GALAXY SPECTRAL ANALYSIS IN THE ERA OF LARGE-SCALE GALAXY SURVEYS

This thesis is submitted in partial fulfilment of the requirements for the award of the degree Doctor of Philosophy of the University of Portsmouth

January 2015

By
Oliver Steele

Institute of Cosmology and Gravitation

Supervisor and Director of Studies: Dr. Daniel Thomas

Second supervisor: Dr. Claudia Maraston
Contents

Acronyms and Abbreviations 10
Abstract 12
Declaration 14
Copyright 15
Acknowledgements 16

1 Introduction 17
  1.1 In the Beginning... 17
  1.2 Observing in the 20th Century 21
  1.3 Galaxy Formation and Evolution 26
  1.4 The Survey Landscape 34
  1.5 Aims 37

2 Analysis Tools 38
  2.1 Pipeline overview 40
  2.2 GANDALF: Gas AND Absorption Line Fitting 42
      2.2.1 pPXF 45
      2.2.2 Stellar population templates 46
      2.2.3 Kinematics calibration 51
      2.2.4 Reddening analysis 52
  2.3 BPT diagnostic diagram 57
  2.4 Absorption line measurements 60
      2.4.1 Stellar population parameters 62
  2.5 Testing the pipeline with SDSS-I/II DR7 64
3 The GAMA survey

3.1 Our data products
3.1.1 Features of the spectra
3.2 External data products
3.2.1 Local environment
3.2.2 Stellar mass estimates
3.2.3 Group catalogue
3.2.4 Gaussian emission line models
3.3 The environmental dependence of star formation and ionisation characteristics in the GAMA survey
3.3.1 Background
3.3.2 Sample selection
3.3.3 $M_\ast - \rho$ correlation
3.3.4 Controlling for a variable
3.3.5 Emission Line Properties
3.3.6 Main sample
3.3.7 Non-/Group samples
3.3.8 Conclusions
3.4 Further science based on our pipeline data
3.4.1 Ching et al. (prep)
3.4.2 Lara-López et al. (2013)
3.4.3 Lara-López et al. (2013)
3.4.4 Hardcastle et al. (2013)
3.4.5 Foster et al. (2012)

4 SDSS-III/BOSS

4.1 Spectroscopic properties of BOSS DR9 galaxies
4.1.1 Stellar velocity dispersions
4.1.2 Emission lines
4.1.3 Target selection colour-colour space
4.1.4 Conclusions
4.2 Stellar population analysis of high-S/N BOSS galaxies
4.2.1 Stacking
4.2.2 Lick indices
4.2.3 Stellar population parameters
4.2.4 Conclusion
4.3 Contributions to the collaboration ........................................... 178
  4.3.1 DR9 ........................................................................... 178
  4.3.2 DR10 ......................................................................... 178
  4.3.3 Beifiori et al. (sub) .......................................................... 178

5 Conclusions ............................................................................. 182

A Catalogue products .................................................................. 195

B Emission lines .......................................................................... 197

C Lick indices .............................................................................. 199

D Hα offset spectra ..................................................................... 201

E Hβ offset spectra ..................................................................... 203
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Tested $\lambda$ ranges for kinematics analysis</td>
<td>53</td>
</tr>
<tr>
<td>3.1</td>
<td>GAMA field information</td>
<td>68</td>
</tr>
<tr>
<td>3.2</td>
<td>Emission lines used to estimate $z$ in GAMA</td>
<td>74</td>
</tr>
<tr>
<td>3.3</td>
<td>GAMA spectra selection results</td>
<td>77</td>
</tr>
<tr>
<td>3.4</td>
<td>GAMA EW distribution stats</td>
<td>79</td>
</tr>
<tr>
<td>3.5</td>
<td>$M_*-\rho$ relation significance</td>
<td>94</td>
</tr>
<tr>
<td>4.1</td>
<td>BOSS BPT/L10 comparison</td>
<td>144</td>
</tr>
<tr>
<td>4.2</td>
<td>Emission line counts in BOSS colour-colour diagram</td>
<td>151</td>
</tr>
<tr>
<td>4.3</td>
<td>BOSS stacked object statistics</td>
<td>157</td>
</tr>
<tr>
<td>4.4</td>
<td>Statistical analysis of SSP parameters vs $\sigma$ and $z$</td>
<td>171</td>
</tr>
</tbody>
</table>
# List of Figures

1.1 A geocentric solar system diagram ........................................... 18
1.2 Hubble Sequence ........................................................................ 20
1.3 SDSS $ugriz$ filter response functions ........................................ 22
1.4 Example high-SN galaxy spectra .................................................. 24
1.5 Model of CDM hierarchical formation .......................................... 27
1.6 Comparing the Millennium simulation with observations ............... 29
1.7 Formation downsizing .................................................................. 31
1.8 Assembly downsizing, Pozzetti (2010) .......................................... 32
1.9 Assembly downsizing, Perez-Gonzalez (2008) ................................. 33
1.10 Large scale structure from the CfA survey ..................................... 35

2.1 Pipeline workflow ....................................................................... 39
2.2 Single stellar population models used in our fits ............................ 48
2.3 Looking at the effect of resolution on templates ............................. 50
2.4 Investigating the optimal $\lambda$ range for kinematics analysis .......... 52
2.5 gandalf vs MPA-JHU EWs with 2d reddening ................................. 54
2.6 E(B-V) histograms for 2d reddening ............................................. 55
2.7 gandalf vs MPA-JHU EWs with 1d reddening ................................. 56
2.8 Example BPT diagnostic diagram .................................................. 59
2.9 Example absorption feature showing the Lick index definitions ........ 61
2.10 Testing out $\sigma$ measurements against those of DR7 .................... 65
2.11 Testing our EW measurements against those of DR7 .................... 66

3.1 GAMA survey cone .................................................................... 69
3.2 GAMA redshift distributions ....................................................... 74
3.3 GAMA $\sigma$ distributions ............................................................ 77
3.4 GAMA survey EW distributions .................................................. 78
3.5 BPT for entire GAMA survey ..................................................... 80
‘Darkness took me, and I strayed out of thought and time... The stars wheeled overhead, and every day was as long as a life age of the earth... But it was not the end. I felt life in me again.’
– Gandalf the White
<table>
<thead>
<tr>
<th>Acronyms and Abbreviations</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAO</td>
</tr>
<tr>
<td>AGE</td>
</tr>
<tr>
<td>AoN</td>
</tr>
<tr>
<td>APO</td>
</tr>
<tr>
<td>BAO</td>
</tr>
<tr>
<td>BOSS</td>
</tr>
<tr>
<td>BPT</td>
</tr>
<tr>
<td>(Λ)CDM</td>
</tr>
<tr>
<td>DMU</td>
</tr>
<tr>
<td>DR</td>
</tr>
<tr>
<td>EW</td>
</tr>
<tr>
<td>FOV</td>
</tr>
<tr>
<td>FWHM</td>
</tr>
<tr>
<td>GAMA</td>
</tr>
<tr>
<td>GANDALF</td>
</tr>
<tr>
<td>IMF</td>
</tr>
<tr>
<td>L10</td>
</tr>
<tr>
<td>LINER</td>
</tr>
<tr>
<td>LOSVD</td>
</tr>
<tr>
<td>LRG</td>
</tr>
<tr>
<td>MILES</td>
</tr>
<tr>
<td>MPA/JHU</td>
</tr>
<tr>
<td>PPXF</td>
</tr>
<tr>
<td>QC</td>
</tr>
<tr>
<td>SAM</td>
</tr>
<tr>
<td>SAURON</td>
</tr>
<tr>
<td>Abbreviation</td>
</tr>
<tr>
<td>--------------</td>
</tr>
<tr>
<td>SDSS</td>
</tr>
<tr>
<td>SED</td>
</tr>
<tr>
<td>SF</td>
</tr>
<tr>
<td>SFR</td>
</tr>
<tr>
<td>S/N</td>
</tr>
<tr>
<td>SSFR</td>
</tr>
<tr>
<td>SSP</td>
</tr>
<tr>
<td>TMJ</td>
</tr>
<tr>
<td>(z)</td>
</tr>
<tr>
<td>(Z)</td>
</tr>
</tbody>
</table>
Abstract

In this work I address two of the big questions in modern astrophysics; the role of environment as a driver of galaxy evolution, and the role of mass in star formation and stellar population evolution. I use one of the most powerful tools available to the astrophysical community, large-scale galaxy spectroscopy, to contribute towards the answers to these dilemmas. I construct a data analysis pipeline based on the public codes GANDALF and pPXF to extract gas and stellar dynamics, emission line statistics, absorption line indices and stellar population parameters from these galaxy spectra. I test and calibrate this pipeline against existing results for the Sloan Digital Sky Survey Data Release 7, and find it to provide accurate measurements.

I use the emission line results from this to probe the dependence of star formation and ionisation characteristics on stellar mass, local environment and global environment in the Galaxy AND Mass Assembly survey. I find that mass is the main driving factor behind the presence of star formation and determining different ionisation sources, and see a trend with increasing mass from star forming objects to those hosting active galactic nuclei via composites of the two. Local density plays a role only at the highest densities, and is considerably less significant than mass; global environment is found to have negligible impact. This suggests that star formation quenching is primarily a mass-driven process, with active galactic nucleus feedback being a likely candidate for the environment independent process involved in our sample.

I stack objects together from the Sloan Digital Sky Survey III: Baryon Oscillation Spectroscopic Survey in order to produce high-signal-to-noise spectra for the purpose of absorption line measurement and the subsequent modelling of stellar population parameters. I use this to investigate the dependence of age, metallicity and $\alpha$/Fe on mass (using stellar velocity dispersion as a proxy for dynamical mass) and redshift. I find that light-averaged age, metallicity and
$\alpha$/Fe all increase with velocity dispersion, which are predictions of the downsizing paradigm, where the least massive galaxies form their stars later, over more extended timeframes and less efficiently than more massive galaxies. Age is also seen to increase with redshift, which is simply the result of everything in the Universe getting older, whilst I see no evidence of metallicity or $\alpha$/Fe changing with lookback time. Investigating how galaxies age when compared to the Universe, I find that more massive galaxies appear to age faster than the Universe whilst less massive galaxies age slower. I hypothesise that this is due to the different star formation histories of galaxies with differing masses, and test this by compiling models with varying stellar histories and comparing them to our observations. I find that as mass decreases, I require more extended periods of star formation that peak more recently. At the high-mass end, the relationship between the most massive bins is best reproduced by a passively evolving population whose stars formed at higher redshift than I observe. This is a clear result of downsizing, and sets tough restrictions on future models of galaxy formation and evolution.
Declaration

Whilst registered as a candidate for the above degree, I have not been registered for any other research award. The results and conclusions embodied in this thesis are the work of the named candidate and have not been submitted for any other academic award.

This research has been supported by the Science and Technology Facilities Council (STFC).

The work in Chapter 3 is partially published in Thomas et al. (2013) and the remainder will be published in Steele et al. (2014a). The work in Chapter 4 will be published in Steele et al. (2014b).

Word count: 33,438
Copyright

Copyright in text of this thesis rests with the Author. Copies (by any process) either in full, or of extracts, may be made only in accordance with instructions given by the Author. This page must form part of any such copies made. Further copies (by any process) of copies made in accordance with such instructions may not be made without the permission (in writing) of the Author.

The ownership of any intellectual property rights which may be described in this thesis is vested in the University of Portsmouth, subject to any prior agreement to the contrary, and may not be made available for use by third parties without the written permission of the University, which will prescribe the terms and conditions of any such agreement.

Further information on the conditions under which disclosures and exploitation may take place is available from the head of the Institute of Cosmology and Gravitation.
Acknowledgements

There are numerous people without whom this thesis would never have existed, and to each of them I owe a debt of gratitude. Words alone cannot show my appreciation, but in this instance they are all I have.

First I would like to thank my supervisors, Daniel and Claudia, for their constant guidance, help and support. Their dedication and belief in my work inspired me to achieve more than I otherwise would have thought possible.

I thank my PhD examiners, David Bacon and Ignacio Ferreras, for their exceptionally useful comments and encouragement. They engaged me in a very enjoyable and informative viva voce, and helped improve and clarify this thesis as a whole.

I also thank the computer support staff, Edd and Gary, without whom I would have floundered long ago. They have helped me overcome innumerable obstacles, and have facilitated all of the work presented here.

I would also like to thank the Institute of Cosmology and Gravitation for giving me the opportunity to conduct my PhD, and its members (past and present) who have helped me throughout the years. Special mentions go to David Wilkinson, Jonas Johansson, Tim Higgs and Rob ‘Stats Deity’ Crittenden.

A special thank you goes out to my parents, Adrienne and James, and my brother, Theo, for always being there. Without them backing me up and supporting me I would not be where I am today. I am fortunate to have close family as great as you.

Finally, I would like to thank my grandmother Donnie and all my close friends, especially Anita, Chris, Robyn, Mo, Olly, Taichi, Jon and Elli, for all their encouragement throughout the years.
Chapter 1

Introduction

1.1 In the Beginning..

Mankind has been fascinated by the heavens for as long as history has been recorded. From the framework of ancient calendars to the materials of mystics and soothsayers, extraterrestrial bodies have been important to civilisation for millennia. The significance that has been given to understanding and ‘harnessing’ celestial motions is perhaps best demonstrated by the age of the oldest known observatory, the Goseck Circle, built in Saxony-Anhalt, Germany around 4,900 BC - nearly 2,500 years before the Great Pyramid of Giza.

Early models of the Solar System were usually based on geocentric concentric spheres, an idea supported by Plato and Aristotle and famously touted by Ptolemy of Alexandria in the 2nd century AD. These models used series of concentric circles (epicycles) placed on larger concentric circles (deferents) to try and explain the complex motions of planets and stars through the sky (shown as a toy model in Fig. 1.1), and were generally accepted for over 1,400 years until Nicolaus Copernicus (1473-1543) produced a viable heliocentric model that challenged Ptolemaic views. This marked the beginning of a revolution in the realm of astronomy and sent a shockwave through the academic world, fuelling many people to challenge commonly held views and change the face of our understanding of the Universe forever.

Two of the most well known and important figureheads of this ‘revolution’ were Galileo Galilei (1564-1642) and Johannes Kepler (1571-1630), who between them spearheaded research into heliocentricity and modern astronomy. Galileo
Figure 1.1: A toy model of a Ptolemaic geocentric solar system; black orbits show the objects’ deferents, or paths around the Earth, and red circles show the objects’ epicycles, which are imposed over the deferents. Note that here distances are not representative. Image taken from http://www.teachastronomy.com/astropedia/article/Copernicus-and-the-Heliocentric-Model.
used his own improvements to telescopes to observe the moons of Jupiter, witness-
ing for the first time that bodies could orbit things other than the Earth. He also
observed the phases of Venus, the only explanation for which was the Copernican
system. Kepler studied the motions of planets in detail, and produced models
based on heliocentricity and, uniquely for the time, elliptical orbits.

These observations and theories culminated in the work of Isaac Newton
(1642-1727), who brought them all together with his theories of gravity and
motion. Now with physical origins to explain Kepler’s laws, the dominance of
heliocentricity was cemented in astrophysics. With increasingly powerful tele-
scopes it became possible to observe more and more of the surrounding Universe,
allowing for the classification of other stars as being similar in structure to our
Sun, and leading to Thomas Wright’s (1711-1786) theory that our solar system
was just one of many existing in a flattened disk (now known as the Milky Way).

However, not all stars were seen to behave in precisely the same way as our
Sun. A star with varying luminosity was observed in 1784 by Edward Pigott
(1753-1825), followed by the documentation of another one a few months later by
John Goodricke (1764-1786). These are what is now known as Type I Cepheids;
stars that cycle between larger, brighter states and smaller, dimmer ones. It was
discovered in 1908 (and published in 1912) by Henrietta Swan Leavitt (1868-
1921) that there was a distinct relation between period and luminosity for these
stars [Leavitt & Pickering, 1912]. This allowed for distances outside of our Solar
System to be calculated for the first time by measuring the period of a Cepheid
and its apparent magnitude, and converting this to absolute magnitude where
the only free parameter is distance.

Another important observation was that of extended, cloudy sources, with un-
known physical origin, known as nebulae. Some people controversially suggested
that these may have been located outside of our galaxy, for example the philoso-
pher Immanuel Kant who believed that each nebula was an ‘Island Universe’
similar to our own galaxy and yet spatially distant (Kant, 1755), challenging
traditionally held views that our galaxy was everything in existence. This was
brought to a head by the discovery in 1913 that the Andromeda nebula had a
velocity of $-300$ km/s [Slipher, 1913], far greater than that observed of stars
at the time, leading to the ‘Great Debate’ of April, 1920. This was a debate
between two of the great astronomers of the time, Harlow Shapley (1885-1972)
and Heber Curtis (1872-1942), regarding the nature of nebulae; Shapley arguing
Figure 1.2: The Hubble sequence, a morphological classification of different galaxy types introduced by Edwin Hubble in 1926. Image taken from http://www.universetoday.com/50428/dark-energy-model-explains-hubble-sequence-of-galaxies/.

for the view that the Milky Way was the entirety of the Universe, whilst Curtis supported the now commonly accepted idea that nebulae were extragalactic in nature. Curtis supported his argument by presenting evidence that the number of novae observed in Andromeda far exceeded that seen in other regions of our galaxy, which would be difficult to explain were it not located outside of our Milky Way, but the debate did not finish conclusively, and the nature of these nebulae was still uncertain.

It was not until the work of Edwin Hubble (1889-1953) that it was conclusively shown that Andromeda is, indeed, extragalactic. He used Cepheid variables to measure the distance from Earth to various nebulae, and found them to be located far outside of the bounds of the Milky Way (Hubble, 1926), finally proving that the Universe extended beyond our own galaxy. He further extended this discovery in the same paper by classifying galaxies by their shape, a system known as the Hubble sequence (Hubble, 1926).

Fig. 1.2 illustrates Hubble’s morphological classification scheme. This is a purely empirical separation of objects based on their shape; on the left are ‘elliptical’ galaxies, where the proceeding number indicates the ellipticity based on the
major and minor axes from a head-on view of the object, with 0 indicating no to minimal elongation and 7 indicating maximal elongation. Galaxies on the right are all those observed to have structure in the form of spiral arms. The top row are those galaxies with spiral arms protruding directly from the central bulge of the galaxy, where the proceeding letter indicates the tightness of the wound arms and the corresponding expanse of the central bulge; ‘a’ indicates tightly wound arms with larger bulges, ‘c’ indicates loosely wound arms with small bulges, and ‘b’ falls in-between. The bottom row on the right follows the same general trend as the top row, but where galaxies are also observed to contain bars extending from the bulge to the beginning of the arms. Elliptical galaxies are generally known as ‘early-types’, and spirals ‘late-types’, due to their positions on this diagram; this definition has nothing to do with formation times or processes. These discoveries paved the way for future astronomers, and the study of these ‘Island Universes’ is now the focus of much modern astrophysics.

1.2 Observing in the 20th Century

Since the discovery of these new structures, one of the main goals in astrophysics has been to understand their formation and evolution. In order to do this it is necessary to know the physical features of these galaxies, and to try and understand how they interact with one another. Since the only current way to investigate a galaxy is through observing its light - obviously one cannot stick Andromeda on a set of scales to obtain its mass, or cut it in half and count its rings to get its age - we have had to come up with clever ways to exploit this resource as best we can. In the near future it is likely that the detection of neutrinos and gravitational waves will play an important part in probing and understanding galaxies, which will open up the field of astrophysics to a wealth of new data.

Due to the complexities and timescales of observing, we essentially have two methods of taking data: we can maximise the spatial coverage of our surveys at the sacrifice of detailed information, or we can maximise the amount of information from individual sources but limit ourselves spatially. The former is known as photometric imaging; the latter, spectroscopic imaging.

In the photometric approach, light is integrated over fixed-width wavelength bands using filters, smoothing over the effects at individual wavelengths but allowing for images of large areas of the sky to be taken at once. The great advantage
of photometric surveys is their ability to cover vast spatial regions with impressive spatial resolution over very short timescales, due to the broad wavelength bands collecting the light (bands of width $\sim 1000\text{Å}$ are not unusual). The Sloan Digital Sky Survey (SDSS, York & SDSS Collaboration 2000) epitomises this ability as the largest all-sky survey ever performed, having now provided 14,555 deg$^2$ of optical imaging data up to the 8th data release (DR8, the final SDSS photometric release Aihara et al., 2011) - over a third of the Celestial Sphere - in five bands ($u, g, r, i, z$) with a point spread function (PSF) full width half-maximum (FWHM) of 1.2-2.3” Stoughton et al. (2002). The bands represent the regions within which light was integrated, and the Sloan $ugriz$ system is shown in Fig. 1.3 as a representative example of photometric filter functions. Whilst the perfect filter function would be a step function, this is clearly not possible due to the realities involved with materials in filters; their chemical make-up does not allow them to perfectly allow through the light that is meant to pass a filter.
CHAPTER 1. INTRODUCTION

To receive the maximum possible amount of information from a source, however, we should observe the flux density - the flow rate intensity of electromagnetic energy - in small wavelength intervals (typically $\sim 1\text{Å}$). This is known as spectroscopy (see Lilly et al., 1995; Falco et al., 1999; York & SDSS Collaboration 2000; Colless et al., 2001; Cannon et al., 2006; Lilly et al., 2007a; Drinkwater et al., 2010; Jones et al., 2009; Driver et al., 2011, for examples). The advantage of spectroscopy is the potentially high spectral resolution obtainable; the main disadvantage is the amount of time required to get a high signal-to-noise spectrum. As a result of this, specific targets have to be selected in order to avoid wasting time investigating space not of interest to a given survey.

Once obtained, however, galaxy spectra are veritable gold mines of information. Fig. 1.4 shows two examples of high signal-to-noise (S/N) spectra, constructed by stacking together individual, lower SN galaxy spectra ($\sim 4400$ for the top figure, and $\sim 16000$ for the bottom). The S/N of the top spectrum is 321.1, and of the bottom spectrum is 9.9. The comparatively high S/N of the top spectrum is due to its constituent objects being selected for having high masses (and therefore high luminosities), and also for being relatively nearby with a mean redshift of $z \sim 0.16$. The galaxies making up the top spectrum were selected as objects with $0.1 < z < 0.2$ and velocity dispersions of $250 < \sigma < 315$ from the SDSS-III/BOSS (Ahn et al., 2013), and those in the bottom spectrum were specifically selected as galaxies with emission lines from the Galaxy And Mass Assembly (GAMA) survey (Driver et al., 2011, see Section 1.4 for more details).

These spectra each clearly display very important, very different features. The top spectrum shows obvious absorption features - small regions of the spectrum where flux is absent, resulting in a dip. The bottom spectrum shows strong emission features - small regions of the spectrum with an abundance of flux, resulting in Gaussian-like peaks. These are both caused by the quantum nature of matter, and can tell us very different, key properties about the source object.

Emission lines are sourced from the gas content of a galaxy, and are caused by the presence of major ionising sources. There are two types of emission lines: recombination lines brought about by photon emission when a free electron recombined with an ion; and collisional excitation lines caused by electrons being stripped from their host atoms, colliding with another atom/ion and exciting them, and then emitting a photon when transitioning back down from this excited state. As both of these are caused by energy level transitions, the photons
Figure 1.4: Example high signal-to-noise (SN) galaxy spectra, made by stacking together \( \sim 4400 \) SDSS-III/BOSS objects (top) and \( \sim 16000 \) GAMA objects selected to have emission lines (bottom).
emitted have very specific energies equal to the energy difference between the levels involved, which is itself dependent on the element being excited and the levels being transitioned. Were the atoms involved not moving at all, the emission features would be infinitely thin due to the specificity of these energies; instead, they are broadened by the motion of atoms in the radial direction. We can therefore use these widths to measure the velocity dispersions of the gas.

Further, since collisional excitation requires a much more powerful ionising source than recombination, we would expect them to be associated with a more powerful engine, such as active galactic nucleus (AGN) activity (energy emitted by the exceptional heat of an accretion disk around a supermassive black hole), whilst recombination lines can be caused by weaker engines such as star formation (Baldwin et al., 1981). We can therefore use the ratios of these lines to differentiate between powerful ionising sources in galaxies, which we go into more detail on in Section 2.3.

Absorption lines, on the other hand, are caused by the absorption of photons by the dense material within stars. Photons are emitted in the centres of stars and some are absorbed by atoms in the stellar atmospheres; when they are re-emitted they are lost from the beam heading directly towards us, resulting in a dip in flux at the energy level required to excite a given atom. This energy is again discrete but widened by motion along the line of sight, allowing us to measure the velocity dispersion of stars in a galaxy. We can also link the absorption features to the specific elements involved, allowing us to measure the abundance ratios of different elements in the stars of galaxies. This is elaborated on in Section 2.4.

Observations alone, however, are not enough to extract all galaxy parameters; a variety of models are required on top. Stellar population models of full spectral energy distributions can be constructed using models of starbursts (intense periods of star formation) simplified to a single stellar age (i.e. with star formation rates modelled as Dirac delta functions) (e.g. Maraston, 1998; Bruzual & Charlot, 2003; Maraston, 2005; Maraston & Strömbäck, 2011), and allow for the estimation of stellar masses, ages and metallicities amongst other things through fitting them to spectra or photometry. More detail on these can be found in Section 2.2.2.

Alternatively, stellar population models of absorption indices (e.g. Worthey et al., 1994; Trager et al., 2000; Thomas et al., 2003; Schiavon, 2007; Thomas et al., 2011) allow for the derivation of element abundance ratios in addition to
stellar ages and metallicities. As different chemicals are formed in different stars, comparing their ratios can tell us about the star formation history of a galaxy (e.g. Greggio & Renzini 1983; Thomas et al. 1998, 2005, 2010), which allows us to constrain galaxy evolution processes - for example, this approach allowed us to discover the ‘archaeological formation downsizing’ we discuss in Section 1.3.

With these tools arrayed around us, we are able to extract a wealth of information. From observations we know that galaxies have many intrinsic and extrinsic properties, such as their shape (morphology) (e.g. Hubble 1926; Dressler 1980; Masters et al. 2011), colour, mass (e.g. Maraston et al. 2013), luminosity (the total energy of their emitted light, e.g. Schechter 1976), dynamics (e.g. Rubin et al. 1980; Bosma 1981), and the characteristics of their stellar populations (e.g. Worthey et al. 1994; Trager et al. 2000; Thomas et al. 2003, 2005, 2010). We also know that they live in varying environments (e.g. Dressler 1980; Bamford et al. 2009; Peng et al. 2010; Rogers et al. 2010; Robotham et al. 2011), which also correlate with various properties.

In addition to all of the baryonic processes observed, we have also detected the presence of another type of matter, known as dark matter, which is believed to make up 84.5% of all matter in the Universe (Planck Collaboration et al. 2013). Originally introduced as a source of gravity to make up for a mass deficit observed when comparing galaxy dynamics to stellar mass estimated through luminosity (Zwicky 1933; Faber & Gallagher 1979) and to provide the required amount of mass to allow large gas clouds to collapse into galaxies (White & Rees 1978), it is now considered to be one of the principal components of galaxy formation due to the belief that baryonic matter follows the distribution of dark matter (e.g. Kauffmann et al. 1993; Ferreras et al. 2005).

1.3 Galaxy Formation and Evolution

As we cannot follow a single galaxy from its creation until the present day, it is necessary to create models of how we would expect a galaxy to evolve and then compare the predictions of this to observations. These comparisons then allow us to constrain the physics inserted into the models, which will eventually lead us to the correct physics behind how the Universe formed and evolved.

The goal of any galaxy formation/evolution model is to reproduce the observed
Figure 1.5: A schematic of CDM halo merging in an N-body simulation, where symbol size represents halo mass and the vertical axis represents time flowing downwards. The horizontal lines represent snapshots of the process, corresponding to timesteps in the simulation. *Image taken from* [Baugh (2006)](http://example.com).
distributions of and relationships between parameters. The most generally accepted view of large-scale galaxy structure evolution is the Cold Dark Matter (CDM) hierarchical galaxy formation model of [White & Rees (1978)]. This model uses purely gravitational interactions to produce the distribution of the mass component. Fig. 1.5 shows a simple schematic merger tree for a CDM halo, where symbol size indicates halo mass and the vertical axis represents time (flowing downwards).

In this model, galaxies form when baryonic gas cools and settles into dark matter haloes, which are subsequently built up by merging with one another (White & Rees, 1978). The hierarchical build-up of dark matter haloes has been modelled using N-body simulations (e.g. Davis et al., 1985; Springel et al., 2005; Guo et al., 2011), which have been very successful in reproducing the observed ‘cosmic web’ of large scale structure using Newtonian gravity. Fig. 1.6 from Springel et al. (2006) shows an example of this, visually comparing the famous Millennium Simulations (Springel et al., 2005) with various observations of large scale structure and showing the astounding accuracy with which a purely gravitational model can reproduce the observed cosmic structure.

Whilst the behaviour of dark matter is relatively easy to model due to its limited methods of interaction, baryonic matter interacts through numerous other modes and as such is considerably more difficult to simulate. There are two main methods of doing so: hydrodynamical models, which model the dynamics of resolved gaseous regions, and semi-analytic models, which use heuristic methods to populate predetermined dark matter haloes.

Hydrodynamical models follow the evolution of gas physics in a gravitational potential (taken from the sum of the baryonic and dark matter halo masses), solving Newton’s equations of motion combined with baryonic physics recipes to simulate galaxy formation and evolution. Comparing the simulation results to observations allows us to constrain the baryonic physics at play in galaxies, by investigating the effects of the baryonic recipes used. Semi-analytic models (SAMs) populate pre-created dark matter haloes using heuristic results and formulae, without modelling individual resolved gas regions - i.e. they take a dark matter halo with a given mass, structure and merger tree, and populate it with galaxies according to heuristic models. In addition to the dark matter halo information, SAMs require a collection of methods to solve baryonic physics problems such as cooling, star formation, stellar feedback, AGN feedback, galaxy mergers
Figure 1.6: Visual comparison of the Millennium simulation [Springel et al., 2005] (red, bottom and right) with observations (blue, left and top). *Image taken from Springel et al. (2006)*
and their effects (including dynamical friction, tidal stripping and ram pressure stripping), disc sizes and stellar population synthesis.

Many hydrodynamical models and SAMs, both based on the hierarchical paradigm, are unable to reproduce some key observational results. For example, the SAMs of Kauffmann et al. (1993), Baugh et al. (1996), Kauffmann & Charlot (1998) and Cole et al. (2000) all predict that the most massive early-type galaxies form their stars more recently than less massive galaxies. This is in contradiction to recent observations (such as Brinchmann & Ellis, 2000a; Ferreras & Silk, 2000; De Lucia et al., 2004; Heavens et al., 2004; Jimenez et al., 2005; Nelan et al., 2005; Thomas et al., 2005; Ziegler et al., 2005; Thomas et al., 2010) that find something known as ‘archaeological formation downsizing’; that more massive galaxies form their stars over short timescales at high redshift, and less massive galaxies have extended star formation at lower redshifts. This is excellently demonstrated in Figure 9 of Thomas et al. (2010), shown here in Fig. 1.7, where the x-axis shows lookback time (or redshift), the y-axis shows specific star formation rate and different coloured histograms represent galaxies of different masses. Clearly we can see here that whilst the most massive galaxies (log $M/M_* \sim 12.0$) have formed most of their stellar populations by $z \sim 4$ and form negligible stars beyond $z \sim 3$, less massive galaxies (log $M/M_* \sim 10.5$) form the bulk of their stars around $z \sim 0.7$ and are still forming stars out to $z \sim 0.4$.

In another example, the hydrodynamical models of Hopkins et al. (2006) and Faber et al. (2007) and the SAMs of Kauffmann et al. (1993), Kauffmann & Charlot (1998) and De Lucia et al. (2006) predict that the most massive galaxies assemble their mass at the lowest redshifts, which would seem a logical result of baryonic matter following dark matter as it forms hierarchically. However, observations contend with this; Ferreras et al. (2009) for example find that the number density of the most massive systems shows no evolution since $z \sim 1.2$. Pozzetti et al. (2010) showed that whilst the number density of the most massive galaxies early-types stays constant out to redshift $\sim 1$, the number density of less massive early-types grows with time. Pérez-González et al. (2008) find a similar result for all morphological types out to a redshift of $z \sim 4$, finding that the most massive galaxies formed first and very rapidly. Figs. 1.8 and 1.9 demonstrate these results, clearly showing that less massive early-types built up their stellar mass at later epochs than more massive early-types. This observation is known as ‘mass-assembly downsizing’. Pozzetti et al. (2010) tentatively put forward
the explanation that star formation quenching is a largely secular process with merging having an insignificant role, with more massive galaxies having more efficient quenching. As such, the late build up of stellar mass in less massive early-types is due to their high star formation rates in an era where the star formation in high-mass galaxies has already stopped.

Several modelling attempts have been made to reconcile observed downsizing with the hierarchical paradigm. For example, De Lucia et al. (2006) try to explain archaeological formation downsizing using a SAM by investigating the differences between assembly times and formation times. They find that with the addition of AGN feedback to their models they are able to produce older stellar populations in the most massive ellipticals, whilst assembling these objects last in accordance with the expected outcome of the hierarchical paradigm. Whilst this is still contradicted by mass assembly downsizing, it is an important step forward in our understanding of how to quench star formation in the most massive objects.

Monaco et al. (2006) use a SAM to try and solve the mass assembly downsizing problem. Their models suggest that the mass assembly downsizing can be somewhat explained in cluster galaxies by a significant fraction ($\sim 30\%$) of
Figure 1.8: The number density of early-type galaxies in bins of mass as a function of redshift. Image taken from Pozzetti et al. (2010).
Figure 1.9: The fraction of the local stellar mass density already assembled at a given redshift for several mass intervals. Image taken from Pérez-González et al. (2008).
stars being ‘lost’ from the galaxy to the diffuse stellar component of the cluster at each merger, where the diffuse stellar component consists of stars not closely bound to any single galaxy. This fraction of stars in the diffuse stellar component is consistent with observations (e.g. Durrell et al., 2002; Neill et al., 2005; Krick et al., 2006), and this process helps to flatten the increase in the number density of large mass galaxies with time, as required by mass assembly downsizing.

1.4 The Survey Landscape

With the expansion of theories to try and understand galaxy formation and evolution, it has become necessary to improve observations in order to add further constraints on the models. Spectroscopic galaxy surveys, i.e. observational programmes covering multiple galaxies, have come a long way in the past 35 years, filling in key areas of missing information. Importantly, they have allowed us for the first time to analyse galaxy populations, rather than just individual objects, and to take a statistical approach to our analysis. With large enough samples we are able to investigate the redshift evolution of galaxy properties and to research the impact of galaxy environments, allowing us to zero in on the main drivers of galaxy evolution with unprecedented insight.

In 1982 the Harvard-Smithsonian Center for Astrophysics (CfA) completed a series of observations that obtained the spectra of \( \sim 2400 \) galaxies with a limiting magnitude of \( m_B < 14.5 \) (Tonry & Davis, 1979), with the objective of measuring their redshifts to map the structure of the Universe and for use as a cosmological probe (e.g. for the measurement of the cosmological density parameter). It was the results from this survey that first showed the distribution of galaxies to be structured and not random, as shown in Fig. 1.10 from de Lapparent et al. (1986).

Surveys continued to progress and improve throughout the 1980s and 1990s. Due to time constraints, surveys have had to make a trade-off between depth and area covered. Some chose to cover large areas of the sky with shallow observations, such as \( \sim 18000 \) galaxies in CfA2 (Geller & Huchra, 1989) and \( \sim 26000 \) in Las Campanas Redshift Survey (Shectman et al., 1996), whilst others go deep but over a smaller area of the sky, for example \( \sim 700 \) objects in the Canada-France Redshift Survey (Lilly et al., 1995). Whilst wide surveys are great for narrowing down cosmological parameters such as the initial conditions of the Universe (Davis & Huchra, 1982), deep surveys have the advantage of being able to track the
CHAPTER 1. INTRODUCTION

Figure 1.10: The positions in RA-DEC-z space for $\sim 1100$ galaxies taken from the CfA redshift survey, showing the presence of structures of galaxies in the Universe for the first time. Image taken from de Lapparent et al. (1986).
progenitors of low-redshift objects further back in time, for example tracking the evolution of the luminosity function with redshift.

In the last 15 years the scale of these surveys has increased by 10-50x due to the creating of the multi-object spectrograph, that allows multiple spectra to be observed in each exposure. The groundbreaking Sloan Digital Sky Survey (SDSS; York & SDSS Collaboration 2000) and Two-Degree Field Galaxy Redshift Survey (2dFGRS; Lewis et al. 2002) obtained spectra for more than 800000 and 220000 galaxies over approximately 8000 and 1500 square degrees respectively. By 2009 the SDSS had produced its 7th data release, containing the spectra of more than 930000 galaxies over 10000 square degrees (Abazajian et al. 2009), and the 2dF team had progressed to the Six-Degree Field survey containing \( \sim 135000 \) spectra over 17000 square degrees at much greater depth than the SDSS (Jones et al. 2009). The scale of these surveys enabled a plethora of avenues of research, for example into AGN feedback (e.g. Schawinski et al. 2007), stellar mass distributions (e.g. Li & White 2009) and how galaxies populate dark matter haloes (e.g. Guo et al. 2010).

Recent surveys have continued this trend of impressive improvement. The SDSS has nearly completed its third phase, SDSS-III (Eisenstein et al. 2011), and produced its 10th data release (Ahn et al. 2013). This release contained \( \sim 928000 \) optical galaxy spectra over 6370 square degrees as part of its Baryon Oscillation Spectroscopic Survey (BOSS) component, and \( \sim 178000 \) optical stellar spectra as part of its the Apache Point Observatory Galaxy Evolution Experiment (APOGEE). The upgraded 2dF instrument AAOmega mounted on the Anglo-Australian Telescope (AAT) has been used to obtain \( \sim 287000 \) galaxy spectra for the Galaxy And Mass Assembly (GAMA) survey, which also incorporates \( \sim 31000 \) SDSS spectra into its sample, over a collection of fields covering around 388 square degrees.

Both of these surveys allow us to analyse different aspects of galaxy formation and evolution. Covering an approximately uniform mass distribution over a redshift range out to \( z \sim 0.7 \) (Dawson et al. 2013), BOSS allows us to research the redshift evolution of galaxy parameters in the low-redshift Universe to an unprecedented level. This will allow us for example to investigate the star formation histories of massive galaxies over this time frame, searching for a difference in stellar history between galaxies of different masses. GAMA on the other hand
has a very high level of completeness down to relatively low magnitudes, meaning that a large fraction of low-mass galaxies will be observed, enabling us to investigate the impact of environment \cite{Driver2011}.

1.5 Aims

I construct a data analysis pipeline to extract dynamical, emission line and absorption line properties from galaxy spectra, and then extract stellar population parameters from the absorption information. This process is detailed in Chapter 2. I find that our pipeline is very good at recreating external results when run over SDSS DR7.

I then run this pipeline over spectra from the GAMA survey, and use the information we gain here to analyse the environmental dependence of emission lines (Chapter 3). We do this with the objective of constraining the processes that quench star formation in galaxies, and investigating the relative importance of mass, local environment and global environment on quenching efficiencies. This can be used to constrain galaxy evolution models, as each model requires specific quenching modes that respond differently to mass and environment; for example, AGN feedback is expected to be largely environment-independent, whilst the effects of satellite infall should not be. We find that mass is by far the most significant driver of star formation quenching in galaxies, with local environment playing a minor role predominantly at high local densities, and global environment having no measurable impact.

I analyse spectra from SDSS-III/BOSS in Chapter 4, looking at absorption features with the aim of constraining stellar ages and star formation timescales as functions of dynamical mass (using stellar velocity dispersion as a proxy) and redshift. This is useful to constrain the archaeological formation downsizing phenomenon, as tracking the ages of stellar populations as a function of redshift will allow us to determine when the last period of star formation occurred, and tracking element ratios will inform us of the timescales involved. We find that our results are consistent with the idea that more massive galaxies have short bursts of star formation at higher redshifts, whilst less massive galaxies have extended star formation at much lower redshifts.
Chapter 2

Analysis Tools

Once a galaxy spectrum has been observed, calibrated, cleaned and sky-subtracted, analysis can be performed on it. Comparing the spectrum to models it is possible to calculate the physical properties of a galaxy, such as its stellar mass, star formation history, metallicity, stellar velocity, stellar velocity dispersion, gas velocities, gas velocity dispersions, chemical abundance ratios, emission line statistics and dust content.

I created a pipeline to extract the kinematics, stellar ages and metallicities, chemical abundance ratios and emission line statistics from input spectra from a variety of surveys. A complete list of the output of my pipeline can be found in Appendix A. Details of the surveys used and the data produced by them are given in Sections 3 and 4. In this chapter we describe the pipeline that I created to analyse galaxy spectra, including the individual components that it consists of and the stellar population models we used. We discuss how I adapted and improved the programs used that were written by other people, and how I calibrated our pipeline. We also introduce the Baldwin et al. (1981) diagnostic diagram (BPT) and the Lick index system of absorption indices. Finally I test the results of our pipeline against those of a known and widely accepted catalogue.

Our pipeline was written in a combination of IDL and FORTRAN. I adopted GANDALF (Sarzi et al., 2006) and pPXF (Cappellari & Emsellem, 2004) for spectral fitting, modifying GANDALF for our purposes as described below. Our Lick index code was written by C. Maraston and slightly altered by D. Thomas and myself, and stellar population parameter code by J. Johansson. A flow chart summarising our pipeline can be seen in Fig. 2.1.
Figure 2.1: The workflow of our pipeline.
The vast majority of work presented in this document was produced on SEP-net’s SCIAMA supercomputer at Portsmouth, where our pipeline is implemented. In order to parallelise the pipeline and enable multiple spectra to be analysed at once, a Perl script was written by E. Edmondson and modified by myself to send blocks of spectra to individual cores, and care was taken to ensure that no files needed by one block were overwritten by another. My parallelisation ensured that up to 20 blocks could be run at any given time, cutting the overall time required to analyse the spectra of a survey by a factor of 20. It took on average 2 minutes to run a single spectrum and compute emission statistics, kinematics, Lick index values and stellar population parameters, without errors on the latter two. Since errors on the absorption indices and stellar population parameters were calculated via Monte Carlo simulations, it could take up to 2 hours per spectrum to calculate these. As such, they were only computed for objects we were definitely interested in the absorption statistics for (see Section 4.2). This means that, for a survey of $\sim 200,000$ objects without errors on absorption indices or stellar population parameters being calculated, a single run of our pipeline would take approximately 2 weeks to complete.

2.1 Pipeline overview

The first task of our pipeline was to reduce the input spectra from all the surveys we were involved in into one common format. This involved obtaining each set of galaxy spectra (the primary spectrum and an error spectrum), turning all error spectra into variances, correcting for Milky Way reddening, log-rebinning all spectra to the same wavelength resolution using the same method, removing all invalid pixels, extracting all required metadata for each object and finally outputting our quantities to an identical format. This homogenising method was used as it resulted in the simple addition of a small chunk of code for each survey in one early location, rather than risking errors propagating through the analysis code from improperly prepared data. This allowed for the quick and easy inclusion of any survey data into our pipeline. These processes are all explained in detail here.

Milky Way foreground dust reddening is corrected for using the Schlegel et al. (1998) Milky Way E($B - V$) dust maps, inserting these values into the O’Donnell
extinction curve and applying this to both the flux and the variance spectra.

The log-binning method used was one of linear interpolation, not flux conservation. Linear interpolation is the method of (theoretically) drawing a straight line between the data points either side of the wavelength desired, where the x-axis is wavelength and y-axis is flux density, finding the location of the desired wavelength on this line and reading the flux value at this location off the y-axis. This was chosen as the readme of the analysis tool we used explicitly stated that flux interpolated binning methods would not work properly with the tools. The invalid pixels removed were those with value ‘NaN’, ‘-NaN’, ‘Inf’, ‘-Inf’ or ‘-9999’, which were found occasionally in early observations for some of the surveys used but were largely absent from the final data releases. These were replaced with linearly interpolated values. The metadata required were the right ascension (RA) and declination (DEC) of each object, i.e. its location in the sky, the redshift of the object and the ‘goodness’ of the object - whether or not the spectrum was deemed good enough for science. The quality of the spectrum was measured in a different way for each survey, and is covered in more detail in the survey sections.

Once this had been done, the newly formatted spectra were sent to GANDALF (Gas AND Absorption Line Fitting) (Sarzi et al., 2006), a fitting code covered in more detail below. In short, GANDALF utilises least-squares chi-squared minimisation methods to first extract stellar kinematics, then fit theoretical stellar spectra (stellar population templates) to the observed stellar continuum simultaneously with Gaussian models to the observed galaxy emission lines. This allows for the extraction of emission line fluxes, continua and equivalent widths (EWs, where the EW is the ratio of a line’s flux to its median continuum flux density), stellar velocity dispersions ($\sigma$), dust reddening and stellar continuum properties. Appendix B contains a list of the emission lines analysed for each survey, and Section 2.3 introduces one of the diagnostic tools for emission lines we used most commonly, the Baldwin, Phillips & Terlevich (Baldwin et al., 1981) diagnostic diagram.

Absorption index values were then measured for each of the 25 Lick indices listed in Appendix C. These values are essentially the equivalent widths (or magnitudes for some lines) of absorption features defined in the Lick system of absorption indices (Worthey et al., 1994; Trager et al., 1998), i.e. they are the measure of an amount of flux absent from the continuum. This was done using a
code primarily written by C. Maraston and modified by D. Thomas, J. Johansson and myself. The primary modification I made was to change integration methods from Simpson’s rule to a simple trapezoid method; tests were made to ensure that the difference in result was less than the measurement error. This change was necessary to enable calculations on noisy spectra, where calculations with Simpson’s rule would often fail to converge.

The absorption indices were then $\chi^2$-fit to the chemical abundance models of Thomas et al. (2011), using code developed by D. Thomas and J. Johansson and introduced in Johansson et al. (2012), with an equal weighting given to each line. This procedure fits for age, metallicity, $[\alpha/\text{Fe}]$ and a number of other element ratios, as detailed in Section 2.4.1. The specific outputs of this that are used by this work are luminosity-weighted age, metallicity and $[\alpha/\text{Fe}]$.

Finally, all of these values were concatenated into a structure, which was further placed into an array of structures with each structure containing the output for one object. This array was written out to a catalogue, which was then used for scientific analysis. The total list and a brief description of all products output into our catalogue(s) can be found in Appendix A.

\subsection*{2.2 GANDALF: Gas AND Absorption Line Fitting}

I use the publicly available codes pPXF (penalised PiXel Fitting) (Cappellari \& Emsellem, 2004) and GANDALF (Sarzi et al., 2006) to calculate stellar kinematics and derive emission line properties. pPXF extracts stellar kinematics from the data in pixel space, whilst GANDALF fits stellar population templates (see Section 2.2.2) and Gaussian emission line templates to the observed spectra simultaneously using a least-squares minimisation method, allowing for the separation of stellar continuum and absorption from ionised gas emission.

In more detail, the vanilla GANDALF version we use is v1.8, and works by first computing the stellar velocity dispersion using pPXF, a penalised pixel-fitting method (see Section 2.2.1). It then performs bound Levenberg-Marquardt least-squares minimisation (a $\chi^2$ minimisation technique) over all input stellar templates having changed their resolution to account for the results of pPXF, any reddening parameters selected, a multiplicative Legendre polynomial used to
correct the template continuum during the fit, and Gaussians placed at specified emission line locations. The stellar continuum and emission line Gaussians are therefore fit simultaneously, and a model spectrum created purely through combining the stellar templates and parameters output by GANDALF, ignoring the Gaussian emission lines, should somewhat accurately recreate any absorption features present in emission line regions.

The Levenberg-Marquardt algorithm is a least-squares minimisation method designed to solve the least squares curve fitting problem: given a set of \( m \) empirical data points in \( x - \) and \( y - \) space with values \((x_i, y_i)\), optimise the parameters \( \beta \) of the model curve \( y = f(x, \beta) \) such that the differences between the model and each data point are minimised; i.e. minimising \( S(\beta) \) in Eq. 2.1 where \( y_i \) and \( x_i \) are the \( y - \) and \( x - \) values of data point \( i \), and \( \beta \) are the model’s parameters. The Levenberg-Marquardt algorithm does this iteratively by replacing the parameter \( \beta \) by \( \beta + \delta \) such that \( f(x_i, \beta + \delta) \approx f(x_i, \beta) + J_i \delta \) where \( J_i = \partial f(x, \beta)/\partial \beta \) (i.e. the gradient of \( f(x, \beta) \) with respect to \( \beta \)). The difference between the LMA and other least-squares algorithms is the usage of a damping factor that scales in the direction of each component by the gradient of its curvature, which pushes the minimisation method between two different approaches (the Gauss-Newton algorithm and the method of gradient descent) and allows for faster convergence on the final values.

\[
S(\beta) = \sum_{i=1}^{m} [y_i - f(x_i, \beta)]^2 \quad (2.1)
\]

Dust reddening is accounted for by adopting a Calzetti (2001) obscuration law. Whilst we acknowledge that this law was calibrated for local starbursts and does not necessarily hold for higher redshifts or in passive galaxies, we opted to keep it (it was used in vanilla GANDALF) as it is popularly used in the literature (e.g. Sheen et al., 2009; Ferreras et al., 2012; Thomas et al., 2013) and no ‘true’ dust law is really known. Since we do not use the shape of the spectrum for any of our final products, and every time we use emission lines we choose to compare ratios of those that are very close together (and therefore minimally affected by reddening), this decision has no impact on the results provided in this work.

GANDALF has the ability to perform a two-part reddening correction, one diffuse correction applied to the entire spectrum to account for reddening between the target galaxy and our observations, and one nebular component applied only
to emission lines to account for dust in the emitting regions themselves. However, I found that the internal reddening component was hugely overestimated when the survey data had a low S/N, and so decided not to include this in our analysis. My work analysing the reddening parameters of GANDALF is listed in Section 2.2.4.

Using this method I am able to extract stellar kinematics, gas kinematics, emission line fluxes and emission line equivalent widths (EWs) from the data. Emission line EWs are calculated as the ratio between line flux and continuum flux density, where the continuum flux density adopted is the median of the flux densities in two continuum windows (one to the blue and one to the red of the emission feature) at a distance of ±200 km/s from the centre of the line and with widths of 200 km/s.

Whilst it would generally be preferable for the gas kinematics of weaker lines to be tied to stronger lines such as Hα at 6563 Å and [NII] at 6583 Å (e.g. Brinchmann et al., 2004; Sarzi et al., 2006; Schawinski et al., 2007), this is not practical for surveys investigating redshifts where the maximum observed wavelength falls below $\lambda_{\text{max}} = \lambda_{\text{line}} \ast (1+z)$ where $\lambda_{\text{line}}$ is the wavelength of the strong line you wish to tie to. For example, the BOSS survey has a median redshift of 0.506 meaning that the maximum observed wavelength is required to be $\lambda_{\text{max}} > 9884$ Å in order to detect Hα in the majority of objects. With a standard deviation on z of 0.151 and maximum observed wavelength of $\lambda_{\text{max}} \sim 10,000$, Hα can clearly not be detected for many objects, making it impossible to tie the kinematics of weaker lines to that of Hα for all objects. We therefore leave emission line velocities and velocity dispersions as free parameters for all lines except for [OIII]λ4959 and [NII]λ6548, which are tied to their nearby, stronger neighbours [OIII]λ5007 and [NII]λ6583 respectively. We follow this approach independently of redshift and survey to guarantee homogeneity within the full samples investigated.

It should be noted that I also corrected the way GANDALF calculated EWs, as the vanilla version corrects emission line flux for redshift effects but not continuum flux density, which would leave EWs out by a factor of $1/(1+z)$. We therefore multiplied the continuum by a factor of $(1+z)$ in order to k-correct it. The wavelengths we use for these lines are those as measured in air, as opposed to the vacuum wavelengths used by SDSS (York & SDSS Collaboration, 2000).
2.2.1 pPXF

The ‘Penalized Pixel-Fitting’ procedure, or pPXF, was developed by Cappellari & Emsellem (2004) to extract stellar kinematics, both stellar velocity and the line-of-sight stellar velocity dispersion (LOSVD), by fitting a theoretical template to an observed spectrum in pixel space using a (parametric) penalised pixel-fitting method. This fit being performed in pixel space rather than using a Fourier-based technique (e.g. Bender, 1990) is important as it allows for easy masking of emission features and bad pixels, and also allows templates with high spectral resolution spectra to be carefully matched to the observed spectrum.

Like for GANDALF, pPXF utilised a $\chi^2$ minimisation technique with the objective of minimising the residuals between the observed spectrum and the fitted template, weighted by the measurement error. The (weighted) stellar population models input into the template used are convolved with a broadening function controlled by the LOSVD, which is allowed to vary between limits in order to achieve a best fit. The LOSDV is expanded as a Gauss-Hermite series, rather than as a pure Gaussian, as explained in van der Marel & Franx (1993) and Gerhard (1993). A further addition is made to the convolved models before being fitted in the form of weighted Legendre polynomials, to account for low frequency differences in shape between the observation and templates. In order for the methods detailed here to work, the weights for both the models and the polynomials are first optimised for a given (initial guess) LOSVD model, allowing the final fit to be driven by variations of the LOSVD model.

Due to the fact that at low signal-to-noise (S/N) the LOSDV is well approximated by a Gaussian, pPXF uses a method whereby at high S/N the actual LOSVD profile is well reproduced by the fit whereas at low S/N the LOSVD tends to a Gaussian shape. In order to do this, the $\chi^2$ is penalised to bias towards a Gaussian shape such that $\chi_p^2 = \chi^2 + \alpha P$, where $P$ is a penalty function given by the integral of the square deviations of the observed line profile from its best fitting Gaussian and $\alpha$ is a term to weight the value of the penalty by the $\chi^2$ value itself. In practise the way $P$ is formulated means that a deviation of the LOSVD from a Gaussian shape will be accepted as an improvement to the fit only if it decreases the variance by an amount related to the deviation - i.e. the decrease has to be greater than an amount determined by the confidence level at which a Gaussian solution can be excluded in order for the fit to be classed as an improvement.
CHAPTER 2. ANALYSIS TOOLS

This penalisation essentially results in only preserving statistically significant non-Gaussian features, otherwise the LOSVD solution is reduced to a Gaussian. The production of a perfectly Gaussian absorption feature would require a continuous, Gaussian distribution of individual stellar velocities along the line of sight centred on zero; this is clearly not always the case. It is therefore important to be able to make these deviations from Gaussian form in order to robustly measure interesting parameters such as stellar velocity dispersion. The parameters of the final LOSVD model give the galaxy’s stellar recessional velocity (as a deviation from the input redshift), stellar velocity dispersion and the shape of the LOSVD via the Hermite polynomials.

2.2.2 Stellar population templates

Theoretical stellar population templates are required during the fitting process in order to model the properties of the real galaxy from its observed light. The templates that we used are based on the models of [Maraston & Strömbäck (2011)], which are those of single starbursts.

The core concept behind the use of these templates is that a galaxy’s stellar population can be modelled by a series of discrete instantaneous starbursts, with negligible star formation occurring between these periods (i.e. modelled as Dirac delta functions). Whilst this is obviously not physically the case, it has been shown (e.g. Tojeiro et al., 2007) that this method can reproduce observed galaxy spectra accurately, and with far fewer template permutations - and as such in a far shorter computing timescale - than when fitting with composite, continuous star formation templates (which is clearly just a convolution of single bursts anyway).

Each template represents a stellar population that was formed a given period of time in the past, with a given initial mass function (IMF) and metallicity, and with individual stellar spectra modelled on a set stellar library. The IMF determines the ratio of stars of different masses that form, and the metallicity impacts on absorption features. For our purposes we selected models with a Salpeter IMF [Salpeter, 1955] and at solar metallicity. The stellar library provides the behaviour of individual stellar spectra as functions of temperature, gravity and chemical abundances, and we opted for models based on the MILES library [Sánchez-Blázquez et al., 2006]. The MILES library is an empirical library of 985 stars of varying evolutionary stages and metallicities, observed in the wavelength range $3500 – 7428 \, \text{Å}$ at a resolution of $\Delta \lambda = 2.54 \, \text{Å}$ [Beifiori et al., 2011], although
recently measurements by Falcón-Barroso et al. (2011) instead suggest a value of 2.51 Å). The extension of models beyond this range was done using the theoretical library UVBLUE (Rodríguez-Merino et al., 2005), as described in Maraston & Strömbäck (2011).

Once the initial starburst has occurred with its predetermined parameters, all stars are evolved using the models set out in Maraston & Strömbäck (2011), and ‘snapshots’ of the spectrum of the entire combined population are taken at set times after the initial formation, known as ages. These snapshots make up the single stellar population (SSP) templates that we use. We used the ages 6.5, 25, 100, 200, 400, 600 and 800 Myr, and 1, 1.5, 2, 3, 5, 6, 7, 8, 9, 10 and 11 Gyr, which we chose as they cover a wide range of possible population spectra whilst limiting too much degeneracy. Fig. 2.2 shows these templates at their base (model) resolution.

I opted not to limit the age of populations used in our fits by the age of the Universe; i.e. it was possible for our analysis to include a model of a 11 Gyr-old galaxy at redshift 0.5, at which time the Universe was around 8.65 Gyr old, therefore making our model galaxy 19.65 Gyr old at redshift zero. I did this as the surveys we used had signal-to-noise ratios too low (see details in Section 3 for GAMA, Section 4 for SDSS-III/BOSS, but average SNR ∼ 2 – 5) to get truly accurate stellar population parameters from, and all we could hope for was a best fit, approximate model continuum from which to extract emission and absorption statistics. Further, we limited the models to those at solar metallicity, since we were not interested in extracting ages or metallicities from the templates, and using a single metallicity and a large range of ages allowed us to take advantage of the age-metallicity degeneracy to reduce the overall number of templates needed to get an acceptable fit.

I also downgraded the spectral resolution of these templates to match the resolution of each individual survey. This is necessarily in order to calculate accurate velocity dispersions, as lowering spectral resolution has exactly the same observable effect as increasing an object’s velocity dispersion, and as such incorrect template resolutions will alter the measured velocity dispersions by the amount the template is off by. This was done by convolving the template spectrum with a smoothing resolution, determined as a function of the original template resolution and that of the observed spectra. Eq. 2.2 shows how the velocity dispersion of the smoothing function was calculated, where $\sigma$ is velocity dispersion, $f$ indicates the
CHAPTER 2. ANALYSIS TOOLS

Figure 2.2: The theoretical single stellar population models we used in our fits, at their inherent resolution (that of the MILES library; $\Delta \lambda = 2.54$ Å).
final template resolution, \( i \) the input (MILES) resolution, and \( s \) the smoothing resolution.

\[
\sigma_s^2 = \sigma_f^2 - \sigma_i^2
\]  

(2.2)

The smoothing resolution is calculated as in Eq. 2.3 where \( \lambda \) is wavelength, \( \Delta \lambda \) is the wavelength step and \( c \) is the speed of light in a vacuum. The templates are then convolved with this smoothing factor to downgrade them to the final spectral resolution required.

\[
\sigma = \frac{\Delta \lambda}{\lambda} \cdot c \cdot \frac{1}{2.35} \quad \text{km/s}
\]  

(2.3)

Fig. 2.3 shows how the downgrading procedure and subsequent log-rebinning alters the template spectrum in the wavelength region \( 3910 - 3990 \) Å for all templates listed above, where the black line is the original template at MILES resolution (\( \Delta \lambda \sim 2.54 \) Å), the red line the template downgraded to BOSS resolution (\( \Delta \lambda \sim 2.17 \) Å, see Chapter 4) and the blue line the template downgraded to GAMA resolution (\( \Delta \lambda \sim 4.5 \) Å, see Chapter 3). Both the red and blue lines have been log-rebinned using a linear interpolation method to a regular grid of \( d\lambda = 0.0001 \) and then had their wavelength indices raised to the power of 10, whilst the black line has just been kept linearly binned.

It is clear simply from the resolutions that whilst the BOSS resolution is actually better than the MILES resolution, the GAMA resolution is considerably worse. Looking at how this transfers onto the templates, we see that the GAMA templates have noticeably fewer features and that the signal is smeared out over a broader range, as expected. The absorption features in the GAMA templates are unsurprisingly shallower and broader (albeit slightly) due to the signal being smeared out. In comparison, the BOSS resolution is higher than that of the base templates but almost undetectable so, and it is likely that any real offsets between the MILES and BOSS templates are likely to be due to minute changes caused by the log-rebinning method. It should be noted that although it is not shown here, the overall broad shape of the template spectra are not changed by any part of this process. It is also important to note that ‘downgrading’ the spectral resolution from 2.54 Å to 2.17 Å does not introduce new information, it essentially just resamples the spectrum to a new grid.
Figure 2.3: A comparison of how altering SSP template resolution changes the shape of the template features. Templates at the MILES resolution (the default resolution of the model templates, $\Delta \lambda = 2.54 \, \text{Å}$) are shown in black, at the BOSS resolution ($\Delta \lambda \sim 2.17 \, \text{Å}$) in red, and at the GAMA resolution ($\Delta \lambda \sim 4.5 \, \text{Å}$) in blue. The ages given for each image are how long the stellar populations were evolved before the spectrum was taken.
### 2.2.3 Kinematics calibration

The region within which stellar kinematics are calculated in vanilla GANDALF was based on calibrations performed for the spectra the tool was originally designed to run over, i.e. those of the SAURON team. Since they were largely arbitrary we opted to explore the parameter space and calibrate it for the spectra we were using - if our spectra are more likely to be affected by fringing, for example, then it makes little sense to include the edges of the spectra in calculations.

Our benchmark values (those that we chose to calibrate to) were those given in the MPA-JHU (Max-Planck institut für Astrophysik/Johns Hopkins University) catalogue of spectroscopic data products based on the SDSS Data Release 7 (DR7 [Abazajian et al., 2009]). We chose this as the MPA-JHU catalogue is widely known and used for science and is generally trusted as being reliable, accurate and bug-free. Although we note that the MPA-JHU catalogue uses different stellar population models (those of [Bruzual & Charlot, 2003]) based on the STELIB library ([Le Borgne et al., 2003]), this should not impact significantly on the velocity dispersions calculated if the methodology is correct.

By default stellar kinematics were calculated across the wavelength range covered by the entire template spectrum. We decided to investigate the kinematics calculated by considering different regions by altering the wavelength range in the observed frame, whilst always keeping the edges of our range bound by the edges of the rest frame spectrum. We also ensured that the range considered never extended further than the template wavelength boundaries.

I investigated the parameter space 4000 – 7500 Å, varying the lower and upper limits within this boundary whilst maintaining a minimum gap of 1500 Å. This overall lower limit was constrained by our attempts to avoid low-λ fringing in the lowest redshift objects, and the upper limit set by the maximum wavelength of our stellar population templates (7428 Å). I then compared the output stellar kinematics values from GANDALF to those in the MPA-JHU catalogue for a sample of 8,000 randomly selected objects. Table 2.1 shows the number of objects with a velocity dispersion error of $< 30\%$, mean and median offsets and standard deviations of the log MPA-JHU stellar velocity dispersions subtracted from the log GANDALF values for the ranges of parameter space investigated.

It can be seen from this table that the optimal values for the ranges in which kinematics are calculated is either 4000 – 6000 Å, as it has a low mean and standard deviation, or 4000 – 6500 Å as the slightly higher standard deviation...
Figure 2.4: The distribution of the differences between the velocity dispersions calculated by GANDALF and those in the MPA-JHU catalogue when different wavelength ranges are investigated. The left panel shows this for the two most optimal ranges considered, 4000–6500 Å (black, with a mean of 0.008±0.095) and 4000–6000 Å (red, with a mean of 0.004±0.065) where the dashed lines represent their respective medians. The right panel shows this for the final range chosen (4000–6500 Å, black) and three other randomly selected ranges: 5000–6500 Å (red, with a mean of 0.014±0.125), 4500–7000 Å (green, with a mean of 0.042±0.121) and 5500–7500 Å (blue, with a mean of 0.018±0.653).

is made up for by an increase in the number of acceptable measurements and a better consistency between the mean and the median differences. The left panel of Fig. 2.4 shows the distribution of the differences for these two ranges where the range 4000–6500 Å is shown in black and 4000–6000 Å in red, and the dashed lines represent their respective medians. Given the similarity between the two distributions, I picked the wavelength range 4000–6500 Å for the consistency between the mean and median differences. The right panel of Fig 2.4 shows the distribution of the differences for the ranges 4000–6500 Å (black), 5000–6500 Å (red), 4500–7000 Å (green) and 5500–7500 Å (blue). It is clear from this comparison that even slight changes to the wavelength range used in kinematics calculations can have significant effects on the results.

We proceeded to use the range 4000–6500 Å when performing analysis on all surveys covering this wavelength range.

2.2.4 Reddening analysis

When we were first running this pipeline on SDSS and GAMA spectra, I noticed that whilst the majority of measurements of the emission line fluxes of objects
<table>
<thead>
<tr>
<th>range (Å)</th>
<th>count</th>
<th>mean (dex)</th>
<th>median (dex)</th>
<th>sigma (dex)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4000 − 5500</td>
<td>6700</td>
<td>0.011</td>
<td>0.015</td>
<td>0.064</td>
</tr>
<tr>
<td>4000 − 6000</td>
<td>6741</td>
<td>0.004</td>
<td>0.010</td>
<td>0.065</td>
</tr>
<tr>
<td>4000 − 6500</td>
<td>6833</td>
<td>0.008</td>
<td>0.008</td>
<td>0.095</td>
</tr>
<tr>
<td>4000 − 7000</td>
<td>6844</td>
<td>0.029</td>
<td>0.011</td>
<td>0.121</td>
</tr>
<tr>
<td>4000 − 7500</td>
<td>5322</td>
<td>0.127</td>
<td>0.114</td>
<td>0.393</td>
</tr>
<tr>
<td>4500 − 6000</td>
<td>6484</td>
<td>0.009</td>
<td>0.016</td>
<td>0.065</td>
</tr>
<tr>
<td>4500 − 6500</td>
<td>6640</td>
<td>0.015</td>
<td>0.009</td>
<td>0.095</td>
</tr>
<tr>
<td>4500 − 7000</td>
<td>6746</td>
<td>0.042</td>
<td>0.018</td>
<td>0.121</td>
</tr>
<tr>
<td>4500 − 7500</td>
<td>4526</td>
<td>−0.003</td>
<td>−0.008</td>
<td>0.470</td>
</tr>
<tr>
<td>5000 − 6500</td>
<td>6511</td>
<td>0.014</td>
<td>0.007</td>
<td>0.125</td>
</tr>
<tr>
<td>5000 − 7000</td>
<td>6624</td>
<td>0.031</td>
<td>0.011</td>
<td>0.148</td>
</tr>
<tr>
<td>5000 − 7500</td>
<td>4486</td>
<td>0.071</td>
<td>0.135</td>
<td>0.515</td>
</tr>
<tr>
<td>5500 − 7000</td>
<td>6062</td>
<td>0.013</td>
<td>−0.008</td>
<td>0.212</td>
</tr>
<tr>
<td>5500 − 7500</td>
<td>3222</td>
<td>0.018</td>
<td>0.259</td>
<td>0.653</td>
</tr>
<tr>
<td>6000 − 7500</td>
<td>4307</td>
<td>0.052</td>
<td>0.163</td>
<td>0.577</td>
</tr>
</tbody>
</table>

Table 2.1: The kinematics ranges investigated for SDSS DR7 with the number of galaxies that have a stellar velocity dispersion error of < 30% (count), and the means, medians and standard deviations of the differences between the GANDALF calculations and those in the MPA-JHU catalogue such that the difference is equal to \( \log( \sigma \text{ (GANDALF) } ) - \log( \sigma \text{ (MPA-JHU) } ) \).
Figure 2.5: The ratio of vanilla (2-dimensional reddening) GANDALF and MPA-JHU measurements against the MPA-JHU measurement of \( \log H\alpha \) EW (left) and \( \log H\beta \) EW (right) for 8,000 randomly selected SDSS DR7 galaxies with AoN > 4 for each line. The blue line represents the 1:1 relation, and the green line a 0.5 dex offset from this relation. Red points are those for which the GANDALF flux measurements were greater than 0.5 dex offset from the MPA-JHU measurement. The offset value of 0.5 dex was visually chosen as an illustrative guide to see highly offset objects, and has no statistical significance.

matched up to previous studies, there was often a small sub-population of fluxes that were grossly overestimated by the GANDALF code. Fig. 2.5 shows a plot of \( H\alpha \) and \( H\beta \) EWs comparing our measurements using default GANDALF settings (including 2d reddening) against those of the MPA-JHU catalogue, for a sample of 8,000 randomly selected galaxies from SDSS DR7 where the AoN of each line was greater than 4.

The blue line represents the 1:1 relation, and the green lines mark a 0.5 dex deviation from the 1:1 relation which we consider a suitable dividing line between ‘accurate’ measurements and those that are clearly inconsistent between the catalogues. Red points are those that make up the ‘offset population’, i.e. those for which the GANDALF measurement is greater than 0.5 dex different to the MPA-JHU measurement. This offset criteria was chosen visually to illustrate highly offset objects, and has no statistical significance. I chose a minimum AoN value of 4, rather than the 2 we typically use for science, in order to remove the expected scatter caused by noisy spectra from my investigation and to only concentrate on scatter caused by improper fits. I found that the offset fraction was 37.0% for \( H\alpha \) and 47.7% for \( H\beta \).

I investigated the cause of this offset, and found that it correlated with the reddening correction applied in the fit. By default GANDALF fits two reddening
components: one to the shape of the entire fit, to take diffuse reddening into account, and one only to emission lines to account for nebular reddening. Fig. 2.6 shows a histogram of the E(B-V) values calculated for objects that matched up well with the MPA-JHU catalogue (black) and those that made up the offset (red) when looking at Hα flux (top) and Hβ flux (bottom), with the left histograms displaying the diffuse component and the right histograms the nebular.

From these histograms it is clear that the distribution of the nebular reddening components for the ‘accurate’ and ‘offset’ objects are markedly different. Whilst the ‘accurate’ measurements have nebular reddening components of the order 0-0.5, offset measurements stretch out much higher and then have a second peak around E(B-V)~5.

To see if this was a cause rather than a side-effect, I removed the nebular reddening component. Fig. 2.7 shows the same as in Fig. 2.5 but where the
Figure 2.7: The ratio of GANDALF and MPA-JHU measurements against the MPA-JHU measurement of log Hα flux (left) and log Hβ flux (right) for 8,000 randomly selected SDSS DR7 galaxies with AoN > 4 for each line, where GANDALF measurements were made taking only diffuse reddening into account. The blue line represents the 1:1 relation, and the red line a 0.5 dex offset from this relation. The red points are those for which the original vanilla GANDALF flux measurements were greater than 0.5 dex offset from the MPA-JHU measurements, and the red triangles are objects that still class as offset with only diffuse reddening taken into account. Red points outside of the 0.5 dex boundary are those that had AoN > 4 when both reddening components were taken into account, but AoN < 4 when only diffuse reddening was used.

GANDALF fits were performed using only a diffuse reddening correction. The points in red are those objects that were declared offsets in Fig. 2.5, and the red triangles represent objects that are offset after the removal of the nebular component. From this figure we can see that the offset sub-population is removed when only one reddening component is considered.

This issue was caused by noisy data being fit best by very large Gaussians combined with unrealistic levels of dust attenuation, as the S/N was often not high enough to be able to differentiate between this and a smaller Gaussian. Although the second dimension of reddening is a highly desirable feature, as it correctly deals with the differences in dust between active regions and the galaxy as a whole, it is not absolutely necessary and only acts as a second-order effect.

The remaining offset objects were investigated by eye. Plots of the spectra and models can be found in Appendix D for Hα offset objects, and Appendix E for Hβ offset objects.

It was found that for the 15 objects with an Hα offset of > 0.5 dex, 5 have good fits in GANDALF, 2 have no visible emission, 6 have poor overall GANDALF fits, and
2 have poor fits to the emission line alone. In the cases where no emission line is visible, the measured EW is found to be low overall as well as lower for GANDALF than for MPA-JHU measurements, and are therefore likely to be accurate. It is therefore safe to surmise that 8 of the objects in my sample have allegedly good GANDALF measurements which are in reality unacceptable. Given that there were 3,909 objects with Hα AoN> 4, this gives us a highly acceptable offset rate for Hα measurements of 0.20%.

For the 34 objects with an Hβ offset of > 0.5 dex, however, there is a different story. Only 2 objects are observed to be featureless, and the remaining 32 have poor fits to the data, with 8 of them being poor overall fits and 24 poor fits to the line specifically. With there being 2,776 objects with Hβ measured with AoN > 4, this gives a slightly higher offset rate of 1.15% for Hβ measurements. However, it is important to note that 26 (76.5%) of the poor fits to the Hβ line specifically were at $z \sim 0.148 \pm 0.003$, at which redshift the Hβ emission line falls close to a sky line. I investigated this, and found that it was due to very high variance levels near the Hβ line location, often approaching the same value as that of the flux. When this was the case, GANDALF often obtained an optimal $\chi^2$ fit by fitting an emission line model to the not-fully-subtracted skyline, where a relatively wide Gaussian was allowed due to the width of the variance spike. When objects with $0.145 < z < 0.150$ were discarded, i.e. where emission lines will be inaccurate anyway due to sky contamination, an acceptable offset rate of 0.33% was measured.

When a limit of AoN > 2 was applied instead - the limit we generally use for science - we found that the offset rate for Hα was 2.5%, and for Hβ was 5.17%. When lines suffering from sky contamination were removed, the Hβ offset rate improved to 4.55%. We consider these to be acceptable giving the low S/N of the data and the relative weakness of the Hβ line.

Having established this, we used only the diffuse reddening component in all future analysis.

### 2.3 BPT diagnostic diagram

Emission line properties, once properly extracted, can be used to determine the energetic processes at work within a galaxy. This is most commonly done using the diagnostic diagram of Baldwin, Phillips & Terlevich (BPT, [Baldwin et al.](#)).
revised by Veilleux & Osterbrock (Veilleux & Osterbrock, 1987). This method uses the logarithms of the ratios of emission lines \([\text{OIII}]\lambda 5007/\text{H}\beta\) and \([\text{NII}]\lambda 6583/\text{H}\alpha\) to separate galaxies whose ionisation processes are dominated by star formation (SF), Seyfert activity, LINER-like activity and composite SF-AGN behaviour.

Whilst the processes behind energetic star forming regions are self-explanatory, Seyferts and LINERs are less self-evident. Seyfert galaxies are AGN with lower luminosities than quasars, typically accepted due to the work of Schmidt & Green (1983) as having a maximum bolometric magnitude of \(M_B > -21.5 + 5 \log h_0\) for the active nucleus in order to distinguish Seyferts from quasars (Peterson, 1997). They have quasar-like nuclei (i.e. very bright, compact nuclei), but the host galaxy is clearly detectable. They are defined by their strong, high-ionisation emission lines. LINERs, or Low-Ionisation Nuclear Emission-line Regions, are very common but somewhat poorly understood. They spectroscopically resemble Seyfert galaxies in many respects, but also have strong low-ionisation lines (e.g. \([\text{OI}]\lambda 6300\) and \([\text{NII}]\lambda\lambda 6548, 6583\)). The link between LINERs and AGN is a hotly debated topic, although a consensus is gradually forming that LINER emission is not caused by a point-like source and is therefore unlikely to be AGN-driven (e.g. Yan & Blanton, 2012; Singh et al., 2013).

Fig. 2.8 shows an example BPT diagnostic diagram using our results for the GAMA spectra (see Section 3 for details), only considering objects an amplitude-over-noise of greater than 3 for all four BPT diagnostic lines. We use the empirical separation between star forming galaxies and AGN from Kauffmann et al. (2003) (dashed line) and the theoretical extreme starburst line from Kewley et al. (2001) (solid curved line) to identify pure star forming and pure AGN emission. We adopt the common practise of assuming the area between these lines is populated by galaxies with composite star forming and AGN spectra. We further use the dividing line defined by Schawinski et al. (2007) (solid straight line) to distinguish between LINER and Seyfert emission based on SDSS galaxy classifications obtained through the \([\text{SII}]\)/\text{H}\alpha ratio.

The restriction of objects based on their line amplitude-over-noise (AoN) values is a necessary step in order to avoid giving emission line classifications to objects without real emission lines. Obviously in extreme cases the classifications would be completely meaningless, as they would be derived from noise; a minimum requirement of AoN = 2 seems to remove the majority of objects without
Figure 2.8: Example BPT diagram for the GAMA survey. We use the empirical separation between star forming galaxies and AGN from [Kaufmann et al. (2003)](#) (dashed line) and the theoretical extreme starburst line from [Kewley et al. (2001)](#) (solid curved line) to identify pure star forming and pure AGN emission. We adopt the common practise of assuming the area between these lines is populated by galaxies with composite star forming and AGN spectra. We further use the dividing line defined by [Schawinski et al. (2007)](#) (solid straight line) to distinguish between LINER and Seyfert emission based on SDSS galaxy classifications obtained through the [SII]/Hα ratio.
proper emission features (according to a visual inspection of objects with AoN values around 2). Here we have chosen a minimum of 3 in order to display only objects with confident measurements of emission lines for our example. In general, we consider objects with emission line AoN values below our chosen limits to be ‘passive’ in the emission line sense.

2.4 Absorption line measurements

In addition to emission lines, which characterise the gas content of a galaxy, absorption lines are also very useful for galaxy evolution studies as they characterise the stellar population. The Lick index system was defined by the Lick group in [Worthey et al. (1994)] for 21 absorption indices with a further 4 higher order Balmer indices added in [Worthey & Ottaviani (1997)], and the original 21 were refined in [Trager et al. (1998)] using better wavelength calibrated spectra. These 25 indices trace the most prominent absorption features in the optical spectrum. Unlike emission lines which are measured against a measured ‘actual’ continuum, the Lick indices are measured against a pseudo-continuum due to the significant broadening of absorption features by stellar velocity dispersions. To this end the pass-band containing the actual absorption feature is flanked by two pseudo-continuum pass-bands, where the pseudo continuum is defined as a line connecting the mean fluxes of each pass-band located at the central wavelengths of each band. Fig. 2.9 illustrates this for Mgb using a randomly selected high-S/N BOSS stack where the solid lines mark the boundaries of the feature pass-band, the vertical dashed lines the boundaries of the pseudo-continuum pass-bands, solid horizontal lines the mean flux values of the pseudo-continuum bands and the horizontal dashed line the pseudo-continuum itself. Appendix C lists the boundary values for each of the 25 indices.

Index strengths are measured in two ways: either as magnitudes for broad molecular features (CN$_1$,CN$_2$,Mg$_1$,Mg$_2$,TiO$_1$,TiO$_2$) or as EWs for narrow atomic features (all other lines). These are defined as

$$\text{Mag} = -2.5 \log_{10} \left( \int_{\lambda_1}^{\lambda_2} \left( \frac{1}{\lambda_2 - \lambda_1} \right) \frac{F_{I\lambda}}{F_{C\lambda}} \, d\lambda \right)$$

(2.4)

$$\text{EW} = \int_{\lambda_1}^{\lambda_2} \left( 1 - \frac{F_{I\lambda}}{F_{C\lambda}} \right) \, d\lambda$$

(2.5)
Figure 2.9: An example absorption feature showing the feature boundaries (solid lines), blue and red pseudo-continua boundaries (vertical dashed lines), mean flux levels in each wing (horizontal solid lines) and final pseudo-continuum (horizontal dashed line). This is the Mgb index for a randomly chosen BOSS stacked spectrum, with a redefined flux unit.
where $\lambda_1$ and $\lambda_2$ mark the boundaries of the feature pass-band, and $F_{I\lambda}$ and $F_{C\lambda}$ are the fluxes per unit wavelength of the feature pass-band and pseudo-continuum respectively.

Three outputs from GANDALF are required for the calculation of the Lick indices for each object: the emission-cleaned spectrum, the best-fit model spectrum and the galaxy velocity dispersion, as well as the redshift of the object and the resolutions of the observed spectrum and the model templates.

First, the observed and model spectra are downgraded to the resolution of the models we wish to compare the indices to; in our case, this was the Lick resolution as we want to compare to the models of (Thomas et al. 2011, henceforth referred to as TMJ. See also Johansson et al. (2010)). The model spectrum is then smoothed by the velocity dispersion of the observed galaxy. All smoothing and downgrading was done using a similar procedure to that described above for the stellar population synthesis models. The required parameters are then measured, and the final index values stored.

Errors on Lick index measurements are produced via Monte Carlo simulations. For each spectrum we create 1000 randomly simulated spectra with fluxes defined as

$$F_{MC} = F_{mod} + \text{RAND} \times \sigma_{res}$$

where $F_{MC}$ is the Monte Carlo-simulated flux, $F_{mod}$ is the flux of the best-fit spectrum, RAND is a set of random numbers (one for each spectral pixel) normalised to a Gaussian centred at 0, and $\sigma_{res}$ is the robust standard deviation of the best-fit residuals (i.e. $\sigma_{res} = \text{ROBUST\_SIGMA}(f_{mod} - f_{obs})$ where $f_{obs}$ is the observed flux). We then measure the Lick indices of each of these simulations, take the standard deviation of the results and treat this as our index error.

### 2.4.1 Stellar population parameters

Lick indices were designed for massive galaxies with significant line broadening through random stellar motion (leading to the stellar velocity dispersion), and as such have very wide pass-bands up to 50 Å across. This width has the advantage of increasing S/N and making measurements robust, whilst unfortunately concurrently combining the absorption lines from a large number of chemical elements. This results in making the derivation of individual elemental abundances
and therefore galaxy ages and other properties a highly non-trivial task. A brief summary of the process will be provided here, but for a detailed description please see Thomas et al. (2010, 2011); Johansson et al. (2012).

Our measured Lick indices are fit to stellar population models using a $\chi^2$ minimisation technique introduced by Proctor & Sansom (2002), that minimises the difference between the observed indices and those of the models. The procedure that does this is an iterative process. First, a best-fit is obtained using all chosen indices. A probability distribution is calculated for different values of $\chi^2$ at its minimum, giving a probability $P$ that the observed $\chi^2$ for a correct model should be less than the $\chi^2$ obtained in the fit (i.e. the probability that the best fit is wrong). If $P \geq 0.999$ then the fit is deemed unacceptable, the index with the highest $\chi^2$ is removed from the fitting process and the entire procedure begins anew. This is repeated until $P < 0.999$. This is the technique used at each stage of the fitting process. The indices Ca4455, Fe5015, Fe5270, Fe5782, NaD, TiO$^1$ and TiO$^2$ were never used in the fits as they were too commonly rejected during the initial tests of the $\chi^2$ fitting process from Johansson et al. (2012). I found that for our stacks no index was required to be removed from the fits other than those that were chosen to be discarded pre-fit.

The models we fit to are those of Thomas et al. (2011) at Lick resolution, which are single stellar population models of Lick absorption line indices with variable element abundance ratios. These are an extension of the models of Thomas et al. (2003) and Thomas et al. (2004), which are based on the models of Maraston (1998, 2005). The globular cluster calibrations used for these models are discussed in Thomas et al. (2011). It is important to note that each fit was made to a single starburst, not a composite spectrum with multiple phases of star formation.

Initially a fit is made varying age, [Z/H] and [O/Fe] using the indices Mgb, the Balmer indices H$\delta_A$, H$\delta_F$ and H$\beta$, and the iron indices Fe4383, Fe5270, Fe5335 and Fe5406. These indices were chosen for this stage as they are well calibrated with galactic globular clusters without element abundance variations Thomas et al. (2011), and are sensitive to the parameters being fitted.

Fixing these initial parameters, further fits are performed to calculate [C/Fe], [N/Fe], [Mg/Fe], [Ca/Fe] and [Ti/Fe] individually, looping over all ratio minimisations up to 5 times to determine an optimal fit. The indices used here were those sensitive to the elements being investigated (CN$^1$, CN$^2$, Ca4227, G4300,
H$_{\gamma}$A, H$_{\gamma}$F, Fe4668, Mg$_1$ and Mg$_2$ for carbon; CN$_1$, CN$_2$ and Ca4227 for nitrogen; Mg$_1$ and Mg$_2$ for magnesium; Ca4227 for calcium and Fe4531 for titanium).

Finally, a fit was performed to re-calculate age, [Z/H] and [O/Fe], with the other abundances fixed at the values calculated above. These parameters were then fed back to the abundance fitting loop to re-calculate abundances with the base parameters fixed at these new values. This whole loop was iterated a maximum of 5 times in order to calculate the best-fit stellar population parameters.

Errors on these values were estimated using Monte Carlo simulations, as we did for the Lick indices calculations. I did this by creating sets of simulated Lick indices, as defined by

$$A_{MC} = A_{obs} + \text{RAND} \times \sigma_A$$  \hspace{1cm} (2.7)

where $A_{MC}$ is the simulated index, $A_{obs}$ is the measured value of that index, RAND is a randomly generated number taken from a Gaussian distribution centred at zero, and $\sigma_A$ is the Lick index error discussed in Section 2.4. We then ran our stellar population parameter code over these simulated indices, and recorded the results. We repeated this 1000 times for each object and took the standard deviation of the results, which we take as our measurement error.

### 2.5 Testing the pipeline with SDSS-I/II DR7

I compare our pipeline results for a subset of Sloan Digital Sky Survey (SDSS) Data Release 7 (DR7) galaxies with the MPA-JHU values published in DR8 ([Aihara et al.] 2011). When looking at emission line statistics we use observed, not dust-corrected, values in order to reduce the impact of differences caused by choice of reddening equation. Furthermore, it should be noted that MPA-JHU line fluxes are rescaled on a per-plate basis such that the mean r-band flux in the spectrum matches the r-band fiber mag from the photometry. We have therefore multiplied our line fluxes with this rescaling factor provided by the ‘spectofiber’ keyword in the MPA-JHU database.

Fig. 2.10 shows the comparison between our stellar velocity dispersion measurements (described above) and those in the MPA-JHU catalogue. The measurements are in good agreement with a negligible median offset in $\sigma$ of 1% and a dispersion of 11%.

Fig. 2.11 shows the comparison for emission line fluxes and EWs of [OII]λ3726+.
Figure 2.10: Velocity dispersions measured in this work for a subset of SDSS galaxies in comparison with the measurements published in the Data Release 7. Colour indicates number of galaxies (scale given by the colour bar on the right-hand side). There is good agreement between the measurements. The median offset is 1% with a dispersion of 11%.

3729, \( H\beta \), \([\text{OIII}]\lambda 5007\), \( H\alpha \), and \([\text{NII}]\lambda 6583\). It can be seen that the measurements generally agree well showing tight correlations with small scatter and only small offsets. Median offsets in emission line flux measurements are below 0.02 dex with a dispersion of \( \sim 0.1 \) dex. Only \([\text{OII}]\lambda\lambda 3726 + 3729\) is slightly more offset by \( \sim 0.1 \) dex with a somewhat larger dispersion of \( \sim 0.2 \) dex. This may not be surprising, as the \([\text{OII}]\lambda\lambda 3726 + 3729\) doublet is barely resolved at SDSS spectral resolution, and the measurement is therefore more uncertain. Similar offsets, even though somewhat larger, are present for the EWs. Median offsets are below 0.04 dex for \([\text{OIII}]\lambda 5007\), \( H\alpha \), and \([\text{NII}]\lambda 6583\), while they increase to 0.1 dex and 0.25 dex for \( H\beta \) and \([\text{OII}]\lambda\lambda 3726 + 3729\), respectively, with a larger dispersion of 0.2 dex. These larger discrepancies in the EW measurements will most likely be caused by differences in the treatment of continuum fitting. As for the fluxes, discrepancies in \([\text{OII}]\lambda\lambda 3726 + 3729\) are further caused by uncertainties in the measurement of this line, as the doublet is not resolved.
Figure 2.11: Emission line fluxes and equivalent widths for [OII]λλ3726 + 3729, H\textbeta, [OIII]λ5007, H\alpha, and [NII]λ6583 measured in this work for a subset of SDSS-I/II galaxies in comparison with the measurements published in the Data Release 7. [OII]λλ3726 + 3729 has been calculated as the sum of [OII]λ3726 and [OII]λ3729. Colour indicates number of galaxies (scale given by the colour bar on the right-hand side). Observed, non dust-corrected, values are used. Emission line fluxes and equivalent widths are slightly higher in this work by \sim 0.1 dex.
Chapter 3

The GAMA survey

The Galaxy And Mass Assembly (GAMA) survey is a multi-wavelength galaxy survey that currently, in phase-II, covers 387.6 square degrees and consists of over 280,000 spectroscopically observed galaxies up to $z \sim 0.4$ (Driver et al., 2011). It has two primary objectives: to utilise the distribution of galaxies in the Universe to probe the Cold Dark Matter (CDM) paradigm, and to investigate the internal structure and evolution of galaxies themselves. Within this survey, we are interested in the latter.

We aim to use the data from this survey to investigate the relationships between different emission line classes and objects with no emission lines, as functions of stellar mass, local environment and larger-scale (global) environment. We look to address the presence and significance of environmental and relatively secular feedback processes, such as satellite and AGN quenching respectively, and to disentangle the roles of mass and environment in overall star formation quenching.

To this end, I have constructed a pipeline to analyse the emission line statistics and stellar kinematics of GAMA galaxies, which has been used in several science papers as detailed in Section 3.4. These works investigate the relationship between star formation, mass and metallicity; the impact of environment on radio galaxy properties; and the impact of radio galaxy properties on star formation processes. Note that unless stated otherwise our data extends to internal GAMA release v21, which includes all of the first five years of observations and part of the sixth.

The survey area covers five regions in the sky which are listed in Table 3.1. Fig. 3.1 shows these fields as orange in RA and redshift space, where RA is plotted
CHAPTER 3. THE GAMA SURVEY

<table>
<thead>
<tr>
<th>Field ID</th>
<th>RA min</th>
<th>RA max</th>
<th>DEC min</th>
<th>DEC max</th>
<th>Sqr. Deg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>G02</td>
<td>29.85</td>
<td>44.15</td>
<td>-36.0</td>
<td>-30.0</td>
<td>85.8</td>
</tr>
<tr>
<td>G09</td>
<td>129.0</td>
<td>141.0</td>
<td>-3.0</td>
<td>3.0</td>
<td>72.0</td>
</tr>
<tr>
<td>G12</td>
<td>174.0</td>
<td>186.0</td>
<td>-3.0</td>
<td>3.0</td>
<td>72.0</td>
</tr>
<tr>
<td>G15</td>
<td>211.5</td>
<td>222.5</td>
<td>-3.0</td>
<td>3.0</td>
<td>72.0</td>
</tr>
<tr>
<td>G23</td>
<td>331.9</td>
<td>346.2</td>
<td>-36.0</td>
<td>30.0</td>
<td>85.8</td>
</tr>
</tbody>
</table>

Table 3.1: The locations, sizes and depths of the five GAMA fields.

as the angle and redshift as the radius. This figure also shows the footprint of the 2dFGRS (cyan), SDSS DR9 (dark blue) and 6dFGS (green), from which it is clear that GAMA extends deeper in redshift than all of them. All five fields are 98% spectroscopically complete within the selected magnitude range, where completeness refers to the fraction of objects assigned at least one fibre over the whole survey compared to the number of objects in the input target catalogue. The spectroscopic requirement of the survey was to achieve an extremely high level of completeness for all objects within its target selection, which it clearly achieved. This depth and completeness makes it an excellent complement to the SDSS as it probes down to lower stellar masses (due to the lower magnitude limit), and the range of data available (such as stellar masses, local environments and group environments) makes it very beneficial for emission line studies.

GAMA combines optical spectroscopy taken with the AAOmega spectrograph (Saunders et al., 2004; Smith et al., 2004; Sharp et al., 2006) mounted on the Anglo-Australian Telescope (AAT, Siding Spring Observatory, NSW, Australia) with spectra taken at Apache Point Observatory (APO, Sunspot, NM, USA) for SDSS DR7 (Abazajian et al., 2009), and with photometry taken at a variety of telescopes. Optical and near-IR imaging is obtained through the SDSS, UKIDSS (UKIRT Infrared Deep Sky Survey), VISTA (Visible and Infrared Survey Telescope for Astronomy), and the VST (VLT Survey Telescope). UV data is taken from the GALEX (Galaxy Evolution Explorer), mid-IR from WISE (Wide-Field Infrared Survey Explorer), far-IR from the Herschel Space Observatory and radio from ASKAP (Australian Square Kilometre Array Pathfinder) and GMRT (Giant Metrewave Radio Telescope).

The spectroscopy input catalogue was constructed using imaging from the SDSS DR7 (Abazajian et al., 2009) and UKIDSS (Dye et al., 2006) and is described in Baldry et al. (2010) (note that Baldry et al. (2010) is based on GAMA-I,
Figure 3.1: Shows the distribution of galaxies in RA and redshift space, where RA is plotted as angle and redshift as distance from the centre. Galaxies observed for the GAMA survey are shown in orange, the 2dFGRS survey in cyan, SDSS DR9 in dark blue and 6dfGS in green. Concentric circles are displayed to show the lookback time (Gyr) at fixed redshifts. *Image taken from the presentation ‘Status of the redshift survey’ by Joe Liske, given at the GAMA Team Meeting, LJMU, September 2013.*
an earlier version of the survey, and as such utilises SDSS DR6 instead of DR7. Further, different magnitude limits are applied for the \( r \)–band than displayed here, as both of these were changed for GAMA-II). Since the objective of GAMA was to obtain a high completeness in a given magnitude range, and not select a particular type of galaxy, the only target selections applied were the completeness limits and star-galaxy separation. The main survey selection magnitudes are given in Eq. 3.1, where \( r_{\text{petro}} \) is the Petrosian \( r \)–band magnitude from the SDSS, \( z_{\text{model}} \) and \( r_{\text{model}} \) are the \( z \)– and \( r \)–band model magnitudes from SDSS, and \( K_{\text{AB,auto}} \) is the \( K \)–band AUTO magnitude in the AB system from UKIDSS. These requirements were taken from Baldry et al. (2010) and the GAMA internal webpage at the time of publishing. An object is only required to satisfy one of these conditions in order to be selected for spectroscopy.

\[
r_{\text{petro}} < 19.8 \quad \text{OR} \quad z_{\text{model}} < 18.2 \quad \text{AND} \quad r_{\text{model}} < 20.5 \quad \text{OR} \quad K_{\text{AB,auto}} < 17.6 \quad \text{AND} \quad r_{\text{model}} < 20.5 \tag{3.1}
\]

A star-galaxy separation partially based on but different to that of Sloan is applied by the GAMA team, detailed in Baldry et al. (2010). The SDSS star-galaxy separation parameter cut is simply \( \Delta_{sg} = r_{\text{psf}} - r_{\text{model}} \), where \( r_{\text{psf}} \) and \( r_{\text{model}} \) are the \( r \)–band PSF and model magnitudes, and a simple limit of \( \Delta_{sg} > 0.24 \) was imposed to select an object for spectroscopic observation. Here, \( \Delta_{sg} \) is a measure of how much an object deviates from appearing as a point source, specifically if an exponential profile fit accounts for more flux than a PSF fit. Analysis has shown, however, that this cut excludes some compact galaxies, especially when observing at deeper magnitudes (Baldry et al., 2010). As such, the first GAMA cut applied is to select objects with \( \Delta_{sg} > 0.05 \), which removes the majority of unresolved objects (mostly stars and quasars). Since this cut includes many binary star systems as well as barely resolved galaxies, it is necessary to provide further cuts to select the latter and remove the former. These cuts are shown in Eq. 3.2 where \( \Delta_{sg,jk} \) is defined in Eq. 3.3 \( f_{\text{focus}}(x) \) in Eq. 3.4 and \( f_{\text{sg,step}}(x) \) in Eq. 3.5 (Baldry et al. 2010). In these equations, \( J_{AB} \) and \( K_{AB} \) are the \( J \)– and \( K \)–band AB magnitudes from UKIDSS scaled to an aperture of \( \sqrt{2}'' \), and \( g \) and \( i \) are the \( g \)– and \( i \)–band model magnitudes from
CHAPTER 3. THE GAMA SURVEY

SDSS.

\[ \Delta_{sg} > 0.25 \quad \text{OR} \]
\[ \Delta_{sg} > 0.05 \quad \text{AND} \quad \Delta_{sg,jk} > 0.20 \quad \text{OR} \quad (3.2) \]
\[ \Delta_{sg} > f_{sg,slope}(r_{model}) \quad \text{AND} \quad \text{no } J-K \text{ measurement} \]

\[ \Delta_{sg,jk} = J_{AB} - K_{AB} - f_{focus}(g-i) \quad (3.3) \]

\[ f_{focus}(x) = -0.89 + 0.615x - 0.13x^2 \quad \text{for} \quad 0.3 < x < 2.3 \]
\[ -0.7172 \quad x < 0.3 \]
\[ -0.1632 \quad x > 2.3 \]

\[ f_{sg,slope}(x) = 0.25 - \frac{1}{15}(x - 19) \quad \text{for} \quad 19.0 < x < 20.5 \quad (3.5) \]
\[ 0.15 \quad x > 20.5 \]

These cuts were made to ensure that as many galaxies as possible were spectroscopically observed within the magnitude limits, with minimal contamination from galactic sources, in GAMA-II. Spectra were taken from the SDSS DR7 where applicable, and were otherwise observed at the AAO. The GAMA-observed AAO spectra cover a wavelength range from 3750 to 8850 Å with a resolution and signal-to-noise that are wavelength dependent, from \( R \sim 1000 \) and \( S/N \sim 1-5 \) per pixel at the blue end to \( R \sim 1600 \) and \( S/N \sim 1-10 \) per pixel at the red end (Hopkins et al., 2013), and account for 286,725 galaxy spectra. The SDSS DR7 spectra cover 3900 − 9100 Å with \( R \sim 2000 \) and \( S/N > 3.9 \) (Abazajian et al., 2009) and account for 30,924 galaxy spectra. Note that these spectrum counts include duplicates and scientifically unsatisfactory spectra.

A full description of the spectroscopic processing applied to the data then taken can be found in Hopkins et al. (2013). We include here a brief summary of
this work.

Software developed at the AAO, 2dFDR [Croom et al. 2004], was used to process the raw 2D spectra observed at the AAT and extract 1D spectra by the GAMA data reduction team. This program performs bias subtraction, flat-fielding, fibre trace fitting and wavelength calibration. Bias subtraction refers to the removal of the small background signal inherently within the CCD itself, caused by the constant DC voltage maintained within the camera electronics in order to stop the signal from measuring negative. Flat-fielding removes any artefacts caused by the pixel-to-pixel sensitivity, both in the dark frame (when the shutter is closed) and when light is being recorded (i.e. the variance in the gains recorded by each pixel from a given amount of light). Fibre trace fitting involves finding the location of all of the flux from a given fibre on the CCD, extracting this flux using a minimum variance Gaussian weighted method, and then correcting this flux for distortions caused by the optics in the spectrograph. Finally, wavelength calibration is the process of assigning the correct wavelength to each flux density measurement, which was done to an accuracy of better than 0.1 Å as measured from key sky line features.

Around 25 fibres were allocated on each plate for sky positions in order to perform sky subtraction, where the sky spectrum was made up of the median of the corresponding pixels in each of the normalised sky fibres bar the two brightest (in order to avoid inadvertent non-sky flux being added at the fibre location). This was followed by a further sky subtraction based on principal component templates, which reduced the amplitude of the sky subtraction emission residuals to below 1% in most cases. Atmospheric telluric absorption in the red part of the spectrum was corrected for by combining all spectra in a given field and then iteratively clipping to remove residual emission or absorption features. This takes advantage of the range of redshifts in a field resulting in actual galaxy features being offset from one another and averaging out, and allows for the average spectrum to be fit by a low-order polynomial around the telluric features whilst ignoring the actual absorption. Dividing through by this fit results in a correction spectrum that is set to unity outside of the telluric absorption bands, and can be used to correct each individual spectrum for telluric absorption by acting as a denominator for it.

Finally the blue and red spectra were spliced together for each galaxy by doing a first-pass flux calibration to match the spectra optimally at the slice-wavelength
of around 5700 Å. The final pixel scale is \( \sim 1 \, \text{Å/pix} \).

Redshifts were then obtained using a GAMA-specific version of \texttt{runz}, described in Driver et al. (2011) and detailed further in Liske et al. (prep). In short, \texttt{runz} identifies an automatic redshift using a cross-correlation approach, but then allows for user selection of an optimal redshift if the automatic one is deemed not good enough and assigns a quality flag to each measurement. The automatic redshift is determined by cross-correlating the spectrum with a range of observed templates, and then fitting Gaussian models to emission lines and searching for multi-line matches. The spectral lines fitted can be found in Table 3.2. This process is repeated by multiple team members in order to obtain a probabilistically defined normalised quality scale. I helped perform this redshifting process for one batch of observations. Redshifts are assigned a subjective, user-defined quantity, ‘Q’ by each team member. The values of Q assigned are used to calculate the probability for each redshifter that they find the correct redshift as a function of Q. These probabilities, combined with Q values, are used to assign a most likely redshift to each object with a normalised quality flag, ‘nQ’. This nQ is what is finally used to assign a quality to each redshift. A small quantity (\( \sim 3\% \)) of the GAMA AAO spectra are affected by fringing, although \( \sim 50\% \) of these still allow for a good quality redshift measurement.

Fig 3.2 shows galaxy redshift distributions for the whole spectroscopic GAMA survey (left, blue), the spectroscopic SDSS DR7 sample (left, green), the AAO observed GAMA objects (right, black) and the APO observed GAMA objects (right, red). We compare against DR7 as this is considered the benchmark for large scale spectroscopic surveys. The mean redshift of all GAMA objects is \( \bar{z} = 0.24 \pm 0.14 \), of AAO observed GAMA objects is \( \bar{z} = 0.25 \pm 0.14 \), of APO observed GAMA objects is \( \bar{z} = 0.14 \pm 0.12 \) and of SDSS DR7 is \( \bar{z} = 0.14 \pm 0.10 \). This means that GAMA probes a mean factor of 1.6 deeper in time than SDSS DR7.

The spectra are then spectrophotometrically flux calibrated following the \texttt{IDLspec2d} pipeline used for the SDSS DR6 (Adelman-McCarthy et al. 2008), primarily to provide an approximate absolute flux calibration and to correct for the wavelength-dependence of the system throughput. Curvature corrections and relative flux calibrations are determined from the standard stars on each plate (of which there are \( \sim 3 \) per plate), and the absolute calibration is determined such that the integrated flux of the GAMA objects over the SDSS filter curves
<table>
<thead>
<tr>
<th>Name</th>
<th>$\lambda$ (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C4</td>
<td>1548.98</td>
</tr>
<tr>
<td>C3</td>
<td>1908.66</td>
</tr>
<tr>
<td>Mg2</td>
<td>2798.70</td>
</tr>
<tr>
<td>[OII]</td>
<td>3727.80</td>
</tr>
<tr>
<td>K</td>
<td>3933.66</td>
</tr>
<tr>
<td>H</td>
<td>3968.47</td>
</tr>
<tr>
<td>Hδ</td>
<td>4101.73</td>
</tr>
<tr>
<td>G</td>
<td>4304.40</td>
</tr>
<tr>
<td>Hγ</td>
<td>4340.46</td>
</tr>
<tr>
<td>Hβ</td>
<td>4861.33</td>
</tr>
<tr>
<td>[OIII]</td>
<td>4958.91</td>
</tr>
<tr>
<td>[OIII]</td>
<td>5006.84</td>
</tr>
<tr>
<td>Mg</td>
<td>5175.30</td>
</tr>
<tr>
<td>Na</td>
<td>5894.00</td>
</tr>
<tr>
<td>Hα</td>
<td>6562.80</td>
</tr>
<tr>
<td>[NII]</td>
<td>6583.46</td>
</tr>
<tr>
<td>[SII]</td>
<td>6716.44</td>
</tr>
<tr>
<td>[SII]</td>
<td>6730.81</td>
</tr>
</tbody>
</table>

Table 3.2: The emission lines used to estimate redshifts, and their corresponding vacuum wavelengths.

Figure 3.2: Left: The redshift distributions of all spectroscopically observed galaxies from the SDSS DR7 with ‘zwarning’ equal to 0 (green), and all objects in the GAMA survey with spectra observed at either the AAO or APO and with scientifically acceptable redshift quality (blue). Right: The redshift distributions of the GAMA objects observed at the AAO (black) and APO (red) with scientifically acceptable redshift quality.
matches the Petrosian magnitude of the SDSS photometry of each object. Standard stars are typed by being compared to theoretical spectra using Kurucz model atmospheres (Kurucz 1992) and the spectral synthesis code spectrum (Gray & Corbally, 1994). A flux correction vector is then formed by taking the ratio of the observed spectrum to its best fit model, and can be used to perform a curvature correction to account for lower CCD response at the extreme wavelengths. An absolute flux calibration is finally obtained by tying the spectrophotometry to the SDSS $r$-band Petrosian magnitudes, accomplished by multiplying individual spectra by the SDSS $r$-band filter response.

Some objects were observed by the SDSS and GAMA, and were used to test the processing applied above. The spectra were found to be in good agreement across the entire wavelength range (to around 10%), but particularly in the central wavelengths.

From here on when talking about spectra, we use the term ‘GAMA’ refer to the combination of AAO and APO spectra in the GAMA fields.

### 3.1 Our data products

One of the notable features of the GAMA collaboration is that any data products produced by the team must be robustly quality tested by an independent group before they can be released and used for science and in publications. Once they have passed the quality control (QC) process, they make up a GAMA data management unit (DMU) which is then made available on the GAMA website. Before submitting the results to the QC process all data types must have a description and units, and an accompanying ‘notes’ file must be provided which contains a detailed description of how all quantities were obtained, any limitations and any caveats. All data products must also not be provided in vector form, i.e. each individual emission line and property must have its own entry for each object. During the QC process all data products are tested to ensure that they behave as expected, all descriptions are checked and great care is taken to ensure that the results are robust and reliable. Any results deemed unacceptable or unreliable will not be passed through the QC process and therefore will not be usable by the collaboration.

We provide a main DMU to the collaboration, the ‘spectroscopic analysis’ DMU, which contains all of our kinematics and emission line analysis for the
GAMA objects (see Appendix A for the complete list). The creation of this DMU involved numerous reiterations of our catalogue and private correspondences with the QC team, especially Joe Liske. A description of the catalogue is included in Hopkins et al. [2013], and we provide our own analysis below.

For this DMU, and our own analysis, I ran the pipeline described in Section 2 over the GAMA spectra observed at both the AAO and the APO. I first created two sets of single stellar population templates downgraded to the AAO and APO resolutions, in order to calculate velocity dispersions accurately. I then ran the APO and AAO spectra separately through the pipeline, and concatenated the results afterwards. It is important to note that whilst the raw AAO spectra are provided already log binned, the APO spectra are provided linearly binned. Our pipeline however takes care of this.

As part of the design, many objects have been observed multiple times and therefore have multiple spectra, to the extent that $\sim17.7\%$ of all galaxy spectra are redundant repeat observations. This is generally done to obtain an accurate redshift when the first spectrum, taken by whichever survey, was too noisy. Whilst we run all spectra through the pipeline and obtain results, it is clearly not a good idea to look at the results of all spectra when performing scientific analysis, rather the results from the best spectrum for each object should be used. Fortunately a flag for this is included in the provided GAMA spectra. Further, not all spectra have decent redshifts attributed to them, even when the best observation for a single object. Clearly those spectra with poor redshifts should also not be used for science, as a bad redshift will make stellar population and emission line fits meaningless.

Bearing this in mind, all of the results and analysis below are done on spectra which are flagged as the best observation for their target object, and have a redshift quality that is good enough for science ($nQ>2$). The breakdown of the counts of AAO- and APO-spectra under various selection criteria can be seen in Table 3.3.

### 3.1.1 Features of the spectra

Fig. 3.3 shows the distribution of the measured velocity dispersions ($\sigma s$) of the GAMA sample, in log space. The left panel shows the distribution for the entire survey in solid black, with the distribution of objects with a velocity dispersion error of less than 30% ($d\sigma/\sigma < 0.3$) in dashed black. The right panel shows the
all spectra  nQ > 2  optimal  optimal & nQ > 2
AAO  286,725  246,683  236,667  222,295
APO  30,954  29,970  24,836  24,464
TOTAL  317,679  276,653  261,503  246,759

Table 3.3: The number of spectra observed at each site and in total which pass through various selection criteria.

Figure 3.3: The distributions of log stellar velocity dispersions for GAMA objects with nQ > 2 and where the best spectrum was selected in cases of multiple observations. Left: the distribution for the entire survey where the solid line represents all objects and the dashed line objects with $d\sigma/\sigma < 0.3$. Right: this distribution is further separated into AAO-observed spectra (black) and APO-observed spectra (red), in the redshift range $0.15 < z < 0.20$, and normalised to 1.

distributions normalised to 1 for the APO spectra (black) and AAO spectra (red) separately, for objects with $d\sigma/\sigma < 0.3$, in the redshift range $0.15 < z < 0.20$. This restriction in redshift is made to separate out the effects of the APO observations probing lower in redshift from the effects of the difference in magnitude limit, which we are choosing to highlight here. The normalisation to 1 is useful to compare the shapes of the APO and AAO distributions, as there is a significant number more unique, scientifically satisfactory AAO spectra than APO spectra with $d\sigma/\sigma < 0.3$ (127,525 compared to 12,500).

Comparing the AAO and APO velocity dispersions we can see that the two distributions peak at different values, with AAO observations at a given redshift probing lower in $(\sigma)$-space. This is to be expected from the difference in magnitudes probed, as the lower magnitudes observed at the AAO allows for the detection of less luminous, and therefore less massive, galaxies. Whilst the higher-sigma shape of the distribution is expected to be physical, as high-sigma
objects are generally more luminous and therefore likely to be detected and observed, the shape of the lower-sigma side is largely caused by selection effects as lower-sigma objects are less luminous and therefore less likely to be detected.

Fig. 3.4 shows the distributions of the EWs of Hβ, [OIII] λ5007, Hα and [NII] λ6583 for all objects in the GAMA survey with non-zero emission line measurements. The black histogram contains all objects, whilst the red contains only those objects with AoN > 2 for the relevant line. The mean values are shown as vertical dashed lines of the corresponding colour. Table 3.4 shows the statistics for each panel including the number of objects in each histogram, their means and their standard deviations. Note that the total number of objects in each is not consistent due to not only some objects having particular line measurements of zero, but also the higher wavelength lines being redshifted out of the observed spectrum at $z \sim 0.5$.  

Figure 3.4: The distributions of the EWs of Hβ (top left), [OIII] λ5007 (top right), Hα (bottom left) and [NII] λ6583 (bottom right) in the GAMA survey for objects with non-zero EW measurements. The black histogram contains all objects, whilst the red contains only those objects with AoN > 2 for the relevant line.
We can see that the mean EWs of objects with AoN > 2 are higher than for those without; this is an obvious result of the signal increasing and as such the S/N increasing as well, and serves as a sanity check. We also find that Hα has the highest ‘good’ (AoN > 2) mean, followed by [NII], Hβ and finally [OIII]. This corresponds nicely to the number of objects found to have AoN > 2, again expected due to its definition, but a welcome result to find.

Looking at the distributions we can see that the overall sample appears to consist of a convolution of two Gaussians, one with a mean located approximately at the mean of the AoN; > 2 distribution and one located at EW = 0; this is due to objects either having measurable emission lines, and therefore having high AoN, or having effectively no emission lines but with a measurement scatter around 0 caused by noise in the data.

Fig. 3.5 shows the BPT diagnostic diagram, described in Section 2.3, for the GAMA data, including all objects with AoN > 2 for all four BPT lines and z < 0.5 due to the loss of [NII] λ6583 beyond this point. We use the empirical separation between star forming galaxies and AGN from Kauffmann et al. (2003) (dashed line) and the theoretical extreme starburst line from Kewley et al. (2001) (solid curved line) to identify pure star forming and pure AGN emission. We adopt the common practise of assuming the area between these lines is populated by galaxies with composite star forming and AGN spectra. We further use the dividing line defined by Schawinski et al. (2007) (solid straight line) to distinguish between LINER and Seyfert emission based on SDSS galaxy classifications obtained through the [SII]/Hα ratio. Contours show the 5th, 25th, 50th, 75th and 95th percentiles. We find that 12.5% of the total sample have all four BPT lines with AoN > 2, of which 80.1% are classified as star forming, 14.0% as composite, 3.8% as Seyfert and 2.1% as LINER. These fractions are heavily
Figure 3.5: BPT diagram for the entire GAMA survey, showing objects with all four BPT lines having AoN > 2 and z < 0.5. We use the empirical separation between star forming galaxies and AGN from Kauffmann et al. (2003) (dashed line) and the theoretical extreme starburst line from Kewley et al. (2001) (solid curved line) to identify pure star forming and pure AGN emission. We adopt the common practise of assuming the area between these lines is populated by galaxies with composite star forming and AGN spectra. We further use the dividing line defined by Schawinski et al. (2007) (solid straight line) to distinguish between LINER and Seyfert emission based on SDSS galaxy classifications obtained through the [SII]/Hα ratio.
biased by selection effects, and should not be considered representative of the galaxy population as a whole.

3.2 External data products

The GAMA collaboration share numerous value-added data products with each other, enabling science to be done with minimal repetition of analysis. Here we introduce the four products, produced by other people, that we use in our work: stellar mass estimates (Taylor et al., 2011), local environmental density measures (Etherington et al. [prep]), group membership and properties (Robotham et al., 2011), and Gaussian models of emission lines (Brough et al., 2011; Hopkins et al., 2013).

3.2.1 Local environment

The intricacies of the definition of local environment can have a significant impact on any trends observed with it, and as such the choice of measure is very important to any study. For the purpose of our work we use the adaptive Gaussian environment (AGE) density as formulated in Schawinski et al. (2007) and used in e.g. Thomas et al. (2010), which was designed to compensate for the “finger-of-God” effect in high density environments, where redshift can be significantly affected by peculiar velocities resulting in galaxies appearing further away from the centre of a density than they really are. The code was run by J. Etherington, and details of the method used can be found in Etherington et al. (prep).

This method involves taking a volume-limited sample, which we create by limiting our redshift range to $z \leq 0.18$, and then creating an adaptive volume around each target galaxy where the length of the volume in the radial direction (i.e. that governed by redshift) is scaled according to the number of galaxies in a fixed volume around the target. This scaling factor ($c_z$) is then employed, along with an arbitrary dispersion factor ($\sigma$) and the angular ($r_a$) and line of sight ($r_z$) distances to each galaxy from the target, to define the AGE parameter $\rho_g$ as the sum over all neighbours within the ellipse defined by

$$\left(\frac{r_a}{3\sigma}\right)^2 + \left(\frac{r_z}{3c_z\sigma}\right)^2 \leq 1$$

(3.6)
i.e. we search out to $3 \sigma$:

$$
\rho_g(\sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left[ -\frac{1}{2} \left( \frac{r_a^2}{\sigma^2} + \frac{r_z^2}{c_z^2\sigma^2} \right) \right].
$$

(3.7)

Following Schawinski et al. (2007) we take an arbitrary value of $\sigma$ of 2.0 Mpc to set the scale length to focus approximately on the scale of large groups and small clusters, although they performed tests to confirm that changing $\sigma$ does not alter results within 1 standard deviation. It should be noted that the target galaxy itself is not included in the total density, and that this method clearly assigns a much higher weight to nearby galaxies.

### 3.2.2 Stellar mass estimates

Photometrically-derived stellar mass estimates were calculated by Taylor et al. (2011) as a part of GAMA data release 2. Only optical imaging was used, as tests were made also using NIR photometry and it was found that this inclusion significantly decreased the quality of population synthesis fits. The reason for this was uncertain, but it was suggested that either there were problems in their stellar library or with the NIR data itself.

Stellar masses were derived by fitting composite stellar populations based on the synthetic single stellar population models of Bruzual & Charlot (2003) assuming a Chabrier (2003) IMF, convolved with a Calzetti et al. (2000) dust law, to the GAMA broadband optical SED $ugriz$-bands, and the $YJHK$-bands for NIR investigations. Template ages ranged from 100 Myr to 13.4 Gyr, and metallicities from 0.0001 to 0.05. Fits were performed using $\chi^2$ minimisation technique, with a normalisation factor applied to raise/lower overall flux. The properties of this composite fit can then be ascribed to the galaxy being modelled, and the normalisation factor used to extract stellar mass.

Note that masses here are calculated with the caveat that the contribution from TP-AGB stars is not considered. Maraston (2005) showed that the inclusion of TP-AGB stars is vital for stellar population synthesis models. They are very bright but short-lived, so for galaxies with stellar ages of 0.2 to 2 Gyr they contribute significantly to the red component of the spectrum (Maraston et al., 2006); therefore, if they are not taken into account, a lot of excess mass in red stars is required to match to the luminosities observed when in reality it is sourced by much less mass in TP-AGB stars. Unfortunately at the time of writing, the
Maraston-based masses are unavailable for the dataset used here; this will be amended for the final publication.

Fig. 3.6 shows the relation between our log $\sigma$ and Taylor’s log stellar mass ($M_*$), with the line of best fit in red. Objects were only considered if they fulfilled the criteria $d\sigma/\sigma < 0.3$. The contours represent the 5th, 25th, 50th, 75th and 95th percentiles. It can be seen from this that, as expected, stellar velocity dispersion and stellar mass are correlated albeit with a large scatter.

Fig. 3.7 shows the log $M_*$-$z$ relation, where the blue dashed line shows the $z = 0.18$ cut-off we apply to get a volume-limited sample. We can clearly see that galaxy mass increases with redshift, a selection effect caused by more massive galaxies generally being more luminous and therefore easier to detect (and more likely to therefore be selected) at greater distance. We also note that there is still some evolution within the $z = 0.18$ volume-limited range; this is caused by the magnitude-limited GAMA target selection method. It has indeed been shown
Figure 3.7: The relationship between redshift and log stellar mass. The blue dashed line shows the $z = 0.18$ cut for our volume-limited sample. The contours represent the 5th, 25th, 50th, 75th and 95th percentiles.
in Lara-López et al. (2013) that there is no evolution in metallicity seen in the GAMA sample until $z \sim 0.2$, confirming that this is not of physical origin.

### 3.2.3 Group catalogue

Galaxies were classified into groups by Robotham et al. (2011) via a robust friends-of-friends grouping algorithm, using galaxy-galaxy linking to define groups as opposed to galaxy-halo linking.

Links between galaxies based on their separation were calculated, treating the comoving projected and radial separations separately to account for line-of-sight effects from peculiar velocities in groups and clusters. Galaxies were then determined to be grouped if they fell within a maximum separation distance for both methods. Fig. 3.8 shows schematically how both the projected and radial separations are used to determine if an object is in a group, underlining how either measurement on its own is not enough to determine a true group.

The quality of the group matching was then tested thoroughly on mock catalogues constructed from the Millenium dark matter simulation (Springel et al., 2005) and populated with galaxies using the GALFORM (Bower et al., 2006) semi-analytic galaxy formation recipe. This quality was measured using two-way (bijective) statistics, as it was important that the data group catalogue was an accurate representation of the mock catalogue, and vice versa. The two global measures used to define quality were how well the groups and the galaxies within them were recovered. Bijective groups were defined as those where the joint galaxy populations of the data and mock groups included $> 50\%$ of their respective group members, and the global halo finding efficiency was defined as

$$E_{tot} = \frac{N_{g_{bij}}}{(N_{F_{FoF}} \ast N_{m_{mock}})}$$

where $N_{g_{bij}}$, $N_{F_{FoF}}$ and $N_{m_{mock}}$ are the number of bijective, data-based (FoF) and mock groups respectively. The ‘purity’ of the matching groups was measured by summing the largest products for the relative membership fractions between the FoF and mock groups multiplied by their memberships. The global grouping purity was defined as

$$Q_{tot} = Q_{F_{FoF}}Q_{m_{mock}}$$

for $Q_{F_{FoF}} = \frac{\sum_{i=1}^{N_{F_{FoF}}} P_{F_{FoF}}[i] \ast N_{m_{FoF}}[i]}{\sum N_{m_{FoF}}}$ and $Q_{m_{mock}} = \frac{\sum_{i=1}^{N_{m_{mock}}} P_{m_{mock}}[i] \ast N_{m_{mock}}[i]}{\sum N_{m_{mock}}}$, where $N_{m_{FoF}}[i]/P_{F_{FoF}}[i]$ and $N_{m_{mock}}[i]/P_{m_{mock}}[i]$ are the number of galaxies in/purity product of the $i^{th}$ FoF and mock group respectively. The purity product was defined as the largest fraction of galaxies in FoF/mock group $i$ that can be found in any mock/FoF group. The final statistic used was $S_{tot} = E_{tot}Q_{tot}$, which spans the range $0 – 1$. 
Figure 3.8: Schematic of the two step process used when associating galaxies via FoF algorithm on redshift survey data. The same set of galaxies are shown in two panels: along the line of sight (left) and projected on the sky (right). Both the radial and projected separations are used to disentangle projection effects and recover the underlying group (galaxies 1, 5 and 6 in this example). The radial linking length has to be significantly larger than the projected one to properly account for peculiar velocities along the line of sight. Image and caption taken Figure 1, Robotham et al. (2011).
CHAPTER 3. THE GAMA SURVEY

Figure 3.9: Comparing the distributions for galaxies found to not be in groups (black) and those found to be in groups (red) of log stellar mass (left) and log density (right). All objects have nQ > 2 and the best spectrum was selected in cases of multiple observations.

Grouping parameters were then optimised by maximising the obtained value of $S_{tot}$ using a standard Nelder-Mead approach. It was found that luminosity is the most fundamentally related parameter to optimal galaxy groups. It was also found that the most successful algorithm was necessarily a conservative one, where haloes are robustly detected and interlopers are kept low in these systems.

Fig. 3.9 compares the distributions of stellar masses (left) and AGE-derived local density measurements (right) for galaxies that are not classified as in groups (black) and those that are (red). A two-sample Kolmogorov-Smirnov (KS) test comparing the group mass sample and the non-group mass sample provides a test statistic value of 0.059, and a corresponding probability that the two samples are drawn from the same distribution of $2.1 \times 10^{-88\%}$. For density the test statistic value is 0.321, again with a probability of being drawn from the same distribution of $< 10^{-323\%}$, although this is to be expected simply from the definition of a group. The difference in mass distributions is primarily caused by the difference in density distributions and a positive density-mass relationship, and is discussed in Section 3.3.3.

Fig. 3.10 shows BPT diagrams for group members (top) and those outside of groups (bottom), from which we see that objects out of groups are more likely to have emission lines, although not by a significant amount. We also see that non-group objects are more likely to be star forming if they have emission lines, whereas galaxies in groups are more likely to be composites, Seyferts or LINERs. Analysis of the emission line properties of group and non-group objects is
Figure 3.10: BPT diagrams for our group objects (top) and nongroup objects (bottom). Symbols and lines have the same meanings as in Fig. 3.5.
performed in Section 3.3.

### 3.2.4 Gaussian emission line models

In addition to the catalogue we provide of emission line measurements, a simple Gaussian model fit was made by Rob Sharp to a selection of common emission lines for each spectrum to obtain basic measurements. This was done by performing a simultaneous iterative $\chi^2$ fitting of positive emission peaks to the common emission lines. The continuum for each line was approximated with a linear fit. A small velocity offset from the measured GAMA redshift was allowed, and a line width common to all lines was fit. Line width was constrained to 3-5 Å.

Flux values for individual lines were rejected if the inclusion of that line in the global fit fails to improve the reduced $\chi^2$ value by a factor of 3. Note that just as in our method, the [OII] doublet was fitted as one line, as the resolution of the AAOmega spectra was not high enough to resolve both lines. Errors on these measurements were those associated with the formal Gaussian fitting process.

I primarily use this product to test our more sophisticated fitting methods, and to ensure that there is no systematic offset in our results. Fig. 3.11 shows the comparison between our EWs and those from the simple Gaussian fits for H$\beta$, [OIII]λ5007, H$\alpha$ and [NII]λ6583, where contours are used to cover areas where point density is too high. We can see that [OIII]λ5007 and H$\alpha$ EW measurements are highly comparable for both methods, whilst gandalf measures slightly ($\sim$ 0.04 dex) lower EWs for [NII]λ6583 and significantly higher ($\sim$ 0.41 dex) EWs for H$\beta$.

The forbidden lines looked at here ([OIII] and [NII]) were chosen as (a) they are largely unaffected by stellar absorption, and (b) they are both used in the BPT diagnostic diagram (see Section 2.3 for a description of the BPT diagnostic). Overall and within errors, the distributions of these lines are consistent with the one-to-one relation. It is likely that the small observed offsets are caused by the different definitions of continuum used by each method.

Looking at the Balmer lines, we see a negligible offset for the H$\alpha$ measurements but a significant offset for the H$\beta$ measurements. This is as expected, as the Balmer lines are strongly affected by stellar absorption which is taken care of inherently by gandalf but treated separately in the Gaussian fit method. The effect is particularly noticeable for the H$\beta$ line, where the absorption can be relatively large when compared to the weak emission line. It is found in...
et al. (2013) that a correction of 2.5 Å applied to the Gaussian-fit Balmer line EWs deals with this absorption discrepancy. Whilst we find that this works for Hα and reduces the offset to < 5%, it does not fully make up for the difference seen for the Hβ line. I consider it likely that a further absorption correction is required for the Hβ line, and find that a value of 5.6 Å is necessary for this correction.

3.3 The environmental dependence of star formation and ionisation characteristics in the GAMA survey

We aim to investigate the dependence of different emission line classes on stellar mass, local density and group membership. I do this in two primary ways.
the entire population I compare different emission line classes to the passive population, in order to look at the mass and local environment effects on quenching. I then compare the same emission line classes in groups and outside of groups, to see how group membership affects the parameter space each class is found in.

3.3.1 Background

The impact of environment on galaxy evolution is a much debated and researched topic (e.g. Dressler 1980; Thomas et al. 2005; Baldry et al. 2006; Schawinski et al. 2007; Bamford et al. 2009; Peng et al. 2010; Rogers et al. 2010; Thomas et al. 2010). It is made difficult by the definition of environment itself being ambiguous; trends could be caused by local environment, such as local density and interactions with nearby galaxies (e.g. Peng et al. 2010), or by global environment, for example the mass of a galaxy’s host dark matter halo (e.g. Rogers et al. 2010). Even within a given type of environment, the method used to determine this parameter can have a major impact on results (for further details see Schawinski et al. 2007; Etherington et al. prep), making it very challenging to get a homogenous view of what’s going on.

The relative importance of mass and environment is also difficult to distinguish, largely due to the degeneracy that exists between the two caused by the mass function depending on environment (see Bundy et al. 2006, for details). When looking at star formation rates (SFRs) for example, some studies favour environment as the major driving factor (e.g. Mendes de Oliveira et al. 2005), whilst others claim that mass is a more significant parameter (e.g. Wake et al. 2005; Bundy et al. 2006; Rogers et al. 2010; Thomas et al. 2010).

Once these effects have been disentangled, it is also important to know which parameters and properties of galaxies to investigate, and how to detect them. Emission lines have long been used to analyse the energetics of galaxies (e.g. Kauffmann et al. 2003; Brinchmann et al. 2004; Tremonti et al. 2004; Sarzi et al. 2006; Schawinski et al. 2007; Thomas et al. 2013), in order to understand the primary ionising source within them. Importantly, as demonstrated by Baldwin et al. (1981) and Veilleux & Osterbrock (1987), they can be used to separate star forming objects, those dominated by AGN and those without emission lines (passive objects) by comparing rates and strengths of high- and low-energy excitation modes. This enables us to investigate the rate of star formation quenching in galaxy populations, and see how (or if) it correlates with AGN activity.
A recent study by Peng et al. (2010) using the SDSS DR7 (Abazajian et al., 2009) and zCOSMOS (Lilly et al., 2007b) data has suggested that there are two different modes of quenching: environmental, which is independent of mass, and mass, which is independent of environment. In this work they use colour to separate star forming objects (blue) and passive objects (red), and show that the fraction of red galaxies is the product of these two modes. They postulate that each mode is powered by a different physical process, termed mass quenching and environment quenching. They put forward the idea that the mass quenching could be caused by AGN activity (e.g. Dekel & Silk, 1986; Silk & Rees, 1998; Silk, 2005; Springel et al., 2005), superwinds from supernovae or some local ionisation model (e.g. Cantalupo, 2010), whilst the environment quenching could be from satellite galaxies falling into larger haloes, although they stress that this is purely conjecture. We aim in part to further this, using emission instead of colours and adding the extra parameter of global environment.

In this section we use the Galaxy And Mass Assembly (GAMA) survey to investigate the impact of environment on emission line properties. We aim to investigate the relationships between mass, environment (both local and global), the fraction of star forming galaxies, the fraction of AGN and the fraction of passive galaxies, and identify where possible the different quenching mechanisms put forward by Peng et al. (2010).

### 3.3.2 Sample selection

We use phase-II of GAMA, in which objects were spectroscopically observed up to an $r$-band magnitude limit of $r_{pet} < 19.8$ to a 98% completeness level in a 280 deg$^2$ survey area. After applying cuts requiring that all objects investigated have $0.01 \leq z \leq 0.18$, $8.5 < \log M_* < 11.5$, $dM_*/M_* < 0.1$ and $0.04 \text{ Mpc}^{-3} < \rho < 8 \text{ Mpc}^{-3}$ where $M_*$ is stellar mass and $\rho$ is local density as described in Section 3.2.1 (requiring that all of our sample was required to have a measurable local environment), we were left with a main sample of 28,697 objects. The limit of $z \leq 0.18$ was a result of our local density calculation requiring a volume limited sample.

Fig. 3.12 shows the log $M_*$ distribution (left) and log $\rho$ distribution (right) of our main sample. I find a mean log $M_*$ value of $10.17 \pm 0.52$, and a mean log $\rho$ of $-0.173 \pm 0.466$. I look into the correlation between these variables in Section 3.3.3.
CHAPTER 3. THE GAMA SURVEY

Looking at group membership, I find 14,362 galaxies in groups and 14,335 galaxies outside of groups. This gives us three final samples: one containing our main galaxy sample, one subset of this containing galaxies in groups and one containing galaxies outside of groups.

Fig. 3.13 shows the distributions of log $M_*$ (left) and log $\rho$ (right) for our group (black) and non-group (red) samples. I perform KS tests comparing the $M_*$ and $\rho$ distributions of these samples, and find that they are both (as is obvious from the figure) significantly different. I find that objects in groups are generally more massive and found in denser environments than those outside of groups, with mean $M_*$ values of $10.26 \pm 0.53$ and $10.08 \pm 0.49$, and mean $\rho$ values of $-0.010 \pm 0.435$ and $-0.336 \pm 0.448$, for group and non-group objects respectively.
### 3.3.3 $M_\ast - \rho$ correlation

When investigating the relationships between different subsets of two variables, it is vital to understand how these two variables depend on each other. An understanding of how these variables are correlated is necessary for the disentangling of their roles and effects.

I first check whether there is an inherent relationship between stellar mass and density for our three samples. Fig. 3.14 shows the distributions of $M_\ast$ vs $\rho$ for our main sample (top), group sample (middle) and non-group sample (bottom) in log-log space. Table 3.5 shows the number of objects ($n$) in each sample and the Pearson’s correlation coefficient ($r$) for the velocity dispersions and local densities of each sample, as well as the t-test statistic ($t_0$). The $t_0$ values were calculated according to the formula

$$t_0 = r \sqrt{\frac{n - 1}{1 - r^2}}$$

and need to have magnitudes higher than $t_{\text{min}}$ in order for the relationship to be significant to 99%.

I find that there is a very significant relationship between $M_\ast$ and $\rho$ for all of our samples, finding that higher mass objects preferentially lie in higher density environments. This is a known phenomenon (e.g. Kauffmann et al. [2004]; Baldry et al. [2006]; Vulcani et al. [2012]), due primarily to low-mass galaxies being found predominantly in sparser environments.
Figure 3.14: The log $M_*$ - log $\rho$ distributions for our main sample (top), group sample (middle) and non-group sample (bottom).
3.3.4 Controlling for a variable

When investigating the effect of an independent variable (e.g. emission line class or group membership) on a dependent variable (e.g. $M_*$) which has a known relationship with a third variable (e.g. $\rho$), it is necessary to control for this third variable in order to disentangle the roles of it and the desired dependent variable.

When controlling for a variable, our objective is to ensure that the distributions of this variable for both samples are the same. My method is detailed here. We call the controlled variable $v_c$, the dependent variable we are looking at $v_d$, and our different samples $a$ and $b$. We first assign uniformly distributed random numbers $u$ in the range $0 - 1$ to $v_c(b)$. We then map the distributions of $v_c(a)$ and $v_c(b)$, which we shall call $d(a)$ and $d(b)$, in $nbin$ bins. The number of bins used depends on the number of objects in each sample. Next we take the fraction $f_{ab} = d(a)/d(b)$, setting $f_{ab} = 0$ where $d(b) = 0$, and normalise $f_{ab}$ to 1. We then loop over each bin, and locate all $v_c(b)$ which both fall into this bin and have $u(b) \leq f_{ab}$; these objects make up our final variable-controlled sample from $b, b_c$. Finally we remove from $v_c(a)$ any objects that fall into an interpolated bin for which no suitable $v_c(b)$ are found, making our variable-controlled sample from $a, a_c$. We can then compare $v_d(a_c)$ and $v_d(b_c)$ to understand whether or not $v_d$ differs between $a$ and $b$ independently of $v_c$.

The effectiveness of this approach was tested using two test samples, both of random numbers generated in a Gaussian distribution, with counts of $n_a = 1000$ and $n_b = 10000$, means of $< v(a) >= -0.012$ and $< v(b) >= 0.404$, and standard deviations of $\sigma(a) = 0.501$ and $\sigma(b) = 1.003$. Applying our variable control method with $nbin = 50$, we end up with counts of $n_c(a) = 998$ and $n_c(b) = 4096$, means of $< v_c(a) >= -0.013$ and $< v_c(b) >= -0.020$, and standard deviations of $\sigma_c(a) = 0.500$ and $\sigma_c(b) = 0.502$. Repeat tests showed the resultant means and standard deviations to be approximately equal to each other, and unbiased. Fig. 3.15 shows the example distributions discussed here, with the original distributions in the left panel and variable-controlled distributions in the right, sample $a$ in red and sample $b$ in black, and the means and standard deviations marked with solid and dashed lines respectively.
3.3.5 Emission Line Properties

The equivalent widths and amplitude-over-noise (AoN) values of the lines [OIII]5007, [NII]6583, Hβ and Hα are measured for all of our sample. The AoN is the ratio of the line amplitude to the standard deviation of the residual spectrum (Sarzi et al., 2006). These lines allow for the usage of the well known BPT diagnostic diagram, introduced in Section 2.3.

Fig. 3.16 shows the BPT diagram for our main sample. We use the empirical separation between star forming galaxies and AGN from Kauffmann et al. (2003) (dashed line) and the theoretical extreme starburst line from Kewley et al. (2001) (solid curved line) to identify pure star forming and pure AGN emission. We adopt the common practise of assuming the area between these lines is populated by galaxies with composite star forming and AGN spectra. We further use the dividing line defined by Schawinski et al. (2007) (solid straight line) to distinguish between LINER and Seyfert emission based on SDSS galaxy classifications obtained through the [SII]/Hα ratio. We find that 33.8% of the total sample have all four BPT lines with AoN > 2, of which 86.5% are classified as star forming, 9.9% as composite, 2.3% as Seyfert and 1.2% as LINER.

Fig. 3.17 shows BPT diagrams for our group (left) and non-group (right) samples, from which we see that objects outside of groups are more likely to have emission lines. This is primarily driven by an increased number of star forming galaxies in non-group galaxies.
Figure 3.16: BPT diagram for our main sample. We use the empirical separation between star forming galaxies and AGN from Kauffmann et al. (2003) (dashed line) and the theoretical extreme starburst line from Kewley et al. (2001) (solid curved line) to identify pure star forming and pure AGN emission. We adopt the common practise of assuming the area between these lines is populated by galaxies with composite star forming and AGN spectra. We further use the dividing line defined by Schawinski et al. (2007) (solid straight line) to distinguish between LINER and Seyfert emission based on SDSS galaxy classifications obtained through the [SII]/Hα ratio.
Figure 3.17: BPT diagrams for our group sample (top) and non-group sample (bottom).
Figure 3.18: The fraction of, from left to right, top to bottom, star forming, composite, Seyfert and LINER objects in log $M_\ast$ - log $\rho$ space for our main sample.

For the purposes of this work we consider galaxies with all four lines having a value of AoN > 2 as emission line galaxies, and all other galaxies as ‘passive’. Taking our samples I investigate how emission line properties depend on both mass and environment.

### 3.3.6 Main sample

Fig. 3.18 shows from left to right, top to bottom, for objects in our main sample, the fractions of star forming, composite, Seyfert and LINER galaxies, in log $M_\ast$ - log $\rho$ space. One of the most striking features of this plot is the apparent importance of mass when compared to environment, when looking at the presence of emission lines. It is clear that at a given environment, mass plays an essential role, whilst at a given mass, environment only plays a role at the highest densities. Another interesting property is that the class of emission line object depends strongly on mass. Composites and Seyferts are clearly found at higher
masses than star forming objects, and LINER’s at higher masses still, in line with the findings of Schawinski et al. (2007). The effect of environment is harder to decipher, mainly due to the small number statistics at play when looking at Seyferts and LINERs.

In order to look at the differences in the importance of mass and environment in a different way, we show the normalised distributions of each variable when controlling for its counterpart in Fig. 3.19, where each emission line class is in a different colour and black indicates the overall distribution of objects. Plots on the left show $M_*$ distributions whilst controlling for $\rho$, and on the right show $\rho$ distributions whilst controlling for $M_*$, each normalised by dividing through by the total number of objects in the distribution. Emission line classes from top to bottom are star forming (blue), composite (purple), Seyfert (green) and LINER (red). Errors shown are Poisson errors, and the values shown are the mean $\pm$ the standard deviation of each distribution in respective colours.

From these it is again clear that the difference in parameter space between each emission line class and the overall distribution is strongest for mass, and minimal for environment. Looking at mass distributions, we can confirm that the mass distributions of each BPT class and the whole population at fixed environments are significantly different. We show unambiguously that the class-with-mass trend tentatively observed in Fig. 3.18 is independent of environment, with star forming objects having the lowest masses, followed by composites, Seyferts and LINERs respectively.

Looking at environment distributions, we see that there is virtually no trend; i.e. we find that local density has a relatively minor effect on emission line class. Whilst the differences are less significant, mass-matched emission line objects are generally found at the lower side of the density distribution. We find that star forming objects exist at the lowest local densities, followed by composites, Seyferts and LINERs respectively; however we emphasise that this is a relatively weak trend.

This is made abundantly clear in Fig. 3.20, which shows Gaussian models of the normalised distributions of each emission line class where the second parameters in each case have been controlled against the overall population. This shows us the relative location of each emission line class with respect to the overall distribution. We see an almost total absence of emission line dependence on environment when mass is controlled for, but a clear progression in mass.
Figure 3.19: The normalised distributions of log $M_*$ (left) and log $\rho$ (right) when controlled for $\rho$ and $M_*$ respectively, for all objects in black and different emission line classes in varying colours. Numbers at the top right show the means and standard deviations of each distribution. From top to bottom: star forming (blue), composites (purple), Seyferts (green) and LINERs (red).
That star forming galaxies are preferentially found at lower masses is a known phenomenon (e.g. Ferreras & Silk, 2000; Brinchmann & Ellis, 2000b; Peng et al., 2010) that is a core property of formation downsizing, which states that lower mass galaxies form their stars later. The presence of AGN in higher-mass galaxies is also well known (e.g. Bower et al., 2006). Bower et al. (2006) suggest that the presence of AGN in the most massive galaxies is partially what causes downsizing, as more massive galaxies are able to host AGN that can therefore use them to quench star formation, resulting in more massive galaxies having older stellar populations and therefore being less likely to be actively forming stars in the present day.

Schawinski et al. (2007) find a similar mass-dependence of emission line classes, and also link it to AGN feedback. They form a mass-dependent evolutionary sequence from star forming to nuclear activity to passive, using photometry and absorption indices to build stellar histories and track the most recent period of star formation. They theorise that star formation is quenched by AGN feedback, before the AGN itself runs out of fuel and the galaxy settles into quiescence, in a mass-dependent process. I extend this work by looking at environmental dependence, and find that our picture agrees with theirs, with environment having minimal importance.

This apparent dominance of mass over environment as a star formation driver is in agreement with other studies of star formation fractions (e.g. Thomas et al., 2010; Rogers et al., 2010). Thomas et al. (2010) find that an increase in the
fraction of galaxies with young stellar populations is observed at lower masses and densities, with mass being the driving factor, due to an increased fraction of ‘rejuvenated’ galaxies (objects which have had recent minor star formation events, i.e. those in the ‘blue cloud’) which are defined as having low ages and $\alpha/Fe$ and high metallicities. Rogers et al. (2010) also find that mass (using stellar velocity dispersion as a proxy) is a more important driver of star formation than local environment. They observe that the fraction of objects with recent star formation is lower in environments with high halo masses, and suggest that the (minor) decrease in star formation fraction at higher environments could be due to the higher efficiency of gas stripping in more massive haloes. It is worth noting that these different approaches all agree with the finding that star formation is driven by mass more than environment, whilst implementing entirely different methods. Thomas et al. (2010) look at stellar histories (i.e. the past of the galaxy) via Lick indices and Rogers et al. (2010) use Principal Component Analysis to determine spectral vectors to project observed spectra onto, again giving the star formation history of the galaxy, whilst we look at the current star formation in galaxies - a different parameter entirely. Despite this, all of our findings are consistent.

Interestingly, the mass around which Seyferts are significant and star forming objects become less significant in number is located at approximately the same mass as Kauffmann et al. (2003) and Ferreras et al. (2004) find a break in star formation efficiency. Kauffmann et al. (2003) investigated age-sensitive parameters and found a bimodality at $\log M_* \sim 10.5$, with lower mass objects indicating a younger population and higher mass objects an increased star formation efficiency. Ferreras et al. (2004) find a clear increase in star formation efficiency at stellar velocity dispersion values of $\sim 140$ km s$^{-1}$. Assuming that AGN quenching is a function of mass, and not environment, then we can combine all of the above into a bigger picture. In this paradigm, the more massive objects form their stars efficiently and quickly and are then quenched by AGN feedback, leaving behind old stellar populations with little current star formation.

This agrees with the quenching models of Peng et al. (2010), who introduce the idea of two separable methods to investigate the mass and environment star formation fraction changes, called mass quenching and environment quenching respectively, utilising galaxy colour as an indication of star formation. They suggest that mass quenching, which is independent of environment, could be
caused by AGN feedback, whilst environment quenching is caused by satellite quenching. They define environment quenching efficiency as

\[ \epsilon_\rho(\rho, \rho_0, m) = \frac{f_{\text{red}}(\rho, m) - f_{\text{red}}(\rho_0, m)}{f_{\text{blue}}(\rho_0, m)} \]

and mass quenching efficiency as

\[ \epsilon_m(m, m_0, \rho) = \frac{f_{\text{red}}(m, \rho) - f_{\text{red}}(m_0, \rho)}{f_{\text{blue}}(m_0, \rho)} \]

where \( m \) represents a mass measure, \( \rho \) environment, \( \rho_0 \) the lowest density environment, \( m_0 \) the lowest mass environment, and \( f_{\text{red}} \) and \( f_{\text{blue}} \) as the fraction of red and blue galaxies respectively at a given location in \( m - \rho \) space. Clearly this defines \( \epsilon_\rho \) as the fraction of galaxies at a given mass that are quenched with respect to objects at a reference environment \( \rho_0 \), and \( \epsilon_m \) as the fraction of galaxies at a given environment that are quenched with respect to objects at a reference mass \( m_0 \). I reproduce these methods, replacing ‘blue fraction’ with star forming fraction and ‘red fraction’ with passive fraction, and present my results in Fig. 3.21.

I find that mass quenching is completely independent of environment except at the highest densities, whilst environment quenching is somewhat dependent of mass. My finding of negligible environmental quenching is in contrast to the findings of Peng et al. (2010), who find a clear signal of a mass-independent quenching mode. I suggest this is due to mass quenching being the dominant process in our sample. This suggests that our data are predominantly quenched by secular processes such as supernovae and AGN feedback, and not by external factors such as satellite infall. The reasons for this are unclear at the present time, and warrant investigations in future work.

### 3.3.7 Non-/Group samples

Having established that local environment is not as significant a parameter for star formation quenching than mass, we turn our investigation to larger-scale environments. Looking at the effect of group membership we control for \( M_* \) and \( \rho \) when looking at \( \rho \) and \( M_* \) respectively, and perform KS tests on the different samples. I find a 8.6 \( \times \) 10\(^{-35}\)% probability that the group and non-group \( M_* \) samples were drawn from the same sample, and a 4.9 \( \times \) 10\(^{-316}\)% probability that the \( \rho \) samples were drawn from the same sample. Fig. 3.22 shows the \( M_* \)
Figure 3.21: Relative mass quenching efficiency (top) and environment quenching efficiency (bottom).
Figure 3.22: The distributions of $\log M_*$ (top) and $\log \rho$ (bottom) when controlling for $\rho$ (left) and $M_*$ (right), for group galaxies (black) and non-group galaxies (red). Means are marked as solid vertical lines, and standard deviations as dashed vertical lines.

distributions (top) and $\rho$ distributions (bottom) when comparing $M_*$ (left) and $\rho$ (right), where the group sample is represented by black and the non-group sample by red. Here we can see that even when controlling for $M_*$, the local density distributions are very different for group and non-group galaxies. We see that galaxies at a given mass tend to reside in denser environments if they are in a group than their non-group counterparts; this is an expected trend caused by the group definition preferentially selecting galaxies in denser environments. The very weak mass trend with group membership is a marginal signal at best; this by-and-large agrees with the findings of Vulcani et al. (2013), who note that the galaxy mass function does not change with global environment.

Fig. 3.23 shows the same as Fig. 3.20 but for our group sample (top) and non-group sample (bottom). The higher-peaked LINER model in our group sample is brought about by LINERs populating a narrower density range in the group sample, likely due to the low number of them present in the sample. We see that the
Figure 3.23: The same as Fig. 3.20 but for our group sample (top) and non-group sample (bottom).

The same class trends are observed in each sample, but with the group sample distributions located at higher environments and slightly higher mass. The differences in mass and environment distributions between the samples for each emission line class is consistent with the difference between the overall distributions for each sample. The primary results from this are that group membership does not influence the emission line class trends observed for the main sample, and that the negligible influence of local density on emission line classes is independent of group membership.

3.3.8 Conclusions

I have extracted emission line statistics from GAMA survey galaxies, and used these to ascertain the ionisation source within these objects via the BPT diagnostic diagram. I have then used stellar mass estimates, a measure of local density
that corrects for the finger of god effect, and a friends-of-friends group catalogue to investigate the dependence of these ionising sources on mass and environment. When I look at mass, I control for local density, and vice versa; this is in order to negate the dependence of the mass function on local environment.

I find that there is a significant mass sequence for emission line classes from star forming objects at the low end, through composites and Seyferts, to LINERs, in agreement with the findings of other works. Schawinski et al. (2007) found a similar result, and use stellar population analysis to determine that this sequence is a progression in time, from star forming to passive via AGN activity. I extend this picture by investigating the effects of local and global environment on this trend. I find that there is no significant difference in local density distribution between different emission line classes when we have controlled for mass. This tells us that local environment has minimal impact on the progression previously observed. If this progression is due to AGN feedback, as suggested in Schawinski et al. (2007), then this would imply that AGN feedback is a primarily mass-driven process.

I further find that this independence of emission line class from local density is not affected by group membership, nor are the mass distributions of galaxies with different ionisation sources. This results in us finding a scale-based trend in the importance of a property on ionisation source, with the most important parameter being mass, followed by the much less significant influence of local density, and then the absolutely negligible influence of global environment.

3.4 Further science based on our pipeline data

Data products from our pipeline have been used for science in six works to date (submitted or close to submission, including my own), and as a test of robustness in one other publication. A brief summary of these papers is included below, along with a small description of my involvement with any of them and where our results were used.

3.4.1 Ching et al. (prep)

Galaxy And Mass Assembly (GAMA): The environments of high- and low- excitation radio galaxies.
Introduction and data

Whilst some radio galaxies are observed to have emission lines due to the presence of a radiatively efficient accretion disk surrounding the central supermassive black hole (high-excitation), other radio galaxies are observed not to have these lines (low-excitation). It is believed that high-excitation objects are able to form radiatively efficient disks due to the accretion of cold gas (cold mode), whilst low-excitation objects are unable to due to accreting hot gas (hot-mode) which is not efficient enough to form a disk (Hardcastle et al., 2007).

Models predict that these two types of radio galaxies suppress star formation in different ways, rendering their understanding very important for the completeness of galaxy formation and evolution theory (Croton et al., 2006). This work investigates the environmental dependence of these two populations down to much smaller environments than previous studies.

Radio sources were selected by cross-matching the Faint Images of the Radio Sky at Twenty-cm (FIRST Becker et al., 1994) catalogue with the photometric catalogue of the SDSS DR6 (Adelman-McCarthy et al., 2008) in all GAMA regions. Emission line statistics were obtained from the MPA-JHU catalogue for SDSS DR7 (Abazajian et al., 2009) spectra in the GAMA regions, and from gandalf in the GAMA-observed spectra.

Low- and high-excitation galaxy classification

Before applying a quantitative diagnostic cut between different galaxy types, a visual inspection was made to separate out broad-line AGN as they may have significant contamination from the non-thermal emission of the AGN. The BPT line diagnostic was then applied to remove galaxies dominated by star formation, requiring all four BPT diagnostic lines with AoN > 3 and the object to fall in the star forming zone for it to be considered star formation dominated. To remove any star formation-dominated objects that avoided this test, the SFR estimates from Hα flux are compared with those estimated from the total FIRST radio flux using the relation from Hopkins et al. (2003), and any galaxies where the SFR is within 3σ of the one-to-one relation were removed from the sample. The remaining galaxies were deemed to be a robust sample of radio-loud AGN.

In order to separate low- and high-excitation radio galaxies, a simple cut in [OIII]λ5007 EW was made. High-excitation radio galaxies (HERGs) were classified as those with AoN([OIII]) > 3 and EW([OIII]) > 5Å, and all other
galaxies were considered low-excitation radio galaxies (LERGs). This cut was
determined by comparing $\text{EW}(\text{OIII})$ with the previous visual classification.

The final sample contained 508 LERGs and 89 HERGs.

**Measurements & findings**

This work uses three primary diagnostic tools: stellar mass, local density, and
global (group) environment. The stellar mass is taken from estimates by [Taylor et al. (2011)](Taylor) which fits SED models to GAMA photometry in order to form this estimate. Local density was measured as the fifth nearest neighbour density, a
projected surface density estimate for galaxies using the formula $\Sigma_5 = 5/(\pi r_5^2)$
where $r_5$ is the projected distance of the galaxy to the fifth nearest neighbour. A
limit of $z < 0.18$ was applied in order to create a volume limited sample. Group
environments were measured in the GAMA Group Catalogue ([Robotham et al. 2011](Robotham)]).

Fig. 3.24 shows $\Sigma_5$ against stellar mass for the whole GAMA sample (grey and contours), LERGs (red) and HERGs (blue). It can be seen from this that there is a relative excess of LERGs in high density environments when compared to both HERGs and the main GAMA sample, and that HERGs and LERGs span a relatively wide range of $\Sigma_5$. Also it is clear that the radio galaxy sample is dominated by galaxies at the higher mass end of the GAMA sample, with some scatter towards the lower end, and that LERGs occupy a higher mass space than HERGs.

It was found to $>99\%$ probability that the LERG population was drawn from a different $\Sigma_5$ distribution to the main sample, but only to 7.8% for the HERGs. This means that for radio-loud AGN with no emission lines, local density plays a role beyond the usual density-stellar mass relationship found for all galaxies, whilst this was not the case for those with emission lines.

Looking at group properties, this work finds that LERGs do not inhabit higher mass haloes than non-radio objects, which is inconsistent with previous studies and also with the higher density environments found above. The reason for this is not found. HERGs are also found to have no group property dependences.

**My contribution**

This work used the BPT diagnostic classifications, $H\alpha$ EWs and $[\text{OIII}]\lambda5007$ EWs
from my catalogue run over v17 of the GAMA spectra. The BPT diagnostic and
Figure 3.24: The local galaxy density \( \langle \Sigma_5 \rangle \) as a function of stellar mass in the full redshift range of the density measurements for the GAMA sample (black) and radio-loud AGN of different classes, LERG (red) and HERG (blue). The black contours replace regions where the number density of GAMA points are too high. Crosses represent objects with spurious stellar mass estimates. *Image provided via personal correspondence with J. Ching; caption taken Figure 3, Ching et al. (prep).*
Ha EWs were used to select star forming objects, whilst the [OIII] EWs were used to separate high- and low-excitation objects.

### 3.4.2 Lara-López et al. (2013)

**Galaxy And Mass Assembly (GAMA): the connection between metals, specific SFR and H I gas in galaxies: the Z SSFR relation.**

**Introduction and data**

This work investigates the relationships between metallicity (Z), specific star formation rate (sSFR) and neutral gas (HI) content, in order to produce a general picture of the gas recycling process. The optical data used are from two surveys: the GAMA survey and the SDSS DR7 (Abazajian et al., 2009), and were used to determine SFRs, stellar mass and metallicities. HI data also came from two surveys: the ALFAFA survey (Haynes et al., 2011) and the GASS survey (Catinella et al., 2010).

From GAMA, SFRs are based on the method described in Gunawardhana et al. (2011) using the Hα emission line, metallicities are estimated using the empirical calibration of Pettini & Pagel (2004) from the oxygen abundance and the O3N2 ( ([OIII]λ5007/Hβ)/(NII)λ6583/Hα ) index, and both are recalibrated to the Bayesian system using the method of Lara-López et al. (2013). The stellar mass measurements are described in Taylor et al. (2011). From the SDSS, metallicity measurements described in Tremonti et al. (2004), SFR estimates described in Brinchmann et al. (2004) and stellar mass estimates from Kauffmann et al. (2003) are used.

From both optical surveys, only star forming (SF) galaxies were selected, as determined by the BPT diagnostic and using the discrimination of Kauffmann et al. (2003), leaving 35,212 GAMA objects and 156,910 from SDSS.

The ALFAFA and GASS surveys provide HI determinations for 4,491 of these objects, all of which came from the SDSS due to survey overlap areas.

**The Z-sSFR relation**

Fig. 3.25 shows the Z-sSFR relation for SDSS and GAMA data. Each panel corresponds to a volume-limited sample of a different redshift, as described in Lara-López et al. (2013). The coloured circles correspond to the median sSFR
in bins of $Z$ for different $M_\ast$. Panels (a) and (b) clearly display opposing trends between low- and high-mass galaxies, with an inflection point at $\log(\text{sSFR}) \sim -9.9$ yr$^{-1}$, where low-mass galaxies show a decrease in sSFR with $Z$ whilst high-mass show the opposite.

The suggested reason for this, working on the assumption that galaxies of at a given stellar mass can have different HI abundances, is downsizing. The left panel of Fig. 3.26 shows a cartoon model explaining the situation, where low mass galaxies are represented by a blue ellipse, and high mass galaxies by a red ellipse.

For low mass objects the differences in sSFR can be explained through differing amounts of HI, as low mass galaxies with high HI abundances will show a higher sSFR than those of the same mass with low HI abundances as they have more fuel for star formation. This is cancelled out in higher mass galaxies by the effect of metallicity however, as massive galaxies with large amounts of HI are shown to have a higher $Z$ than their lower mass counterparts. This is again a product of downsizing, as $Z$ is being driven by the amount of HI. Low mass galaxies with large amount of HI have a lower metallicity as they are processing and enriching their gas slower than their lower HI counterparts with the same mass, whilst high mass galaxies with large HI abundances have higher metallicities as they processed and enriched their gas faster in the past, and as such already have high $Z$, and also have a high sSFR due to the large amount of HI fuel available.
The model is tested using the 4,491 galaxies with optical and HI measurements, using the definitions that gas mass $M_{\text{gas}} = 1.32xM_{\text{HI}}$ and the gas mass fraction is $M_{\text{gas}}/(M_* + M_{\text{gas}})$. The right panel of Fig. 3.26 shows the Z-sSFR relation for this sample where colour corresponds to the log gas mass fraction, which only contains the low mass branch of this relationship. It is clear from this that the gas mass fraction increases as sSFR increases and metallicity decreases, as predicted by the cartoon model. The black triangles in this figure represent SDSS galaxies in the ALFAFA and GASS fields with no HI detection in the ALFAFA and GASS fields for $\log(M_*) > 11.0$ dex. *Image and caption taken Figure 2, Lara-López et al. (2013a).*

**HI scaling relations**

This work presents the primary finding that neutral gas content is a driving factor behind the interplay between metallicity and sSFR at a given stellar mass. Low mass galaxies with high neutral gas contents will have high sSFRs and low metallicities, whilst those with low neutral gas contents will have lower sSFRs and higher metallicities. In contrast, high mass galaxies with high neutral gas contents will have moderate sSFRs and high metallicities, whilst those with small
amounts of gas will have low sSFRs and low metallicities.

This finding is important in the context of galaxy evolution as it suggests that the HI content is an important driver of this evolution, not just stellar mass.

**My contribution**

I provided the H$\alpha$ measurements used to determine the star forming population and their star formation rate, and the metallicities used throughout the work. The metallicity values were calculated specifically for this project and the work on these relations, and all values were taken from v17 of the GAMA data.

### 3.4.3 Lara-López et al. (2013)

**Galaxy And Mass Assembly (GAMA): a deeper view of the mass, metallicity and SFR relationships.**

**Introduction and data**

It is known that the stellar mass ($M_*$), SFR and gas metallicity ($Z$) are related through the mass-metallicity and mass–SFR relationships, but there is no strong correlation between the metallicity and SFR alone. A fundamental plane has, however, been found between the three (Lara-López et al., 2010). This work investigates the $M_*$–SFR–$Z$ plane using emission line data from the spectra of GAMA and SDSS DR7, using the MPA-JHU catalogue for SDSS DR7 objects and a combination of our GANDALF catalogue and a simple Gaussian fit approach for GAMA objects.

Only SF emission line galaxies were considered, which were defined as those with all four BPT lines detected with a SNR of over 3$\sigma$ for [NII], H$\alpha$ and H$\beta$, as well as line ratios that placed them in the SF region of the diagnostic diagram. Metallicities were estimated by the method of Tremonti et al. (2004), using the four BPT lines. SFRs were calculated using the method of Brinchmann et al. (2004) for SDSS and Hopkins et al. (2003) for GAMA, using the H$\alpha$ EW, and the GAMA measurements were calibrated to work in the same system as the SDSS estimates. Masses came from Taylor et al. (2011).
Findings

Over the redshift range $0.330 < z < 0.365$, a maximum evolution for metallicity of $\sim 0.1$ dex was found, as was a maximum SFR evolution of $\sim 0.4$ dex and sSFR evolution of $\sim 0.56$ dex.

The fundamental plane found in this work is shown in Fig. 3.27 which shows a projection of the 3D distribution formed by $M_\star$, log(SFR) and 12+log(O/H) ($Z$) for the sample. The orange plane shows the fundamental plane described in equation 3.9. The vertical axis always shoes $M_\star$, and the cube is rotated clockwise from the upper-left to the bottom-right panel, with the final panel showing the edge-on projection. The grey dots are those points above the plane, and black those below.

The equation for the fundamental plane is calculated as

$$log(M_\star/M_{\odot}) = \alpha[12 + log(O/H)] + \beta[log(SFR)] + \gamma$$ (3.9)

where $\alpha = 1.3764 \pm 0.006$, $\beta = 0.6073 \pm 0.002$ and $\gamma = -2.5499 \pm 0.058$. This relation is found to recover the $M_\star$ of the entire sample with $\sigma = 0.2$ dex. This exists as the current mass of stars in a galaxy is a measure of the gas currently being converted into stars (the SFR) and the star formation history (here represented by $Z$). A lack of evolution of the plane was found to $z < 0.365$, as a consequence of SFR and $Z$ evolving in different directions.

Finally, a bimodality was found between high- and low-mass galaxies when looking at metallicities and SFR at a fixed stellar mass. For massive galaxies at fixed stellar mass, the median metallicity was found to be higher/lower for high/low SFR galaxies, whereas for low mass galaxies the median metallicity was found to be lower/higher for high/low SFR galaxies. This is attributed to downsizing and differing amounts of neutral gas. More massive galaxies process their gas quickly, so a galaxy with a larger amount of neutral gas will use this supply to form more stars and further enrich its metallicity than a galaxy with less neutral gas. Conversely, for less massive galaxies that process their gas slowly, higher metallicities can be explained by bursty star formation history that exhausted the galaxy’s gas and increased its metallicity - hence galaxies with lower sSFRs having higher metallicities now, as they exhausted their gas in the past.
Figure 3.27: Projections of the 3D distribution formed by $M_*$, log(SFR) and 12+log(O/H) (Z) for GAMA and SDSS galaxies. The orange plane shows the fundamental plane described in Eq. 3.9. The vertical axis shows $M_*$ in all panels. The cube is rotated clockwise from the upper-left to bottom-right panel. The last panel shows the edge-on projection of the derived fundamental plane. Grey and black dots show galaxies above and below the plane, respectively. Image and caption taken Figure 10, Lara-López et al. (2013).
My contribution

I provided the BPT diagnostic classifications, Hα AoNs, [NII] AoNs and Hβ AoNs used to determine which objects were star forming. The Hα, Hβ, [NII] and [OIII] EWs used to calculate metallicities were taken from my catalogue, and my Hα measurements were further used to determine star formation rates. All of these were taken from v17 of the GAMA data.

3.4.4 Hardcastle et al. (2013)

Herschel-ATLAS/GAMA: a difference between star formation rates in strong-line and weak-line radio galaxies.

Introduction and data

There are considered two types of radio-loud galaxy: those with emission lines (high-excitation radio galaxies, HERGs) and those without (low-excitation radio galaxies, LERGs). LERGs show no evidence of radiatively efficient AGN outside of what would be expected from a nuclear jet, whilst HERGs display like AGN with additional radio jets and lobes. Models predict that these two types of radio galaxies suppress star formation in different ways, rendering their understanding very important for the completeness of galaxy formation and evolution theory. This work investigates the difference in SFR between these two subclasses.

A sample of radio galaxies was made by cross-matching the FIRST (Becker et al., 1994) radio catalogue with optical sources from SDSS DR6 (Adelman-McCarthy et al., 2008) in all GAMA regions. Emission line statistics were obtained from the MPA-JHU catalogue for SDSS DR7 spectra (Abazajian et al., 2009) in the GAMA regions, and from GANDALF in the GAMA-observed spectra. FIR flux densities were obtained from the H-ATLAS survey (Eales et al., 2010).

Low- and high-excitation galaxy classification

Before applying a quantitative diagnostic cut between different galaxy types, a visual inspection was made to separate out broad-line AGN as they may have significant contamination from the non-thermal emission of the AGN. The BPT line diagnostic was then applied to remove galaxies dominated by star formation, requiring all four BPT diagnostic lines with AoN > 3 and the object to fall in the star forming zone for it to be considered star formation dominated. To remove
any star formation-dominated objects that avoided this test, the SFR estimates from Hα flux are compared with those estimated from the total FIRST radio flux using the relation from Hopkins et al. (2003), and any galaxies where the SFR is within $3\sigma$ of the one-to-one relation were removed from the sample. The remaining galaxies were deemed to be a robust sample of radio-loud AGN.

In order to separate low- and high-excitation radio galaxies, a simple cut in [OIII]λ5007 EW was made. High-excitation radio galaxies (HERGs) were classified as those with AoN([OIII]) > 3 and EW([OIII]) > 5Å, and all other galaxies were considered low-excitation radio galaxies (LERGs). This cut was determined by comparing EW([OIII]) with the previous visual classification.

**Findings**

It is found that rest-frame 250−μm luminosities are systematically higher for HERGs than LERGs for all radio luminosities sampled, and further that this measurement is higher for HERGs than normal galaxies at matched absolute magnitude whilst being lower for LERGs than normal galaxies at matched absolute magnitude.

Dust masses are found to be comparable between the two samples, but HERGs are found to have higher temperatures. It is argued that this provides strong evidence for HERGs having higher SFRs on average, and the higher HERG FIR luminosities are indications of this, although this is a statistical relationship only and not a one-to-one relation between AGN type and SFR.

**My contribution**

This work used the BPT diagnostic classifications, Hα EWs and [OIII]λ5007 EWs from my catalogue run over v17 of the GAMA spectra. The BPT diagnostic and Hα EWs were used to select star forming objects, whilst the [OIII] EWs were used to separate high- and low-excitation objects.

**3.4.5 Foster et al. (2012)**

Galaxy And Mass Assembly (GAMA): The mass-metallicity relationship.
CHAPTER 3. THE GAMA SURVEY

Introduction and data

Whilst the presence of a mass-metallicity relationship (MMR) describing a correlation between stellar mass and gas-phase metallicity is well-established, the exact shape and its dependence on other variables is still unknown. This is partially because the precise cause of this relationship is unknown, and likely a product of many different factors such as outflow dependence on mass, the interplay between outflows and inflows, star formation rate (SFR), etc.

The aim of this work is to measure the MMR in the GAMA survey, compare this measure to that of the SDSS, and study its dependence on selection criteria. GAMA galaxies are separated into AGN, composites and star forming objects via the BPT diagnostic diagram. SFRs are calculated from the Hα flux, masses from the data product introduced above, and metallicities using three different calibrations ([Kewley & Dopita 2002; Pettini & Pagel 2004; Kobulnicky & Kewley 2004]) that require measures of the EWs of [NII]λ6583, the [OII] doublet, [OIII]λ5007, Hα and Hβ variously. The effect of different selection methods, such as requiring different lines to have a given AoN, is also investigated.

Findings

Fig. 3.28 shows the MMR for the GAMA sample with different selection criteria, where red contours represent the results of the selection criteria used in this work and black contours represent the results obtained using the selection criteria of [Brinchmann et al. 2004; Tremonti et al. 2004; Kewley & Ellison 2008; Lara-López et al. 2010; Kobulnicky & Kewley 2004; Mannucci et al. 2010] respectively. From this it is clear that the choice of selection criteria strongly influences the shape of the MMR.

It is also found that the metallicity calibration used has a significant impact on the shape and position of the MMR, but that when a robust calibration is used, the MMR of GAMA objects is in reasonable agreement with that of SDSS objects despite the difference in the luminosity ranges (and therefore masses) probed. Selecting based on the [OIII]λ5007 line is seen to have the biggest impact, and it is also noted that for monotonic metallicity calibrations such as that of Pettini & Pagel (2004) a selection based on metallicity uncertainties can be applied without biasing the sample.
Figure 3.28: GAMA MMR measured using a variety of selection criteria taken mostly from the recent literature, as labelled, for the PP04 abundance estimate. The MMR obtained using the selection criteria of Brinchmann et al. (2004); Tremonti et al. (2004); Kewley & Ellison (2008); Lara-López et al. (2010); Kobulnicky & Kewley (2004); Mannucci et al. (2010) (respective black 1, 2 and 3σ contours) are compared to that obtained using the selection criteria used in this work (red 1, 2 and 3σ contours). The MMR measured varies significantly (> 0.05 dex) if one selects on the [OIII]λ5007 line. Image and caption taken Figure 4, Foster et al. (2013).
My contribution

My catalogue provided the H\(\alpha\) fluxes and EWs, as well as the H\(\beta\), [NII], [OII] doublet and [OIII] EWs. These were used to determine the emission line classifications of objects, as well as calculate their metallicities and star formation rates. This work was done using v17 of the GAMA data.
Chapter 4

SDSS-III/BOSS

BOSS is the largest of the SDSS-III collaboration’s four surveys (Eisenstein et al., 2011), utilising an upgrade of the multi-object spectograph on the 2.5m SDSS telescope (Gunn et al., 2006) to collect galaxy and quasar spectra over 10,060 deg$^2$ of sky. Its primary aim is to use the baryon acoustic oscillation (BAO) feature in large-scale structure as a standard ruler to measure cosmological distances, in order to determine the expansion history of the Universe to a high precision (Dawson et al., 2013). It is designed to measure the BAO scale in the clustering of matter over a volume larger than all previous spectroscopic large scale structure surveys combined. BOSS consists of two separate surveys simultaneously covering the same area with the same objective, one targeting galaxies and the other quasars. We are only interested in the galaxy sample, however, and so only include this survey in our discussions and considerations throughout. In order to achieve the primary aim, the final survey is expected to obtain $\sim 1,500,000$ galaxy spectra out to redshifts of $z \sim 0.8$.

We aim to use the data from this survey to investigate the mass- and redshift-dependence of star formation histories, in order to constrain the archaeological downsizing phenomenon at high lookback times. For this I stack objects together and perform stellar population analysis on the resulting absorption line spectra. We also want to test the robustness of the BOSS colour-colour target selection cuts against the presence of emission line activity, specifically star formation. With this in mind I have constructed a pipeline to analyse the emission line statistics, stellar kinematics and absorption statistics of BOSS galaxies for both data releases, which has been used already in a science paper as detailed in Section 4.3.
To date there have been two public releases of the BOSS data, SDSS-III DR9 and DR10, containing 535,995 and 927,844 galaxy spectra respectively. The BOSS footprint is shown in Fig. 4.1 for both (currently) publicly available data releases, as well as for the next public release of DR11, although the results of DR11 are beyond the scope of this work. The top panels show observations in the North Galactic Cap, whilst lower panels show those in the South Galactic Cap. Colours represent the spectroscopic completeness within each circle, as indicated in the key in the lower right panel. The grey area shows the expected BOSS footprint by the completion of the survey. The total sky coverage in DR9, DR10 and DR11 is 3,275 deg$^2$, 6,161 deg$^2$ and 8,377 deg$^2$ respectively (Anderson et al., 2013).

Fig. 4.2 shows a projected light cone for galaxies included in DR9 (white), the SDSS DR7 main galaxy sample (yellow) and the SDSS DR7 LRG sample (red), clearly showing the significant increase in depth obtained by BOSS over previous SDSS projects.

Galaxy target selection is based on the colour and magnitude evolution of a passively evolving model by Maraston et al. (2009) and previous selections for DR7 (Eisenstein et al., 2001); for details see Eisenstein et al. (2011). We provide a brief summary of the selection details here. The model magnitude
Figure 4.2: Projected light cone of SDSS DR7 main galaxies (yellow), SDSS DR7 LRGs (red) and BOSS galaxies included in DR9 (white) (A. Ross, private correspondence).
colour combinations used for target selection are given in Eq. (4.1) \cite{eisenstein2011}. $c_{\text{perp}}$ and $c_{\parallel}$ are designed to select galaxies below $z \sim 0.4$, and $d_{\perp}$ at higher redshift.

$$c_{\parallel} = 0.7(g - r) + 1.2(r - i) - 0.18$$
$$c_{\perp} = (r - i) - (g - r)/4 - 0.18$$
$$d_{\perp} = (r - i) - (g - r)/8$$

(4.1)

Within the galaxy survey of BOSS there are again two separate sub-samples: ‘LOWZ’ and ‘CMASS’ (‘Constant MASS’). The LOWZ selection aims to select similar luminous red galaxy samples as in SDSS-I/II but fainter, therefore improving number density. CMASS selection aims to achieve a constant mass sample as a function of redshift, which is well achieved as shown in \cite{maraston2013}. Taking the colour combinations defined above, the LOWZ sample is defined as having $16 < r < 19.5$, $r < 13.6 + c_{\parallel}/0.3$ and $|c_{\perp}| < 0.2$, where $r$ is the cmodel magnitude in the SDSS imaging catalogue. The CMASS sample is defined by $17.5 < i < 19.9$, $d_{\text{perp}} > 0.55$ and $i < 19.86 + 1.6 \times (d_{\text{perp}} - 0.8)$ where $i$ again is a cmodel magnitude. This selection gives rise to the LOWZ sample targeting the redshift interval $0.15 < z < 0.43$, and CMASS the interval $0.43 < z < 0.7$, although in practise both samples extend beyond these boundaries. Further analysis of this target selection procedure is given in Section 4.1.3. Both of these samples are expected to be 98.8% complete within their limits by the completion of the survey.

Spectroscopic observations were made using a multi-object spectrograph mounted on the SDSS dedicated wide-field 2.5 m telescope \cite{gunn2006} at Apache Point Observatory (APO, Sunspot, NM, USA). The spectra cover the wavelength range $3600 \text{ Å} < \lambda < 10,000\text{ Å}$ with a resolution of $R \sim 1300$ at $\lambda \sim 3600\text{ Å}$ and $R \sim 3000$ at $\lambda \sim 10,000\text{ Å}$, and with signal to noise values of $\sim 10$ per pixel at $g = 22$ in the blue cameras and $\sim 22$ per pixel at $i = 21$ in the red cameras.

Spectra were processed and reduced the day after being observed, and the full details of this procedure can be found in \cite{dawson2013}. A summary is presented here.

Data processing is performed by a software known as IDLSPEC2D. This software first operates on the raw 2D CCD frames, performing bias subtraction - the
removal of the small signal inherently caused by the CCD electronics; flat-fielding - corrections for any pixel-to-pixel differences both in the dark and light frames; and fibre trace fitting - the determination of which flux came from which fibre. 1D spectra are then extracted from the 2D frames. The spectra are then wavelength calibrated, the process of assigning wavelengths to each pixel; sky subtracted, the removal of atmospheric interference; and flux calibrated, the correction for varying spectral responses as determined using models fit to the spectra of standard stars. Since these processes are very similar (and in some cases identical) to those applied to the GAMA data, further detail can be found in Section 3. Finally the spectra are resampled onto a log(\(\lambda\)) grid and spectra from the red and blue cameras are combined.

The resultant spectra are classified into specific object types and have their redshifts measured (Aihara et al., 2011; Bolton et al., 2012). Redshifts are measured by first fitting star, quasar and galaxy templates to the spectra, and then re-fitting targeted galaxies with only stellar and galaxy templates. Classifications are based on the constituent of the best-fit template. Errors detected in the fitting process are reported by a bitmask keyword, and it is found that only \(\sim 0.2\%\) of confidently reported CMASS sample redshifts and classifications are incorrect.

Fig. 4.3 shows the redshift distributions of BOSS DR10 (black), DR9 (blue) and SDSS-I DR7 (green). We compare against DR7 as this is considered the benchmark for large scale spectroscopic surveys. The mean redshift of all BOSS DR10 objects is \(\bar{z} = 0.48 \pm 0.15\), of BOSS DR9 objects is \(\bar{z} = 0.49 \pm 0.15\) and of SDSS DR7 is \(\bar{z} = 0.14 \pm 0.10\). This means that BOSS DR10 probes a mean factor of 2.8 deeper in time than DR7.

### 4.1 Spectroscopic properties of BOSS DR9 galaxies

Here I perform a spectroscopic analysis of 492,350 galaxy spectra from SDSS DR9, the ninth data release of the SDSS collaboration, which is the first BOSS public data release and contains spectra from the first two years of BOSS observations. I do this using the code detailed in Sections 2.1 and 2.2, and Fig. 4.4 shows example fits for some typical BOSS galaxies. Despite the fact that the typical signal-to-noise ratio of BOSS galaxies is only \(\sim 5\ \text{Å}^{-1}\), we show in Section 4.1.1 that it is adequate to measure stellar velocity dispersions and emission line fluxes.
Figure 4.3: The redshift distributions of all spectroscopically observed galaxies from the SDSS DR7 with ‘zwarning’ equal to 0 (green), and from the BOSS survey with ‘zwarning’ equal to 0 for DR9 (blue) and DR10 (black).
Figure 4.4: Example spectra with and without emission lines for BOSS galaxies around \( z \sim 0.3 \) in observed-frame wavelength with the best fit spectrum composed of the stellar population and the emission line templates overplotted (red line). The bottom spectrum is the variance, the residual from the fit is plotted in the bottom sub-panel.

for individual objects. The typical stellar velocity dispersion of a BOSS galaxy is \( \sim 240 \text{ km/s} \), with an error of \( \sim 14\% \). We find that 93\% of BOSS galaxies have an error of below 30\% on stellar velocity dispersion, and also that the stellar velocity dispersion distribution is redshift independent in the range \( 0.15 < z < 0.7 \). This redshift independence is inherently caused by the survey’s targeting of massive galaxies in this range with a near-uniform mass distribution. It should be noted that despite the fact that at wavelengths blueward of \( \sim 4,500 \text{ Å} \) our MILES-based stellar population templates have a slightly worse spectral resolution than our data, which would affect our velocity dispersion measurements, our values are reliable as we consider only the region \( 4,500 - 6,500 \text{ Å} \) when calculating stellar kinematics (see Section 2.2.3 for details on why this region was chosen).

Looking at emission lines, I show in Section 4.1.2 that although the BOSS resolution is sufficient to measure them, only around 4\% of BOSS galaxies have detections for the common lines \([\text{OII}]\lambda\lambda3726 + 3729, \text{H}\beta, [\text{OIII}]\lambda5007, \text{H}\alpha\) and \([\text{NII}]\lambda6583\). This is again to be expected, as massive galaxies are generally not as active than their less massive counterparts. I analyse these five lines in diagnostic diagrams in order to separate out star forming objects and AGN. I find that emission lines properties in the BOSS sample are strongly redshift dependent,
Figure 4.5: Distributions of the relative errors in the velocity dispersion measurements of BOSS galaxies (left) and the distribution of velocity dispersions with error smaller than 30% (right).

and that there is a clear correlation between observed frame colours and emission line properties. The main redshift split is between the LOWZ sample in the range $0.15 < z < 0.3$, where around half of the objects with detectable emission lines have LINER-like emission line ratios and most of the rest have Seyfert-AGN dominated spectra, with only a small scattering of purely star forming galaxies, and the CMASS sample in the range $0.4 < z < 0.7$, where more than half of the emission line objects are star forming. This is a pure selection effect caused by the weak H$\beta$ emission lines in the BOSS spectra. Finally, in Section 4.1.1 I display the clear separation of star forming, AGN and emission line free galaxies in the $g-r$ vs $r-i$ target selection diagram.

### 4.1.1 Stellar velocity dispersions

Even though the typical signal-to-noise ratio of BOSS galaxy spectra is considerably lower than that provided by SDSS-I spectroscopy due to BOSS targets being relatively fainter, I find that measurements of their line-of-sight stellar velocity dispersions can still be made to a reasonable accuracy. Fig. 4.5 shows the distributions of formal measurement errors (left) and measured stellar velocity dispersions for objects with formal uncertainties of under 30% (right). I see that the typical relative error is $\sim 14\%$, and that 93% of objects have a formal measurement uncertainty of below 30%. Although these measurements are clearly quite large, I find through comparison with independent measurements that they are robust (see below). I find that the distribution of stellar velocity dispersion
Comparison with other measurements

I test the robustness of these measurements by comparing them to two independent measurements: those of [Bolton et al. 2012] based on the original SDSS pipeline, and those of [Chen et al. 2012] based on principal component analysis fitting of [Bruzual & Charlot 2003] models. Fig. 4.6 shows the comparison between our measurements and those of the two independent measurements. I find that our measurements are $\sim 4\%$ higher than those of [Bolton et al. 2012].
and $\sim 7\%$ lower than those of Chen et al. (2012), with dispersions of 19% and 16% respectively. Both of these offsets are below the typical measurement error of 14%, suggesting that velocity dispersion measurements can be made reliable and robustly on BOSS spectra despite their low S/N levels.

**Comparison with higher S/N spectra**

I further test the robustness of our measurements by investigating the dependence on S/N ratio. In order to do this I made use of repeated observations and constructed a sample of 574 BOSS galaxies with at least six 1-hour observations. I then stacked the individual spectra of these objects to create a single, high-S/N spectrum for each object. The top panel of Fig. 4.7 shows the individual spectra and final stack for one BOSS object, demonstrating the improvement in S/N from $\sim 5 \text{ pix}^{-1}$ to $\sim 20 \text{ pix}^{-1}$, a factor of approximately 4.

The bottom panel of Fig. 4.7 shows the ratio of the measured stellar velocity dispersions of the stacks and the individual spectra for all 574 objects as a function of the measurements on individual spectra. We see a generally good agreement with no systematic offset when comparing measurements on low S/N and high S/N spectra, and in fact found there to be no trend with S/N ratio. We consider this to be a good argument that our measured velocity dispersions are robust, and are as such reliable within their errors despite the low S/Ns of individual spectra.

**Distributions with redshift**

Fig. 4.8 shows the distributions of stellar velocity dispersions with an error of under 30% in different redshift bins. The redshift intervals shown here are $0 - 0.15$, $0.15 - 0.3$, $0.3 - 0.45$, $0.45 - 0.6$, $0.6 - 0.75$ and $> 0.75$, and the red lines represent the mean position of the peaks of the four central redshift intervals. Clearly the distribution is largely independent of redshift within the range $0.15 < z < 0.75$, with the peak position staying largely constant whilst the width increases with redshift due to slightly larger measurement errors. This independence is not displayed for $z < 0.15$ or $z > 0.75$ however, with the distributions skewed and slightly offset towards lower and higher velocity dispersions respectively.

This uniform behaviour is consistent with the BOSS target selection objective of selecting objects with reasonably constant mass across the redshift range $0.15 < z < 0.7$. This is a very useful feature of the BOSS galaxy sample, as it establishes
Figure 4.7: Top panel: Individual spectra and their stack of a single BOSS galaxy with repeated 1-hour observations. The S/N ratio per resolution element is given by the labels. The S/N ratio of the stacked spectrum is higher by about a factor four compared to the individual spectra. Bottom panel: Median velocity dispersion measured on individual BOSS spectra versus the ratio between the median of measurements on individual and stacked spectra. Each stacked spectrum is the sum of individual spectra of the same object from repeated plate observations. The blue symbols are the median in bins of velocity dispersion, error bars indicate Poisson errors. The agreement is very good, hence velocity dispersion measurements are reliable even at typical BOSS S/N.
Figure 4.8: Distribution in velocity dispersion for spectra with an error of less than 30 per cent in the redshift intervals $0 < z \leq 0.15$, $0.15 < z \leq 0.3$, $0.3 < z \leq 0.45$, $0.45 < z \leq 0.6$, $0.6 < z \leq 0.75$, and $z > 0.75$. The red line indicates the position of the peak for the four central redshift intervals. The distribution is approximately Gaussian and there is no evolution within $0.15 < z < 0.75$. At redshifts below $z \sim 0.15$ and above $z \sim 0.75$, the distributions are skewed and slightly offset toward lower and higher velocity dispersions, respectively.
a direct link between the galaxy samples at various redshifts and allows for an accurate probe of the redshift evolution of massive galaxies. Our findings on the success of this target selection method are similar to the conclusions drawn in Maraston et al. (2013), who show that the distribution of photometric stellar masses is approximately uniform up to $z \sim 0.6$.

### 4.1.2 Emission lines

Emission line statistics are measured as described in Chapter 2. We assume an emission line is measured if the amplitude-over-noise (AoN) ratio is higher than 2, implying that a reasonable line strength is required for detections to be made. We note that line fluxes of the OII doublet [OII] $\lambda\lambda 3726 + 3729$ are calculated as the sum of the individual lines [OII] $\lambda 3726$ and [OII] $\lambda 3729$. As discussed in Chapter 2, we leave the emission line velocities and velocity dispersions as free parameters despite the low S/N of the observed spectra, due to the redshifts investigated in the survey resulting in many of the stronger lines being lost outside of the observed wavelengths. We follow this approach regardless of redshift in order to guarantee homogeneity within the full sample.

The aim of our emission line analysis is to investigate the prevalence of various galaxy types and emission line classes in the BOSS target selection algorithm. We split our work into two parts. In Section 4.1.2, we look at galaxies with $z < 0.45$ in order to have the full set of key diagnostic emission lines ([OII], H$\beta$, [OIII], H$\alpha$ and [NII]) observable at the BOSS wavelengths. This sample consists of 140,596 galaxy spectra, out of the 492,450 available in the full redshift range. We then extend our analysis to higher redshifts in Section 4.1.2, using alternative but more restrictive methods based on bluer emission lines that fall within the BOSS observed wavelengths.

#### BPT classification

Baldwin, Phillips & Terlevich (BPT, Baldwin et al., 1981) diagnostic classifications are determined as described in Section 2, using the emission lines H$\beta$, [OIII] $\lambda 5007$, H$\alpha$ and [NII] $\lambda 6583$. This classification scheme uses emission lines to indicate the ionisation source of the interstellar gas within galaxies, and has been proven to be very powerful at separating star forming galaxies from objects containing AGN. Applying an AoN limit of 2 for all BPT-diagnostic lines yields a
Figure 4.9: Emission line classification [Baldwin et al. 1981] for BOSS galaxies at various redshift bins (redshift increasing from top to bottom) for CMASS galaxies (left panels) and LOWZ galaxies (right panels). Objects at $z < 0.45$ are selected to warrant detectability of Hα and [NII]λ 6583. The fraction of objects for which all four emission lines are detected (requiring the amplitude-over-noise ratio of all four lines to be larger than two) is given in the bottom-left corner of each panel. The empirical separation between star forming galaxies and AGN (dashed line) is from [Kauffmann et al. 2003], and the theoretical extreme star burst line from [Kewley et al. 2001] to identify pure AGN emission (solid curved line). A dividing line defined by [Schawinski et al. 2007] is used to distinguish between LINER and Seyfert emission (solid straight line). The fractions of galaxies in each of the emission line classes are given by the labels in the panels. Overall, the fraction of objects with detected emission lines is small in BOSS. There is a marked difference in the emission line properties between the BOSS galaxy samples at high and low redshift, mainly caused by selection effects. Furthermore, there is a striking difference between the LOWZ and CMASS samples (see text for details).
sample of 4,887 galaxy spectra (3.5% of the total in this redshift range), although we note that applying this limit to only Hα and [NII]λ6583 gives a considerably larger count (∼ 30%). In all of our analysis we require all four lines to meet this requirement for us to consider something a galaxy with emission lines, and therefore make it identifiable on the BPT diagram.

Fig. 4.9 shows the BPT diagram for our low redshift sample in the redshift bins 0 < z < 0.15, 0.15 < z < 0.3 and 0.3 < z < 0.45 for both the CMASS (top) and LOWZ (bottom) samples. The dashed line is the empirical star forming galaxies—AGN separation line of [Kauffmann et al. (2003)], the solid curved line is the theoretical extreme starburst line of [Kewley et al. (2001)], and the solid straight line distinguishes between LINER and Seyfert emission based on SDSS galaxy classifications obtained through the [SII]/Hα ratio [Schawinski et al. (2007)]. We assume the area between the two curved lines is populated by galaxies with a composite of star burst and AGN spectra.

Clearly there is a strong dependence of the emission line properties of BOSS galaxies on redshift, in contrast to the distributions of galaxy masses and stellar velocity dispersions. This is primarily caused by selection effects. We also see a clear difference between the LOWZ and CMASS samples. Whilst the number of emission line galaxies in LOWZ sample obviously decreases with redshift, the number in the CMASS sample increases; this is due to the design of the target selection. The overall fraction of galaxies with detected emission lines also differs between the samples and with redshift. We also see that there is a relatively low abundance of star forming galaxies, which we’d generally expect as the BOSS survey tries to target massive galaxies at all redshifts.

The sample with the highest fraction of emission line galaxies in BOSS is the LOWZ sample at redshifts below 0.15, where 15% of all galaxies in the sample are found to have AoN > 2 for the BPT lines. This is not surprising as this is the regime with the highest level of contamination by lower-mass galaxies, as shown in Fig. 4.8. We see that within the LOWZ sample, the majority of emission-line galaxies have some AGN component, with the overall fraction of star forming galaxies being only 20%. We also see within this sample that the fraction of emission-line galaxies drops dramatically with increasing redshift to only a few percent, and that the relative fraction of AGN and LINER-like emission also increases with redshift.

Despite target selection cuts designed to limit CMASS galaxies to those with
CHAPTER 4. SDSS-III/BOSS

Figure 4.10: Stellar velocity dispersion distributions of the entire CMASS sample (grey), the CMASS sample with $z < 0.15$ (red striped) and the CMASS sample with $z < 0.15$ that are designated as star forming in the BPT (blue striped).

For $z > 0.4$, we see that there are still a number of objects in the CMASS sample with $z < 0.15$. Most of these objects with emission lines are star forming galaxies that fell into the CMASS colour selection cut despite being at low redshift, probably due to dust reddening. Fig. 4.10 shows the velocity dispersion distribution for the entire CMASS sample (grey), all CMASS objects with $z < 0.15$ (red striped) and the CMASS objects that have $z < 0.15$ as well as being classified as star forming on the BPT (blue striped), with all distributions normalised to 1. Clearly this reveals that CMASS objects with $z < 0.15$ are relatively low mass, especially those that are found to be star forming, with velocity dispersions at the low tail of the distribution. With increasing redshift this population progressively disappears from the CMASS sample.

Closer to the redshift range actually targeted by the CMASS sample, for $z > 0.3$, we find that only a small fraction of CMASS objects have emission lines, and those objects were mostly AGN. Whilst we find that most of these objects are Seyferts, we stress the caveat that AGN fractions derived through optical...
emission line ratios are strongly dependent on S/N ratios. LINERs in particular typically have lower emission line fluxes, and tend to drop out of the sample as S/N decreases. We therefore argue that the relatively low fraction of BOSS galaxies with LINER emission at high redshifts is most likely a selection effect, and needs to be considered carefully in any scientific analysis looking at this.

**Emission line properties at higher redshifts**

Due to the problem of the H\(\alpha\) and [NII]\(\lambda6583\) emission lines shifting out of the observed optical wavelengths at redshifts \(z > 0.45\), an alternative method of analysis is required to determine AGN and star forming fractions above this limit. Fortunately, much of the separation between AGN and star forming galaxies is provided by the [OIII]\(\lambda5007/\)H\(\beta\) ratio, which is observable over the entire BOSS redshift range. The main difficulty then becomes the identification of LINER-like emission and SF-AGN composites with low [OIII]/H\(\beta\), which places them where most star forming objects lie, and which are usually separated by their enhanced [NII]/H\(\alpha\) in the BPT diagram.

Numerous alternative diagnostics have been suggested in the literature. One suggestion by Lamareille (2010) is to replace [NII]/H\(\alpha\) by [OII]\(\lambda\lambda3726+3729/\)H\(\beta\), which helps to identify LINER-like emission by their enhanced [OII]/H\(\beta\) ratios. Alternative approaches have used additional information from the stellar population properties of the host galaxy to separate AGN from star formation. The Mass-Excitation (MEx) method by Juneau et al. (2011) makes use of the fact that AGN are typically hosted by relatively massive galaxies, whilst the approaches of Yan et al. (2011) and Trouille et al. (2011) are potentially very effective methods that identify AGN through the relatively red \(U - B\) and \(g - z\) rest-frame colours of their host galaxies.

The MEx method only works over a large range in galaxy mass, hence cannot be applied to the BOSS sample which covers a very narrow range at the high-mass end (Maraston et al., 2013). A potential problem of the approaches by Yan et al. (2011) and Trouille et al. (2011) is that we may not necessarily know how host galaxy properties change with redshift. By tapping into the host galaxy properties we may introduce a bias or some contamination in the classification calibrated at low redshifts. Also, the translation into rest-frame colours requires stellar population modelling, which has the risk of introducing further uncertainties if the models are imperfect. In particular, the rest-frame z-band wavelength range is
I therefore decided to focus on using the blue emission line diagnostics by Lamareille (2010) (hereafter L10), which seems to be most comparable to the BPT approach we use elsewhere for redshifts $z < 0.45$. In detail, this diagnostic defines $y = \frac{\text{[OIII]} \lambda 5007}{H\beta}$ and $x = \frac{\text{[OII]} \lambda 3726 + 3729}{H\beta}$, and uses the dividing lines $y_1 = 0.11/(x - 0.92) + 0.85$ for $x < 0.92$, $y_2 = 0.3$ for $x < 0.72$ and $y_3 = (0.95 \times x) - 0.4$ for $x > 0.72$ to separate between star forming objects, Seyferts, LINERs, and a region of mixed star forming and Seyfert objects. Star forming objects are defined as those with $y < y_1$ and $y < y_2$, Seyferts with $y > y_1$ and $y > y_3$, LINERs with $y > y_1$ and $y < y_3$ and the mixed star forming/Seyfert region with $y < y_1$ and $y > y_2$. We demonstrate this diagnostic in Fig. 4.11 for all objects in BOSS with all L10 diagnostic lines having AoN $> 2$, where contours cover the 1st, 5th, 25th, 50th, 75th and 95th percentiles.

The major drawback of the L10 diagnostic is that, while the blue emission not covered by SDSS photometry at BOSS redshifts and requires near-IR imaging which we do not have available.
line diagnostic diagram does help to identify LINER-like emission through the enhanced [OII]/H$\beta$ ratio, the LINER region is still significantly contaminated with star forming objects and star forming-AGN composites. We show how L10 classifications compare with BPT classifications in Table 4.1. In this table, the columns represent BPT classes and the rows L10 classes, with the bold numbers in each cell being the number of objects that fall into both classifications, the percentages being the those of objects in that cell of the total L10-classified (left) and BPT-classified (right) objects per classification (i.e. the percentages on the left should add to 100 horizontally, and those on the right vertically). This is then visualised in Fig. 4.12 where the left panel shows the BPT-classified star forming and LINER-like objects in blue and red respectively, and the right panel BPT-classified composites and Seyferts in purple and green respectively, as shown on the L10 diagnostic diagram.

The large number of BPT-classified ‘Passive’ galaxies falling into other L10 classes (∼3,000 in total) is an unsurprising finding caused by objects with weak [NII] or H$\alpha$ having detectable [OII] emission. We see that the majority of these objects are classed by L10 as star forming or having star forming/Seyfert emission, making this likely to be caused by objects having low [NII] which would otherwise have appeared on the left side of the BPT diagnostic diagram. We see that it is impossible to identify a pure LINER-emission sample, with around 40% of L10 LINERS being detected as star forming or composite by the BPT. On the other hand we see that the L10 star forming region is contaminated by only 2% of objects identified by the BPT as having no star formation, and the L10 SF/Sy region is dominated by BPT star forming objects and those declared passive in the BPT. The L10 Seyfert classification is shown to be a fairly good representation of BPT Seyferts, with ∼80% of objects in this category agreeing with their counterpart BPT classification.

Overall, this makes the identification of a pure LINER sample and star forming-AGN composites impossible using L10 alone. Still, the diagnostic is useful to identify overall fractions in star forming and (Seyfert-type) nuclear activity.

I analyse the emission line properties of CMASS galaxies in the redshift range $0.3 < z < 0.75$. In total there are 373,924 CMASS galaxies at $z > 0.3$, for 10,238 of which (2.7%) all three blue diagnostic emission lines ([OII]λλ3726 + 3729, H$\beta$, [OIII]λ5007) could be detected at an AoN larger than 2. Fig. 4.13 presents the results for the redshift intervals $0.3 < z < 0.45$, $0.45 < z < 0.6$, and $0.6 <$
Figure 4.12: The L10 diagnostic diagram showing top: BPT-classified star forming objects (blue) and LINER-like objects (red), and bottom: BPT-classified composites (purple) and Seyferts (green).
Table 4.1: Object counts with $z < 0.45$ and BPT/L10 classifications shown by column/row. Numbers in brackets are the percentage of objects in that cell of the total L10-classified (left) and BPT-classified (right) objects per classification.

<table>
<thead>
<tr>
<th>L10</th>
<th>SF</th>
<th>Comp</th>
<th>Seyfert</th>
<th>LINER</th>
<th>Passive</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>SF</td>
<td>6,232</td>
<td>1,654</td>
<td>5</td>
<td>184</td>
<td>1,206</td>
<td>9,281</td>
</tr>
<tr>
<td></td>
<td>67%</td>
<td>78%</td>
<td>18%</td>
<td>0%</td>
<td>0%</td>
<td>2%</td>
</tr>
<tr>
<td>SF/Sy</td>
<td>1,221</td>
<td>54</td>
<td>153</td>
<td>49</td>
<td>1,009</td>
<td>2,486</td>
</tr>
<tr>
<td></td>
<td>49%</td>
<td>15%</td>
<td>2%</td>
<td>6%</td>
<td>7%</td>
<td>2%</td>
</tr>
<tr>
<td>Seyfert</td>
<td>143</td>
<td>16</td>
<td>1,758</td>
<td>101</td>
<td>258</td>
<td>2,276</td>
</tr>
<tr>
<td></td>
<td>6%</td>
<td>2%</td>
<td>1%</td>
<td>0%</td>
<td>77%</td>
<td>4%</td>
</tr>
<tr>
<td>LINER</td>
<td>351</td>
<td>1,483</td>
<td>291</td>
<td>2,009</td>
<td>420</td>
<td>4,554</td>
</tr>
<tr>
<td></td>
<td>8%</td>
<td>4%</td>
<td>33%</td>
<td>46%</td>
<td>6%</td>
<td>13%</td>
</tr>
<tr>
<td>Passive</td>
<td>18</td>
<td>42</td>
<td>16</td>
<td>18</td>
<td>270,396</td>
<td>270,490</td>
</tr>
<tr>
<td></td>
<td>0%</td>
<td>0%</td>
<td>1%</td>
<td>0%</td>
<td>1%</td>
<td>0%</td>
</tr>
<tr>
<td>Total</td>
<td>7,965</td>
<td>3,249</td>
<td>2,257</td>
<td>2,361</td>
<td>273,303</td>
<td>–</td>
</tr>
</tbody>
</table>
Figure 4.13: Emission line classification Lamareille (2010) for CMASS BOSS galaxies at higher redshift bins (redshift increasing from left to right). The fraction of objects for which all three emission lines are detected (requiring the amplitude-over-noise ratio of all three lines to be larger than two) is given in the bottom-left corner of each panel. The fractions of galaxies in each of the emission line classes are given by the labels in the panels. The relative fractions of star forming and AGN in the redshift interval $0.3 \leq z \leq 0.45$ agree well with the results from the BPT classification in Fig. 4.9. Beyond those redshifts, the fraction of star forming CMASS galaxies increases considerably with increasing redshift.
$z < 0.75$, equivalent to the BPT diagnostic of Fig. 4.9 at lower redshifts. The fractions of galaxies in each of the emission line classes are given by the labels in the panels.

The relative fractions of star forming/star forming-AGN composite and AGN in the redshift interval $0.3 \leq z \leq 0.45$ agree well with the results from the BPT classification in Fig. 4.9. Beyond those redshifts, the fraction of star forming CMASS galaxies increases considerably with increasing redshift. At redshifts above 0.6, almost two thirds of the CMASS galaxies with detected emission lines are star forming or SF-AGN composites. This trend is most probably due to the fact that the H$\beta$ emission line is the weakest among those three, so that objects with low H$\beta$ fluxes drop out first with increasing redshift. These are AGN, while star forming galaxies with stronger H$\beta$ emission stay above the AoN threshold. A scientific analysis aimed at studying relative fractions of star forming and AGN with redshift will need to perform a careful assessment of this selection effect. Still, the BOSS sample is certainly useful for the identification and selection of massive star forming galaxies at redshifts around $z \sim 0.6$.

4.1.3 Target selection colour-colour space

I have shown that a small but significant fraction of BOSS galaxies contains emission lines, encompassing all ionisation classes from star forming to AGN. It is interesting to investigate whether these emission line properties follow some distinct pattern in the colour-colour diagram used for target selection in BOSS. In the following we split in the redshift ranges $z < 0.45$ and $0.45 < z < 0.75$, so that we can benefit from the full BPT classification where accessible for the separation between emission-line free galaxies, star forming and AGN. In Fig. 4.14 we present the colour-colour plot $g-r$ vs $r-i$ of all DR9 BOSS galaxies. The top panel is for $z < 0.45$, and the emission line classes are derived from the BPT classification. The bottom panel shows BOSS galaxies in the redshift range $0.45 < z < 0.75$, and the emission line classification is based on the blue diagnostic diagram of Lamareille (2010).
Figure 4.14: BOSS target selection colour-colour diagram (based on galactic extinction-corrected $\text{modelmags}$). Objects below redshift $z = 0.45$ and above $z = 0.45$ are plotted in the left-hand and right-hand panels, respectively. Emission line classification is based on the BPT diagram of Fig. 4.9 ($z < 0.45$, left-hand panel) and the Lamareille (2010) diagnostics of Fig. 4.13 ($z > 0.45$, right-hand diagram). Emission line free galaxies are plotted as grey symbols. Galaxies with emission line detections at an amplitude-over-noise ratio larger than 2 in all diagnostic lines are overplotted as coloured symbols, following the colour-coding of Figs. 4.9 and 4.13. The solid line labelled $d_{\parallel}$ separates the high-$z$ sample CMASS from the low-$z$ sample LOWZ ($d_{\parallel} \equiv (r - i) - (g - r)/8 = 0.55$). Galaxies above this line are generally at $z > 0.4$. The solid lines labelled ‘LRG’ and ‘2SLAQ’ are the selection cuts used in Eisenstein et al. (2001) and Cannon et al. (2006). The dashed line is the dividing line between early-type and late-type from Masters et al. (2011). Total galaxy numbers and emission line class fractions are given in Table 4.2.
The $(g-r)$ vs $(r-i)$ diagram

Galactic extinction-corrected *modelmags* are used. The solid line labelled $d_\perp$ indicates the major colour selection cut

$$d_\perp \equiv (r - i) - (g - r)/8 = 0.55$$

(4.2)

that has been used to separate the high-redshift CMASS and the low-redshift LOWZ samples (Dawson et al. 2013). The colour $r - i$ is an excellent redshift indicator around redshifts of $z = 0.4$, because here the 4000Å break passes from the $g$-band into the $r$-band. As a consequence, galaxies become significantly redder in $r - i$ and slightly bluer in $g - r$ with increasing redshift beyond a redshift of $z = 0.4$ (Eisenstein et al. 2001). This pattern is displayed by the BOSS galaxy data in Fig. 4.14. Galaxies above the solid line ($d_\perp \geq 0.55$) define the CMASS sample and typically have redshifts $z \geq 0.4$, while everything below that line is in the LOWZ sample with typical redshifts of $z \leq 0.4$. The solid lines labelled ‘LRG’ and ‘2SLAQ’ are the selection cuts used by Eisenstein et al. (2001) to define the SDSS LRG sample and by Cannon et al. (2006) to define the 2SLAQ LRG sample, respectively. The dashed line is the dividing line between early-type and late-type from Masters et al. (2011) (‘M11’).

The correlation seen in the LOWZ sample at $z < 0.45$ (top panel) is a sequence of varying redshift in which both $g - r$ and $r - i$ become redder with increasing redshift up to $z = 0.4$. The sequence is relatively tight, because redshift effects dominate the observed colours. This is demonstrated in Fig. 4.15 which shows the $g - r$ against $r - i$ diagram with symbol colour representing redshift, and all objects with $z > 0.9$ plotted as red. The horizontal line represents the $d_\perp$ line, separating the LOWZ and CMASS samples. The redshift distribution of CMASS galaxies in the $g-r$ vs $r-i$ colour-colour plane is quite different to that of LOWZ. As the observed $g-r$ colour depends strongly on the star formation history of the galaxy, a wide range in $g-r$ colours is covered. The observed $r-i$ colour, instead, is generally red with a relatively small dynamical range and much less dependent on the star formation history. Passively evolving galaxies populate the top right-hand section of the diagram, as indicated by the LRG selection line.
Figure 4.15: $g - r$ vs $r - i$ where colour indicates redshift. All objects with $z > 0.9$ are displayed as red. The horizontal line represents $d_\perp$, separating the LOWZ and CMASS samples.
Correlations between observed frame colours and emission line properties

The full BOSS sample is plotted as grey symbols, galaxies for which emission lines have been detected are plotted as coloured symbols. The various emission-line classes derived in the present work are shown using the colour-code of Fig. 4.9 (top panel) and Fig. 4.13 (bottom panel).

As is to be expected, there is a large overlap of the various emission line classes for the LOWZ sample at \( z < 0.45 \). Still, galaxies with current star formation activity (blue points) are slightly shifted towards bluer \( g-r \) colour by \( \sim 0.1 \) mag, while no measurable offset is observed for the AGN populations (green and red symbols). Galaxies with SF-AGN composite spectra can be found in between those two extremes (purple points). This pattern of star-forming galaxies in the ‘blue cloud’ and transition objects in the ‘green valley’ between the blue cloud and the red sequence has previously been seen in colour-magnitude diagrams of SDSS-II galaxies (Graves et al., 2007; Schawinski et al., 2007; Salim et al., 2007).

For the CMASS sample at redshifts above \( z \sim 0.4 \), the separation of the various emission-line classes is far more pronounced. There are clear correlations between the observed frame colours and the emission-line properties of CMASS galaxies. There is a strong bias such that CMASS galaxies without emission lines are predominantly found in the classical LRG section towards the right-hand side of the diagram at red \( g-r \) colours (most evident in the right-hand panel). CMASS galaxies with emission lines, instead, are more likely to be found at blue \( g-r \) colour, and galaxies with star formation activity (blue points), in particular, populate the blue end of the \( g-r \) colour space, typically having \( (g-r) \leq 1.0 \). AGN, both Seyfert and LINER, preferentially occupy intermediate \( g-r \) colour in between those two extremes. This pattern can be observed in both the low-z and the high-z versions of the diagram. It should be noted that a large fraction of the star forming galaxies in CMASS/BLUE at blue \( g-r \) colours at \( z < 0.45 \) (left-hand diagram) are low redshift \( (z \sim 0.1) \) interlopers that fell into the CMASS colour cut region most likely because of dust reddening as discussed in Section 4.1.2 (see Fig. 4.9).

Table 4.2 quantifies the fractions of emission line galaxies and the various emission line classes as well the total number of objects for the four sections in Fig. 4.14 defined by the solid lines: CMASS/LRG (above \( d_\perp \), right of ‘LRG’), CMASS/M11 (above \( d_\perp \), between ‘LRG’ and ‘M11’), CMASS/2SLAQ (above \( d_\perp \),
between ‘M11’ and ‘2SLAQ’), CMASS/BLUE (above $d_\perp$, left of ‘2SLAQ’), LOWZ (below $d_\perp$). CMASS/BLUE is most abundant in emission-line objects (23% at $z < 0.45$ and 7% at $z > 0.45$) and star forming galaxies (around half of emission line spectra). The classical LRG section (CMASS/LRG), instead, is devoid of emission line objects (less than per cent). CMASS/2SLAQ, instead, contains the largest fraction of galaxies with Seyfert-like AGN emission lines (60%).

### Star formation histories and morphologies

The dependence of star formation activity on observed $g-r$ colour for the CMASS galaxies as found here is not surprising. As discussed above, it was to be expected that the blue $g-r$ colour is driven by the presence of young stellar populations caused by recent or current star formation episodes. [Tojeiro et al., 2012] analyse the star formation histories of CMASS galaxies through spectral SED fitting and find extended star formation in blue CMASS galaxies (see also [Chen et al., 2012]). The present results confirm this finding.

If the CMASS galaxies with blue $g-r$ are dominated by young stellar populations and have a large fraction of star forming galaxies, one would expect to find mostly late-type galaxies. This is indeed the case. [Masters et al., 2011] analyse the morphologies of BOSS galaxies through HST/COSMOS imaging and

---

**Table 4.2: Percentages of emission line classes from Fig. 4.14**

<table>
<thead>
<tr>
<th></th>
<th>CMASS</th>
<th>CMASS</th>
<th>CMASS</th>
<th>CMASS</th>
<th>LOWZ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LRG</td>
<td>M11</td>
<td>2SLAQ</td>
<td>BLUE</td>
<td></td>
</tr>
<tr>
<td>$z &lt; 0.45$ (left-hand panel in Fig. 4.14)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>17,390</td>
<td>518</td>
<td>8,685</td>
<td>6,890</td>
<td>104,317</td>
</tr>
<tr>
<td>Emission</td>
<td>0.84</td>
<td>6.76</td>
<td>4.17</td>
<td>23.35</td>
<td>2.46</td>
</tr>
<tr>
<td>SF</td>
<td>19.86</td>
<td>42.86</td>
<td>17.96</td>
<td>69.48</td>
<td>12.94</td>
</tr>
<tr>
<td>SF/AGN</td>
<td>13.70</td>
<td>14.29</td>
<td>15.19</td>
<td>10.69</td>
<td>36.18</td>
</tr>
<tr>
<td>Seyfert</td>
<td>53.42</td>
<td>40.00</td>
<td>61.60</td>
<td>17.65</td>
<td>15.59</td>
</tr>
<tr>
<td>LINER</td>
<td>13.01</td>
<td>2.86</td>
<td>5.25</td>
<td>2.18</td>
<td>35.28</td>
</tr>
</tbody>
</table>

| $0.45 < z < 0.75$ (right-hand panel in Fig. 4.14) |       |       |       |       |      |
| Total  | 211,389 | 55,220 | 53,609 | 28,777 | 4,085 |
| Emission | 0.33 | 0.79  | 4.78  | 18.95 | 3.30  |
| SF     | 15.20 | 12.79 | 26.86 | 60.22 | 60.74 |
| SF/AGN | 3.55  | 2.74  | 6.95  | 11.57 | 5.19  |
| Seyfert| 50.43 | 65.98 | 53.69 | 22.38 | 17.04 |
| LINER  | 30.82 | 18.49 | 12.50 | 5.83  | 17.04 |
find that most CMASS galaxies with blue $g - r$ colour are in fact late type systems. Their morphology-driven dividing line between early-type and late-type at $g - i = 2.55$ (dashed line in Fig. 4.14) separates quite well between star forming and passive galaxies.

### 4.1.4 Conclusions

I show that the typical signal-to-noise ratio of BOSS spectra, despite being low ($\sim 5\AA^{-1}$), is sufficient to measure simple dynamical quantities such as stellar velocity dispersion for individual objects. I verify the reliability of our measurements on individual BOSS spectra through comparison with high signal-to-noise spectra from repeat-plate observations in BOSS, using a sub-sample of 574 BOSS galaxies with at least six 1-hour observations each. The agreement is very good. There is no systematic offset between measurements on individual (low S/N) and stacked (high S/N) spectra. I also do not find any systematic trend with S/N ratio. Finally, I compare our measurements with independent measurements within the BOSS collaboration by Bolton et al. (2012) and Chen et al. (2012), and find good agreement. The typical error in the velocity dispersion measurement is 14%, and 93% of BOSS galaxies have velocity dispersions with an accuracy better than 30%. I show that the typical velocity dispersion of a BOSS galaxy is $\sim 240$ km/s. The distribution in velocity dispersion is nearly Gaussian and is redshift independent between redshifts 0.15 and 0.7. At redshifts below $z \sim 0.15$ and above $z \sim 0.75$, the distributions are skewed and slightly offset toward lower and higher velocity dispersions, respectively. This reflects the survey design targeting massive galaxies with an approximately uniform mass distribution in the redshift interval $0.15 < z < 0.75$.

I show that emission lines can be measured on BOSS spectra, but the majority of BOSS galaxies lack detectable emission lines, as is to be expected because of the target selection design toward massive galaxies.

I analyse the emission line properties for a subsample of 140,596 galaxies below $z = 0.45$, so that the full set of key diagnostic emission lines Hβ, [OIII]λ5007, Hα, and [NII]λ6583 are accessible in the rest-frame spectra. All four diagnostic lines are detected at an amplitude-over-noise ratio above two for 4,887 spectra (3.5%). For these, I present the classical diagnostic diagrams (Baldwin et al., 1981) to divide star-forming objects from AGN separately for the high-z sample CMASS and the low-z sample LOWZ. I find that the emission line properties are strongly
redshift dependent. Furthermore, there is a clear correlation between observed frame colours and emission line properties. In general, the fraction of star forming galaxies decreases and the fraction of AGN increases with increasing redshift, mostly owing to selection effects. Within in the LOWZ sample, the majority of emission-line galaxies has some AGN component, the fraction of purely star forming galaxies only being a few per cent at $z > 0.15$. The CMASS sample, instead, contains bluer galaxies and the fraction of star forming galaxies is as high as 20 per cent at $z > 0.3$. Interestingly, there are some CMASS galaxies at low redshifts ($z < 0.15$) that are star forming and fell into the CMASS colour selection cut most probably due to dust reddening.

To assess the emission line properties of BOSS galaxies at higher redshifts, I additionally study the 373,924 CMASS galaxies in the redshift range $0.3 < z < 0.75$, containing 10,238 objects (2.7%) with significant emission line detections. For this purpose I use the blue diagnostic diagram of Lamareille (2010) based on the emission lines $\text{OII} \lambda \lambda 3726 + 3729$, $\text{H} \beta$, and $\text{OIII} \lambda 5007$. For this sample, the fraction of star forming galaxies is considerably higher, which is most probably due to the fact that the $\text{H} \beta$ emission line is the weakest among the three diagnostic lines, so that objects with low $\text{H} \beta$ fluxes, hence AGN, drop out first with increasing redshift. Therefore, the BOSS sample turns out to be instrumental for the identification and selection of massive star forming galaxies at redshifts around $z \sim 0.6$.

Finally, we show that CMASS galaxies whose emission lines are produced by star formation activity have blue observed $g - r$ colours and are well separated in the $g - r$ vs $r - i$ target selection diagram.

To conclude, BOSS offers spectra of a large sample of galaxies up to redshifts $\sim 0.8$. The quality of BOSS spectroscopy, even though designed for redshift determination, allows the measurement of simple quantities on individual BOSS spectra for a wealth of galaxy evolution studies on dynamical, gas, and stellar population properties.

### 4.2 Stellar population analysis of high-S/N BOSS galaxies

Whilst emission lines probe gas kinematics, current star formation and AGN activity and originate from the gas in a galaxy, absorption features give us a detailed
view of the stellar history of a galaxy. They are able to do this as they originate from stellar atmospheres and therefore mostly from the parent gas clouds of stars themselves, with a minor addition from internal convective mixing. They are also useful as they are largely insensitive to dust attenuation (MacArthur 2005), can be measured and calibrated to a common system (Burstein et al. 1984; Faber et al. 1985), and can provide unique information on the individual element abundances in galaxies which can be used to set constraints on their chemical enrichment and star formation histories (Thomas et al. 1999, 2003).

Element abundances are interesting to know as other than H and the majority of He, which were formed in the primordial nucleosynthesis, all elements were formed in stellar nucleosynthesis. As such, knowledge of the elements present in a galaxy gives an understanding of the stellar activity prevalent within it in previous stellar generations. Further, since different elements are produced in different stellar evolutionary phases, some of which do not occur for all stars (e.g. supernovae and their different types), varying formation histories can be traced by comparing abundance ratios (Matteucci 1994; Thomas et al. 1999). This can lead to a deeper understanding of the initial mass function (IMF), galaxy formation in general and galaxy evolution (e.g. Thomas et al. 2005, 2010; La Barbera et al. 2013).

The most widely used absorption line system is the Lick index system (see Section 2.4 for a detailed description of the Lick indices and how they are calculated). Modelling of the Lick indices allows for the derivation of galaxy ages, formation epochs, star formation histories and element abundances (Trager et al. 2000; Thomas et al. 2005; Bernardi et al. 2006; Thomas et al. 2010), effectively allowing for observations of the history of galaxies, known as the archaeological approach at $z \sim 0$. Our method for extracting galaxy properties from these indices is covered in Section 2.4.1.

The results presented here are the first to present absorption line index and from these age, metallicity and $\alpha$/Fe measurements to such a high redshift ($z \sim 0.7$) for a reasonably numbered galaxy sample with high S/N. We provide conclusive evidence that massive galaxies have only experienced minor evolution in their metal content over half the age of the Universe, and have evolved passively with minimal star formation. The downsizing pattern with lower-mass galaxies exhibiting more extended star formation histories (e.g. Brinchmann & Ellis 2000a; Ferreras & Silk 2000; De Lucia et al. 2004; Heavens et al. 2004; Jimenez 2005).
et al., 2005; Nelan et al., 2005; Thomas et al., 2005; Ziegler et al., 2005; Thomas et al., 2010) is clearly visible from the data analysed here.

4.2.1 Stacking

The typical S/N of a BOSS object is $\sim 5 \text{ pix}^{-1}$, which is considerably too low to obtain accurate measurements of absorption line index strengths. We therefore found it necessary to create high-S/N stacks of BOSS objects in bins of redshift and stellar velocity dispersion, in order to investigate how stellar populations evolve as a function of these variables.

Stacks were created by summing the flux of BOSS galaxies in bins of redshift and stellar velocity dispersion, with standard deviations propagated in the normal way, resampled to a linear wavelength grid covering 2500 – 6500Å (rest frame) in steps of 1Å. I require that all BOSS galaxies added to the stack have a minimum S/N of 1.5 pix$^{-1}$, in order to avoid just adding noise to the spectrum, and also a maximum error in velocity dispersion of 30% in order to minimise contamination between velocity dispersion bins. I calculate the S/N of the final stack using the formula

$$S/N = \frac{\text{MEDIAN} (f_b)}{\text{ROBUST\_SIGMA} (f_b - f_s)}$$

(4.3)

where $f_b$ is the flux of the best fit model and $f_s$ is the flux of the stack. I use the fit residuals rather than propagated variance to calculate the S/N as it is a more robust measurement of the noise of the spectrum. It is typically twice as large as the combined variance from the individual spectra.

Table 4.3 shows, respectively, the $z$- and log($\sigma$)-bins used for stacking objects, the number of objects in each stack, the S/N of the stack, the mean redshift of objects in the stack, the mean stellar velocity dispersion of the objects making up the stacks, and the measured stellar velocity dispersion of the stacked spectrum. Subscript $i$ indicates a measurement from the input spectra making up the stack; $s$ indicates a measurement on the stack itself. The redshift of the stack was not adjusted within GANDALF as all its constituent spectra were shifted to $z = 0$ in the stacking process. Clearly from this table the stellar velocity dispersions of the stacks agree well with the mean dispersions of their constituents, with the notable example of the lowest $\sigma$ bins in the highest $z$ bin. This offset is likely due to the relatively low S/N of high-$z$ objects, combined with the stack
containing relatively few objects, and also low-σ objects being generally fainter (and therefore with a lower S/N) than their high-σ counterparts.

The stacks themselves can be seen in Fig. 4.16 along with one randomly selected galaxy that contributes to the make-up of each stack. The flux densities in these plots have been normalised for comparison. The increase in S/N is obvious, especially at high redshifts where individual S/N values are low.

The S/N values of these stacks are sufficient to accurately measure Lick absorption-line indices. Fig. 4.17 demonstrates the quality of fit associated with these S/N values; we show each stack (top panels, black) along with its best-fit model (top panels, red), and the residuals of that fit (bottom panel, black). The blue line in the bottom panel indicates the zero point.

4.2.2 Lick indices

Here I present a subset of key indices to display the major trends in the data. The procedure used to derive population parameters uses the full set of indices (see Section 2.4.1 for further detail), and the results of this are described in Section 4.2.3. All errors on the measured indices were estimated through Monte Carlo simulations of the stacked spectra.

Fig. 4.18 shows the ratio of Mgb to Fe for each stack, where Fe = Fe5270 + Fe5335, in the stacks’ redshift bins (i.e. each panel shows stacks that fell into the same redshift bin when being created). Symbol size represents the stellar velocity dispersions of the stacks. The stellar population models of Thomas et al. (2011) are overplotted for various metallicities and α/Fe ratios, with the age for the models in each panel set at the mean age of the galaxies in that bin. The ages of the stacked objects were calculated using the stellar population fitting code detailed in Section 2.4.1. Models with different metallicities are displayed as sloping down to the right, with metallicities [Z/H] of −0.33, 0, 0.35 and > 0.35 from bottom left to top right respectively. Model contours going up and to the right have varying α/Fe ratios of 0, 0.3 and 0.5 from top left to bottom right. The red line in each panel indicates solar metallicity (Z= 0).

Here we clearly see from individual panels that metallicity increases with stellar velocity dispersion. There is also a tentative sign that α/Fe increases with σ, but this is inconclusive without full SSP parameter fitting (performed below). Overall we see that α/Fe falls between 0.3 and 0.5 for almost all objects, and metallicity is consistently super-solar. By comparing panels we see that both
Table 4.3: Statistics for each BOSS stacked object. $z$-bin shows the redshift range of galaxies in each stack; log($\sigma$)-bin shows the stellar velocity dispersion range (in log units) of galaxies in each stack. N is the number of objects in the stack, $S/N_s$ the stack signal-to-noise in units of pix$^{-1}$, $< z_i >$ and $< \sigma_i >$ the mean redshift and stellar velocity dispersion of the galaxies making up the stack (the latter in units of km s$^{-1}$), and $\sigma_s$ is the measured stellar velocity dispersion of the stack in units of km s$^{-1}$.

<table>
<thead>
<tr>
<th>$z$-bin</th>
<th>log($\sigma$)-bin</th>
<th>N</th>
<th>$S/N_s$</th>
<th>$&lt; z_i &gt;$</th>
<th>$&lt; \sigma_i &gt;$</th>
<th>$\sigma_s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1-0.2</td>
<td>2.0-2.2</td>
<td>2451</td>
<td>117.9</td>
<td>0.15</td>
<td>137.2</td>
<td>145.2</td>
</tr>
<tr>
<td>0.1-0.2</td>
<td>2.2-2.3</td>
<td>5208</td>
<td>137.2</td>
<td>0.15</td>
<td>181.3</td>
<td>185.8</td>
</tr>
<tr>
<td>0.1-0.2</td>
<td>2.3-2.4</td>
<td>8919</td>
<td>113.3</td>
<td>0.16</td>
<td>225.0</td>
<td>228.9</td>
</tr>
<tr>
<td>0.1-0.2</td>
<td>2.4-2.5</td>
<td>4430</td>
<td>87.2</td>
<td>0.16</td>
<td>273.0</td>
<td>276.3</td>
</tr>
<tr>
<td>0.1-0.2</td>
<td>2.5-2.6</td>
<td>288</td>
<td>81.6</td>
<td>0.17</td>
<td>336.4</td>
<td>338.0</td>
</tr>
<tr>
<td>0.2-0.3</td>
<td>2.0-2.2</td>
<td>3267</td>
<td>61.4</td>
<td>0.25</td>
<td>135.0</td>
<td>133.4</td>
</tr>
<tr>
<td>0.2-0.3</td>
<td>2.2-2.3</td>
<td>9128</td>
<td>67.6</td>
<td>0.25</td>
<td>183.4</td>
<td>186.0</td>
</tr>
<tr>
<td>0.2-0.3</td>
<td>2.3-2.4</td>
<td>28985</td>
<td>68.3</td>
<td>0.26</td>
<td>227.0</td>
<td>228.6</td>
</tr>
<tr>
<td>0.2-0.3</td>
<td>2.4-2.5</td>
<td>18614</td>
<td>66.8</td>
<td>0.26</td>
<td>273.7</td>
<td>273.3</td>
</tr>
<tr>
<td>0.2-0.3</td>
<td>2.5-2.6</td>
<td>1453</td>
<td>62.2</td>
<td>0.26</td>
<td>336.9</td>
<td>333.7</td>
</tr>
<tr>
<td>0.3-0.4</td>
<td>2.0-2.2</td>
<td>4587</td>
<td>75.8</td>
<td>0.35</td>
<td>139.1</td>
<td>139.7</td>
</tr>
<tr>
<td>0.3-0.4</td>
<td>2.2-2.3</td>
<td>16433</td>
<td>85.5</td>
<td>0.35</td>
<td>183.1</td>
<td>186.7</td>
</tr>
<tr>
<td>0.3-0.4</td>
<td>2.3-2.4</td>
<td>48604</td>
<td>80.8</td>
<td>0.35</td>
<td>226.7</td>
<td>229.7</td>
</tr>
<tr>
<td>0.3-0.4</td>
<td>2.4-2.5</td>
<td>33040</td>
<td>79.0</td>
<td>0.35</td>
<td>274.7</td>
<td>276.7</td>
</tr>
<tr>
<td>0.3-0.4</td>
<td>2.5-2.6</td>
<td>3670</td>
<td>67.2</td>
<td>0.36</td>
<td>338.0</td>
<td>338.9</td>
</tr>
<tr>
<td>0.4-0.5</td>
<td>2.0-2.2</td>
<td>16634</td>
<td>111.7</td>
<td>0.47</td>
<td>138.5</td>
<td>139.9</td>
</tr>
<tr>
<td>0.4-0.5</td>
<td>2.2-2.3</td>
<td>39955</td>
<td>132.9</td>
<td>0.46</td>
<td>181.5</td>
<td>183.6</td>
</tr>
<tr>
<td>0.4-0.5</td>
<td>2.3-2.4</td>
<td>78976</td>
<td>114.6</td>
<td>0.46</td>
<td>225.7</td>
<td>228.4</td>
</tr>
<tr>
<td>0.4-0.5</td>
<td>2.4-2.5</td>
<td>57991</td>
<td>100.7</td>
<td>0.46</td>
<td>276.8</td>
<td>279.4</td>
</tr>
<tr>
<td>0.4-0.5</td>
<td>2.5-2.6</td>
<td>11356</td>
<td>99.1</td>
<td>0.46</td>
<td>341.7</td>
<td>342.9</td>
</tr>
<tr>
<td>0.5-0.6</td>
<td>2.0-2.2</td>
<td>27159</td>
<td>127.2</td>
<td>0.55</td>
<td>138.1</td>
<td>139.3</td>
</tr>
<tr>
<td>0.5-0.6</td>
<td>2.2-2.3</td>
<td>57722</td>
<td>144.8</td>
<td>0.54</td>
<td>181.1</td>
<td>183.0</td>
</tr>
<tr>
<td>0.5-0.6</td>
<td>2.3-2.4</td>
<td>103740</td>
<td>138.4</td>
<td>0.54</td>
<td>225.7</td>
<td>227.6</td>
</tr>
<tr>
<td>0.5-0.6</td>
<td>2.4-2.5</td>
<td>79537</td>
<td>109.3</td>
<td>0.55</td>
<td>276.8</td>
<td>280.9</td>
</tr>
<tr>
<td>0.5-0.6</td>
<td>2.5-2.6</td>
<td>21614</td>
<td>92.6</td>
<td>0.55</td>
<td>341.7</td>
<td>348.3</td>
</tr>
<tr>
<td>0.6-0.7</td>
<td>2.0-2.2</td>
<td>8384</td>
<td>136.4</td>
<td>0.64</td>
<td>138.1</td>
<td>137.5</td>
</tr>
<tr>
<td>0.6-0.7</td>
<td>2.2-2.3</td>
<td>17151</td>
<td>147.7</td>
<td>0.64</td>
<td>181.1</td>
<td>179.6</td>
</tr>
<tr>
<td>0.6-0.7</td>
<td>2.3-2.4</td>
<td>32181</td>
<td>140.1</td>
<td>0.64</td>
<td>225.3</td>
<td>224.6</td>
</tr>
<tr>
<td>0.6-0.7</td>
<td>2.4-2.5</td>
<td>31043</td>
<td>125.3</td>
<td>0.64</td>
<td>278.1</td>
<td>278.9</td>
</tr>
<tr>
<td>0.6-0.7</td>
<td>2.5-2.6</td>
<td>13179</td>
<td>104.4</td>
<td>0.64</td>
<td>341.4</td>
<td>348.4</td>
</tr>
<tr>
<td>0.7-0.8</td>
<td>2.0-2.2</td>
<td>965</td>
<td>74.2</td>
<td>0.73</td>
<td>138.4</td>
<td>127.5</td>
</tr>
<tr>
<td>0.7-0.8</td>
<td>2.2-2.3</td>
<td>2011</td>
<td>86.8</td>
<td>0.73</td>
<td>181.2</td>
<td>173.9</td>
</tr>
<tr>
<td>0.7-0.8</td>
<td>2.3-2.4</td>
<td>3761</td>
<td>88.8</td>
<td>0.73</td>
<td>226.5</td>
<td>219.4</td>
</tr>
<tr>
<td>0.7-0.8</td>
<td>2.4-2.5</td>
<td>4291</td>
<td>81.5</td>
<td>0.73</td>
<td>281.4</td>
<td>277.6</td>
</tr>
<tr>
<td>0.7-0.8</td>
<td>2.5-2.6</td>
<td>2654</td>
<td>67.5</td>
<td>0.73</td>
<td>349.9</td>
<td>346.3</td>
</tr>
</tbody>
</table>
Figure 4.16: Normalised stacked spectra (red) and one randomly selected component galaxy of each stack (black).
z = 0.46  \( \sigma = 181 \)

z = 0.46  \( \sigma = 225 \)

z = 0.46  \( \sigma = 276 \)

z = 0.46  \( \sigma = 341 \)

z = 0.55  \( \sigma = 138 \)

z = 0.54  \( \sigma = 181 \)

z = 0.54  \( \sigma = 225 \)

z = 0.55  \( \sigma = 278 \)
metallcity and $\alpha$/Fe appear to be largely independent of redshift.

Fig. 4.19 shows the relationships between MgFe and the Lick indices $H\beta$, $H\delta_A$, and $H_\gamma A$, where MgFe = $\sqrt{Mgb \times (0.72 \times Fe5270 + 0.28 \times Fe5335)}$. Symbol size again represents stellar velocity dispersion, and colour indicates redshift. The stellar population models are overplotted, this time with varying age and metallicity, with $\alpha$/Fe fixed at 0.3 (left) and 0.5 (right). The dotted lines represent models with ages of 10 to 15 Gyr.

From these the most obvious (and unsurprising) conclusion is that galaxy age depends on redshift. We can also see here that metallicity is mostly independent of redshift, although this is less clear and again requires population modelling in order to be properly constrained. We see from the top panel that models using $H\beta$ and [MgFe] give ages older than the age of the Universe; this is a known and common problem with the $H\beta$ absorption line (Schiavon, 2007; Poole et al., 2010; Thomas et al., 2011). Comparing models with low (left) and high (right) $\alpha$/Fe ratios, we can see that models based on $H\delta_A$ and $H_\gamma A$ shift to give galaxies lower metallicities and higher ages, whilst those based on $H\beta$ are largely independent of $\alpha$/Fe.
Figure 4.17: Stacked spectra (top panels, black) and their best-fit models (top panels, red), along with the residuals of this fit (bottom panels, black). The blue lines indicate the zero point for the residuals. This demonstrates the quality of fits associated with each S/N value.
4.2.3 Stellar population parameters

I derived the stellar population parameters age, metallicity and $\alpha$/Fe as described in Section 2.4.1 following the procedure of Thomas et al. (2010) and Johansson et al. (2012) for all 35 BOSS stacked objects. Fig. 4.20 shows these properties plotted against redshift (left) and stellar velocity dispersion (right). Colours are explained in the bottom panels for each column, but indicate different $\sigma$ values (left) and different redshifts (right). Errors are estimated through Monte Carlo simulations of the Lick indices. I perform correlation analysis on each line (of constant $\sigma$ in the left plots, constant $z$ in the right) and present our results in Table 4.4. Here, $r$ is the Pearson’s $r$ value, $t_0$ is the t-test statistic, and $t$ is the threshold t-test value required for an 80% confidence sign of a non-zero slope. The absolute value of $t_0$ must exceed the value of $t$ in order for there to be a significant non-negligible slope.

Correlations with velocity dispersion

We see relationships between all three stellar population parameters with velocity dispersion at all redshifts. Our results generally reproduce the well-known pattern
Figure 4.18: Lick indices $M_{gb}$ vs $\langle Fe \rangle$ where $Fe = Fe_{5270} + Fe_{5335}$, in panels of redshift for all of our BOSS stacked objects. The lines represent the models of Thomas et al. (2011) at metallicities $[Z/H]$ of $-0.33$, $0$ and $0.35$, $\alpha$/Fe of $0$, $0.3$ and $0.5$, and ages set for each panel by the mean age of the stacks. Metallicity increases from bottom left to top right, and $\alpha$/Fe increases from top left to bottom right. The age of the stacks was calculated using a stellar population fitting code, as explained in Section 2.4.1. The red line indicates solar metallicity.
Figure 4.19: MgFe vs the Lick indices H\text{\textbeta} (top), H\delta A (middle) and H\gamma A (bottom) where MgFe = \sqrt{Mgb \times (0.72 \times Fe5270 + 0.28 \times Fe5335)} for all of our BOSS stacked objects. The lines represent the models of [Thomas et al. (2011)] at \alpha/Fe of 0.3 (left), 0.5 (right), metallicities [Z/H] of −0.33, 0, 0.35 and 0.67 (vertical lines from left to right), and ages of 1, 2, 4, 7, 10 and 15 Gyr (from top to bottom). Dotted lines represent ages of 10 and 15 Gyr. Colours indicate redshift (black being the lowest and red the highest), and symbol size indicates stellar velocity dispersion.
Figure 4.20: Derived ages (top), metallicities (middle) and $\alpha$/Fe ratios (bottom) plotted against redshift (left) and stellar velocity dispersion (right) for all stacked objects (black) and mean values and standard deviations in respective $z$ and $\sigma$ bins (red).
### Table 4.4: Statistical analysis of Fig. 4.20 looking at the significance of slopes.

The panels of the table represent the panels in Fig. 4.20, and the different samples each line. $r$ is the Pearson’s coefficient for each line, $t_0$ is the t-test statistic for that line, and $t$ is the threshold t value required for a slope to be 80% significant.

<table>
<thead>
<tr>
<th>sample</th>
<th>$r$</th>
<th>$t_0$</th>
<th>$t$</th>
<th>sample</th>
<th>$r$</th>
<th>$t_0$</th>
<th>$t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$100 &lt; \sigma &lt; 160$</td>
<td>-0.88</td>
<td>-4.65</td>
<td>1.476</td>
<td>$0.1 &lt; z &lt; 0.2$</td>
<td>1.00</td>
<td>36.20</td>
<td>1.638</td>
</tr>
<tr>
<td>$160 &lt; \sigma &lt; 200$</td>
<td>-0.97</td>
<td>-10.43</td>
<td>1.476</td>
<td>$0.2 &lt; z &lt; 0.3$</td>
<td>0.99</td>
<td>12.71</td>
<td>1.638</td>
</tr>
<tr>
<td>$200 &lt; \sigma &lt; 250$</td>
<td>-0.96</td>
<td>-10.46</td>
<td>1.476</td>
<td>$0.3 &lt; z &lt; 0.4$</td>
<td>1.00</td>
<td>39.09</td>
<td>1.638</td>
</tr>
<tr>
<td>$250 &lt; \sigma &lt; 315$</td>
<td>-0.98</td>
<td>-11.11</td>
<td>1.476</td>
<td>$0.4 &lt; z &lt; 0.5$</td>
<td>0.80</td>
<td>2.64</td>
<td>1.638</td>
</tr>
<tr>
<td>$315 &lt; \sigma &lt; 400$</td>
<td>-0.98</td>
<td>-12.23</td>
<td>1.476</td>
<td>$0.5 &lt; z &lt; 0.6$</td>
<td>0.97</td>
<td>7.80</td>
<td>1.638</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\alpha/Fe$ vs $z$</th>
<th>$[Z/H]$ vs $\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$100 &lt; \sigma &lt; 160$</td>
<td>0.42</td>
</tr>
<tr>
<td>$160 &lt; \sigma &lt; 200$</td>
<td>-0.85</td>
</tr>
<tr>
<td>$200 &lt; \sigma &lt; 250$</td>
<td>-0.60</td>
</tr>
<tr>
<td>$250 &lt; \sigma &lt; 315$</td>
<td>-0.01</td>
</tr>
<tr>
<td>$315 &lt; \sigma &lt; 400$</td>
<td>0.80</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\alpha/Fe$ vs $\sigma$</th>
<th>$[Z/H]$ vs $z$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$100 &lt; \sigma &lt; 160$</td>
<td>-0.72</td>
</tr>
<tr>
<td>$160 &lt; \sigma &lt; 200$</td>
<td>-0.62</td>
</tr>
<tr>
<td>$200 &lt; \sigma &lt; 250$</td>
<td>-0.12</td>
</tr>
<tr>
<td>$250 &lt; \sigma &lt; 315$</td>
<td>0.18</td>
</tr>
<tr>
<td>$315 &lt; \sigma &lt; 400$</td>
<td>0.86</td>
</tr>
</tbody>
</table>

$0.6 < z < 0.7$ | 0.94 | 5.51 | 1.638 |
$0.7 < z < 0.8$ | 0.86 | 3.36 | 1.638 |
of age, metallicity and $\alpha$/Fe to increase with galaxy velocity dispersion (see e.g. [Thomas et al., 2010]). Our analysis shows that these correlations are independent of redshift within the redshift range probed by BOSS, hence have been in place since at least $z = 0.7$. Evidence for this trend has already been noticed by [Ziegler et al., 2005] based on a small galaxy sample observed with FORS at the VLT. While the results in [Ziegler et al., 2005] only provided tentative evidence, the present study puts these findings on more robust statistical ground.

The correlation between metallicity and $\sigma$ is expected due to $\sigma$ being a tracer of the dynamical mass, and as such the potential well a galaxy is in. More massive galaxies have a deeper potential well, which leads to more efficient star formation - i.e. a higher fraction of the galaxy’s gas content is converted into stars - which results in a higher rate of gas processing and therefore metal production (e.g. [Thomas et al., 1999]). This efficiency is largely due to the increased energy required for supernovae ejections to escape the galaxy, meaning that gas that would otherwise have been expelled from the galaxy is instead retained, enriches the ISM and provides more gas for future star formation.

We also see that $\alpha$/Fe increases with $\sigma$, as has already been found in galaxies at $z \sim 0$ ([Thomas et al., 2005, 2010]). This is expected within the downsizing scenario due to the $\alpha$-elements and Fe being produced by different types of supernovae ([Matteucci, 1994; Thomas et al., 1999, 2005]). The $\alpha$-elements are formed by SN II, which act on very short timescales, whilst Fe is produced by SN Ia which act on very long timescales. Since in the downsizing paradigm (and according to our results) more massive galaxies have shorter star formation timescales, whilst SN II explosions happen early enough for their ejected gas to be used in the next generation of stars, enriching the metallicity of the stellar population, SN Ia explosions happen over a longer timescale and as such a smaller fraction of SN Ia have exploded before the most recent starburst than for their less massive counterparts. We would therefore expect to see that more massive galaxies have lower iron abundances but similar $\alpha$-element abundances, and therefore higher $\alpha$/Fe ratios, which is precisely what we observe.

Finally, galaxy age is found to correlate with velocity dispersion at high statistical significance, confirming the trend that more massive galaxies host older stellar populations ([Proctor & Sansom, 2002; Thomas et al., 2005, 2010] and references therein), in line with the higher $\alpha$/Fe ratios discussed above.
Evolution with redshift

The left column in Fig. 4.20 shows the evolution of the stellar population parameters age, metallicity and $\alpha$/Fe with redshift for five velocity dispersion bins as indicated by the labels. We see a clear relationship between age and redshift, which is an obvious result of galaxies getting older as the Universe ages. There is an interesting dependence on velocity dispersion, such that the evolution with redshift is steepest for the most massive galaxies (a factor of four from 8 to 4 Gyr), while the objects in the lowest $\sigma$ bin show considerably less evolution (factor two from 2 to 1 Gyr). Considering that the quantity measured here is light-averaged stellar population age, this pattern fits well into the downsizing picture: ongoing residual star formation activity in the lower-mass objects across the redshift interval probed maintains the light-averaged age at a low value at all look-back times. The objects in the $\sigma$-bin representing the most massive galaxies evolve passively and hence show the steepest evolution in light-averaged age. Metallicity and $\alpha$/Fe ratio, instead, show no significant sign of redshift evolution, at least within the measurement accuracy of this study.

The evolution of the light-averaged age can be better studied by direct comparison with the age of the universe at each epoch assuming a given cosmology. To this end Fig. 4.21 shows the difference between galaxy ages and the age of the Universe ($\Delta$age), with colours indicating $\sigma$ as shown by the labels. The age of the Universe was calculated assuming a standard, flat, $\Lambda$CDM cosmology with $H_0 = 71$ and $\Omega_m = 0.27$ (calculating using Ned Wright’s cosmology calculator, Wright (2006)).

A very interesting pattern emerges separating objects very clearly by galaxy mass (or velocity dispersion). The highest $\sigma$ bin ($315 < \sigma < 400$ km s$^{-1}$) shows clear evolution with $\Delta$age increasing as a function of redshift and look-back time from $\sim 3$ to $\sim 5$ Gyr. The lowest $\sigma$ bin ($100 < \sigma < 160$ km s$^{-1}$), instead, exhibits the opposite evolution $\Delta$age decreasing with increasing redshift and look-back time from $\sim 10$ to $\sim 6$ Gyr. Objects between these two extremes show very little or no evolution. This implies that massive galaxies appear to be ageing faster and low-mass galaxies appear to be ageing slower than the universe.

To understand this result, it is important to recall that we are measuring light-averaged age of the galaxy stellar populations which we then compare with the age of the universe at each given epoch. The light-averaged age of a galaxy (population) will change and may also decrease as a function of cosmic time
Figure 4.21: Difference between the ages of the stacked objects and that of the Universe according to a standard, flat, ΛCDM cosmology with $H_0 = 71$ and $\Omega_m = 0.27$ (calculating using Ned Wright’s cosmology calculator, Wright (2006)), plotted as a function of redshift.
depending on the star formation activity along the galaxy’s evolutionary path. To illustrate this, we have constructed toy star formation histories such that the evolution seen in Fig. 4.21 is reproduced. Using a variety of different star formation histories as input, we calculated the light-averaged age at each given cosmic time along the redshift interval covered by the BOSS data. The light-averaged age $\langle \text{age} \rangle_{lw}$ of the resulting composite stellar population at each time step $t$ is the integral from formation time $t_{\text{form}}$ to $t$ over the star formation rate $\psi$ weighted by the optical $M/L$ ratio of the stellar population depending on its age (see Eqn. 4.4).

$$\langle \text{age} \rangle_{lw}(t) = A \int_{t_{\text{form}}}^{t} \psi(t') \left( \frac{M}{L} \right)^{-1}(t') \, dt'$$  (4.4)

Metallicity is assumed solar for the sake of simplicity. The stellar population models by Maraston (2005) have been adopted to calculate $M/L$, and we have assumed a Salpeter (1955) IMF.

Fig. 4.22 shows the resulting plot. The BOSS data from Fig. 4.21 is now shown in grey, $\Delta_{\text{age}}$ as a function of redshift resulting from this modelling exercise is shown by the coloured lines. The corresponding star formation histories are shown as insets in the figure, and in more detail in Fig. 4.23. It can be seen that a decreasing $\Delta_{\text{age}}$ with cosmic time is best reproduced by a passively evolving stellar population (red line and histogram). The ages as derived from the BOSS data require some fraction of secondary star formation at a look-back time of $6 - 7$ Gyr close to the maximum redshift observed ($z \sim 0.7$) on top of an old population in order to recover the exact slope. This trajectory corresponds to the BOSS galaxy stack in the highest mass ($\sigma$) bin ($315 < \sigma < 400$ km s$^{-1}$).

The objects in the subsequently lower $\sigma$ bins ($200 < \sigma < 250$;km s$^{-1}$ and $250 < \sigma < 315$;km s$^{-1}$) show very little evolution in $\Delta_{\text{age}}$. The reason for this is that some residual star formation extending into the observed redshift interval flattens the decrease of the light-averaged galaxy age with time, and hence the evolution of $\Delta_{\text{age}}$. This trend continues for objects in the even lower $\sigma$ bins ($100 < \sigma < 160$;km s$^{-1}$ and $160 < \sigma < 200$;km s$^{-1}$), and the evolution of $\Delta_{\text{age}}$ with cosmic time even gets reversed. The star formation histories required to recover this reverse evolution are even more extended with significant star formation within the past $\sim 6$ Gyr (since $z \sim 0.6$). This continues supply of young stellar populations as the galaxy evolves leads to a very shallow evolution
Figure 4.22: Grey lines and colour coding as in Fig. 4.21. Light-averaged ages of model composite stellar populations with different underlying star formation histories as indicated by the line colours are overplotted. The corresponding star formation histories are indicated by the inset histograms following the colour coding of the lines. See also Fig. 4.23.

Figure 4.23: Colour coding as in Fig. 4.21. The various panels show the underlying star formation histories of the model composite stellar populations in Fig. 4.23.
of light-averaged age with time, which makes the universe appear to age much faster seen from the relatively steep rise of $\Delta_{\text{age}}$ with time.

**Manifestation of downsizing**

The result reported here is a further, very clear manifestation of downsizing. The lower the galaxy mass, the more extended its star formation history. The star formation histories derived here explicitly from the redshift evolution of galaxy spectra agree well with the archaeological approach of Thomas et al. (2010). The evolution of light-averaged galaxy ages derived here sets stringent constraints on galaxy formation modelling.

**4.2.4 Conclusion**

I have stacked BOSS spectra together in bins of redshift and stellar velocity dispersion ($\sigma$). I then ran our data analysis pipeline over these stacked objects, and measured their absorption line indices in the Lick system. I used these to derive stellar population parameters with the models of Thomas et al. (2011). I find that light-averaged age, metallicity and $\alpha/\text{Fe}$ all increase with $\sigma$, which are predictions of the downsizing paradigm, where the least massive galaxies form their stars later, over more extended timeframes and less efficiently than more massive galaxies. Age also increases with redshift, which is simply the result of everything in the Universe getting older, whilst I see no evidence of metallicity or $\alpha/\text{Fe}$ changing with lookback time.

Investigating how the light-averaged age of galaxies compares with the age of the Universe, we find that more massive galaxies appear to age faster than the Universe whilst less massive galaxies age slower. We hypothesise that this is due to the different star formation histories of galaxies with differing masses, and test this by compiling models with varying star formation histories and constraining them with our modelled observations. We find that as mass decreases, we require more extended periods of star formation that peak more recently. At the high-mass end, the relationship between the most massive bins is best reproduced by a passively evolving population whose stars formed at higher redshift than we observe. This is a clear result of downsizing, and sets tough restrictions on future models of galaxy formation and evolution.
4.3 Contributions to the collaboration

In addition to performing my own science, I have contributed a data product to the collaboration as a whole and co-written papers. The data product provided is a value-added catalogue, containing the majority of the outputs from our pipeline. This has been published in both DR9 and DR10. The paper most significantly contributed to to-date is Beifiori et al. (ApJ submitted), which investigates the evolution of galaxies’ dynamical properties with redshift.

4.3.1 DR9

For DR9 I produced a value-added catalogue named Portsmouth_emlinekin that contains emission line statistics and kinematics Ahn et al. (2012). This entailed giving the BOSS data pipeline team a version of our pipeline that worked on their servers, and testing every value created on our side with the values created on theirs to ensure a one-to-one relation. At this stage our output did not include robust error measurements on emission line values. Results from this pipeline were used in Comparat et al. (2013); Maraston et al. (2013); Thomas et al. (2013); Beifiori et al. (sub).

4.3.2 DR10

I provided a value-added catalogue for DR10 as well under the same name, with some improvements Ahn et al. (2013). Our DR10 pipeline included emission line value errors, as well as a vital fix for emission line measurements that were incorrect by a factor of $1 + z$ due to bugs in ‘vanilla’ GANDALF. Results from this pipeline will be used in Steele et al. (prep) to describe stellar population parameters, with work on this paper so far included above.

4.3.3 Beifiori et al. (sub)

Redshift Evolution of the Dynamical Properties of Massive Galaxies from SDSS-III/BOSS

Introduction and data

It has long been known that elliptical galaxies have tight correlations between their dynamical and stellar properties such as their size (effective radius, $R_e$),
surface brightness and stellar velocity dispersion ($\sigma_e$) (the fundamental plane, Dressler et al., 1987). This relation is a useful tool to study the co-evolution of baryonic and dark matter in galaxies. Importantly the evolution with time of this plane, dynamical and stellar properties can be used to help break the known degeneracy between dark matter fraction and IMF (Bender et al., 1998; Thomas et al., 2011), which is a vital step towards understanding of the true IMF. This paper aims to investigate the link between the dynamical and stellar properties of local and higher redshift galaxy populations, hereby contributing towards breaking this degeneracy.

The sample used was a combination of the main galaxy population of SDSS-III/BOSS DR9 covering a redshift range of $0.2 < z < 0.7$, combined with a local sample of massive galaxies at $z \sim 0.1$ taken from SDSS-II. The selection of SDSS-II galaxies was of early types following Hyde & Bernardi (2009).

Method

The fundamental properties investigated are the stellar masses from Maraston et al. (2013), the stellar velocity dispersions derived in this thesis, and effective radii derived using the SDSS-III DR8 pipeline (Aihara et al., 2011) and calibrated using COSMOS imaging. The dynamical mass was calculated following Beifiori et al. (2012) using the virial mass estimator $M_{\text{dyn}} = \beta(n) R_e \sigma_e^2 / G$ where $\beta$ is a dimensionless constant that depends on galaxy structure and $G$ is the gravitational constant.

The impacts of aperture, rotation, galaxy type and galaxy structure are all investigated and found to be either negligible or have no trend with redshift. The influence of effective radius used for dynamical mass calculations was looked at by comparing $M_{\text{dyn}}$ with SDSS and COSMOS $R_e$ values, and was found to be negligible. Progenitor bias was corrected for by removing objects that would not have had time to evolve into passive objects (3 Gyr) according to their age and redshift (i.e. young, low-z objects).

Results

Fig. 4.24 shows the redshift dependence of $R_e$ (left), $\sigma_e$ (centre) and $\log(M_{\text{dyn}}/M_\ast)$ (right), separated into $M_\ast -$ (top) and $M_{\text{dyn}} -$ (bottom) selected galaxies. The mass selection was designed to keep objects within $1\sigma$ of the mean mass in order to avoid selection effects as a function of $z$, and finer cuts were found to have no
Figure 4.24: Left panels: Effective radius $R_e$ as a function of redshift. Central panels: Stellar velocity dispersion $\sigma_e$ as a function of redshift. Right panels: Ratio between dynamical and stellar mass $M_{\text{dyn}}/M_*$ as a function of redshift. Top and bottom panels are for galaxies selected using $M_*$ and $M_{\text{dyn}}$, respectively (within $\pm 1\sigma$ of the mass distributions, total number of galaxies given by the labels). The shaded contour region indicates the full sample after correction for progenitor bias. Contours show 10 equally-spaced density levels showing the percentage of galaxies compared to the peak value of each plot. The coloured filled circles are the median values in four redshift bins. The green solid line is a linear fit, the green dashed line is a fit with zero slope for comparison. The black dotted line is a linear fit to the sample if no size correction is applied. Error bars of the median points indicate the $1\sigma$ uncertainty derived from Monte Carlo simulations. The red-continuous lines and arrows indicate the range where we fit our data. Image and caption taken from Beifiori et al. (sub).
significant impact on results. The shaded contour region shows the full sample, the coloured circles are the median values in four redshift bins, the solid green line is the linear fit and the green dashed line has zero slope for comparison. The black dotted line is the linear fit to the sample if no size correction is applied (i.e. no calibration using COSMOS). The red line and arrow indicates the range where the data are fit, with objects falling into the last redshift bin not accounted for due to the larger uncertainty in their radius calibration.

Galaxy radius decreases with redshift at $1.5\sigma$ significance, which agrees with the literature. Stellar velocity dispersion is found to mildly increase with redshift at $>2\sigma$ significance, again generally consistent with the literature although somewhat milder than has been found, possibly due to the narrow redshift range observed. This is in contention, however, with the models of Hopkins et al. (2009). The dynamical-to-stellar mass ratio shows a decrease with redshift at $>2\sigma$ significance, driven by the decrease in size which is not balanced out by the mild increase in velocity dispersion. This is again consistent with the literature. It is noted that at fixed $M_\star$, $M_{\text{dyn}}$ decreases with $z$ whereas at fixed $M_{\text{dyn}}$, $M_\star$ increases with $z$. Finally this ratio is investigated at greater redshift by utilising data from other studies, and is found to be decreasing with a roughly consistent slope out to $z < 2.18$.

It is suggested that this change in ratio is likely due to a decrease in dark matter fraction rather than a variable IMF, i.e. the dark matter fraction in massive galaxies within the half-light radius increases with cosmic time, as the study probes a well selected, passively evolving galaxy population consisting of low-$z$ massive galaxies and their progenitors. This is well understood through size growth from minor mergers, through which the effective radius increases significantly - increasing the area within which we measure dark matter in this case - whilst only marginally increasing stellar mass.

My contribution

I provided the stellar velocity dispersion measurements used by this paper to calculate $M_{\text{dyn}}$, advised the lead author on where to apply these measurements and when caution should be taken, and engaged in private discussion about the paper.
Chapter 5

Conclusions

In this work we address two of the big questions in modern astrophysics; the role of environment as a driver of galaxy evolution, and the role of mass in star formation and stellar population evolution. We use one of the most powerful tools available to us, large-scale galaxy spectroscopy, to contribute towards the answers to these dilemmas.

I have constructed a data analysis pipeline based on the public codes GANDALF and pPXF to extract gas and stellar dynamics, emission line statistics, absorption line indices and stellar population parameters from these galaxy spectra, as detailed in Chapter 2. I test and calibrate this pipeline against existing results for the SDSS DR7, and find it to provide accurate measurements. We then use the outputs from this to investigate galaxy evolution.

We utilise the GAMA survey to look at the impact of environment on emission line classes (Chapter 3). I use the emission line diagnostic of Baldwin et al. (1981) to separate between star forming objects, composites, Seyferts, LINERs and emission-free objects (passives), and look at the distributions of these as functions of stellar mass, local environmental density and global environment (group membership). I control for mass when investigating environment, and vice versa, in order to avoid confusion caused by a correlation between masses and environments. I find that mass is the main dividing feature between these classes, with local environment playing a minimal role, but showing some importance at high densities when comparing with passives. I further find that global environment has no significant impact on these distributions, either with mass or local density.

Putting this into the bigger picture, the trend with mass agrees with previous studies (e.g. Schawinski et al., 2007) that generally suggest it is a result of...
AGN feedback; i.e. the quenching of star formation by radiation emanating from accretion around the central supermassive black hole. We extend this by finding this feature to be independent of environment, a result expected if AGN activity is largely caused by internal processes and properties.

I then investigate the properties of the SDSS-III/BOSS survey and use it to look at the effect of emission line sources on target selection (Section 4.1). I find that the BOSS survey observes a fairly uniform mass distribution in the range $0.15 < z < 0.7$, in line with survey objectives, and that the S/N of the spectra is adequate to measure stellar velocity dispersions and emission line fluxes for individual objects. We then find that star forming objects, Seyferts and LINERs are all well separated above $z \sim 0.45$ on the $g - r$ vs $r - i$ diagram. We further see that the presence of star formation in low-mass, low-$z$ objects can cause them to shoot up into the region of this diagram typically considered to be populated by $z > 0.45$ objects.

I then stack these objects into bins of stellar velocity dispersion and redshift in order to accurately measure absorption line indices and convert these (via comparisons to the models of Thomas et al. 2011) into stellar population parameters, with the objective of looking into the dependence of stellar population parameters on mass and redshift (Section 4.2). I find that the S/N of our stacked objects is sufficient to measure the Lick indices with a simulated error of $\sim 5\%$.

I find that light-averaged age, metallicity and $\alpha$/Fe all increase with $\sigma$, which are predictions of the downsizing paradigm, where the least massive galaxies form their stars later, over more extended timeframes and less efficiently than more massive galaxies. Age also increases with redshift, which is simply the result of everything in the Universe getting older, whilst I see no evidence of metallicity or $\alpha$/Fe changing with lookback time.

Investigating how the light-averaged age of galaxies compares to the age of the Universe, we find that more massive galaxies appear to age faster than the Universe whilst less massive galaxies age slower. We hypothesise that this is due to the different star formation histories of galaxies with differing masses, and test this by compiling models with varying stellar histories and comparing them to our observations. We find that as mass decreases, we require more extended periods of star formation that peak more recently. At the high-mass end, the relationship between the most massive bins is best reproduced by a passively evolving population whose stars formed at higher redshift than we observe. This
is a clear result of downsizing, and sets tough restrictions on future models of galaxy formation and evolution.
Bibliography


Baugh C. M., 2006, Reports on Progress in Physics, 69, 3101


Beifiori A., Thomas, D. Maraston, C. et al. sub.
Ching J., Hardcastle, M. et al. in prep.
Faber S. M., Gallagher J. S., 1979, ARA&A, 17, 135
BIBLIOGRAPHY


Kant I., 1755, Allgemeine Naturgeschichte und Theorie des Himmels


Leavitt H. S., Pickering E. C., 1912, Harvard College Observatory Circular, 173, 1
Liske J., GAMA Collaboration et al. in prep.
BIBLIOGRAPHY

Peterson B. M., 1997, An Introduction to Active Galactic Nuclei


Slipher V. M., 1913, Lowell Observatory Bulletin, 2, 56


Steele O., Thomas D., Maraston C., Johansson J., et al. in prep.


Tonry J., Davis M., 1979, AJ, 84, 1511


Zwicky F., 1933, Helvetica Physica Acta, 6, 110
Appendix A

Catalogue products

This Appendix contains the data products released in our catalogues, along with their units, a brief description and a flag to cover which catalogue the product was released in. The values of the flags are: ‘A’ means all catalogues, ‘G’ means exclusive to the GAMA collaboration catalogues, ‘B9’ for BOSS DR9, ‘B10’ for BOSS DR10, ‘M’ for the MaNGA catalogue and ‘S’ for all SDSS catalogues.
<table>
<thead>
<tr>
<th>Name</th>
<th>Units</th>
<th>Flag</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objname</td>
<td></td>
<td>A</td>
<td>Unique name for each object, based on its name in the input survey</td>
</tr>
<tr>
<td>CATA</td>
<td></td>
<td>G</td>
<td>Unique index value for each object from the GAMA database</td>
</tr>
<tr>
<td>RA</td>
<td>deg</td>
<td>S</td>
<td>Right ascension</td>
</tr>
<tr>
<td>DEC</td>
<td>deg</td>
<td>S</td>
<td>Declination</td>
</tr>
<tr>
<td>BPT</td>
<td></td>
<td>A</td>
<td>Primary energetic process as determined from the Baldwin, Phillips and Terlevich (Baldwin et al., 1981) diagnostic diagram</td>
</tr>
<tr>
<td>A</td>
<td></td>
<td>A</td>
<td>Relative strength of emission lines if in a multiplet</td>
</tr>
<tr>
<td>V</td>
<td>km/s</td>
<td>A</td>
<td>Emission line velocity</td>
</tr>
<tr>
<td>V_Err</td>
<td>km/s</td>
<td>A</td>
<td>Error on emission line velocity</td>
</tr>
<tr>
<td>Sig</td>
<td>km/s</td>
<td>A</td>
<td>Velocity dispersion of emission lines</td>
</tr>
<tr>
<td>Sig_Err</td>
<td>km/s</td>
<td>A</td>
<td>Error on velocity dispersion of emission lines</td>
</tr>
<tr>
<td>Amp</td>
<td>erg/cm²/km/Å</td>
<td>A</td>
<td>Amplitude of emission lines</td>
</tr>
<tr>
<td>Amp_Err</td>
<td>erg/cm²/km/Å</td>
<td>A</td>
<td>Error on amplitude of emission lines</td>
</tr>
<tr>
<td>Flux</td>
<td>erg/s/cm²</td>
<td>A</td>
<td>Emission line flux</td>
</tr>
<tr>
<td>Flux_Err</td>
<td>erg/s/cm²</td>
<td>A</td>
<td>Error on emission line flux</td>
</tr>
<tr>
<td>Amp_Obs</td>
<td>erg/cm²/km/A</td>
<td>A</td>
<td>Observed (red) amplitude of emission lines</td>
</tr>
<tr>
<td>Flux_Obs</td>
<td>erg/s/cm²</td>
<td>A</td>
<td>Observed (red) flux of emission lines</td>
</tr>
<tr>
<td>Flux_C</td>
<td>erg/s/cm²</td>
<td>A</td>
<td>Continuum flux of emission lines</td>
</tr>
<tr>
<td>Flux_C_Err</td>
<td>erg/s/cm²</td>
<td>A</td>
<td>Error on emission line continuum flux</td>
</tr>
<tr>
<td>EW</td>
<td>Å</td>
<td>A</td>
<td>Equivalent Widths of emission lines (Flux_Obs / Flux_C)</td>
</tr>
<tr>
<td>EW_Err</td>
<td>Å</td>
<td>A</td>
<td>Error on EW of emission lines</td>
</tr>
<tr>
<td>AoN</td>
<td>Å</td>
<td>A</td>
<td>Amplitude over Noise of emission lines</td>
</tr>
<tr>
<td>EBMV</td>
<td></td>
<td>A</td>
<td>E(B-V) value for model fit (a measure of attenuation)</td>
</tr>
<tr>
<td>EBMV_Err</td>
<td></td>
<td>A</td>
<td>Error on E(B-V)</td>
</tr>
<tr>
<td>Dust</td>
<td></td>
<td>A</td>
<td>Dust attenuation for each emission line, calculated by applying E(B-V) to the Catzetti dust law (Calzetti, 2001)</td>
</tr>
<tr>
<td>Vel_Stars</td>
<td>km/s</td>
<td>A</td>
<td>Recessional velocity of the stellar population as a whole</td>
</tr>
<tr>
<td>Redshift</td>
<td></td>
<td>A</td>
<td>Spectroscopic redshift of object</td>
</tr>
<tr>
<td>SigStars</td>
<td>km/s</td>
<td>A</td>
<td>Stellar velocity dispersion</td>
</tr>
<tr>
<td>SigStars_Err</td>
<td>km/s</td>
<td>A</td>
<td>Error on Sigma_Stars</td>
</tr>
<tr>
<td>ChiSq</td>
<td></td>
<td>A</td>
<td>Chi-squared value of the template fit</td>
</tr>
</tbody>
</table>
Appendix B

Emission lines

A list of the emission lines analysed, their wavelengths, whether or not they were part of a multiplet (M), the relative strength of that line if part of a multiplet (A) and a flag indicating which survey they were used in. When looking at whether or not a line was treated as part of a multiplet, ‘s’ means it was not and ‘mn’ means it was where n indicates the index value of the strongest line in the multiplet. The values of the flags are ‘A’ for all surveys, ‘G’ for the GAMA survey and ‘S’ for all SDSS surveys.

Note that the only GAMA-specific line is a combination of the only two SDSS-specific lines, combined as GAMA resolution was unable to disentangle the lines.
<table>
<thead>
<tr>
<th>I</th>
<th>Name</th>
<th>λ (Å)</th>
<th>M</th>
<th>A</th>
<th>Flag</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HeII</td>
<td>3203.15</td>
<td>s</td>
<td>1.00</td>
<td>A</td>
</tr>
<tr>
<td>2</td>
<td>[NeV]</td>
<td>3345.81</td>
<td>s</td>
<td>1.00</td>
<td>A</td>
</tr>
<tr>
<td>3</td>
<td>[NeV]</td>
<td>3425.81</td>
<td>s</td>
<td>1.00</td>
<td>A</td>
</tr>
<tr>
<td>4</td>
<td>[OII]</td>
<td>3726.03</td>
<td>s</td>
<td>1.00</td>
<td>S</td>
</tr>
<tr>
<td>5</td>
<td>[OII]</td>
<td>3727.20</td>
<td>s</td>
<td>1.00</td>
<td>G</td>
</tr>
<tr>
<td>6</td>
<td>[OII]</td>
<td>3728.73</td>
<td>s</td>
<td>1.00</td>
<td>S</td>
</tr>
<tr>
<td>7</td>
<td>[NeIII]</td>
<td>3868.69</td>
<td>s</td>
<td>1.00</td>
<td>A</td>
</tr>
<tr>
<td>8</td>
<td>[NeIII]</td>
<td>3967.40</td>
<td>s</td>
<td>1.00</td>
<td>A</td>
</tr>
<tr>
<td>9</td>
<td>Hδ</td>
<td>3970.07</td>
<td>s</td>
<td>1.00</td>
<td>A</td>
</tr>
<tr>
<td>10</td>
<td>Hγ</td>
<td>4101.73</td>
<td>s</td>
<td>1.00</td>
<td>A</td>
</tr>
<tr>
<td>11</td>
<td>[OIII]</td>
<td>4340.46</td>
<td>s</td>
<td>1.00</td>
<td>A</td>
</tr>
<tr>
<td>12</td>
<td>[OIII]</td>
<td>4363.15</td>
<td>s</td>
<td>1.00</td>
<td>A</td>
</tr>
<tr>
<td>13</td>
<td>HeII</td>
<td>4685.74</td>
<td>s</td>
<td>1.00</td>
<td>A</td>
</tr>
<tr>
<td>14</td>
<td>[ArIV]</td>
<td>4711.30</td>
<td>s</td>
<td>1.00</td>
<td>A</td>
</tr>
<tr>
<td>15</td>
<td>[ArIV]</td>
<td>4730.74</td>
<td>s</td>
<td>1.00</td>
<td>A</td>
</tr>
<tr>
<td>16</td>
<td>Hβ</td>
<td>4861.32</td>
<td>s</td>
<td>1.00</td>
<td>A</td>
</tr>
<tr>
<td>17</td>
<td>[OII]</td>
<td>4958.83</td>
<td>s</td>
<td>1.00</td>
<td>A</td>
</tr>
<tr>
<td>18</td>
<td>[OIII]</td>
<td>5006.77</td>
<td>s</td>
<td>1.00</td>
<td>A</td>
</tr>
<tr>
<td>19</td>
<td>[Ni]</td>
<td>5197.90</td>
<td>s</td>
<td>1.00</td>
<td>A</td>
</tr>
<tr>
<td>20</td>
<td>[Ni]</td>
<td>5200.39</td>
<td>s</td>
<td>1.00</td>
<td>A</td>
</tr>
<tr>
<td>21</td>
<td>HeI</td>
<td>5875.60</td>
<td>s</td>
<td>1.00</td>
<td>A</td>
</tr>
<tr>
<td>22</td>
<td>[OI]</td>
<td>6300.20</td>
<td>s</td>
<td>1.00</td>
<td>A</td>
</tr>
<tr>
<td>23</td>
<td>[OI]</td>
<td>6363.67</td>
<td>s</td>
<td>1.00</td>
<td>A</td>
</tr>
<tr>
<td>24</td>
<td>[NII]</td>
<td>6547.96</td>
<td>s</td>
<td>1.00</td>
<td>A</td>
</tr>
<tr>
<td>25</td>
<td>Heα</td>
<td>6562.80</td>
<td>s</td>
<td>1.00</td>
<td>A</td>
</tr>
<tr>
<td>26</td>
<td>[NII]</td>
<td>6583.34</td>
<td>s</td>
<td>1.00</td>
<td>A</td>
</tr>
<tr>
<td>27</td>
<td>[SII]</td>
<td>6716.31</td>
<td>s</td>
<td>1.00</td>
<td>A</td>
</tr>
<tr>
<td>28</td>
<td>[SII]</td>
<td>6730.68</td>
<td>s</td>
<td>1.00</td>
<td>A</td>
</tr>
</tbody>
</table>
Appendix C

Lick indices

The Lick indices and the boundaries of their feature (Fb, Fr), blue pseudo-continuum (Bb, Br) and red pseudo-continuum (Rb, Rr) boundaries in units of Å.
<table>
<thead>
<tr>
<th>I</th>
<th>Name</th>
<th>Fb</th>
<th>Fr</th>
<th>Bb</th>
<th>Br</th>
<th>Rb</th>
<th>Rr</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Hδₐ</td>
<td>4083.500</td>
<td>4122.250</td>
<td>4041.600</td>
<td>4079.750</td>
<td>4128.500</td>
<td>4161.000</td>
</tr>
<tr>
<td>2</td>
<td>Hδₜ</td>
<td>4091.000</td>
<td>4112.250</td>
<td>4057.250</td>
<td>4088.500</td>
<td>4114.750</td>
<td>4137.250</td>
</tr>
<tr>
<td>3</td>
<td>Hγₐ</td>
<td>4319.750</td>
<td>4363.500</td>
<td>4283.500</td>
<td>4319.750</td>
<td>4367.250</td>
<td>4384.750</td>
</tr>
<tr>
<td>4</td>
<td>Hγₜ</td>
<td>4331.250</td>
<td>4352.250</td>
<td>4283.500</td>
<td>4319.750</td>
<td>4354.750</td>
<td>4384.750</td>
</tr>
<tr>
<td>5</td>
<td>CN₁</td>
<td>4142.125</td>
<td>4177.125</td>
<td>4080.125</td>
<td>4117.625</td>
<td>4244.125</td>
<td>4284.125</td>
</tr>
<tr>
<td>6</td>
<td>CN₂</td>
<td>4142.125</td>
<td>4177.125</td>
<td>4083.875</td>
<td>4096.375</td>
<td>4244.125</td>
<td>4284.125</td>
</tr>
<tr>
<td>7</td>
<td>Ca₄227</td>
<td>4222.250</td>
<td>4234.750</td>
<td>4211.000</td>
<td>4219.750</td>
<td>4241.000</td>
<td>4251.000</td>
</tr>
<tr>
<td>8</td>
<td>G4300</td>
<td>4281.375</td>
<td>4316.375</td>
<td>4266.375</td>
<td>4282.625</td>
<td>4318.875</td>
<td>4335.125</td>
</tr>
<tr>
<td>9</td>
<td>Fe₄383</td>
<td>4369.125</td>
<td>4420.375</td>
<td>4359.125</td>
<td>4370.375</td>
<td>4422.875</td>
<td>4455.375</td>
</tr>
<tr>
<td>10</td>
<td>Ca₄455</td>
<td>4452.125</td>
<td>4474.625</td>
<td>4445.875</td>
<td>4454.625</td>
<td>4477.125</td>
<td>4492.125</td>
</tr>
<tr>
<td>11</td>
<td>Fe₅453</td>
<td>4514.250</td>
<td>4559.250</td>
<td>4504.250</td>
<td>4514.250</td>
<td>4560.500</td>
<td>4579.250</td>
</tr>
<tr>
<td>12</td>
<td>Fe₆68</td>
<td>4634.000</td>
<td>4720.250</td>
<td>4611.500</td>
<td>4630.250</td>
<td>4742.750</td>
<td>4756.500</td>
</tr>
<tr>
<td>13</td>
<td>Hβ</td>
<td>4847.875</td>
<td>4876.625</td>
<td>4827.875</td>
<td>4847.875</td>
<td>4876.625</td>
<td>4891.625</td>
</tr>
<tr>
<td>14</td>
<td>Fe₅015</td>
<td>4977.750</td>
<td>5054.000</td>
<td>4946.500</td>
<td>4977.750</td>
<td>5054.000</td>
<td>5065.250</td>
</tr>
<tr>
<td>15</td>
<td>Mg₁</td>
<td>5069.125</td>
<td>5134.125</td>
<td>4895.125</td>
<td>4957.625</td>
<td>5301.125</td>
<td>5366.125</td>
</tr>
<tr>
<td>16</td>
<td>Mg₂</td>
<td>5154.125</td>
<td>5196.625</td>
<td>4895.125</td>
<td>4957.625</td>
<td>5301.125</td>
<td>5366.125</td>
</tr>
<tr>
<td>17</td>
<td>Mg₈₉</td>
<td>5160.125</td>
<td>5192.625</td>
<td>5142.625</td>
<td>5161.375</td>
<td>5191.375</td>
<td>5206.375</td>
</tr>
<tr>
<td>18</td>
<td>Fe₅270</td>
<td>5245.650</td>
<td>5285.650</td>
<td>5233.150</td>
<td>5248.150</td>
<td>5285.650</td>
<td>5318.150</td>
</tr>
<tr>
<td>19</td>
<td>Fe₅335</td>
<td>5312.125</td>
<td>5352.125</td>
<td>5304.625</td>
<td>5315.875</td>
<td>5353.375</td>
<td>5363.375</td>
</tr>
<tr>
<td>20</td>
<td>Fe₅406</td>
<td>5387.500</td>
<td>5415.000</td>
<td>5376.250</td>
<td>5387.500</td>
<td>5415.000</td>
<td>5425.000</td>
</tr>
<tr>
<td>21</td>
<td>Fe₅709</td>
<td>5696.625</td>
<td>5720.375</td>
<td>5672.875</td>
<td>5696.625</td>
<td>5722.875</td>
<td>5736.625</td>
</tr>
<tr>
<td>22</td>
<td>Fe₅782</td>
<td>5776.625</td>
<td>5796.625</td>
<td>5765.375</td>
<td>5775.375</td>
<td>5797.875</td>
<td>5811.625</td>
</tr>
<tr>
<td>23</td>
<td>NaD</td>
<td>5876.875</td>
<td>5909.375</td>
<td>5860.625</td>
<td>5875.625</td>
<td>5922.125</td>
<td>5948.125</td>
</tr>
<tr>
<td>24</td>
<td>TiO₁</td>
<td>5936.625</td>
<td>5994.125</td>
<td>5816.625</td>
<td>5849.125</td>
<td>6038.625</td>
<td>6103.625</td>
</tr>
<tr>
<td>25</td>
<td>TiO₂</td>
<td>6189.625</td>
<td>6272.125</td>
<td>6066.625</td>
<td>6141.625</td>
<td>6372.625</td>
<td>6415.125</td>
</tr>
</tbody>
</table>
Appendix D

Hα offset spectra

Images of spectra with measurements of log Hα EW from our pipeline that were different by 0.5 dex to those from the MPA-JHU catalogue when an amplitude-over-noise minimum requirement of 4 was applied. The black lines are the observed spectra, red lines are our best-fit models, and blue dashed lines show the location of the Hα emission line.
Appendix E

H$\beta$ offset spectra

Images of spectra with measurements of log H$\beta$ EW from our pipeline that were different by 0.5 dex to those from the MPA-JHU catalogue when an amplitude-over-noise minimum requirement of 4 was applied. The black lines are the observed spectra, red lines are our best-fit models, and blue dashed lines show the location of the H$\beta$ emission line.
APPENDIX E. H\(\beta\) OFFSET SPECTRA