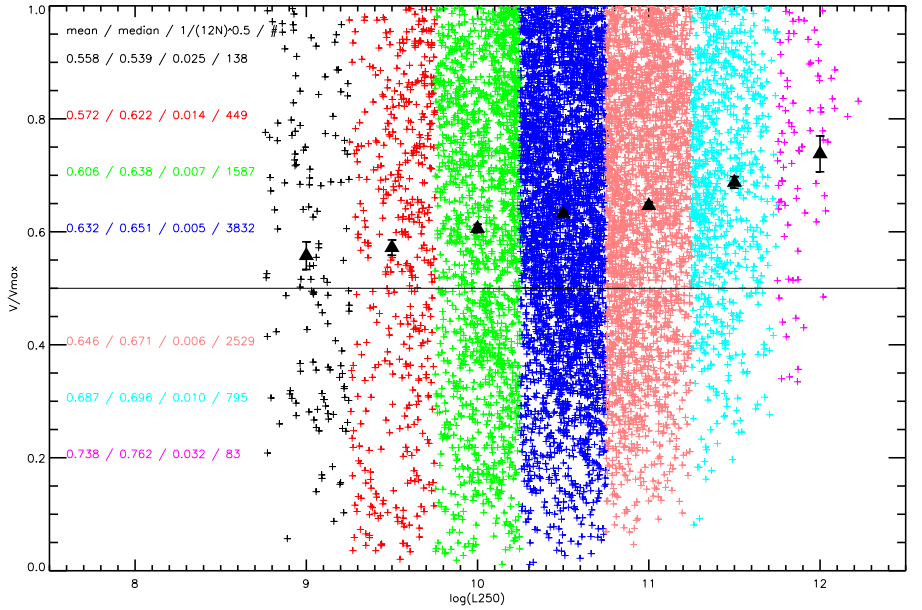

Figure 7.12: V/V_{max} for SPIRE Sources in the COSMOS Field

7.4.2 Evolution of Multi-Wavelength Luminosity Functions

Starting from the considerations made in the previous Section, to actually map the evolution of Herschel sources we need to directly investigate the evolution of their luminosity functions. To proceed with this intent we developed two separate approaches both based on the idea to give a parametric estimate of the evolution of the luminosity function with redshift.

The first approach adopts a parametric functional form whose best fitting values and uncertainties have been found by a non linear least square fitting procedure applied separately to each redshift bin $1/V_{max}$ estimates for the IR bolometric luminosity functions (shown in Fig. 7.8). This method presents some possible intrinsic biases, as detailed in the following, thus we develop a second approach that estimates the best fitting parameter by using all the redshift-binned $1/V_{est}$ estimates (shown in Fig. 7.9) together. Besides, this approach allows us to consider together the results coming from the HerMES Wide Fields sample in the local Universe, described in Chap. 6, and those at higher redshift by using COSMOS data. This method is more similar to adopt a model to describe the whole evolutionary trend of the luminosity functions and has been applied to both the SPIRE 250 μm and IR bolometric luminosity functions. The choice of using $1/V_{max}$ and $1/V_{est}$ respectively has no really incidence in the results we obtain, since the two methods give very similar results in those high luminosity bins of the LFs where the majority of sources are found and thus effectively guide the fit in both cases (this is shown in Fig. 7.9). This choice can then be thought as a cross-check of the procedures implemented along the work. In the text, in order to distinguish the two approaches, we refer to the first LF fitting procedure as the the *single-z* fitting procedure and to the second as the *global* procedure.

Single-z Fitting Procedure Results

Fig. 7.13 illustrates the *single-z* fitting procedure results. In this plot the evolution of the IR Bolometric LF is shown as a function of redshift estimated using the classical $1/V_{max}$ estimator in the COSMOS field, overplotting its measurements at different redshifts along with the least-square best fits performed using a log-gaussian best-fit shape partly inspired by the IR Bolometric LLF determined in Chap. 6. Namely, to perform this fit we assume that while in the first z-bin all the parameters have been estimated, starting from the second z-bin, the values of α and σ have been frozen at the values found at lower redshift, leaving only L_* and Φ_* free to vary.

While this method is particularly straightforward to apply, it presents some limitations when the data used presents some incompleteness with redshift. As we already pointed out, due to the small sky area sampled by COSMOS, the $1/V_{max}$ estimator (and thus also the $1/V_{est}$ estimates, which are effectively a slightly improved version of $1/V_{max}$ estimates and thus have almost the same level of uncertainty), have problems in really describing the low luminosity trend of the luminosity functions. Consequently, proceeding with a least square fit separately in each redshift bin could give us some over-extrapolation of the predicted shape of the LF. To better constrain these low luminosity bins, we would better consider together the HerMES Wide Fields and the COSMOS samples and use a more sophisticated approach to fit all data together. For these reasons we decided to perform the already introduced *global* fit procedure explained below. In Fig. 7.17 we illustrate the evolution of L_* and ϕ_* as a function of redshift obtained with the *single-z* fitting procedure as a comparison to those we obtain

from the following *global* fit procedure.

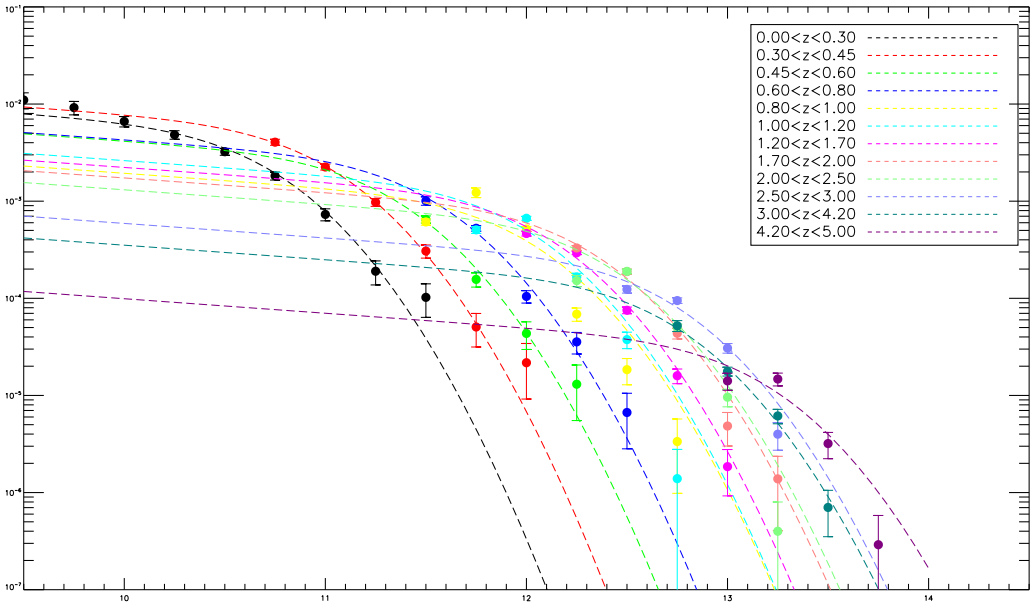

Figure 7.13: IR Bolometric Luminosity Function Estimates obtained from $1/V_{max}$ analysis fitted by the *single-z* fitting procedure described in the text.

Global Fitting Procedure Results

As a guideline to compare LFs from the two (Wide Fields and COSMOS) selections and to constrain in a simple parametrized form the evolution patterns emerging from our analyses, we have fitted the whole combined data points in Figs. 7.14 and 7.15, for the 250 mum and the IR Bolometric LF respectively, with simple models. The two datasets turn out to be nicely complementary in terms of their coverage of the luminosity-redshift plane (see Fig. 6.12 for the Wide Fields and Fig. 7.10 for the COSMOS sample), the Wide Fields needed to measure the low-z LFs where COSMOS does not have enough sampled volume, the latter to achieve the depth needed to get good source statistics at high redshifts. Direct overlap in redshift space is achieved at about $0.2 < z < 0.4$ where the two samples provide comparable statistics, though in slightly different luminosity intervals. Because of potential incompleteness in the Wide Fields at redshifts higher than $z \sim 0.6$ (see Chap. 6 and Fig. 6.13, last panel), we do not consider these data beyond this redshift limit in the subsequent analysis.

We adopted a representation of the local LF form in terms of the Log-Gaussian function (as described in Chap. 6) defined in Sec. 2.3.3, Eq. 2.62, and allowed it to evolve in the simplest possible way. We also adopted the alternative double power-law shape (Eq. 2.66), obtaining essentially undistinguishable results. Since there appears to be very marginal evidence in Fig. 7.13 for an evolution in shape of the LFs as a function of redshift, we have assumed here too that the LF shape parameters keep constant in time, while evolving in both luminosity (PLE) and comoving number density (PDE) according to simple power-law dependences on $(1+z)$ as defined in Eq. 2.83 and Eq. 2.82. In practical terms, this means that the L_* and ϕ_*

parameters change as a function of redshift as:

$$L_* = L_{*,0}\ (1+z)^{\alpha_L}, \qquad \phi_* = \phi_{*,0}\ (1+z)^{\alpha_D} \tag{7.5}$$

with the α_D parameter value allowed to change at $z = z_D$ and α_L to change at two redshifts, $z = z_{L1}$ and $z = z_{L2}$.

We report in Table 7.2 the values of the parameters providing best fits to the rest-frame 250 μm and to the IR bolometric LFs in Figs. 7.14 and 7.15, respectively. Best-fit values and confidence limits are obtained with the use of a χ^2 test. Comparisons of the fitting functions with observational data are reported in the same figures as green solid lines.

250 μm rest-frame Luminosity Functions

Redshift range	$\log(L_*)$ $[L_\odot]$	α	σ	ϕ_* $[\mathrm{Mpc}^{-3}\mathrm{dex}^{-1}]$	α_D	α_L
$0 < z < 0.1$	9.3	1.1	0.32	-2.02	-	-
$0.1 < z < z_D(z_D = 0.72)$	$9.3(1+\langle z\rangle_i)^{\alpha_L}$	1.1[1]	0.32[2]	$-2.02(1+\langle z\rangle_i)^{\alpha_D}$	-1. 57	3.90
$z_D < z < z_{L1}(z_{L1} = 0.74)$	$9.3(1+\langle z\rangle_i)^{\alpha_L}$	1.1[3]	0.32[4]	$-2.02(1+\langle z\rangle_i)^{\alpha_D}$	-5.24	3.90
$z_{L1} < z < z_{L2}(z_{L2} = 2.72)$	$9.3(1+\langle z\rangle_i)^{\alpha_L}$	1.1[5]	0.32[6]	$-2.02(1+\langle z\rangle_i)^{\alpha_D}$	-5.24	3.84
$z_{L2} < z < 4.0$	$9.3(1+\langle z\rangle_i)^{\alpha_L}$	1.1[7]	0.32[8]	$-2.02(1+\langle z\rangle_i)^{\alpha_D}$	-5.24	-2.11
IR Bolometric Luminosity Functions						
$0 < z < 0.1$	10.1	1.17	0.34	-2.32	-	-
$0.1 < z < z_{L1}(z_{L1} = 0.20)$	$10.1(1+\langle z\rangle_i)^{\alpha_L}$	1.17[9]	0.34[10]	$-2.32(1+\langle z\rangle_i)^{\alpha_D}$	-1. 49	6.66
$z_{L1} < z < z_D(z_D = 1.17)$	$10.1(1+\langle z\rangle_i)^{\alpha_L}$	1.17[11]	0.34[12]	$-2.32(1+\langle z\rangle_i)^{\alpha_D}$	-1.49	4.17
$z_D < z < z_{L2}(z_{L2} = 1.95)$	$10.1(1+\langle z\rangle_i)^{\alpha_L}$	1.17[13]	0.34[14]	$-2.32(1+\langle z\rangle_i)^{\alpha_D}$	-4.3	4.17
$z_{L2} < z < 4.0$	$10.1(1+\langle z\rangle_i)^{\alpha_L}$	1.17[15]	0.34[16]	$-2.32(1+\langle z\rangle_i)^{\alpha_D}$	-4.3	1.63

Table 7.2: Best-Fit parameters of the LF evolution obtained by the *global* fitting procedure described in the text. With $\langle z\rangle_i$ we intend the mean of the redshift bins plotted in Fig. 7.14 and 7.15.

Statistical consistency of the Wide Fields and COSMOS Luminosity Function Estimates

A first consideration emerging from the inspection of these fits is that there appears to be a good consistency between LFs derived from the Wide Fields and from COSMOS at the low-redshift end. This is quite reassuring about the statistical validity of our analyses.

The LF's Low-Luminosity Behaviour

A second important consideration concerns the behaviour of the LFs at the low-luminosity end in the various redshift bins. This part of the LFs is very well sampled at $z = 0.05$ by the Wide Fields and by COSMOS at z=0.25 (particularly for the 250 μm function). Otherwise, incompleteness effects related to the survey flux limits prevent us a direct estimate of them. So a question emerges about how well our formal best-fit solutions based on the assumption of a constant LF shape may describe the statistical properties of the Herschel galaxy populations at such low luminosity values. Or if alternative solutions, like those suggested by the *semi-parametric* test in Fig. 7.9 with a large excess in the number densities at the low-luminosity end, might be more appropriate.

A decisive answer to these questions comes from calculations we have performed by assuming evolutionary LFs at 250 μm as in Fig. 7.14 and Tab. 7.2 and a representative average IR spectral shape (for the latter we used SED # 31 in Tab. 5.1 which has got a k-correction largely representative of our Herschel sample at large). With these ingredients, we computed the source number counts at 250 μm and compared them to published data in Fig. 7.16. In spite of our rough assumption of a single average SED when computing k-corrections, this simple model offers a decent fit to the number count data and a value of the contribution by these sources to the CIRB background ($\nu I_\nu \simeq 10\ nW/m^2/sr$) entirely consistent with the DIRBE measurements (Hauser et al. (1998) ; Dole et al. (2006b)). The latter is an essential and vital constraint to any evolutionary model of IR galaxy populations.

The opposite would happen if we assumed different forms of the LFs like those in Fig. 7.9, that would dramatically overpredict the counts at the faint-flux end the overall background intensity by a large factor.

The bottom line is that the heuristic model reproduced in Figs. 7.14 and 7.15 is consistent with all available information and best-fits our new LF data. Any significant deviations at the low-luminosities with respect to our current estimate would make it inconsistent with data.

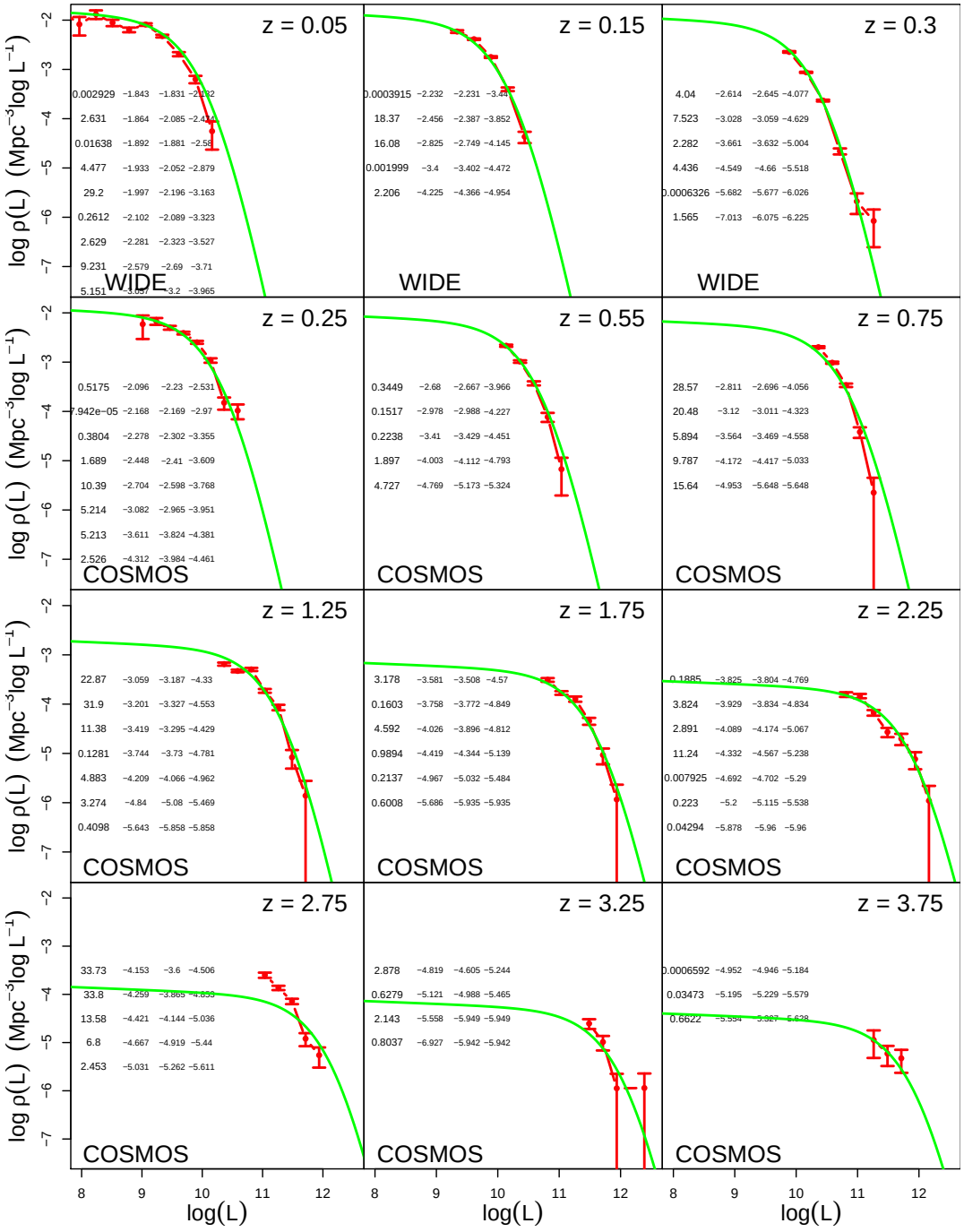

Figure 7.14: The COSMOS SPIRE 250 μm luminosity function estimate fitted by a modified Schechter function using the *global* fitting procedure explained in the text. Red points: $1/V_{est}$ estimates; green line: result of the fitting procedure.

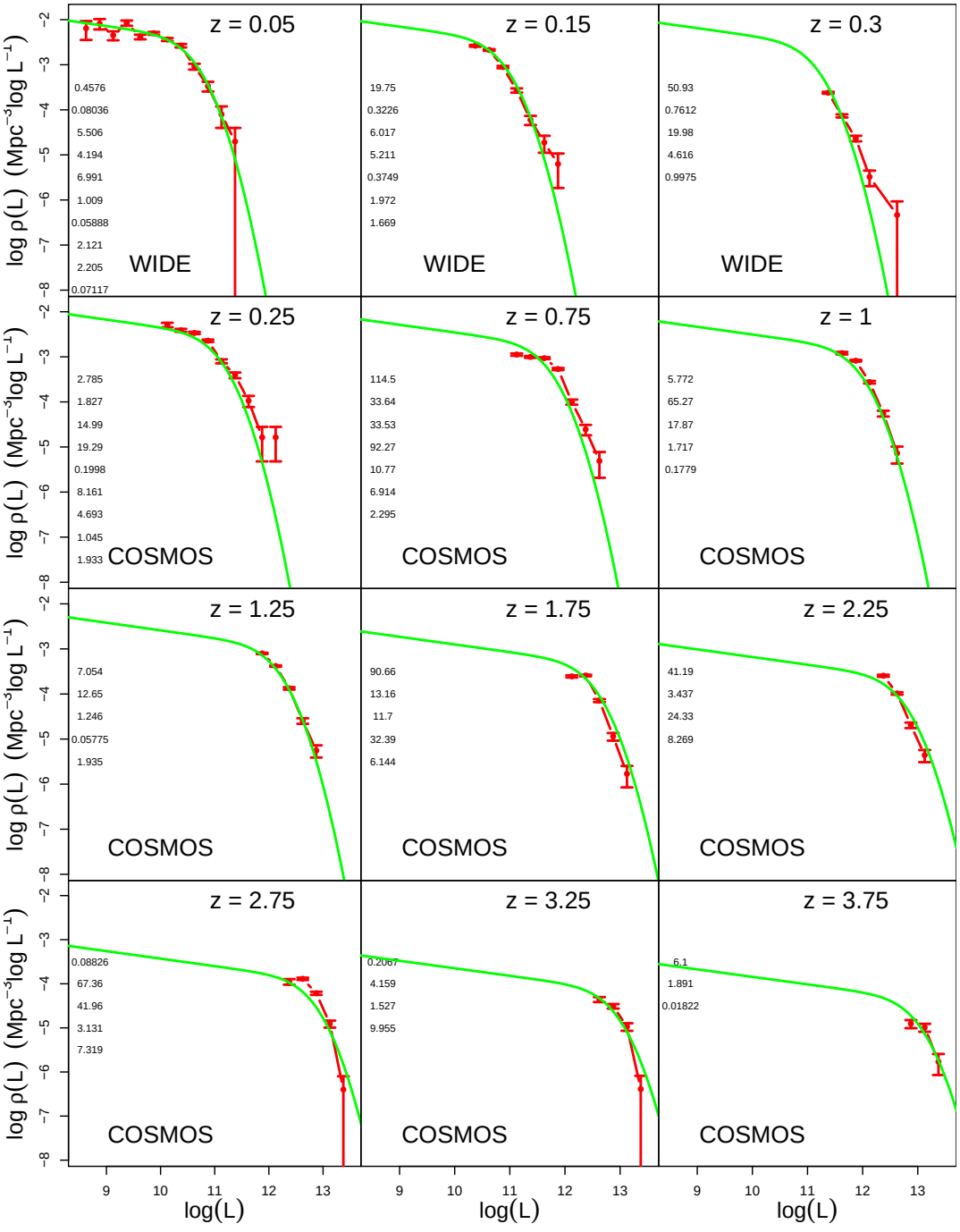

Figure 7.15: The COSMOS IR bolometric luminosity function estimate fitted by a modified Schechter function using the *global* fitting procedure explained in the text. Red points: $1/V_{est}$ estimates; green line: result of the fitting procedure.

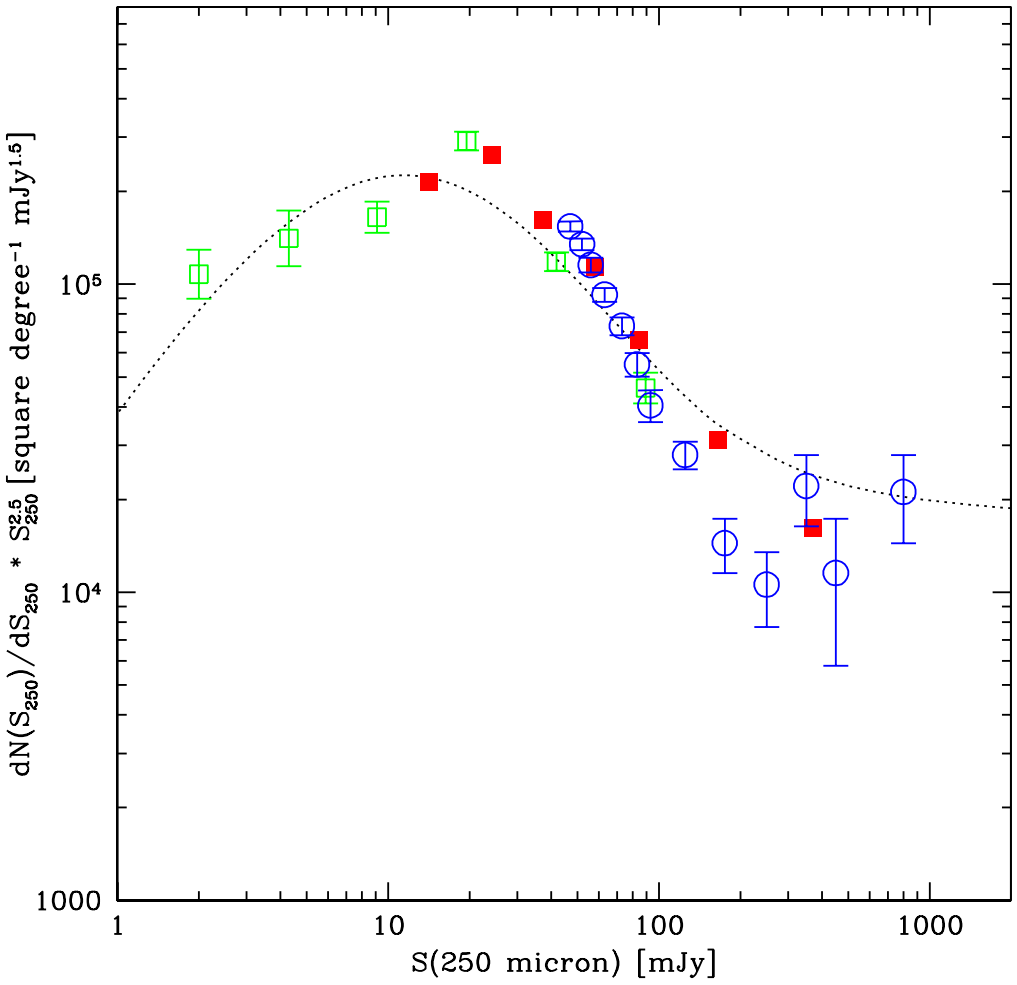

Figure 7.16: The COSMOS SPIRE 250 μm number counts as a function of redshifts. Black dotted result is our estimate assuming a mean SED to describe the whole sample of galaxies; in red the result by Oliver et al. (2010) and in green the result by Glenn et al. (2010) both obtained using the SPIRE 250 μm data observed during the SPIRE Science Demonstration Phase within HerMES; in blue Clements et al. (2010) obtained from H-ATLAS observations of the first $\sim$ 14 deg^2 of the entire H-ATLAS survey.

Evolutionary patterns of Herschel sources

A representative summary of the evolutionary properties of Herschel sources can be inferred from Fig. 7.17 and 7.18, reporting two independent assessments of how the two basic parameters of the IR bolometric luminosity functions, the characteristic luminosity L_* and the (comoving) number density normalization parameter ϕ_*, evolve with time.

Fig. 7.17 is based on the *single-z* fits to individual LFs in redshift bins reported in Fig.7.13, while Fig. 7.18 refers to the *global* solution to the bolometric LFs in Fig. 7.15. While the details from the two analyses show some differences (the *single-z* fiitting solutions have a change in the luminosity evolution rate at $z = 0.9$, while the *global* fit sees a similar change at z=1.95), due to the different approach already explained, the general pattern is extremely well and coherently defined.

The evidence is for a strong negative evolution in the source comoving number density traced by the ϕ_* parameter, paralleled by an even stronger increase in the characteristic luminosity L_*. The corresponding variations are of a couple of orders of magnitude decrease in number and two and a half orders increase in luminosity from local to $z = 4$. The logarithm of number density is inversely proportional to redshift, $log\phi_* \sim -0.5\ z$, while the average luminosity shows a fast increase up to $z \sim 2$ and a subsequent flattening.

Some consequences of these evolutionary patterns on physical models of galaxy formation and evolution will be discussed in Sec. 7.5.

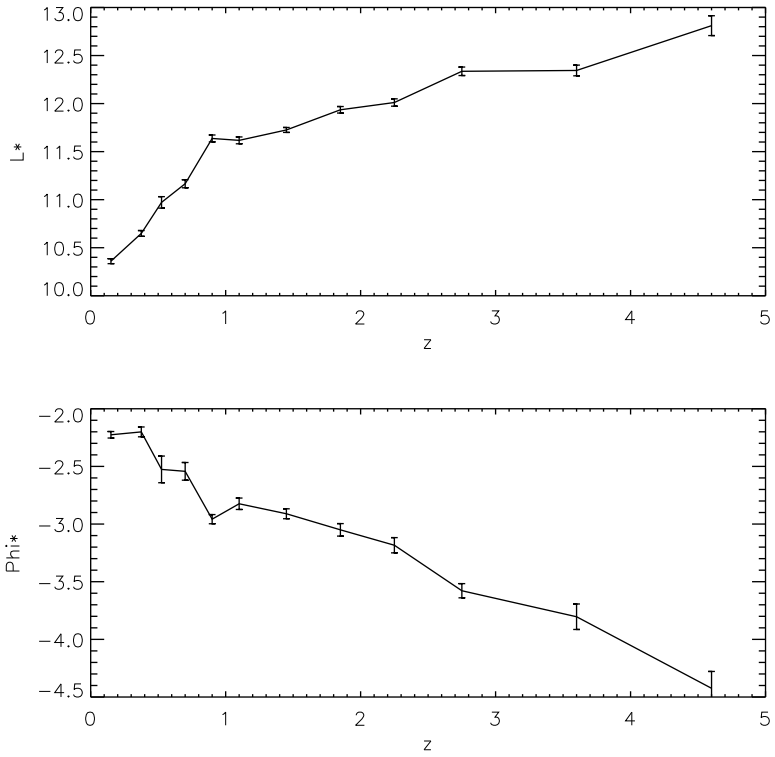

Figure 7.17: Evolution of IR Bolometric Luminosity Function Parameters obtained from $1/V_{max}$ analysis fitted by using the *single-z* fitting procedure explained in the text. **top**: the luminosity evolution as $L_*(1+z)^{\alpha_L}$. **Bottom**: the density evolution as $\phi_*(1+z)^{\alpha_D}$.

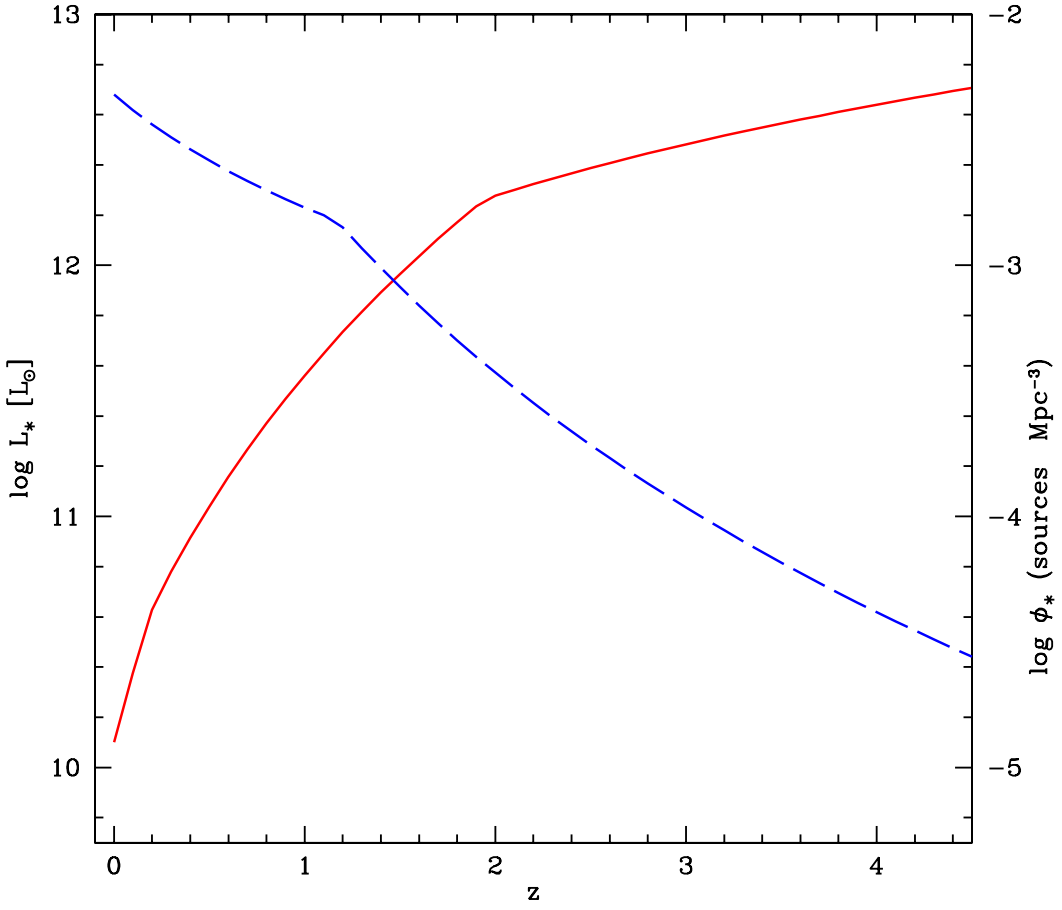

Figure 7.18: Evolution of IR Bolometric Luminosity Function Parameters obtained from $1/V_{est}$ analysis fitted by using the *global* fitting procedure explained in the text. The red filled line represent the luminosity evolution as $L_*(1+z)^{\alpha_L}$; the blue dashed line represents the density evolution as $\phi_*(1+z)^{\alpha_D}$. The luminosity and density evolution are plotted following the left and the right y axis respectively.

7.5 Evolution of Infrared Emissivity and Comoving Star Formation Rate Density

Since, as seen in previous Sections, we believe that we control with sufficient precision both the high-luminosity and the low-luminosity ends of the LFs (the latter particularly thanks to our HerMES Wide Fields LF analysis and available constraints from the CIRB diffuse background as shown in Fig. 7.16 by the SPIRE 250 μm source counts) over a wide range of redshifts, we are confident that we can provide reliable determinations of the galaxy bolometric IR comoving emissivity ρ_{IR} from integrations of the LFs in various redshift bins.

Results from this computation are reported in Fig. 7.19 based on our *single-z* fits in individual redshift bins (Fig. 7.13). Those from analogous calculations for the *global* fit in Fig. 7.15 are shown in Fig. 7.20. The two independent determinations are quite consistent with each other, even if the first is probably affected by greater uncertainties.The galaxy emissivity increases very steeply from $z = 0$ to 0.5, with a lower rate to $z \sim 1.5$, while a clear convergence sets in at $z > 2$.

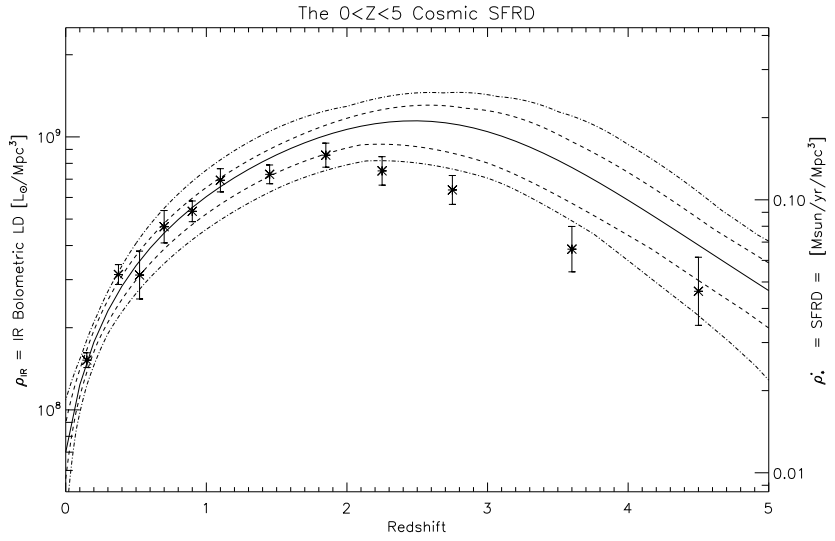

Figure 7.19: the SFRD estimates from COSMOS IR bolometric luminosity function integrated over different z-bins as obtained by the *single-z* fitting procedure explained in the text.

Fig. 7.20 provides also a comparison of our determinations with literature data. At $z < 0.5$ we find excellent agreement of our estimated ρ_{IR} values with those reported by Rodighiero et al. (2010) and by Vaccari et al. (2010), the latter based on the analysis of a preliminary sample of Herschel sources from the SDP programme. At higher redshifts our new determinations tend to keep lower than previous publications, while still marginally consistent within the statistical errors. At $z > 2$ the convergence shown by our data was not seen before, and our analysis becomes inconsistent with previous determinations based of aggressive extrapolations of 24 μm data (as reported by Pérez-González et al. (2003), Pérez-González et al. (2005b) and Reddy et al. (2008), in particular). We do not consider here

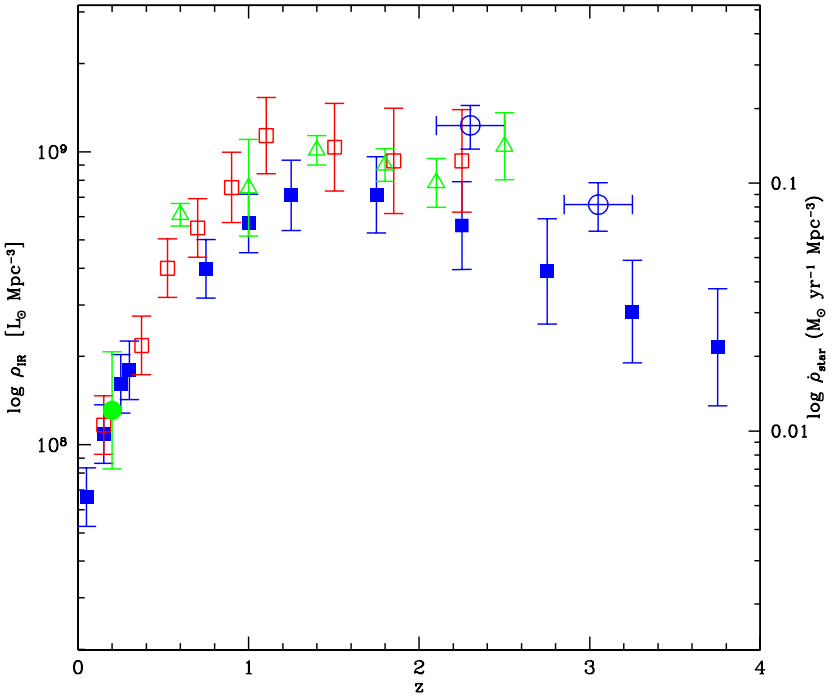

Figure 7.20: The SFRD and the IR Luminosity density estimates from COSMOS IR bolometric luminosity function integrated over different z-bins as obtained by the *global* fitting procedure explained in the text. Blue filled squares: this work; red empty squares: Rodighiero et al. (2010); blue empty circles: Reddy et al. (2008); green empty triangles: Pérez-González et al. (2005b); green filled circle: Vaccari et al. (2010).

other previous determinations at very high redshifts, e.g. those based on SCUBA data, that required enormous extrapolations in wavelength and luminosity.

A decade of studies of IR galaxy populations have coherently demonstrated that the primary power source in these objects are young massive stars, while a minor contribution of about 10-15% may be attributed to obscured AGN emission (e.g. Fadda et al. (2002), Fadda et al. (2010); Alexander et al. (2005)), although the detailed value of the AGN contribution is still uncertain. Consequently, with small correction factors, the plot of Fig. 7.20 may be viewed as the evolution with cosmic time of the comoving rate of star formation, concerning the part of it obscured by dust.

To this IR we have to add the fraction of star formation unobscured by dust, as traceable by UV observations, if we want the complete history to be described. We report in Figure 7.21 a comparison of the (comoving) star-formation rate history recovered by Herschel and Spitzer/MIPS observations with that inferred from UV-selected high-z galaxy samples. The two contributions are completely independent one from the other (the UV-selected one is not corrected for extinction), so that the total star-formation history is the sum of the two. It is interesting to note that, while the dust-extinguished and un-extinguished phases provide

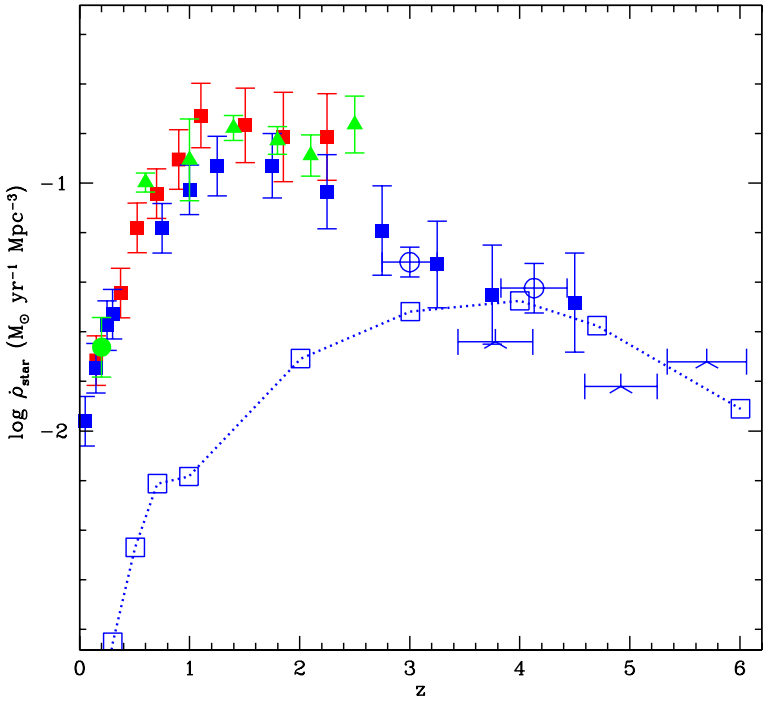

Figure 7.21: Comparison of estimates of the comoving star-formation rate density contributed by IR-selected galaxies with that by UV-selected objects. Filled blue squares are our own estimate from Herschel sources, open squares and circles and stars trace the UV-selected star-formation rate estimated by Shimasaku et al. (2005). For other datapoints, see caption to Fig. 7.20.

comparable contributions at high redshifts, $z \sim 3$ to 5, the dusty phase that we have measured in the Herschel surveys completely dominates at lower redshifts.

7.6 Constraints on Physical Models of Galaxy Formation

The Herschel Space Observatory FIR and SMM observations have offered us for the first time a chance to measure, in combination with shorter wavelength data, the bolometric emission of very numerous, complete samples of cosmic sources over a large fraction of the Hubble time. Precise measurements of FIR/SMM light is particularly relevant to track phases of enhanced star-formation (and to a similar extent, of gravitational accretion in AGNs), because these typically happen in strongly dust-extinguished environments (e.g. Sanders & Mirabel (1996)).

With all these data in hand, we have now a complete census of the cosmic sites where star-formation has happened. And here comes the surprise illustrated in Fig. 7.19 and 7.20. Going back in time, at least a naive rendition of the hierarchical clustering paradigm would expect star formation to happen in progressively more numerous, but lower luminosity (and lower mass) objects. Subsequent merging would then produce more massive and luminous galaxies at increasing time.

In fact, our results in Fig. 7.19 and 7.20 reveal the exactly opposite trend: going back in cosmic time, star-formation occurs in more and more luminous objects, becoming progressively rarer. This behaviour is quite more reminiscent of galaxy formation models considering fast monolithic collapse of primordial gas to form already massive galaxies at relatively high-redshifts. (e.g. Eggen et al. (1962); Larson & Tinsley (1978); Renzini (2006)).

Indeed, a number of more or less strong indications have accumulated during the last 15 years that the above mentioned hierarchical paradigm has some fundamental flaws, starting from the seminal paper by Cowie et al. (1996) concerning the *downsizing* effect about the evidence of the emergence of lower mass and luminosity galaxies at progressing cosmic time (see also, among a numerous literature, Gavazzi et al. (1996), Franceschini et al. (1998), Cimatti et al. (2004) and Franceschini et al. (2006)).

In our view, one of the most impressive counter-arguments against the simple hierarchical model are offered by our present findings summarized by our Fig. 7.19 and 7.20.

To better quantify this possible disagreement, we have performed a preliminary systematic comparison of our data with predictions of semi-analytical models of galaxy formation as reported in Fontanot et al. (2012) using the MORGANA code by Monaco et al. (2007). These include standard descriptions of the evolution of the dark-matter field, of the collapse of baryons and the formation of stars, and include a treatment of dust extinction and photon re-processing in the FIR.

We report in Fig. 7.22 and 7.23 results of such comparison for two representative selection wavelengths at 250 and 100 μm respectively.

While the Fontanot et al. (2012) models appear to reproduce well our data on the multi-wavelength local LFs, the disagreement with the data becomes progressively more serious at increasing redshift, and is a double-faced problem. On one side the model is unable to generate galaxies more luminous than $L_{250} > 10^{11}$ and $L_{250} > 10^{12}$ $L_{\odot}$, while the data show them to dominate more and more at redshifts $z > 1.5$. On the other hand, the model predicts a number density of sources only slightly increasing with redshift, such that the expected number of moderate-luminosity objects remains roughly constant at $z \sim 1.5$ to 4, and results into a conflict of almost 2 orders of magnitude compared to what is allowed by our data (see e.g. LF fits in Fig. 7.15 and 7.14).

It should be stressed that the results of Fontanot et al. (2012), as discussed in the paper,

are entirely representative of and equivalent to several others based on a variety of models appeared in the literature (e.g. Monaco et al. (2007); Wang et al. (2008); Somerville et al. (2012). The difficulties described in the current section appear common to all attempts to frame the formation of galaxies into a self-consistent cosmological context starting from first principles (the Big Bang, and known physics). Our present data do not certainly go in the direction of easing such difficulties.

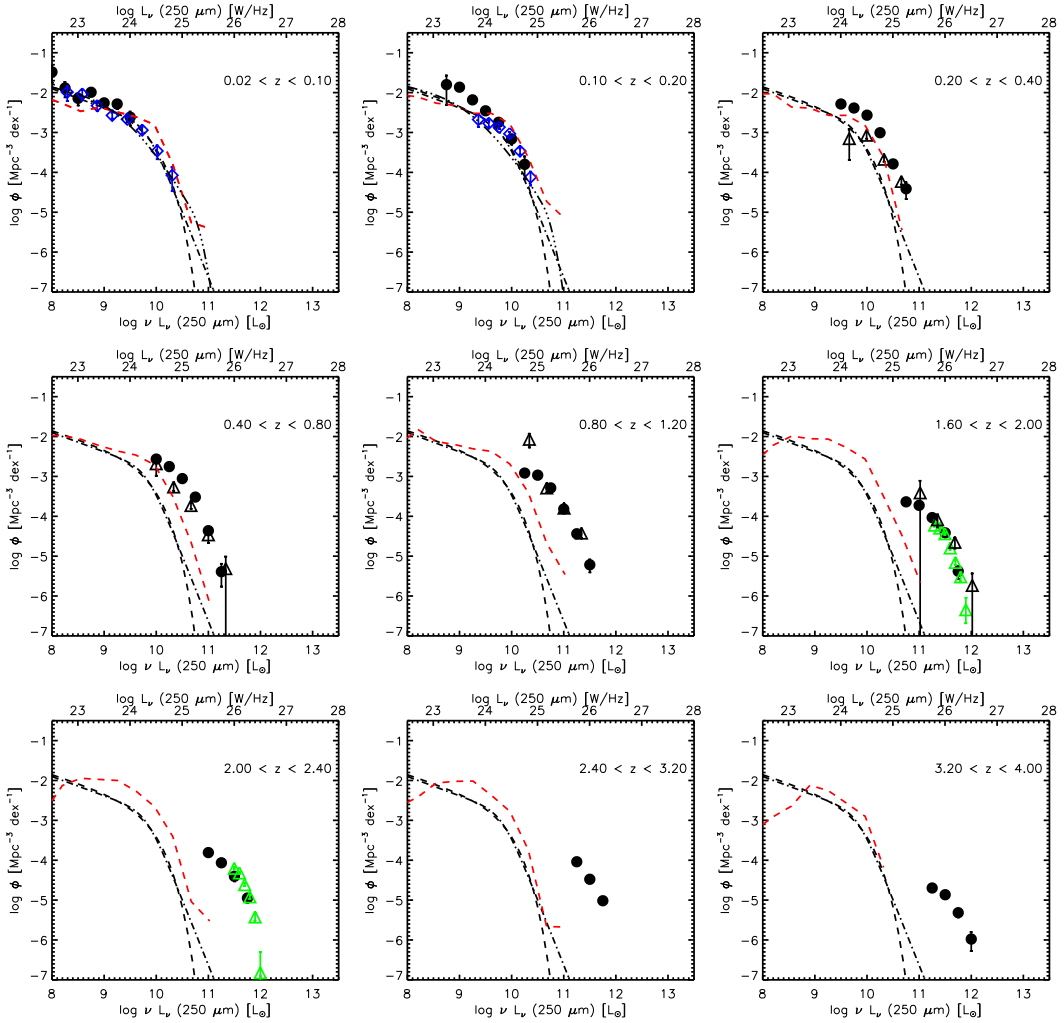

Figure 7.22: The SPIRE 250 μm luminosity function estimate using data in the COSMOS SPIRE 250 μm selected sample compared to Fontanot et al. (2012) models. Black filled point are the $1/Vmax$ result compared with the model's results by Fontanot et al. (2012) (red dashed line). Black empty triangles: SPIRE 250 μm luminosity function by Eales et al. (2010b); blue empty diamonds: Dye et al. (2010) SPIRE 250 μm local luminosity function estimates using the first $\sim$ 14 deg^2 observed within the Herschel-ATLAS projects; black dashed dot dot dot line: Negrello et al. (2007) model.

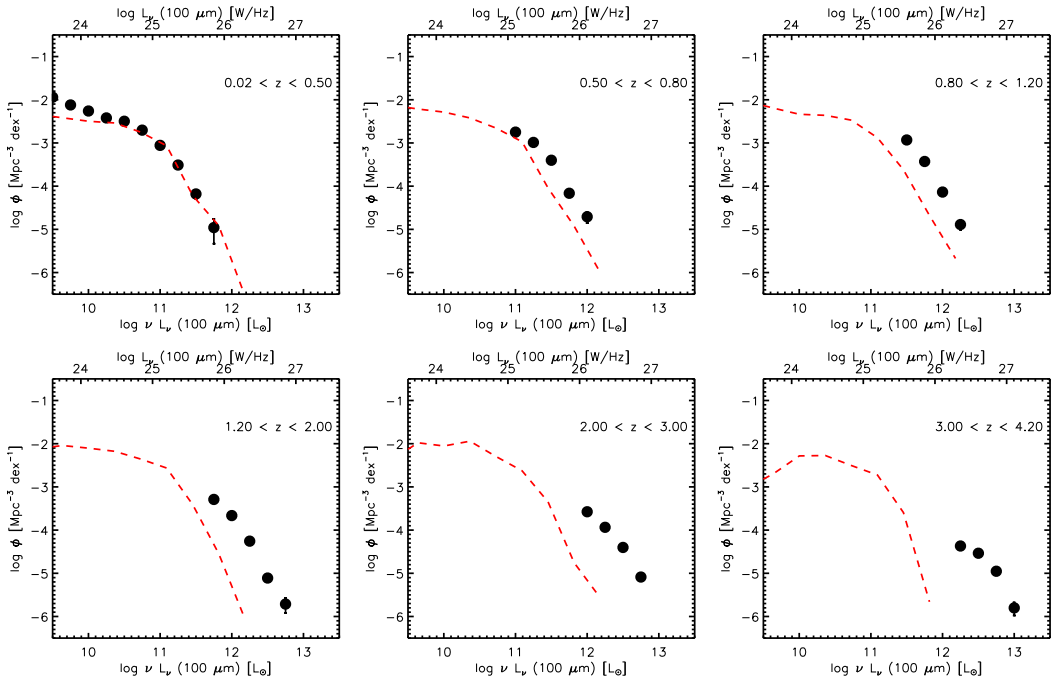

Figure 7.23: The SPIRE 100 μm luminosity function estimate using data in the COSMOS SPIRE 250 μm selected sample compared to Fontanot et al. (2012) model. Black filled point are the $1/Vmax$ result compared with the model's results by Fontanot et al. (2012) (red dashed line).

Conclusions

Infrared Galaxies in the Low-Redshift & High-Redshift Universe

Infrared wavelengths contain a substantial amount of information on the origin of galaxies and Active Galactic Nuclei and about the evolutionary history of star formation and gravitational accretion. They offer a widely complementary view with respect to more classical surveys in the optical. This Thesis work was devoted to contributing new estimates of the multi-wavelength luminosity functions of infrared sources and their evolution with cosmic time. To this end, we have carried out a detailed investigation of statistical properties of infrared galaxies detected by the Spitzer and Herschel infrared space satellites over a large fraction of cosmic time with the aim of establishing in a most accurate way their evolutionary properties.

Spitzer & Herschel Deep & Wide Cosmological Surveys

With the above purposes in mind, we have exploited some of the deepest and widest-area sky surveys in the mid-infrared, far-infrared and sub-mm obtained with the Spitzer Space Telescope and with the Herschel Space Observatory. These two space missions jointly cover more than two decades in wavelengths, roughly spanning from from 3 to 600 micron, and, thanks to their versatile instrumentation, have allowed us to make during the last ten years orders-of-magnitude improvement in the depth and area of infrared extragalactic surveys.

The Spitzer Wide Area Infrared Extragalactic (SWIRE) survey in six fields, the Spitzer Deep Wide Field Survey (SDWFS) in the Bootes field and the Spitzer Extragalactic First Look Survey (XFLS) in the XFLS field, jointly covering 70 $\deg^2$, were re-reduced and carefully combined with the most extensive ultraviolet, optical and near-infrared photometric and spectroscopic surveys currently available in these fields. The resulting Spitzer Data FusionSpitzer Data Fusion is an unrivalled resource for infrared galaxy formation and evolution studies in the Local Universe, and particularly so for multi-wavelength studies of dust-obscured star formation and black hole accretion and their dependence on environment.

This well-documented and easy-to-use database has proved invaluable in detecting and characterizing Herschel sources within the shallow Wide Fields observed by the Herschel Multi-tiered Extragalactic Survey (HerMES). Furthermore, the versatile nature of the procedures developed to produce the database also allowed us to adapt them to the needs of the Spitzer Exploratory Representative Volume Survey (SERVS) and will be useful to VISTA and VST survey programs in the near future. Thanks to agreements with other consortia, we were then able to access and exploit proprietary multi-wavelength databases in the HerMES deeper and

narrower COSMOS field, which allowed us, along with PACS data provided by the PACS Extragalactic Probe (PEP) survey, to follow the evolution of Spitzer and Herschel sources from the low-redshift to the high-redshift Universe.

Spitzer and Herschel Luminosity Functions

Armed with the above datasets, we have embarked on an extensive study of Spitzer and Herschel luminosity functions. We first exploited the Spitzer Data Fusion to determine a reliable estimate of the Spitzer local luminosity functions at 24, 70, and 160 μm based on MIPS 24, 70 and 160 μm samples, respectively, thus providing an important local benchmark for studies exploiting Spitzer deeper observations.

We then used the HerMES Wide Fields and the PEP/HerMES COSMOS field, covering almost 30 and 2 deg^2, and studied their SPIRE 250 μm population to probe the low-redshift and high-redshift ends of the SPIRE luminosity function, respectively.

A prerequisite for these studies was to carry out a fully-fledged SED fitting analysis of about 20 thousand IR sources. Most obviously, this was required for obtaining precise k-corrections and and thus rest-frame luminosities at various FIR/SMM wavelengths (250/160/90/60 μm) and IR bolometric luminosities. Furthermore, in the medium term we plan to exploit this analysis for a physical characterization of individual sources, their classification into sub-populations and studying their statistical properties.

Luminosity functions and cosmological evolution have been explored using a variety of statistical methods and tests, from the classical $1/V_{max}$, to the $1/V_{est}$ (including a version slightly modified by us) and a *semi-parametric* estimator. These analyses have coherently found evidence for strong evolution of the IR bolometric luminosity function, both in comoving number density and in luminosity. The surprising novelty consisted, in particular, in the discovery with high statistical significance of negative evolution back in time of the galaxy comoving volume density, by a factor about one hundred over the redshift interval $0 < z < 4$. In spite of this, the galaxy IR emissivity increases by about a factor ten from local to $z \sim 2$, due to very strong increase of the galaxy average luminosity with redshift.

Our results then imply a kind of *inversion* of the hierarchical clustering paradigm, i.e. a process going from top to bottom with time, instead of the bottom-up expected trend. Indeed, from a comparison with state-of-art physical models of galaxy formation and evolution, our results appear to set a radical challenge to our physical understanding of the origins of cosmic structure.

Future Perspectives

While the Spitzer and Herschel missions will be coming to a close over next the few years, their data reduction and combination with improved ground-based optical and near-infrared photometry and spectroscopy will remain a fruitful realm of study for years to come. ALMA will soon be completed and will probe Herschel (and SCUBA2) sources at much greater angular resolution and flux sensitivity, and JWST will provide similar improvements at wavelengths previously accessible (from space) only by Hubble and/or Spitzer, albeit over a limited field of view. Eventually, Euclid and SKA deep and wide surveys will provide orders-of-magnitudes

improvements in survey speeds, ideally bridging the gap between galaxy evolution and observational cosmology with much improved measures of stellar mass and star formation rate budgets across cosmic space and time. The resources and techniques developed as part of this Thesis will thus be of the uttermost importance in readily handling the increasing data rates to be provided by future instrumentation, and thus in timely answering deeper questions about the nature of infrared galaxies, their formation and evolution.

Appendix A

Markov Chain Monte Carlo & The Metropolis-Hastings Algorithm

Monte Carlo methods are computational algorithms that rely on random sampling and that are guided by certain rules designed to give the desired outcome. An important subclass of Monte Carlo methods consists of the Markov Chain Monte Carlo (MCMC) methods, defined as those in which a given step in the sequence depends only upon the previous step. The sequence of steps is known as a Markov chain. Each step corresponds to some particular value of the parameters, for which the likelihood is evaluated. The Markov chain can therefore be viewed as a series of steps (or jumps) around the parameter space that investigates the likelihood function shape as it goes. The simplest task that a Markov chain in this situation could carry out would be to find the maximum of the likelihood: Choose a random starting point, propose a random jump to a new point, accept the jump only if the new point has a higher likelihood, return to the proposal step, and repeat until the highest point has been found. Even this simple algorithm needs some tuning: If the steps are too large, the algorithm may soon find it difficult to successfully find a higher-likelihood point to jump to, whereas if the steps are too small, the chain may get stuck in a local maximum that is not the global maximum. The latter problem may be overcome if a series of chains are run from different starting points. The maximum alone is not of great interest; what we want to know is the region around the maximum that is compatible with the data. To this end, we need an algorithm in which the Markov chain elements correspond to random samples from the posterior parameter distribution of the parameters; that is, each chain element should represent the probability that those particular parameter values are the true ones. The simplest algorithm that achieves this goal is the Metropolis-Hastings algorithm, a remarkably straightforward modification of the algorithm described by these steps (Liddle 2009):

1. Choose a starting point within the parameter space.

2. Propose a random jump. Any function can be used to determine the probability distribution for the length and direction of the jump, as long as it satisfies the "detailed balance" condition that a jump back to the starting point is as probable as the jump away from it. This condition is most easily met by using a symmetric proposal function, such as a multivariate Gaussian about the current point. Evaluate the likelihood at the new point and hence the probability by multiplying by the prior at that point. (If the

prior is flat, the probability and likelihood become equivalent.).

3. If the probability at the new point is higher, accept the jump. If it is lower, accept the jump with the probability given by the ratio of the probabilities at the new and old points. If the jump is not accepted, stay at the same point, creating a duplicate in the chain.

4. Repeat Steps 2 and 3 until satisfied that the probability distribution is well mapped out, for instance by comparing several chains run from different starting points and/or by using convergence statistics, among which the Gelman-Rubin test (used in Patel et al. (2012), see Gelman et al. (2004) for details on the test) is the most common.

When a chance of moving to a lower probability point is introduced, the algorithm can explore the shape of the posterior in the vicinity of the maximum. Generically, the algorithm starts in a low-likelihood region and migrates toward the high-likelihood "mountains". Once the algorithm nears the top, most of the possible jumps are in the downward direction, and the chain meanders around the maximum mapping out its shape. Accordingly, all the likelihood evaluations, which are where the CPU time is spent, are carried out in the region in which the likelihood is large enough to be interesting. The exception is the early stage, which is not representative of the posterior distribution as it maintains a memory of the starting position. This so-called burn-in phase is then deleted from the list of chain points. Although any choice of proposal function that satisfies the detailed balance condition ultimately yields a chain sampling from the posterior probability distribution, in practice – as with the simple example above – the algorithm needs to be tuned to work efficiently. The tuning process is referred to as the convergence of the chain (to the posterior probability). The proposal function should be tuned to the scale of variation of the likelihood near its maximum, and if the usual choice of a Gaussian is made, its axes should ideally be aligned to the principal directions of the posterior (so as to navigate quickly along parameter degeneracies). Usually, a short initial run is carried out to roughly map out the posterior distribution, which is then used to optimize the proposal function for the actual computation.1 The resulting acceptance rate of the new points tends to be around 25%.

Schematic diagram illustrating the random walk MHA is reported in Fig. A.1 with θ as first guess on posterior probability.

The Metropolis-Hastings algorithm achieves the desired goal, but it may not do so efficiently if it is difficult to find a good proposal function or if the assumption of a fixed proposal function proves disadvantageous. This tends not to be a problem in the common astonomical applications, but it is nevertheless worthwhile to know that there are alternatives that may be more robust. Some examples are:

1. *Slice sampling.* This method allows the proposal function to change during the calculation, tuning itself to an appropriate scale, although there is an additional computational cost associated with enforcing the detailed balance condition. The steps are made in a single parameter direction at a time (hence the name), and they cycle through the parameter directions either sequentially or randomly.

2. *Gibbs sampling.* This method relies on obtaining a proposed step by sampling from conditional probability distributions; to step in the θ_1 direction one samples from $P(\theta_1|\theta_2)$ and vice versa. It turns out that such proposals are always accepted, enhancing the

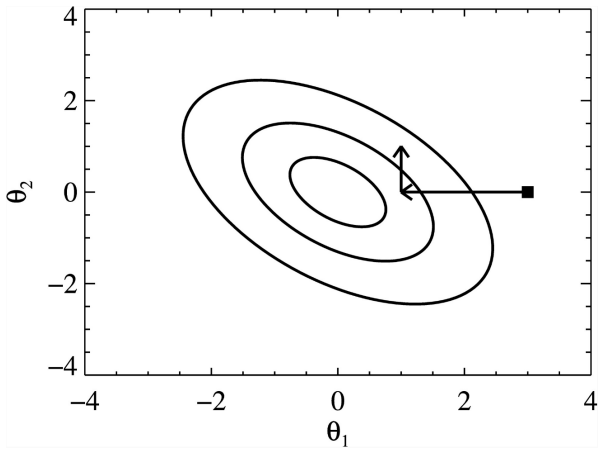

Figure A.1: Schematic diagram illustrating the random walk MHA (Kelly et al. 2008). The posterior probability distribution is illustrated by the contours, and the random walk is initially at the position marked with a square. A new proposed value of θ_1 is randomly drawn, marked by the arrow pointing to the left. Because the proposed value of θ_1 is at a location with higher posterior probability, the new value of θ_1 is saved, and the random walk "jumps" to the position marked by the arrow. Then, a new proposal for θ_2 is randomly drawn, marked by the arrow pointing upward. Because this proposed value of θ_2 is at a location with lower posterior probability, it is only accepted with probability equal to the ratio of the values of the posterior at the proposed position and the current position. If the proposed value is kept, then the new value of θ_2 is saved; otherwise, the current value of θ_2 is saved. Next, a proposed value of θ_1 is randomly drawn, and the process repeats, creating a random walk through the parameter space. Because the amount of time that the random walk spends in any given bin in θ_1 and θ_2 is proportional to the posterior probability distribution, after the random walk has converged, the values of θ_1 and θ_2 from the random walk may be treated as a random draw from the posterior distribution.

efficiency. The method can, however, struggle to make progress along highly correlated parameter directions, traversing the diagonal through a series of short steps parallel to the axes.

3. *Hamiltonian sampling.* This more sophisticated approach uses an analogy with Hamiltonian dynamics to define a momentum from derivatives of the likelihood. The momentum associated with a point enables large proposal steps to be taken along trajectories of constant energy. This method is particularly well adapted to very high dimensionality problems.

Appendix B

Spitzer Data Fusion Database of Instrumental Transmissions

The Spitzer Data Fusion and its extension to deep fields such as ADFS, COSMOS, ECDFS, EGS, EHDFN, GOODS-N, GOODS-S, LHE, LHN, UDS, VVDS merges observations obtained with a wide variety of ground-based and space-based telescopes, and the SED template fitting process required to make sense of the physical properties of detected sources therefore combines photometric observations obtained from most of these instruments. It is thus mandatory to assemble accurate Instrumental Transmissions (incorporating the Telescope, Filter and Detector transmissions) for the various datasets. The filter database originally distributed with the `Hyperz` package, using the same filter specification format as the `LePhare` package, was extensively updated and old and new filters were researched in the literature and online to provide a reliable instrumental transmission to be used in SED template fitting in Chap. 5. The full list of filters is included in Tab. B.1. The latest pubic release of the database can be obtained at `http://www.astro.unipd.it/background/df/filters`.

Table B.1: List of Filters included in the Spitzer Data Fusion Database of Instrumental Transmissions. Filters indicted with an asterisk are from the filter database originally distributed assembled with the `Hyperz` package.

ID	Description	NP	DF
0001	Koo-Kron U+ filter (Koo's thesis)	13	*
0002	Koo-Kron J+ filter (Koo's thesis)	23	*
0003	Koo-Kron F+ filter (Koo's thesis)	25	*
0004	Koo-Kron N+ filter (Koo's thesis)	28	*
0005	Koo-Kron R band (=127+RG610, data from Koo, Durham)	15	*
0006	Couch and Newell (80) BJ (photographic) filter	43	*
0007	Couch and Newell (80) RF (photographic) filter	23	*
0008	Koo-Kron U+ filter (Bruzual's thesis)	13	*
0009	Koo-Kron J+ filter (Bruzual's thesis)	25	*
0010	Koo-Kron F+ filter (Bruzual's thesis)	25	*
0011	Koo-Kron N+ filter (Bruzual's thesis)	23	*

0012	Buser's U filter	24	*
0013	Buser's B2 filter	40	*
0014	Buser's B3 filter	40	*
0015	Buser's V filter	54	*
0016	Matthews and Sandage U filter	13	*
0017	Matthews and Sandage B filter	22	*
0018	Matthews and Sandage V filter	25	*
0019	Sandage and Smith B filter	43	*
0020	Sandage and Smith V filter	38	*
0021	Sandage and Smith R filter	43	*
0022	ST-UV14 filter	33	*
0023	ST-UV17 filter	42	*
0024	ST-UV22 filter	53	*
0025	ST-UV27 filter	64	*
0026	OAO-UV1 filter	33	*
0027	OAO-UV2 filter	47	*
0028	OAO-UV3 filter	59	*
0029	OAO-UV4 filter	68	*
0030	OAO-UV5 filter	58	*
0031	OAO-UV6 filter	70	*
0032	Johnson's R filter	23	*
0033	Johnson's I filter	26	*
0034	Johnson's J filter	31	*
0035	Johnson's K filter	17	*
0036	Johnson's L filter	14	*
0037	Butcher's r filter	51	*
0038	Butcher's i filter	36	*
0039	Butcher-Oemler R filter (10/75 1978, data from Koo, Durham)	31	*
0040	Butcher-Oemler R filter (5/76 1978, data from Koo, Durham)	31	*
0041	Bessell u filter	23	*
0042	Bessell g filter	46	*
0043	Bessell r filter	22	*
0044	UKIRT H FILTER (Leiden, 1983)	33	*
0045	R. S. Ellis U(PE) filter	53	*
0046	R. S. Ellis J filter	105	*
0047	R. S. Ellis R filter	57	*
0048	R. S. Ellis N filter	111	*
0049	C. MacKay and P. Hall KG3 filter (Cambridge)	189	*
0050	C. MacKay and P. Hall I filter (Cambridge)	126	*
0051	Gunn g filter + four-shooter Ti CCD + Palomar 200" atmospher	29	*
0052	Gunn r filter + four-shooter Ti CCD + Palomar 200" atmospher	47	*
0053	Gunn i filter + four-shooter Ti CCD + Palomar 200" atmospher	75	*
0054	Gunn z filter + four-shooter Ti CCD + Palomar 200" atmospher	81	*
0055	IR J filter + Palomar 200 IR detectors + atmosphere	21	*
0056	IR H filter + Palomar 200 IR detectors + atmosphere	27	*
0057	IR K filter + Palomar 200 IR detectors + atmosphere	33	*

0058	NOAO CTIO 4m ISPI J#186	46	
0059	NOAO CTIO 4m ISPI H#187	76	
0060	NOAO CTIO 4m ISPI K'#188	639	
0061	A. Tyson J filter	40	*
0062	A. Tyson R filter	41	*
0063	A. Tyson I filter	65	*
0064	ANS 1550 Wide Filter (J. Koorneef)	4	*
0065	ANS 1800 Filter (J. Koorneef)	18	*
0066	ANS 2200 Filter (J. Koorneef)	4	*
0067	ANS 2500 Filter (J. Koorneef)	4	*
0068	ANS 3300 Filter (J. Koorneef)	23	*
0069	Approximate U band for Lilly and Cowie	13	*
0070	Approximate I band for Lilly and Cowie	61	*
0071	IRAS 12 micron, Neugebauer etal 1984,ApJL,278,L1	35	*
0072	IRAS 25 micron, Neugebauer etal 1984,ApJL,278,L1	28	*
0073	IRAS 60 micron, Neugebauer etal 1984,ApJL,278,L1	46	*
0074	IRAS 100 micron, Neugebauer etal 1984,ApJL,278,L1	32	*
0075	H filter Bessell and Brett PASP 100, 1134, 1988	20	*
0076	J filter Bessell and Brett PASP 100, 1134, 1988	21	*
0077	K filter Bessell and Brett PASP 100, 1134, 1988	25	*
0078	L (3.5 microns) filter Bessell and Brett PASP 100, 1134, 198	21	*
0079	L' (3.8 microns) filter Bessell and Brett PASP 100, 1134, 19	20	*
0080	M filter Bessell and Brett PASP 100, 1134, 1988	19	*
0081	IRAM MAMBO-1 1.2 mm, 37 channel (winter 99/00 -today)	247	
0082	IRAM MAMBO-2 1.2 mm,117 channel	221	
0083	g Gunn (original)	17	*
0084	r Gunn (original)	22	*
0085	i Gunn (original)	27	*
0086	z (original)	22	*
0087	z + RCA	22	*
0088	CCD RCA ESO (JPP reference)	15	*
0089	CCD RCA CAHA (Manual d'utilisateurs)	8	*
0090	B CAHA (original manuel)	16	*
0091	B Bessell	24	*
0092	V Bessell	18	*
0093	R Bessell	22	*
0094	I Bessell	12	*
0095	K Prime CFHT Redeye	110	*
0096	CCD RCA2 CFHT (Manuel utilisateurs)	24	*
0097	Bj TYSON (orig. filter AT, private com.)	22	*
0098	CCD TEK#25 (ESO, Manuel Utilisateurs)	17	*
0099	CCD LORAL#34 (ESO, Manuel Utilisateurs)	17	*
0100	CCD SAIC#1 (CFH, Manuel Utilisateurs)	15	*
0101	CCD Lick2 CFHT (CFH, Manuel Utilisateurs)	8	*
0102	ESO NTT SUSI B Bessell#639	200	*
0103	ESO NTT SUSI V Bessell#641	200	*

0104	ESO NTT SUSI R Bessell#642	200	*
0105	ESO NTT EMMI V#606	200	*
0106	B#4402 CFHT	23	*
0107	R#4609 CFHT	18	*
0108	B#1412 CFHT FOCAM	148	*
0109	B#1414 CFHT B Tyson selon JB	24	*
0110	V#1504 CFHT	113	*
0111	V#1510 CFHT FOCAM	165	*
0112	R#1611 CFHT	29	*
0113	I#1808 CFHT FOCAM	65	*
0114	I#1809 CFHT FOCAM	65	*
0115	Thomson THX 31156 CCD#17 ESO	17	*
0116	Thomson THX 31156 CCD#18 ESO	17	*
0117	R#585 Bessell ESO	30	*
0118	K#6 UKIRT	36	*
0119	Passe-tout	41	*
0120	F555W + WFPC2 normalized	28	*
0121	F814W + WFPC2 normalized	142	*
0122	F300W + WFPC2 normalized	156	*
0123	F450W + WFPC2 normalized	154	*
0124	F606W + WFPC2 normalized	145	*
0125	F702W + WFPC2 normalized	28	*
0126	F675W + WFPC2 normalized	18	*
0127	F336W + WFPC2 normalized	17	*
0128	ESO NTT 3.6m SOFI Js	262	
0129	ESO NTT 3.6m SOFI J	959	
0130	ESO NTT 3.6m SOFI H	493	
0131	ESO NTT 3.6m SOFI Ks	365	
0132	KPNO IRIM 2.12 Filter	23	*
0133	KPNO IRIM 2.14 Filter	26	*
0134	KPNO IRIM 2.16 Filter	29	*
0135	KPNO IRIM H Filter	43	*
0136	KPNO IRIM J Filter	38	*
0137	KPNO IRIM K Filter	62	*
0138	KPNO IRIM K' Filter	46	*
0139	VLT Test Camera Detector's Quantum Efficiency	19	*
0140	B-band filter of the VLT Test Camera	21	*
0141	V-band filter of the VLT Test Camera	24	*
0142	R-band filter of the VLT Test Camera	24	*
0143	I-band filter of the VLT Test Camera	23	*
0144	SUSI2's CCDs Quantum Efficiency	8	*
0145	SUSI Bessell U#810	200	*
0146	SUSI Bessell B#811	196	*
0147	SUSI Bessell V#812	200	*
0148	SUSI Bessell R#813	200	*
0149	SUSI Bessell I#814	200	*

0150	FORS Standard U (including instrument + CCD)	79	*
0151	FORS Standard Bessell B (including instrument + CCD)	77	*
0152	FORS Standard Bessell V (including instrument + CCD)	79	*
0153	FORS Bessell R (including instrument + CCD)	79	*
0154	FORS Bessell I (including instrument + CCD)	79	*
0155	FORS Gunn G (including instrument + CCD)	79	*
0156	ESO 2.2m WFI U#841+CCD57+wfi_2p2_optics (U/38 AKA U38)	106	
0157	ESO 2.2m WFI B#842+CCD57 (old B/99, for new see B/123)	101	*
0158	ESO 2.2m WFI V#843+CCD57+wfi_2p2_optics (V/89)	249	
0159	ESO 2.2m WFI Rc#844+CCD57+wfi_2p2_optics (Rc/162)	466	
0160	ESO 2.2m WFI Ic#845+CCD57+wfi_2p2_optics (Ic/lwp)	512	
0161	ESO 2.2m WFI Z#846+CCD57+wfi_2p2_optics (Z+/61)	382	
0162	ESO 2.2m WFI U#877+CCD57+wfi_2p2_optics (U/50 AKA U35)	200	
0163	ESO 2.2m WFI B#878+CCD57+wfi_2p2_optics (B/123)	602	
0164	SDSS u	47	
0165	SDSS g	89	
0166	SDSS r	75	
0167	SDSS i	89	
0168	SDSS z	141	
0169	ESO VST OmegaCAM u	39	
0170	ESO VST OmegaCAM g	83	
0171	ESO VST OmegaCAM r	83	
0172	ESO VST OmegaCAM i	93	
0173	ESO VST OmegaCAM z	76	
0174	CFHT CFH12k B (Mould)	63	
0175	CFHT CFH12k V (Mould)	147	
0176	CFHT CFH12k R (Mould)	87	
0177	CFHT CFH12k I (Mould)	130	
0178	CFHT CFH12k Z (Prime)	246	
0179	JCMT SCUBA 450 micron	1299	
0180	JCMT SCUBA 850 micron	822	
0181	AzTEC 1.1 mm	275	
0182	Infamous 2.2m UH8K B filter + loral3 + MK atmosphere	61	*
0183	2.2m UH8K V filter + loral 3 + atmosphere	64	*
0184	2.2m UH8K I filter + MK atmosphere	126	*
0185	KPNO B, from AAT Users Manual	20	*
0186	H+K filter	150	*
0187	Wyin filter U (filter + CCD reponse)	13	*
0188	Wyin filter B (filter + CCD reponse)	24	*
0189	ESO VLT ISAAC J (ESO web pages)	169	
0190	ESO VLT ISAAC H (ESO web pages)	167	
0191	ESO VLT ISAAC Ks (ESO web pages)	173	
0192	ESO VLT ISAAC L (ESO web pages)	132	
0193	ESO VLT ISAAC M (ESO web pages)	131	
0194	Palomar 200" WIRC J	42	
0195	Palomar 200" WIRC K	66	

0196	Calar Alto 3.5m Omega2000 J	109	
0197	Calar Alto 3.5m OmegaPrime K	44	
0198	Spitzer IRAC CH1 (3.6 micron)	391	
0199	Spitzer IRAC CH2 (4.5 micron)	401	
0200	Spitzer IRAC CH3 (5.8 micron)	311	
0201	Spitzer IRAC CH4 (8.0 micron)	350	
0202	Spitzer MIPS CH1 (24 micron)	128	
0203	Subaru SuprimeCam U	217	*
0204	Subaru SuprimeCam B	200	
0205	Subaru SuprimeCam V	200	
0206	Subaru SuprimeCam r	200	
0207	Subaru SuprimeCam i	200	
0208	Subaru SuprimeCam z	200	
0209	UH 2.2m QUIRC H+K (AKA HK')	108	
0210	CFHT MegaCam i2 AKA y (new, after October 2007)	38	
0211	NOAO FLAMINGOS J (J-2000toJuly2003)	46	
0212	NOAO FLAMINGOS H (H-2000toJuly2003)	76	
0213	NOAO FLAMINGOS Ks (Ks-2000toJuly2003)	699	
0214	Spitzer MIPS CH2 (70 micron)	111	
0215	Spitzer MIPS CH3 (160 micron)	400	
0216	250 micron SPIRE/Herschel	426	
0217	350 micron SPIRE/Herschel	431	
0218	500 micron SPIRE/Herschel	399	
0219	2MASS J	109	
0220	2MASS H	57	
0221	2MASS Ks	76	
0222	UKIRT WFCAM (UKIDSS) J	581	
0223	UKIRT WFCAM (UKIDSS) H	581	
0224	UKIRT WFCAM (UKIDSS) K	581	
0225	INT WFC u	401	
0226	INT WFC g	401	
0227	INT WFC r	500	
0228	INT WFC i	499	
0229	INT WFC z	101	
0230	NOAO KPNO 4m MOSAIC1 U	751	
0231	NOAO KPNO 4m MOSAIC1 g	411	
0232	NOAO KPNO 4m MOSAIC1 r	391	
0233	NOAO KPNO 4m MOSAIC1 i	441	
0234	NOAO KPNO 4m MOSAIC1 z	801	
0235	NOAO CTIO 4m MOSAIC2 u	888	
0236	NOAO CTIO 4m MOSAIC2 g	888	
0237	NOAO CTIO 4m MOSAIC2 r	888	
0238	NOAO CTIO 4m MOSAIC2 i	888	
0239	NOAO CTIO 4m MOSAIC2 z	888	
0240	NOAO CTIO 4m MOSAIC2 U	787	
0241	CFHT MegaCam u*	61	

0242	CFHT MegaCam g	61
0243	CFHT MegaCam r	55
0244	CFHT MegaCam i1 AKA i (old,before October 2007,for new see i2)	65
0245	CFHT MegaCam z	63
0246	AKARI N60	100
0247	AKARI WIDE-S	100
0248	AKARI WIDE-L	100
0249	AKARI N160	100
0250	PACS 70 Instrument Simulator as of Herschel Launch	45
0251	PACS 100 Instrument Simulator as of Herschel Launch	52
0252	PACS 160 Instrument Simulator as of Herschel Launch	39
0253	NOAO FLAMINGOS J from 2003 to present (MAN240B)	260
0254	NOAO FLAMINGOS H from 2003 to present (MAN109A)	874
0255	NOAO FLAMINGOS Ks from 2003 to present (MAN306A)	1012
0256	NOAO FLAMINGOS K (2000toPresentDaytNOAO-OCLI)	50
0257	Subaru SuprimeCam Rc	200
0258	Subaru SuprimeCam Ic	200
0259	Subaru SuprimeCam g	200
0260	Spitzer IRS 16 micron (bluePUtrans)	371
0261	Spitzer IRS 22 micron (redPUtrans)	215
0262	ESO VLT VIMOS U (transmission is average of 4 quadrants)	506
0263	ESO VLT VIMOS B (transmission is average of 4 quadrants)	512
0264	ESO VLT VIMOS V (transmission is average of 4 quadrants)	701
0265	ESO VLT VIMOS R (transmission is average of 4 quadrants)	843
0266	ESO VLT VIMOS I (transmission is average of 4 quadrants)	843
0267	ESO VLT VIMOS z (transmission is average of 4 quadrants)	843
0268	NOAO KPNO 4m MOSAIC1 R	1201
0269	UKIRT WFCAM (UKIDSS) Z	581
0270	UKIRT WFCAM (UKIDSS) Y	581
0271	CFHT WIRCam J (cfh8101)	1077
0272	CFHT WIRCam H (cfh8201)	828
0273	CFHT WIRCam Ks (cfh8302)	910
0274	NOAO KPNO 4m MOSAIC1 Bw	901
0275	NOAO KPNO 4m MOSAIC1 B	1201
0276	NOAO KPNO 4m MOSAIC1 V	1301
0277	NOAO KPNO 4m MOSAIC1 I	351
0278	NOAO CTIO 4m MOSAIC2 B	888
0279	NOAO CTIO 4m MOSAIC2 V	888
0280	NOAO CTIO 4m MOSAIC2 R	887
0281	NOAO CTIO 4m MOSAIC2 I	888
0282	TIFKAM/ONIS J	133
0283	TIFKAM/ONIS H	108
0284	TIFKAM/ONIS K	139
0285	90prime SDSS-u	901
0286	90prime SDSS-z	902
0287	90prime U	146

0288	90prime B	157	
0289	90prime V	102	
0290	90prime R	66	
0291	90prime I	99	
0292	90prime_Washington_M	676	
0293	NEWFIRM J	601	
0294	NEWFIRM H	1021	
0295	NEWFIRM Ks	1041	
0296	GALEX NUV	1321	
0297	GALEX FUV	471	
0298	MMT Megacam u	150	
0299	MMT Megacam g	369	
0300	MMT Megacam r	431	
0301	MMT Megacam i	498	
0302	MMT Megacam z	604	
0303	Subaru MOIRCS Y	616	
0304	Subaru MOIRCS J	1201	
0305	Subaru MOIRCS H	941	
0306	Subaru MOIRCS Ks	569	
0307	Subaru MOIRCS K	987	
0308	APEX SABOCA 350 micron	146	
0309	APEX LABOCA 850 micron	116	
0310	HST NIC3 F110W (J)	739	
0311	HST NIC3 F160W (H)	672	
0312	HST NIC3 F222M (K)	870	
0313	HST ACS/WFC F435W (B)	469	
0314	HST ACS/WFC F606W (V)	745	
0315	HST ACS/WFC F814W (I)	763	
0316	HST ACS/WFC F475W (g)	762	
0317	HST ACS/WFC F625W (r)	471	
0318	HST ACS/WFC F775W (i)	619	
0319	HST ACS/WFC F850LP (z)	660	
0320	Subaru SuprimeCam IA427	423	
0321	Subaru SuprimeCam IA445	483	
0322	Subaru SuprimeCam IA464	421	
0323	Subaru SuprimeCam IA484	501	
0324	Subaru SuprimeCam IA505	501	
0325	Subaru SuprimeCam IA527	532	
0326	Subaru SuprimeCam IA550	301	
0327	Subaru SuprimeCam IA574	611	
0328	Subaru SuprimeCam IA598	799	
0329	Subaru SuprimeCam IA624	668	
0330	Subaru SuprimeCam IA651	360	
0331	Subaru SuprimeCam IA679	760	
0332	Subaru SuprimeCam IA709	643	
0333	Subaru SuprimeCam IA738	638	

0334	Subaru SuprimeCam IA767	323	
0335	Subaru SuprimeCam IA797	392	
0336	Subaru SuprimeCam IA827	572	
0337	Subaru SuprimeCam IA856	141	
0338	Subaru SuprimeCam IA907	74	
0339	Subaru SuprimeCam NA656	344	
0340	Subaru SuprimeCam NB711	651	
0341	Subaru SuprimeCam NB816	520	
0342	Subaru SuprimeCam NB921	338	
0343	LBT-LBC blue Uspec	749	
0344	LBT-LBC blue U	150	
0345	LBT-LBC blue B	300	
0346	LBT-LBC blue V	600	
0347	LBT-LBC blue g (#1)	748	
0348	LBT-LBC blue r (#1)	988	
0349	LBT-LBC red V	730	
0350	LBT-LBC red R	899	
0351	LBT-LBC red I	747	
0352	LBT-LBC red r	699	
0353	LBT-LBC red i	893	
0354	LBT-LBC red z	800	
0355	LBT-LBC red F972N20	33	
0356	LBT-LBC red Y	795	
0357	ISO CAM LW2 (6.7/7 micron)	27	
0358	ISO CAM LW10 (12 micron)	31	
0359	ISO CAM LW3 (14.3/15 micron)	33	
0360	ISO PHT C100-DETECTOR C90-FILTER (90/5 micron)	300	
0361	ISO PHT C200-DETECTOR C160-FILTER (170/5 micron)	300	
0362	VISTA VIRCAM Z	149	
0363	VISTA VIRCAM Y	338	
0364	VISTA VIRCAM J	452	
0365	VISTA VIRCAM H	607	
0366	VISTA VIRCAM Ks	548	

Bibliography

Abazajian, K., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2004, AJ, 128, 502

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, ApJS, 182, 543

Aihara, H., Allende Prieto, C., An, D., et al. 2011, ApJS, 193, 29

Alexander, D. M., Bauer, F. E., Chapman, S. C., et al. 2005, ApJ, 632, 736

Allamandola, L. J., Tielens, A. G. G. M., & Barker, J. R. 1989, ApJS, 71, 733

Allen, C. W. 1976, Astrophysical Quantities

Alonso-Herrero, A., Rieke, G. H., Rieke, M. J., et al. 2006, ApJ, 650, 835

Altieri, B., Metcalfe, L., Kneib, J. P., et al. 1999, A&A, 343, L65

Arnouts, S., Cristiani, S., Moscardini, L., et al. 1999, MNRAS, 310, 540

Ashby, M. L. N., Stern, D., Brodwin, M., et al. 2009, ApJ, 701, 428

Aussel, H., Cesarsky, C. J., Elbaz, D., & Starck, J. L. 1999, A&A, 342, 313

Avni, Y. & Bahcall, J. N. 1980, ApJ, 235, 694

Avni, Y., Soltan, A., Tananbaum, H., & Zamorani, G. 1980, ApJ, 238, 800

Babbedge, T. S. R. 2004, PhD thesis, Imperial College London

Babbedge, T. S. R., Rowan-Robinson, M., Vaccari, M., et al. 2006, MNRAS, 370, 1159

Baldry, I. K. & Glazebrook, K. 2003, ApJ, 593, 258

Bell, E. F. 2003, ApJ, 586, 794

Bell, E. F., Papovich, C., Wolf, C., et al. 2005, ApJ, 625, 23

Benítez, N. 2000, ApJ, 536, 571

Berta, S. 2005, PhD thesis, Dipartimento di Astronomia, Univ. di Padova, Vicolo dell'Osservatorio 2, I-35122, Padova, Italy

Berta, S., Fritz, J., Franceschini, A., Bressan, A., & Pernechele, C. 2003, A&A, 403, 119

Berta, S., Magnelli, B., Nordon, R., et al. 2011, A&A, 532, A49

Berta, S., Rubele, S., Franceschini, A., et al. 2006, A&A, 451, 881

Berta, S., Rubele, S., Franceschini, A., et al. 2008, VizieR Online Data Catalog, 348, 80533

Bertin, E. & Arnouts, S. 1996, A&AS, 117, 393

Bertin, E., Dennefeld, M., & Moshir, M. 1997, A&A, 323, 685

Bertoldi, F., Carilli, C., Aravena, M., et al. 2007, ApJS, 172, 132

Béthermin, M., Dole, H., Cousin, M., & Bavouzet, N. 2010, A&A, 516, A43

Béthermin, M., Le Floc'h, E., Ilbert, O., et al. 2012, A&A, 542, A58

Bicker, J. & Fritze-v. Alvensleben, U. 2005, A&A, 443, L19

Bidelman, W. P. 1988, Bulletin d'Information du Centre de Donnees Stellaires, 35, 52

Blain, A. W. & Longair, M. S. 1993, MNRAS, 264, 509

Blanton, M. R., Brinkmann, J., Csabai, I., et al. 2003, AJ, 125, 2348

Boggess, N. W. 1992, JRASC, 86, 282

Bolzonella, M., Miralles, J.-M., & Pelló, R. 2000, A&A, 363, 476

Boselli, A. 2011, A Panchromatic View of Galaxies

Bothwell, M. S., Kenicutt, R. C., Johnson, B. D., et al. 2011, MNRAS, 415, 1815

Bouchet, P., Lequeux, J., Maurice, E., Prevot, L., & Prevot-Burnichon, M. L. 1985, A&A, 149, 330

Bower, R. G., Benson, A. J., Malbon, R., et al. 2006, MNRAS, 370, 645

Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, MNRAS, 351, 1151

Brisbin, D., Harwit, M., Altieri, B., et al. 2010, MNRAS, 409, 66

Bruzual, G. & Charlot, S. 2003, MNRAS, 344, 1000

Buat, V., Giovannoli, E., Burgarella, D., et al. 2010, MNRAS, 409, L1

Budavári, T., Szalay, A. S., Connolly, A. J., Csabai, I., & Dickinson, M. 2000, AJ, 120, 1588

Burgarella, D., Heinis, S., Magdis, G., et al. 2011, ApJ, 734, L12

Calzetti, D. 1997, AJ, 113, 162

Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682

Calzetti, D., Kennicutt, R. C., Engelbracht, C. W., et al. 2007, ApJ, 666, 870

Calzetti, D., Kinney, A. L., & Storchi-Bergmann, T. 1994, ApJ, 429, 582

Caputi, K. I., Lagache, G., Yan, L., et al. 2007, ApJ, 660, 97

Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245

Cava, A., Rodighiero, G., Pérez-Fournon, I., et al. 2010, MNRAS, 409, L19

Chabrier, G. 2001, ApJ, 554, 1274

Chapman, S. C., Blain, A. W., Smail, I., & Ivison, R. J. 2005, ApJ, 622, 772

Charlot, S. & Fall, S. M. 2000, ApJ, 539, 718

Chary, R. & Elbaz, D. 2001, ApJ, 556, 562

Cimatti, A., Daddi, E., Renzini, A., et al. 2004, Nature, 430, 184

Clements, D. L., Rigby, E., Maddox, S., et al. 2010, A&A, 518, L8

Coil, A. L., Blanton, M. R., Burles, S. M., et al. 2011, ApJ, 741, 8

Coleman, G. D., Wu, C.-C., & Weedman, D. W. 1980, ApJS, 43, 393

Condon, J. J. 1974, ApJ, 188, 279

Condon, J. J. 1989, ApJ, 338, 13

Condon, J. J. 1992, ARA&A, 30, 575

Condon, J. J., Cotton, W. D., Greisen, E. W., et al. 1998, AJ, 115, 1693

Cowie, L. L., Songaila, A., Hu, E. M., & Cohen, J. G. 1996, AJ, 112, 839

Croton, D. J., Farrar, G. R., Norberg, P., et al. 2005, MNRAS, 356, 1155

Csabai, I., Budavári, T., Connolly, A. J., et al. 2003, AJ, 125, 580

Csabai, I., Dobos, L., Trencséni, M., et al. 2007, Astronomische Nachrichten, 328, 852

Cunha, C. E., Lima, M., Oyaizu, H., Frieman, J., & Lin, H. 2009, MNRAS, 396, 2379

Cutri, R. M., Skrutskie, M. F., van Dyk, S., et al. 2003, 2MASS All Sky Catalog of point sources.

Dale, D. A., Barlow, R. J., Cohen, S. A., et al. 2010, ApJ, 712, L189

Davis, M. & Huchra, J. 1982, ApJ, 254, 437

de Graauw, T., Helmich, F. P., Phillips, T. G., et al. 2010, A&A, 518, L6

Devlin, M. J., Ade, P. A. R., Aretxaga, I., et al. 2009, Nature, 458, 737

Dickinson, M., Bergeron, J., Casertano, S., et al. 2004, Spitzer Proposal, 169

Dole, H., Lagache, G., & Puget, J. 2006a, in Astronomical Society of the Pacific Conference Series, Vol. 357, Astronomical Society of the Pacific Conference Series, ed. L. Armus & W. T. Reach, 290–+

Dole, H., Lagache, G., Puget, J.-L., et al. 2006b, A&A, 451, 417

Dole, H., Rieke, G. H., Lagache, G., et al. 2004, ApJS, 154, 93

Doyle, D., Pilbratt, G., & Tauber, J. 2009, IEEE Proceedings, 97, 1403

Dye, S., Dunne, L., Eales, S., et al. 2010, A&A, 518, L10

Dye, S. & Eales, S. A. 2010, MNRAS, 405, 2406

Eadie, W. T., Drijard, D., & James, F. E. 1971, Statistical methods in experimental physics

Eales, S. 1993, ApJ, 404, 51

Eales, S., Dunne, L., Clements, D., et al. 2010a, PASP, 122, 499

Eales, S. A., Raymond, G., Roseboom, I. G., et al. 2010b, A&A, 518, L23

Efstathiou, G., Ellis, R. S., & Peterson, B. A. 1988, MNRAS, 232, 431

Egami, E., Dole, H., Huang, J.-S., et al. 2004, ApJS, 154, 130

Eggen, O. J., Lynden-Bell, D., & Sandage, A. R. 1962, ApJ, 136, 748

Eisenstein, D. J., Weinberg, D. H., Agol, E., et al. 2011, AJ, 142, 72

Elbaz, D. 2000, in Bulletin of the American Astronomical Society, Vol. 32, Bulletin of the American Astronomical Society, 1527–+

Elbaz, D., Cesarsky, C. J., Chanial, P., et al. 2002, A&A, 384, 848

Elbaz, D., Cesarsky, C. J., Fadda, D., et al. 1999, A&A, 351, L37

Elbaz, D., Dickinson, M., Hwang, H. S., et al. 2011, A&A, 533, A119

Engelbracht, C. W., Blaylock, M., Su, K. Y. L., et al. 2007, PASP, 119, 994

Faber, S. M., Willmer, C. N. A., Wolf, C., et al. 2007, ApJ, 665, 265

Fadda, D., Flores, H., Hasinger, G., et al. 2002, A&A, 383, 838

Fadda, D., Lari, C., Rodighiero, G., et al. 2004, A&A, 427, 23

Fadda, D., Yan, L., Lagache, G., et al. 2010, ApJ, 719, 425

Fazio, G. G., Hora, J. L., Allen, L. E., et al. 2004, ApJS, 154, 10

Felten, J. E. 1977, AJ, 82, 861

Fernandez-Conde, N., Lagache, G., Puget, J.-L., & Dole, H. 2008, A&A, 481, 885

Fisher, K. B., Strauss, M. A., Davis, M., Yahil, A., & Huchra, J. P. 1992, ApJ, 389, 188

Fisher, R. A. 1912, Messenger of Mathematics, 41, 155

Fitzpatrick, E. L. 1986, AJ, 92, 1068

Fixsen, D. J., Dwek, E., Mather, J. C., Bennett, C. L., & Shafer, R. A. 1998, ApJ, 508, 123

Fontanot, F., Cristiani, S., Santini, P., et al. 2012, MNRAS, 421, 241

Franceschini, A. 2000, ArXiv Astrophysics e-prints `arXiv:astro-ph/0009121`

Franceschini, A., Aussel, H., Cesarsky, C. J., Elbaz, D., & Fadda, D. 2001, A&A, 378, 1

Franceschini, A., Danese, L., de Zotti, G., & Xu, C. 1988, MNRAS, 233, 175

Franceschini, A., Mazzei, P., de Zotti, G., & Danese, L. 1994, ApJ, 427, 140

Franceschini, A., Rodighiero, G., Cassata, P., et al. 2006, A&A, 453, 397

Franceschini, A., Rodighiero, G., Vaccari, M., et al. 2010, A&A, 517, A74

Franceschini, A., Silva, L., Fasano, G., et al. 1998, ApJ, 506, 600

Frayer, D. T., Fadda, D., Yan, L., et al. 2006a, AJ, 131, 250

Frayer, D. T., Huynh, M. T., Chary, R., et al. 2006b, ApJ, 647, L9

Frayer, D. T., Sanders, D. B., Surace, J. A., et al. 2009, AJ, 138, 1261

Fukugita, M., Ichikawa, T., Gunn, J. E., et al. 1996, AJ, 111, 1748

Gallego, J., García-Dabó, C. E., Zamorano, J., Aragón-Salamanca, A., & Rego, M. 2002, ApJ, 570, L1

Gallego, J., Zamorano, J., Aragon-Salamanca, A., & Rego, M. 1995, ApJ, 455, L1

Gavazzi, G., Pierini, D., & Boselli, A. 1996, A&A, 312, 397

Gelman, A., Chew, G. L., & Shnaidman, M. 2004, Biometrics, 60, 407

Genzel, R. & Cesarsky, C. J. 2000, ARA&A, 38, 761

Genzel, R., Lutz, D., Sturm, E., et al. 1998, ApJ, 498, 579

Glenn, J., Conley, A., Béthermin, M., et al. 2010, MNRAS, 409, 109

Gonzalez, A. H., Brodwin, M., Brown, M. J. I., et al. 2010, in American Astronomical Society Meeting Abstracts, Vol. 216, American Astronomical Society Meeting Abstracts, 415.13

González-Solares, E. A., Irwin, M., McMahon, R. G., et al. 2011, MNRAS, 416, 927

Gordon, K. D., Engelbracht, C. W., Fadda, D., et al. 2007, PASP, 119, 1019

Graham, A. W. & Driver, S. P. 2005, Publications of the Astronomical Society of Australia, 22, 118

Gregg, M. D., Lacy, M., White, R. L., et al. 2002, ApJ, 564, 133

Greve, T. R., Bertoldi, F., Smail, I., et al. 2005, MNRAS, 359, 1165

Griffin, M. J., Abergel, A., Abreu, A., et al. 2010, A&A, 518, L3

Gruppioni, C., Pozzi, F., Andreani, P., et al. 2010, A&A, 518, L27

Guiderdoni, B., Bouchet, F. R., Puget, J.-L., Lagache, G., & Hivon, E. 1997, Nature, 390, 257

Gunn, J. E., Siegmund, W. A., Mannery, E. J., et al. 2006, AJ, 131, 2332

Hacking, P., Houck, J. R., & Condon, J. J. 1987, ApJ, 316, L15

Hambly, N. C., Collins, R. S., Cross, N. J. G., et al. 2008, MNRAS, 384, 637

Hanish, D. J., Meurer, G. R., Ferguson, H. C., et al. 2006, ApJ, 649, 150

Hastings, W. 1970, Biometrika, 57, 97

Hatziminaoglou, E., Omont, A., Stevens, J. A., et al. 2010, A&A, 518, L33

Hauser, M. G. 2001, in IAU Symposium, Vol. 204, The Extragalactic Infrared Background and its Cosmological Implications, ed. M. Harwit, 101–+

Hauser, M. G., Arendt, R. G., Kelsall, T., et al. 1998, ApJ, 508, 25

Hauser, M. G. & Dwek, E. 2001, ARA&A, 39, 249

Heyl, J. S. & Hernquist, L. 1997, ApJ, 491, L95

Hjort, N. L. & Jones, M. C. 1996, Ann. Statist., 24, 1619

Hogg, D. W. 1999, ArXiv Astrophysics e-prints `astro-ph/9905116`

Hogg, D. W. 2001, AJ, 121, 1207

Hogg, D. W., Cohen, J. G., Blandford, R., & Pahre, M. A. 1998, ApJ, 504, 622

Holland, W. S., Robson, E. I., Gear, W. K., et al. 1999, MNRAS, 303, 659

Hopkins, A. M. 2004, ApJ, 615, 209

Hopkins, A. M. & Beacom, J. F. 2006, ApJ, 651, 142

Houck, J. R., Roellig, T. L., van Cleve, J., et al. 2004, ApJS, 154, 18

Hubble, E. P. 1926, ApJ, 64, 321

Huchra, J. P. 1977, ApJS, 35, 171

Hughes, D. H., Serjeant, S., Dunlop, J., et al. 1998, Nature, 394, 241

Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, A&A, 457, 841

Ilbert, O., Capak, P., Salvato, M., et al. 2009, ApJ, 690, 1236

Ilbert, O., Salvato, M., Le Floc'h, E., et al. 2010, ApJ, 709, 644

Irwin, M. & Lewis, J. 2001, New Astronomy Review, 45, 105

Ivezić, Ž., Lupton, R. H., Schlegel, D., et al. 2004, Astronomische Nachrichten, 325, 583

Ivezić, Ž., Smith, J. A., Miknaitis, G., et al. 2007, AJ, 134, 973

Ivison, R. J., Greve, T. R., Dunlop, J. S., et al. 2007, MNRAS, 380, 199

Jannuzi, B. T., Dey, A., Brown, M. J. I., et al. 2004, in Bulletin of the American Astronomical Society, Vol. 36, American Astronomical Society Meeting Abstracts, 1478

Johnson, H. L. 1962, ApJ, 135, 975

Johnston, R. 2011, A&A Rev., 19, 41

Kafka, P. 1967, Nature, 213, 346

Kelly, B. C., Fan, X., & Vestergaard, M. 2008, ApJ, 682, 874

Kennicutt, Jr., R. C. 1983, ApJ, 272, 54

Kennicutt, Jr., R. C. 1998, ApJ, 498, 541

Kessler, M. F., Steinz, J. A., Anderegg, M. E., et al. 1996, A&A, 315, L27

Koranyi, D. M. & Strauss, M. A. 1997, ApJ, 477, 36

Kormendy, J. & Sanders, D. B. 1992, ApJ, 390, L53

Kron, R. G. 1980, ApJS, 43, 305

Kroupa, P. 1998, MNRAS, 298, 231

Kroupa, P. 2001, MNRAS, 322, 231

Lacy, M., Storrie-Lombardi, L. J., Sajina, A., et al. 2004, ApJS, 154, 166

Lacy, M., Wilson, G., Masci, F., et al. 2005, ApJS, 161, 41

Lagache, G., Dole, H., & Puget, J.-L. 2003, MNRAS, 338, 555

Lagache, G., Dole, H., Puget, J.-L., et al. 2004, ApJS, 154, 112

Lagache, G., Puget, J.-L., & Dole, H. 2005, ARA&A, 43, 727

Lagage, P. O., Claret, A., Ballet, J., et al. 1996, A&A, 315, L273

Lapi, A., González-Nuevo, J., Fan, L., et al. 2011, ApJ, 742, 24

Larson, R. B. 1998, MNRAS, 301, 569

Larson, R. B. & Tinsley, B. M. 1978, ApJ, 219, 46

Le Borgne, D., Elbaz, D., Ocvirk, P., & Pichon, C. 2009, A&A, 504, 727

Le Floc'h, E., Aussel, H., Ilbert, O., et al. 2009, ApJ, 703, 222

Le Floc'h, E., Papovich, C., Dole, H., et al. 2005, ApJ, 632, 169

Leger, A. & Puget, J. L. 1984, A&A, 137, L5

Levenson, L., Marsden, G., Zemcov, M., et al. 2010, MNRAS, 409, 83

Liddle, A. R. 2009, Annual Review of Nuclear and Particle Science, 59, 95

Lilly, S. J., Le Fèvre, O., Renzini, A., et al. 2007, ApJS, 172, 70

Loader, C. R. 1996, Ann. Statist., 24, 1602

Longair, M. S. 2008, Galaxy Formation

Lonsdale, C., Polletta, M. d. C., Surace, J., et al. 2004, ApJS, 154, 54

Lonsdale, C. J., Hacking, P. B., Conrow, T. P., & Rowan-Robinson, M. 1990, ApJ, 358, 60

Lonsdale, C. J., Smith, H. E., Rowan-Robinson, M., et al. 2003, PASP, 115, 897

Lupton, R. H., Gunn, J. E., & Szalay, A. S. 1999, AJ, 118, 1406

Lupton, R. H., Ivezic, Z., Gunn, J. E., et al. 2002, in Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, Vol. 4836, Survey and Other Telescope Technologies and Discoveries. Edited by Tyson, J. Anthony; Wolff, Sidney. Proceedings of the SPIE, Volume 4836, pp. 350-356 (2002)., ed. J. A. Tyson & S. Wolff, 350–356

Lutz, D., Poglitsch, A., Altieri, B., et al. 2011, A&A, 532, A90

Ly, C., Malkan, M. A., Kashikawa, N., et al. 2007, ApJ, 657, 738

Magnelli, B., Elbaz, D., Chary, R. R., et al. 2009, A&A, 496, 57

Mainzer, A. K., Eisenhardt, P., Wright, E. L., et al. 2005, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 5899, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, ed. H. A. MacEwen, 262–273

Mandelbaum, R., Seljak, U., Hirata, C. M., et al. 2008, MNRAS, 386, 781

Marleau, F. R., Fadda, D., Appleton, P. N., et al. 2007, ApJ, 663, 218

Marleau, F. R., Fadda, D., Storrie-Lombardi, L. J., et al. 2004, ApJS, 154, 66

Martin, D. C., Fanson, J., Schiminovich, D., et al. 2005a, ApJ, 619, L1

Martin, D. C., Seibert, M., Buat, V., et al. 2005b, ApJ, 619, L59

Mathis, J. S., Rumpl, W., & Nordsieck, K. H. 1977, ApJ, 217, 425

Mauduit, J.-C., Lacy, M., Farrah, D., et al. 2012, ArXiv e-prints

McMahon, R. G., Walton, N. A., Irwin, M. J., et al. 2001, New Astronomy Reviews, 45, 97

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. 1953, Journal of Chemical Physics, 21, 1087

Mobasher, B., Dickinson, M., Ferguson, H. C., et al. 2005, ApJ, 635, 832

Monaco, P., Fontanot, F., & Taffoni, G. 2007, MNRAS, 375, 1189

Moshir, M., Kopman, G., & Conrow, T. 1992, IRAS Faint Source Survey, Explanatory supplement version 2 (Pasadena: Infrared Processing and Analysis Center, California Institute of Technology, 1992, edited by Moshir, M.; Kopman, G.; Conrow, T. et al.)

Moster, B. P., Somerville, R. S., Newman, J. A., & Rix, H.-W. 2011, ApJ, 731, 113

Negrello, M., Perrotta, F., González-Nuevo, J., et al. 2007, MNRAS, 377, 1557

Nguyen, H. T., Schulz, B., Levenson, L., et al. 2010, A&A, 518, L5

Norberg, P., Cole, S., Baugh, C. M., et al. 2002, MNRAS, 336, 907

Oliver, S., Mann, R. G., Carballo, R., et al. 2002, MNRAS, 332, 536

Oliver, S., Rowan-Robinson, M., Alexander, D. M., et al. 2000, MNRAS, 316, 749

Oliver, S. J., Bock, J., Altieri, B., et al. 2012, ArXiv e-prints

Oliver, S. J., Rowan-Robinson, M., & Saunders, W. 1992, MNRAS, 256, 15P

Oliver, S. J., Wang, L., Smith, A. J., et al. 2010, A&A, 518, L21

Oyaizu, H., Lima, M., Cunha, C. E., et al. 2008, ApJ, 674, 768

Padmanabhan, N., Schlegel, D. J., Finkbeiner, D. P., et al. 2008, ApJ, 674, 1217

Page, M. J. & Carrera, F. J. 2000, MNRAS, 311, 433

Papovich, C., Dole, H., Egami, E., et al. 2004, ApJS, 154, 70

Papovich, C., Rudnick, G., Le Floc'h, E., et al. 2007, ApJ, 668, 45

Partridge, R. B. & Peebles, P. J. E. 1967, ApJ, 148, 377

Pascale, E., Ade, P. A. R., Bock, J. J., et al. 2009, ApJ, 707, 1740

Patanchon, G., Ade, P. A. R., Bock, J. J., et al. 2009, ApJ, 707, 1750

Patel, H., Clements, D. L., Vaccari, M., et al. 2012, ArXiv e-prints

Peacock, J. A. & Gull, S. F. 1981, MNRAS, 196, 611

Pearson, C. & Rowan-Robinson, M. 1996, MNRAS, 283, 174

Peebles, P. J. E. 1993, Principles of physical cosmology (Princeton Series in Physics, Princeton, NJ: Princeton University Press, 1993)

Pérez-González, P. G., Gallego, J., Zamorano, J., et al. 2003, ApJ, 587, L27

Pérez-González, P. G., Rieke, G. H., Egami, E., et al. 2005a, ApJ, 630, 82

Pérez-González, P. G., Rieke, G. H., Egami, E., et al. 2005b, ApJ, 630, 82

Petrosian, V. 1976, ApJ, 209, L1

Pier, J. R., Munn, J. A., Hindsley, R. B., et al. 2003, AJ, 125, 1559

Pilbratt, G. L., Riedinger, J. R., Passvogel, T., et al. 2010, A&A, 518, L1

Poggianti, B. M. & Wu, H. 2000, ApJ, 529, 157

Poglitsch, A., Waelkens, C., Geis, N., et al. 2010, A&A, 518, L2

Polletta, M., Tajer, M., Maraschi, L., et al. 2007, ApJ, 663, 81

Pozzi, F., Gruppioni, C., Oliver, S., et al. 2004, ApJ, 609, 122

Press, W. H. & Schechter, P. 1974, ApJ, 187, 425

Prevot, M. L., Lequeux, J., Prevot, L., Maurice, E., & Rocca-Volmerange, B. 1984, A&A, 132, 389

Puget, J.-L., Abergel, A., Bernard, J.-P., et al. 1996, A&A, 308, L5+

Reddy, N. A., Steidel, C. C., Pettini, M., et al. 2008, ApJS, 175, 48

Renzini, A. 2006, ARA&A, 44, 141

Richards, G. T., Hall, P. B., Vanden Berk, D. E., et al. 2002, in Bulletin of the American Astronomical Society, Vol. 34, Bulletin of the American Astronomical Society, 1309–+

Richards, G. T., Strauss, M. A., Fan, X., et al. 2006, AJ, 131, 2766

Rieke, G. H., Alonso-Herrero, A., Weiner, B. J., et al. 2009, ApJ, 692, 556

Rieke, G. H., Young, E. T., Cadien, J., et al. 2004, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 5487, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, ed. J. C. Mather, 50–61

Rigopoulou, D., Magdis, G., Ivison, R. J., et al. 2010, MNRAS, 409, L7

Roche, N. & Eales, S. A. 1999, MNRAS, 307, 111

Rodighiero, G., Cimatti, A., Franceschini, A., et al. 2007, A&A, 470, 21

Rodighiero, G., Vaccari, M., Franceschini, A., et al. 2010, A&A, 515, A8

Roseboom, I. G., Oliver, S., Parkinson, D., & Vaccari, M. 2009, MNRAS, 400, 1062

Roseboom, I. G., Oliver, S. J., Kunz, M., et al. 2010, MNRAS, 409, 48

Rowan-Robinson, M. 2001, ApJ, 549, 745

Rowan-Robinson, M. 2009, MNRAS, 394, 117

Rowan-Robinson, M., Babbedge, T., Oliver, S., et al. 2008, MNRAS, 386, 697

Rowan-Robinson, M., Babbedge, T., Surace, J., et al. 2005, AJ, 129, 1183

Rowan-Robinson, M., Hughes, J., Vedi, K., & Walker, D. W. 1990, MNRAS, 246, 273

Rowan-Robinson, M., Lari, C., Perez-Fournon, I., et al. 2004, MNRAS, 351, 1290

Rowan-Robinson, M., Roseboom, I. G., Vaccari, M., et al. 2010, MNRAS, 409, 2

Rush, B. & Malkan, M. A. 1993, in Bulletin of the American Astronomical Society, Vol. 25, Bulletin of the American Astronomical Society, 1362–+

Sacchi, N., La Franca, F., Feruglio, C., et al. 2009, ApJ, 703, 1778

Sajina, A., Scott, D., Dennefeld, M., et al. 2006, MNRAS, 369, 939

Salpeter, E. E. 1955, ApJ, 121, 161

Sandage, A., Tammann, G. A., & Yahil, A. 1979, ApJ, 232, 352

Sanders, D. B. 2003, Journal of Korean Astronomical Society, 36, 149

Sanders, D. B. & Mirabel, I. F. 1996, ARA&A, 34, 749

Santini, P., Fontana, A., Grazian, A., et al. 2009, A&A, 504, 751

Saunders, W., Rowan-Robinson, M., Lawrence, A., et al. 1990, MNRAS, 242, 318

Saunders, W., Sutherland, W. J., Maddox, S. J., et al. 2000, MNRAS, 317, 55

Scalo, J. 1998, in Astronomical Society of the Pacific Conference Series, Vol. 142, The Stellar Initial Mass Function (38th Herstmonceux Conference), ed. G. Gilmore & D. Howell, 201

Scalo, J. M. 1986, Fundamentals of Cosmic Physics, 11, 1

Schafer, C. M. 2007, ApJ, 661, 703

Schechter, P. 1976, ApJ, 203, 297

Scheuer, P. A. G. 1957, Proceedings of the Cambridge Philosophical Society, 53, 764

Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525

Schmidt, M. 1968, ApJ, 151, 393

Schmidt, M. & Green, R. F. 1983, ApJ, 269, 352

Scott, K. S., Stabenau, H. F., Braglia, F. G., et al. 2010, ApJS, 191, 212

Scott, S. E., Fox, M. J., Dunlop, J. S., et al. 2002, MNRAS, 331, 817

Scoville, N., Aussel, H., Brusa, M., et al. 2007, ApJS, 172, 1

Seaton, M. J. 1979, MNRAS, 187, 73P

Serjeant, S., Gruppioni, C., & Oliver, S. 2002, MNRAS, 330, 621

Serjeant, S. & Harrison, D. 2005, MNRAS, 356, 192

Serjeant, S., Oliver, S., Rowan-Robinson, M., et al. 2000, MNRAS, 316, 768

Shimasaku, K., Ouchi, M., Furusawa, H., et al. 2005, PASJ, 57, 447

Shupe, D. L., Rowan-Robinson, M., Lonsdale, C. J., et al. 2008, AJ, 135, 1050

Silva, L., Granato, G. L., Bressan, A., & Danese, L. 1998, ApJ, 509, 103

Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163

Smail, I., Ivison, R. J., & Blain, A. W. 1997, ApJ, 490, L5

Smith, A. J., Wang, L., Oliver, S. J., et al. 2012, MNRAS, 419, 377

Smith, J. A., Tucker, D. L., Allam, S. S., & Jorgensen, A. M. 2002, in Bulletin of the American Astronomical Society, Vol. 34, Bulletin of the American Astronomical Society, 1272–+

Soifer, B. T., Helou, G., & Werner, M. 2008, ARA&A

Soifer, B. T. & Neugebauer, G. 1991, AJ, 101, 354

Soifer, B. T., Sanders, D. B., Madore, B. F., et al. 1987, ApJ, 320, 238

Somerville, R. S., Gilmore, R. C., Primack, J. R., & Domínguez, A. 2012, MNRAS, 423, 1992

Springel, V. & White, S. D. M. 1998, MNRAS, 298, 143

Stansberry, J. A., Gordon, K. D., Bhattacharya, B., et al. 2007, PASP, 119, 1038

Starikova, S., Berta, S., Franceschini, A., et al. 2012, ApJ, 751, 126

Stern, D., Eisenhardt, P., Gorjian, V., et al. 2005, ApJ, 631, 163

Strauss, M. A. 2002, in Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, Vol. 4836, Survey and Other Telescope Technologies and Discoveries. Edited by Tyson, J. Anthony; Wolff, Sidney. Proceedings of the SPIE, Volume 4836, pp. 1-9 (2002)., ed. J. A. Tyson & S. Wolff, 1–9

Strauss, M. A. & Willick, J. A. 1995, Physics Reports, 261, 271

SubbaRao, M., Frieman, J., Bernardi, M., et al. 2002, in Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, Vol. 4847, Astronomical Data Analysis II. Edited by Starck, Jean-Luc; Murtagh, Fionn D. Proceedings of the SPIE, Volume 4847, pp. 452-460 (2002)., ed. J.-L. Starck & F. D. Murtagh, 452–460

Sullivan, M., Treyer, M. A., Ellis, R. S., et al. 2000, MNRAS, 312, 442

Surace, J. A., Shupe, D. L., Fang, F., et al. 2005, in Bulletin of the American Astronomical Society, Vol. 37, Bulletin of the American Astronomical Society, 1246–+

Sutherland, W. & Saunders, W. 1992, MNRAS, 259, 413

Takeuchi, T. T., Ishii, T. T., Dole, H., et al. 2006, A&A, 448, 525

Tielens, A. G. G. M. 2008, ARA&A, 46, 289

Treffers, R. & Cohen, M. 1974, ApJ, 188, 545

Tresse, L. & Maddox, S. J. 1998, ApJ, 495, 691

Trichas, M., Georgakakis, A., Rowan-Robinson, M., et al. 2009, MNRAS, 399, 663

Tucker, D. L., Annis, J. T., Lin, H., et al. 2006, ArXiv Astrophysics e-prints `astro-ph/0611137`

Vaccari, M., Lari, C., Angeretti, L., et al. 2005, MNRAS, 358, 397

Vaccari, M., Marchetti, L., Franceschini, A., et al. 2010, A&A, 518, L20

Valiante, E., Lutz, D., Sturm, E., Genzel, R., & Chapin, E. L. 2009, ApJ, 701, 1814

Voges, W., Aschenbach, B., Boller, T., et al. 1999, A&A, 349, 389

Walcher, J., Groves, B., Budavári, T., & Dale, D. 2011, Ap&SS, 331, 1

Walmsley, C. M., Bertout, C., Combes, F., et al. 2010, A&A, 518, 1

Wang, J., De Lucia, G., Kitzbichler, M. G., & White, S. D. M. 2008, MNRAS, 384, 1301

Weinberg, S. 1972, Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity (Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity, by Steven Weinberg, pp. 688. ISBN 0-471-92567-5. Wiley-VCH , July 1972.)

Werner, M., Fazio, G., Rieke, G., Roellig, T. L., & Watson, D. M. 2006, ARA&A, 44, 269

Werner, M. W., Roellig, T. L., Low, F. J., et al. 2004, ApJS, 154, 1

Westra, E., Geller, M. J., Kurtz, M. J., Fabricant, D. G., & Dell'Antonio, I. 2010, ApJ, 708, 534

Williams, R. E., Blacker, B., Dickinson, M., et al. 1996, AJ, 112, 1335

Wright, E. L. 2004, New Astronomy Review, 48, 465

Yahil, A., Strauss, M. A., Davis, M., & Huchra, J. P. 1991, ApJ, 381, 348

York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, AJ, 120, 1579