Physical Properties of Massive, Star-Forming Galaxies When the Universe Was Only Two Billion Years Old

by

Nicole Christina Fu, B.Sc.

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
In partial fulfillment of the requirements
For the MSc degree in Physics

Department of Physics
Faculty of Science
University of Ottawa

The undersigned recommend to
the Faculty of Graduate and Postdoctoral Studies
acceptance of the thesis

Physical Properties of Massive, Star-Forming Galaxies When the Universe Was Only Two Billion Years Old

submitted by **Nicole Christina Fu, B.Sc.**
in partial fulfillment of the requirements for
the MSc degree in Physics

Chair, Department of Physics

Thesis Supervisor

External Examiner

University of Ottawa

Date _____

Abstract

Due to the finite speed of light and a vast, expanding universe, telescopes are just now receiving the light emitted by galaxies as they were forming in the very early universe. The light from these galaxies has been redshifted (stretched to longer, redder wavelengths) as a result of its journey through expanding space. Using sophisticated techniques and exceptional multi-wavelength optical and infrared data, we isolate a population of 378 galaxies in the process of formation when the Universe was only two billion years old. By matching the distinctive properties of the light spectra of these galaxies to models, the redshift, age, dust content, star formation rate and total stellar mass of each galaxy are determined. Comparing our results to similar surveys of galaxy populations at other redshifts, a picture emerges of the growth and evolution of massive, star-forming galaxies over the course of billions of years.

For Elisabeth Shawn

Acknowledgments

I would like to thank the University of Ottawa for the graduate entrance scholarship and exceptional funding that made possible my pursuit of this research. The field of astrophysics lies outside the mainstream of physics research at the University of Ottawa, and I am grateful for the willingness of the department to allow me to engage in this non-traditional stream of study. I am indebted to Dr. Chris Willott, for providing me with an inherently beautiful and challenging project, for sharing his time, knowledge and expertise, and for continuing to supervise this project long after moving ahead in his career and taking on full-time research with the CADC. I owe a debt of gratitude to Dr. Bela Joos, Dr. Ivan L'Heureux, Dr. Peter Watson, Dr. Thomas Brabec, Dr. Richard Hodgson and Dr. Darren Grant for their support of my work. The cheerful assistance offered to me by Madeleine Thomas and my office mates, and my interactions with professors, classmates and students, at both the University of Ottawa and Carleton University, have made my time here rewarding and enjoyable.

The completion of this thesis has been a long and difficult undertaking, for which much has been sacrificed. I have not borne this alone; my lovely daughter Maeghan has shared in the loss of simple pleasures: time together watching movies, home cooked meals, working plumbing, walks with Emmett and Hank. Through it all, I have been deeply touched by her unwavering focus on the big picture, by her insight

and encouragement when I faltered, and by the grace with which she embraced the financial hardships. Her fierce love and support have seen me through to the end.

The encouragement and generosity of my family and friends has meant a great deal to me and I am truly thankful. My longtime friends, Anita, Carol, Eliza and Peter, have offered me warm hospitality and diversion, and remained faithful despite my neglect of our friendship in the name of the thesis. For the love and friendship of my sister Karen, my Indiana Jones, I am always grateful... even though she has a home in Hawaii and has been to the observatories while I have not! In opening her heart, she has opened my world. My friend Darren has been a constant in my life, providing not only his friendship in many a wee hour, but his physics expertise and his enthusiasm for experimental research. His generosity and kindness have made my days beautiful and rescued me in many a desperate hour. We have big plans...!

Finally and foremostly, this work would not have been possible without the steadfast love, encouragement and financial support of my Mum, Elisabeth, to whom this thesis is dedicated. For this and so much more, I shall always be grateful.

Statement of Originality

This thesis comprises the analysis of existing, publicly available data from the astronomical community. The contributions of the author extend to all aspects of the presented analysis and were done independently, although some of the software tools and methods used were developed by others working in the field. The author did not have a part in the image or data collection.

Project Data The original data for the thesis consists of:

- optical images obtained as part of the Canada-France-Hawaii Telescope Legacy Survey (CFHTLS) and the associated data available in the 2007 CFHTLS Terapix T0004 catalogues, including object selection and stellarity values;
- mid-infrared Spitzer Space Telescope Infrared Array Camera (IRAC) images, data and photometric catalogues obtained as part of the Spitzer Legacy Cosmic Evolution Survey (S-COSMOS);
- near-infrared K band images and data obtained as part of the CFHT Wide-field InfraRed Camera (WIRCam) Deep Survey (WIRDS) and the Sanders-led PI program to map COSMOS with WIRCam;

- spectroscopic data obtained as part of the VIsible Multi-Object Spectrograph (VIMOS) Very Large Telescope (VLT) Deep Survey (VVDS), available in the VVDS-DEEP V1.0 catalogue.

Data Calibration For the purposes of this project, catalogues of 2" diameter aperture magnitudes and aperture correction values for objects in the optical and K band images were provided by Chris Willott. The routine APER, an adaptation of the Dominion Astrophysical Observatory Photometry program DAOPHOT, was used to compute mid-infrared aperture photometry from the Spitzer IRAC images. Standard S-COSMOS aperture correction values were applied to the routine-generated IRAC magnitudes. Spitzer IRAC and K band limiting magnitudes were based on data generated with the APER routine. Galactic dust extinction and reddening of the optical data were corrected using the dust maps by Schlegel, Finkbeiner and Davis (1998). Galactic extinction correction factors for the optical filters were provided by Chris Willott.

The author was responsible for compilation of the ten-band photometric data and image set, configuration and generation of concentric aperture photometry, aperture photometry corrections, the development of the routine to apply dust map data to correct for galactic extinction, and determination of limiting magnitudes and associated uncertainties. The blended source analysis for the elimination of targets with compromised photometry was designed, developed and implemented by the author.

Isolation of Galaxies The high redshift galaxy sample was isolated by the author. Steps involved the matching of photometric and spectroscopic data from different surveys, colour-redshift and colour-colour analysis of target objects, colour corrections, establishment of optical selection criteria, and the analysis of selection

efficiency using infrared colours and galaxy model tracks by Charlot and Bruzual (CB2008).

Model Fitting An initial algorithm for the model fitting procedure was provided by Chris Willott. Fits were based on: CB2008 galaxy model spectra; a quasar composite spectrum produced by Daniel Fryer and Chris Willott and based on Sloan Digital Sky Survey (SDSS) data; and star spectra from the Model Atmosphere Radiative and Convective Scheme (MARCS) library. Spectra were customized for dust extinction using the method of CB2008 and corrected for Lyman forest absorption using the transmission curves determined for the Canada-France High-z Quasar Survey (CFHQS) by Willott *et al.* (2005) based on original data and equations by Songaila (2004).

The author was responsible for the compilation and customization of model spectra, and the development of the model fitting code. Steps included: generation of dust extinction curves and determination of the mean Lyman forest transmission spectrum specific to each model's redshift; correction of models for Lyman forest absorption and dust extinction; generation of $\sim$ 36,500 spectral templates; determination of photometric and model flux densities; fit generation based on a floating normalization routine and minimization of the reduced χ^2 statistic; calculation of star formation rates and stellar masses based on fit normalization values; analysis of model fits; identification and removal of interlopers; and plotting of model fit spectra.

Data Set Refinement Cross-correlation analysis of target objects was performed by Chris Willott using complementary data collected by: the Hubble Space Telescope (HST) Treasury Project COSMOS, the spectroscopic redshift survey zCOSMOS, the X-ray Multi-Mirror Mission XMM-COSMOS, and the Spitzer Legacy project

S-COSMOS. The author was responsible for removal of misidentified objects, point sources and poor model fits; visual inspection and elimination of damaged images; and selection of the final data set.

Ensemble Properties The ensemble properties of the galaxy population were determined based on the technique described by Papovich *et al.* (2001) and Verma *et al.* (2007). The ensemble test was performed by the author and included analysis of model fitting issues, $\Delta\chi^2$ mapping of the multi-dimensional parameter space of each galaxy, Monte Carlo generation of synthetic galaxy data, probability transformations based on synthetic fits, determination of individual and composite probability distributions and confidence intervals, and the analysis and description of ensemble properties.

Results Interpretation of the results and formulation of an evolutionary picture by the author would not have been possible without the exemplary research established in previous studies: Papovich *et al.* (2001), Shapley *et al.* (2001), Eyles *et al.* (2007), Verma *et al.* (2007), Yabe *et al.* (2008), and Stark *et al.* (2009).

Final Notes A number of IDL routines, publicly available through the Goddard Space Flight Center IDL Astronomy User's Library, were applied in the analysis of the photometric data. The author gratefully acknowledges this contribution from the astronomical community of IDL programmers. The use of other's work is cited in the text where appropriate. Any omissions or errors in the thesis are solely the responsibility of the author.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Modern cosmology is focused on one goal: to understand the beginnings of the Universe. A foundational field of knowledge contributing to this understanding is the formation and evolution of galaxies [19] [30]. To date, very little is known about the primary constituents of the Universe, let alone the physical mechanisms by which large-scale structures such as galaxies and galaxy clusters are formed. The formation of these structures has progressed from the genesis of the very first stars to the assembly of the billions of galaxies that populate the Universe today. If we are to understand the physics at every scale, star-forming galaxies must be studied at their beginnings.

Standard cosmological theory is based on the currently favoured ΛCDM model [11] [20]. The principal constituents of this model are: a relatively unknown component called dark energy, or the cosmological constant (Λ), and cold dark matter (CDM), which refers to low-velocity non-baryonic matter that may interact gravitationally on a variety of scales. Baryonic matter, photons and neutrinos comprise minor constituents. The parameters of ΛCDM are derived from measurements in observational astrophysics, including results from the Wilkinson Microwave Anisotropy

Probe (WMAP) experiment [58]. A global fit to cosmological data establishes the current energy density proportions of the components at 4% baryons, 24% dark matter and 72% dark energy, acting against gravity.

Within the ΛCDM model, minute density fluctuations in the matter distribution of the early universe provide the seeds of formation of the first structures [11] [19] [30]. These fluctuations cause gravitational instabilities that lead to their collapse into spheroidal dark matter halos. The halos act as gravitational wells and accrete increasing amounts of dark matter. In addition, baryonic matter trapped in the wells collects and cools, forming the first generation of stars. As dark matter structures increase in scale, stellar populations form. Over time, gravitational merging assembles these stellar populations into galaxies. Formation according to this model is referred to as the hierarchical bottom-up scenario.

Observational evidence indicates that, as a result of hierarchical processes, some massive galaxies have formed over billions of years [12] [35] [53]. However, the discovery of massive, evolved galaxies in existence in the early universe is inconsistent with the hierarchical scenario. Instead, it appears that galaxy assembly can reach a massive threshold as early as a billion years after the Big Bang.

The source of this puzzle, and its answer, seem to lie in our understanding of the physics of baryonic matter. Unlike the dark matter discussed above, which interacts only through gravity, the physics of baryonic matter is difficult to model due to the need to consider self-interactions [19] [30] [37]. Galaxy-scale computations are necessary and must deal with the simultaneous interaction of factors such as gas flow, dust content, magnetic fields, chemical enrichment, black hole formation, galaxy mergers, and evolutionary timescales. Within such an environment, it is difficult to reliably predict the effect that evolving parameters will have on the activity of star formation. In view of such complexity, any advancement of theories of galaxy evolution will rely on observational studies of primordial galaxies at high redshift.

The universe of high redshift galaxies became accessible in the 1990s with the development of the Lyman-break technique [19] [34] [57]. The Lyman-break technique searches for a discontinuity in the rest-frame ultra-violet (UV) region of a galaxy's spectral energy distribution (SED). The discontinuity is caused by an attenuation of flux at wavelengths shortward of the 912 Å Lyman limit. UV photons at these wavelengths are absorbed in the process of photoionization of hydrogen in stellar atmospheres and in the interstellar medium (ISM). The result is a distinct spectral break. For galaxies at redshifts $z \gtrsim 2$ (see Equation 2.1), the Lyman-break is redshifted into the optical band. The position of the Lyman-break in the observed-frame is used to determine redshift. High redshift galaxies identified by this method are known as Lyman-break galaxies (LBGs).

Rest-frame UV emission in galaxies is indicative of active star formation [19]. The Lyman-break technique is therefore an effective method for locating starburst galaxies at high redshifts. While galaxies of this type had previously been identified up to $z \sim 1$, the application of the Lyman-break technique brought within reach young star-forming galaxies with redshifts up to $z \sim 4$.

Although the Lyman-break technique lacks the precision of spectroscopically-determined redshifts, it efficiently generates large samples of galaxies within broad redshift ranges [19]. From these initial ranges, we can improve the redshift estimate as well as determine other defining parameters of the LBG such as the mix of stellar populations, star formation rate (SFR), total stellar mass, age and dust content. This refinement is accomplished by comparing LBG photometry to theoretical models of galaxy spectra. LBGs are today being identified with redshifts reaching $z \sim 8$ [6], thus providing a glimpse into the very early universe.

The continued advancement of telescope and detector technologies has improved the search for LBGs [17] [30] [53]. These advances are: broader wavelength coverage, deeper imaging for the detection of fainter objects, and larger redshift surveys for

reduced cosmic variance. With regard to broader wavelength coverage in particular, ground-based observations are being complemented by imaging from space-based instruments. The result is exceptional multi-wavelength data sets. Data in the near-infrared is especially important because it represents the rest-frame optical emission of LBGs [34]. By extending optical SEDs into the near-infrared, it is possible to place tighter constraints on the aforementioned LBG parameters.

Much of our view of galaxy formation in the early universe is based on studies of LBGs [17] [30] [53]. The LBG population is composed of the earliest star-forming galaxies with masses on the order of 10^{10} solar masses ($M_{\odot}$) by $z \sim 3$. An evolutionary link is suggested between LBGs and the ancient, massive galaxies currently in existence. By observing LBGs over a complete range of redshifts, we can elucidate the physical mechanisms and timescales responsible for their formation, growth and diversity.

This thesis details the isolation of a population of $z \sim 4$ LBGs and the identification of the properties of this population. In Chapter 2, the reader will find a discussion of preliminary astrophysical concepts, including the Lyman-break technique. A description of the data used in this project, including sources, photometric methods and corrections, is covered in Chapter 3. Chapter 4 outlines the selection of a population of LBGs at $z \sim 4$ based on the Lyman-break technique. Chapter 5 familiarizes the reader with galaxy modeling by stellar population synthesis. Also in Chapter 5 are the details of the procedure to fit target object photometry to model galaxy spectra. Chapter 6 presents the refinement of the LBG population including elimination of: contaminating objects, compromised data, and inadequate model fits. As well, Chapter 6 describes the details of the ensemble test used to generate composite probability distributions of the parameters of the LBG population as a whole. Finally, in Chapter 7, results are compared with other studies of LBG populations at different redshifts. The goal is a clear picture of the defining properties

of LBGs as they evolve in the high redshift universe.

Chapter 2

Astrophysical Concepts

2.1 Cosmological Redshift

We live in an expanding universe, a relict effect of the Big Bang [11] [38] [59]. The wavelength of the light emitted from cosmically distant sources is lengthened over the course of its journey due to the expansion of space itself. This effect is known as cosmological redshift. To say that a galaxy is redshifted is to describe the lengthening in the rest-frame wavelength of the light emitted by that galaxy. Redshift is defined by the parameter z:

$$z = \frac{\lambda_{obs} - \lambda_{rest}}{\lambda_{rest}} = \frac{\Delta\lambda}{\lambda_{rest}}, \tag{2.1}$$

where λ_{obs} is the wavelength of the radiation detected by an observer, and λ_{rest} is the wavelength of the radiation at the time of its emission (i.e. the rest-frame wavelength). The rate of expansion of the Universe has changed over time and, therefore, determining actual distances to galaxies is not straightforward. This has resulted in several different measurement scales that diverge with increasing distance [60]. Redshift, a measure of this expansion, provides a consistent alternative and thus statements of galaxy redshift have become standard.

Rearranging Equation 2.1, we may define the fractional change in wavelength as a function of redshift (z) [11]:

$$\frac{\lambda_{obs}}{\lambda_{rest}} = (1 + z). \tag{2.2}$$

where λ_{obs} is the observed wavelength of the radiation, and λ_{rest} is the rest-frame wavelength. Thus, the observed wavelength of light emitted by a galaxy at $z = 3$, for example, will be detected at a wavelength that is four times the rest-frame wavelength. This increase in wavelength provides a direct measure of the amount by which space has expanded over the same period of time. Correspondingly, since the light left the galaxy at $z = 3$, the Universe has expanded by a factor of four.

As well as distance, redshift gives us a measure of time [11]. A plot of the age of the Universe as a function of redshift is shown in Figure 2.1. Standard cosmological parameters, Hubble constant $H_0 = 70$ km s^{-1} Mpc^{-1}, matter density $\Omega_m = 0.28$, dark energy density $\Omega_\Lambda = 0.72$, curvature density $\Omega_k = 0.0$ (flat) and deceleration $q_0 = -0.58$ [58], were used in the generation of the redshift-cosmic age relationship, and have been assumed throughout this thesis. The WMAP results estimate the current age of the Universe to be 13.75 $\pm$ 0.13 billion years [58]. From Figure 2.1, a galaxy detected at $z = 3$ appears as it did when the Universe was just 2.2 billion years old ($\sim$ 15% its current age) and progressively higher redshifts indicate even earlier states of the Universe.

2.2 Photometric Redshift Technique

Galaxy redshifts may be determined spectroscopically and photometrically. Within the light spectrum of a distant galaxy, the shift of absorption or emission lines of known rest-frame wavelength reveals the galaxy's redshift spectroscopically [11]. Obtaining spectroscopic redshifts is an ideal scenario. However, the faintness of high

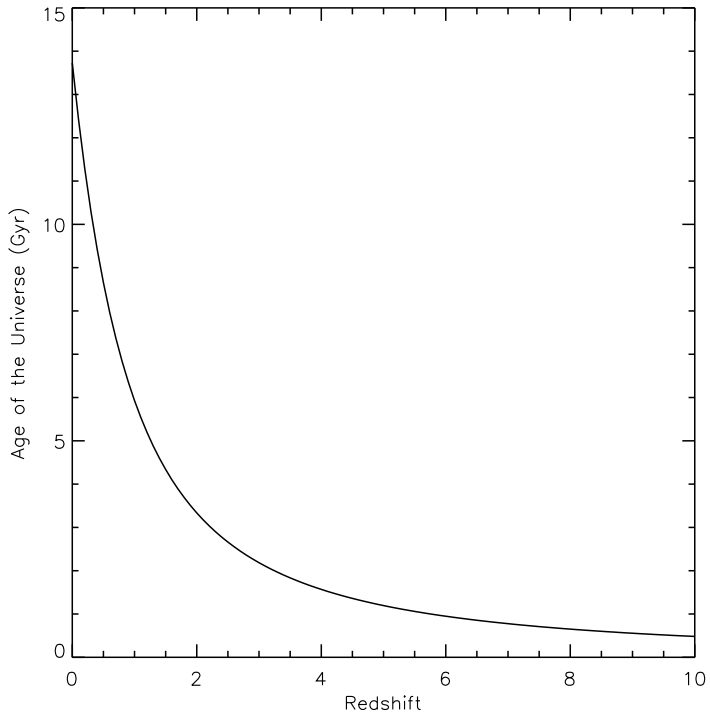

Figure 2.1: A plot of the age of the Universe as a function of redshift. Cosmological parameters used: $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_m = 0.28$, $\Omega_\Lambda = 0.72$, $\Omega_k = 0.0$ (flat), q_0 = -0.58 [58] [61].

redshift galaxies and the hundreds of galaxies currently being surveyed limit the usefulness of this approach [31]. Photometric redshifts provide a practical alternative.

To determine redshift photometrically, a galaxy is imaged in a series of broad-band filters [24] [31]. The galaxy's light emission, or flux, is measured over the wavelength range of each filter. This photometry provides a low resolution spectrum of the galaxy's energy distribution which is compared to models of galaxy spectra covering a range of redshifts. A successful match to a model depends on the capture of distinctive features within the galaxy's spectrum. The galaxy is assigned the redshift of the best-fit model. This technique is based largely on computational routines

and therefore, despite the lower resolution compared to spectroscopic redshifts, its strength is the ability to efficiently determine large numbers of redshifts.

2.3 The Hertzsprung-Russell Diagram

At a fundamental level, a galaxy's evolution is that of its stellar components. Stellar evolution theory describes the changes that occur in the structure and composition of a star over the course of its lifetime [11] [26]. These changes may be divided into several distinct evolutionary phases. The study of these phases relies heavily on the Hertzsprung-Russell (HR) diagram of stellar properties (see Figure 2.2).

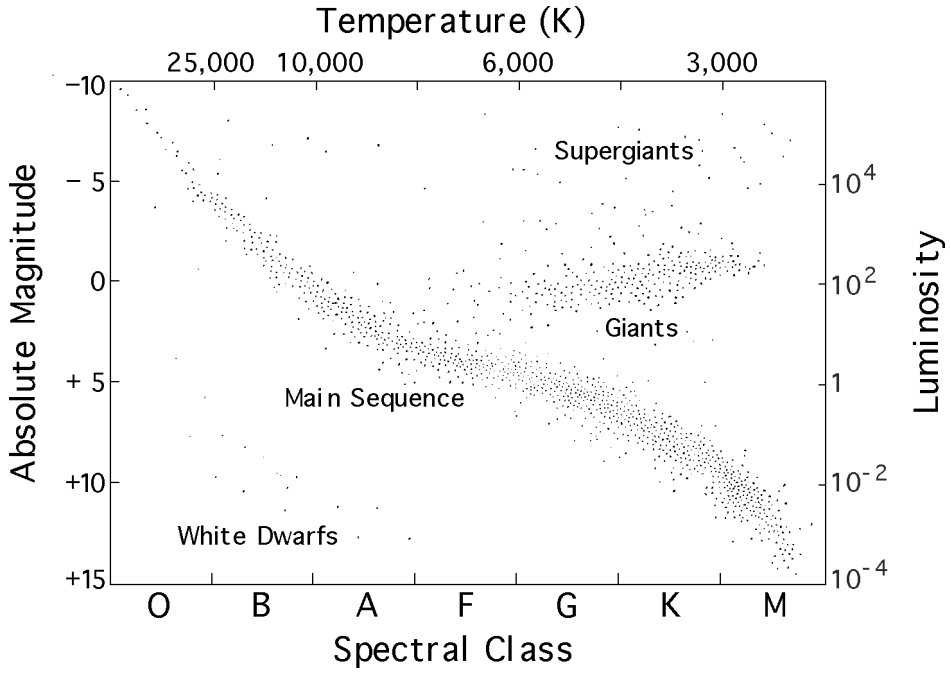

Figure 2.2: An example of a Hertzsprung-Russell diagram. Stars are distributed within the plot based on fundamental parameters. Trends in the physical properties of stars are evident and indicate evolutionary phases in the lives of stars [62].

The HR diagram (Figure 2.2) depicts the distribution of a large sample of stars based on the parameters of absolute magnitude, effective temperature, luminosity and spectral class [11] [26]. Trends in the distribution reveal the physical possibilities for stellar formation and evolution. The dominant trend is called the stellar main sequence. Stars in the steady hydrogen-burning phase of their lives occupy the main sequence. This evolutionary phase represents the longest part of the life of a star and, therefore, the vast majority, $\sim$ 80% to 90%, of stars lie on the main sequence. In the diagram, the upper left main sequence contains hot, bright, massive stars; the lower right contains cool, dim, low-mass stars.

The position of a star along the main sequence, and the evolutionary path it will follow, are predetermined by the star's mass at the time of its formation [11] [26]. In nature, massive stars are less likely to be formed than low mass stars. Massive stars also burn hotter and brighter, and have shorter main sequence lives, than lower-mass stars. This trend is revealed in the HR diagram by the predominance of low mass stars on the main sequence (see Figure 2.2). Once their stores of hydrogen are exhausted, stars move off the main sequence and enter advanced stages of evolution. The various substructures off the main sequence represent these different phases of advanced stellar evolution, including white dwarfs, red giants and supergiants.

Galaxies are comprised of stars at various stages of evolution and each of these stars contributes to the SED of a galaxy [34] [53]. Hot, young, massive stars emit light in the UV region of the spectrum, while cool, long-lived, low-mass stars emit light in the visible to near-infrared. The integrated light of a galaxy reveals the evolutionary mix of its stellar components.

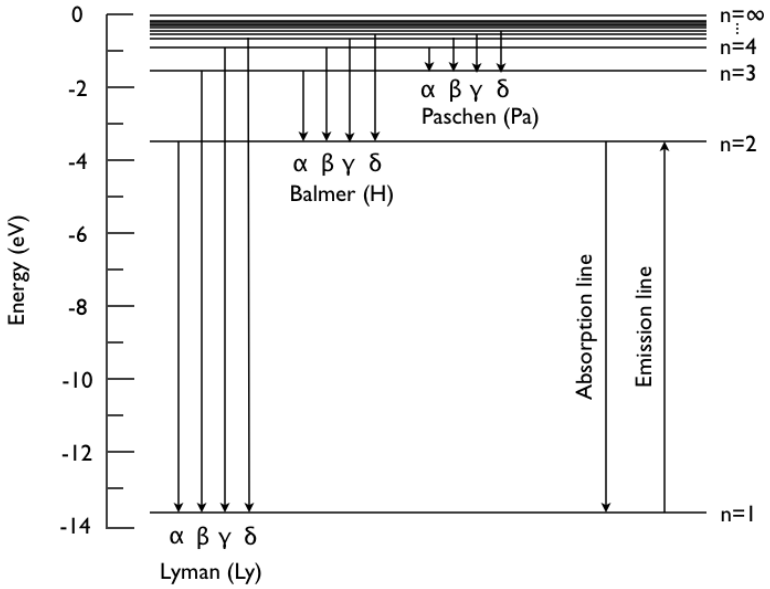

Figure 2.3: Transitions in the hydrogen atom including the Lyman series [11].

2.4 Lyman Series

The Lyman series consists of electron transitions in the hydrogen atom that occur with respect to the ground state ($E_n \leftrightarrow E_1$), as shown in Figure 2.3 [26]. The wavelengths of the Lyman series (in Å) are 1216, 1026, 972, 950, 938 ... 912, for quantum numbers n = 2,3,4,5,6 ... ∞, respectively. These wavelengths lie in the UV portion of the electromagnetic spectrum.

Galaxy spectra exhibit Lyman transition features due to hydrogen absorption of the UV radiation emitted by stars [11] [26]. The interstellar medium (ISM), which fills the space between stars, is composed primarily of hydrogen gas ($\sim$ 70% by mass). Clouds of hydrogen gas are also found within the intergalactic medium (IGM), the space between galaxies. In the low temperature, low density environment of the ISM and the IGM, hydrogen exists in its neutral ground state. Hydrogen absorption of

UV radiation at a wavelength of 1216 Å forms the prominent Lyman-α line ($E_1 \rightarrow E_2$). Lyman-α transitions occur within the ISM of a galaxy and, predominantly, within neutral hydrogen clouds in the IGM at redshifts less than the galaxy redshift. These transitions produce the Lyman-α forest absorption feature that is evident in the spectra of high redshift galaxies. The series of transitions that occur for permitted wavelengths between 912 Å and 1216 Å (i.e. Lyman-β, Lyman-γ, etc.) also contribute to the Lyman-α forest.

The Lyman series converges towards the hydrogen ionization limit of 912 Å [11] [26]. The absorption of UV radiation at wavelengths smaller than 912 Å, or energies in excess of 13.6 eV, ionizes the hydrogen atom. Massive, newly-formed stars are observed to emit vast amounts of ionizing UV radiation. In the high temperature environment surrounding these stars, hydrogen gas is almost entirely ionized. This background of UV radiation also ionizes the ISM and the IGM. Given the small interaction cross-section within these diffuse media, the probability of recombination is very low and therefore the gas remains ionized. Within hot, dense stellar environments, the rates of ionization and recombination may reach an equilibrium. Together, the intrinsic absorption of ionizing photons in the stellar atmosphere, absorption within the ISM of the galaxy, and absorption by neutral hydrogen in the IGM produce the 912 Å Lyman-break spectral feature.

2.5 Lyman Spectral Features

The photometric identification of high redshift star-forming galaxies relies on the UV spectral absorption features caused by Lyman series transitions [19] [45]. These features comprise the Lyman-break and the Lyman-α forest. The Lyman-break at rest-frame 912 Å is recognizable as a well-defined drop in the UV emission of a galaxy. For $z > 3$, absorption within the IGM reduces a galaxy's continuum

emission, shortward of 912 Å, to essentially zero. Lyman series absorption also affects the continuum spectrum from 912 Å to the 1216 Å Lyman-α line. At high redshifts, this selective absorption of UV radiation is evident in galaxy spectra as a dense series of absorption lines, reminiscent of a forest. The strength of the absorption depends on the redshift of the galaxy. Details of the Lyman-α forest are discussed in Chapter 5.

2.6 Lyman-Break Technique

For galaxies at redshifts $z \gtrsim 2$, Lyman spectral features at rest-frame UV wavelengths are redshifted into the optical band ($\sim$ 3000 Å to 8000 Å) [19]. The Lyman-break technique captures these features using three optical broad-band filters. The technique is illustrated in Figure 2.4 for a galaxy at a redshift of $z = 3.15$. In the figure, the SED of the galaxy is overlain by transmission curves for the optical filters U_n, G and R. For a redshift of $z = 3.15$, the 912 Å Lyman-break is observed at a wavelength of 3785 Å and the Lyman-α absorption line at a wavelength of 5046 Å (see Equation 2.2). These two features divide the galaxy's spectrum into three distinct regions: the ionization zone shortward of the Lyman-break ($\lambda < 3785$ Å), the region of lowered flux that is the Lyman-α forest (3785 Å $< \lambda <$ 5046 Å), and the continuum spectrum longward of the Lyman-α line ($\lambda > 5046$ Å). At a redshift of $z \sim 3$, these three regions of the galaxy's spectrum coincide with the three filter bands. The flux in each filter provides a photometric measurement of the total emission of the galaxy over the wavelength range of the filter. The difference in flux between adjacent filters defines the galaxy's colours in those filters. Every galaxy has unique identifying colours, however, due to the position of the Lyman-break, these colours will be similar for galaxies of the same redshift.

The primary goal of the technique is to identify a large number of LBGs within a

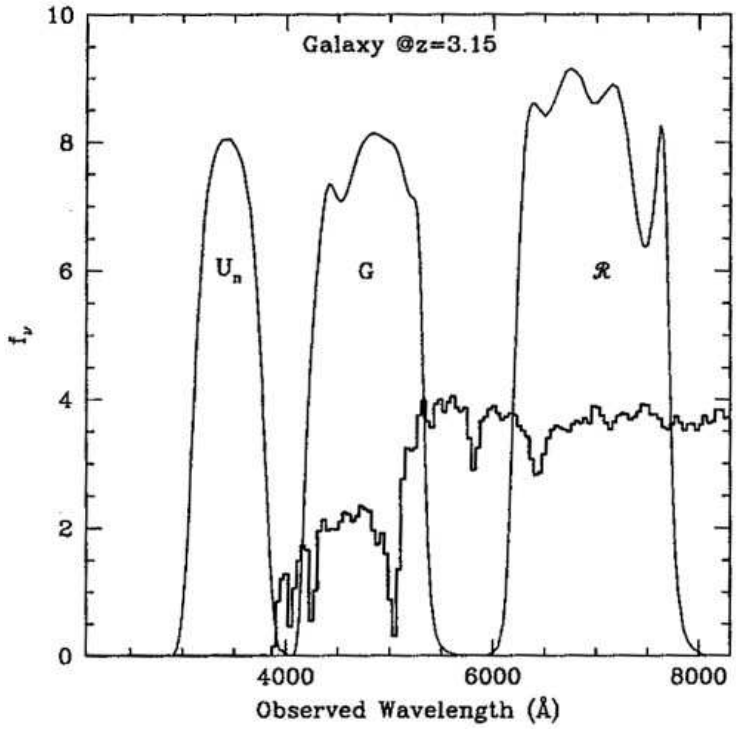

Figure 2.4: Plot illustrating the Lyman-break technique. The SED of a $z = 3.15$ galaxy is overlain by the transmission curves of the optical filters U_n, G and R. The Lyman-break occurs at a wavelength of 3785 Å and the Lyman-α forest is the diminished flux region between 3785 Å and 5046 Å [63].

specified redshift range (e.g. $3 < z < 4$) [19]. Galaxies in different ranges are isolated by adopting filter sets that cover the expected wavelength region of the redshifted Lyman-break. LBGs within the targeted redshift range are selected on the basis of position in a colour-colour plot. LBGs with similar redshifts will be localized within a specific region of the plot. The fitting of galaxy photometric data to model spectra is required to further refine a galaxy's redshift estimate.

Chapter 3

Photometric Data and Calibrations

3.1 Optical Data

3.1.1 CFHT MegaCam Data

The observational data for this project is limited to galaxies located within the four Deep Fields of the Canada-France-Hawaii Telescope Legacy Survey (CFHTLS) [64] [65]. The five year survey is a joint effort of the Canadian and French astronomical communities. The primary purpose of the survey is the detection and monitoring of supernovae and the study of high redshift galaxy distributions. These efforts are aimed towards achieving a clearer picture of the early universe and refining estimates of dark energy parameters. Analysis of galaxy and quasar populations will provide the data necessary to place constraints on theories of galaxy formation and evolution. The work of this thesis will contribute to these goals.

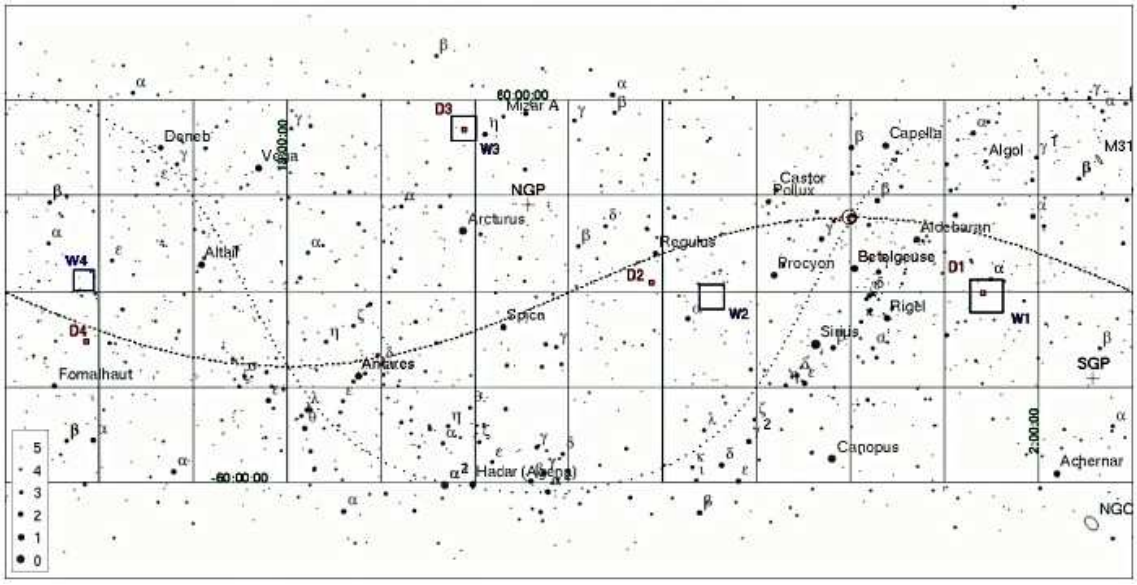

Figure 3.1: A full sky map displaying the location of the CFHTLS Deep and Wide Fields in relation to the ecliptic (darker curve) and galactic (lighter curve) planes. The one square-degree Deep Fields are labeled D1 through D4, in red [64].

Location of Fields

The CFHTLS Deep Field positions are displayed in Figure 3.1 and the Deep Field coordinates are provided in Table 3.1 [64]. Field locations were selected primarily for their low dust extinction, telescope visibility over the course of a year, and overlap with other survey projects. Each Deep Field covers one square-degree of sky.

Field	Right Ascension	Declination
D1	02° 26’ 00”	- 04° 30’ 00”
D2	10° 00’ 28”	+ 02° 12’ 21”
D3	14° 19’ 28”	+ 52° 40’ 41”
D4	22° 15’ 32”	- 17° 44’ 06”

Table 3.1: CFHTLS Deep Fields central coordinates [64].

MegaPrime/MegaCam Imaging Facility

Optical wavelength imaging at the 3.6-meter Canada-France-Hawaii Telescope (CFHT) is obtained with MegaPrime, a wide field imager, which has been equipped with MegaCam, a 36 CCD 1 degree $\times$ 1 degree field of view camera [64] [66]. The MegaPrime/MegaCam imaging facility is designed to take full advantage of the exceptional image quality afforded by the Mauna Kea environment. The complete set of 2048 $\times$ 4612 pixel CCDs provide 340 megapixel images which cover a full square-degree of sky with a resolution of 0.187 arcseconds (") per pixel.

Data collection, pre-processing, and photometric and astrometric calibrations are completed by the CFHT corporation [64] [66]. The Paris-based Terapix data processing center is responsible for fine astrometric calibration, data stacking and compilation of catalogues. Archiving of raw data and release of data products is undertaken by the Canadian Astronomy Data Centre (CADC).

3.1.2 Photometric System

The optical data for the CFHTLS Deep Fields are available in the u*, g′, r′, i′ and z′ filter set (Figure 3.2) [67]. The MegaCam 'ugriz' system is designed to meet the photometric goals of the survey. Each filter allows only a specific wavelength band to enter the telescope. By imaging an astronomical source through a sequence of filters, a low-resolution SED can be constructed. Characteristics of the source can then be determined by comparison of the SED to spectral templates and models.

When imaging a source, one measures the energy incident on the detector surface over a period of time [26]. This is a measure of flux density (F), the power of radiation per unit area at a given frequency. Flux densities are usually expressed in Jansky units, where one Jansky (Jy) is equivalent to 10^{-26} W m^{-2} Hz^{-1}. The observed flux density of a source through a certain filter is a measure of the brightness or

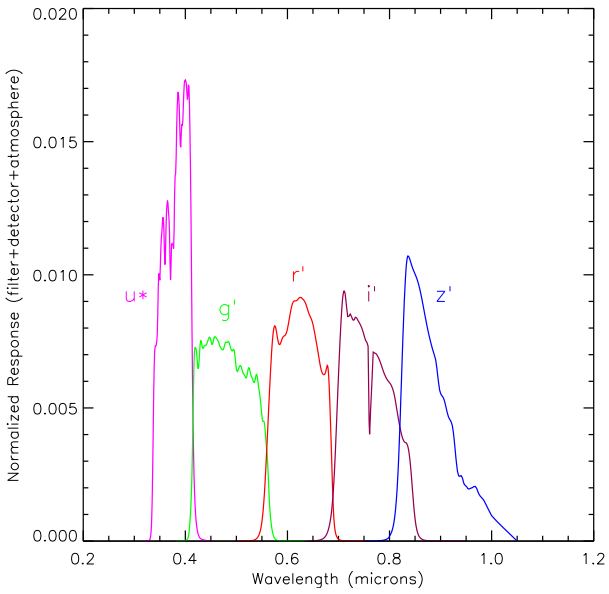

Figure 3.2: CFHTLS MegaCam optical filter response curves including typical atmospheric absorption and the CCD detector response [67].

magnitude (m) of the object at that wavelength. Magnitude values are determined with respect to a preselected flux density (F_0), defined to have a magnitude of zero. This 'zero point' differs according to the magnitude system used. Magnitudes may be defined as a function of F_0 by the equation:

$$m = -2.5log_{10}\left(F/F_0\right), \tag{3.1}$$

where F is the observed flux density of the source.

MegaCam photometry is based on the standard AB magnitude system [66] [70]. AB magnitude is defined as:

$$m_{AB} = -2.5log_{10}(F) - 48.60, \tag{3.2}$$

where F is the flux density in erg sec^{-1} cm^{-2} Hz^{-1} and a magnitude of zero is equivalent to a flux density of 3631 $\times$ 10^{-23} erg sec^{-1} cm^{-2} Hz^{-1} or 3631 Jy. All

MegaCam images have a zero point of 30 AB magnitudes. At the time of this writing, the Deep Field portion of the CFHTLS is complete [71]. The average exposure times and magnitudes depths of the completed survey are provided in Table 3.2. Note that the greater the magnitude depth, the larger the number of fainter objects that can be detected in an image.

Filter	u*	g′	r′	i′	z′
Average exposure time (hours)	19	23	42	67	47
Magnitude depth	26.5	26.1	25.5	25.1	24.9

Table 3.2: The average exposure time per filter across the four Deep Fields of the CFHTLS and the corresponding magnitude depths for the completed survey [71].

3.1.3 Colour Index

The difference between two magnitudes is referred to as a colour index, or just colour [26]. If the observed flux densities of a source through different filters are F_1 and F_2, the colour index is given by:

$$m_1 - m_2 = -2.5 log_{10} \left(F_1 / F_2 \right), \tag{3.3}$$

where m_1 and m_2 are the respective magnitudes of the source in each filter. Colour index is normally calculated between successive filters with the bluer filter written first. The smaller or more negative a colour index, the bluer the source. The larger or more positive a colour index, the redder the source. Colours defined by the MegaCam ugriz system are u* - g′, g′ - r′, r′ - i′, and i′ - z′ (see Figure 3.2).

3.1.4 Stellarity

Each Deep Field image contains a mixture of foreground stars, distant quasars and galaxies at various redshifts. To isolate a high redshift galaxy selection, all other

objects are considered contaminants. Deep Field objects are categorized in terms of stellarity, a parameter used to differentiate stars from galaxies [68]. Galaxies are spatially resolved (extended objects) and have low values of stellarity, in the range of 0 to 0.80. Stars are spatially unresolved (point-like objects) and have high values of stellarity, in the range of 0.80 to 1.00. This catalogue parameter can be applied as a first cut to separate interlopers from a high redshift galaxy selection.

3.1.5 Project Data

The optical data for this project is based on the 2007 CFHTLS Terapix T0004 catalogues and 85% best seeing images for the four Deep Fields: D1, D2, D3 and D4 [68] [69]. Seeing refers to the limit of a telescope's resolving power due to atmospheric turbulence. Best seeing images are stacks of select images chosen so as to optimize image quality while increasing magnitude depth [67]. Catalogue objects are i′-selected (objects identified in the i′ images). Each Deep Field catalogue contains more than 300,000 entries. Object magnitudes are aperture-based, determined by summing pixel values within a circular aperture centered on each i′-selected object. We used catalogues of 2" diameter aperture magnitudes generated by Chris Willott using SExtractor source extraction software [3] [54].

3.1.6 Aperture Correction

Even for point-like objects such as stars, the aperture method of determining magnitudes results in some loss of light [23]. An aperture correction is used to compensate for flux that is present but outside the bounds of the chosen aperture. Corrections are based on the loss of flux measured for a bright, unresolved object, with a further correction for galaxies due to their greater spatial extent. We make the reasonable approximation that flux is underestimated by the same amount for

all high redshift galaxies in the Deep Fields. All 2" diameter aperture magnitudes were adjusted by -0.26 mag to achieve total magnitudes [54] [68].

3.1.7 Galactic Extinction Correction

The presence of interstellar dust causes extinction and reddening of optical light from distant galaxies [26] [79]. Extinction is the loss of light primarily due to absorption and scattering by dust grains with diameters on the order of the wavelength of incident light. Light may be absorbed by the dust and re-radiated at infrared wavelengths in the range of 100-250 μm. Scattering by dust particles diminishes light intensity along the original direction of propagation. Reddening of optical light is due to the higher extinction by scattering of bluer wavelengths of light than redder wavelengths.

Schlegel et al. Dust Maps

The extinction and reddening of our optical data were corrected using the dust maps by Schlegel *et al.* [40] [79]. They have constructed a full sky 100 μm map from the combined and re-analyzed infrared maps generated by the satellite missions COBE/DIRBE and IRAS/ISSA. The Cosmic Background Explorer (COBE) and the onboard Diffuse Infra-Red Background Experiment (DIRBE) have performed measurements of the infrared and microwave radiation originating from the early universe [81]. The Infrared Astronomical Satellite (IRAS) has completed full sky surveys at wavelengths of 12, 25, 60 and 100 μm and identified $\sim$ 350,000 infrared point sources [82]. The IRAS Sky Survey Atlas (ISSA) is a compilation of moderate resolution images derived from the survey data.

Schlegel *et al.* dust maps are corrected for infrared point source contamination, zodiacal foreground (the light emitted by dust grains within our solar system) and

cosmic infrared background [79]. The maps are calibrated using a standard reddening law and the predicted colors of elliptical galaxies. Corrections to optical data are made by extracting the dust opacity and corresponding level of galactic extinction along a specified line of sight. The applicability of map corrections becomes limited in regions of high extinction.

Applying Dust Maps

Since the resolution of the dust maps is $\sim$ 2 arcminutes (’), each square-degree Deep Field was divided into $\sim$ 1’ $\times$ 1’ sections and the same correction was applied to all objects in the section. One square-degree at the celestial equator = 60’ $\times$ 60’ = 3600 1’ $\times$ 1’ squares. Further from the celestial equator, the area described by one degree of right ascension (RA; analogous to longitude) by one degree of declination (Dec; analogous to latitude) does not equate to the same one square-degree area found at the equator. The spherical coordinate system introduces a factor, the cosine of declination, such that distance in RA is correctly described by ΔRA cos (Dec) [11]. This factor was incorporated into the sectioning of each field.

The mean RA and Dec of the Deep Field objects within each $\sim$ 1’ $\times$ 1’ square were used to determine the center of each small region. A coordinate transformation, from equatorial to galactic, was performed and the Schlegel *et al.* dust map value at the center of each square was retrieved [61]. Since extinction is wavelength, and hence filter, dependent, the dust extinction magnitudes are based on dust value and a multiplicative factor specific to each filter band. The extinction factors for each filter are listed in Table 3.3, and extinction corrected magnitudes in each filter were calculated according to:

$$\text{extinction corrected magnitude} = \text{magnitude} - (\text{filter factor} \times \text{dust value}) \quad (3.4)$$

The typical magnitudes of dust extinction at r′ are provided in Table 3.4 for each of

Filter	u*	g′	r′	i′	z′
Correction factor	5.155	3.793	2.751	2.086	1.479

Table 3.3: Galactic extinction correction factors for MegaCam optical filters [54].

Deep Field	1	2	3	4
Typical dust extinction magnitude at r′	0.07	0.05	0.03	0.07

Table 3.4: Typical magnitudes of dust extinction at r′ for each Deep Field.

the Deep Fields.

Galactic extinction affects the colours of objects beyond our galaxy, however, by applying the correction to all objects in a field, the magnitudes of stars in our own galaxy are also adjusted. This introduces a higher degree of scatter in the colours of stars within the field. We considered this an acceptable error and made no further adjustment.

3.1.8 Magnitude Limits

Limiting magnitude is a measure of the faintest object detectable in an image [23]. Determining the limiting magnitude, however, depends on the threshold above which a detection is believed to be real compared to the expected noise in the signal of the instrument. In this analysis, a detection is considered to be real when the signal-to-noise ratio (S/N) is greater than or equal to 2, or more probable than 95%. If the measured magnitude is greater than this 2-sigma (2σ) limit (i.e. dimmer), then the signal is considered a non-detection.

For the ugriz filters, a real object may be detectable in some filters, but not visible in others (Figure 3.3). This is in fact the basis for the Lyman-break technique for

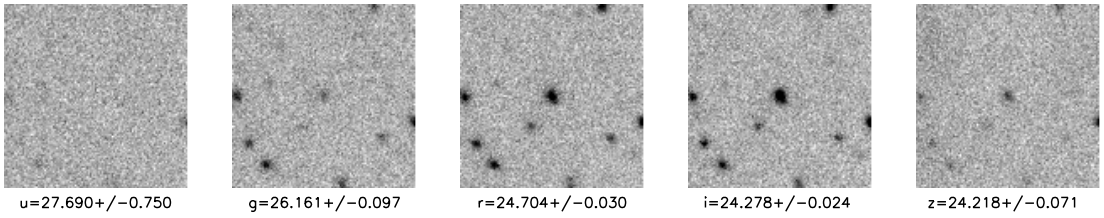

Figure 3.3: Sample D2 Field object imaged in the optical filters. The target is centered in the image.

estimating photometric redshifts, described in Section 2.6 [57]. The flux density in a certain filter is reduced due to the Lyman spectral break at rest-frame UV wavelength being redshifted into that filter (see Figure 2.4). These non-detections are referred to as 'dropouts' [5].

Although objects may not be detectable in certain filters, this information is still considered. Objects that are fainter than the detection limit are assigned the limiting magnitude of that filter (i.e. a 2σ lower limit is given). In that way, limits on magnitude also place limits on colour.

To determine the limiting magnitude, we define two magnitudes such that $m_2 = m_1 + \Delta m$ [18]. Rearranging and incorporating Equation 3.3 gives:

$$\Delta m = m_2 - m_1 = -2.5 log_{10} \left(\frac{F_2}{F_1} \right). \tag{3.5}$$

Setting,

F_1 = flux of the detection limit = 2σ

F_2 = flux of the detection limit + 1σ uncertainty = 3σ

and substituting into Equation 3.5 gives:

$$\Delta m = -2.5 log_{10} \left(\frac{3\sigma}{2\sigma} \right) = -0.44. \tag{3.6}$$

The limiting magnitude is reached when the error on the magnitude, Δm, has a value of 0.44 mag. Optical limiting magnitudes were determined from plots of magnitude

versus error on magnitude for the D1 Field data. Individual plots were generated for objects imaged in each filter band. A sample plot is provided in Figure 3.4 and limiting magnitude values are listed in Table 3.5. The limiting magnitudes determined for each filter band apply to all four Deep Fields.

The photometric uncertainties of objects whose magnitudes are set to the limiting magnitude of a filter must be adjusted accordingly. Equation 3.7 can be used to determine the uncertainty of an aperture photometry measurement when the signal-to-noise ratio (S/N) is defined, as in the case of limits, [18]:

$$SE = -2.5log_{10}(1 - (S/N)^{-1}), \tag{3.7}$$

where SE is the conservative 1σ standard error in units of magnitude. For objects at an optical limiting magnitude of 2σ (i.e. a S/N $\geq$ 2), photometric uncertainties were set to the standard error of 0.75 mag.

3.2 Infrared Data

The optical data for this project is extended into the infrared with the addition of near-infrared CFHT WIRCam K band data and mid-infrared four-channel Spitzer IRAC data. This enhanced data set will be crucial in eliminating the degeneracies inherent in galaxy model fitting based solely on optical measurements. Infrared measurements also make possible the identification and elimination of contaminants within the data set.

3.2.1 Spitzer Space Telescope Data

The Spitzer Legacy Cosmic Evolution Survey (S-COSMOS) [39] is a deep infrared imaging survey covering the two square-degree COSMOS field, centered at RA = 10:00:28.6, Dec = + 02:12:21.0 [73] [84]. Within the bounds of the COSMOS

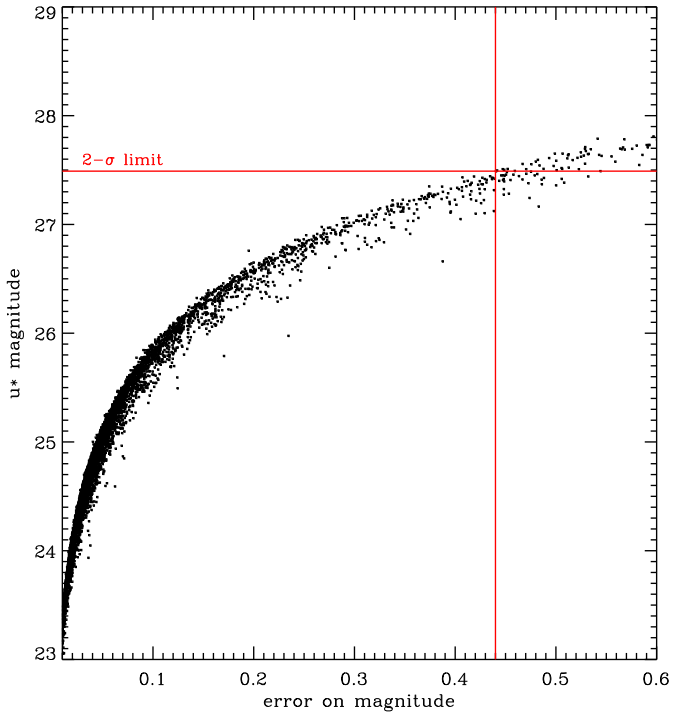

Figure 3.4: Magnitude versus error on magnitude plot for determining the 2σ limiting magnitude in the u* filter using D1 Field data.

Filter	Magnitude limit (original data)	Magnitude limit (dust & aperture corrected data)
u*	27.87	27.49
g′	28.21	27.85
r′	27.96	27.62
i′	27.70	27.42
z′	26.49	26.21

Table 3.5: Optical filter limiting magnitudes for the T0004 data release. Limiting magnitudes are based on a 2σ limit which corresponds to an uncertainty on the measured magnitude of 0.44. Limits are given for both the original and corrected D1 Field data. Limits apply to all four Deep Fields.

field is the D2 Field imaged by the CFHTLS. Data for S-COSMOS are obtained with the Spitzer Space Telescope, a 0.85-meter telescope carrying three cooled and shielded instruments: the Infrared Array Camera (IRAC), the Multiband Imaging Photometer for Spitzer (MIPS) and the Infrared Spectrograph (IRS) [85]. Infrared images are captured in a range of broad-band filters: 4-channel IRAC data centered at wavelengths of 3.6, 4.5, 5.8 and 8.0 μm and far-infrared MIPS data at 24, 70 and 160 μm. The survey has compiled a series of photometric catalogues based on sources detected at 3.6 and 24 μm. Image mosaics are available for all seven filter bands. The primary goal of S-COSMOS is the study of large galaxy populations for evidence of: stellar mass and star formation rate evolution, active galactic nuclei (AGN), and the presence of large-scale structure [73] [86]. IRAC data in particular will be used to determine photometric redshifts for S-COSMOS galaxies in the range of $1 < z < 5$.

3.2.2 IRAC Project Data

The Spitzer infrared data for this project consists specifically of IRAC images at 3.6, 4.5, 5.8, and 8.0 μm [83] [84] [86]. Each channel detector is comprised of a 256 $\times$ 256 pixel array with a resolution of $\sim$ 1.2" per pixel, producing images 5.12' $\times$ 5.12' in size. S-COSMOS images are resampled to a scale of 0.6" per pixel. The S-COSMOS archive provides Spitzer data in the form of cutouts, small image clips centered on the coordinates of each object of interest [90]. Spitzer IRAC data were accessed for the 5765 D2 Field objects matching our optical selection criteria (discussed in detail in Chapter 4). Cutouts of $\sim$ 40" per side (67 square pixels at 0.6"/pixel) were generated for each D2 Field object in all four IRAC channels (3.6, 4.5, 5.6, 8.0 μm).

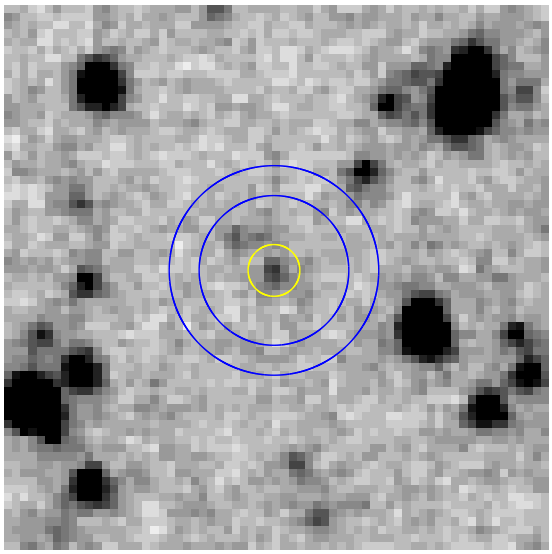

Figure 3.5: Spitzer IRAC image showing sample aperture placement for concentric aperture photometry. The target aperture is shown in yellow. The surrounding annulus for determining the sky brightness is shown in blue.

3.2.3 Concentric Aperture Photometry

The routine APER, adapted from the Dominion Astrophysical Observatory photometry program DAOPHOT [50], was used to compute concentric aperture photometry from the image cutouts of the D2 Field objects [92]. In aperture-based photometry, fluxes are obtained by summing the image pixel values within a circular or elliptical aperture centered on each object [2] [23]. The brightness of the background sky is estimated by summing the pixels within an annulus placed concentrically around the target aperture (see Figure 3.5). The background sky level is subtracted from the target aperture flux to obtain the flux of the object. Background noise in the image is calculated from pixel-to-pixel variations between adjacent sky pixels within the annulus.

A standard IRAC 1.9" (3.14 pixel) radius was used for the target aperture in all

four IRAC channels. Fifteen objects at bright, mid-range and dim i′ magnitudes ($\sim$ 15.6 to 24.7 mag) were selected and their Spitzer cutouts examined to determine the placement of the sky annulus. An annulus was configured with an inner radius of 9" (15 pixels) and an outer radius of 18" (30 pixels). The APER routine returned the magnitude for each object, the error on the magnitude, the background sky value and the error on the sky value.

3.2.4 IRAC Data Corrections

Aperture Corrections

An aperture correction specific to each IRAC channel was applied to the magnitude data to compensate for the source flux that falls outside the 1.9" aperture radius. Table 3.6 shows the correction factors that are applied to obtain the total flux for each IRAC channel and aperture. The stated aperture corrections are implemented as multiplicative factors to the flux. We converted the 1.9" aperture correction factors to magnitude differences (Ch1 : 0.291, Ch2 : 0.327, Ch3 : 0.510, Ch4: 0.591). These differences were then subtracted from the magnitudes measured in each channel to obtain total magnitudes.

Aperture radius	Channel 1 3.6 μm	Channel 2 4.5 μm	Channel 3 5.8 μm	Channel 4 8.0 μm
1.4"	0.610	0.590	0.490	0.450
1.9"	0.765	0.740	0.625	0.580
2.9"	0.900	0.900	0.840	0.730
4.1"	0.950	0.940	0.940	0.910

Table 3.6: Aperture correction factors for Spitzer IRAC channels. Fractional values represent the amount of the source flux that is measured within the aperture. Measured flux values are divided by the aperture correction factor in order to obtain total fluxes [84].

Magnitude Limits

The variation in background sky values measured by the APER routine was used to determine the limiting magnitude at the location of each object. Limiting magnitudes were derived using an algorithm for the RMS noise in the target aperture, adapted from DAOPHOT [50] [91]:

$$\begin{aligned} \text{RMS noise} &= (\text{variance in the local sky})^{1/2} \\ &= (\text{variance of the sky brightness} \times \text{aperture pixel count})^{1/2} \\ &= \text{error on the sky value} \times (\text{aperture pixel count})^{1/2}, \end{aligned} \tag{3.8}$$

where aperture pixel count is the aperture area in square pixels, and RMS noise is measured in units of flux.

Equation 3.8 assumes that the background noise properties are characterized by Poisson statistics, however, this is only valid when the image pixels are uncorrelated [2] [4] [30]. For correlated sky values in adjacent pixels, noise is artificially lowered and 2σ noise fluctuations occur more often than expected. Data reduction procedures, such as image stacking, can introduce correlated signal effects and lead to an underestimation of photometric uncertainties. To account for this effect in the Spitzer IRAC imaging, a 3σ limit was chosen to define the limiting magnitude. Limits for each object in each channel were calculated according to:

$$\text{limiting magnitude} = -\ 2.5\ \log_{10}(3 \times \text{RMS noise}) + 25, \tag{3.9}$$

where 25 is the default zero point magnitude of the APER routine. If the channel magnitude of an object was greater than the 3σ limit, the object was classified as a non-detection in that channel and its magnitude changed to the limiting magnitude. The photometric uncertainties of objects at the 3σ magnitude limit were set to 0.44 mag (see Equation 3.7).

Zero Point

IRAC images have a zero point of 21.58 AB magnitudes calculated according to Equation 3.10:

$$\text{zero point} = 2.5 \log_{10} \left(\frac{3631 \times 10^6}{8.4616} \right) = 21.58, \tag{3.10}$$

where the factor of 8.4616 converts IRAC photometry in units of MJy/sr to μJy/pixel [84]. Since the APER routine assumes a zero point of 25, a correction of 3.42 mag was subtracted from all routine-derived magnitudes and limits to account for the difference in zero points.

3.2.5 CFHT WIRCam Data

K band data is a product of two surveys: the Wide-field InfraRed Camera (WIRCam) Deep Survey (WIRDS) [87] [88], which provides deep near-infrared imaging of the Deep Fields, and the Sanders-led PI program to map COSMOS with WIRCam [73]. Data is obtained at the CFHT with WIRCam for the J, H, K_s filter set ($\sim$ 1.0 to 2.2 μm) [89]. WIRCam consists of four 2048 $\times$ 2048 pixel detectors with a 20' $\times$ 20' field of view and a resolution of 0.3" per pixel. WIRCam images are subsequently resampled to the MegaCam pixel scale of 0.186" per pixel to allow for direct pixel correlation between WIRDS and CFHTLS images. Image processing and distribution is again undertaken by the CFHT, Terapix and the CADC.

3.2.6 K Band Project Data

The WIRCam K band image was obtained for the COSMOS/D2 Field. WIRCam images, at one-ninth the field of view of CFHT's MegaCam, were available as a single 19354 $\times$ 19354 pixel image covering the full one square-degree D2 Field [88]. A catalogue of 2" aperture K band magnitudes was compiled by Chris Willott using

the locations of the i′-selected D2 Field objects previously catalogued in the optical band [54]. A discrepancy between the indexing systems of the K band and optical catalogues resulted from using different versions of the source extraction software, SExtractor [3], to generate the catalogues. Corresponding objects in the optical and K band were matched by performing a pixel-based position search between catalogues.

3.2.7 K Band Data Corrections

Due to pixel-to-pixel sensitivity variation across the field, K band limiting magnitudes were individually determined for each object [72]. The background sky values near each object were measured with the APER routine used previously for the IRAC photometry. A target aperture with a 2.0" (10.75 pixel) radius was selected to be consistent with the 2.0" aperture magnitudes generated for the K band catalogue. Since object images in the K band are compact, the sky annulus was configured with a tight inner radius of 4" (21.51 pixels) and an outer radius of 10" (53.76 pixels).

Limiting magnitudes were calculated using Equations 3.8 and 3.9, following the procedure used for the IRAC data. Limits were again set by a 3σ flux. A value of 6.40 mag was added to each limit to correct for the APER routine default zero point of 25 mag and the K band image zero point of 31.40 mag [54] [88]. Where the K band magnitude of an object was greater than the limiting magnitude, the magnitude of the non-detection was changed to the limiting magnitude. Photometric uncertainties of objects at the 3σ magnitude limit were set to 0.44 mag (see Equation 3.7). All K band catalogue magnitudes and 2.0" aperture limits derived from the image were adjusted by -0.17 mag to obtain total magnitudes.

3.2.8 Summary

A summary of the multi-wavelength data used in this project is provided in Table 3.7. The filter response curves for the WIRCam K band and Spitzer IRAC infrared filters are shown in Figure 3.6.

3.3 Spectroscopic Data

3.3.1 VIMOS VLT Deep Data

The VIMOS VLT Deep Survey (VVDS) is a large spectroscopic survey covering 16 square-degrees of sky in four fields of 2×2 square-degrees [28] [94]. The survey is a French-Italian project designed to study the formation and evolution of galaxies, galaxy clusters, AGN and large-scale structure, with the goal of constraining current cosmological models. A total of 150,000 spectroscopic redshifts are to be measured for galaxies in the range of $z = 0$ to $z = 5$.

The survey is made possible through the use of VIMOS, a VIsible Multi-Object Spectrograph, installed at the 8.2-meter Melipal unit telescope, part of the European Southern Observatory (ESO) Very Large Telescope (VLT) at the Paranal Observatory in Chile [28] [95]. VIMOS is a four-channel spectrograph, each channel covering a $\sim$ 7' $\times$ 8' field of view for a total of $\sim$ 218 square-arcminutes. Slit masks are used to obtain as many as 1000 spectra in a single exposure, collected on a 2048×2048 square-pixel CCD at a scale of 0.205" per pixel. Depending on the number of slits per mask, the spectral resolution ranges from $\sim$ 200 to $\sim$ 2500.

Data	Source	Filter	Central wavelength (μm)	Wavelength range (μm)
Optical	CFHT MegaCam	u*	0.38	0.318 - 0.550
		g′	0.48	0.374 - 0.606
		r′	0.63	0.512 - 0.720
		i′	0.77	0.660 - 0.908
		z′	0.89	0.780 - 1.020
Near-infrared	CFHT WIRCam	K	2.15	1.927 - 2.382
Mid-infrared	Spitzer IRAC	Ch 1	3.6	3.081 - 4.010
		Ch 2	4.5	3.723 - 5.222
		Ch 3	5.8	4.744 - 6.623
		Cn 4	8.0	6.151 - 10.497

Table 3.7: Optical and infrared data sources, filters and transmission wavelengths [64] [89] [83].

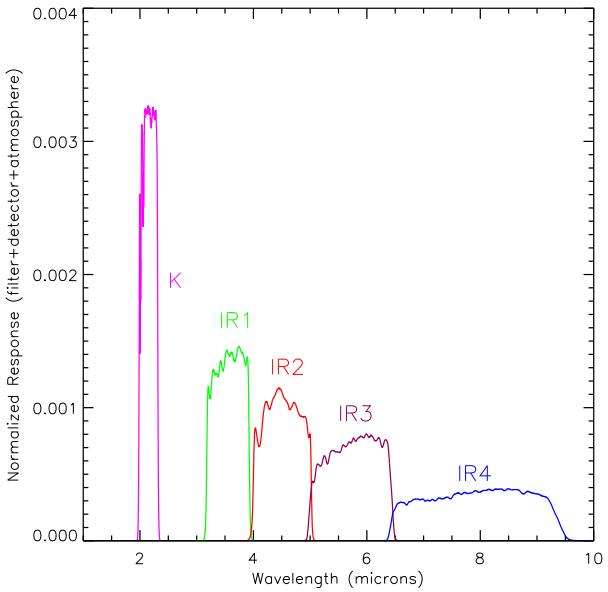

Figure 3.6: CFHTLS WIRCam K band and Spitzer 4-Channel IRAC infrared filter response curves including typical atmospheric absorption and the CCD detector response [83] [89].

3.3.2 Project Data

The VVDS-DEEP V1.0 catalogue contains 8981 spectroscopic redshifts for objects in the VVDS-F02 Deep Field (RA = 02:26:00, Dec = -04:30:00) [24] [96]. The VVDS-F02 Deep Field has an area of ~ 0.4 square-degrees in common with the CFHTLS D1 Deep Field, thus providing redshift data for a small subset of D1 Field galaxies. The VVDS spectroscopic redshifts are provided for a magnitude-limited set of objects in the range of $17.5 \leq I_{AB} \leq 24$ and have a maximum redshift of 5.228. Flags and confidence classes are assigned to VVDS redshift estimates based on the quality of the spectroscopic data. These are given in Table 3.8.

Flag	Description
0	no redshift could be assigned to the 1D spectrum
1	50% confidence in redshift
2	75% confidence in redshift
3	95% confidence in redshift
4	100% confidence in redshift
9	single isolated emission line spectra

Table 3.8: VVDS-DEEP redshift data flags indicating the confidence levels of spectroscopic redshift estimates [96].

3.4 Blended Source Elimination

A possible source of contamination of the photometric data is flux within the target aperture that originates from a bright nearby object. This effect is largest for the Spitzer data which has the poorest resolution and requires the largest apertures. A catalogue of S-COSMOS sources [84] may be used to locate and exclude all targets that have a Spitzer source close enough to significantly compromise the target photometry. The random nature of the distribution of targets with blended Spitzer

data makes it possible to eliminate them from the analysis with zero or negligible bias.

Targets are eliminated based on the level of flux contamination within the target aperture. By considering a Gaussian distribution of flux from possible contaminating sources, one may vary the distance between the target and the contaminating object according to the amount of contributed flux. Thus, the fainter of the bright objects may be closer to the target before being considered as contaminating sources. In taking this approach, the number of targets that would be eliminated based solely on a minimum target-source separation is reduced. This analysis is applied to the target galaxy selection (see Chapter 4), with details as follows.

3.4.1 Contaminant Flux

The S-COSMOS source catalogue includes IRAC 4-channel photometry for all objects that have a measured flux above 1 μJy in IRAC Channel 1 (3.6 μm) [84]. Flux measurements for Spitzer objects close to bright stars or positioned near the edge of the image (flags 1 and 2) are not trustworthy. Therefore, Spitzer sources were limited to those with reliable flux measurements (flag 0 in all 4 IRAC channels).

Of the four Spitzer IRAC channels, Channel 1 has the deepest data and the best resolution for resolving distinct sources and the blended source analysis is therefore based on Channel 1 photometry [16]. The Spitzer map has fairly uniform noise such that the level of contaminating flux was selected to be 2 times the typical uncertainty on the Spitzer source flux in Channel 1 (i.e. a 2σ detection). This level of contamination is equivalent to a magnitude of 24, or a flux of 0.912 μJy. Thus, if a nearby Spitzer source contributes a flux of 0.912 μJy or greater to a target aperture, the target photometry is considered contaminated.

It should be noted that contamination is greater in the other IRAC channels, in

particular at Channels 3 and 4, due to their broader point spread functions (PSFs) [16]. A PSF describes the distribution of flux that results from the distortion of a point source within an image. The broader the PSF, the larger the number of pixels over which the image is spread. Consequently, a source centered just outside the photometry aperture boundary will have a larger percentage of its flux falling within the target aperture. The level of contamination will depend on the proximity of the source to the target aperture. Given the much lower point source sensitivities in these channels, however, signals of $\sim$ 24 mag are not detected and the uncertainties due to noise are much greater than the flux of contaminants.

3.4.2 Source Flux Profile

To apply a Gaussian approach, one must be able to estimate the amount of flux contained within any-sized aperture centered on a Spitzer source. Spitzer IRAC correction factors to the total flux are shown in Table 3.6 for a limited set of apertures. Using this information, a plot was produced of the IRAC apertures (1.4", 1.9", 2.9" and 4.1") versus the stated aperture correction factors for Channel 1. The resulting curve can be approximated by an exponential function that defines the correction factor as follows:

$$\text{correction factor} = [[\text{A} \times \exp(\text{-radius}^{B})]+\text{C}]^{-1}, \tag{3.11}$$

where radius is the aperture radius in arcseconds and A, B and C are best-fit coefficients [93]. A function fitting routine was applied to estimate the values of the best-fit coefficients for our data set [20]. Results are provided in Table 3.9. Using these values with Equation 3.11, a formula to estimate the amount of source flux contained within an aperture of any radius was obtained. Note that flux estimates based on the fit are more accurate for larger radii.

Coefficient	A	B	C
Best-fit value	1.521 ± 0.020	0.945 ± 0.020	0.998 ± 0.006

Table 3.9: Estimated Spitzer IRAC Channel 1 best-fit coefficients for Equation 3.11, used to determine the correction factor to the total flux for an aperture of any radius centered on a Spitzer source.

3.4.3 Defining a Blended Source

By determining the distance between each target object and the closest bright source, a 'zero' separation of 0.31" was set to avoid the object itself appearing as a contaminant. If the separation between a bright source and a target object was found to be ≤ 0.31", they were considered to be the same object. (This step compensated for a slight position offset between the two data sets.) To define a source as non-blended, a minimum separation of source and target was set as follows. Target fluxes were measured using an IRAC 1.9" aperture (see Section 3.2.3). From Table 3.6, an IRAC 2.9" aperture centered on a bright source contains 90% of the flux in Channel 1. Adding these two radii, a minimum target-source separation of 4.8" was set as the limit for consideration as non-blended. At this distance, a bright source would contribute no more that one percent of its flux as contaminating flux within the target aperture. Using these criteria, any possible blended source must therefore lie within 0.31" and 4.8" of the target object.

To estimate the amount of contaminating flux from a bright nearby source that is contained within the 1.9" aperture of a target object, it was necessary to consider the position of the target aperture within the flux distribution of the bright source (see Figure 3.7). The radial distances from the bright source to the closest and furthest edges of the target aperture were determined. Using Equation 3.11, the portion of the bright source flux contained within the annulus formed by these two

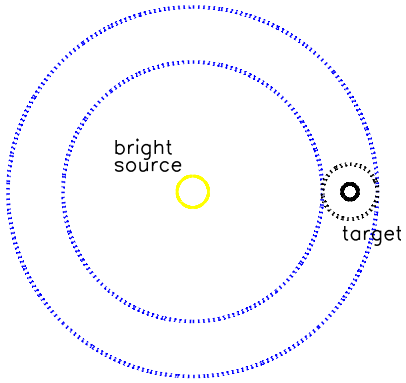

Figure 3.7: Blended source illustration showing the position of a target object within the flux field of a bright source. The dotted blue circles represent the bright source apertures that enclose the target aperture (dotted black circle). Diagram is not to scale.

radii was calculated, as well as the percentage of that flux falling within the target aperture area. If the contaminating flux was $\geq$ 0.912 μJy, the target and source were considered blended and the target was removed from the data set. If the contaminating flux was less than 0.912 μJy, a search was made for multiple sources of contamination. If the integrated contaminating flux of multiple sources became greater than 0.912 μJy, then the target object was eliminated. All other objects were considered non-blended and the target photometry reliable.

3.4.4 Results

The blended source analysis was applied following the isolation of a target sample of D2 Field LBGs (described in Chapter 4). Of the 1359 LBG candidates initially isolated, 351 or approximately 26% were found to have photometry blended with that of a nearby Spitzer source (see Figure 3.8). Those LBG candidates matching the criteria for a blended source were excluded from the data set, leaving 1008 targets for further analysis.

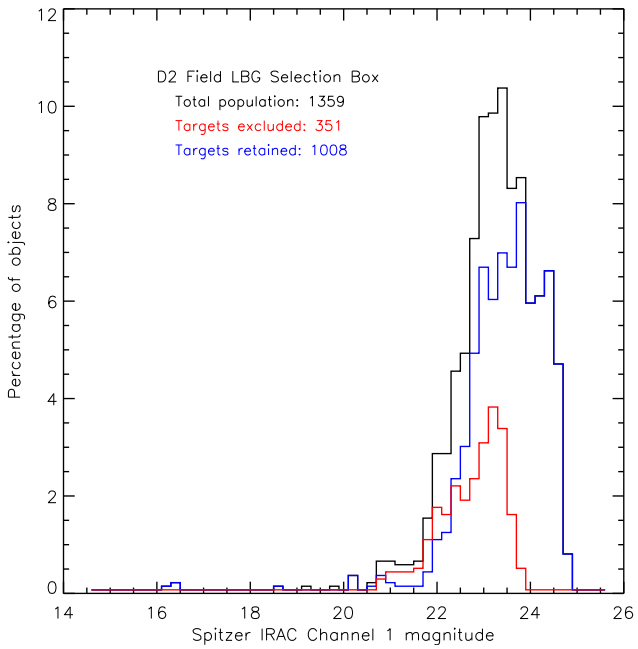

Figure 3.8: Elimination statistics for target LBGs with blended Spitzer photometry. Targets are excluded based on a flux contamination of 0.912 μJy or greater within the target aperture that results from a nearby Spitzer source.

Chapter 4

Lyman-Break Galaxy Sample

High redshift galaxies are isolated according to the principle of Lyman-break colour selection [43] [57]. Using this technique, galaxies at various redshifts are identified by the specific regions they occupy in colour-space. We take advantage of the photometric and spectroscopic redshift data for the CFHTLS D1 Deep Field to locate an initial region of interest. Once the colour-space of a LBG sample has been identified, further discrimination and analysis may then take place.

4.1 Combining Survey Data

The region covered by the VVDS-DEEP spectroscopic survey partially coincides with the CFHTLS D1 Field, permitting a match of magnitude and redshift data for select objects within the region of overlap [24] [96]. Slight astrometric discrepancies generally exist between experimental data sets from different projects. Therefore, a position match of objects was necessary in order to combine the photometric and spectroscopic data from the two surveys.

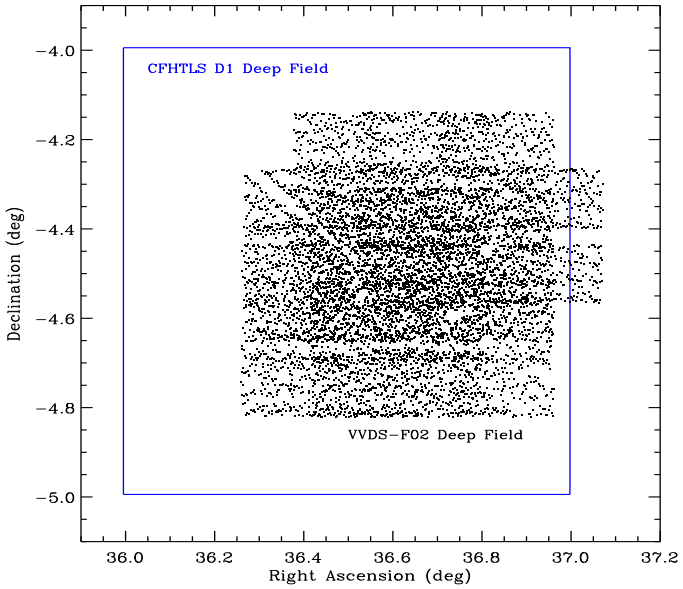

Figure 4.1: Regions covered by the CFHTLS D1 Deep Field and the VVDS-F02 Deep Field [68] [96].

4.1.1 Position Match

The VVDS-DEEP data set contains spectra for 8981 objects with I_{AB} magnitudes ≤ 24 [96]. The region covered by both surveys includes 196,134 i′-selected D1 Field objects and 8806 VVDS objects. A plot of the overlapping fields is provided in Figure 4.1. To locate the same object in each field, the separation in RA and Dec was determined between all spectroscopic and photometric catalogue objects within the region of overlap. The range of offsets for each coordinate is expected to have a median value of approximately zero. Deviation from a zero-centered distribution was corrected by adjusting one field by the median offset in RA and Dec. Median offsets were determined through an iterative technique whereby objects were matched and shifted until a centered distribution was achieved. RA and Dec offsets of 0.0176" and 0.2772", respectively, were applied to the data.

The object separation was defined by:

$$\text{total offset} = [(\Delta\text{RA})^2 + (\Delta\text{Dec})^2]^{1/2}, \tag{4.1}$$

where ΔRA and ΔDec are the offsets in RA and Dec, respectively. The elements with the smallest total position offset were considered matched and total position offsets were adjusted by -0.1697" to account for the remaining offset. A plot of the distribution is shown in Figure 4.2, along with a set of corresponding colour-coded histograms showing the number of matching objects versus distance from the center of the distribution (i.e. zero offset in position). To better determine the true matches, objects are binned and colour-coded according to Δmag, the difference between the I_{AB} and i′ magnitudes of the objects. Mismatched objects can be expected to have larger magnitude differences than true matches. Magnitudes were adjusted by -0.0742 to correct for median offset. The histogram plot inset in Figure 4.2 displays normalized histograms covering the large offset tail of the distribution.

Overall, objects with a small position offset are also well matched in terms of magnitude. The histogram inset (Figure 4.2) reveals the expected increase in magnitude difference for objects with large position offsets. Approximately 90% of all matches have a position offset ≤ 0.5" as well as a magnitude difference ≤ 0.5 mag. However, matches with offsets greater than 0.5" may also be valid. To maximize the value of our data and retain all reasonable matches, a cutoff distance of 0.655" was selected to define an acceptable match.

4.1.2 Secure Spectroscopic Redshifts

In all, 8472 objects ($\sim 96\%$) were found to match with a separation difference of no more than 0.655". From these objects, cuts were made for spectroscopic redshift confidence level and value. A description of the VVDS-DEEP redshift flags is given in Table 3.8. A total of 4038 objects with low quality redshift flags (0, 1, 2, 9) were

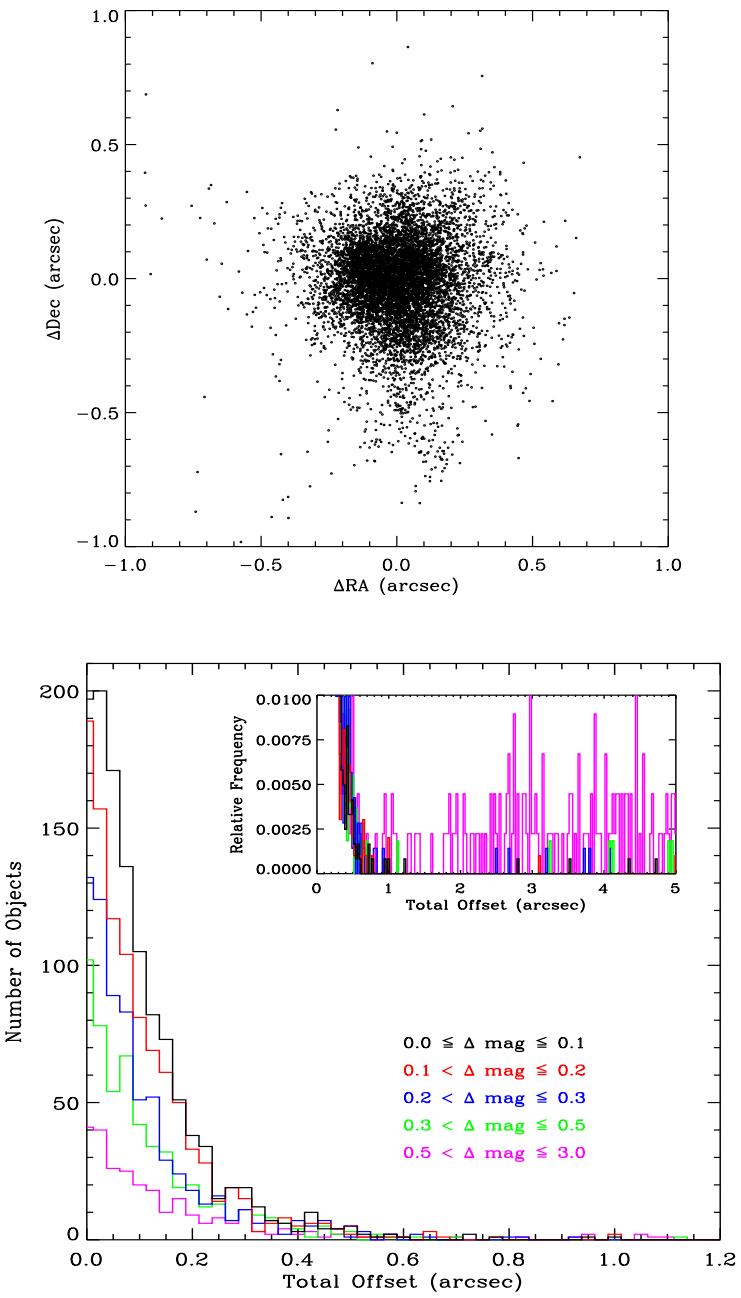

Figure 4.2: Scatter plot (top) and histogram (bottom) of position offsets between CFHTLS D1 Deep Field objects and VVDS-F02 Deep Field objects within the region of overlap. Histograms are colour-coded for offset in I_{AB} - i′ magnitude. Inset plot displays normalized histograms covering the tail of the distribution.

removed. As well, 279 objects with two-digit redshift flags, designating the presence of a secondary target in the slit or a quasi-stellar object (QSO), were eliminated. Also excluded were 672 stars (redshift 0) of any confidence level. The remaining objects with secure spectroscopic measurements (flags 3 & 4) provided a minimum 95% confidence level and a final set of 3673 D1 Field galaxies with reliable spectroscopic redshifts.

4.2 Isolating a LBG Sample

4.2.1 Colour-Redshift Plots

A series of colour-redshift plots was generated for the D1 Field galaxies with secure spectroscopic redshifts from the VVDS. The colour-redshift relationship can be used to determine the redshift at which the colour transitions from blue to red due to the Lyman-break passing through the filters [24] [57]. Plots of the four optical colours, u* - g′, g′ - r′, r′ - i′ and i′ - z′, versus redshift are shown in Figure 4.3. The main population of the sample is comprised of low redshift galaxies at $z \lesssim 1.5$, with a scattering of higher redshift galaxies evident out to $z \sim 4.3$. The u* - g′ plot reveals a trend from blue to red colour (i.e. from smaller to larger colour index) over the redshift range of $2.5 \lesssim z \lesssim 4$. A number of u*-dropouts are also indicated. In the plot of g′ - r′, colour clearly transitions from blue to red over the redshift range of $3 \lesssim z \lesssim 4+$. The plot of r′ - i′ shows a slight tendency towards redder colour at $z \gtrsim 4$ but primarily the plots of r′ - i′ and i′ - z′ exhibit galaxies with flat spectrums and a blue colour over the redshift range of $2 \lesssim z \lesssim 4$. These results indicate the redshifting of the Lyman-break through the u* and g′ filters. A high redshift LBG sample may therefore be isolated by its red g′ - r′ colour and its blue r′ - i′ colour. Selecting the reddest objects in g′ - r′ would provide a sample of LBGs at $z \sim 4$.

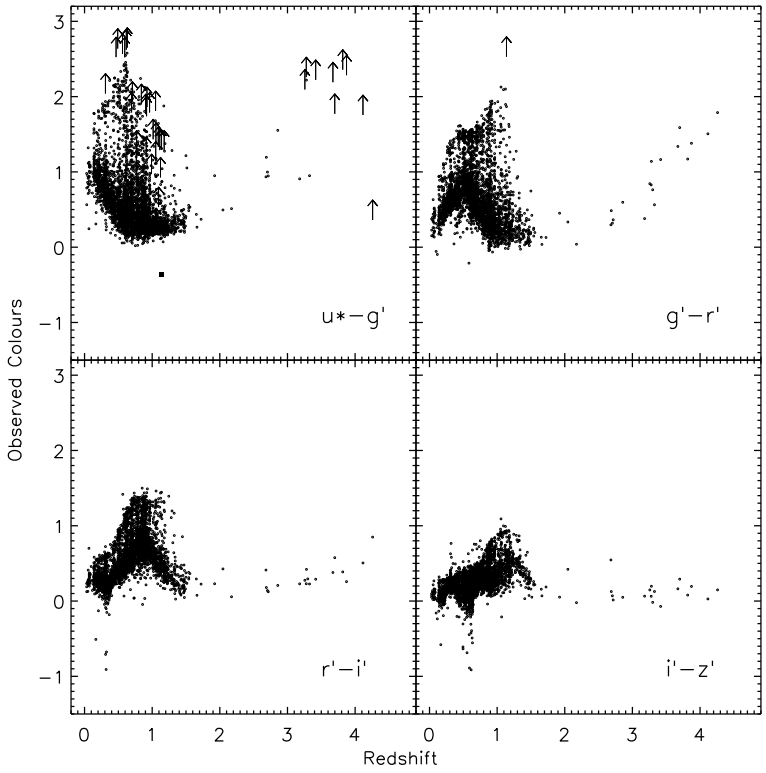

Figure 4.3: Colour-redshift plots of CFHTLS D1 Deep Field objects with secure spectroscopic redshifts from the VVDS. For objects that drop out of the bluer filter, a 2σ lower limit is given and indicated by an arrow. A solitary object that drops out of both the u^* and g' filters is indicated by a square.

Also evident in Figure 4.3 is the oscillation of colour at low redshifts. This effect can be attributed to redshifting of the Balmer/4000Å break, which is present in most galaxy spectra [1] [27] [47]. The Balmer series arises from electron transitions ($E_2 \leftrightarrow E_n$) in atomic hydrogen (see Figure 2.3) [26]. The Balmer-break at ~ 3700 Å is prominent in the spectra of hot stars where hydrogen atoms are present in an excited initial state. The 4000Å-break results from the absorption lines of ionized metals in the spectra of older, cooler stars. Considering galaxy stellar populations, the Balmer-break is a stronger feature for younger star-forming galaxies while the 4000Å-break

is more pronounced for older galaxies that have metal-rich stellar populations. The presence of these spectral features at filter wavelengths means that low redshift galaxies can exhibit the same red colour used to isolate high redshift galaxies. Low redshift galaxies are therefore a potential source of contamination of the LBG sample.

4.2.2 Colour-Colour Plots

A series of colour-colour plots was generated for the D1 Field galaxies with secure spectroscopic redshifts from the VVDS (see Figure 4.4). Galaxies are binned according to redshift using ranges determined from the colour-redshift plots of Figure 4.3. Redshift bins are adjusted so as to optimize the colour-space isolation of a sample of high redshift galaxies spanning a redshift range of $\Delta z \sim 1$. From the three plots of Figure 4.4, it can be seen that the isolation of a specific LBG sample is made difficult by the overlap of galaxies of various redshifts. Ideally, we wish to separate a complete LBG sample within a select redshift range while minimizing sample contamination by galaxies outside this range. Of the three plots, u^* - g' vs. g' - r', g' - r' vs. r' - i' and r' - i' vs. i' - z', an isolated population of high redshift galaxies is most easily located in the plot of g' - r' vs. r' - i' (lower plot, Figure 4.4). These are LBGs in the range of $3.4 < z \leq 4.4$ (shown in purple) that exhibit a red g' - r' colour and a blue r' - i' colour, as discussed Section 4.2.1.

A selection box is outlined that surrounds this target galaxy population and defines the colour-space of the LBG sample (lower plot, Figure 4.4). The selection box is designed so as to maximize the completeness of the sample while avoiding contamination by the main distribution of lower-redshift foreground galaxies [43] [57], shown in green. Due to the proximity in colour-space of different galaxy populations, however, there is always a need for compromise between completeness and contamination. The selection box currently surrounds six LBGs with redshifts in the range

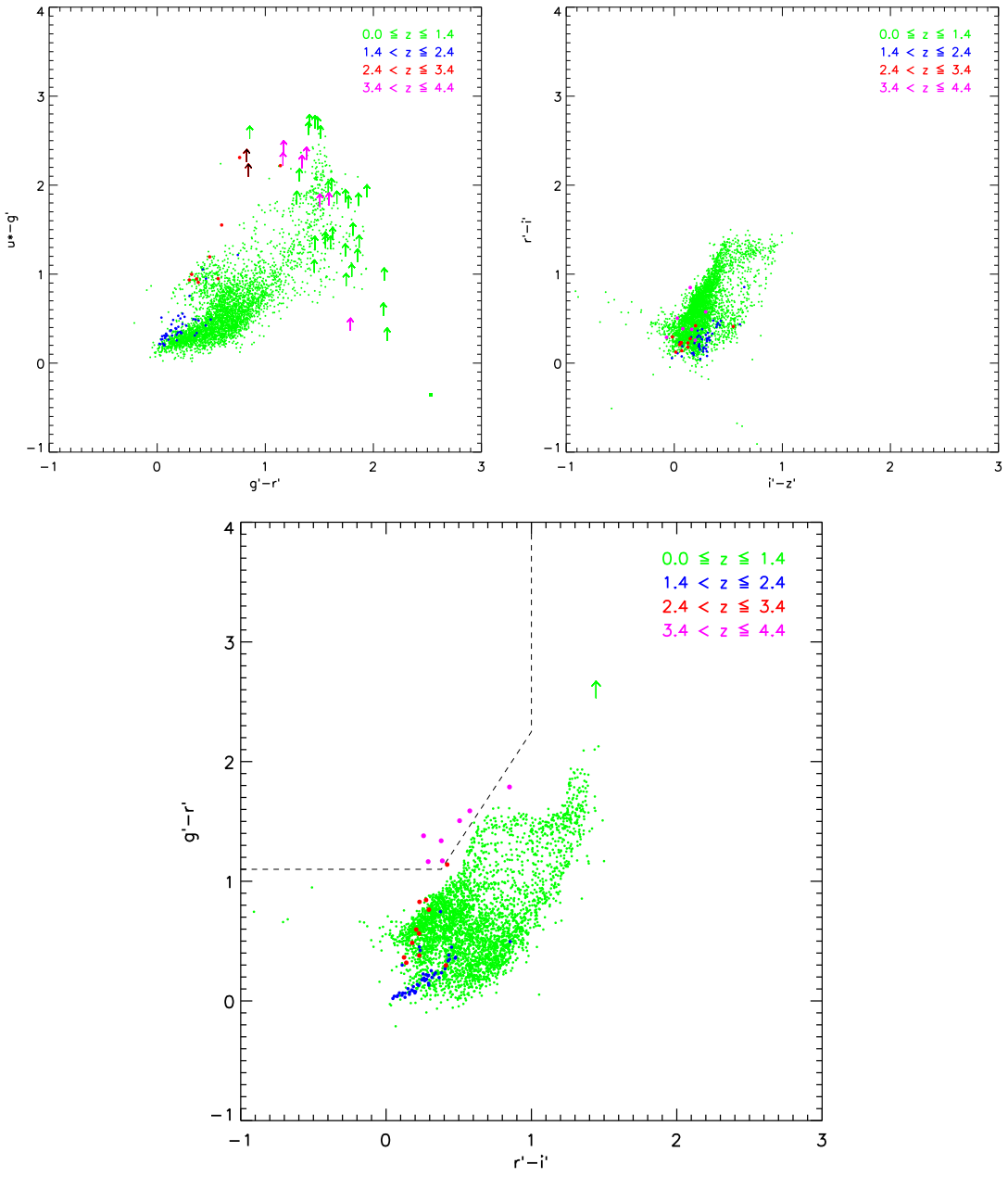

Figure 4.4: Colour-colour plots of D1 Field galaxies with secure spectroscopic redshifts from the VVDS. Galaxies are binned and colour-coded according to redshift. A 2σ lower limit is given for dropouts and indicated by an arrow. The colour-space of a LBG sample in the range of $3.4 < z \leq 4.4$ is indicated by the selection box outlined in the plot of g′ - r′ vs. r′ - i′.

of $3.4 \lesssim z \lesssim 4.4$, establishing our initial region of interest. Using full Deep Field data, the selection box will contain a much larger population of LBG candidates. Optimal placement of the selection box is further analyzed in Chapter 5.

4.3 Deep Fields Colour Selection Box

4.3.1 Sample Selection and Contamination

Colour-colour plots of g′ - r′ vs. r′ - i′ were generated for the complete set of i′-selected objects in each of the four CFHTLS Deep Fields. The magnitude limits of $i' \leq 24.0$, $i' \leq 24.5$ and $i' \leq 24.8$ were used to produce a series of plots for each field. Figure 4.5 displays the four Deep Field plots to a magnitude depth of $i' \leq 24.8$. The colour selection box of LBG candidates is outlined in each of the fields. Objects are colour-coded based on stellarity index. Objects with low stellarity (≤ 0.80) indicate galaxies and are plotted in purple. Objects with high stellarity (> 0.80) indicate stars, quasars or compact galaxies and are plotted in green. In addition, by selecting only those objects with the highest stellarity (> 0.98), we separate the local stars from the quasars and compact galaxies. These stars form a blue swath in each plot that traces the position in colour-space of the stellar main sequence (see Section 2.3). The differentiation of stars within the high stellarity population reveals that the main distribution of each field consists primarily of a large number of compact, low redshift galaxies and possibly some quasars. The structure within the plots is indicative of the physical limitations that exist on the colour of astronomical objects, analogous to the trends seen in the HR diagram of stellar properties (Figure 2.2). In the plots of Figure 4.5, the overlap of the various populations makes it difficult to see the underlying galaxy distribution. For clarity, a plot displaying only the galaxy populations of the D1 field is provided in Figure 4.6.

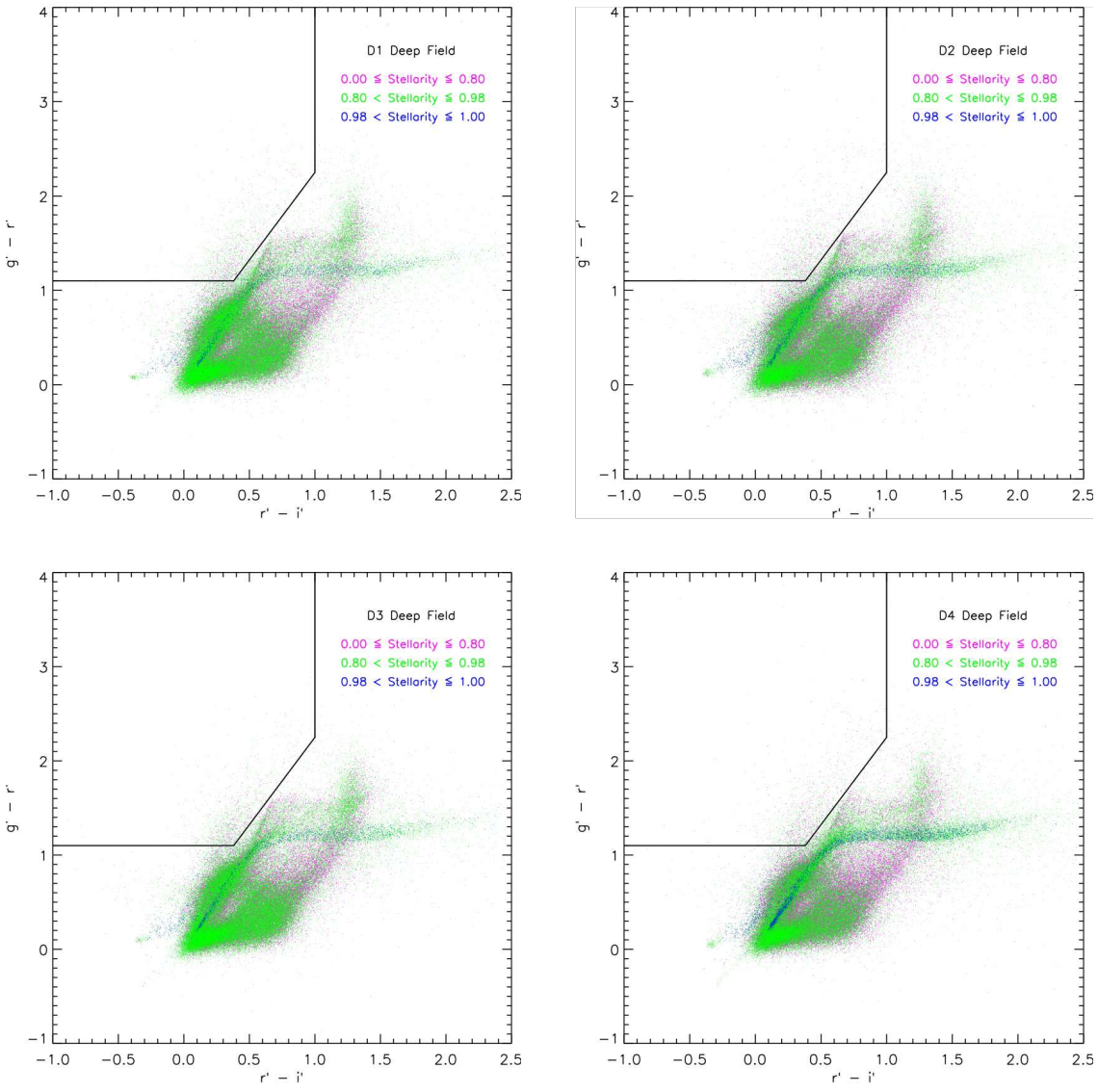

Figure 4.5: CFHTLS Deep Field colour-colour plots showing the selection box for LBG candidates in the range of $3.4 \lesssim z \lesssim 4.4$. Objects have a magnitude limit of $i' \leq 24.8$ and are colour-coded according to stellarity index. Galaxies have low stellarity and are plotted in purple; quasars and compact galaxies have high stellarity and are plotted in green; stars have the highest stellarity and are plotted in blue.

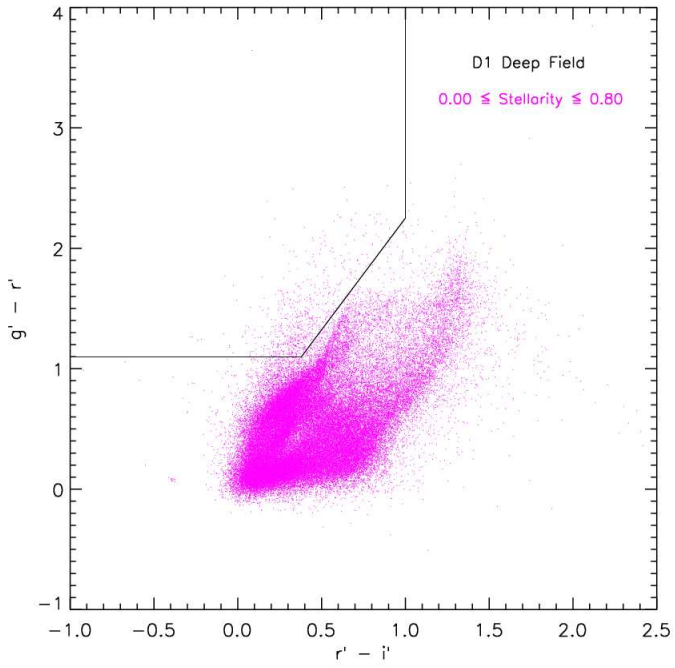

Figure 4.6: CFHTLS D1 Deep Field colour-colour plot of galaxy populations. High stellarity populations comprised of stars, quasars and compact galaxies are omitted from the D1 field plot in order to show the distribution of the low stellarity galaxy populations.

The Deep Field plots in Figure 4.5 display narrow, highly detailed features that reveal the excellence of the photometric data. The low degree of scatter contributes greatly to our ability to successfully avoid the large number of contaminating low redshift galaxies at the fringe of the main distribution. Despite its proximity in colour-space to the high redshift region, the edge of the low redshift population is well-defined, allowing for the precise positioning of the selection box. The boundary of the box closely skirts the dense region of foreground stars and low redshift galaxies in each field, fostering completeness of the high redshift sample while minimizing low redshift contamination. The exceptional photometry provides an advantage

compared to other LBG surveys which suffer higher contamination rates as a result of lower S/N data.

Within the Deep Field LBG selection regions of Figure 4.5, there is evidence of some (green) point-source contamination. These objects are most likely compact elliptical galaxies at redshifts of $z \sim 0.5$, exhibiting high redshift colours due to the 4000Å-break rather than the Lyman-break, and some quasars at the same redshift as the LBGs [57]. Photometric errors may also account for some of the scatter of objects into the selection boxes from the high population regions. Rates of contamination are determined following photometric redshift analysis of the LBG sample in Chapter 5.

4.3.2 Colour Offset Correction

A comparison of the plots in Figure 4.5 reveals the high level of colour similarity between each of the Deep Fields. In particular, the colour evolution of stars along the main sequence is expected to be statistically identical for all fields [97]. An overlap of the stellar main sequences of each field uncovered slight colour offsets between the fields. The discrepancies occurred in only r′ - i′ colour, not g′ - r′ colour, and therefore only i′ magnitudes were adjusted. Corrections were made with respect to the stellar main sequence of the D2 Field which was centrally positioned. Magnitude correction values for each field are given in Table 4.1.

Field	D1	D2	D3	D4
Correction to i′ magnitude	+0.01	-	−0.01	−0.02

Table 4.1: CFHTLS Deep Fields i′ magnitude corrections based on offset in stellar main sequence colour.

Field	Magnitude limit $i' \leq$	Number of objects in selection box	Number of objects in field
D1	24.0	218	77,219
	24.5	722	113,680
	24.8	1431	142,154
D2	24.0	184	82,313
	24.5	679	118,478
	24.8	1359	145,753
D3	24.0	179	75,808
	24.5	704	111,921
	24.8	1392	140,024
D4	24.0	212	79,503
	24.5	805	115,616
	24.8	1567	143,541

Table 4.2: The number of objects contained within the LBG selection box of each Deep Field to various i' magnitude limits (see Figure 4.5). Total field numbers are given for comparison.

4.4 Optical Selection Criteria

Formally, we set the selection criteria for our LBG sample as:

$$\begin{aligned} g' - r' &\geq 1.1 \\ r' - i' &\leq 1.0 \\ g' - r' &\geq 1.8551(r' - i') + 0.3951. \end{aligned} \tag{4.2}$$

These values define the colour bounding box for LBG candidates in the redshift range of $3.4 \leq z \leq 4.4$, as outlined in Figure 4.5. The number of objects in each Deep Field that meets the colour selection criteria is listed in Table 4.2. We limit our selection to objects at $i' \leq 24.8$.

An important factor influencing number counts in galaxy fields is the presence of underlying large-scale density fluctuations in the local universe [33] [44] [51]. The

uncertainty introduced by large-scale structure is referred to as cosmic variance and is exhibited as a field-to-field variation in observational number counts. Cosmic variance is a particular issue in high redshift studies that tend to be focused on small fields of view. Uncertainties in galaxy abundances and densities limit the conclusions that can be drawn regarding the properties of high redshift galaxy populations and their evolution. Further, Trenti and Stiavelli (2007) state that, "Relatively little effort has been so far devoted at quantifying the impact of cosmic variance on the determination of properties of high redshift galaxies."

Estimates of cosmic variance are currently based on statistical data for similar galaxy populations and on the galaxy clustering predictions of CDM theory [44]. As the volumes of space surveyed become larger, cosmic variance becomes less of an issue. The Deep Field number counts from Table 4.2 show a high consistency from field to field (i.e. within 10% for full field counts). This indicates that a large enough volume of the Universe is being sampled and that our results are not significantly affected by cosmic variance.

4.5 Extended Selection Box

In an effort to optimize boundary placement and investigate the nature of objects close to the selection box which may be photometrically scattered in, we extend the boundary to include those objects close to the edge of the box (see Figure 4.7). This extended selection box now includes the spur of objects on the main galaxy distribution. The extended selection box is defined by:

$$\begin{aligned} g' - r' &\geq 1.0 \\ r' - i' &\leq 1.0 \\ g' - r' &\geq 1.80(r' - i') + 0.10. \end{aligned} \tag{4.3}$$

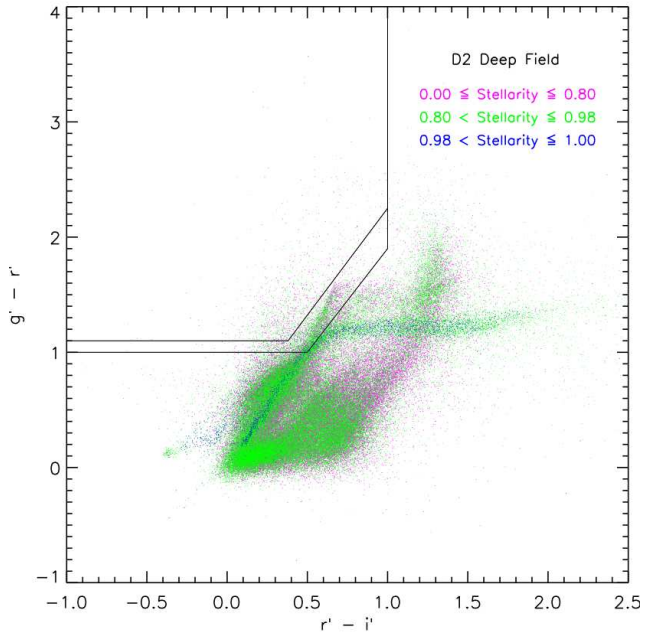

Figure 4.7: CFHTLS D2 Deep Field colour-colour plot showing the extended selection box that includes objects close to the edge of the nominal LBG selection box. The extended selection box for the D2 Field contains 5765 objects.

A histogram of spectroscopic redshifts for the 76 VVDS galaxies located within the extended region of the selection box is provided in Figure 4.8. The population consists predominantly of low redshift ($z \lesssim 0.5$) galaxies, as expected, but also includes two galaxies at higher redshifts ($z > 3$). The observed infrared colours of objects within the extended selection box and the predicted colours of various galaxy models will be used to further refine the colour-space selection of LBG candidates (see Chapter 5).

Note: While this preliminary work in the optical was completed for all four of the Deep Fields, the extended multi-wavelength data necessary for the final analysis became available at a later date, and only for the D2 Field. As a result, the final results of the thesis are based solely on the analysis of galaxies within the D2 Field.

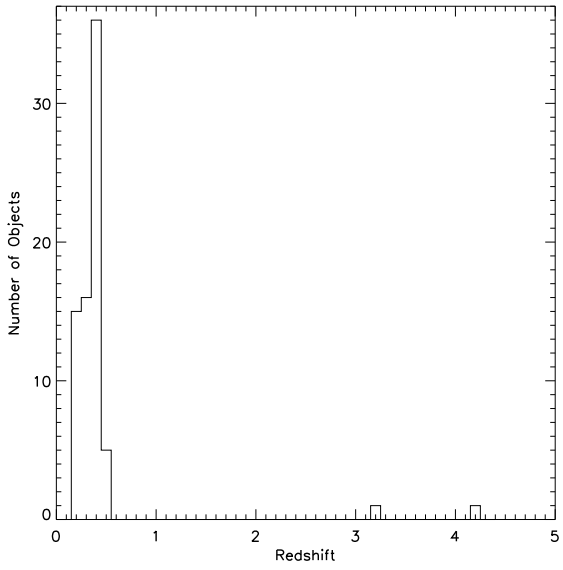

Figure 4.8: Histogram of spectroscopic redshifts for the 76 VVDS galaxies located within the extended region of the selection box (see Figure 4.7).

Chapter 5

Galaxy Fitting

The Lyman-break technique is used to identify galaxies within a broad redshift range. For more precise redshift estimates, galaxy models may be used. In addition, fitting photometric data to galaxy model spectra can provide estimates of galaxy properties such as star formation rate (SFR), stellar mass, dust content and age. We use the world-leading models of Charlot and Bruzual (CB2008) (unpublished) to determine the properties of our LBG selection.

CB2008 galaxy models are designed using stellar population synthesis, a technique founded on the concept that, "*The star formation history of galaxies is imprinted in their integrated light* [10]." As such, CB2008 models are built from the level of individual stars into galaxies, producing high resolution model spectra. Contained within these highly detailed spectra are the underlying properties of individual galaxies without which a definitive picture of LBGs would not be possible. The construction and application of CB2008 models is discussed in this chapter.

5.1 Stellar Population Synthesis

5.1.1 Astrophysical Components

Population synthesis modelling incorporates two lines of astrophysical research: stellar evolutionary tracks and stellar spectra [9] [10]. The path, or track, of a star's evolution is uniquely determined by its initial mass and chemical composition, which fall within a broad range of possibilities. The phases that comprise these tracks are outlined in the HR diagram (Figure 2.2). Complementary to evolutionary tracks are the spectra of stars in each phase of evolution. Libraries of synthetic and observational spectra provide the typical SEDs of stars at each phase of the evolutionary process. Observational collections of spectra offer increased resolution while theoretical models fill in the observational gaps to produce uniform sampling. Together, these components allow one to model the spectral evolution of individual stars and combine them to build model galaxies.

5.1.2 Isochrone Synthesis

To study galaxy evolution, we would like to model the SED of a stellar population and follow its spectral evolution as a function of time. This may be accomplished through the technique of isochrone synthesis [9] [10]. At a given point in time, a stellar population is comprised of stars in various phases of evolution. The number of stars in each phase is determined by the distribution of stellar masses at the time of formation of the population. This mass distribution is estimated using an empirically-based initial mass function (IMF). The addition of new stars to the population may be estimated by introducing a SFR and by considering the rate of chemical enrichment resulting from stellar evolution. An isochrone connects the evolutionary tracks of stars of different initial masses at a specific point in time.

It is essentially a snapshot in time of the distribution of evolutionary states in a stellar population. Spectra are assigned to stars in each evolutionary phase along the isochrone. By summing the individual spectral contributions of all stars at each point along the isochrone, the SED of a stellar population is determined. The evolution of a population can then be modelled as a series of evolving SEDs, determined by computing isochrones as a function of time.

5.1.3 The Model

For their model, Bruzual and Charlot combine a number of theoretical prescriptions of stellar evolution, based on the agreement of observational data for stars in different phases of evolution [9] [10]. Assembling stellar evolutionary tracks from several libraries, according to these prescriptions, produces a set of tracks spanning the physical range of initial stellar masses. Each phase of stellar evolution along the tracks incorporates observational or synthetic spectra. These spectra have been updated, enhanced and extended at ultraviolet and infrared wavelengths. The result is improved model fitting of high redshift galaxies.

The galaxy SEDs generated by the model span the wavelength range of 91 Å to 160 μm [9] [10]. The model produces a datapoint every 3 Å between 3200 Å and 9500 Å. This provides the maximum resolution and exceptional detail in the model spectra. An evolutionary time span of 20 billion years is covered by the model, with spectra generated at 221 select ages within this range. Different evolutionary scenarios are defined by the modeler's choice of IMF, SFR and rate of chemical enrichment.

In application, the model has been successful in reproducing the colour-magnitude and integrated colour characteristics of star clusters of different ages and chemical compositions [9] [10]. The model-generated SEDs provide excellent fits to observed

galaxy spectra, including the distribution and strengths of prominent spectral features. However, despite the current sophistication of population synthesis models (PSMs), there remain limitations to the interpretation of galaxy spectra. The advanced phases of stellar evolution are particularly challenging to model due to the combined effects of large-scale changes in stellar properties during this stage. These phases are a crucial component of PSMs since the presence of even a small number of old stars within a population can contribute significantly to the integrated flux. In addition, the reddening of galaxy colours may be attributable to any combination of increased age, increased metallicity, greater dust extinction. The inability to distinguish between these influences leads to degeneracies in the model fitting.

To facilitate the interpretation of observational data, a limited number of spectral templates are produced from the model [9] [10]. Each template is the result of a linear combination of physical parameters that have been chosen to define the fit. For model fitting to high redshift galaxies, the model spectra must be individually modified to account for the attenuation of flux by intervening dust and Lyman forest absorption. We apply Bruzual and Charlot PSMs to the analysis of our LBG sample using models in the form of galaxy colour tracks (Section 5.2) and galaxy spectral templates (Section 5.3).

5.2 Infrared Colours and Selection Efficiency

In Chapter 4, the optical selection criteria that defines our LBG sample was established, with an extended region to include those objects close to the colour limit, which we expect to be mostly lower redshift ($z < 0.5$) galaxies (see Figure 4.7). Located within this extended selection box are 5765 D2 Deep Field objects. Infrared K band and Spitzer IRAC data can be used to investigate the nature of these objects. Colour-colour plots that incorporate infrared data allow one to separate various

redshift populations within the sample. Colour-boundary plots may then be used to determine which of the populations fall within the nominal LBG selection region. In addition, model tracks are included to map the expected colours of different stellar populations as a function of redshift. The goal is to ensure the nominal LBG selection boundary is optimally placed to contain a sample of $z \sim 4$ LBGs.

5.2.1 Model Colour Tracks

The plots in this section are overlain with CB2008 model tracks of galaxy colour as a function of redshift. We compare three different evolutionary scenarios:

- A galaxy that formed 13 billion years ago and has a constant SFR with no dust reddening ($\tau = 0.0$).
- A galaxy that formed 13 billion years ago and has a constant SFR with dust reddening ($\tau = 0.5$).
- A simple stellar population in which all the stars formed in a single burst 13 billion years ago, with no dust reddening ($\tau = 0.0$).

These model choices allow us to bracket the range of possible star formation histories and to investigate the influence of dust on colour.

The presence of dust within a galaxy causes a shift in galaxy colour that is dependent upon wavelength [26]. The amount of reddening is introduced to the model by selecting an optical depth of dust extinction (τ), which is approximately equal to the number of magnitudes of light extinction. Dust extinction effects are discussed in detail in Section 5.3.2.

A simple stellar population (SSP) is defined as a passively aging population in which no new stars are formed following an initial instantaneous burst [9]. In a SSP, UV emission from bright, massive stars is significant only when the population is

very young [53]. For models with a constant SFR, massive stars and UV emission can be produced at any age.

5.2.2 Colour-Colour Plots

Most galaxies are dominated by cool stars with a rest-frame peak emission at a wavelength of approximately 1.5 μm [11]. For galaxies at a redshift of $z = 0.5$, this peak emission falls into the K band (at ~ 2.25 μm), whereas for galaxies at $z = 2$, the peak is redshifted into the IRAC Channel 2 band (at ~ 4.5 μm). The difference in a galaxy's magnitude between these two bands results in a negative colour index for low redshift galaxies and a positive colour index for high redshift galaxies. This property is applied to separate the different redshift populations within the LBG sample.

Optical and infrared data were used to generate colour-colour plots for the objects located within the extended selection box of the D2 Field. These included K - IR2 vs z′ - K and IR1 - IR2 vs z′ - IR1, where IR1 and IR2 are the Spitzer IRAC Channel 1 (3.6 μm) and Channel 2 (4.5 μm) magnitudes, respectively. A high degree of scatter was evident in the plots due to the large photometric uncertainties that are associated with faint objects at infrared wavelengths. This scatter was reduced by plotting only those objects detected with small magnitude errors (i.e. Δm < 0.15 mag) in IR1 and IR2. As a result, only about half of the 5765 objects in the extended selection box were included in the plots. The distributions were overlain with CB2008 model tracks for the three evolutionary scenarios described above.

Of the colour-colour plots produced, the plot of K - IR2 vs z′ - K (Figure 5.1) provides the best discrimination of low redshift and high redshift objects. Evident in the plot are: a grouping of stars at K - IR2 ~ -1.5, a dense population of low redshift galaxies at $-1.1 \lesssim$ K - IR2 $\lesssim -0.3$, and a broad scattering of high redshift

galaxies at K - IR2 $\gtrsim$ 0.2. The differences in colour index cleanly separate the objects at $z \ll 1$ from those at $z \gg 1$. This technique is especially effective since we expect most of the contamination of the LBG sample to be in the form of stars and galaxies at redshifts of $z \sim 0.5$ (see Section 4.3.1). The colour-boundary plots discussed in the next section provide further information on the position of these objects relative to the nominal LBG selection box boundary.

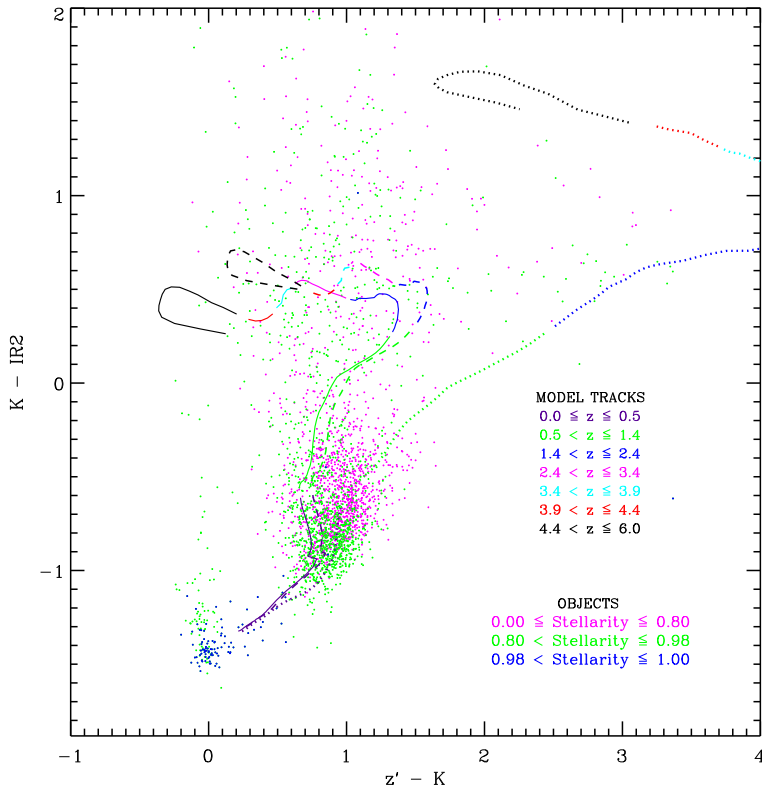

Figure 5.1: A colour-colour plot of objects located within the extended selection box of the D2 Field. Plotted objects are limited to those with magnitude errors less than 0.15 in IR1 and IR2, colour-coded according to stellarity index. Lines represent model tracks: the dust-free constant SFR galaxy model (solid), the dust-reddened constant SFR galaxy model (dashed), and the dust-free SSP model (dotted). Track segments are colour-coded for redshift.

As mentioned above, model tracks also overlie the distributions of Figure 5.1, tracing the predicted colours of different stellar populations as a function of redshift. The plot shows that the constant SFR models (solid and dashed lines) provide a better fit to our data than the SSP model (dotted line), which deviates significantly from the main distribution of galaxies at high redshifts. This result is expected since the selection of high redshift galaxies at i′ band preferentially selects galaxies bright in the rest-frame UV. A galaxy with a constant SFR will have a larger number of young massive stars, and therefore be intrinsically brighter in the UV, than an SSP, which experiences a single burst of star-formation. SSP galaxies would therefore need to be incredibly massive to be detectable by our i′ selection limit.

Further evident in Figure 5.1 is a lack of galaxies in agreement with the dust-free constant SFR model at redshifts $z > 4.4$ (i.e. the solid black end-of-track segment). This is a selection effect which occurs because not all of our sample of g′r′i′ galaxies is detected at infrared wavelengths. Most high redshift galaxies in our sample have magnitudes close to the i′ selection limit (i.e. $24.3 < \mathrm{i}' < 24.8$). In this region of the plot, with z′ - K $\sim$ 0 and K - IR2 $\sim$ 0.5, these galaxies would have magnitudes of $\sim$ 24 in K and IR2, fainter than the detection limits of these bands. With deeper K and IR2 data, galaxies should be visible in this high redshift region of Figure 5.1.

5.2.3 Colour-Boundary Plots

To further examine the properties of objects close to the colour limit, the perpendicular distance in colour-space was calculated from each extended box object to the nominal LBG selection box boundary. With a focus on the sloped portion of the selection box (see Figure 4.7), a number of plots were produced of object colour versus distance to the boundary. Colours were based on a combination of optical and infrared data and included: i′ - IR1, z′ - IR1, K - IR2 and IR1 - IR2. As above,

scatter was reduced by plotting only those objects detected with magnitude errors less than 0.15 in IR1 and IR2, and the plots were overlain with CB2008 model tracks.

Similar to the colour-colour plots in the previous section, the choice of K - IR2 colour provided the best discrimination of objects according to redshift. The plot of K - IR2 vs boundary distance (Figure 5.2) displays a clear separation of the low redshift (K - IR2 $\lesssim 0$) and high redshift (K - IR2 $\gtrsim 0$) populations. In addition, with

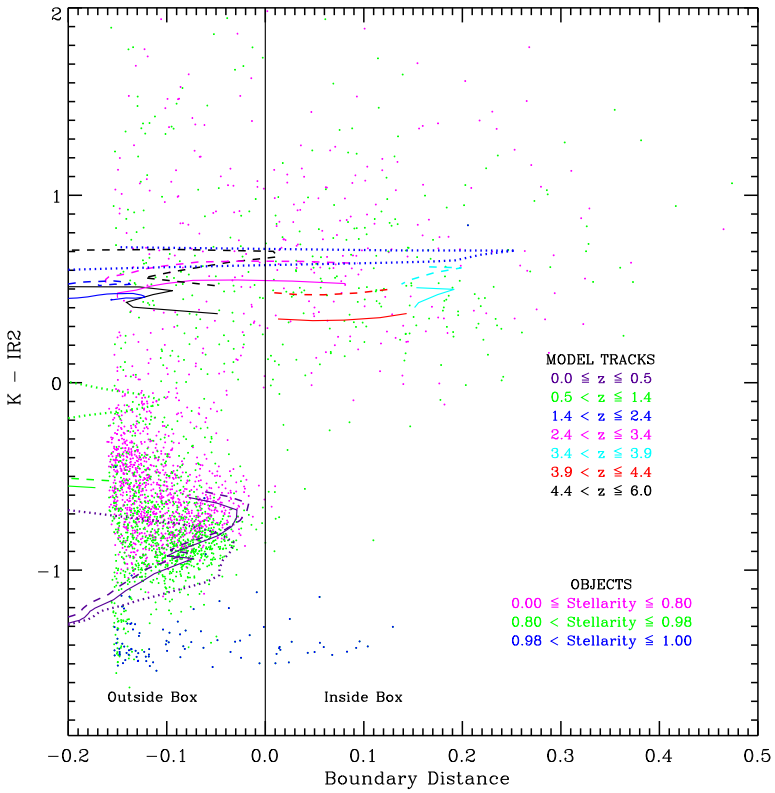

Figure 5.2: A colour-boundary plot of objects located within the extended selection box of the D2 Field. Plotted objects are limited to those with magnitude errors less than 0.15 in IR1 and IR2, colour-coded according to stellarity index. Lines represent model tracks: the dust-free constant SFR galaxy model (solid), the dust-reddened constant SFR galaxy model (dashed), and the dust-free SSP model (dotted). Track segments are colour-coded for redshift.

the boundary line as reference, the plot indicates that the main distribution of low redshift galaxies and stars is well confined to the region outside the nominal LBG selection box. Except for a small scattering of stars, there is very little contamination of the LBG sample and the objects within the nominal LBG selection box consist predominantly of high redshift galaxies at K - IR2 $\gtrsim 0$. This indicates a well-placed boundary for containing the desired population.

Several model track segments overlie the high redshift region of Figure 5.2. The constant SFR model tracks for galaxies within our target redshift range of $3.4 < z < 4.4$ (i.e. the solid and dashed lines in pale blue and red) are located precisely within the boundary of the nominal LBG selection box. Again, this suggests that the boundary is well-placed. Note that both the reddened and unreddened constant SFR models lie approximately the same distance from the colour boundary due to the boundary line being chosen with the same slope as the reddening vector in g' - r' vs. r' - i'. The nominal LBG selection box is enclosed at the high redshift end by a vertical selection limit that intersects the reddening vector at $r' - i' = 1.0$ (see Figure 4.7). This limit restricts the objects included in the selection box to those predominantly within our redshift range. Reddened LBGs at higher redshifts are, however, lost as they pass through the vertical selection line.

Some contamination of the LBG sample by galaxies at redshifts $2.4 < z < 3.4$ is also indicated by the model tracks in Figure 5.2. Placing a more restrictive boundary on the nominal LBG selection box could potentially reduce the contamination by this lower redshift population. The chosen boundary is in fact more relaxed than has been adopted in previous studies of LBGs. Hildebrandt *et al.* [22], for example, use much more restrictive colour criteria to define large high-quality samples of LBGs for the CFHTLS Deep Fields. Applying the same $g'r'i'$ selection criteria as Hildebrandt *et al.* to our LBG sample would, however, result in a loss of over 40% of our LBG candidates. A less restrictive selection region provides not only a significant increase

in the number of potential LBGs but also avoids bias in the selection process. A tighter allowed region limits the selected galaxies to those that are bluest in their UV continuum slope. These are galaxies, for example, that have low dust reddening and/or strong Lyman-α emission lines, which fall into the r′ filter at $z > 3.7$ making them bluer in (r′ - i′). In addition to the contamination, the placement of the boundary results in part of the distribution of high redshift galaxies being located outside of the nominal LBG selection box. The model tracks suggest, however, that these are galaxies in the range of $1.4 < z < 3.4$ and $z > 4.4$, which is outside the target redshift range of this study.

5.2.4 Conclusion

The above analysis confirms that the nominal LBG selection boundary is well-positioned to contain a sample of $z \sim 4$ LBGs and thus further analysis will be limited to the objects located in the nominal LBG selection box. Due to the overlap of various redshift populations, the complete elimination of low redshift interlopers within the selection region is not feasible at this stage. In the next section, model spectra are used to individually determine the properties of each target in the LBG sample. This process enables one to identify and eliminate the remaining contaminants and obtain a refined sample of LBGs within the target redshift range.

5.3 Model Fitting Procedure

Utilizing the optical selection criteria defined in Section 4.4, and excluding targets contaminated by blended photometry (Section 3.4), 1008 LBG candidates remain in the D2 Field. The properties of these candidate objects may be determined through a procedure developed to fit target object photometry to model galaxy spectra in the

following manner. The ten-point photometry of each LBG target is used to generate a low resolution spectrum. Galaxy models are customized to produce a suite of spectral templates covering a range of redshifts, ages and dust extinction levels. A chi-square minimization is applied to determine the best model fit to the photometry of each target. The properties of the best-fit model constrain the redshift, age, SFR, stellar mass and dust content of each LBG. The details of this fitting procedure are provided below.

5.3.1 Library of Model Spectra

Galaxy Spectra

CB2008 population synthesis models provide the galaxy spectra used in the fitting procedure. Based on the colour track analysis described in Section 5.2, we limited our model selection to a galaxy with constant SFR. The galaxy may evolve over a time span of one million to two billion years within which models for 34 galaxy ages were retrieved [10]. Dust extinction effects were introduced by selecting two versions of the model: dust-obscured and dust-free. Generating extinction curves based on these two models, it was possible to further differentiate the model galaxy spectra for a range of dust extinction levels (see Section 5.3.2).

Quasar and Star Spectra

There is an expectation for some contamination of our LBG sample by high redshift quasars and local stars. The model selection therefore includes not only the spectra of galaxies but those of stars and quasars as well. A quasar composite spectrum, based on the unpublished work of Daniel Fryer and Chris Willott, provided the SED of a typical quasar [54]. Original data for the spectrum were collected by the Sloan Digital Sky Survey (SDSS) [99]. Star spectra were obtained from the Model

Atmosphere Radiative and Convective Scheme (MARCS), a library containing $\sim$ 30,000 stellar atmospheric models [21] [98]. The models consist of photospheric fluxes measured at $\sim$ 100,000 select wavelength positions between 1300 Å and 20 μm. Model spectra are available for stars with effective temperatures ranging from 2500 K to 8000 K, in increments of 100 K below 4000 K and 250 K otherwise. We determined the temperatures of stars most likely to be scattered into our z $\sim$ 4 LBG selection box by examining colour-colour diagrams of model stars as a function of temperature. Accordingly, the spectra of stars with temperatures of 3900 K, 4000 K and 4250 K were included in our suite of models.

5.3.2 Customization of Model Spectra

Redshift

Model galaxy and quasar spectra, produced in the rest-frame, were redshifted into the observed-frame according to [11]:

$$\lambda_{obs} = \lambda_{rest}(1 + z), \tag{5.1}$$

where λ_{obs} is the observed wavelength of the radiation for an object at redshift z and λ_{rest} is the rest-frame wavelength of the radiation, as detailed in Section 2.1. Redshifts ranging from $z = 0$ to $z = 5$, in increments of 0.1, were selected for the models, for a total of 51 possible redshifts.

Dust Extinction

The extinction and reddening of light from galaxies is caused by interstellar dust that is typically found in star-forming environments (see Section 3.1.7). The result is a decrease in a galaxy's observed flux (L). This change in flux (dL) depends on

the distance (dr) the radiation has traveled from the source and is given by [26]:

$$dL = -\alpha L dr = -L d\tau, \tag{5.2}$$

where α [m^{-1}] is the opacity, a measure of the effectiveness of the dust in diminishing radiation, and τ [dimensionless] is the optical depth of dust extinction, such that $d\tau = \alpha dr$. Integrating Equation 5.2 from the source (at L = L_0 and $\tau = 0$) to the observer gives:

$$L = L_0 e^{-\tau}. \tag{5.3}$$

Thus, observed flux decreases exponentially with increasing optical depth. In terms of magnitude, the extinction (A) is given by [26]:

$$A = -2.5 log_{10}(e^{-\tau}) = 1.086\tau. \tag{5.4}$$

Therefore, optical depth corresponds approximately to the number of magnitudes of extinction.

Model galaxy spectra were modified to include the effects of flux attenuation due to the presence of interstellar dust. In many previous studies, this involved the application of a basic foreground dust extinction curve that assumes each galaxy is obscured behind a uniform screen of dust (e.g. [41] [56]). In a more sophisticated approach, we account for dust extinction using the method of CB2008 that considers the specific interstellar environments of different stars. For instance, young massive stars are more likely to be surrounded by dust than old low mass stars [100]. To implement this approach, pairs of dusty and dust-free models are compared and an extinction law is determined for each type of galaxy.

Specifically, modifications are based on two versions of the constant SFR galaxy model: one unreddened with optical depth $\tau = 0$; the other reddened at optical depth $\tau = 1$, where tau is defined at wavelength 0.55 μm (V band). For each model, a spectrum was produced for each of the 34 galaxy ages chosen for the suite. Dust

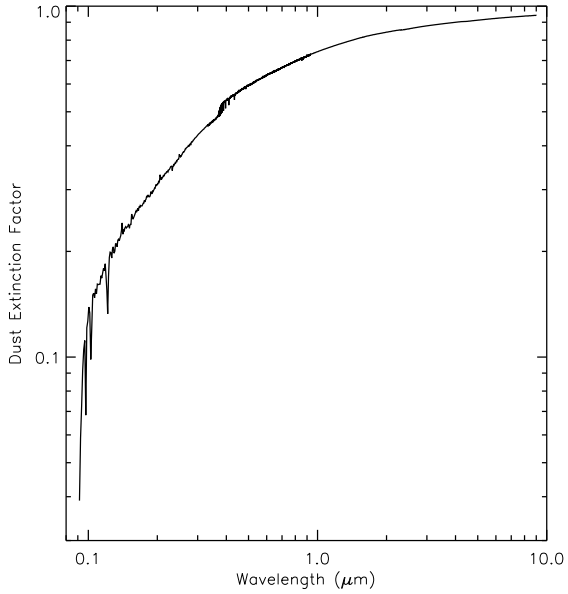

Figure 5.3: An example of a typical dust extinction curve.

extinction curves were generated by dividing each dust-bound model by the dust-free model of the same age. Each dust-free model was then convolved with its associated extinction curve raised to various powers of tau (τ) (i.e. different optical depths), producing galaxy spectra with a range of dust extinction levels. For the model suite, spectra incorporating the effects of dust extinction were generated at each model age for 21 optical depths ranging from $\tau = 0$ to $\tau = 2$. A typical dust extinction curve is shown in Figure 5.3.

Lyman Forest Absorption

Introduced in Chapter 2, the term Lyman forest refers to the dense region of hydrogen absorption lines present at rest-frame UV wavelengths in the spectra of high redshift galaxies and quasars [11]. The high concentration of absorption lines results primarily from hydrogen transitions in the IGM at various stages along the

path from source to observer. Multiple sets of absorption lines are present, each redshifted according to the position along the path at which the transitions occurred.

Specifically, Lyman forest transitions lead to the partial absorption of wavelengths between 912 Å and 1216 Å in the rest-frame [11] [45] [55]. Below the Lyman limit of 912 Å, hydrogen ionization causes essentially complete wavelength absorption. The strength of Lyman forest absorption is, however, redshift dependent. Partial absorption becomes more complete at increasingly higher redshifts. For redshifts $z > 6$, this is attributed to the high fraction of neutral hydrogen present in the early universe [29]. At redshifts $z < 6$, the increase in Lyman forest absorption with redshift is due to more efficient absorption in a smaller, denser early universe. As the number of ionizing sources increased, the ionizing intensity became more uniform so that by $z \sim 5$ most of the hydrogen in the Universe was ionized. However, pockets of high-density, self-shielded neutral hydrogen remained, giving rise to Lyman forest absorption at lower redshifts. The IGM remains today in a highly ionized state. Hence, the Lyman forest feature is prominent only in high redshift galaxy and quasar spectra.

The analysis of high resolution quasar spectra has been used to quantify Lyman forest transmission at high redshifts [45] [46]. The results are based in part on a study by Songaila [45] of 50 quasars with redshifts between $z = 2$ and $z = 6.3$. Formulae have also been developed for the mean Lyman-α and Lyman-β transmissions as a function of redshift. These formulae incorporate the approximation of a uniformly evolving hydrogen ionization rate, which reflects the observed trends in transmission with redshift.

Willott *et al.* [55] have previously determined the Lyman forest IGM transmission curves for quasars of the Canada-France High-z Quasar Survey (CFHQS), using the data and equations of Songaila [45]. Accordingly, the Lyman-β transmission function is determined from the Lyman-α transmission using the optical depth relation: τ_β

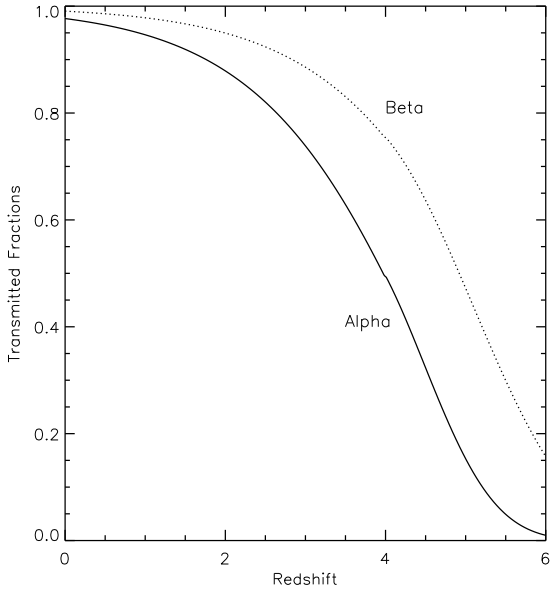

Figure 5.4: Lyman-α and Lyman-β transmission as a function of redshift [55].

$= 0.4\ \tau_\alpha$. Smaller, higher order absorption effects, such as Lyman-γ, are neglected, and all flux below the Lyman limit of 912 Å is set to zero. Figure 5.4 displays the Lyman-α and Lyman-β transmitted fractions as a function of redshift.

Following the prescription of Willott *et al.* [55], we developed a routine to calculate the mean transmission spectrum shortward of 1216 Å in the rest-frame, specific to each model's redshift. Each transmission spectrum was computed as the convolution of the redshift-dependent Lyman-α (rest-frame 912 Å to 1216 Å) and Lyman-β (rest-frame 912 Å to 1026 Å) transmission curves. Each redshifted model galaxy and quasar spectrum was then convolved with the mean transmission spectrum at that redshift. The resultant model spectra incorporate the effects of Lyman forest absorption. Figure 5.5 provides an example of a model spectrum before and after corrections for dust extinction and Lyman forest absorption.

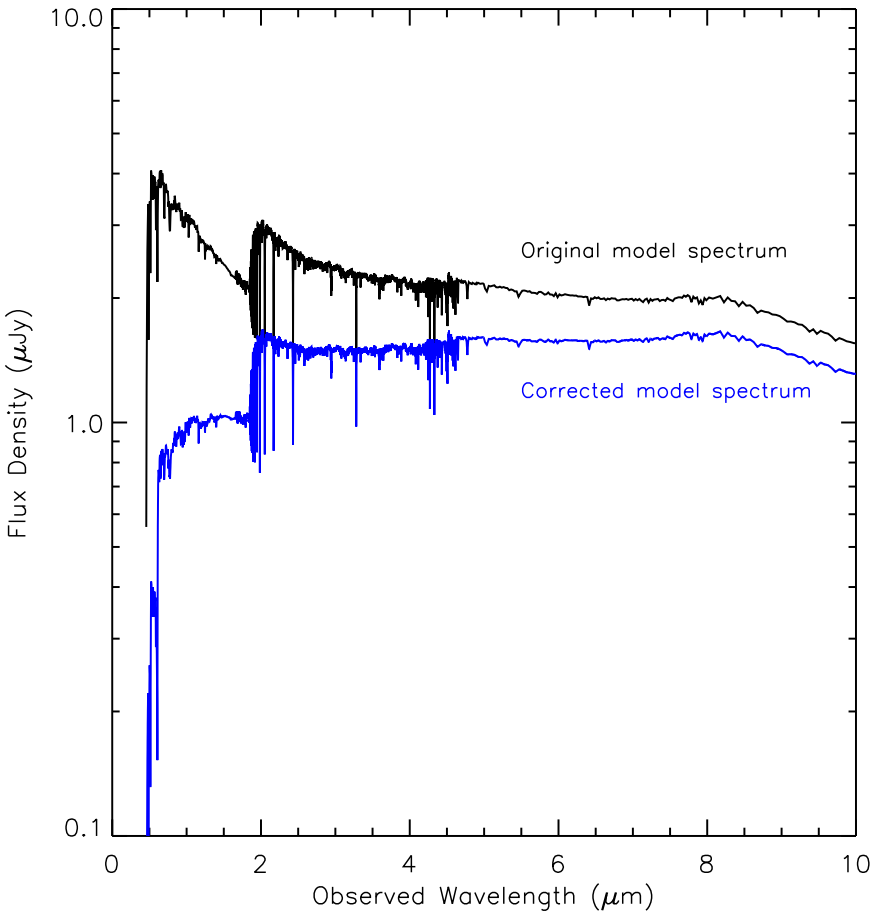

Figure 5.5: A sample model spectrum before and after corrections for dust extinction and Lyman forest absorption. The model is of a galaxy ~ 100 million years old at a redshift $z = 4$, with an optical depth of dust extinction $\tau = 1$.

Model type	Number of models	Parameters
Galaxy	36414	age, redshift, dust extinction
Quasar	51	redshift
Star	3	effective temperature

Table 5.1: Components of the model suite.

Parameter	Range	Number of levels
Age	1×10^6 to 2×10^9 years	34
Redshift	$z = 0$ to $z = 5$	51
Dust extinction	$\tau = 0$ to $\tau = 2$	21

Table 5.2: Summary of the parameters used to define the galaxy model suite.

5.3.3 Model Suite

In summary, the library of models for the fitting routine consists of a constant SFR galaxy spectrum produced at 34 different ages, a single quasar spectrum, and three star spectra. After introducing 51 possible redshifts and 21 levels of dust extinction, we obtain a model suite consisting of 36,468 models. These models serve as the templates for the galaxy fitting routine. A breakdown of the model suite and its parameters is provided in Tables 5.1 and 5.2, respectively.

5.3.4 Fitting Routine

Photometric Data Points

For the purposes of the fitting routine, the following changes were made to the photometric data of the LBG sample:

1. A minimum uncertainty of 0.10 mag, accounting for systematic uncertainties,

was assigned to all photometry measurements. Establishing a minimum level of uncertainty ensures that the chi-square values generated by the fitting routine are not dominated by filters with very small errors [34].

2. The magnitudes in all filters were converted to flux densities via [84]:

$$m_{AB} = -2.5log_{10}(F) + 23.9, \tag{5.5}$$

where m_{AB} is the AB magnitude and F is the flux density in μJy. Note that the magnitude-to-flux conversion results in asymmetric uncertainties. To maintain a single value per filter in the chi-square calculation, we retained the more conservative (i.e. positive) flux uncertainties.

3. The fluxes for non-detections (i.e. objects fainter than the 2σ-optical or 3σ-infrared limiting magnitude) were set to zero.

Model Data Points

Each model in the suite is in the form of a high resolution spectrum, whereas the spectrum of each LBG candidate is defined by ten photometric data points representing the observed flux density in each of the ten filters. To compare the photometric and model data, it was necessary to determine the corresponding filter flux densities of each model. To do this, each model spectrum was convolved with the individual filter transmission profiles (Figures 3.2 and 3.6) and integrated. This provided the (non-normalized) mean flux densities of the model templates in each of the filters.

The flux densities of each model template were initially normalized using the sum of the filter flux densities of the galaxy being fitted. This normalization was adequate except where galaxies had filters with non-detections. In such cases, using zero photometric flux with positive model flux caused too low a normalization for the

model. Instead, normalization was treated as a free parameter in the chi-square minimization with the initial normalization used as the starting value. The normalized mean flux densities of the model template served as the model data points.

Determining the Best Fit

The best-fit model to each LBG was determined through minimization of a reduced chi-square test statistic [34] [53]. The chi-square test statistic is defined as [31] [36]:

$$\chi^2 = \sum_i \frac{1}{{\sigma_i}^2} \left(f_i \left(p_1 ... p_n \right) - x_i \right)^2 , \tag{5.6}$$

where

- the index i represents the number of data points (i.e. filters)
- f_i and x_i are, respectively, the model template and observed flux densities in a given filter
- σ_i is the uncertainty in the observed flux density
- p_n are the n adjustable parameters of the fit.

We define the reduced χ^2 statistic as the χ^2 per degree of freedom, where the number of degrees of freedom is the number of data points minus the number of adjustable parameters of the fit [34] [36]. Galaxy models, a function of redshift, age, dust extinction and normalization, have six degrees of freedom. Quasar models, a function of redshift and normalization only, have eight degrees of freedom. Star models, adjusted only for normalization, have nine degrees of freedom. As in standard Gaussian statistics, a reduced χ^2 value near 1 is indicative of a good fit.

A reduced χ^2 statistic was generated by comparing the observed flux density per filter of each LBG candidate to each model template. The model with the smallest

reduced χ^2 statistic is the best-fit model to the target object photometry. The model fit separates the galaxies, quasars and stars within the sample and, for each LBG identified, the model parameters provide estimates of the galaxy's redshift, age and dust content.

Star Formation Rate and Stellar Mass

The normalization value of the best-fit model to each LBG was used to determine the SFR and stellar mass of the observed galaxy [34]. Calculations were based on the SFR versus redshift relationship for non-normalized CB2008 model data (Figure 5.6). Note that Figure 5.6 is used purely to relate the flux of a model galaxy with its SFR and does not represent an actual physical relationship. The increase in SFR with redshift indicated by the curve results primarily because of the rela-

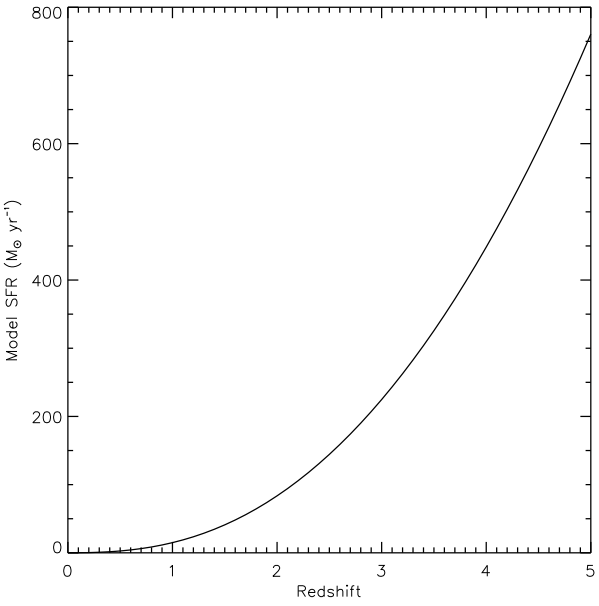

Figure 5.6: Plot of model SFR versus redshift for non-normalized CB2008 models, used to determine the SFRs and stellar masses of model-fitted galaxies [10] [54].

tionship between luminosity and redshift for a constant flux. Specifically, given two equivalently luminous galaxies, the galaxy at higher redshift will appear to be dimmer and therefore to have a lower luminosity-based SFR than it actually has. This effect is accounted for by allowing a higher SFR for a galaxy if it is located at a higher redshift. The curve also incorporates corrections that allow filter-based fluxes from galaxies at different redshifts to be directly compared (i.e. K-corrections) [11]. That is, different wavelength regions of a galaxy's spectrum are redshifted into the same filter for galaxies at different redshifts and this must be compensated for if the measurements are to be compared.

Using the redshift of the best-fit model, the model's SFR was determined from the function (Figure 5.6). The SFR (in $M_\odot$ per year) of the observed galaxy was calculated by multiplying the model's SFR by the normalization factor. The total stellar mass (in $M_\odot$) of the galaxy was estimated as the SFR multiplied by the best-fit age of the galaxy.

5.3.5 Model Fits

Model fits were generated for the 1008 candidate LBGs in the D2 Field and indicate that the selection contains 807 galaxies, 99 quasars and 102 stars. Of the 807 galaxies, 705 have best-fit redshifts within the target range $3.4 \leq z \leq 4.4$. Outside the target range are 58 and 44 galaxies at lower and higher redshifts, respectively. The magnitude and stellarity properties of each class of model fit are provided in Table 5.3. Stars and quasars within the sample are expected to be generally brighter and to have higher stellarity values than galaxies. While this is the case for stars, Table 5.3 indicates that the galaxies and quasars are of approximately the same mean i' magnitude and mean stellarity. This may be explained if a number of quasars have been misidentified as galaxies by the fitting routine or, likewise, if galaxies have been

Model fit	Mean i′ magnitude	Mean stellarity
Galaxy	24.37	0.73
Quasar	24.40	0.73
Star	23.91	0.77

Table 5.3: Magnitude and stellarity properties of model fits by class.

misidentified as quasars. This may reasonably occur given that galaxy and quasar SEDs are similar and this issue will be further addressed in Chapter 6.

A sample selection of galaxy model fits is provided in Figure 5.7, arranged according to best-fit model age. Each plot shows the best-fit model spectrum in the observed frame overlain with the photometric and model data points. Photometric data points are placed at the center of the filter bins, displaying horizontal error bars representing the width of the filter at half the maximum transmission (FWHM). Data describing the parameters of the fit are provided in a legend. All shown examples are of galaxies with reasonably good fits in the target redshift range and various levels of dust extinction.

The main features of the galaxy spectra in Figure 5.7 include the Lyman-break at approximately g′ band (~ 0.48 μm) and the Lyman forest between approximately g′ band and r′ band (~ 0.63 μm). The Balmer/4000Å break (see Section 4.2.1) is also evident in the spectra at approximately K band (~ 2.2 μm). A comparison of the different galaxy SEDs reveals the influence of age and dust on the shape of a galaxy's spectrum. Image cutouts of these galaxies at each of the filter wavelengths are provided in Figure 5.8.

From the best-fit model results, a number of contaminant objects have been identified in the LBG selection, including stars, quasars and low redshift galaxies. A typical spectrum of each of these interlopers is provided in Figure 5.9. Each

interloper has a distinct SED that clearly differentiates it from that of a high redshift galaxy. In particular, low redshift galaxies appear brighter (i.e. redder) than high redshift galaxies at near- and mid-infrared wavelengths (see Section 5.2.2) [53]. The addition of K band and IRAC data has therefore improved the ability to identify and eliminate low redshift galaxies within the high redshift sample.

We note that some of the reduced χ^2 values ($\gg 1$) indicate that LBG selection contains galaxy fits that are poor. In Chapter 6, the model fits of LBGs within the target redshift range will be further analyzed for goodness-of-fit, robustness of photometry, misidentification, and agreement with data from other surveys. The goal is to identify a population of $z \sim 4$ LBGs for which we have accurately constrained the redshift, age, dust content, SFR and total stellar mass of each galaxy.

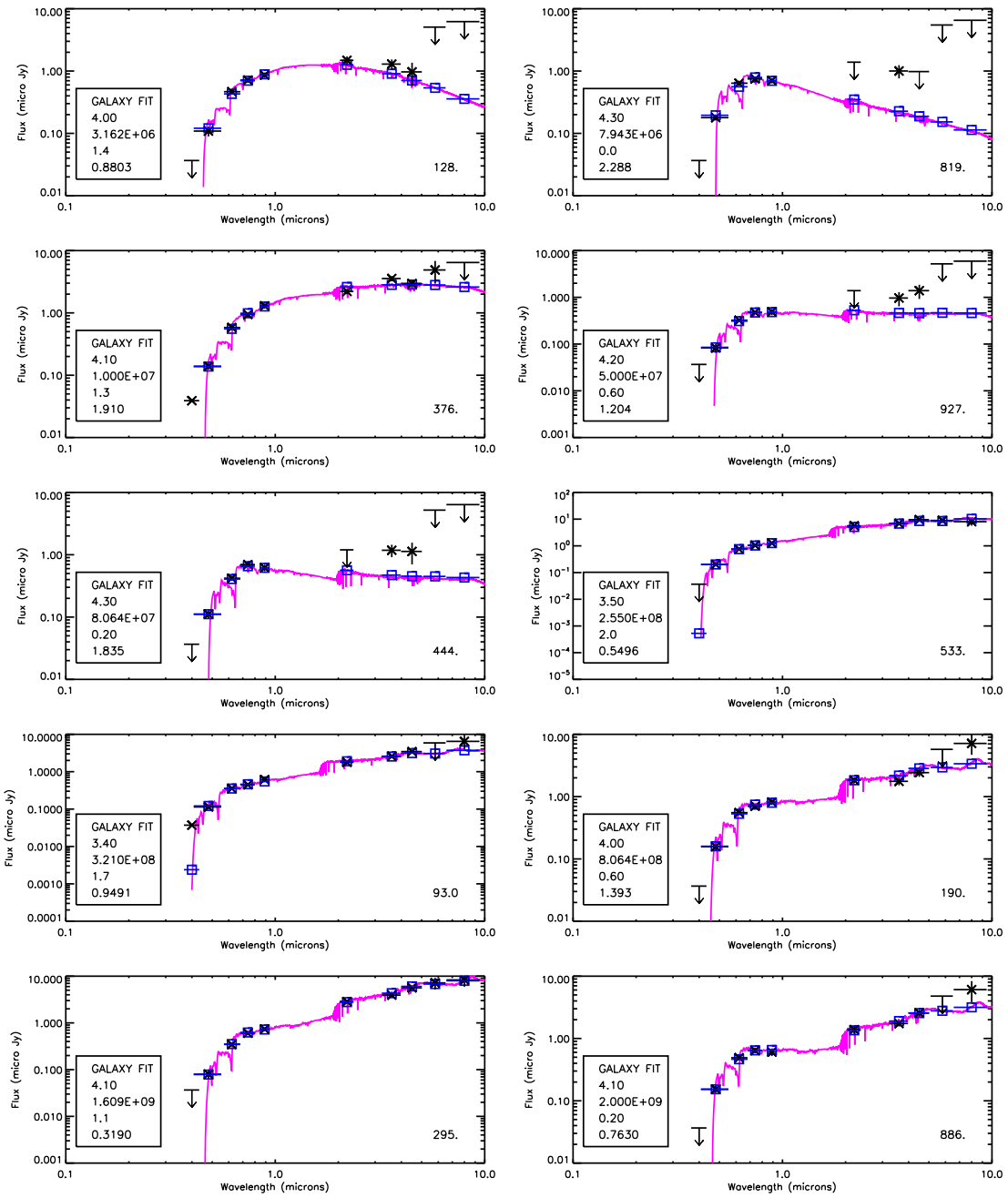

Figure 5.7: A sample selection of galaxy model fits. In each plot, the best-fit model spectrum is shown in purple, overlain by model data points (blue squares) and photometric data points (black crosses). An upper limit is given for photometric non-detections (black arrows). Each legend lists, from top to bottom, the redshift, age in years, optical depth and reduced χ^2 value of the best-fit model. The spectra are arranged in order of increasing age (across each row and down) and identified by an index number in the lower right. Image cutouts of these galaxies are provided in Figure 5.8.

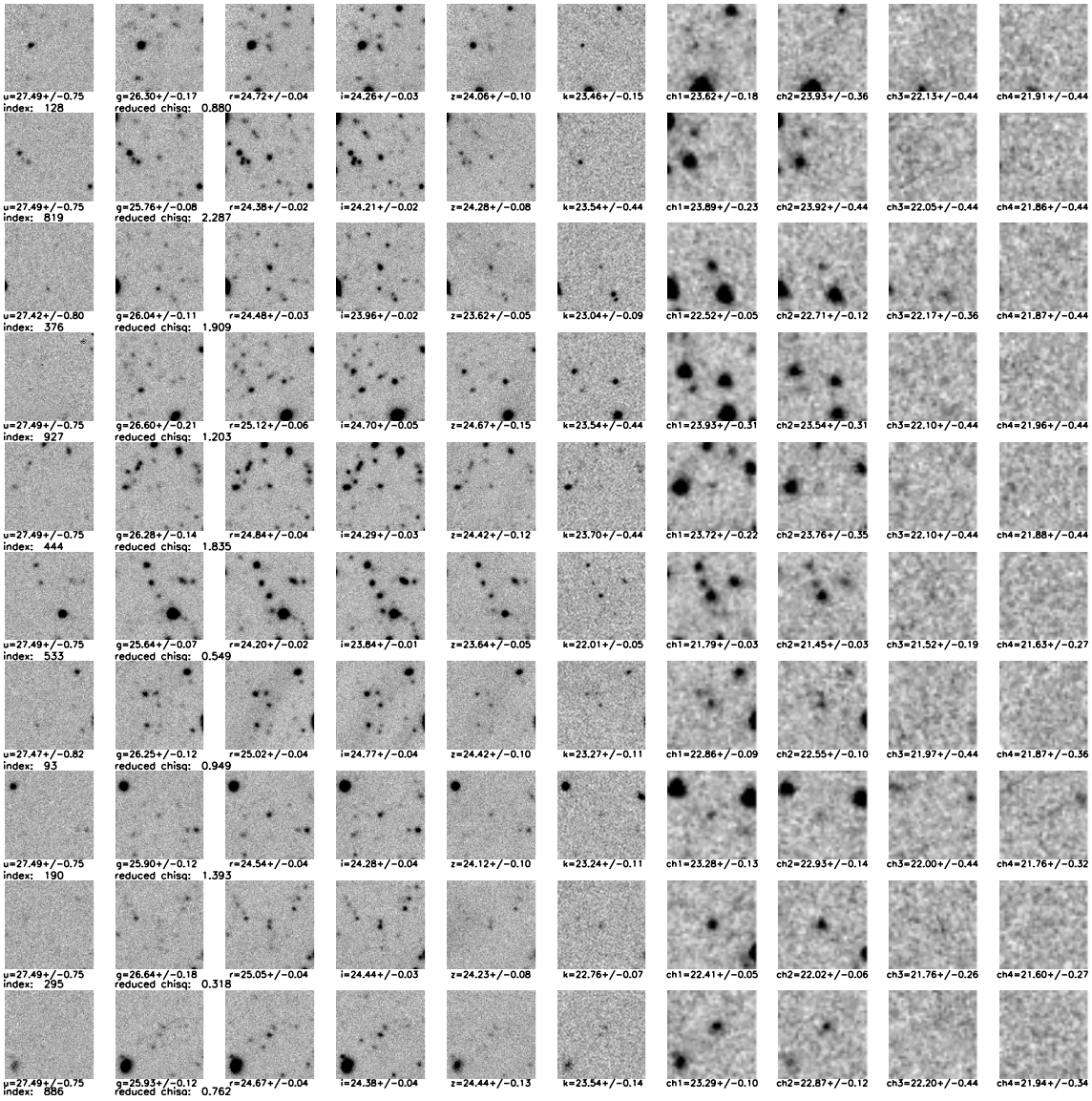

Figure 5.8: Image cutouts of the LBGs fitted in Figure 5.7. Each row shows the image of a galaxy in each of the ten filters. Galaxies are arranged according to the ordering in Figure 5.7, increasing in age (from top to bottom), and identified by an index number. The target galaxy is centered in each cutout and the magnitude of the galaxy in the filter is provided below the cutout.

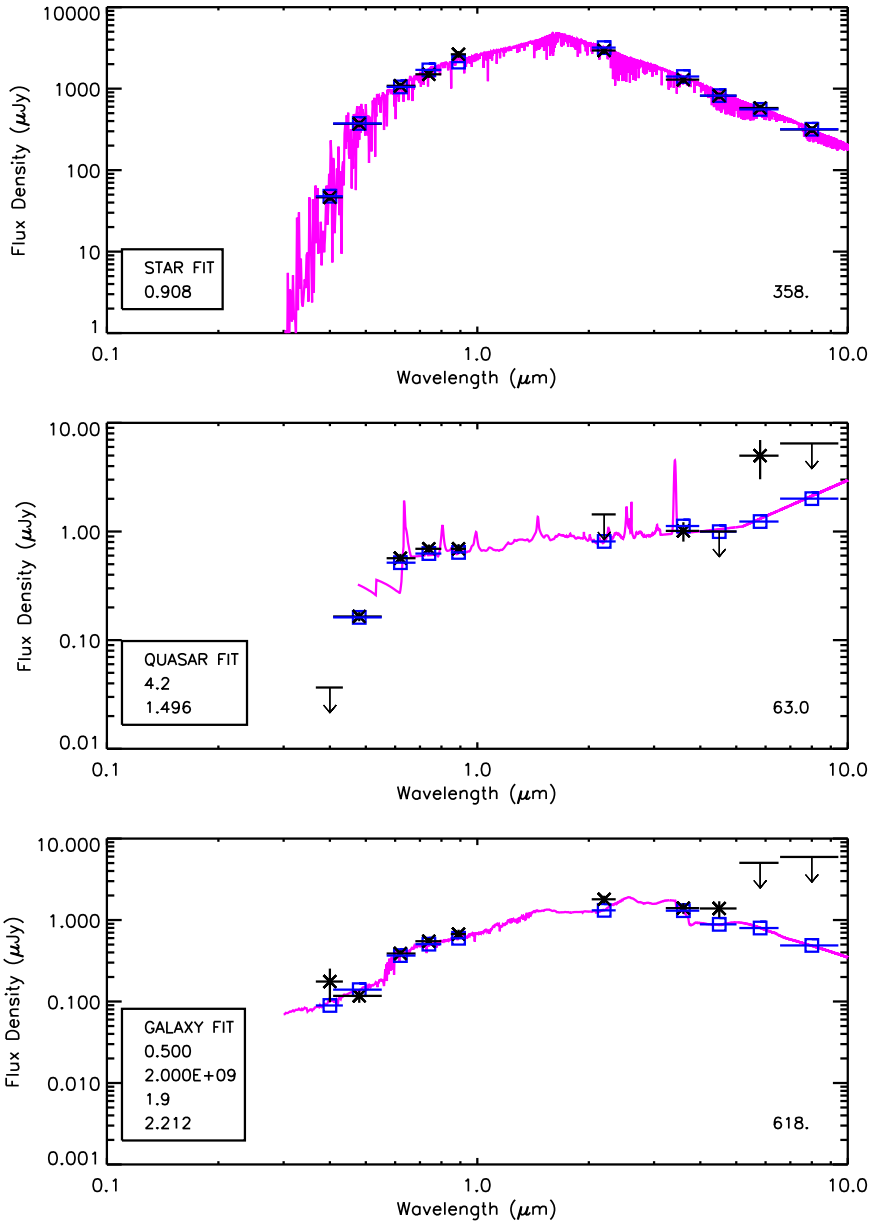

Figure 5.9: Typical spectra of interlopers within the LGB selection. In each plot, the best-fit model spectrum is shown in purple, overlain by model data points (blue squares) and photometric data points (black crosses). An upper limit is given for photometric non-detections (black arrows). The galaxy legend lists, from top to bottom, the redshift, age in years, optical depth and reduced χ^2 value of the best-fit model. The quasar legend lists only redshift and reduced χ^2 value, and the star legend, only reduced χ^2 value.

Chapter 6

The Galaxy Population

As described in Chapter 5, the model fitting routine categorized each of the 1008 LBG candidates in the D2 Field, identifying 807 galaxies, 99 quasars and 102 stars. Each galaxy has an estimated redshift, age, dust extinction, stellar mass and star formation rate. Of the 807 identified galaxies, 102 have best-fit redshifts outside the target range $3.4 \leq z \leq 4.4$, with 58 and 44 galaxies at lower and higher redshifts, respectively. Candidates identified as a quasar or star by the best-fit model were removed from the data set, as well as the galaxies with a best-fit redshift outside our target redshift range.

Any remaining contamination of the sample would be problematic for our LBG survey. At $\sim 7\%$, our fraction of contaminating low redshift galaxies is considered low in comparison to other studies and is discussed in the first section of this chapter. The remainder of the chapter details the final isolation of a robust set of LBGs from the 705 remaining candidates, and the determination of the physical properties of this population.

6.1 Contamination Rates

In general, LBG surveys employ a number of approaches to isolate high redshift galaxies, while avoiding significant contamination by stars and low redshift galaxies. These may include colour selection, brightness, stellarity, S/N limits for detections and non-detections in certain filters, photometric and/or spectroscopic redshifts, when available. Programs of spectroscopic follow-up and deeper imaging can be used to confirm the effectiveness of the selection criteria and, where this is limited, simulations can be used to estimate rates of contamination.

Despite the advantages presented by spectroscopic data, its use is still limited. In a study similar to ours, Steidel *et al.* [49] successfully obtained spectroscopic redshifts for 60 of 207 galaxies in a $z \sim 4$ GRI photometric sample to $I_{AB} = 25.0$. The estimated rate of contamination by low redshift galaxies was $\sim 20\%$ (12/60), with the limited number of spectroscopic confirmations ($\sim 29\%$) attributed to their low S/N data at this redshift. Vanzella *et al.* [52] performed an extensive program of spectroscopic follow-up of LBGs at redshifts $z > 4$, selected from optical images obtained as part of the Great Observatories Origins Deep Survey (GOODS) [101]. From the 85 BVz-selected ($z > 3.1$) LBG candidates, they were able to secure redshifts for 48 sources ($\sim 56\%$), finding one star and one low redshift interloper, or 96% efficiency. It was, however, difficult to determine redshifts for faint sources and galaxies at $z < 3.6$ (due to experimental set-up), suggesting that the rate of contamination may be much higher. In our study, we circumvent this issue by utilizing near-infrared and mid-infrared data. Contaminating low redshift galaxies and stars have unique colours into the infrared making it possible to identify and eliminate them from the data set without the need for spectroscopy.

Most contamination of colour-selected samples is in the form of low redshift galaxies scattered into the selection region due to photometric errors [15]. Eyles *et*

al. [15] required robust detections of $z' < 26.9$ ($> 8\sigma$) to ensure secure ($i' - z'$) colours for their study of $z \sim 6$ LBGs, citing high contamination rates for studies using lower S/N detections. In a comparable study, based on 5σ detections of $z \sim 6$ LBGs in the GOODS fields, the contamination rate of $\sim 45\%$, suggested by simulations, was increased to $\sim 75\%$ on deeper imaging [7] [13]. The high S/N of our optical data ensures that the colours of our sample are very well determined, thus minimizing the contamination due to photometric errors.

As mentioned in Section 5.2.3, Hildbrandt *et al.* [22] used conservative colour selection criteria to define their CFHTLS Deep Fields LBG samples. A low contamination rate was crucial to their study and thus high efficiency colour-space regions were selected and the contamination rates estimated based on simulations. Using model templates and observational data, they generated an extensive colour catalogue to effectively emulate populations of high redshift galaxies, low redshift interlopers and stars. Contamination rates were then based on the number of randomly selected objects meeting their colour selection criteria. Consequently, they applied a brightness cut of $i' > 23.5$ that reduced their $g'r'i'$ selection by $\sim 52\%$, with the remaining contamination from stars and low redshift interlopers estimated at less than 10%. In comparison, in the absence of infrared data, most of the star contamination of our sample, $\sim 10\%$ (102/1008), could be similarly eliminated by brightness and/or stellarity cuts. For instance, selecting either stellarity < 0.90 or $i' > 24.5$ removes $\sim 55\%$ of the stars and $\sim 40\%$ - 50% of the low redshift galaxies within our sample, with the magnitude cut being slightly more effective. However, our potential candidate selection would also be reduced by $\sim 50\%$ using this method and, therefore, the use of infrared data is clearly advantageous. Further, given the magnitude-dependence of contamination rates (see e.g. [25] [42]), the exceptional quality and depth of the CFHTLS-Deep data results in higher counts of LBGs at fainter magnitudes and thus an overall enhanced ability to generate large samples of

high redshift galaxies with low rates of contamination.

6.2 Refining the Data Set

6.2.1 Cross-Correlations

As mentioned in Chapter 3, the CFHTLS D2 Field is located within the two square-degree Cosmic Evolution Survey (COSMOS) region [73]. COSMOS is a Hubble Space Telescope (HST) Treasury Project and the largest of the HST surveys, potentially capable of detecting over 35,000 LBGs. The equatorially-located COSMOS field is accessible to telescopes world-wide and has the best multi-wavelength coverage available on a square-degree scale. Observations include HST optical [74], Spitzer far-infrared [75] and XMM-Newton x-ray satellite imaging [76], as well as optical spectroscopic data from a number of ground-based observatories.

Data collected of the D2 Field by these COSMOS-based surveys provide an opportunity to test the robustness of our model fitting routine. Spectroscopic redshifts and imaging at various wavelengths are used to confirm the galaxy classifications and photometric redshift estimates. The cross-correlation and analysis of target objects was performed by Chris Willott [54] using several source catalogues and is detailed in this section.

Spectroscopic Sources

The COSMOS project includes contributions by several ground-based observatories. Optical spectra of 678 COSMOS targets have been obtained using the IMACS multi-object spectrograph mounted on the Magellan Telescope, Las Campanas Observatory, Chile [77]. Observations for the zCOSMOS redshift survey were carried out with the VIMOS spectrograph of the Very Large Telescope (VLT), Cerro Paranal

Observatory, Chile [78]. The zCOSMOS-*bright* portion of the survey has obtained spectra of more than 10,000 galaxies at redshifts $z \leq 1.2$.

The above spectroscopic surveys target bright objects such that any survey objects matching those in our data set are likely to be quasars or low redshift galaxies, rather than fainter galaxies at high redshift. Two broad line spectrum quasars and two low redshift galaxies were found to match with the survey objects. If we compare the results of the model fitting routine for these objects, the quasars, at spectroscopic redshifts of 3.35 and 4.25, had been fitted as galaxy spectra with photometric redshifts of 4.0 and 4.4, respectively; the low redshift galaxies ($z < 0.4$) had been fitted by galaxy spectra with photometric redshifts of 3.1 and 3.3, just outside the target redshift range.

It should be noted that the model suite is limited to a single quasar spectrum at 51 different redshifts, but contains a total of 36,414 galaxy templates. Given the high number of free parameters for the galaxy fitting, and the differences that exist between quasar spectra, it is expected some quasars may be incorrectly classified. It is therefore possible that other quasars may be part of our galaxy sample. Since the misidentified quasars in our sample were mistaken for galaxies at the high end of the age and mass scales, one may expect the inclusion of such fits to bias estimates of the old, massive galaxy population at high redshift. Similarly, the incorrect classification of low redshift galaxies suggests that other low redshift interlopers may be present within the galaxy sample, contaminating the low redshift end of our selection.

X-Ray Sources

Cross-correlations were also performed using the XMM-Newton X-ray Point-like source catalogue (XMM-COSMOS) [76]. XMM-Newton is a space-based multi-mirror x-ray telescope operated by the European Space Agency (ESA). X-rays pro-

vide a window to the high-energy universe, such that x-ray sources in the data set are likely to be high redshift quasars or low redshift galaxies with active galactic nuclei (AGN). Five high redshift sources were found: the two quasars already discussed above and three additional quasars.

24 Micron Sources

The Multiband Imaging Photometer for Spitzer (MIPS) provides deep imaging at the far-infrared wavelengths of 24, 70 and 160 μm, extending the reach of the IRAC Spitzer data [75]. MIPS is sensitive to dust-obscured radiation in regions of star formation and AGN since energy released from these sources is absorbed by surrounding dust and re-emitted in the far-infrared. Therefore, 24 μm emission may reveal luminous, dust-obscured sources not visible at optical or x-ray wavelengths. However, MIPS imaging is not sensitive enough to detect star-forming galaxies within our redshift range and any MIPS sources in our data set are likely to be dusty, low redshift starburst galaxies or high redshift AGN. Cross-correlation studies revealed seven sources: two of the x-ray quasars and one of the low redshift galaxies found previously, as well as four new high redshift AGN.

Point-Like Sources

The COSMOS deep imaging survey takes advantage of the Advanced Camera for Surveys (ACS), a third-generation HST instrument [74]. The ACS consists of three types of cameras: wide field, high resolution, and solar-blind. The SDSS filters g', r', i' and z' are a part of the ACS filter set [99]. To further eliminate likely high redshift quasars in our data set, a cut was applied for bright sources of high stellarity index. Like stars, quasars are spatially unresolved and appear point-like. Eighteen objects with $i' < 23.5$ and stellarity > 0.90 were found. In addition, high resolution HST

ACS images were used to locate fainter quasar candidates. A list of twenty-nine new i′ band point-like sources was compiled. In summary, a total of 53 cross-correlated and point-like objects were identified and ultimately removed from the LBG data set.

6.2.2 Goodness-of-Fit

Press *et al.* (1994) state that, "...unless the goodness-of-fit is credible, the whole estimation of parameters is suspect [36]." Therefore, as a further refinement of the candidate selection, galaxies with poor model fits were removed from the data set. While a reduced χ^2 of ~ 1 is considered a good fit, the majority of our fits had a reduced χ^2 value greater than this. The chance probability (Q) of finding a χ^2 value larger than a certain limit can be determined from the χ^2 distribution for a specific number of degrees of freedom. If Q is small, it is unlikely that a poor fit is due to some random influence, and instead may be attributed to an inadequate model, underestimation of measurement errors, or measurement errors that deviate from a normal distribution. General practice finds acceptable model fits for $Q > 0.001$. With six degrees of freedom for galaxy fits, and a confidence level (P, where P=1-Q) of 0.999, a maximum χ^2 value of 22.458 is considered an acceptable fit [8]. The corresponding reduced χ^2 value of 3.743 was used as the cut-off limit for acceptable fits, effectively removing 232 galaxies from the data set.

6.2.3 Poor Quality Images

For each galaxy, image cutouts were generated in all ten filter bands and each set of images was visually inspected. An additional 42 galaxies, with image quality affected by artifacts such as non-uniform background, regions of distortion, missing data, diffraction spikes, and flux contamination from nearby sources, were removed

from the data set. This was a necessary step since a poor χ^2 fit is not a definitive indicator of damaged images due to the higher level of background noise in these cases.

6.2.4 Summary

Of the 1008 objects initially selected for the fitting procedure, 705 potential LBG candidates were identified within the target redshift range. This set of candidates was further refined based on cross-correlated data, goodness-of-fit and quality of photometry. The final data set comprises a refined sample of 378 $z \sim 4$ LBGs.

6.3 Ensemble Test

6.3.1 Rationale

A study of the galaxy spectra identified some important issues with respect to the model fits. Despite allowing the normalization to vary as a free parameter in the fit, the best-fit models for a number of galaxies were predicting IRAC 3.6 and 4.5 μm fluxes well below those actually observed. Similarly, galaxies without K and IRAC detections were being fitted to very young models (on the order of $\sim 10^6$ years), which again were lying well below the detection limits (Figure 6.1). In such cases, the fits were most likely driven by the optical data components. However, the extension of optical data to infrared wavelengths is integral in differentiating between divergent models [34]. It is therefore possible that other models may provide fits to the data with χ^2 values that are similar to that of the best-fitting model. Alternatively, fit quality may be an artifact of the simple constant SFR galaxy models considered. More complex star formation histories, such as those involving multiple episodes of star formation, are likely for some galaxies and would require a different set of models

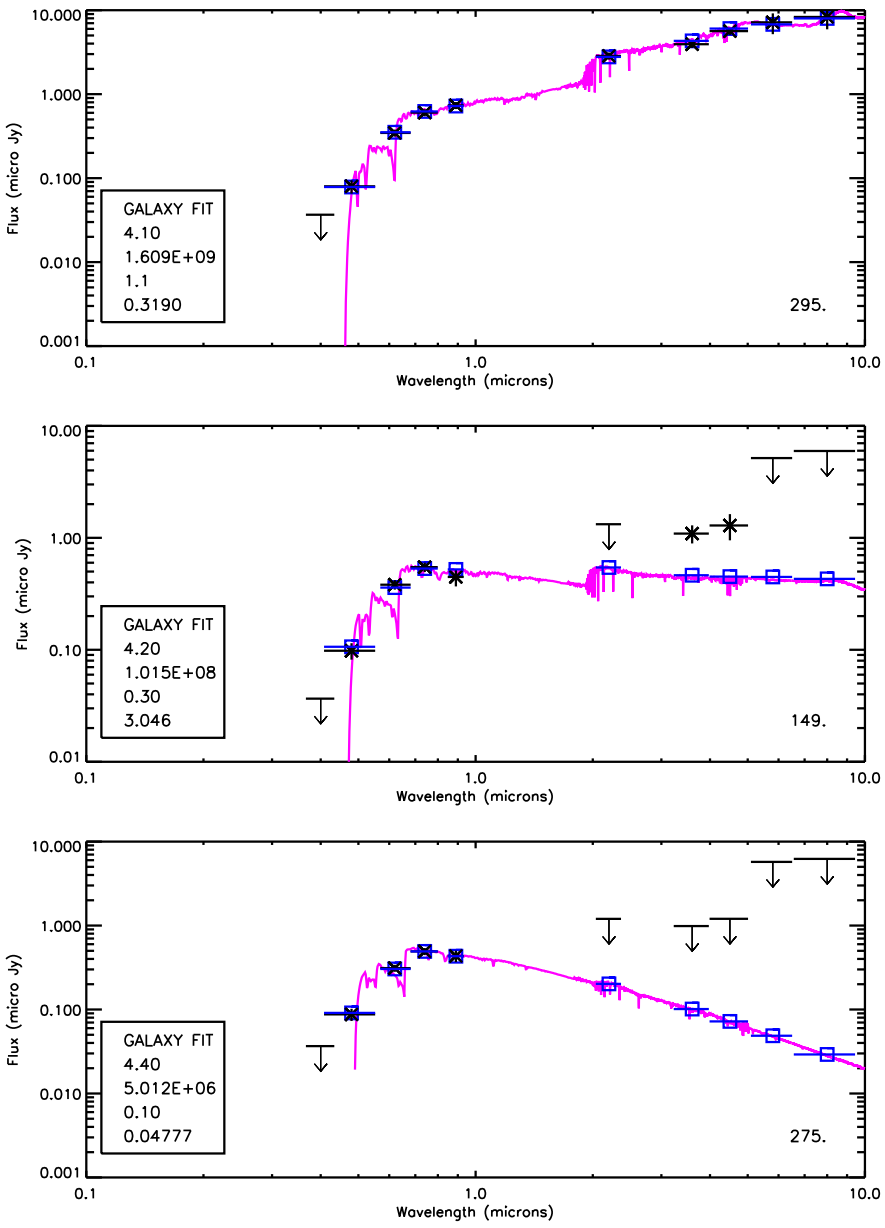

Figure 6.1: An example of best-fit models that under-predict K and IRAC measurements. For comparison, the upper plot shows a best-fit model spectrum well fitted to the observational data and indicating a best-fit age of $\sim$ 1.6 billion years. The middle plot shows a best-fit model spectrum fitted below the galaxy's IRAC 3.6 and 4.5 μm detections, with a best-fit age of $\sim$ 100 million years. The lower plot shows the best-fit model spectrum for a galaxy that is not detected in K and IRAC bands. The best-fit spectrum lies well below the detection limits and indicates a very young best-fit age of $\sim$ 5 million years.

[34] [48] [53]. In either case, the best-fit models alone may not accurately describe the properties of our galaxy population. To correctly determine these properties, an ensemble test is used to establish confidence intervals for the parameters of the best-fit models. Following the work of Papovich *et al.* (2001) [34] and Verma *et al.* (2007) [53], we utilize probability weighted ensemble properties to describe our LBG population.

6.3.2 Approach

Galaxy best-fit parameters are determined by minimizing the χ^2 test of the model fit to the galaxy SED. The $\Delta\chi^2$ between that of the best-fitting model and every other potential model may be used to define regions of constant $\Delta\chi^2$ in parameter space. These regions form the basis of multi-dimensional confidence intervals for the parameters of the galaxy fit [34] [36] [53]. Using Monte Carlo simulations of galaxy data, a relationship between $\Delta\chi^2$ and probability can be defined, thereby specifying the probability confidence levels. By combining individual galaxy probability distributions, the ensemble properties of a galaxy population may be studied. Details are provided in the following sections.

6.3.3 Delta Chi-Square Mapping

For each galaxy, the reduced χ^2 value of the fit to each of the 36,468 models was recorded. The difference in reduced χ^2 value was determined between that of the best-fitting model and every other model, resulting in a $\Delta\chi^2$ mapping of the full multi-dimensional parameter space of each galaxy. Contours of constant $\Delta\chi^2$ could then defined with respect to the galaxy best-fit position.

6.3.4 Designing Synthetic Galaxy Data

Model fitting of photometric data is used to determine a set of parameters for each galaxy that, with some confidence, is likely to represent the "true" parameters of that galaxy [36]. Due to the stochastic uncertainties associated with any measurement, the actual data set obtained is one of many possible data sets. Limited resources make it impractical to obtain multiple sets of real data. Instead, Monte Carlo simulations are used to generate "synthetic" versions of galaxy data which emulate statistical variations in experimental data sets [34] [36] [53]. Applying the same model fitting routine to each of these simulated data sets could result in a parameter set that differs from the original fit. The best-fit parameters from a large number of these synthetic data sets may be used to determine a probability distribution of possible parameter fits for each galaxy.

To generate a set of synthetic galaxies, we used a version of the Bootstrap Method known as Bootstrap Resampling [34] [36] [53]. Random numbers were drawn from a normal distribution and multiplied by the uncertainty in each filter. The original photometric data was then adjusted by the amount of the random uncertainty. Each filter was treated in the same manner, resulting in a new set of "measurements" for the galaxy, slightly altered from the original. The uncertainties on the synthetic flux values were kept consistent with those of the original flux values. Each synthetic galaxy was then subjected to the same χ^2 minimization routine as the original data set, providing new "best-fit" parameters.

Five hundred synthetic realizations were created for each galaxy in the final set. As was done for the original data, magnitude upper limits were treated as zero flux values. The addition of Gaussian noise to these zero flux data points created, in some instances, a negative total flux for the simulation, leading to complications in the fitting normalization and subsequent determinations of stellar mass and SFR. To

correct this, realizations for which the total galaxy flux became negative were handled selectively and the flux value was reset to zero each time a negative random number was drawn. A comparison of several synthetic best-fit spectra and the galaxy's original best-fit model are provided in Figure 6.2.

6.3.5 Probability Transformation

The synthetic galaxy set was culled using the same criteria as previously discussed: removal of realizations with best-fit parameters indicating a star, quasar, out-of-bin redshift or poor reduced χ^2 fit. As a consequence, the number of synthetic fits used to determine the probability distribution of each galaxy was somewhat less than that of the original set of 500 simulations. In no case, however, was the number of synthetic realizations fewer than 80 for an individual galaxy.

The location of each synthetic best-fit model within a galaxy's full parameter space can be associated with the $\Delta\chi^2$ value of that model from the $\Delta\chi^2$ mapping. Probabilities are determined for each synthetic fit by considering the entire distribution for the galaxy and the fraction of synthetic fits that occur outside the region enclosed by a certain constant $\Delta\chi^2$ contour. That is, for each synthetic fit, one sums all fits that have a value of $\Delta\chi^2$ greater than $\Delta\chi^2_0$. The ratio of this sum to the total number of synthetic fits for the galaxy determines the probability assigned to the limiting $\Delta\chi^2_0$ value. The probability values range from 0 (at the edge of the distribution) to 1 (at the best-fit location). The resultant probability distribution has a value at each synthetic fit position within the galaxy's multi-dimensional parameter space. Using this procedure, a probability distribution was determined for each galaxy based on its synthetic best-fits.

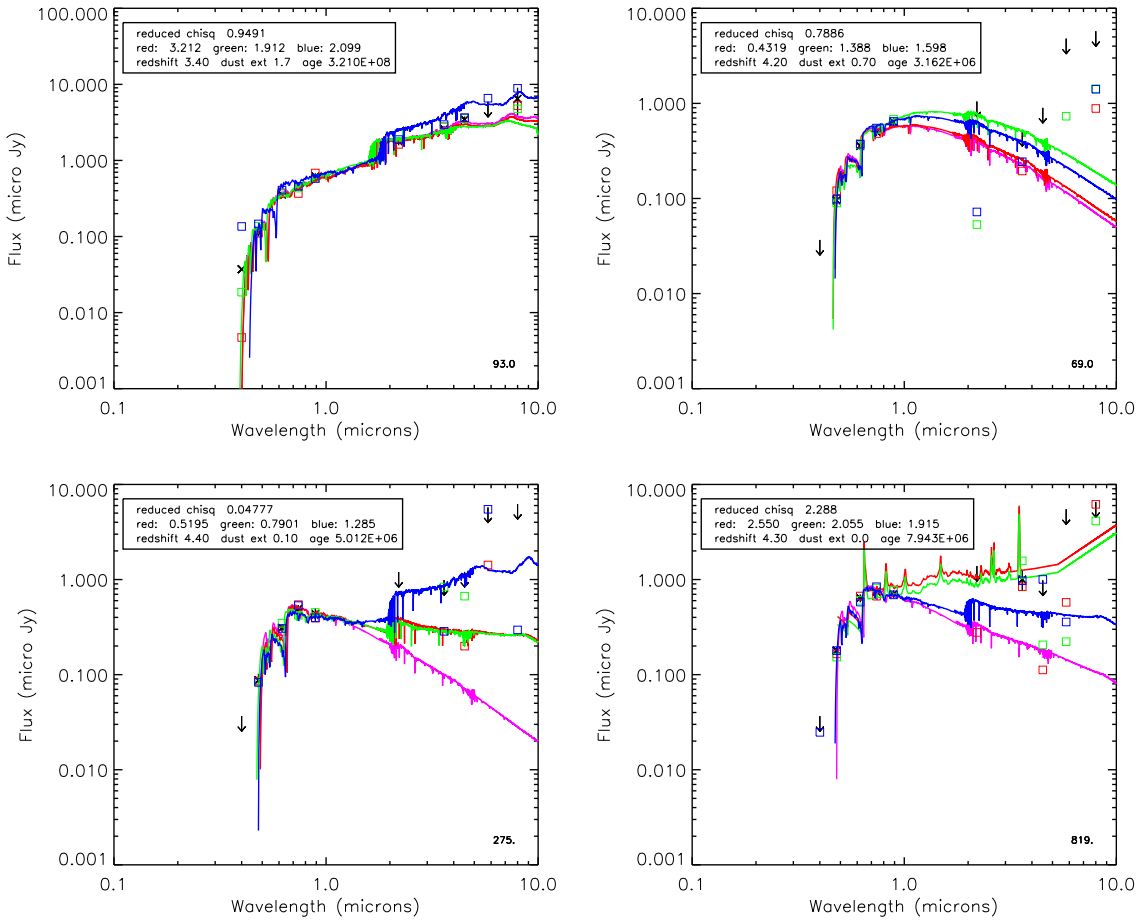

Figure 6.2: A comparison of several synthetic best-fit spectra and the galaxy's original best-fit model. Synthetic spectra are plotted in red, green and blue. The corresponding synthetic data points are plotted as open squares of the same colour. The original best-fit spectrum is plotted in purple with data points in black. Note that the lower-left plot is an example of galaxy data, originally best fit by a young model, being well fit to older models, as discussed in Section 6.2.1.

6.4 Composite Probability Distributions

As per Verma *et al.* (2007) [53], we would like to determine the probability distributions for age, stellar mass, SFR and dust extinction for the entire LBG population. The Monte Carlo simulation results have provided a full parameter space probability distribution for each galaxy. The composite probability distributions are obtained by combining the individual probability distributions of all 378 galaxies at each point of the multi-dimensional parameter space.

6.4.1 Stellar Mass and Star Formation Rate

Galaxy stellar masses and SFRs are calculated from the normalization values of the fitting procedure and are unique to each galaxy. As a result, synthetic best-fit stellar masses and SFRs lie on irregular grids in parameter space that differ for each galaxy. To provide a consistent framework independent of individual galaxy variations, we defined a grid for each parameter and binned the probabilities at each position within the grid. Prior to calculating the binned composite probabilities, each galaxy's probability space was normalized. Galaxies with tightly constrained fits had non-zero probabilities within a small region of the full parameter space, whereas galaxies with poorly constrained fits had non-zero probabilities distributed over a much broader region of parameter space. Normalizing the probability space ensures that the composite probability distribution is not dominated by poorly constrained fits. The composite probability distributions of stellar mass versus age and SFR versus age were determined by averaging the individual galaxy distributions at each point in the defined parameter space. One- and two-sigma contour intervals were determined by integrating outwards from the regions of peak probability, so as to enclose 68.3% and 95.4% of the total probability, respectively.

6.4.2 Dust Extinction

Optical depths (τ) for dust extinction are limited in the fitting routine and forced to a regular grid spacing that is constant from galaxy to galaxy, thus eliminating the need for binning. It was, however, necessary to consider that models may have a different redshift but the same age and optical depth. As a result, these models, and their associated probabilities, share the same location in the optical depth versus age parameter space. The 36,414 model parameter space collapses to a 714 point parameter space, with the optical depth and age parameters forming a 21 x 34 grid. The normalized individual probability distributions were averaged at each point in the grid to determine the composite probability distribution. Confidence contours were determined by integrating outwards from the regions of peak probability, so as to enclose 68.3% and 95.4% of the total probability.

6.5 Ensemble Properties

The composite probability plots of stellar mass, SFR and optical depth versus age are provided in Figures 6.3, 6.4, and 6.5, respectively. Each distribution represents the probability weighted ensemble properties of the set of 378 $z \sim 4$ LBGs within the respective parameter space. Histograms of the integrated composite probabilities along each parameter axis are provided in Figure 6.6. The median and range values of each parameter are summarized in Table 6.1.

6.5.1 Stellar Mass and Age

The ensemble distributions (Figures 6.3, 6.4, and 6.5) indicate that the ages of galaxies within the population range from approximately 1 million to 2 billion years. Within this range, however, are two distinct groups of galaxies (see Figure 6.3): one

Parameter	Median	Range
Age [yr]	4.90×10^7	1×10^6 - 2×10^9
Stellar mass [$M_\odot$]	2.31×10^{10}	5×10^8 - 1×10^{12}
SFR [$M_\odot$ yr^{-1}]	2.93×10^2	1×10^1 - 1×10^5
Optical depth	0.90	0 - 2

Table 6.1: Median and range values of the ensemble distributions. Note that optical depth (τ) is defined at wavelength 0.55 μm (V band).

group for which the majority of the galaxies are relatively young ($<$ 100 Myr) and low mass ($\sim$ 10^9 - 10^{11} $M_\odot$); the other, old ($>$ 100 Myr) and high mass ($\sim$ 10^{11} - 10^{12} $M_\odot$) galaxies. Beyond 100 million years, there is an overlap of the group distributions in terms of age that suggests the presence of galaxies at intermediate stages along the evolutionary path from young and low mass to old and high mass.

Figure 6.3 shows that the old, high mass group of LBGs is better constrained than the young, low mass group, which has a broader distribution in terms of both stellar mass and age. As discussed in Section 6.3.1, galaxies fit to young models are generally those with non-detections in the Spitzer IRAC bands. This may be explained by the fact that galaxies dominated by young stellar populations are brighter in the observed-frame optical than in the near- and mid-infrared (i.e. they possess very blue spectra) [34]. Galaxies bright in the observed-frame near- and mid-infrared contain older, redder stellar populations that comprise the bulk of a galaxy's stellar mass. With fewer data points to limit the model fit, the properties of the young, low mass LBG population are poorly constrained. In particular, the stellar mass distribution appears to have a clear upper limit ($< 10^{11}$ $M_\odot$) but non-zero probabilities down to $\sim 5 \times 10^8$ $M_\odot$. However, the majority of the young best fits are concentrated at the high stellar mass end of the distribution and lower mass galaxies are less likely.

The stellar mass and age histograms (Figure 6.6) both display the high probabil-

ity regions corresponding to the two groups of LBGs within the population. Stellar mass probability peaks broadly at $\sim 10^{10}$ $M_{\odot}$, with a smaller peak at $\sim 2 \times 10^{11}$ $M_{\odot}$. Age probability peaks at $\sim$ 60 million and 2 billion years, with comparable peak probabilities. The overlap in the age distributions of the two groups is also evident as a broad lower peak between the two high probability regions of the age histogram at $\sim$ 150 million years.

6.5.2 Star Formation Rate and Age

The ensemble distribution for SFR and age (Figure 6.4) reveals that SFRs for the LBG population vary from $\sim$ 10 to 100,000 $M_{\odot}$ per year, with a median SFR of 293 $M_{\odot}$ per year. Indications of the two population groups are less evident in this plot, however, the distribution reveals high SFRs ($>$ 100 $M_{\odot}$ per year) for very young galaxies ($<$ 30 Myr) and fairly constant lower SFRs in the range of 20 to 1000 $M_{\odot}$ per year for galaxies older than 30 million years. There is evidence for a steady decline in SFR for galaxies as they increase in age from 1 million to 40 million years, from the maximum of $\sim$ 100,000 $M_{\odot}$ per year to a minimum of $\sim$ 100 $M_{\odot}$ per year. This negative correlation between SFR and age for young galaxies is the result of our selection criteria. Younger galaxies need to be producing stars at a higher rate to be detectable at the same UV luminosity as older galaxies that contain larger numbers of young stars. Since most of the UV flux is due to stars that are less than 100 million years old, the distribution flattens for older galaxies as they surpass this age.

The SFR histogram (Figure 6.6) displays a large peak at $\sim$ 200 $M_{\odot}$ per year, corresponding to the region of constant SFR for older galaxies in Figure 6.4. Beyond this peak, the probabilities of higher SFRs show a continual decline towards the maximum SFR. Based on the above discussions, the two LBG populations may now

be described as a group of young, low mass galaxies with high SFRs and a group of old, massive galaxies that are producing stars at much lower rates.

6.5.3 Dust and Age

The ensemble distribution for optical depth and age (Figure 6.5) reveals a wide spread in the dust reddening for these LBGs, with non-zero probabilities over most of the parameter space. Contained within the one-sigma contour, however, are two regions of higher probabilities: young galaxies ($<$ 20 Myr) with high dust extinction ($\tau > 1$) and old galaxies ($>$ 300 Myr) with low dust extinction ($\tau <$1). Recall that optical depth (τ) is defined at wavelength 0.55 μm (V band). A large number of best-fits are concentrated in the region defined by ages between 3 and 10 million years and optical depths in the range of $1 \leq \tau \leq 1.7$. The region of highest probability is defined by very old galaxies ($>$ 1 Gyr) with little or no dust extinction. A significant number of best-fits are also located in this high probability region of the plot. These results suggest that younger galaxies are likely to be dustier, and experiencing greater dust extinction, than older galaxies. This effect is seen in low redshift starburst galaxies, which are dustier when they are young, as dust is created in the starburst event but then dispersed or destroyed over time [14].

The optical depth histogram (Figure 6.6) indicates that zero dust extinction is the most probable level of dust extinction for the LBG population. Low levels of dust extinction ($\tau <$1) appear to be fairly equal in probability, while higher levels of dust extinction become increasingly less probable. The dust extinction distribution has a median optical depth of 0.90 at 0.55 μm (V band), equivalent to extinction $A_V \sim 1$ mag (see Equation 5.4).

The combined properties of the composite probability plots suggest that LBGs at $z \sim 4$ may be comprised primarily of two distinct LBG populations: young, dusty,

low mass galaxies with high SFRs and old, massive galaxies with low dust extinction and low SFRs. There is evidence also of galaxies with properties intermediate to both of these main population groups. If so, these intermediary galaxies could provide an evolutionary link between the young, low mass galaxy population and old, high mass population. In Chapter 7, the properties of our $z \sim 4$ LBG population are compared to those of other LBG populations at different redshifts in an effort to formulate a clear picture of the evolution of LBGs in the early universe.

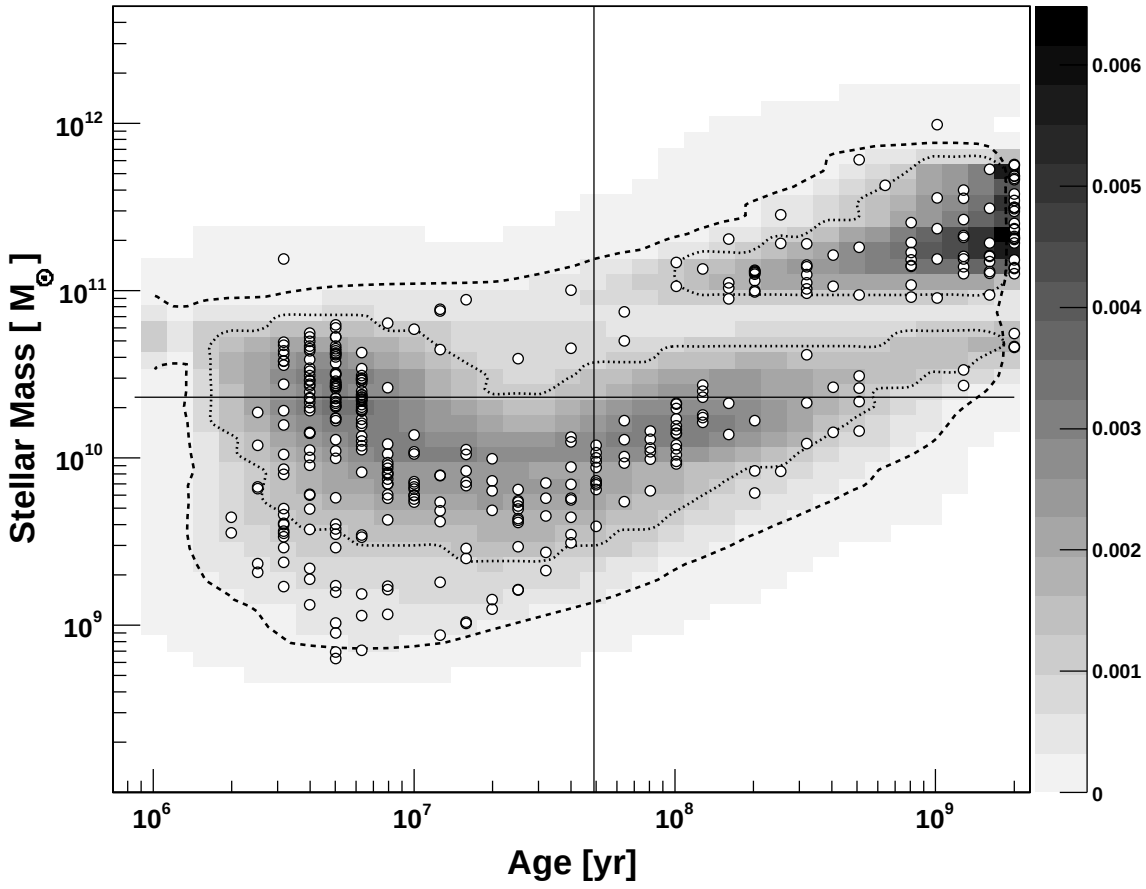

Figure 6.3: Composite probability distribution of stellar mass versus population age for the sample of 378 $z \sim 4$ LBGs. Probabilities are displayed as grayscale shading where the darkest regions correspond to highest probability. Overlaid contours outline the 68.3% (dotted) and 95.4% (dashed) confidence intervals. Open circles indicate the best-fit parameters of each galaxy in the sample and crossed lines mark the median properties of the ensemble distribution.

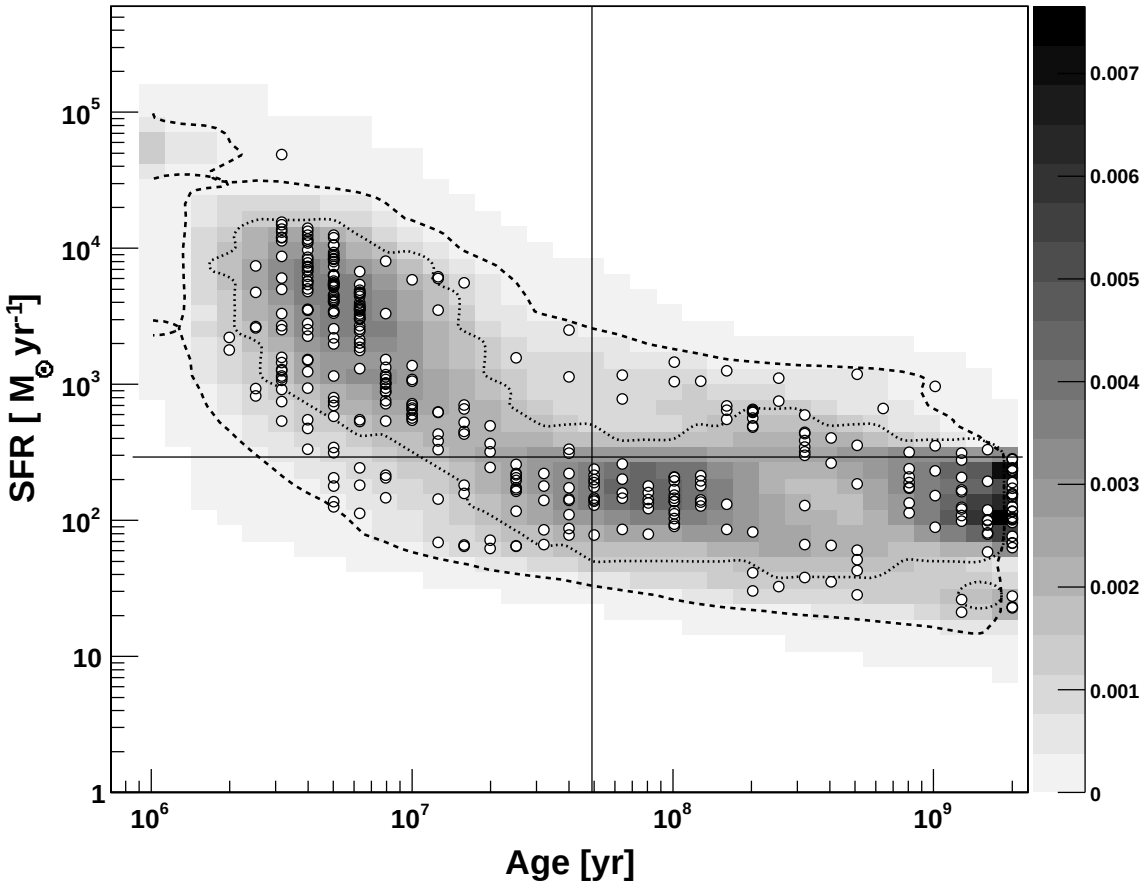

Figure 6.4: Composite probability distribution of SFR versus population age for the sample of 378 $z \sim 4$ LBGs. Probabilities are displayed as grayscale shading where the darkest regions correspond to highest probability. Overlaid contours outline the 68.3% (dotted) and 95.4% (dashed) confidence intervals. Open circles indicate the best-fit parameters of each galaxy in the sample and crossed lines mark the median properties of the ensemble distribution.

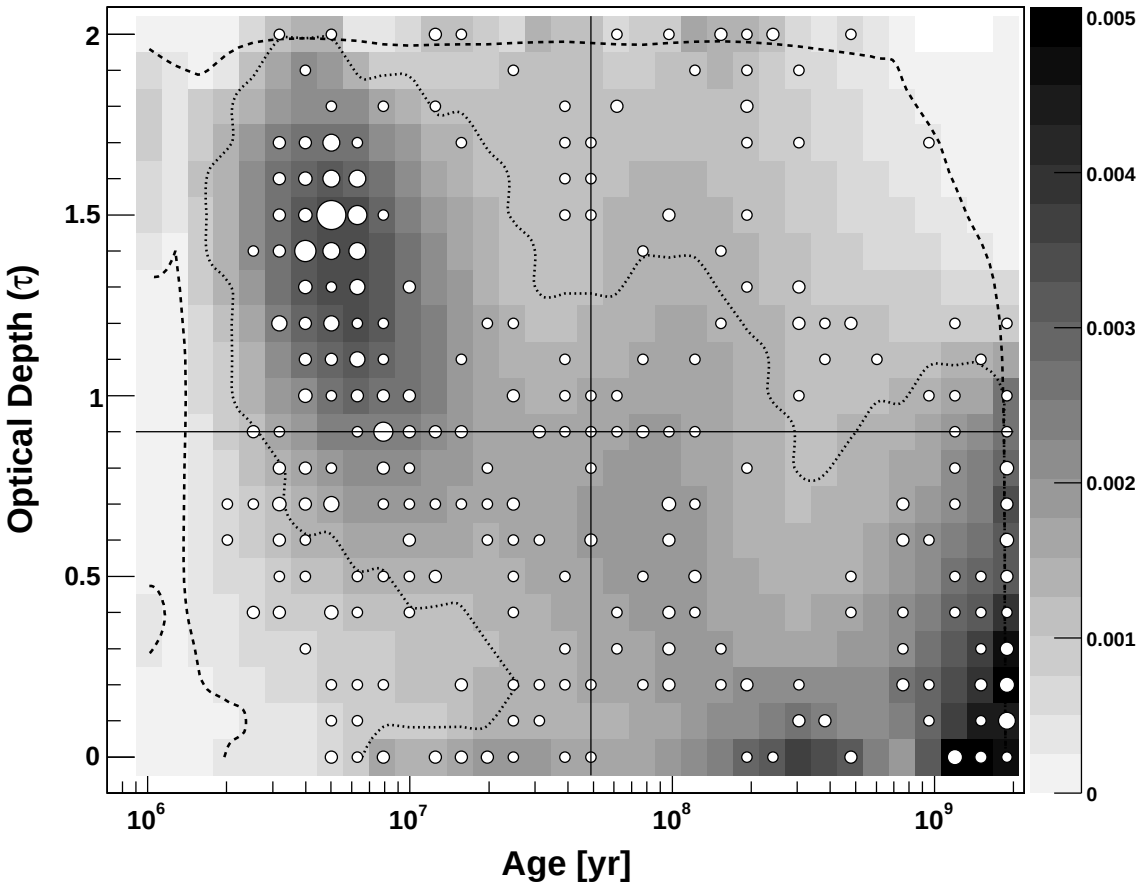

Figure 6.5: Composite probability distribution of optical depth of dust extinction versus population age for the sample of 378 $z \sim 4$ LBGs. Optical depth (τ) is defined at wavelength 0.55 μm (V band). Probabilities are displayed as grayscale shading where the darkest regions correspond to highest probability. Overlaid contours outline the 68.3% (dotted) and 95.4% (dashed) confidence intervals. Open circles indicate the best-fit parameters of each galaxy in the sample. Circle area is proportional to the number of galaxy best-fits occupying the same position in the parameter space and crossed lines mark the median properties of the ensemble distribution.

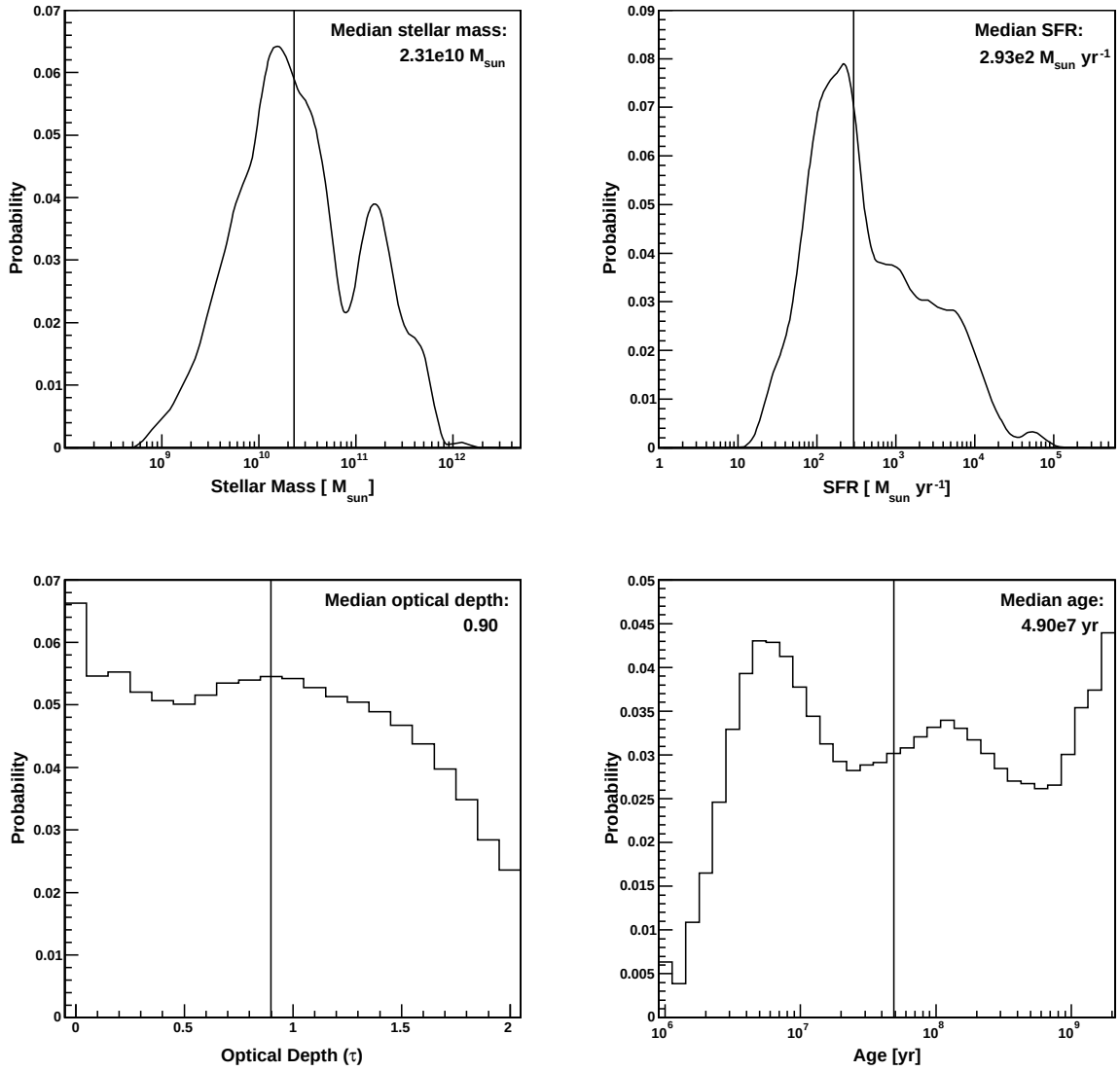

Figure 6.6: Composite probability histograms of stellar mass, SFR, optical depth and population age for the sample of 378 $z \sim 4$ LBGs. Histograms display the integrated probabilities along each parameter axis of the composite probability plots (Figures 6.3, 6.4 and 6.5). Vertical lines specify the median parameter values of the plotted distributions. Optical depth (τ) is defined at wavelength 0.55 μm (V band).

Chapter 7

Discussion and Conclusions

7.1 Summary of Results

Using rest-frame UV-to-optical SEDs, we have isolated a population of 378 LBGs in the redshift range $3.4 \lesssim z \lesssim 4.4$. These are star-forming galaxies selected to $i' \leq 24.8$ from the one square-degree CFHTLS-Deep D2 Field and identified based on distinctive colours produced by the Lyman-break in their rest-frame UV spectra. The physical properties of this LBG sample were determined by fitting population synthesis models to the photometric data. Monte Carlo simulations, designed to emulate galaxy spectra, were used to establish confidence intervals for the best-fit parameters of the ensemble distribution. The median values for the LBG population are an age of ~ 50 million years, stellar mass of ~ 2 x 10^{10} $M_{\odot}$, SFR of ~ 300 $M_{\odot}$ yr^{-1} and dust extinction of ~ 1 mag at 0.55μm (V band). There is evidence in this $z \sim 4$ LBG population of two distinct sub-populations of LBGs: young, dusty, low mass galaxies that are producing stars prolifically and old, massive, less reddened galaxies for which star formation proceeds at a much slower pace. In the next section, we compare these findings with the results of other LBG surveys at different redshifts.

7.2 Interpretation of Results

7.2.1 Comparison of LBG Studies

The physical properties of LBGs at different redshifts have been studied by a number of surveys. The specifics of six leading LBG surveys covering the redshift range $3 \lesssim z \lesssim 6$, as well as the details of our study, are provided in Table 7.1. The corresponding SFR, stellar mass, age and dust extinction parameters determined by these surveys are provided in Table 7.2. Together, these studies form a probable picture of LBG evolution over the course of more than a billion years.

A direct comparison of results from LBG surveys at successive redshifts is, however, complicated due to variations between the studies [53] [56]. These variations may include magnitude depth, size of field, cosmic variance, number of galaxies, wavelength coverage and selection criteria. For example, Stark *et al.* (2009) [48] investigate LBGs in the smaller GOODS fields, but to a much greater magnitude depth than is used in our study of the D2 Field (see Table 7.1). Deeper imaging leads to detecting objects that are intrinsically less luminous. Consequently, the UV luminosities of the population they study are distinctly different from those of our study. In addition, LBG surveys may employ a variety of models in the SED fitting and the resulting parameters may depend significantly on the choice of model ingredients, such as star formation history (SFH), dust extinction law, IMF, SFR and chemical enrichment. This leads to the possibility of systematic differences that may arise in comparisons between studies.

Several authors attempt to overcome these issues by adopting uniform selection criteria and applying the same modelling strategy used in previous studies (e.g. [53] [56]). Differences between LBG samples of varying redshift may then be attributed to evolutionary changes in the properties of the populations. Without the benefit of

complementary analyses, we compare our results to those of other studies in terms of general observed trends rather than specifics.

7.2.2 Star Formation Models

Our assumption of a constant SFR for our models is a simplification of the actual star-forming processes within a galaxy, which are almost certainly more complex. Other studies have modeled a variety of SFHs in an effort to gauge the best-fit properties for a given LBG population [48] [56]. These have included the standard instantaneous burst (or SSP) and constant SFR, as discussed in Chapter 5, linearly increasing SFR, and exponentially decreasing SFR with e-folding timescales (i.e. the period over which the SFR declines by $1/e$). To improve the estimates of the age and total stellar mass of a galaxy, one may apply models incorporating two distinct stellar populations. Such a model may comprise, for example, a young population of stars formed at a constant rate and a passively evolving population of older stars produced in an instantaneous burst of star formation at some time in the distant past.

Model fits to star-forming galaxies are primarily influenced by the flux of massive young stars that dominate the emitted light in the rest-frame UV [34] [53]. The inferred age of a galaxy is thus based on the age of this population of newly-formed stars, measured as the length of the current or most recent star-forming event. If, however, the galaxy contains an older stellar population, as a result of previous episodes of star formation, then the extracted age may not reflect the age of the population that comprises, instead, the bulk of the mass of the galaxy. Similarly, the luminosity-based stellar mass estimate conveys only the mass of the galaxy's youngest population, in effect, setting a lower bound for the total stellar mass of the galaxy.

7.2.3 Evolutionary Scenario

The results from the surveys described in Table 7.1 essentially describe a sequence of evolutionary stages for LBGs. For each measured epoch, we discuss this evolution, starting with $z \sim 6$ when the Universe was just 914 million years old [15]. Based on the findings of Eyles *et al.* (2007) [15], the LBG population at this redshift comprises new young galaxies (< 100 Myr), with masses on the order of $\sim 5 \times 10^8$ $M_\odot$, that are in the process of producing their first generation(s) of stars. Also present are mature galaxies, with ages ranging from $\sim$ 200 - 700 Myr and masses on the order of $\sim 10^{10}$ $M_\odot$, that are dominated by old stars from much earlier eras of star formation. With ages comparable to the age of the Universe at $z \sim 6$, it is reasonable to assume these galaxies were already in existence at much higher redshifts. The inferred SFRs of these galaxies vary from $\sim$ 0 - 140 $M_\odot$ per year. However, given the high stellar masses assembled by these galaxies over such short timescales, it is estimated that SFRs may have been as high as $\sim$ 200 $M_\odot$ per year in the past [15]. Extinction by dust, a by-product of previous generations of star formation, is found to be negligible at this redshift.

At $z \sim 5$, the Universe is approximately 1.2 billion years old [53]. Verma *et al.* (2007) [53] reason that at this epoch we are viewing recently formed galaxies in the process of amassing significant stellar mass. Star formation is observed to be occurring on a large-scale and, while evidence indicates the presence of old, massive galaxies in the population, the majority ($\sim$ 2/3) are young ($\lesssim$ 100 Myr) active starbursts [53]. Typical SFRs of $\sim$ 40 $M_\odot$ per year are estimated and the stellar masses of these galaxies are on the order of $\sim 10^9$ $M_\odot$. Yabe *et al.* (2008) [56] determine that their $z \sim 5$ LBG population has median age and stellar mass values of 25 Myr and $\sim 10^9$ $M_\odot$, respectively, and moderate levels of dust extinction, each consistent with Verma *et al.* (2007). The median SFR for their sample is, however,

higher at 141 $M_\odot$ per year, indicating more intense star formation activity.

By redshift $z \sim 3$, observed LBGs have become dustier and more massive [34]. The Universe is approximately 2.2 billion years old, previous generations of star formation have increased the levels of interstellar dust and rapid star formation is still an ongoing process. Papovich *et al.* (2001) [34] find no evidence of young ($\lesssim$ 100 Myr), unreddened galaxies in their sample, as was observed for emerging populations at earlier redshifts. Their model-dependent best-fit results also indicate that the population at this redshift is comprised of galaxies covering a wide range of ages (30 Myr to $\sim$ 1 Gyr), stellar masses ($\sim 10^9$ - 10^{11} $M_\odot$) and SFRs ($\sim$ 1 - 1000 $M_\odot$ per year). Dust extinction is highest for the brightest, star-forming galaxies and SFRs are, in turn, highest for the youngest galaxies.

Shapley *et al.* (2001) [41] find similar trends in the best-fit properties of their sample of $z \sim 3$ LBGs. In addition, they describe their distribution of galaxies as a continuum of ages, stellar masses and extinction levels representing different evolutionary stages of the population. At each end of this distribution are well-defined sub-populations: at the low mass end ($\sim 10^9$ $M_\odot$) are young ($\lesssim$ 35 Myr), dust-obscured galaxies that are intensely star-forming ($\sim$ 425 $M_\odot$ per year); at the massive end ($\sim 10^{10}$ $M_\odot$) are old ($\gtrsim$ 1 Gyr), dust-reddened galaxies that are forming stars at much lower rates ($\sim$ 50 $M_\odot$ per year). Moderately evolved galaxies populate the parameter space between these two extremes. Proposed is an evolutionary scenario whereby SFRs decline over a period of $\sim$ 50 to 100 Myr to less active levels of star formation that may be maintained over the course of a billion years or more.

Our study focuses on the redshift range $3.4 \lesssim z \lesssim 4.4$, over which the Universe ages from $\sim$ 1.4 Gyr to 1.9 Gyr. Given that our study falls directly between the surveys at $z \sim 3$ and $z \sim 5$, we anticipate our results will describe properties of LBG evolution that link these two redshift epochs. The details of this comparison are provided below:

- At 2×10^{10} $M_{\odot}$, the median stellar mass of our LBG population is higher than the typical stellar masses of $\sim 10^{9}$ $M_{\odot}$ seen at $z \sim 5$, and consistent with those at $z \sim 3$ (see Table 7.2). This comparison, based on the median values of full distributions, may, however, be misleading given the presence of a bimodal distribution within our population. Accordingly, we establish medians for the young (< 100 Myr) and old (> 100 Myr) stellar populations within our distribution of 1×10^{10} $M_{\odot}$ and 2×10^{11} $M_{\odot}$, respectively. These medians do not alter our previously stated observations. The median stellar mass of the young, low-mass population is still higher than that seen typically at $z \sim 5$ and both the young and old median stellar masses remain consistent with those at $z \sim 3$. Although the $z \sim 5$ surveys by Verma *et al.* (2007) [53] and Yabe *et al.* (2008) [56] comprise a smaller number of galaxies at fainter magnitudes, making it difficult to directly compare populations, the stellar masses we derive are significantly higher and this effect is unlikely to be due to sampling inconsistencies (e.g. our brighter population of LBGs). In the same manner, the similarity in median stellar masses at $z \sim 4$ and $z \sim 3$ suggests that stellar mass accumulation has not been significant between these epochs.

- The median age of our population, 49 Myr, shows an increase from the 25 Myr median best-fit age found at $z \sim 5$, and is well below the $z \sim 3$ median of 320 Myr (see Table 7.2). However, our value does fall within the broad range of LBG ages (30 Myr to 1 Gyr) determined by Papovich *et al.* (2001) [34] at $z \sim 3$. The general trend of increasing median age from one epoch to the next suggests that the fraction of newly-formed stellar populations declines from $z \sim 5$ to $z \sim 3$, with the most significant decrease from $z \sim 4$ to $z \sim 3$. The substantial aging of observed stellar populations from $z \sim 4$ to $z \sim 3$ agrees with the inferred decline in stellar mass production between these redshifts,

discussed above. However, we are hesitant to conclude this trend is substantive given the bimodal age distribution in our population.

- At 293 $M_\odot$ per year, the median SFR of our $z \sim 4$ LBG population is significantly higher than the median values found at either $z \sim 5$ or $z \sim 3$, although it is consistent with the broad range defined by Papovich *et al.* (2001) [34] (see Table 7.2). Since the CFHTLS is a relatively bright survey over a large area, the galaxies in our sample are typically more UV-luminous than those of deeper surveys and, as a result, we can expect to observe much higher SFRs. With this in consideration, it is unlikely that we are witnessing an era of increased and very intense star formation at $z \sim 4$, as suggested by the median values. Determining the significance of this parameter based on its extracted median value is therefore difficult and, instead, it may be more reasonable to examine the trends in SFR for the population as a whole, discussed below.

- Although not well constrained, the median V band dust extinction of our LBG sample of ~ 1 mag suggests that dust extinction levels have increased from typically moderate values (e.g. 0.3 mag) at $z \sim 5$, and that LBGs have become dustier still by $z \sim 3$ (see Table 7.2). This trend is consistent with the accumulation of interstellar dust as a result of successive generations of star formation in massive, aging stellar populations.

- Perhaps most interesting is the bimodal distribution in our LBG population, similar to that observed by Shapely *et al.* (2001) [41] at $z \sim 3$. We observe the same general trends in the two dominant populations: young, dusty, low mass galaxies that are intensely star-forming and old, massive, less reddened galaxies that are producing stars at much lower rates (see Figures 6.3, 6.4, 6.5). Also present in our population are galaxies with properties intermediate

to these two evolutionary states. We note, however, that the typical stellar masses and SFRs of the two distinct $z \sim 3$ populations are comparable to lower limits, rather than typical values, for our $z \sim 4$ sub-populations (based on 68% confidence intervals). Further, the observed decline in the SFRs of young galaxies occurs over the course of ~ 40 Myr in our population, compared to the ~ 50 - 100 Myr timescale suggested by Shapley *et al.* (2001). Again, these stellar mass and SFR trends may be attributed to our survey of a brighter, more massive population of galaxies at $z \sim 4$.

Summarizing the above analysis, we propose the following tentative scenario of LBG evolution over the redshift range $3 \lesssim z \lesssim 5$. From $z \sim 5$ to $z \sim 4$, young, actively star-forming LBGs begin to build up significant stellar mass, galaxies become dustier and populations begin to mature. From $z \sim 4$ to $z \sim 3$, evolutionary changes proceed more slowly and LBGs may achieve a relatively stable state, such that there is limited growth. Star formation remains vigorous, dust levels continue to increase, and distinct sub-populations evolve as the populations age. Our survey provides consistent information for a redshift range that had previously not been well sampled and thus establishes a valuable link between existing surveys and the LBG population at $z \sim 4$.

7.2.4 Episodic Star Formation

Crucial to the study of galaxy evolution is an understanding of how star formation proceeds within galaxies [48]. The star formation history of a galaxy may be deduced from the evolution of its physical properties over a range of redshifts. If stars are produced at a constant rate, galaxies of similar luminosity should show substantial evolution over cosmic time, with increased ages, greater accumulated stellar mass and higher dust extinction. If SFRs were much lower in the past, then galaxies

are unlikely to show significant evolution between redshifts. Alternatively, if star formation occurs over short time scales separated by long periods of inactivity, then each redshift should be populated by newly emerged stellar populations.

The general consensus supports the view that LBGs evolve by way of episodic bursts of star formation [34] [41] [48] [53]. In this scenario, periods of intense star formation persist for several hundred million years, at which point SFRs decline and star formation proceeds at much lower rates over potentially very long timescales ($\sim 10^9$ years). Given the intermittent nature of this process, it is believed that episodes of star formation are triggered by random large-scale events such as the merger of galaxies, the accretion of massive quantities of gas or the tidal influence of nearby galaxies. In accordance with the hierarchical scenario, these types of interactions may be expected to occur relatively frequently in the small, denser, early universe.

The consistent dominant presence of young galaxies in the LBG populations at each successive redshift, including our own survey, agrees well with the episodic star formation hypothesis [34]. Further, Stark *et al.* (2009) [48] find that the typical stellar masses and population ages do not evolve significantly in their survey of LBGs at $z \sim 4$, 5 and 6 (see Table 7.2). This trend would be unexpected for galaxies producing stars at a constant rate. Instead, they conclude that episodic star formation processes are likely, ensuring new starburst populations are dominant at each redshift. If this is indeed the case, then substantial populations of galaxies in the quiescent phase of episodic evolution must also exist. If SFRs are very low, the host galaxies may not be detectable by LBG selection criteria, which targets actively star-forming galaxies.

7.3 Future Studies

A number of additional steps could be taken to further optimize the analysis of our $z \sim 4$ LBG population. For instance, an upper bound for the total stellar mass could be determined by refitting our SEDs to a two-component model that includes both young and old stellar populations [34]. Further, a galaxy containing older stars will have greater emission in the observed-frame infrared, such that the use of deeper infrared data would allow us to better constrain total stellar masses as well as model fits for young galaxies [48]. As a natural progression in this research field, high redshift galaxies are now being selected at infrared and submillimetre (200 μm - 1 mm) wavelengths to complement and extend the studies of UV-selected populations [102]. Much of the light emitted by galaxies that are actively star-forming is absorbed by dust and reradiated at longer wavelengths. Submillimetre observations are sensitive to the radiation emitted by cold gas and dust typical of star-forming environments [103]. The SCUBA 2 (Submillimetre Common-User Bolometer Array) wide-field camera, at the James Clerk Maxwell Telescope in Hawaii, operates at 450 and 850 μm and provides access to dust-obscured, star-forming galaxies with a sensitivity that is undiminished over the redshift expanse of $1 < z < 10$. Since galaxies bright in the submillimetre are faint at other wavelengths, submillimeter astronomy has opened a new frontier in extragalactic research.

Perhaps most compelling in the future of these LBG studies is a unified picture of the global SFR and stellar mass densities. The calculations towards this end remain quite preliminary [15] [53]. To accurately determine these parameters, the populations of star-forming galaxies, as well as post-starburst, dust-obscured, and those below detection limits, must be estimated. Developing an understanding of the inter-relationship between these various populations of high redshift galaxies will be fundamental to achieving a comprehensive picture of the evolution of LBGs in

the early universe.

Paper	Papovich *et al.* (2001)	Shapley *et al.* (2001)	this thesis	Verma *et al.* (2007)	Yabe *et al.* (2008)	Eyles *et al.* (2007)	Stark *et al.* (2009)
Redshift	$z \sim 3$	$z \sim 3$	$z \sim 4$	$z \sim 5$	$z \sim 5$	$z \sim 6$	$z \sim 4$, 5, 6
Field(s)	HDF-N	various incl. HDF WESTPHAL	CFHTLS-D2	CDF-S	GOODS-N GOODS-FF	GOODS-S	GOODS-N GOODS-S
Total survey area [arcmin^2]	~ 720	~ 30	3600	~ 400	~ 400	~ 160	~ 320
Magnitude limit	$H_{160} < 26.5$	$R \leq 25.5$	$i' \leq 24.8$	$i_{AB} < 26.3$	$z'_{AB} < 26.5$	$z'_{AB} < 26.9$	$z'_{850} < 27.0$
Number of dropouts	33 U	81 U_n	378 g′	21 V	105 V	16 i′	2443 B 506 V′ 137 i′
Filters	(U) B V I J H K	U_n G R J K_s	u* g′ r′ i′ z′ K IRAC Ch1-4	B V i z J K_s IRAC Ch1-4	(V) I_c z′ IRAC Ch 1-2	B v i′ z′ J K_s IRAC Ch1-2	U B V i′ z′ J K IRAC Ch 1-2
Spectroscopic confirmation	yes	yes	no	yes	yes	yes	yes
Telescope	HST, KPNO	WHT, PHT KPNO	CFHT, VLT Spitzer	HST, VLT Spitzer	HST, Subaru	HST, VLT Spitzer	HST, Subaru KPNO, VLT Spitzer
BC model	BC2000	BC1996	CB2008	BC2003	BC2003	BC2003	CB2008
Modelling scenarios	CSFR, DSFR 2-comp	CSFR SSP, DSFR	CSFR	CSFR, DSFR SSP, 2-comp	CSFR, DSFR SSP, 2-comp	CSFR, DSFR SSP, 2-comp	CSFR, ISFR DSFR
Notes				Compare to $z \sim 3$	Compare to $z \sim 2$, 3, 5, 6		

Table 7.1: Comparison table of current LBG research from $z \sim 3$ to $z \sim 6$. Note: CSFR (constant SFR), SSP (instantaneous burst), ISFR (linearly increasing SFR), DSFR (exponential decreasing SFR with e-folding timescale), 2-comp (two-component model incorporating both young and old stellar populations). Bracketed filters are not included in the fitting.

Redshift	Paper	SFR [$M_\odot$ yr^{-1}]	Stellar mass [$M_\odot$]	Age [Myr]	Dust extinction [mag]
$z \sim 3$	Papovich *et al.* (2001)	10^{-2} - 10^{3}	$\sim 10^{9}$ - 10^{11}	30 - 1000	A_{1700} 0 - 4
	Shapley *et al.* (2001)	92	2.45×10^{10}	320	A_{1600} 1.62
$z \sim 4$	this thesis	293	2.31×10^{10}	49	$A_V \sim 1$
	Stark *et al.* (2009)	-	7.3×10^{8} - 2.23×10^{10}	143 -203	-
$z \sim 5$	Verma *et al.* (2007)	40	2×10^{9}	~ 25	A_V 0.3
	Yabe *et al.* (2008)	141	4.1×10^{9}	25	E(B -V) 0.22
	Stark *et al.* (2009)	-	3.6×10^{9} - 5.20×10^{10}	181 - 286	-
$z \sim 6$	Eyles *et al.* (2007)	0 - 140	~ 1.0 - 3.0×10^{10}	$\sim$ 200 - 700	E(B -V) 0 - 0.16
	Stark *et al.* (2009)	-	2.7×10^{9} - 2.26×10^{10}	102 - 286	-

Table 7.2: Comparison table of the median best-fit parameters determined by the LBG surveys described in Table 7.1. Where medians were not available, the range of best-fit parameters is given.

Bibliography

[1] Adelberger K.L. *et al.*, **Optical Selection of Galaxies at Redshifts 1 < z < 3**, astro-ph/0401445v1 (2004).

[2] Barro G. *et al.*, **On the nature of the extragalactic number counts in the K-band**, astro-ph/0811.3104v1 (2008).

[3] Bertin E. and Arnouts S., **SExtractor: Software for source extraction**, Astronomy & Astrophysics Supplement 317, 393, 1996.

[4] Blanc G.A. *et al.*, **The Multiwavelength Survey by Yale-Chile (MUSYC): Wide K-band Imaging, Photometric Catalogs, Clustering and Physical Properties of Galaxies at z ∼ 2**, astro-ph/0803.0763v2 (2008).

[5] Bouwens R.J., Illingworth G.D., Franx M. and Ford H., **UV Luminosity Functions at z ∼ 4, 5, and 6 from the HUDF and Other Deep HST ACS Fields: Evolution and Star Formation History**, astro-ph/0707.2080v1 (2007).

[6] Bouwens R.J. *et al.*, **UV Luminosity Functions from 113 z ∼ 7 and z ∼ 8 Lyman-Break Galaxies in the Ultra-Deep HUDF09 and Wide-Area ERS WFC3/IR Observations**, astro-ph/1006.4360v2 (2010).

[7] Bouwens R.J., Illingworth G.D., Thompson R.I. and Franx M., **Constraints on z $\sim$ 10 Galaxies from the Deepest HST NICMOS Fields**, astro-ph/0503116v2 (2005).

[8] Brandt S., **Data Analysis: Statistical and Computational Methods for Scientists and Engineers**, 3rd Edition, Springer-Verlag New York Inc., 1998.

[9] Bruzual G., **Stellar population synthesis models at low and high redshift**, astro-ph/0011094v1 (2000).

[10] Bruzual G. and Charlot S., **Stellar population synthesis at the resolution of 2003**, astro-ph/0309134v1 (2003).

[11] Carroll B.W. and Ostlie D.A., **An Introduction to Modern Astrophysics**, 2nd Edition, Addison-Wesley Publishing Company, Inc., 2007.

[12] Conselice C.J., **The Assembly History of Massive Galaxies: What Do We Know?**, astro-ph/0706.3482v1 (2007).

[13] Dickinson M. *et al.*, **Color-selected galaxies at z $\approx$ 6 in the Great Observatories Origins Deep Survey**, astro-ph/0309070v1 (2003).

[14] Dwek E., **The Evolution of the Elemental Abundances in the Gas and Dust Phases of the Galaxy**, astro-ph/9707024v2 (1997).

[15] Eyles L.P. *et al.*, **The stellar mass density at z $\approx$ 6 from Spitzer imaging of i′-drop galaxies**, MNRAS 374, 910-930 (2007).

[16] Fazio G.G. *et al.*, **The Infrared Array Camera (IRAC) for the Spitzer Space Telescope**, astro-ph/0405616v1 (2004).

[17] Forster Schreiber N.M. *et al.*, **A substantial population of red galaxies at z > 2: Modeling of the spectral energy distributions of an extended sample**, astro-ph/0408077v1 (2004).

[18] Gary B.L., **Derivation of S.E. Using SNR**, http://reductionism.net.seanic.net/Astrophotos/SE/snr.htm (2003).

[19] Giavalisco M., **Lyman-Break Galaxies**, Annu. Rev. Astron. Astrophys. 40:579-641(2002).

[20] Grant D.R., private communication.

[21] Gustafsson B. *et al.*, **A grid of MARCS model atmospheres for late-type stars I. Methods and general properties**, astro-ph/0805.0554v1 (2008).

[22] Hildebrandt H. *et al.*, **CARS: the CFHTLS-Archive-Research Survey II. Weighing dark matter halos of Lyman-break galaxies at z = 3 - 5**, A&A 498, 725-736 (2009).

[23] Howell S.B., **Handbook of CCD astronomy**, Cambridge University Press, 2006.

[24] Ilbert O. *et al.*, **Accurate photometric redshifts for the CFHT Legacy Survey calibrated using the VIMOS VLT Deep Survey**, astro-ph/0603217v1 (2006).

[25] Iwata I. *et al.*, **Differential Evolution of the UV Luminosity Function of Lyman Break Galaxies from z $\sim$ 5 to 3**, astro-ph/0701841v1 (2007).

[26] Karttunen H., Kroger P., Oja H., Poutanen M. and Donner K.J. (Eds.), **Fundamental Astronomy**, 2nd Edition, Springer-Verlag, 1994.

[27] Kriek M. *et al.*, **Direct Measurements of the Stellar Continua and Balmer/4000Å Breaks of Red z > 2 Galaxies: Redshifts and Improved Constraints on Stellar Populations**, astro-ph/0603063v1 (2006).

[28] Le Fevre O. *et al.*, **The VIMOS VLT Deep Survey. First epoch VVDS-Deep survey: 11564 spectra with 17.5 $\leq$ I_{AB} $\leq$ 24, and the redshift distribution over 0 $\leq$ z $\leq$ 5**, astro-ph/0409133v3 (2005).

[29] Loeb B. and Barkana R., **The Reionization of the Universe by the First Stars and Quasars**, astro-ph/0010467v1 (2000).

[30] McCracken H.J. *et al.*, **The COSMOS-WIRCam Near-Infrared Imaging Survey: I: BZK Selected Passive and Star Forming Galaxy Candidates at z $\gtrsim$ 1.4**, astro-ph/0910.2705v1 (2009).

[31] Mobasher B. *et al.*, **Photometric Redshifts of Galaxies in COSMOS**, astro-ph/0612344v1 (2006).

[32] Mobasher B. and Mazzei P., **A Self-Consistent Photometric Redshift Technique**, astro-ph/0008027v1 (2000).

[33] Moster B.P., Somerville R.S., Newman J.A. and Rix H-W., **A Cosmic Variance Cookbook**, astro-ph/1001.1737v1 (2010).

[34] Papovich C., Dickinson, M. and Ferguson, H.C., **The Stellar Populations and Evolution of Lyman Break Galaxies**, astro-ph/0105087v1 (2001).

[35] Pozzetti L. *et al.*, **The VIMOS VLT Deep Survey. The Assembly History of the Stellar Mass in Galaxies: from the Young to the Old Universe**, astro-ph/0704.1600v1 (2007).

[36] Press W.H. *et al.*, **Numerical Recipes in C: the Art of Scientific Computing**, Cambridge University Press, 1994.

[37] Renzini A., **Very Massive Galaxies: A Challenge For Hierarchical Models?**, astro-ph/0702148v1 (2007).

[38] Riess A.G. *et al.*, **Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant**, astro-ph/9805201v1 (1998).

[39] Sanders D.B. *et al.*, **S-COSMOS: The Spitzer Legacy Survey of the Hubble Space Telescope ACS 2 deg^2 COSMOS Field I: Survey Strategy and First Analysis**, ApJS, 172:86-98 (2007).

[40] Schlegel D.J., Finkbeiner D.P. and Davis M., **Maps of Dust Infrared Emission for Use in Estimation of Reddening and Cosmic Microwave Background Radiation Foregrounds**, ApJ, 500, 525 (1998).

[41] Shapley, A.E. *et al.*, **The Rest-frame Optical Properties of z $\simeq$ 3 Galaxies**, astro-ph/0107324v1 (2001).

[42] Shim H., Myungshin I., Choi P., Yan L. and Storrie-Lombardi L., **Massive Lyman Break Galaxies at z $\sim$ 3 in the Spitzer Extragalactic First Look Survey**, astro-ph/0708.0885v1 (2007).

[43] Siana B. *et al.*, **High-Redshift QSOs in the SWIRE Survey and the z $\sim$ 3 QSO Luminosity Function**, astro-ph/0711.0211v1 (2007).

[44] Somerville R.S. *et al.*, **Cosmic Variance in the Great Observatories Origins Deep Survey**, astro-ph/0309071v1 (2003).

[45] Songaila A., **The Evolution of the Intergalactic Medium Transmission to Redshift Six**, astro-ph/0402347v1 (2004).

[46] Songaila A. and Cowie L.L., **Approaching reionization: the evolution of the Lyman alpha forest from redshifts four to six**, astro-ph/0202165v1 (2002).

[47] Stanway E.R., Bunker A.J. and McMahon R.G., **Lyman break galaxies and the star formation rate of the Universe at z $\approx$ 6**, MNRAS 342, 439-445 (2003).

[48] Stark, D.P. *et al.*, **The evolutionary history of Lyman break galaxies between redshift 4 and 6: observing successive generations of massive galaxies in formation**, astro-ph/0902.2907v1 (2009).

[49] Steidel, C.C., Adelberger K.L., Giavalisco M., Dickinson M. and Pettini M., **Lyman break galaxies at z $\gtrsim$ 4 and the evolution of the UV luminosity density at high redshift**, astro-ph/9811399v2 (1999).

[50] Stetson P.B., **DAOPHOT: A Computer Program for Crowded-Field Stellar Photometry**, PASP 99: 191-222 (1987).

[51] Trenti M. and Stiavelli M., **Cosmic Variance and Its Effect on the Luminosity Function Determination in Deep High z Surveys**, astro-ph/0712.0398v1 (2007).

[52] Vanzella E. *et al.*, **Spectroscopic Observations of Lyman-Break Galaxies at Redshift $\sim$ 4, 5 and 6 in the GOODS-South Field**, astro-ph/0901.4364v1 (2009).

[53] Verma A., Lehnert M.D., Forster Schreiber N.M., Bremer M.N. and Douglas L., **Lyman-break galaxies at z $\sim$ 5 - I. First significant stellar mass**

assembly in galaxies that are not simply z $\sim$ 3 LBGs at higher redshift, astro-ph/0701725v1 (2007).

[54] Willott C.J., private communication.

[55] Willott C.J., Delfosse X., Forveille T., Delorme P. and Gwyn S.D.J., **First Results from the Canada-France High-z Quasar Survey: Constraints on the z = 6 Quasar Luminosity Function and the Quasar Contribution to Reionization**, astro-ph/0507183v1 (2005).

[56] Yabe K. *et al.*, **The Stellar Populations of Lyman Break Galaxies at z $\sim$ 5**, astro-ph/0811.2041v1 (2008).

[57] Yoshida M. *et al.*, **Luminosity Functions of Lyman-Break Galaxies at z $\sim$ 4 and 5 in the Subaru Deep Field**, astro-ph/0608512v1 (2006).

[58] http://map.gsfc.nasa.gov/news/5yr_release.html

[59] http://wmap.gsfc.nasa.gov/universe/uni_expansion.html

[60] http://www.atlasoftheuniverse.com/redshift.html

[61] http://idlastro.gsfc.nasa.gov/

[62] http://edu-observatory.org/mcc/homework/homework.ch.18-19/homework.ch.18-19.html

[63] http://nedwww.ipac.caltech.edu/level5/Giavalisco/Giav1.html

[64] http://www.cfht.hawaii.edu/Science/CFHLS/

[65] http://www.ast.obs-mip.fr/article204.html

[66] http://www.cfht.hawaii.edu/Instruments/Imaging/MegaPrime/

[67] http://www2.cadc-ccda.hia-iha.nrc-cnrc.gc.ca/community/CFHTLS-SG/docs/extra/filters.html

[68] http://terapix.iap.fr/

[69] http://www3.cadc-ccda.hia-iha.nrc-cnrc.gc.ca/cfht/T0004.html

[70] http://www.astro.utoronto.ca/∼patton/astro/mags.html

[71] http://terapix.iap.fr/cplt/table_syn_T0006.html

[72] http://www.starlink.rl.ac.uk/star/docs/sc5.htx/node15.html

[73] http://cosmos.astro.caltech.edu/

[74] http://acs.pha.jhu.edu/instrument/overview/

[75] http://irsa.ipac.caltech.edu/data/COSMOS/images/spitzer/mips/MIPS_24 _GO3_INFO.txt

[76] http://irsa.ipac.caltech.edu/data/COSMOS/gator_docs/cosmos_xmm_colDescriptions.html

[77] http://irsa.ipac.caltech.edu/data/COSMOS/spectra/magellan/MAGELLAN_INFO.txt

[78] http://irsa.ipac.caltech.edu/data/COSMOS/spectra/z-cosmos/Z-COSMOS_INFO.html

[79] http://www.astro.princeton.edu/∼schlegel/dust/

[80] http://spider.ipac.caltech.edu/staff/jarrett/irsa/dust.html

[81] http://lambda.gsfc.nasa.gov/product/cobe/

[82] http://irsa.ipac.caltech.edu/IRASdocs/iras.html

[83] http://ssc.spitzer.caltech.edu/irac/

[84] http://irsa.ipac.caltech.edu/data/COSMOS/tables/scosmos/scosmos_irac_200706_colDescriptions.html

[85] http://www.spitzer.caltech.edu/

[86] http://www.ifa.hawaii.edu/∼ilbert/S-Cosmos/

[87] http://www2.cadc-ccda.hia-iha.nrc-cnrc.gc.ca/cfht/WIRDST0001.html

[88] http://terapix.iap.fr/rubrique.php?id_rubrique=261

[89] http://www.cfht.hawaii.edu/Instruments/Imaging/WIRCam/

[90] http://irsa.ipac.caltech.edu/data/COSMOS/index_cutouts.html

[91] http://www.ipac.caltech.edu/2mass/releases/sampler/ancillary/ptsource.format.html

[92] http://idlastro.gsfc.nasa.gov/ftp/pro/idlphot/aper.pro

[93] http://web.ipac.caltech.edu/staff/jarrett/irac/calibration/ext_apercorr.html

[94] http://www.oamp.fr/virmos/vvds.htm

[95] http://www.eso.org/sci/facilities/paranal/instruments/vimos/index.html

[96] http://cencosw.oamp.fr/EN/index.en.html

[97] http://www3.cadc-ccda.hia-iha.nrc-cnrc.gc.ca/megapipe/docs/photometry.html

[98] http://marcs.astro.uu.se/

[99] http://www.sdss.org/

[100] www.astro.uvic.ca/∼hsiao/talks/hsiao_dust.pdf

[101] http://www.stsci.edu/science/goods/

[102] http://www.roe.ac.uk/ukatc/projects/scubatwo/

[103] http://astro.uwaterloo.ca/SCUBA2/

Appendix A

Terminology

α	opacity [m^{-1}] , a measure of the effectiveness of dust in diminishing radiation
Å	angstrom, 10^{-10} m
A	extinction of light [mag]
AB magnitude	standard magnitude system in which an object with a constant flux per unit frequency exhibits zero colour
absolute magnitude	intrinsic brightness of an object measured as the apparent magnitude at a distance of 10 parcecs ($\sim 3.1 \times 10^{17}$ m)
ACS	Advanced Camera for Surveys
AGN	active galactic nuclei, centers of galaxies emitting vast amounts of energy
APER	a routine used to compute concentric aperture photometry from astronomical images
apparent magnitude	a measure of the brightness of an astronomical object as seen from Earth, based on a logarithmic scale

arcmin (')	arcminute, 1/60 of a degree or 60 arcsec
arcsec (")	arcsecond, 1/3600 of a degree or 1/60 arcmin
BC model	Bruzual and Charlot galaxy models, specified by year
χ^2	chi-square test statistic
CADC	Canadian Astronomy Data Centre
CB2008	Charlot and Bruzual galaxy models, 2008 series
CCD	charge-coupled device, converts photons to electrons
CDF	Chandra Deep Field (-S south)
CDM	cold dark matter
CFHQS	Canada-France High-z Quasar Survey
CFHT	Canada-France-Hawaii Telescope
CFHTLS	CFHT Legacy Survey
Ch 1(2,3,4)	Spitzer IRAC Channels
COBE	Cosmic Background Explorer
colour index	the difference in magnitude between two filters; a small colour index indicates a bluer source, a large colour index indicates a redder source
cosmic variance	the field-to-field variation in observational number counts of galaxies due to the presence of large-scale structure
COSMOS	Cosmic Evolution Survey
D1, D2, D3, D4	CFHTLS Deep Fields
DAOPHOT	Dominion Astrophysical Observatory (DAO) Photometry program
Dec	declination [deg], astronomical coordinate analogous to latitude

deg (°)	degree, 1/360 of a circle
DIRBE	Diffuse Infra-Red Background Experiment
dropout	an object which is not detected in a certain filter due to the absorption of light emitted by that object at filter wavelengths
E(B-V)	selective extinction [mag], A(B) - A(V)
ecliptic plane	the plane of the Earth's orbit around the Sun; the apparent annual path of the Sun across the sky
effective temperature	the surface temperature of a star, for example, expressed as the temperature of a black body with the same total flux density
ESA	European Space Agency
ESO	European Southern Observatory
F	flux density [Jy], the power of radiation per unit area for a given frequency or wavelength
F_0	flux density that defines a magnitude of zero
FWHM	full width at half the maximum peak height
galactic plane	the plane of the disk of the Milky Way Galaxy
GOODS	Great Observatories Origins Deep Survey (-N north, -S south, -FF flanking fields)
HDF	Hubble Deep Field (-N north)
HR diagram	Hertzsprung-Russell diagram of stellar properties
HST	Hubble Space Telescope
I_{AB}	flux in a standard I (infrared) filter based on the AB magnitude system
IDL	Interactive Data Language

IGM	intergalactic medium, the matter found in the space between galaxies
IMACS	Inamori Magellan Areal Camera and Spectrograph
IMF	initial mass function, the distribution of stars formed as a function of mass
interloper	an unwanted object within the high redshift galaxy data set - i.e. a star, quasar or low redshift galaxy
IR	infrared
IR1(2)	Spitzer IRAC Channel 1 (2)
IRAC	Infrared Array Camera
IRAS	Infrared Astronomical Satellite
IRS	Infrared Spectrograph
isochrone	a connection along the evolutionary track of stars of different initial masses at a specific point in time
ISM	interstellar medium, the matter found in the space between stars
ISSA	IRAS Sky Survey Atlas
Jy	Jansky, unit of flux density, 10^{-26} W m^{-2} Hz^{-1}
KPNO	Kitt Peak National Observatory
λ	wavelength [m]
Λ	cosmological constant, dark energy parameter
ΛCDM	standard cosmological model incorporating dark energy (Λ) and cold dark matter (CDM)
L	observed flux [W]
LBG	Lyman-break galaxy

limiting magnitude	the magnitude of the faintest object that can be reliably detected in an image
longward	at wavelengths longer than the given wavelength
luminosity	the rate at which energy is emitted by an astronomical object
Lyman series	the spectral lines that result from electron transitions with respect to the ground state of the hydrogen atom
m	apparent magnitude [mag]
$M_{\odot}$	solar mass, a measure in units of the mass of the Sun, $\sim 1.989 \times 10^{30}$ kg
M_{sun}	solar mass, as above
$M_{\odot}$ yr^{-1}	solar masses per year, unit for star formation rate
mag	unit of magnitude, a difference of 1 mag corresponds to a factor of 2.512 in brightness
MARCS	Model Atmosphere Radiative and Convective Scheme
MegaCam	CFHT one square-degree field of view camera for optical data
MegaPrime	CFHT wide field imager
micron	micrometer (μm), 10^{-6} m
MIPS	Multiband Imaging Photometer for Spitzer
non-detection	an object that is fainter than the detection limit
P	confidence level, P=1-Q
PHT	Palomar Hale Telescope
PI	principal investigator
PSF	point spread function, the distribution of flux as a result of distortion of a point source in an image

PSM	population synthesis model
Q	quantile representing chance probability
quasar	quasi-stellar radio source, a distant galaxy with an extremely bright and active core
QSO	quasi-stellar object (radio-quiet)
RA	right ascension [deg], astronomical coordinate analogous to longitude
redshift	the lengthening of the rest-frame wavelength of light emitted by cosmically distant sources (see Equation 2.1)
RMS	root mean square
S-COSMOS	Spitzer Legacy Cosmic Evolution Survey
SCUBA 2	Submillimeter Common-User Bolometer Array 2
SDSS	Sloan Digital Sky Survey
SE	standard 1σ error associated with photometric measurements [mag]
SED	spectral energy distribution
SExtractor	Source Extraction software by Bertin and Arnouts (1996)
SFH	star formation history
SFR	star formation rate
shortward	at wavelengths shorter than the given wavelength
S/N	signal-to-noise ratio
spectral class	classification system for stars based on spectral properties
Spitzer	Spitzer Space Telescope
sr	steridian, unit of solid angle

SSP	simple stellar population, passively aging population in which no new stars are formed following an initial instantaneous burst
stellarity	a parameter used to separate stars from galaxies; stars are point-like (spatially unresolved) with high stellarity; galaxies are extended objects (spatially resolved) with low stellarity
Subaru	Subaru Telescope
τ	optical depth of dust extinction [dimensionless]
ugriz	CFHTLS MegaCam standard optical u*, g′, r′, i′, z′ filter set
UV	ultra-violet radiation
V band	optical broad-band filter with central wavelength $\sim$0.55 μm
VIMOS	VIsible Multi-Object Spectrograph
VLT	Very Large Telescope
VVDS	VIMOS VLT Deep Survey
WHT	William Herschel Telescope
WIRCam	Wide-field InfraRed Camera
WIRDS	WIRCam Deep Survey
WMAP	Wilkinson Microwave Anisotropy Probe
XMM-COSMOS	catalogue of COSMOS point-like x-ray sources
XMM-Newton	X-ray Multi-Mirror Mission - Newton space-based telescope
z	redshift parameter [dimensionless]
zCOSMOS	COSMOS spectroscopic redshift survey