THE UNIVERSITY *of* EDINBURGH

The Hidden Growth of Black Holes Across Cosmic Time

Manika Kaur Sidhu

Doctor of Philosophy
The University of Edinburgh
August 2023

Abstract

It is now universally accepted that galaxies grow their central supermassive black holes through active phases of accretion that produces an Active Galactic Nucleus (AGN). Using multiwavelength surveys, the demographics of these AGN and their host galaxies have been studied up to high redshifts, helping illuminate how accretion and the associated physics at all scales evolves from late times to the earlier universe. However, the extent of obscured black hole growth over cosmic time is not yet fully understood due to the difficulty in identifying AGN down to low-luminosities and accurately determining their obscuration properties.

Focusing on the lower redshift universe, I present my work on measuring the independent X-ray luminosity functions of the broad-line and narrow-line (unobscured and obscured) AGN populations at z $\sim$ 0.2 – 1 to determine the extent of obscured black hole growth, relative to unobscured, at these cosmic times. I take X-ray and optical spectroscopic data from the 80 month NuSTAR Serendipitous Survey catalogue, defining a sample of $\sim$ 100 AGN at z $\sim$ 0.2 – 1 detected in the hard (8 - 24 keV) band and $\sim$ 200 AGN at z $\sim$ 0.2 – 1 detected in the full (3 - 24 keV) band. Accounting for incompleteness in the X-ray survey and counterpart identification, I use a Maximum Likelihood Fitting method to generate new measurements of the obscured, unobscured and combined AGN XLFs at z $\sim$ 0.2 – 1. With these results, I determine new measurements of the fraction of obscured SMBH growth as a function of X-ray Luminosity, at these lower redshifts.

I then shift my focus to the early universe and carry out a higher-z, complementary and more in-depth analysis to measure the independent X-ray luminosity functions of the broad-line (unobscured) and narrow-line (obscured) AGN population at z $\sim$ 2 - 4. These measurements allow me to determine the extent of obscured black growth relative to unobscured at these earlier cosmic times. An optical colour pre-selection approach coupled with X-ray imaging from Chandra

and XMM-Newton was used to identify a sample of ∼137 X-ray selected AGN at z ∼ 2-4. Optical spectroscopic follow-up data were obtained using LRIS on the Keck telescope, enabling us to probe both faint and obscured sources. An automated spectral fitting code was developed, which classifies our AGN sources as either broad-line or non-broad-line and accurately identifies their redshift. This code is also applied to large samples of simulated data, allowing us to establish the incompleteness in our optical follow-up and classifications. These completeness corrections are applied to my results to recover X-ray luminosity functions and obtain a robust measurement of the distribution of obscured and unobscured black hole growth in the early universe.

My new XLF measurements show an independent evolution for the obscured and unobscured AGN populations across $0.2 < z < 4$. We see an increase in the number density of both the obscured and unobscured AGN populations with redshift with the majority being the faint and obscured AGN population; the rate of change of this increase is larger for the mid-luminosity unobscured AGN population than it is for their obscured counterparts. My results show the obscured and unobscured AGN XLFs to become comparable at higher redshifts, contrary to other studies. This suggests relatively equal numbers of obscured and unobscured AGN at z ∼ 2 - 4, followed by an increase in the number density of the faintest obscured and unobscured AGN, and a decrease in the number density of the intermediate-luminosity and the brightest AGN as we shift to intermediate redshifts. Shifting then to more local times, the number density of obscured and unobscured AGN decreases across all luminosities. Another key result of mine is an increase in the incidence of faint AGN at high redshift when compared to previous work. This is due to the addition of careful completeness corrections incorporated throughout my XLF calculations, which highlights the importance of such methods.

My obscured AGN fraction results show a mild evolution of the obscured AGN fraction as a function of X-ray luminosity, with redshift; a minor obscured fraction-luminosity dependency at intermediate redshifts, followed by no substantial dependence of obscured fraction on luminosity at higher redshifts. Combined with results from comparable local redshift studies, my work shows the following evolutionary picture: (i) a strong luminosity dependence in the obscured AGN fraction in the local universe, with high fractions of the faint AGN population being obscured and low fractions of the bright AGN population being obscured. (ii) pushing out to intermediate redshifts, we see the same luminosity dependence

but to a lesser extent, with a slight increase in the fraction of faint, obscured AGN. (iii) Arriving at cosmic noon, we see equal fractions of obscured and unobscured AGN across X-ray luminosity. The lack of luminosity dependence on the obscured AGN fraction seen at higher redshifts could go to support the unified AGN model. However, the evolution we see as we move to lower redshifts supports a receding torus model; the decrease in obscured fraction with increasing X-ray luminosity could be due to the more powerful AGN sublimating their obscuring material and becoming less obscured over cosmic time. When considered together, these results point to the possibility of there being evolutionary phases of AGN obscuration over time.

The combination of my low-redshift and high-redshift analysis has allowed me to paint an updated picture of how the obscured AGN fraction evolves as a function of X-ray luminosity, over a broad cosmic time period, in an effort to determine the role obscuration may play in AGN growth and evolution.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(*Manika Kaur Sidhu, August 2023*)

Acknowledgements

This thesis would not have been possible without the help and support of several individuals who have all, in their own way, helped bring this work to completion; it is with great pleasure that I thank those who made this possible.

I am indebted to my supervisor, Dr. James Aird. His guidance, expertise and endless understanding throughout have been the cornerstones of this work and my PhD. I would also like to thank Professor Philip Best and Dr. Mike Watson for their invaluable advice and encouragement during both halves of my PhD. Last, but by no means least, I would like to show my gratitude to both the University of Edinburgh and the University of Leicester. Having started my PhD at the latter and finished it at the former, the wonderful experiences I have had at both universities have led me to where I am today.

From all of the amazing people I have the pleasure of calling my friends, I would like to thank a few in particular. A huge thank you is extended to Nisha, Harry, The Beach Crew, The Board games Crew, The Cafe Club, Les Tres Amigas and The Sisterhood of Sisters (now Doctors!) - thank you for all the love and laughs. I would also like to thank my housemates Jack and Cass for all the time spent playing Elden Ring, watching ridiculous movies and videos, exploring Edinburgh and having deep chats. A particular note of thanks goes to Jack, who has willingly been my best friend throughout some of the most dramatic years of my life.

Without the ongoing support of a few special people, completing this PhD would have felt impossible. I would like to thank my siblings and my best friend Cathye for their unconditional love and understanding over the years. I would especially like to thank my older sister. Penji - you have been a pillar of strength for me and I am infinitely proud and thankful that I get to call you my sister. The most significant acknowledgement is reserved for my parents, for it is with their lifelong love and support that I have been able to achieve everything that I have and become everything that I am. It is with great pride and a feeling of immense accomplishment, that I dedicate this thesis to my Mum and Dad.

Lay Summary

Our Universe is made up of a multitude of systems; planets, stars, galaxies and every bit of gas and dust in between, are all in motion under the influence of forces invisible to our eyes. Like our orbiting moon causes the tides here on Earth, many astronomical systems can be shaped and influenced by each other and, quite often, these interactions have been found to play an important role in their formation and growth over time. Understanding how everything in our observable Universe interacts with each other, is a key goal for many (if not all) astronomers.

Our planet lies within a system of planets, all orbiting our Sun - the Solar System. Lying on the outskirts, ours is just one of millions of other such systems hurtling round in our galaxy, the Milky Way. But, if we are supposedly on the outer edge of our Galaxy, what is at the center? It is now widely accepted in the scientific community that, at the center of our galaxy (and most galaxies) lies a supermassive black hole, a term often shortened to SMBH. These astronomical objects can grow to be millions (or even billions) times the mass of our Sun, and have such a strong gravitational pull that everything that gets too close is dragged inwards and ripped apart - it is this process that causes the SMBH to grow over time. During such active phases of growth, SMBHs emit huge amounts of radiation, across a wide range of energies. By using sophisticated ground and space based telescopes, we can observe the radiation emitted by these active SMBHs (called Active Galactic Nuclei, or AGN for short), and study their growth and activity over a large range of the Universe's history.

Unfortunately, the majority of these AGN are very faint and are often obscured by gas and dust surrounding them. This makes them very difficult to observe and, consequently, less is known about the growth and evolution of the faint and obscured AGN population. In this thesis, I aim to uncover some of the hidden growth of this elusive population; using combinations of data from leading ground and spaced based observatories, I undertake careful calculations and use thorough statistical analytical methods to generate updated measurements of faint and obscured AGN growth over cosmic time. These results go to further our knowledge on these obscured SMBHs, giving us an updated understanding on the growth and evolution of these hidden, gravitational giants.

Some of the data presented herein were obtained at Keck Observatory, which is a private 501(c)3 non-profit organization operated as a scientific partnership among the California Institute of Technology, the University of California, and the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation. The authors wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Maunakea has always had within the Native Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain.

Contents

Publications:

The X-ray Luminosity Functions of Obscured and Unobscured AGN at z 2 – 4 (Sidhu et al in preparation)

Extended Analysis of the NuSTAR Serendipitous Survey: the AGN X-ray Luminosity Function at lower redshifts (z: 0.2 – 1) (Sidhu et al in preparation)

Chapter 1

Introduction

The Universe as we know it is a vast and complex expanse comprised of a multitude of structures and substructures, with some evolving and influencing others. From the galactic clusters and superclusters that populate the dark matter filaments stretching out to the edges of our observable universe, to the individual stars, planets and black holes which make up our known galaxies, every system has its own part to play. It is now accepted that at the centre of most of these galaxies lies a supermassive black hole (SMBH). These black holes, with masses of $\sim 10^6$ - 10^{10} $M_\odot$, are thought to directly influence the growth and evolution of their host galaxies (King & Pounds, 2015; Magorrian et al., 1998; Vietri et al., 2022). As understanding the atom and its constituent quarks was key to furthering our knowledge of matter, unravelling how these supermassive black holes develop and, subsequently, affect their hosts is fundamental if we truly want to perceive the workings of the universe.

1.1 Supermassive Black Holes and Active Galactic Nuclei

Supermassive black holes reside in the centre of most galaxies. These black holes grow through active phases of accretion, during which they are known as 'Active Galactic Nuclei' (AGN for short). It is this accretion process that fuels AGN, resulting in powerful emission across the entire electromagnetic spectrum; this emission can out-shine the host galaxy at certain wavelengths and gives rise to

the 'star-like' appearance of some AGN. On the other hand, when central SMBHs are dormant, no accretion takes place as there is no matter in a close enough orbit to the black hole. In these dormant phases they remain relatively silent and undetected, with little emission. The SMBH at the middle of our Galaxy, the Milky Way, is one such black hole: Sagittarius A* (Sgr A*). Lynden-Bell & Rees (1971) first hypothesized that the center of the Milky Way galaxy would contain such a massive black hole. Sgr A* was first observed by Balick & Brown (1974) using data from the Green Bank Interferometer of the National Radio Astronomy Observatory (Balick & Brown, 1974). They observed a central radio source which emits synchrotron radiation, found to be dense and immobile because of its gravitation. This was the first indication that a supermassive black hole exists in the center of the Milky Way. Since then, Sgr A* has been studied numerous times; one of the more recent and notable results being the release of the first image of its accretion disk, by the Event Horizon Telescope Collaboration, confirming it to be a black hole (Event Horizon Telescope Collaboration et al., 2022).

As AGN showcase the phases of SMBH growth, it is the active population of black holes that are focused on in this Thesis. The exact structure of an AGN is still not completely known, with variation in the standard model across the board. There are, however, some defining features that can be attributed to the central engine of an active galaxy.

The central SMBH is surrounded by an accretion disk, comprised of matter that has orbited too close to the black hole. If an astrophysical object (a star, for example) enters this radius, it breaks apart and joins/forms the disk. When a particle enters the accretion disk, loss of angular momentum through viscous heating in the disk causes it to spiral inwards. The change in gravitational potential energy experienced as a result of this lower orbit is converted to kinetic energy. i.e. the velocity at which the particle is travelling increases. Hence, matter in an accretion disk continuously loses angular momentum and spirals in, whilst increasing in velocity. The resulting increase in friction between particles heats the accretion disk (Peterson, 1997). This hot accretion disk has the energy to photoionise surrounding areas, effectively producing an ionising spectrum which is incident on, and interacts with, surrounding regions.

A corona of high energy particles occupies the area above the accretion disk. The strong X-ray emission that is universally seen in AGN is thought to originate from this region. The hot electrons in the corona and low energy photons from the

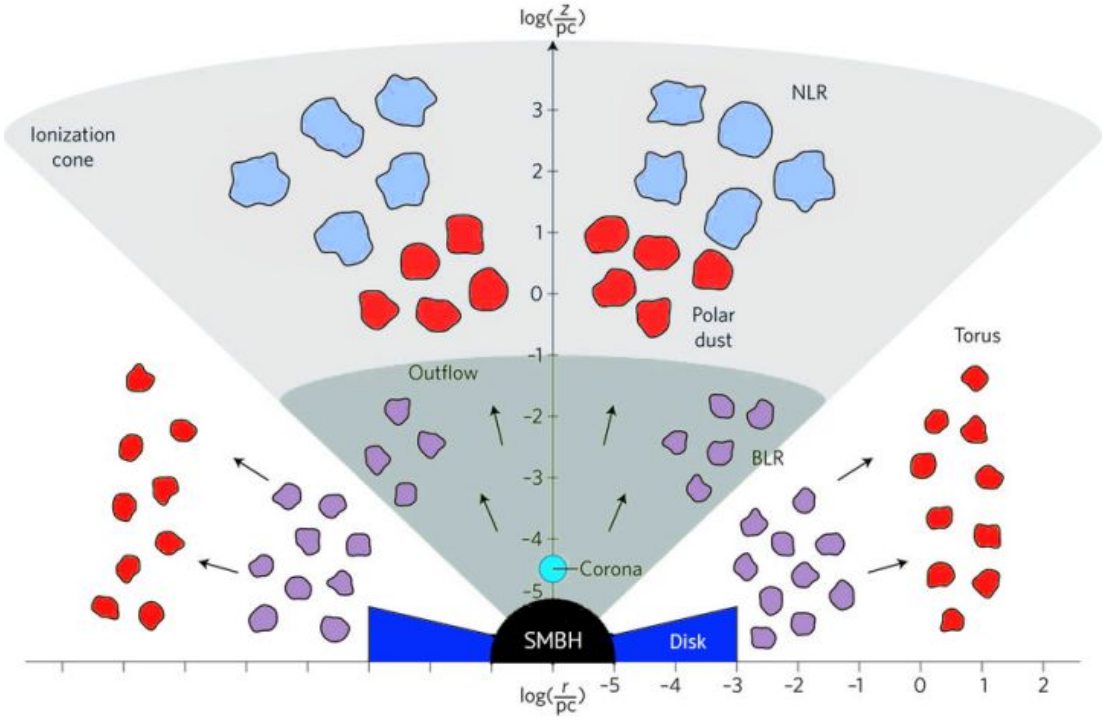

Figure 1.1 *Taken from Hickox & Alexander (2018): "Schematic representation of the AGN physical model, illustrating the broad scales of the key regions. The accretion disk, corona, broad-line region (BLR), and the dusty torus reside within the gravitational influence of the SMBH. The narrow-line region (NLR) is on a larger scale and under the gravitational influence of the host galaxy."*

accretion disk experience Inverse-Compton Scattering[1]; the photons are scattered to higher energies by these relativistic corona-electrons, generating power-law emission. This mechanism is shown by Figure 1.2. This figure illustrates just one possible geometry for the X-ray emitting region of AGN; the exact geometrical structure of the corona is still unclear. Another example is discussed in Brandt and Alexander (2015), which is the suggestion that the corona lies perpendicular to the accretion disk plane and stretches out for 10s of gravitational radii. In this way, X-rays that are 'Compton reflected'[2] off the rest of the nuclear region could also be scattered by the accretion disk and obscuring matter around the SMBH. Figure 1.3 shows an example model AGN X-ray spectrum, with all of its components, taken from Aird et al. (2015). The assumed intrinsic X-ray spectrum, illustrated by the solid purple line, is a power law with photon index

[1]In inverse-Compton scattering, highly energetic electrons transfer energy to photons through collisions. In normal Compton scattering, it is energetic photons which transfer energy to stationary electrons through collisions.

[2]Compton reflection: recognizable emission due to a gas or plasma's Compton scattering. The term used for the spectral signature observed in AGN electromagnetic radiation. This 'Compton reflection' spectral feature is characterised by a 'hump' in an AGN's observed spectral energy distribution (SED).

$\Gamma = 1.9 \pm 0.2$. The absorption component, illustrated by the red dashed line, is attributed to their adopted intervening column density of $N_H = 5 \times 10^{22}$ cm^{-2}. The scattered component, illustrated by the blue dot-dashed line, is attributed to a fraction of the intrinsic X-ray spectrum which is scattered, unabsorbed, into the line-of-sight (Aird et al. (2015) adopt a 2% scattered fraction). The reflection component, illustrated by the green dotted line, is attributed to Compton-reflection from cold, optically thick matter (e.g. an accretion disk or obscuring torus). The resulting total observed X-ray spectrum is illustrated by the solid black line, and the grey region illustrates the 95% confidence interval on their spectrum (Aird et al., 2015). This figure aptly illustrates the X-ray energy signature, and all relevant components, of an AGN.

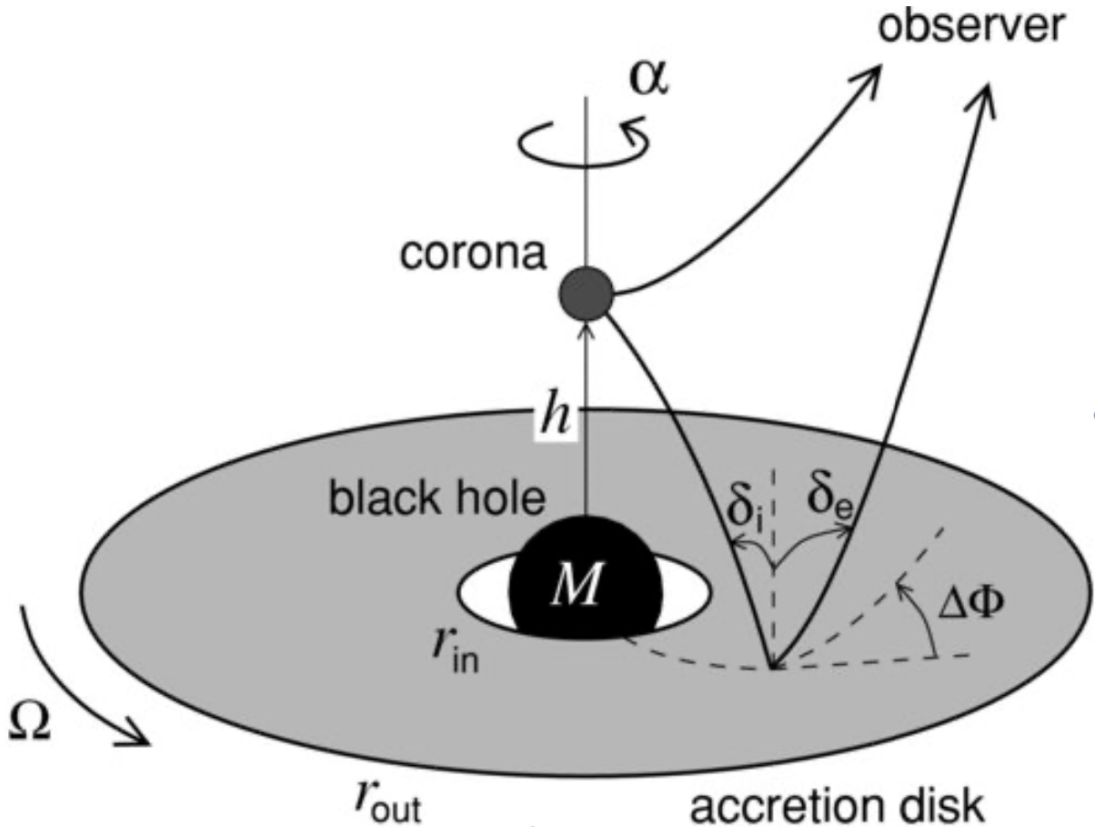

Figure 1.2 *Diagram showing the way in which X-ray emission is reflected off an accretion disk from a coronal source at height h above a black hole, taken from Dovčiak et al. (2011). An observer receives unpolarised power-law radiation from the source in addition to the reflected, polarised light from the disk.*

Around the accretion disk of an AGN are areas of gas, shown in Figure 1.1 as the Broad Line Region and the Narrow Line Region (BLR and NLR respectively). The BLR, lying closer (0.01 – 1 pc) to the central black hole, is comprised of high-density ($n_e \sim 10^9$ cm^{-3}) ionised gas. The NLR, which lies further out than the BLR, is comprised of low-density ($n_e \sim 10^3$ - 10^6 cm^{-3}) ionised gas. These

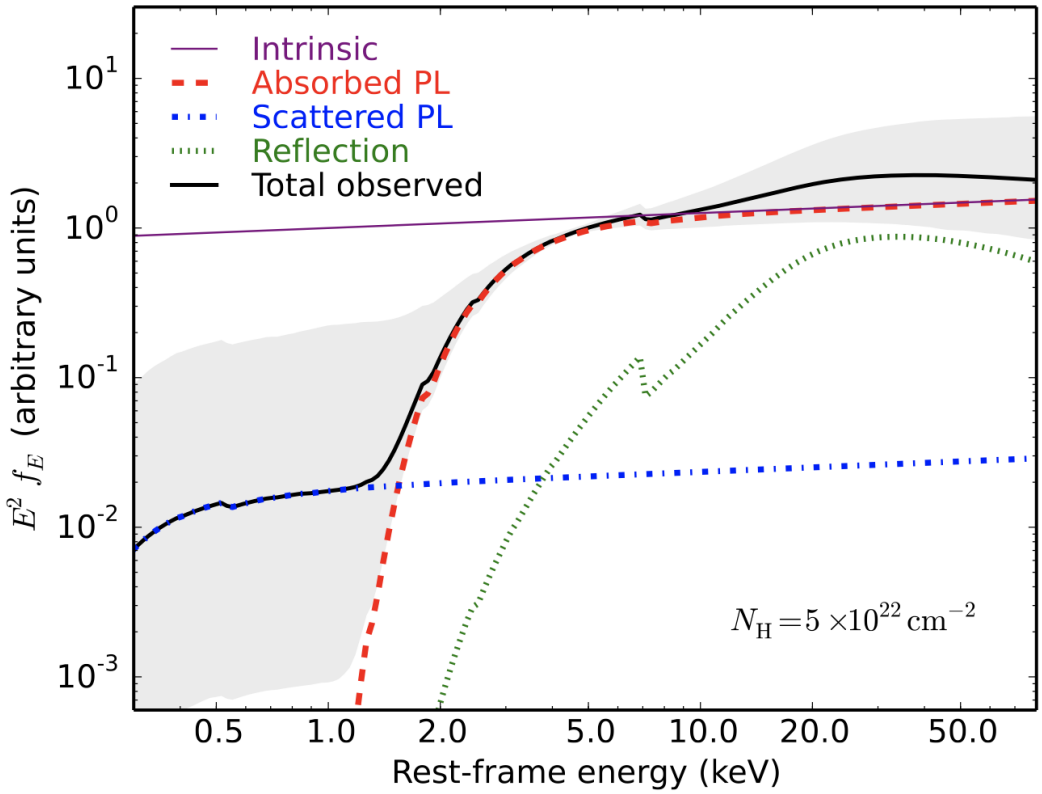

Figure 1.3 *Taken from Aird et al. (2015), an example AGN X-ray spectral model. The solid purple line is their assumed intrinsic X-ray spectrum, characterised by a power law with photon index* $\Gamma = 1.9 \pm 0.2$. *The red dashed line is the absorption component attributed to the intervening column density of molecular hydrogen (fixed at* $N_H = 5 \times 10^{22}$ cm^{-2} *in their example). The blue dot-dashed line is their adopted scattered, unabsorbed component amounting to 2% of the intrinsic X-ray power law, which is subsequently scattered into the line-of-sight. The green dotted line is their adopted reflection component, accounting for Compton-reflection from cold and optically thick matter, such as an accretion disc or obscuring torus. The black line is their total observed AGN X-rau spectrum, with the grey region indicating the 95% confidence interval on their spectrum.(Aird et al., 2015)*

regions are photoionised by the hot accretion disk through the excitation and de-excitation of electrons in the gas, resulting in photon emission over a range of energies. This 'ionising spectrum' from the accretion disk, which is incident on the broad line and narrow line region gas, produces the emission lines we see in spectra across a range of wavelengths: 'broad' emission lines which arise from transitions happening in the high density BLR, with line widths corresponding to velocities ranging from $\sim 10^3$ - 10^4 km s^{-1}; and 'narrow' emission lines which arise from transitions happening in the low density NLR, with line widths corresponding to velocities ranging $\sim$ 300 - 1000 km s^{-1}. On a spectrum, these lines correspond to electron transitions in various elements existing in the gas. An example of a broad emission line is H-α, with a vacuum wavelength of 6563 Å (Morton,

1991), which is a spectral line resulting from electron transitions in hydrogen gas localised in the BLR. An example of narrow emission lines is the OIII doublet, with vacuum wavelengths 4959 and 5007 Å (Morton, 1991), which is a spectral doublet resulting from electron transitions in oxygen gas localised in the NLR. The energy distribution of these spectra fluctuates with time, as a varying AGN accretion rate alters their shape and intensity (Peterson, 1997).

An obscuring, axisymmetric, dusty structure encompasses the BLR, with dimensions of 0.1 – 10 pc (see Figure 1.1). Its exact shape is contended to this day, but it is generally referred to as a torus. Gas situated at the inner radius of the torus is ionised by the central active nucleus, with dusty molecular gas lying deeper into the torus as you work your way out from the central BH. This dusty, molecular torus has an obscuring nature in that, if an AGN is viewed from certain angles, its molecular torus can 'hide/obscure' the BLR and central SMBH - features which are both thought to exist inside the torus. The NLR on the other hand lies externally to the torus, extending from just outside the torus to hundreds/thousands of parsecs along the general direction of the torus' opening. This concept is supported by the variety of spectra obtained from AGN: the 'unobscured' AGN display both broad and narrow emission lines, whilst 'obscured' spectra only display narrow emission lines. This observing angle/obscuration connection is touched on more in Section 1.3.

In its simplest and earliest form, the matter in the torus is modelled as smooth. However, the uncertainty surrounding the obscuring torus has continued over the years, with additional theories proposed to describe its structure and explain observations. The observations of today favor a more complex structure; more recent studies suggest that, contrary to being smooth, this obscuring material forms clumpy, elongated structures (Elitzur & Shlosman, 2006; Khim & Yi, 2017). How far this obscuring material extends has been linked to the luminosity of the AGN itself, with the 'receding torus' model of Lawrence (1991) implying that AGN with high enough X-ray luminosities can actually heat and sublimate the surrounding obscuring, dusty material. The observational effects of such a structure existing around a central black hole and accretion disk, coupled with the contention surrounding the nature of this obscuring medium, presents AGN obscuration as an interesting astronomical issue to further explore.

1.2 AGN and their Host Galaxies

It is important to not focus solely on the immediate surroundings of galactic supermassive black holes, but also how they fit in and interact with their host galaxies. The nuclei of active galaxies are contained in the galactic bulge: a central region primarily comprising gas, dust and old stars. Some (but not all) galaxies then have a flattened galactic disk surrounding the bulge. This disk is similar to the bulge in that it also contains dust and gas, however, it is in this region that younger and bluer stars are found. Finally, there is the galactic halo; a mixture of dust, clusters of old stars and dark matter which encompasses the whole galaxy.

There has been a great deal of research into the ways in which supermassive black holes in the active nuclei of galaxies are linked with and influence their hosts. One of the first results, put forward by Magorrian et al. (1998), was a positive correlation found between the mass of the black hole and the mass of the surrounding galactic bulge. This relationship was re-examined multiple times with similar results obtained (Häring & Rix, 2004; Marconi & Hunt, 2003). Another, more widely acknowledged, relationship is the positive correlation found between the black hole mass and the stellar velocity dispersion of the bulge: the M - σ relation (Ferrarese & Merritt, 2000; Gebhardt et al., 2000; Marsden et al., 2020).

Despite these results, when calculating the region over which the black hole directly influences, its scope is seen to be virtually insignificant in comparison to the scale of the host galaxy. However, upon comparing the binding energies of the black hole and the bulge, the black hole's is significantly larger. Hence, the bulge may be influenced by the presence of the black hole as a consequence of the considerable amounts of energy it generates through the process of accretion. This translates into a potential relationship and interaction between central, active SMBHs and their galaxies; the black hole grows through the influx of accreting matter, it then releases large amounts of energy into the host galaxy (against the direction of accretion), effectively stifling the growth of both the galaxy and the SMBH (Silk & Rees, 1998). Note that this outflow of energy from the black hole cannot be efficient - the energy discharged is much greater than the binding energy of the bulge, and would be enough to disrupt its entire structure, possibly even tearing it apart. This begs the question of how any of this energy is actually 'transmitted' to the bulge. There have been many proposed ideas for how this

AGN feedback is conveyed: for example, near-isotropic fast molecular outflows (momentum or energy driven) which drive away bulge gas; or through a series of shocks which heat and cool the medium (King & Pounds, 2015). Another form of feedback via radio jets, sometimes present in AGN, heating the hot intraclustar gas (Fabian, 2012).

In spite of this uncertainty in the exact mechanisms which govern AGN feedback, its importance and astrophysical role is now widely acknowledged. Large simulations created to model galaxy evolution have shown that they actually *require* AGN feedback to shut down star formation in the most massive galaxies (Booth & Schaye, 2009; Croton et al., 2006; Dubois et al., 2012; Richardson et al., 2016); otherwise, galaxies and SMBHs would keep growing indefinitely, reaching masses which have not currently been observed. Figure 1.4, taken from a paper written by Silk & Mamon (2012), illustrates this idea that AGN feedback must play a key role in galaxy evolution; CDM-motivated theory (the prevailing theory where the assembly of structure is primarily governed by the gravitational interaction of cold dark matter, with galaxies themselves forming at the centres of the dark matter overdensities), when used alone, leads to an overestimate of both ends of the galaxy luminosity function, as illustrated by the red line in Figure 1.4. This means something (AGN feedback) must happen galaxy-wide that limits galactic growth along its evolutionary track at high masses, in order for theory to align with observations (blue line in Figure 1.4). The necessity of this AGN feedback, specifically how SMBH growth effects and shapes that of its galaxy, decidedly demonstrates the importance of understanding AGN evolution carefully and completely.

1.3 The Unified Model and Obscured AGN

AGN as we observe them appear to exhibit a range of general (e.g. radio loudness) and spectral characteristics (e.g. presence of certain optical emission lines). These distinctions have led to the classification of AGN into specific 'types', as shown by Figure 1.6. However, various studies in the 1990s suggested a way to unite all AGN types: the Unified model of AGN (Antonucci, 1993; Urry, 2003; Urry & Padovani, 1995). This heavily debated model proposes all types of AGN to essentially be the same, with the distinction between them arising from an orientation effect with respect to the observer's line of sight; this effectively suggests that the large variety of AGN features can be explained

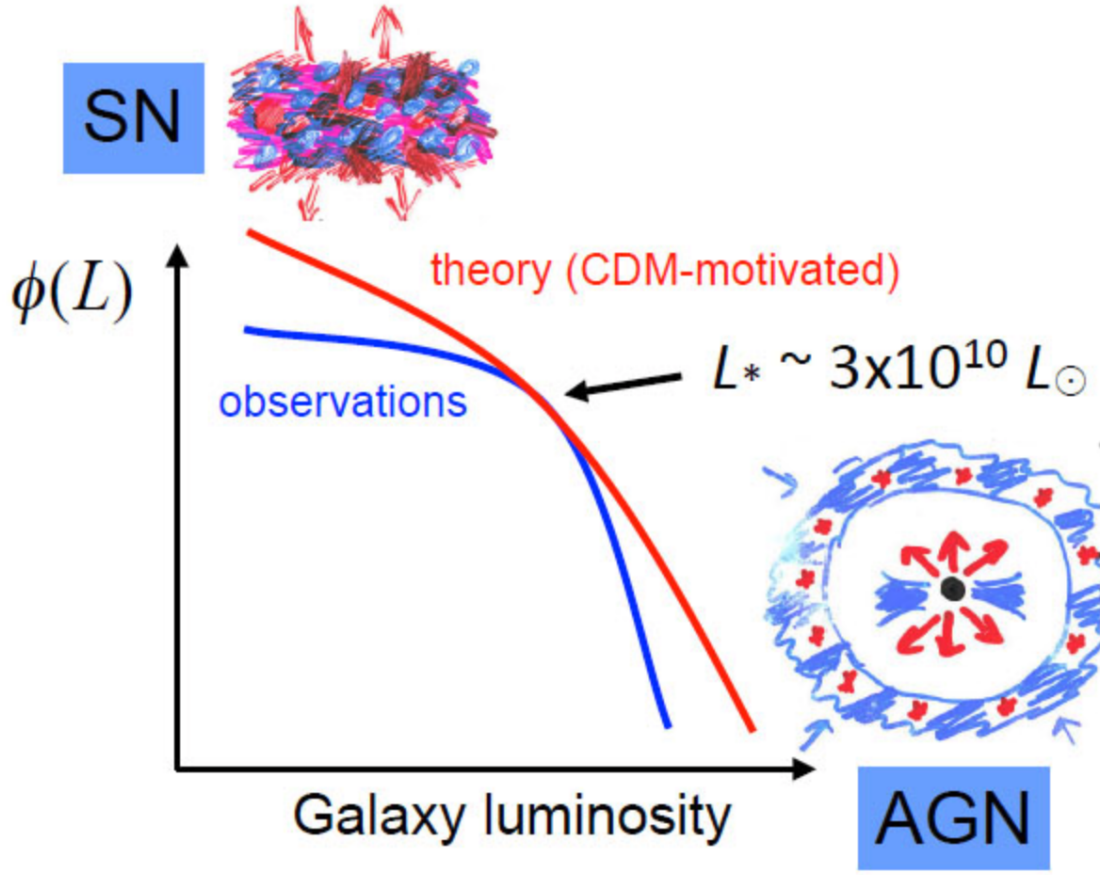

Figure 1.4 *Diagram, taken from Silk & Mamon (2012), illustrating the role of AGN feedback in modifying the galaxy luminosity function. The solid red line illustrates the shape of the galaxy luminosity function predicted by CDM-motivated theory alone. The solid blue line illustrates the observed shape of the galaxy luminosity function. This figure highlights how AGN feedback is a necessary addition to models to align theory with observations.*

by just a few physical parameters. The older unified model described in the aforementioned papers tries to unite all AGN subsets into a generic model with two main specifications: the luminosity of the source, and the inclination of the torus to the observer's line of sight. Figure 1.5 illustrates this model, with green labels indicating the viewing angles respective to the AGN necessary to result in these classifications.

To put the unified model into the context of varying AGN classifications, let us focus on two particular types of AGN: Seyfert (Sy) I and Seyfert (Sy) IIs, often called Type I and Type II AGN. Sy I and II galaxies differ only in the emission lines seen in their UV-optical spectra; Sy I spectra contain broad and narrow emission lines, whilst Sy II spectra show narrow lines, showing broad lines *only* in polarised light. This is one of the key pieces of evidence for the unified model, and reinforces the idea that Type I and Type II AGN are only classed as different

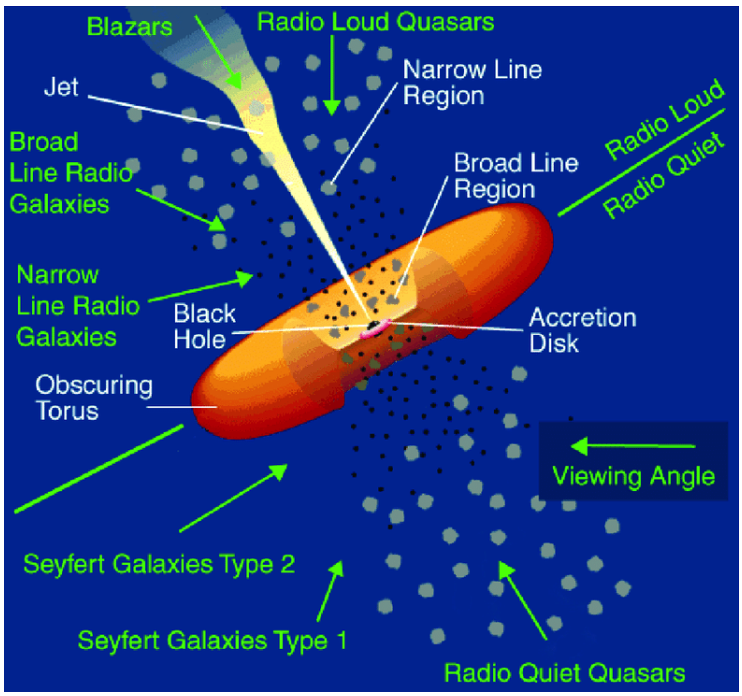

Figure 1.5 *Illustrative model showing the defining features of the active nuclear region of an active galaxy, suggested by (Urry & Padovani, 1995) in their 1995 paper on Unified Schemes for Radio-Loud Active Galactic Nuclei. The original diagram has been revised, with the addition of annotations demonstrating the idea behind the geometrically dependant 'Unified Model of AGN'; all AGN types could effectively be the same, if you view them from a specific angle. White labels are AGN features, and green labels are AGN types.*

due to obscuration of the BLR by the dusty torus structure. Thus, Sy/Type IIs are galaxies viewed along the toroidal plane so only narrow lines are visible in their spectra (the broad lines are 'obscured'), and Sy/Type Is are galaxies viewed perpendicular to the toroidal place, leaving the BLR unobscured so broad lines are seen in their spectra. NLRGs and BLRGS (Narrow and Broad line radio galaxies) are then the radio loud analogues of Sy Is and IIs, and blazars are simply viewed along the plane of the AGNs jets, giving rise to the severe, short-term variability of their continuum.

One of the pitfalls of the 'oversimplified' Unified Model, is that it ultimately relies on the obscuring dusty structure around the BLR to be non-axisymmetric and uniform. Over the last decade, research has shown the need for and potential proof of a more complicated geometry or non-uniform medium for this absorbing region. This has been examined through the use of different forms of modelling: theoretical modelling, and modelling based on observational evidence.

GALAXY TYPE	ACTIVE NUCLEI	NARROW EMISSION LINES	BROAD EMISSION LINES	X-RAYS	UV EXCESS	FAR-IR EXCESS	STRONG RADIO	JETS	VARIABLE	RADIO LOUD
NORMAL (NON-AGN)	NO	WEAK	NO	WEAK	NO	NO	NO	NO	NO	NO
LINER	UNKNOWN	WEAK	WEAK	WEAK	NO	NO	NO	NO	NO	NO
SEYFERT I	YES	YES	YES	SOME	SOME	YES	FEW	NO	YES	NO
SEYFERT II	YES	YES	NO	SOME	SOME	YES	FEW	NO	YES	NO
QUASAR	YES	YES	YES	YES	YES	YES	SOME	SOME	YES	SOME
BLAZAR	YES	NO	SOME	YES	YES	NO	YES	YES	YES	YES
BL LAC	YES	NO	NO/FAINT	YES	YES	NO	YES	YES	YES	YES
OVV	YES	NO	STRONGER THAN BL LAC	YES	YES	NO	YES	YES	YES	YES
RADIO GALAXY	YES	SOME	SOME	SOME	SOME	YES	YES	YES	YES	YES

Figure 1.6 *Table outlining the distinctions between types of AGN and inactive galaxies. Note: OVV = Optically Violent Variables.*

1.4 Modelling of the Obscuring Torus

Theoretical modelling has aided the exploration of specific features and physical mechanisms that could be behind the formation/maintenance/evolution/dissolution of the dusty torus. One such example is the extensive semi-analytical model described in Vollmer et al. (2008). This model details the formation of a clumpy, central disk via gas infall from the host galaxy. This model considers the formation of such tori via three evolutionary phases: (i) a short and massive infall of gas causes a turbulent disk, driven by stellar winds, to form. This initiates star formation processes. (ii) the intercloud medium is then discharged via stellar supernovae processes. This leaves a massive, geometrically thick disk, with a high mass accretion rate. This mass accretion rate decreases over time. (c) As this rate becomes greatly decreased, the thick collisional torus becomes thin and transparent, much alike the disk around the nuclear region in the middle of our Milky Way (Alig et al., 2013).

There have been many mechanisms proposed to support the maintenance of a thick gaseous torus with large height-to-radius ratio, as rotation alone cannot sustain such a structure. Examples of these include: (i) a clumpy medium in which cloud collisions are frequent, to promote high-velocity turbulent/outflow motion (Beckert & Duschl, 2004); (ii) UV-optical (from the central accretion disk) and IR (from inside the torus) radiation pressure (in this mechanism, toroidal clumps are driven radially by radiation from the active nucleus (Beckert & Duschl, 2004; Hönig & Beckert, 2007; Wada, 2012). Dust present in these clumps is also heated by radiation from the nucleus, which causes IR emission. The outward diffusion, absorption and subsequent reemission of this IR radiation

between adjacent clumps, results in a radiation pressure with enough force to counteract (or overcome) the gravitational force of the active nucleus (Krolik, 2007). It is important to note that an *active* galactic nucleus is essential to generate and preserve a thick torus in this model); (iii) star formation activity in inflowing gas; and (iv) magnetic winds (in this mechanism, the obscuring torus is described as a region in the wind which presents as a clumpy toroidal structure (Elitzur & Ho, 2009; Elitzur & Shlosman, 2006; Emmering et al., 1992; Konigl & Kartje, 1994). Any clumps which are inside the black hole's sublimation radius are void of dust and seen as BLR clouds, and clumps at larger radii are dusty and seen as a component of the torus itself).

The two main underlying results of the aforementioned theoretical torus models are as follows: (i) some sphere of influence of the central black hole over its host galaxy, described by the link between the host's gravitational potential and the toroidal dimensions; (ii) constant gas flow, from parsec radii to the innermost stable circular orbit (ISCO) of matter around the central black hole, is necessary for these models to work - this would, observationally, obstruct sections of the BLR behind the disk in Type-I AGN which are viewed at near face-on angles (Netzer, 2015). It is scenarios like this which call for observation-based modelling in addition to theoretical modelling.

Observational modelling aims to create hypotheses regarding the composition, geometry and size of the torus, in an effort to describe observational data such as AGN spectral energy distributions (SEDs) and spatially resolved images (Hönig & Kishimoto, 2010; Pier & Krolik, 1992; Stalevski et al., 2012).

There are three common observational torus models, all of which demonstrate axisymmetry, anisotropy, have their own effective covering factors, and assume that the inner toroidal walls exist at the dust sublimation radius: (i) models assuming a smooth/continuous gas distribution (Dullemond & van Bemmel, 2005; Fritz et al., 2006; Pier & Krolik, 1992), models assuming a clumpy torus (Dullemond & van Bemmel, 2005; Hönig & Kishimoto, 2010; Nenkova et al., 2008), and composite models which assume a mixture of both (Stalevski et al., 2012). The continuous models parameterise the inner/outermost toroidal radii, the toroidal opening angle, and its density profile. Clumpy models involve further parameters, such as clump column density, clump radial distribution, and clump density distribution (Feltre et al., 2012). The composite models, which assume a combination of smooth and clumpy material in the torus, are the more realistic of the three models. These models conserve the key characteristics of the clumpy

models, but also consider further attenuation by intercloud gas and dust (Netzer, 2015).

It is important to note the possibility of AGN obscuration from structures other than the torus. Absorption on galactic scales is prevalent for many AGN, and has been an area of continued scientific study since the 1980s (Lawrence & Elvis, 1982; Malkan et al., 1998; Prieto et al., 2014). Effects due to galactic extinction can sometimes be considerable enough to completely obscure the active nucleus at optical wavelengths (Goulding & Alexander, 2009). This type of obscuration is more prevalent faint Type-II AGN, of which highly spatially resolved observations have exposed dusty filaments, on kiloparsec scales, which are connected to more minor dusty components close to the central black hole. Conclusions such as this prompted the suggestion that, at lower luminosities, central tori are not mandatory for AGN unification (e.g. Prieto et al. (2014)). However, studies such as Lagos et al. (2011) have shown that extinction via galactic dust absorption cannot solely justify the lack of broad emission lines in Type-II AGN.

The presence of a tilted disk could also serve to obscure central regions of AGN, instead of a torus, and possibly settle differences between tori models vs. observations (Lawrence & Elvis, 2010). This theory suggests the existence of misaligned inner and outer disks around the AGN - the larger outer disk is dusty, geometrically thin and potentially warped, and absorbs some UV-optical radiation, reemitting it at IR wavelengths. Observations have been found to relatively concur with this theory (Roseboom et al., 2013), but the results were shown to also support other possible obscuration scenarios. Furthermore, when compared to other results (e.g. hydrodynamic simulations, IR interferometry and SEDs to name a few), the tilted-disk model is inconsistent and there is often no realistic comparison (Netzer, 2015). Clearly, the discussion and analyses around the unification of observational and theoretical AGN obscuration models must be a continued effort.

1.5 The relationship between AGN obscuration, luminosity and redshift

The potential dependency between AGN obscuration, luminosity and redshift has been another field of scientific interest over the years, along with the attempt to

unify the results from different obscuration classifications whilst exploring these dependencies. By classifying sources as Type-I/unobscured or Type-II/obscured, one can determine the Type-I/Type-II ratio and use this as an estimate of the average obscured fraction of a given population. This is generally done in two ways: (i) by measuring X-ray absorbing column densities and determining the AGN type (absorbed or unabsorbed) relative to a fiducial column which separates the desired groups (Brightman & Ueda, 2012; Gilli et al., 2007; Hasinger, 2008; Malizia et al., 2012; Ueda et al., 2003); and (ii) classifying AGN based on the presence of broad and/or narrow emission lines in their UV-optical spectra. Results using both of these methods would concur if the standard unified model holds, but this is unfortunately not the case.

Merloni et al. (2014) undertook a detailed study, superseding previous works (Gilli et al., 2007; Hasinger, 2008; Malizia et al., 2012; Ueda et al., 2003), of X-ray obscuration in AGN, calculating obscured AGN fractions as a function of 2 - 10 keV X-ray luminosity and redshift for both optically and X-ray classified AGN. In this work they use a sample of 1310 sources at z $\sim$ 0.3 - 3.5 based on observations of the XMM-COSMOS field, complementing Lusso et al. 2013's extensive analysis of Type-I sources in the very same field. Figure 1.7 below, taken from Netzer (2015), shows the obscured AGN fraction results from Merloni et al. (2014).

If we look at the left panel of Figure 1.7 we can see that, when using optical methods of AGN classification, it shows a notably steep decrease in obscured AGN fraction as a function of X-ray luminosity, across redshifts. Comparing this to the right panel, using X-ray classification shows a much flatter, weaker dependence in obscured AGN fraction on source luminosity, across redshifts. It seems to show that obscuration based on X-ray classification misses a bulk of low-luminosity Type-II AGN. This also highlights how optically classified Type-I and Type-II AGN cannot be precisely distinguished by their X-ray obscuring column densities, despite Type-II sources being more commonly obscured. Using a separating N_H value of $\sim 10^{21.5}$ cm^{-2}, Merloni et al. (2014) reported roughly 30% of optically classified Type-I AGN to have obscuring columns $> 10^{21.5}$ cm^{-2} (could be the result of X-ray obscuration via ionised gas outflows in the toroidal opening, or BLR gas) and roughly 30% of optically classified Type-II sources have obscuring columns $< 10^{21.5}$ cm^{-2}.

Looking once more at Figure 1.7, the redshift evolution of obscured and unobscured AGN across the full range of X-ray luminosities is hard to discern.

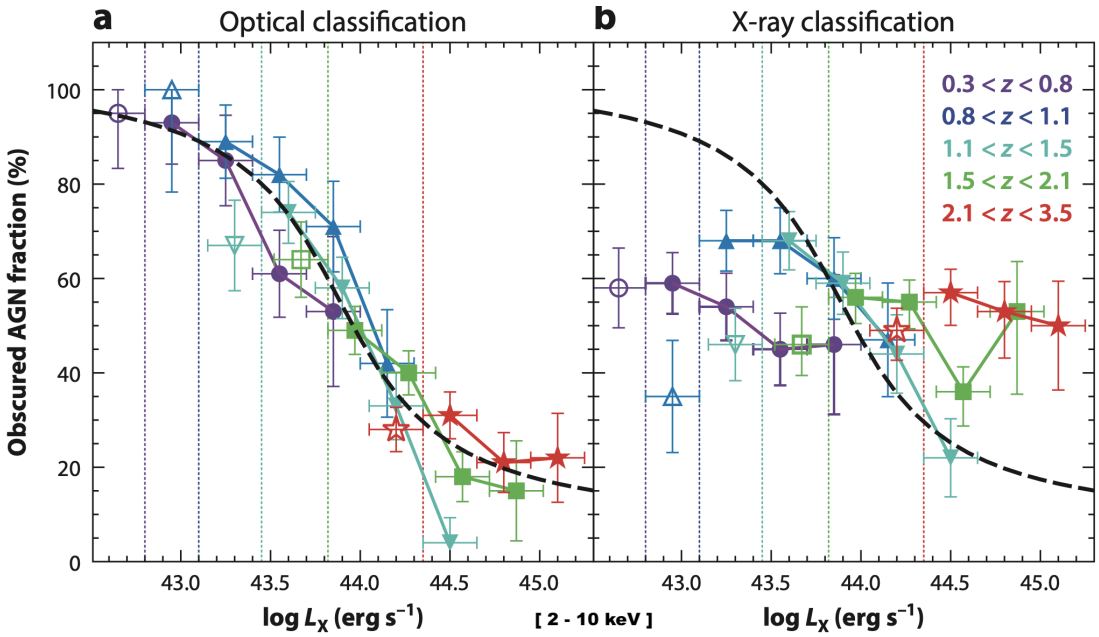

Figure 1.7 *The fraction of obscured AGN across various redshifts, as a function 2 - 10 keV X-ray luminosity, for obscuration based on optical classifications (left panel) and X-ray classifications (right panel). The X-ray classified obscured AGN fractions were based on X-ray observation derived column densities, with a* $N_H = 10^{21.5} cm^{-2}$ *dividing column density between the groups. The optical classification of unboscured/obscured is based on the presence/lack of broad emission lines (> 2000 km* s^{-1} *FWHM) in the optical spectrum. This figure is taken from Netzer (2015), where it is based on data from Merloni et al. (2014).*

This is due to the observational challenges encountered when identifying and classifying obscured, high - z AGN down to low luminosities. Previous studies have found the fraction of obscured AGN to increase with redshift (Hasinger, 2008), whilst some have challenged these results, claiming inadequate consideration of K-corrections. As mentioned previously, even Merloni et al. (2014) shows inconclusive redshift evolution for low to intermediate luminosity AGN. Looking at the brighter sources however, their results demonstrate a factor ~ 2 increase in the obscured AGN fraction between redshifts 1 - 3; the incidence of obscured AGN at higher z being at least comparable to that at local redshifts, with obscured fractions potentially increasing with z, is a result which is echoed in and supported by various other studies (Aird et al., 2015; Brightman & Ueda, 2012; Buchner et al., 2015; Ueda et al., 2014; Vito et al., 2013).

The lack of concurrence between optical/X-ray obscuration classification statistics, and the absence of more recent high-z, low luminosity, completeness corrected optically classified AGN obscuration analyses (when compared to X-ray absorbed AGN populations), strengthens the need for more in-depth and careful X-ray/optical

analyses of obscured AGN.

Efforts have been made to explain an AGN obscuration-luminosity dependence, using torus geometry. One such idea is the 'receding torus' model, first outlined by Lawrence (1991), and suggests that the luminosity of the AGNs photoionising radiation correlates to the radius at which evaporation of the toroidal dust occurs - powerful enough AGN could sublimate toroidal dust over time, causing the obscuring torus to recede. However, this original model could be described as oversimplified - comprehensive analyses of large AGN samples have tested the receeding torus model, by testing prediction that the obscured fraction depends on luminosity, with conflicting results. This variation can be due to (i) the way in which AGN are distinguished and defined as obscured and unobscured, (ii) the energy band used for AGN selection, which essentially effects the sensitivity towards obscuration and optical depth, and (iii) the parameter space over which AGN are selected (e.g. the luminosity-redshift-mass plane) (Hickox & Alexander, 2018).

Despite this, the increased inner toroidal radius with increasing luminosity is a result which is indeed predicted by all models of AGN tori. When combined with the well indicated AGN luminosity-obscuration dependence (confirmed by numerous studies in the X-ray, optical, infrared, and radio wavebands e.g. Lusso et al. (2013); Mateos et al. (2017); Simpson (2005)), and its possible evolution with redshift, this brings forth implications of obscuration linking to AGN evolution in general.

1.6 Obscuration and AGN-galaxy evolution

Models of SMBH - galaxy co-evolution have implemented vital time-varying obscuration to describe the interaction between SMBHs and their hosts, and observed relations between starburst and AGN activity (Di Matteo et al., 2005; Hickox & Alexander, 2018; Hopkins et al., 2008). Some of these models (Figure 1.8, for example) hypothesise rapid phases of merger-triggered SMBH and galaxy growth. These phases could also be provoked by other violent instabilities or interactions which would disrupt the gas and dust in the galaxy, and conceal the central AGN (Hopkins et al., 2008; Sanders et al., 1988). They also predict the bulk of galaxy and SMBH growth to happen in an 'obscured' phase at early

times (Blecha et al., 2018). After this, a 'blowout' phase is predicted to take place, during which SMBH growth is stifled and gas is ejected from the galactic potential well, due to AGN feedback. Star formation is subsequently quenched by this process (Ishibashi & Fabian, 2016). Obscured AGN activity plays a key role in the aforementioned evolutionary model, and is highlighted in the phase during which SMBH mass is built up, whilst the luminosity restricts growth and defines the interplay between BH and host (AGN are observed to influence surrounding gas through heating, outflows, and turbulence in ionised/molecular/atomic gas - Feruglio et al. (2015); Harrison et al. (2014)).

Yet, it is uncertain if SMBH and galaxy evolution is mainly induced by such distinct and powerful events such as those mentioned above. It has been proposed that the evolutionary interplay between SMBHs and their hosts develops on slower scales over cosmic time, with AGN obscuration and fuelling happening stochastically (Cisternas et al., 2011; Mullaney et al., 2012). This would mean obscured AGN do not symbolise a vital evolutionary stage of SMBH feedback or growth. To help reveal the function of AGN obscuration in the context of the cosmological SMBH growth, it is important to discern if obscuration is linked with physical processes on small scales (such that it is disconnected from more extensive galaxy formation/evolution, i.e. localised to the nuclear torus), or if it is linked with galactic-scale processes. Further analyses are necessary to calculate obscured AGN fractions at these scales.

1.7 Motivation for studying obscured AGN

All of this evidence proves the obscured AGN population to the be one of particular interest. AGN obscuration appears to be more extensive in intrinsically low luminosity sources. In general, these less luminous sources make up the majority of the AGN population (only adding to the significance of studying obscured SMBH growth), with brighter AGN being the rarer of the two. How obscured a source is directly impacts its observability; the more obscured sources are less luminous and consequently harder to observe.

However, whilst faint and obscured AGN may show poor signatures (if any, in the case of the more heavily obscured sources) at some wavelengths, looking for these objects across different wavelengths can prove more fruitful. For example,

while UV-optical photometry is ineffective in identifying obscured AGN (due to (i) dust-induced high optical depths at UV wavelengths obscuring accretion disk emission (Padovani et al., 2017) and (ii) high host galaxy emission from stars at UV-optical wavelengths diluting the comparatively weak UV-optical emission from obscured AGN) the selection of obscured AGN in the X-ray waveband is incredibly reliable and one of the most complete methods used, for a variety of reasons (Brandt & Alexander, 2015). X-rays can penetrate through even the most obscuring material (depending on the energy of the X-rays and the amount of AGN obscuration). Additionally, the near ubiquitous X-ray emission from AGN, coupled with the fact that X-ray emission from other astrophysical processes in host galaxies are typically weak by comparison, makes observing in this waveband highly desirable when wanting to select for AGN, as opposed to other astrophysical objects. There are limitations when looking for obscured AGN at X-ray energies. For example, lower energy X-ray emission (often called 'soft' X-ray emission) cannot penetrate more obscured systems and, as such, only utilising soft X-ray selection can lead to more obscured AGN being missed when constructing samples. However, this can be rectified by selecting sources based on higher energy X-ray emission ('hard' X-ray emission) which, at present, only fails for the most heavily obscured systems. Contamination at X-ray wavelengths from host-galaxy processes can also raise issues, but this is again only the case for the most heavily obscured, least luminous AGN. Hence, through careful X-ray selection, it is possible to find obscured AGN with little to no optical signature at all.

Whilst using UV-optical photometry to select for obscured AGN is less efficient, the use of UV-optical spectroscopy can be a fundamental tool when identifying and classifying and obtaining the redshifts for obscured AGN. The UV-optical waveband is incredibly rich in emission lines and, through measuring emission line widths and ratios, renders the study of AGN spectra in this waveband an incredibly effective tool when identifying obscured AGN. Using emission-line diagnostics also provides a way to identify large samples of AGN sources, and allows us to recognise systems with extreme nuclear obscuration (emission from the NLR will not be heavily absorbed by obscuration on small scales). However, despite UV-optical spectroscopy being a powerful probe of AGN obscuration, it is incredibly expensive and unrealistic to be able to obtain spectroscopy for large surveys, using instruments with the ability to detect the faintest and most obscured sources, especially at higher redshifts where UV-optical spectral emission lines can be scarce, signal-to-noise is poor, contamination from

stellar absorption is high, and obscuration/absorption/reddening effects can be heightened. For example, large surveys like SDSS consist of thousands of AGN with redshift of up to z $\sim$ 7, but the aforementioned selection effects hinder and bias AGN classification, leading to incomplete samples. As a result, there are few spectroscopically complete surveys which truly encapsulate obscured AGN and consequently there are a limited number of studies which use UV-optical spectroscopy as a basis for extended obscured AGN anaylses; those which do must create a spectroscopic sample using data from multiple, potentially incomplete, surveys. This can lead to inconsistency when trying to account for completeness corrections and instrumental effects during observational analyses and calculations. All of this strengthens the need for methods and studies which focus on the congruent selection, observation, classification, and completeness corrected analyses of the obscured AGN population.

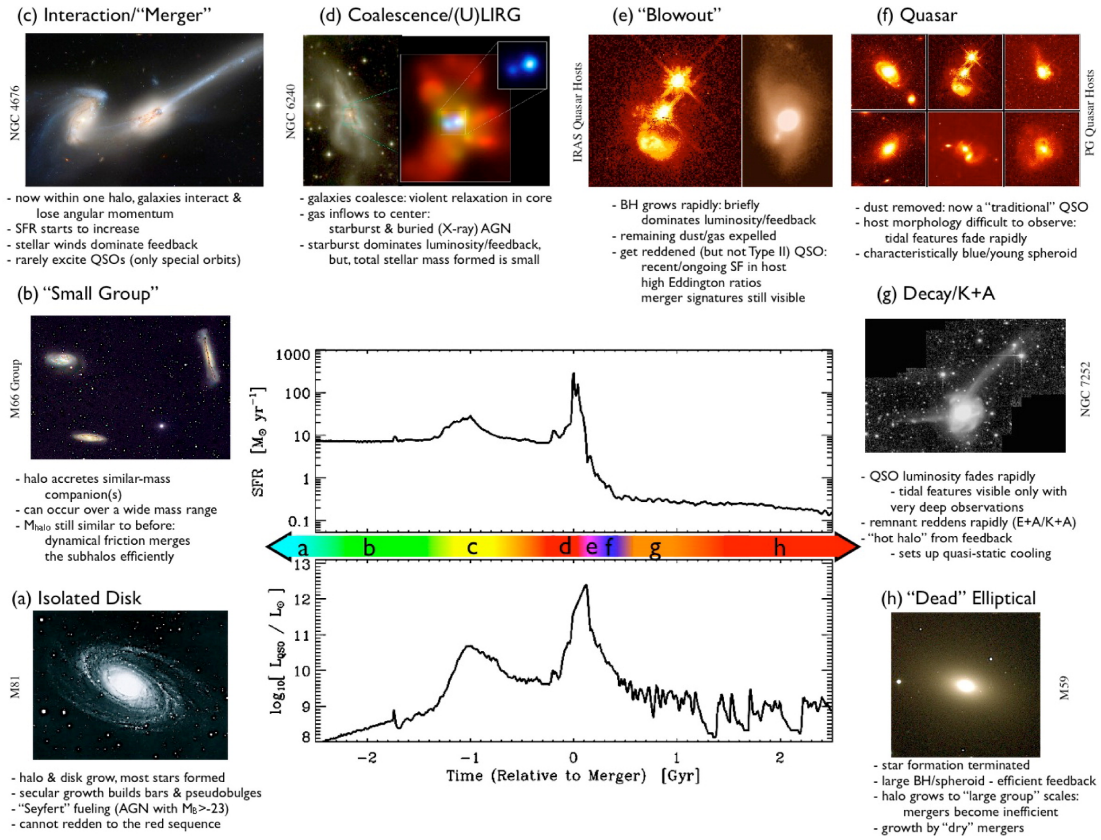

Figure 1.8 *Schematic outline of the growth phases in a 'typical' galaxy undergoing a major, gas-rich merger, (merger-driven AGN model) taken from Hopkins et al. (2008). All image credits are detailed in said paper.*

Another argument for studying the obscured AGN population, is the prevalence of these objects at early times. If AGN obscuration is linked to host-galaxy gas content, as discussed before, then it could be proposed that there is a

higher incidence of obscured AGN at higher redshifts; at early times, galactic atomic or molecular mass fractions were much greater than in the local Universe (Carilli & Walter, 2013). It was shown in a paper written by Treister & Urry (2006) that the intrinsic fraction of obscured sources increases at earlier redshifts (z). This is illustrated in the bottom panel of 1.9. Furthermore, as previously mentioned, obscured fractions based on X-ray observations have also indicated an increase of the obscured AGN fraction with redshift; in the case of galaxies with very high gas masses, large scale clouds can heavily absorb X-ray emission, leading to X-ray observations missing larger fractions of AGN growth at high - z (Hickox & Alexander, 2018). A conclusion can be drawn from this: obscured, low-luminosity sources could have dominated the AGN population at some point in time, corresponding to the majority of BH growth at earlier cosmic epochs. This dependence of obscuration on redshift, and even luminosity, has still not yet been fully explored; the standard unified model predicts there should be no luminosity dependence on obscuration, the receding torus model predicts luminosity dependence but no dependence on redshift, but if z - dependent obscuration did exist it would hint once again at the need for an evolutionary model.

All of this highlights the importance of correctly understanding how AGN obscuration ties in with AGN evolution; studying obscured, low luminosity AGN sources is important when attempting to characterise AGN evolution. It has the potential to prove or disprove running theories, and could lead to updated and more complex arguments on the matter.

1.8 The X-Ray Luminosity Function

The AGN population is one that steadily evolves, with typical luminosities and number densities seen to decrease from redshift 1 to 0. This means that SMBH accretion was more widespread in the universe when it was approximately half its present age. One method of constraining and tracing such a population's growth and evolution is by looking at the evolution of the AGN X-ray Luminosity Function (XLF). The AGN X-ray Luminosity function is a measure of the number of AGN per unit volume with X-ray luminosities that lie within the range L_x to $L_x + dL_x$ (i.e. the AGN number density as a function of AGN luminosity), and is a fundamental statistical measure for the AGN population, and for the understanding of SMBH accretion history. The shape of the AGN XLF,

typically described as a double power law with a 'break' at a certain luminosity, traces the combination of the underlying SMBH mass distribution and accretion rates/Eddington ratios[3] (Aird et al., 2015; Ananna et al., 2022; Schulze et al., 2015). Hence, precise and robust measurements of the evolution of the XLF for different populations can present vital insights into SMBH growth over cosmic time, and the physical processes which stimulate such evolution.

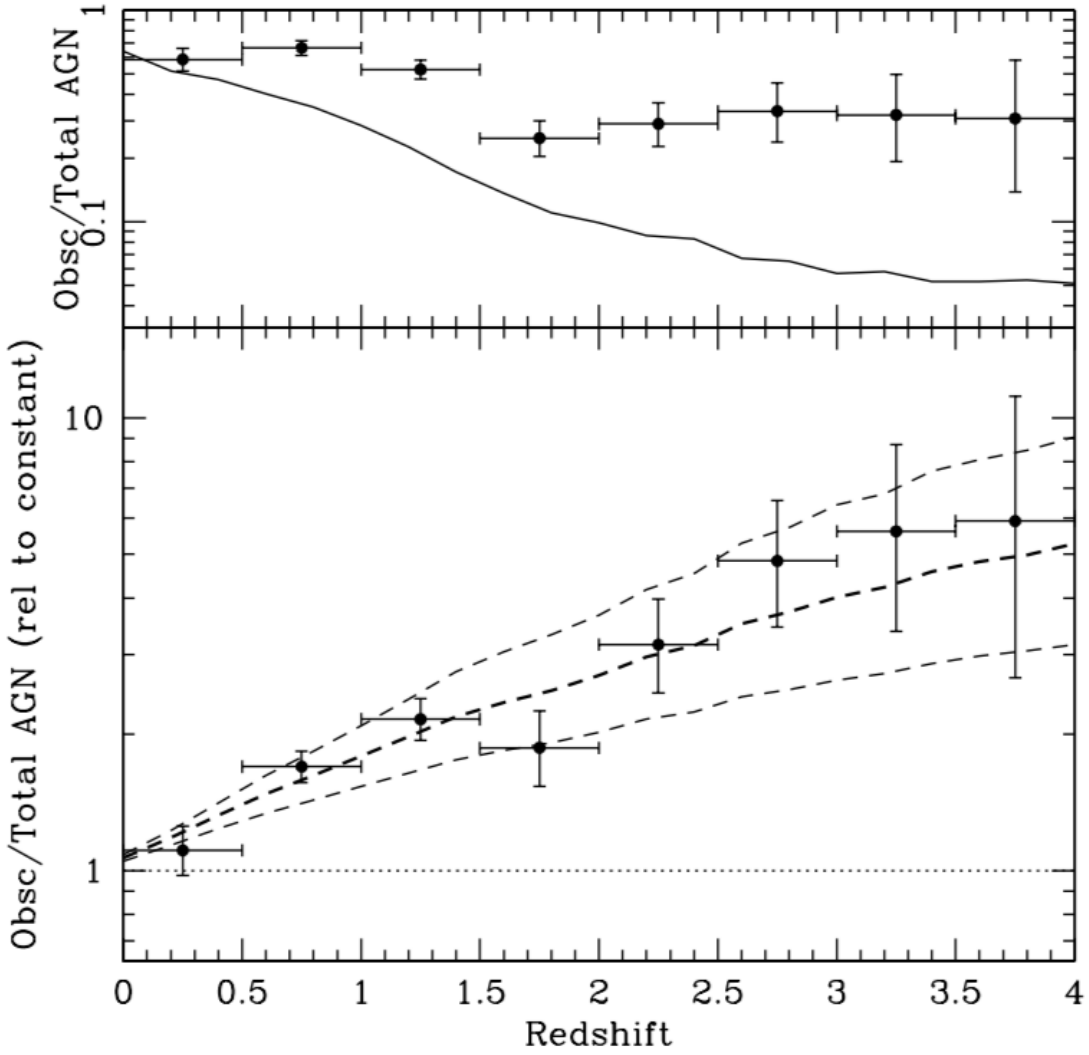

Figure 1.9 ***Top****: Observed obscured fraction vs redshift for optically-selected AGN, from Treister & Urry (2006); a decrease in obscured fraction with increasing redshift is seen. Solid line shows the expected obscured fraction observed if the intrinsic fraction of obscured sources does not evolve, when accounting for spectroscopic incompleteness.* ***Bottom****: Fraction of obscured AGN relative to the expected number observed if intrinsic fraction does not evolve. A significant increase in fraction with redshift is seen*

[3]The Eddington luminosity/limit, L_{Edd}: the maximum luminosity a BH can achieve when maintaining hydrostatic equilibrium, during which radiation pressure generated by accreted infalling matter balances the gravitational force of said BH for spherically symmetric time-invariant accretion. The Eddington ratio is the ratio of the AGN's luminosity output across all wavelengths (bolometric luminosity, L_{bol}) and the Eddington luminosity: L_{bol}/L_{Edd}.

A collection of works have constrained AGN XLFs out to relatively high - z, utilising the efficiency and selection power which samples of deep and wide X-ray surveys offer (Aird et al., 2015; Ananna et al., 2022; Buchner et al., 2015; Fotopoulou et al., 2016; Miyaji et al., 2015; Peca et al., 2023; Ueda et al., 2014). Results from these works show that, as a population, AGN evolve strongly with redshift. AGN number densities are shown to distinctly increase from local times to z $\sim$ 1 - 2; more luminous X-ray selected AGN concur with their optically selected counterparts in that their number densities peak at around z $\sim$ 2, whereas fainter AGN number densities are shown to peak at later times (z $\sim$ 1) and show a shallower decrease to more local times. The potential evolutionary patterns hinted at by these studies have resulted in the following parameterisations, which aim to describe AGN XLF evolution over cosmic time: (i) Pure luminosity evolution (PLE) - this was the first evolutionary adjustment made to the double power law XLF model. First investigated in 1968 by Schmidt, this evolution has continued to be examined by studies over the years (Ebrero et al., 2009; La Franca et al., 2005; Miyaji et al., 2000; Ueda et al., 2003), and manifests as a shift in the XLF to higher luminosities with an increase in redshift. In this model, the XLF's shape is presumed to stay unchanged and, as such, PLE can be understood to show a change in luminosity of the overall AGN population, with redshift. (ii) Pure Density Evolution (PDE) - this was also investigated in the initial stages of XLF analyses (Schmidt 1968), and manifests as change in the normalisation of the XLF with redshift. This can be understood as a change in AGN number densities, but with unchanged luminosities. Physically, this would be possible if changes in phases of AGN activity (i.e. going from active to inactive stages) were swift (and therefore undetectable). (iii) Independent luminosity density evolution (ILDE) - utilised by Yencho et al. (2009), ILDE characterises AGN XLF evolution at z $<$ 1.2, and predicts a synchronous evolution in luminosity and AGN number density. (iv) Luminosity and density evolution (LADE) - made known by Aird et al. (2010), this evolutionary model facilitates independent XLF evolution in both AGN number density and X-ray luminosity, whilst including a critical z value. In this mechanism (used in Aird et al. (2010); Ross et al. (2013)) a broken power law describes the luminosity evolution, and a power law describes the number density evolution. (v) Luminosity-dependent density evolution - proposed by Schmidt (1983a) and formalised in works including Miyaji et al. (2000); Ueda et al. (2003) and Ueda et al. (2014), this XLF evolutionary model has been widely used by studies to investigate the evolution of the AGN XLF out to high redshifts (Aird et al., 2015; Fotopoulou et al., 2016; Schmidt, 1983b; Ueda et al., 2014). LDDE

describes a more complicated evolution which summarises the changing AGN number density with z, whilst also accounting for the luminosity - dependence of the critical redshift (due to faint and bright AGN displaying different evolutionary timescales). This culminates in an AGN XLF which changes shape over cosmic time (Aird et al., 2015). Over the years, works have favoured LDDE over PLE at softer and harder X-ray regimes (Aird et al., 2015; Ebrero et al., 2009; Miyaji et al., 2015; Ueda et al., 2014).

One can attempt to measure the AGN XLF in many ways. The $1/V_a$ estimate (Miyaji et al., 2000), which in itself is a general form of the $1/V_{max}$ method brought forward by Schmidt (1968), is a method which is frequently used. These approaches generate binned differential luminosity function estimates at desired redshift intervals. However, despite their common use, such methods can be biased at (i) faint fluxes close to flux sensitivity limits, and (ii) large luminosity bins (Miyaji et al., 2001; Page & Carrera, 2000). As such, other methods have been crafted to attempt to rectify such limitations (Ananna et al., 2022; Miyaji et al., 2001).

Developing methods for the constraint of unbinned XLF estimates has also been an area of particular scientific interest in recent years. The use of a fully Bayesian statistical approach, which allows for the consideration of a range of X-ray spectral shapes and corrections in various forms (e.g. accounting for absorption effects, flux uncertainties, redshift uncertainties, luminosity uncertainties), has been crucial to unveiling more robust measurements of AGN XLFs across all redshifts (Aird et al., 2010, 2015; Fotopoulou et al., 2016). Mathematical methods such as modelling probability distribution functions and likelihood functions on a source-by-source basis (Aird et al., 2015) and using tools like Markov Chain Monte Carlo samplers (Ananna et al., 2022) to robustly probe the full XLF parameter space, have been a critical addition to AGN XLF analyses in recent years. These methods truly come into their own when attempting to deliver robust, completeness corrected results which include information from sources where obvious data is scarce (i.e. faint and obscured AGN at high z). It is the aim of this thesis to build on the current and develop new statistical methods to determine the obscured and unobscured AGN XLFs from the local universe to high - z, comprehensively correcting for incompleteness at multiple stages.

Figures 1.10 and 1.11 (taken and adapted from Fotopoulou et al. (2016) and Peca et al. (2023)) show examples of AGN XLF results constrained using some of the methods specifically referenced above. A few key features are clear to

see: (i) the AGN XLF appears to shift to the right with increasing redshift (showing luminosity evolution); (ii) the faint end slope appears to flatten with increasing redshift; (iii) the normalisation of the XLF appears to decrease with increasing redshift (showing density evolution); (iv) the incidence of low luminosity AGN is seen to decrease with increasing redshift, whilst the incidence of highly luminous AGN is seen to increase with increasing redshift, and (v) studies are in most agreement in the intermediate to high luminosity, low redshift regime. This highlights important discussion points. The fact that more is currently known/well-constrained regarding the very luminous AGN population, with a definite increase in number density seen at higher redshifts, means that the population of *lower* luminosity AGN unfortunately tends to be overlooked and unaccounted for due to a combination of reasons. We need the deepest X-ray data in order to investigate these lower luminosity objects at such high redshifts, leaving us with a small number of samples from an even smaller number of appropriate fields. Whats more, identification through spectroscopy can be difficult for these faint X-ray sources, as they also tend to have incredibly faint optical counterparts and can be obscured in multiple bands. This lead to fields becoming spectroscopically incomplete, which rendered established high-z XLFs dangerously biased against faint, obscured sources. Instances where this incompleteness has been corrected for has ultimately led to a higher AGN space density found at these higher redshifts (K.Nandra & Steidel, 2005), proving that including the faint AGN population is crucial in these studies. Fortunately, photometry has progressed in recent years and, where photometric redshift estimates used to fail around z: 2 - 3, studies have now filled in some of the 'high-redshift blank space' using photometric redshift estimates with increased reliability (given enough bands of data). However, despite the improved accuracy of the process, quantifying the success (modelling the uncertainty) of photometric redshifts is difficult; this can inevitably lead to false estimates. Furthermore, photometry is insufficient to characterise AGN obscuration - complete and high quality spectroscopy is needed to achieve this. This begs for further elucidation on the following questions: how can we account for these faint, obscured sources? How much does the faint and obscured population contribute to the shape of the total AGN XLF, and extent of SMBH growth over cosmic time? It is the aim of the work carried out in this thesis, to aid in answering these questions.

The discussion around the evolution of the obscured AGN population and obscured AGN fractions from Section 1.3 can be further explored through independently constraining AGN XLFs for the obscured and unobscured AGN

populations, across a range of redshifts. As previously mentioned, this can be done by using either X-ray absorption (using X-ray data to look at source column densities) or optical obscuration signatures (using spectroscopy to identify broad and narrow emission lines in source spectra) as a means of classifying obscured/absorbed vs. unobscured/unabsorbed sources.

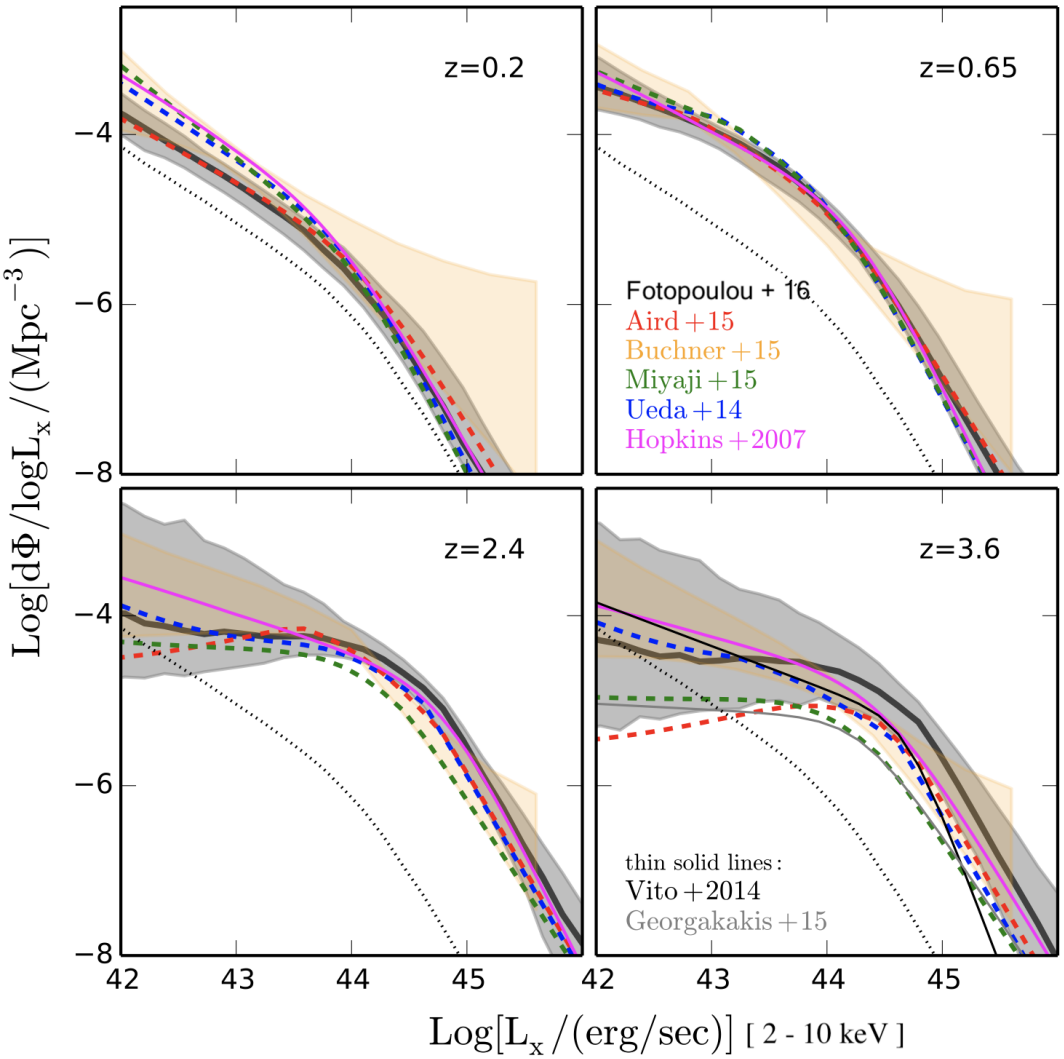

Figure 1.10 *Taken and adapted from Fotopoulou et al. (2016), showing their 2 - 10 keV (transformed from 5 - 10 keV for ease of comparison) AGN XLF results, across redshift; AGN XLFs generated from previous studies, mentioned in this introduction, are included for comparison and referenced.*

Previous works have constrained X-ray absorbed and unabsorbed AGN XLFs at varying redshifts, as showcased by Figures 1.12 and 1.14 (Aird et al., 2015; Peca et al., 2023). Upon examining these results we can see that both the absorbed and unabsorbed AGN populations experience distinct luminosity evolution (shifting right to higher luminosities with increasing redshift) and an overarching density

evolution with redshift. They show a dynamic ratio of absorbed to unabsorbed AGN at different luminosities and redshifts, illustrating a complex evolution in the AGN XLF's shape. Looking at the absorbed vs the unabsorbed AGN XLFs, these results also show that absorbed AGN dominate at low luminosities, and may potentially dominate at the brighter luminosities as we get to higher redshifts. However, once more, the difficulties with and lack of data for faint, obscured sources at high redshifts limits observational constraints in these regimes, this potentially biasing any parameterisation of the obscured and unobscured XLFs.

A study by Ananna et al. (2022) has provided one of the few, up to date examples of AGN XLFs divided based on optical obscuration properties (derived from optical spectroscopy), the results of which are shown in Figure 1.13. However, it is worth noting that this obscured and unobscured XLF study is limited to the local universe as it is relient on the relatively shallow Swift-BAT survey (and associated optical spectroscopic follow-up). Again, the difficulties with procuring and lack of faint, obscured, high-z data limits observational constraints in these regimes, and could potentially skew any parameterisation of the obscured and unobscured XLFs, as well as insights into the evolution of the XLF over cosmic time. Furthermore, the lack of alignment between X-ray classified and optically-classified obscured/unobscured AGN (mentioned previously in Section 1.3) and specifically the lack of updated, completeness-corrected and viable optically-classified measures of obscured/unobscured AGN growth in the aforementioned faint, high-z regime, leaves current population analyses of these sources incomplete.

The goal in this thesis is to attempt to fill in this 'blank space' by, not only determining the AGN XLF for faint AGN across a range of redshifts, but also establish how much of this emission is from optically classified obscured and unobscured sources separately, whilst also carefully accounting for incompleteness where possible. This will give us an updated and robustly constrained insight into the growth and evolution of some of the most difficult-to-detect sources in our universe.

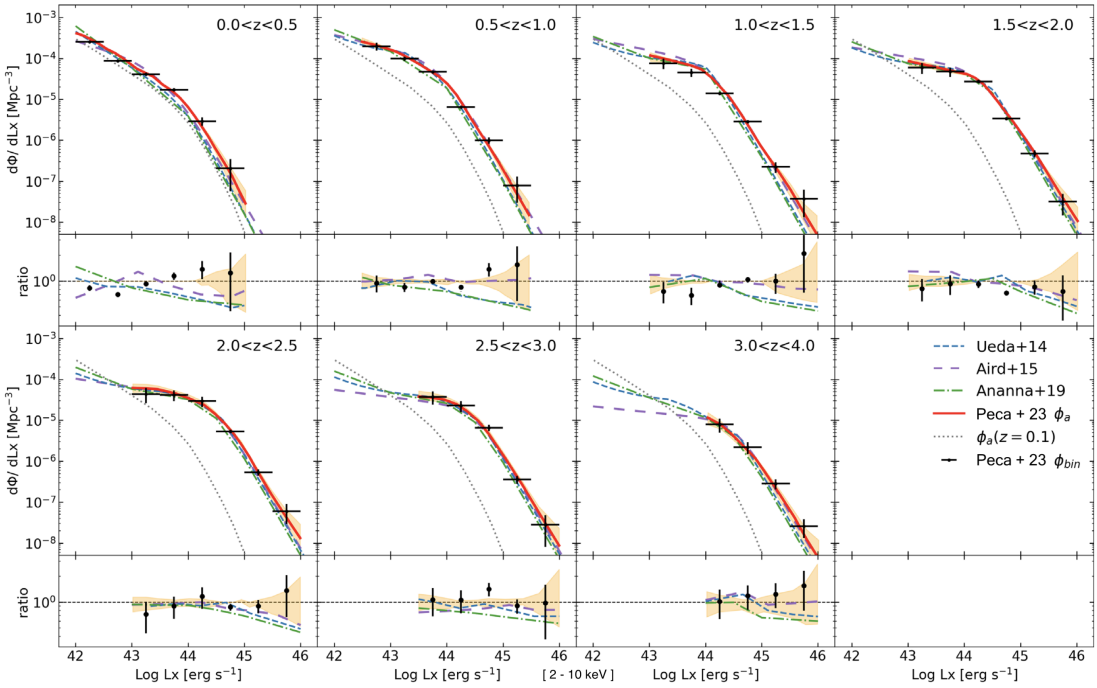

Figure 1.11 *Taken and adapted from Peca et al. (2023), showing their 2 - 10 keV AGN XLF results across redshift; XLFs from previous studies, mentioned in this introduction, are included for comparison and referenced.*

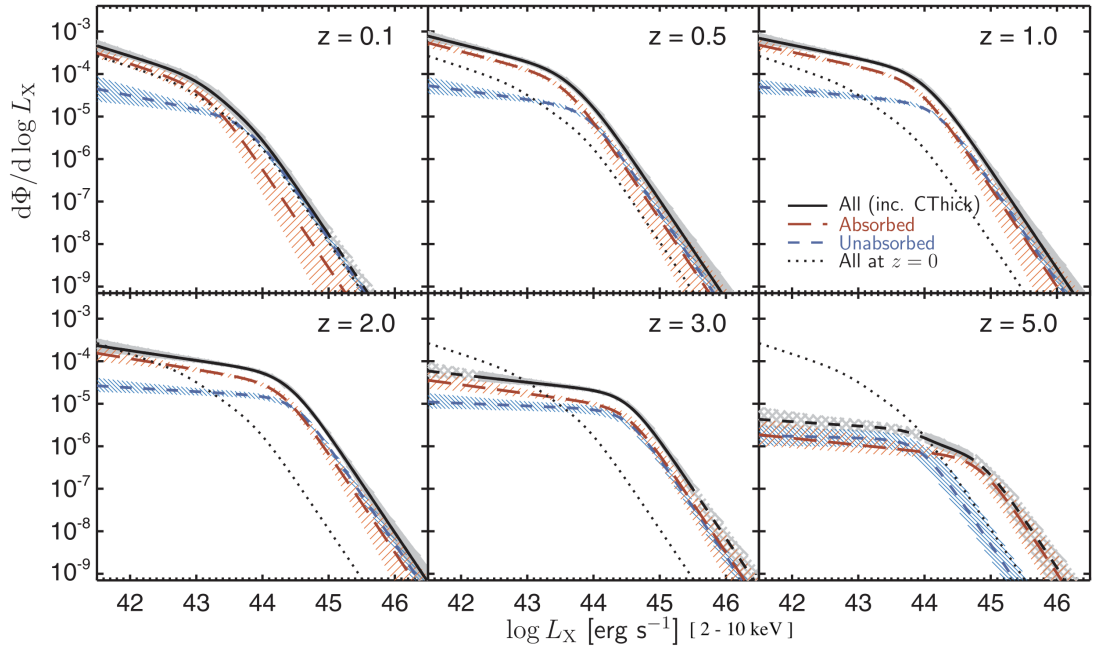

Figure 1.12 *Taken from Aird et al. (2015), showing their total (black solid line), unabsorbed ($20 < logN_H < 22$, blue dashed line) and absorbed ($22 < logN_H < 24$, red dashed line) 2 - 10 keV AGN XLFs for a range of redshifts, based on X-ray classifications of obscuration.*

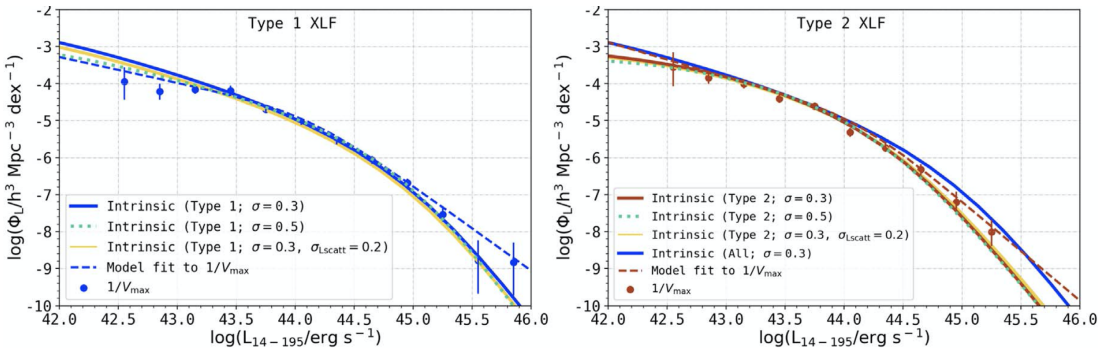

Figure 1.13 *Taken from Ananna et al. (2022), showing their 14 - 195 keV Type I (unobscured, left panel) and Type II (obscured, right panel) AGN XLFs for $z \sim 0.01$ - 0.3, based on optical classifications of obscuration. Solid lines and shaded areas show their final, bias corrected intrinsic distribution functions, with colours indicating various levels of assumed uncertainty.*

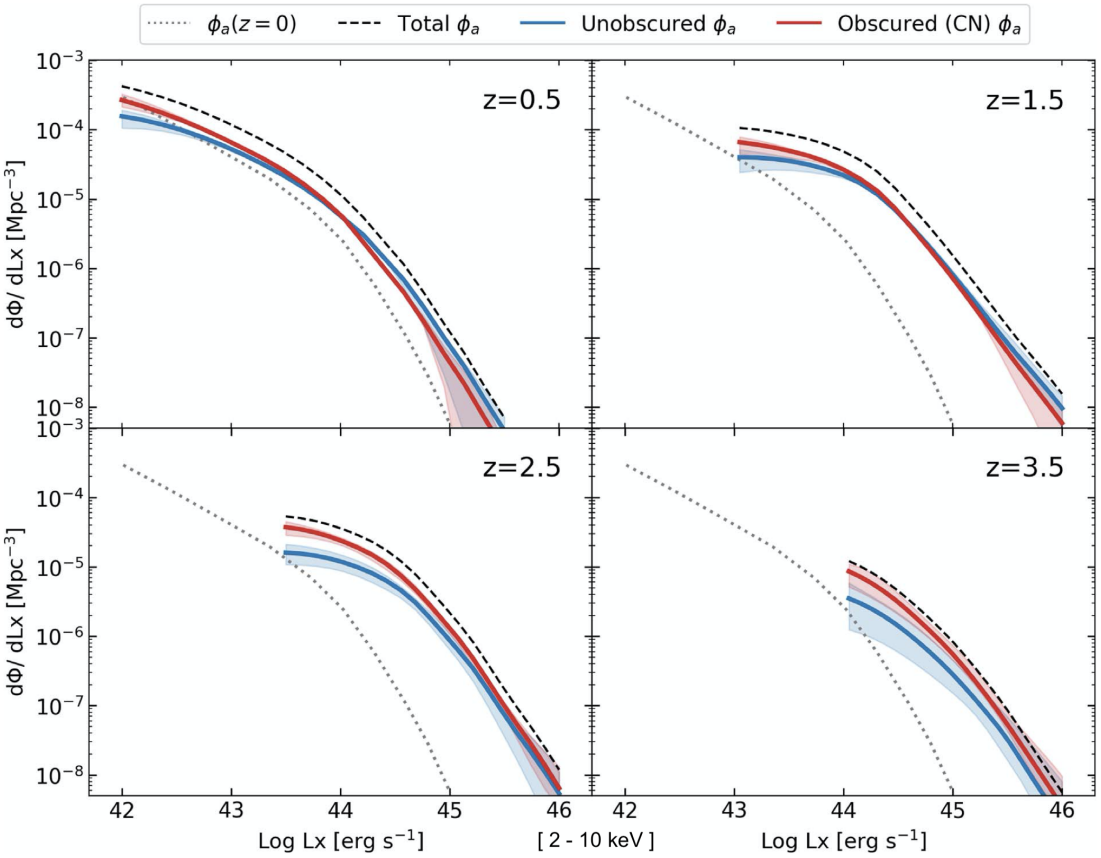

Figure 1.14 *Taken from Peca et al. (2023), showing their total (black dashed line), unabsorbed ($logN_H\ cm^{-2} < 22$, solid blue line) and absorbed ($22 < logN_H\ cm^{-2} < 24$, solid red line) 2 - 10 keV AGN XLFs at various redshift bins, based on X-ray classifications of obscuration. Shaded regions illustrate 3σ uncertainties.*

1.9 Research Aims

The ultimate aim of this project is to accurately determine the amount of obscured black hole growth taking place at differing points in the universe's history: in the early universe (at redshifts 2 - 4), and in the more local universe (redshifts 0 - 1). This will allow us to obtain new and robust measurements of SMBH growth and evolution, tracked across cosmic time. This is a multilayered process, which can be broken down into discrete objectives:

Low redshift analysis:

- Construct a suitable sub-sample of z $\sim$ 0.2 - 1.1 X-ray AGN sources, with accurately matched multiwavelength data, from the NuSTAR Serentipitous Survey's 80-month catalogue.
- Calculate the independent Broad Line and Non-Broad Line AGN X-Ray Luminosity Functions at redshifts 0.2 - 1.1 from our sources, and produce an updated and more robust estimate of obscured black hole growth at lower redshifts.

High redshift analysis:

- Obtain spectra from appropriately selected AGN sources - sources must be within a defined luminosity limit to be classed as 'low luminosity', in addition to existing in the correct redshift range ($2 < z < 4$).
- Construct a method of accurately estimating the redshift and classification of these sources as obscured and unobscured - the two categories being Non-Broad Line AGN and Broad Line AGN (NBLAGN and BLAGN respectively).
- Automate this process, so it can be applied en masse to simulated sources - from this, corrections for completeness can be inferred and subsequently applied to our sample.
- Calculate the independent Broad Line and Non-Broad Line AGN X-Ray Luminosity Functions at redshifts 2 - 4 from our sources, and produce an updated and more robust estimate of obscured black hole growth at higher redshifts.

Discussion:

- Tie together lower and higher redshift AGN analysis and obtain new and robust measurements of obscured and unobscured SMBH growth and evolution, tracked across cosmic time.

Chapter 2

Analysis of the NuSTAR Serendipitous Survey: calculations of the AGN X-ray Luminosity Function at redshifts z $\sim$ 0.2 – 1

In the following chapter, I use data from the NuSTAR Serendipitous Survey's upcoming 80-month catalogue (hereafter known as NSS80) to provide new measurements of the obscured and unobscured AGN X-ray luminosity functions at low redshifts (z $\sim$ 0.2 - 1). I start by discussing NuSTAR as an observatory (Section 2.1), describing the NuSTAR Serendipitous Survey programme and what it entails (Section 2.2), before moving on to the the updated NSS80 catalogue (Sections 2.3 to 2.3.2). I then talk through how I apply various cuts to said catalogue to select highly spectroscopically-complete subsamples of hard (8-24 keV) and full (3-24 keV) band detected BL (unobscured) and NL (obscured) AGN at z $\sim$ 0.2 - 1, with reliable multi-wavelength counterpart information (Sections 2.4 to 2.5). Following this, I move on to describing how I determine the independent obscured and unobscured AGN XLFs for these hard and full band NSS80 subsamples, using a maximum likelihood fitting approach (Sections 2.6 to 2.6.2). I finally using these results to calculate obscured AGN fractions at z $\sim$ 0.2 - 1 (Section 2.7).

2.1 The NuSTAR Observatory

The Nuclear Spectroscopic Telescope Array (hereafter referred to as NuSTAR; Harrison et al. (2013)) brought about a great breakthrough in studying the high-energy X-ray - emitting population in our universe. It was the first orbiting observatory which had the ability to focus X-rays at energies $>$ 10keV; this enables two orders of magnitude greater sensitivity, as well as one order or magnitude improvement in angular resolution, than any of its non-focusing predecessors. With its broad energy coverage, extending to such high energies (3 - 79 keV), NuSTAR has provided us with the opportunity to study and observe large, cleanly selected samples of high-energy-emitting AGN out to high redshifts (Aird et al., 2015; Harrison et al., 2016; Lansbury et al., 2017b; Mullaney et al., 2015). It has resolved $\approx$ 35% of the CXB at 8 – 24keV ((Harrison et al., 2016), provided the first measurements of the rest frame 10 - 40 keV AGN luminosity functions at z $\geq$ 0.1 (Aird et al., 2015), and identified extreme heavily obscured AGN (Civano et al., 2015; Lansbury et al., 2017a; Masini et al., 2018). Our view of the universe in the 3 - 79 keV spectral window has been limited because previous orbiting telescopes have not employed true focusing optics; they have instead used coded apertures that have intrinsically high backgrounds, limited sensitivity and lack angular resolution. High-energy X-ray astronomy is particularly useful when observing more obscured objects; at higher (harder) energies, X-rays are able to penetrate absorbing dust and gas obscuring such systems, allowing us to study all but those most heavily obscured. The 3 - 24 keV band in particular is the most sensitive for relatively faint sources; this broad band pass maximises the potential source signal, whilst minimising soft and hard X-ray background components. Figure 2.1 illustrates how NuSTAR's 3-24 keV energy range (shown by the red and blue solid bars) probes different parts of a moderately absorbed (N_H = 3 x 10^{23} cm^{-2}) AGN X-Ray spectrum at z $\sim$ 1 (shown by the solid black line). The decrease in emission seen at lower energies is the result of absorption; higher levels of emission are seen at higher energies for obscured systems such as this.

It is for all these reasons that NuSTAR is an incredibly useful and versatile observatory for the study of obscured and unobscured AGN across a broad redshift range, favourable for an analysis such as mine.

The NuSTAR observatory, orbiting the Earth once every 97 minutes, consists of two separate, co-aligned, identical telescopes. These telescopes (A and B), each

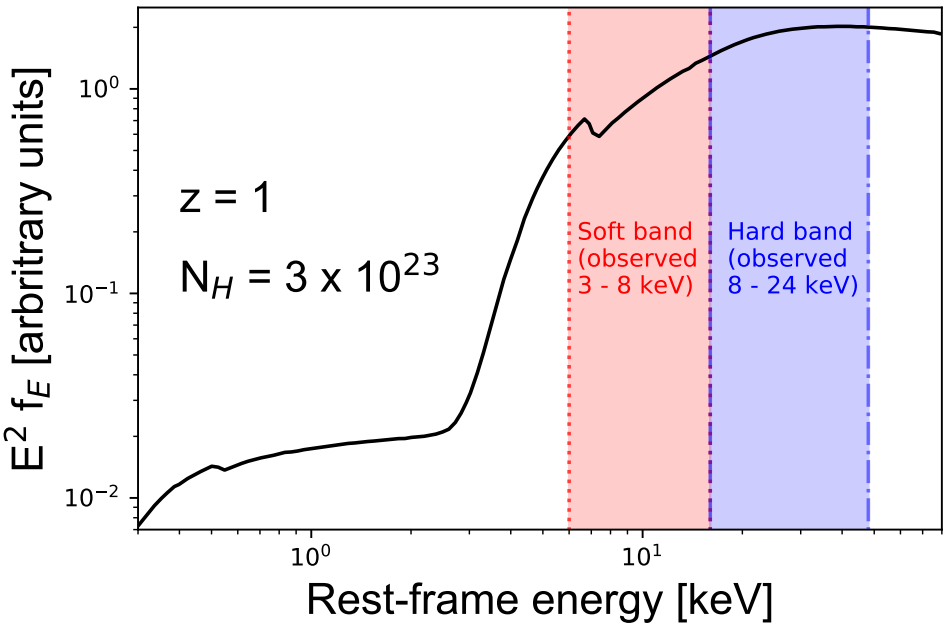

Figure 2.1 *Showing how, at $z \sim 1$, the broad 3 - 24 keV energy range probed by the NuSTAR observatory illuminates different parts of the AGN X-ray Spectrum.*

have a focal plane module (FPMA and FPMB) with $\approx$ 12' $\times$ 12' field of view; these FPMs overlap in sky coverage (Klindt et al. 2023 in prep.). NuSTAR provides $\sim$ one order of magnitude improvement in angular resolution and $\sim$ two orders of magnitude improvement in sensitivity, when compared to its hard X-ray survey predecessors.

NuSTAR's two-year long primary mission phase involved mapping selected sky areas, to enable progress towards the following science goals, as stated on the official NuSTAR website[1]: 1) Taking a census of collapsed stars and black holes of varying sizes, by performing deep extragalactic observations and surveying regions around the center of the Milky Way. 2) Mapping material in young supernova remnants. 3) Understanding what powers the relativistic jets emitting from extreme AGN. The first of these, namely the creation of the NuSTAR extragalactic survey, is the largest scientific program in terms of time investment that has been undertaken with NuSTAR, and is one of the highest priorities of the mission.

The broad range of additional science investigations which NuSTAR has given researchers the opportunity to accomplish (e.g. studying the extreme physics around collapsed stars, or mapping microflares on the Sun's surface) only furthers

[1]https://www.nustar.caltech.edu/page/about

its case as an indispensable astronomical tool for probing the X-ray Universe.

2.2 The NuSTAR Serendipitous Survey

There are two main components to the aforementioned NuSTAR extragalactic survey programme. The first involves carrying out dedicated surveys of blank fields[2], amounting to $\approx$ 3 deg^2 of sky coverage (Civano et al., 2015; Masini et al., 2018; Mullaney et al., 2015). The second is the NuSTAR serendipitous survey (NSS), a wide-area 'serendipitous survey' performed by searching almost all NuSTAR observations over its lifetime for backround X-ray sources (Alexander et al., 2013; Lansbury et al., 2017b). In this chapter, we use data from said NuSTAR serendipitous survey to generate new measurements of the z $\sim$ 0.2 - 1.1 obscured and unobscured AGN XLFs.

Whilst the NuSTAR serendipitous survey is the largest component of NuSTAR's extragalactic survey programme (providing $\approx$ 75–80% of NuSTAR's detected sources), it required zero dedicated time to perform as it all comes from pre-existing NuSTAR observations; the survey is constructed by searching the background regions from almost every non-survey NuSTAR pointing for background sources which are unassociated with the original science target. Any observations which form part of other NuSTAR surveys are excluded, to ensure that the serendipitous survey catalogue is made up of truly serendipitous sources. These include: the dedicated NuSTAR extragalactic surveys (COSMOS, ECDFS, EGS, GOODS-N, UDS); the Galactic surveys (Hong et al., 2016; Mori et al., 2015), i.e. all fields within a 2 degree radius of the Galactic Centre - this specifically decreases the presence of Galactic sources, to maintain an extragalactic sample; the Norma Arm survey (Fornasini et al., 2017); fields where the total counts exceed 106 within 120" of the on-axis position due to a bright science target, to decrease the chance of contamination from such bright science targets (Klindt et al. 2023 in prep.). In addition, Solar system fields (solar, lunar and planetary pointings), nebular fields (supernova remnants) and fields of nearby galaxy nuclei (e.g., the nucleus of M31) are excluded, again to decrease the chances of galactic contamination (Klindt et al. 2023 in prep.). Fields found to have bad exposure maps [3] (post data processing) are excluded, as well as those

[2]Regions of the sky that are devoid of stars down to a given threshold magnitude

[3]A map of the total exposure as a function of position on the sky. It is an image describing the variation (decrease) in effective area away from the optical axis (vignetting), accounting for

where more than two thirds of the field is contaminated by excess background emission from things such as stray light, ghost rays, aperture background and bright science targets.

The NuSTAR serendipitous survey is comprised of a combination of deep and shallow wide-area coverage, allowing for source detection across a large range of X-ray luminosities and redshifts, as well as the identification of more rare cosmic X-ray background source populations which may not be detected in dedicated, smaller-area surveys. For example, the first full 40-month serendipitous survey catalogue (Lansbury et al. (2017b), hereafter referred to as NSS40) consisted of 497 sources found over a sky area of 13 deg^2 - this itself was already $>$ 4 times larger in volume than the dedicated survey components of NuSTAR's extragalactic survey programme . With the imminent finalisation of the 80-month serendipitous survey catalogue (hereafter referred to as NSS80), the NuSTAR serendipitous survey maintains the largest areal coverage of all NuSTAR extragalactic surveys; the upcoming NSS80 catalogue paper, referenced in Klindt et al. 2023 in prep., confirms the catalogue is based on $\approx$ 62 Ms of effective exposure time over 894 unique fields, with an areal coverage of $\sim$ 36 deg^2 . This corresponds to a factor of $\sim$3 increase in the sky coverage with NuSTAR between the 40-month and 80-month catalogues, due to the addition of 947 individual NuSTAR exposures over 563 unique fields (NSS40 was comprised of 510 exposures over 331 unique fields).

Figure 2.2 depicts an all-sky map of the full NSS80 catalogue, illustrating how NSS80 sources are spread across the whole sky.

As part of the NuSTAR extragalactic survey programme, the NuSTAR serendipitous survey has already facilitated valuable AGN science. For example, Aird et al. (2015) made the first direct measurements of the $\geq$ 10 KeV AGN XLF at $0.1 < z < 3$ using a sample of 97 X-ray sources from the NuSTAR extragalactic survey programme, 33 of which were from NSS40. In this chapter we use the most recent NSS source catalogue, NSS80, along with its extensive optical spectroscopic follow up, to generate updated and robust measurements of the z $\sim$ 0.2 - 1.1 obscured and unobscured AGN XLFs.

gaps between CCDs, bad pixels and columns, and other instrumental artifacts. By dividing an image by the exposure map, these effects can be corrected for.

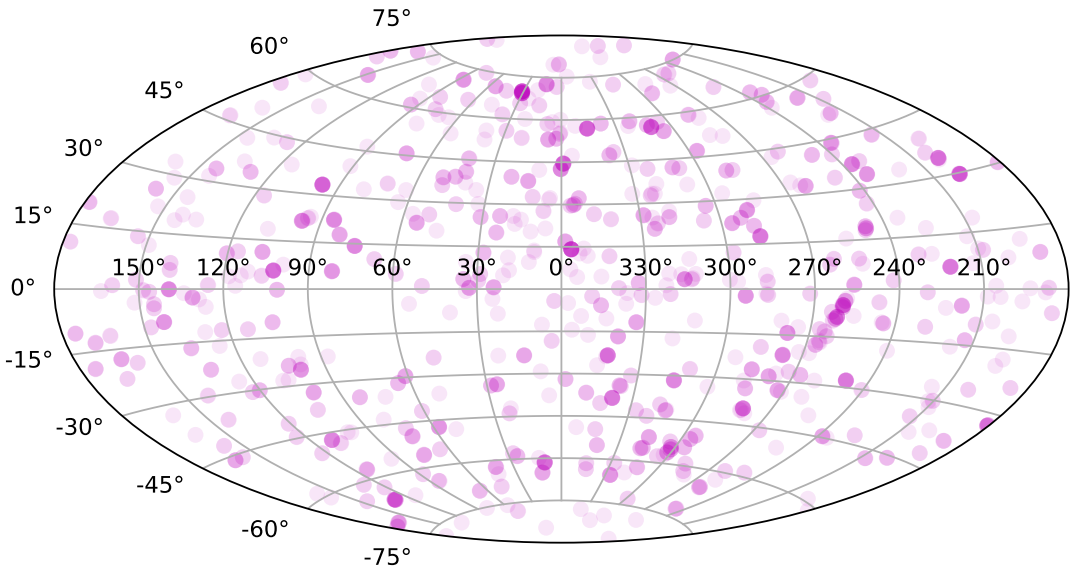

Figure 2.2 *Aitoff projection showing the distribution of the NSS80 survey sources on the sky, in equatorial coordinates (RA, Dec). Excluded from the serendipitous survey are NuSTAR fields containing bright targets and the Legacy surveys (not shown in this figure). Adapted from Klindt et al. 2023 in prep.*

2.3 The 80 month serendipitous survey source catalogue

The full NSS80 source list is made up of 1457 individual NuSTAR exposures that cover 894 unique fields, i.e. regions of the sky. These fields amount to an overall sky coverage of 36 deg^2, with a cumulative exposure time of 62.0 Ms. Klindt et al. 2023 in prep. describes the NSS80 source detection procedure. They first produced false probability maps for each of their images, to determine the probability at each pixel position that the observed counts within a 20" radius circular aperture were produced by a background fluctuation. Then, using the SExtractor software (Bertin & Arnouts, 1996), they created source lists independently on each of the energy bands by identifying regions where the false probability was less than 10^{-6}. A master source list was then created by merging the source lists for each of the energy bands and removing sources within 90" radius of original science target positions, as well as removing extended areas with high background rates.

The total survey area sensitive to a given flux (i.e. area curve) for the full NSS80 sample is shown by the solid orange line in Figure 2.3.

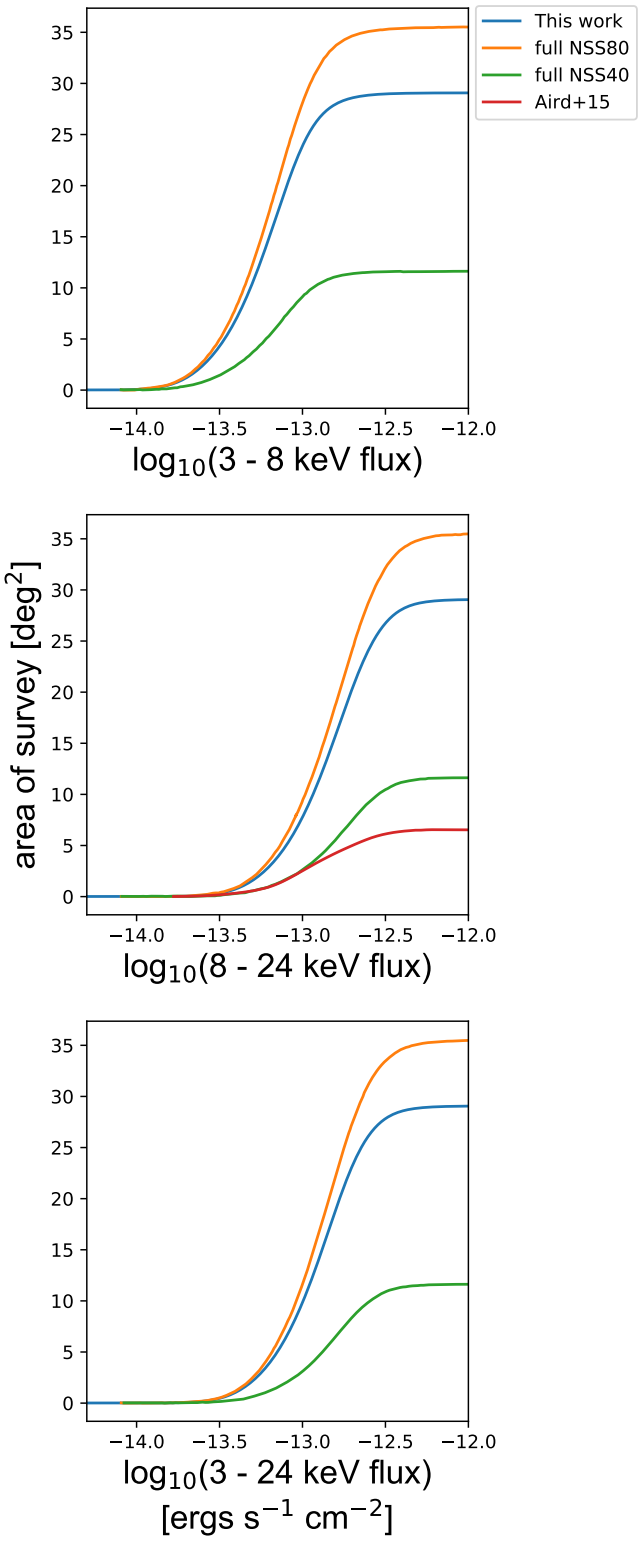

Figure 2.3 *Soft (3 - 8 keV, top panel), hard (8 - 24 keV, middle panel) and full (3 - 24 keV, bottom panel) band area curves for the following samples: the full NSS80 sample (solid orange line), the full NSS40 sample (solid green line), the Aird et al. (2015) sample (solid red line) and the sample used in this work (solid blue line).*

To determine this total NSS80 area curve, individual sensitivity curves were determined for a given background and exposure map, by calculating the flux limit at the detection threshold for every point in the NuSTAR image (with the exclusion of peripheral regions[4].) The total area curve for the NSS80 sample was then determined by summing the sensitivity curves of the 894 unique fields for each energy band (Klindt et al. 2023 in prep.).

Figure 2.3 compares the soft, hard and full band (3 - 8, 8 - 24 and 3 - 24 keV respectively) area curves for the full NSS80 sample (solid orange line) with that of the older NSS40 sample (solid green line). The area curve of the sample used in Aird et al. (2015) (a paper which provided the first direct measurements of the $>$ 10 keV AGN XLF using NSS40 data) is shown in solid red, and the area curve determined for the NSS80 subsample used in this work is shown in solid blue (how we isolate our sample and determine its resultant area curve is detailed in Section 2.5). This figure illustrates a factor $\sim$3 increase in the sky coverage between the NSS40 and the NSS80 surveys. It additionally illustrates a factor $\sim$5 increase in the sky coverage between the hard band area curve used in Aird et al. (2015), and that constrained in this chapter's work.

The master serendipitous survey catalogue source list consists of 1488 serendipitous NuSTAR sources that are significantly detected in at least one energy band. This master source list is split up into two catalogues: the primary NSS80 catalogue, dominated by AGN residing in field galaxies, and the secondary NSS80 catalogue, comprised of X-ray detected NuSTAR sources residing in highly extended optical galaxies and galaxy clusters. All analysis detailed in this chapter uses the primary NSS80 catalogue; all mention of NSS80 hereafter refers to this primary catalogue.

The NSS80 catalogue provides 1274 hard X-ray sources in the 3 – 24 keV band. As previously mentioned, this is the most sensitive band for the relatively faint sources detected in the NuSTAR extragalactic surveys. Table 2.1 (taken and adapted from Klindt et al. 2023 in prep.) and Figure 2.4 provide the source statistics for the NSS80 catalogue. The total number detected in the full band, soft band, and hard band are 1078 (85%), 706 (55%), and 406 (32%), respectively. The analysis detailed in this chapter will focus on using the full band and hard band NSS80 detections; as previously mentioned, higher X-ray energies are able to penetrate absorbing dust and gas in obscured systems, so using this energy range is particularly beneficial for our obscured AGN analysis. Once more referring to

[4]There is a steep drop-off in exposure in peripheral regions. Furthermore, regions close to the FoV edge have poorly characterised background estimates.

Figure 2.1 to illustrate how NuSTAR's 3-24 keV energy range (shown by the red and blue solid bars) probes different parts of a moderately absorbed (N_H = 3 x 10^{23} cm^{-2}) AGN X-Ray spectrum at z $\sim$ 1 (shown by the solid black line); the decrease in emission seen at lower energies is the result of absorption; higher levels of emission are seen at higher energies for obscured systems such as this. This highlights why using the full and hard band NSS80 detections is beneficial for our study of obscured AGN at z $\sim$ 0.2 - 1.

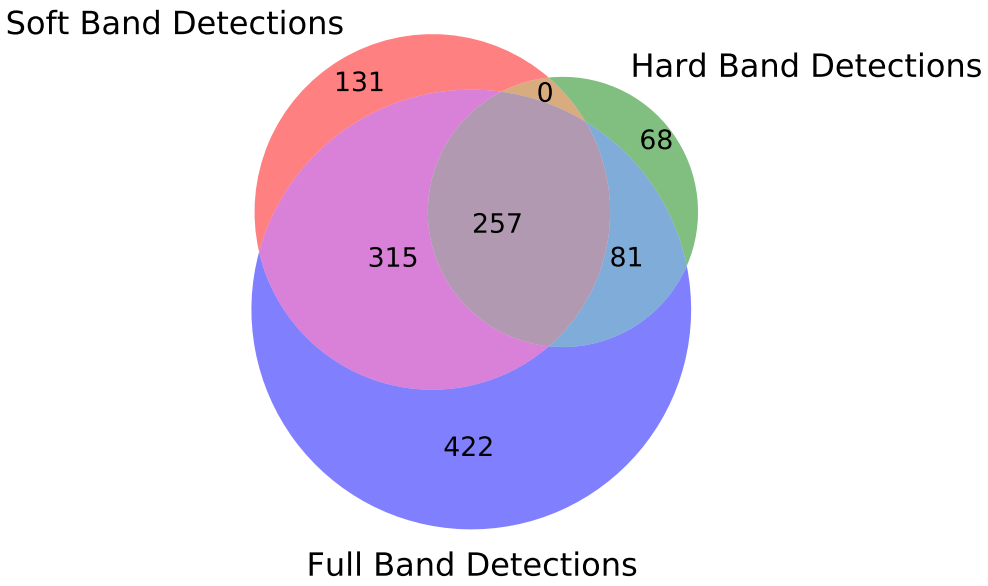

Figure 2.4 *Venn diagram showing the number of sources detected in each band in the primary NSS80 catalogue.*

2.3.1 Identifying counterparts

First, lower-energy (soft) X-ray counterparts to the NSS80 sources were searched for; this was done as NuSTAR's positional accuracy ranges from $\approx$ 8" to $\approx$ 20" for bright to faint sources, so it is beneficial to have more accurate X-ray positions to search for reliable optical/IR counterparts. The NuSTAR sources were cross-matched to (1) the Chandra Source Catalog Release 2.0 (CSC2.0; Evans (2020)), (2) the Fourth XMM-Newton Serendipitous Source Catalog, Tenth Data Release (4XMM-DR10; Webb et al. (2020)) and its stacked version (4XMM-DR10s; Traulsen et al. (2020)), and (3) the Swift XRT Point Source Catalogue (2SXPS; Evans et al. (2020)). In total, 964 NSS80 sources (76%) were identified to have identified soft X-ray counterparts.

Band	N	N_{spec}	N_z	$N_{z,failed}$	$N_{r,det}$	$N_{r<21}$
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Any band	1274	594	551	43	1058	623
F + H	81	35	32	3	71	41
F only	422	166	153	13	336	164
H only	68	10	8	2	37	18

Table 2.1 *Source statistics for the NuSTAR serendipitous survey.* ***Columns:*** ***(1)*** *F, and H refer to sources detected in the full (3–24 keV) and hard (8–24 keV) energy bands, respectively. "F + H", for example, refers to sources detected in the full and hard bands only, but not in the soft band, and "F only" refers to sources detected exclusively in the full band.* ***(2)*** *The number of sources detected post-deblending, for a given band or set of bands.* ***(3)*** *The number of sources for which (ground-based) optical spectroscopic observations were undertaken.* ***(4)*** *The number of sources with spectroscopic redshift measurements and the associated percentage (including robust and uncertain counterpart associations based on our Nway analysis).* ***(5)*** *The number of sources for which spectroscopic observations were undertaken, but lack a reliable redshift measurement (the majority of which have faint, red continuum spectra).* ***(6)*** *The number of sources with an associated optical counterpart detected in the r-band; magnitudes are obtained from SDSS, PanSTARRS, USNOB1 and NSC (NOAO Source Catalog).* ***(7)*** *The number of sources with an associated optical counterpart brighter than r = 21. This table only includes information on the full and hard band detections as these are the detection bands used in the following analysis. Taken and adapted from Klindt et al. 2023 in prep.*

In the absence of a reliable soft X-ray counterpart to provide accurate source positions for optical counterpart identification, another reliable approach is required to distinguish between potential optical counterparts. Looking at AGN emission at MIR wavelengths is another reliable way to trace AGN activity, particularly for the obscured population (obscuring dust largely emits at these wavelengths, whereas host galaxy star formation peaks in the FIR). In this vein, NuSTAR sources were matched to the IR all-sky WISE[5] survey (specifically CatWISE20, Marocco et al. (2021), and NEO-WISE) and the optical Pan-STARRS[6] survey (specifically PS1-DR2), to obtain more reliable source positions and matching multiwavelength counterparts.

[5]WISE has a 2" positional uncertainty

[6]Pan-STARRS has a 1" positional uncertainty

Multiple rounds of matching using NWAY[7] were undertaken in order to robustly find multiwavelength counterparts for the NSS80 X-ray sources. The preliminary round involved creating separate IR and optical base catalogues from CatWISE20 and PS1-DR2 respectively. This involved compiling all IR CatWISE20 and optical PS1-DR2 positions within 40" NuSTAR positions, extracting these IR/optical sub-catalogues around each NuSTAR source, and then matching them to each other with NWAY.

Following this preliminary round, another round of matching is carried out. This round uses sky-density and astrometric information to locate a 'best' counterpart for the NSS80 sources, with the aim of generating reliable photometric priors which can be used in another round of matching. Expected sky densities for NSS80 sources were calculated using Harrison et al. 2016b's logN-logS curves, whereas PS1-DR2 and CatWISE20 sky densities were determined from number counts within 40" of the NuSTAR sources. Using NWAY, the following counterpart probabilities were calculated: $\rho_{Cat,best}$ and $\rho_{PS,best}$, the probabilities that each WISE or Pan-STARS source is the correct counterpart to a specific X-ray source; $\rho_{Cat,any}$ and $\rho_{PS,any}$, the probabilities that *any* WISE or Pan-STARRS source is the correct counterpart to a specific X-ray source. The following constraints were then applied to the resultant matching, to ensure only sources with a high 'correct-match' probability were kept in the catalogue: $\rho_{Cat/PS,any} > 0.5$, $\rho_{Cat/PS,best} > 0.8$, and sources at high Galactic latitudes > 20 deg (source confusion is higher closer to the Galactic plane due to high densities of Galactic sources). The high quality matching results of NSS80 sources at high galactic latitudes from this round were then used to generate IR and optical magnitude and colour distributions, to be used as photometric priors for the next round of NWAY matching.

In this next round, the X-ray catalogue is split into two: one sample containing sources at low Galactic latitudes ($|b| \leq 20$ deg), and the other containing sources at high Galactic latitudes ($|b| > 20$ deg). The sample of sources at low Galactic latitudes is then matched geometrically to the base optical/IR catalogues, whereas the sources at high Galactic latitudes are matched using the photometric priors from the previous round. Any counterpart match with a resulting $\rho_{Cat/PS,best} >$

[7]nway is a source cross-matching tool for astronomical catalogs. It provides Bayesian match probabilities based on astronomical sky coordinates (RA, DEC), and can handle matching between multiple catalogs, as well as varying errors. nway can also use additional prior information, such as the magnitude or color distributions of the sources to match, to aid the matching process Buchner et al. (2021).

10% is then stored and, from this, a refined catalogue is created which lists only the 'best' Pan-STARRS and CatWISE20 match for each NSS80 source; this catalogue includes X-ray information, multi-band positions and photometry, and key Nway information (e.g. $\rho_{Cat/PS,best/any}$). The matching statistics of this refined catalogue are defined as follows: 1247/1274 NSS80 sources have at least one CatWISE20/PS1-DR2 match with a probability $> 10\%$ (1221/1274 from CatWISE20, 867/1274 from PS1-DR2); 912/1121 of the CatWISE20 and 647/867 PS1-DR2 NWAY sources were matched to a soft X-ray counterpart, and therefore have more reliable counterpart associations than those solely with NuSTAR X-ray positions; of the 912 soft X-ray counterparts with a CatWISE20 association, 648 have a PS1-DR2 association. For the 264 CatWISE + soft X-ray counterpart sources which do not have a PS1-DR2 association, the NSS80 team matched to the following catalogues: (1) the SDSS photometric catalogue DR12 (Alam et al., 2015), (2) the second data release of the NOIRLab Source Catalog (NSC2; Nidever et al. (2021)), (3) the Dark Energy Survey Data Release 2 (DES-DR2; Abbott et al. (2016)), and (4) the USNOB1 catalogue (Monet et al., 2003). They additionally matched to the aforementioned surveys for the following sources: the 26 NuSTAR sources with soft X-ray counterparts but no CatWISE20 or PS1-DR2 counterparts; the 116 NuSTAR sources with CatWISE20 counterparts but no PS1-DR2 or soft X-ray counterparts ; the 193 NuSTAR sources with CatWISE20 and PS1-DR2 counterparts, but no soft X-ray counterparts. The result of all this multiwavelength cross-matching is an optical completeness of 1165/1274 (91%) for the NSS80 sample, comprised of 918/964 soft X-ray and 247/310 CatWISE20 associations.

2.3.2 Spectroscopic follow-up

The NSS80 team undertook an extensive campaign of ground-based spectroscopic follow-up to obtain new spectroscopic source redshifts and classifications for NSS80 sources. They achieved this by utilising a variety of telescopes located at different geographical latitudes (covering the wide range in declination of the sources) at different times of the year (covering the wide range in right ascension of the sources). In the Northern hemisphere, they used a combination of the Hale Telescope at the Palomar Observatory and Keck I at the W. M. Keck Observatory. In the Southern hemisphere they used a combination of the Very Large Telescope (VLT) at the European Southern Observatory and the Southern African Large Telescope (SALT) which is part of the South African Astronomical Observatory

in Sutherland.

Spectroscopic redshifts were obtained by identifying and approximately measuring Gaussian profiles of emission and absorption lines in their observed-frame wavelengths. These profiles were then cross-matched with a look-up table of wavelength ratios based on the emission and absorption lines observed in AGN and galaxy spectra. The manual and automatic redshifting software (MARZ; Hinton (2016)) was also used to cross-match redshifts with the identified emission and absorption lines. This information was also used to aid spectral classification by visual inspection of the NuSTAR sources.

The resultant NSS80 spectroscopic catalogue contains 612 sources. 594 of these are unique spectroscopic entries, along with 18 additional entries. The results of this campaign and existing archival spectroscopy are redshifts for 551 NSS80 sources: 58 spectroscopically confirmed Galactic sources, 298 classified broad emission line objects (BL), 181 narrow emission line objects (NL), 10 galaxies (Gxy), one galaxy cluster, and three unclassified (though with $z > 0$ estimates). In addition, the NSS80 team identified 4 dual AGN pairs and one BL Lac candidate. The criteria used to identify BL, NL and Galaxy objects are as follows: **BL:** if any permitted line is broader than the forbidden lines, or if a single-line measurement satisfies FWHM > 1000 kms^{-1}; **NL:** if the permitted lines are of equivalent width to the narrow forbidden lines; **Gxy:** if only absorption lines are detected. There are also a few optical classifications with a "?" symbol appended (12 BL? and 17 NL?). These are cases where the NSS80 team had, by eye, deemed it ambiguous whether the permitted lines are broad or not. However, regardless of the optical classification, the vast majority of the sources are expected to be AGN due to the detection of luminous X-ray emission at high X-ray energies of > 3 keV, which is further confirmed for most sources by the identification of strong optical emission lines often superimposed on a power-law spectrum. Figure 2.5 shows the classification information for the NSS80 catalogue.

Figures 2.6, 2.7 and 2.8 show the distribution of the NSS80 sample in luminosity, redshift, optical and X-ray flux space. The broad and hard energy coverage of NuSTAR, in addition to NSS80's high spectroscopic coverage, robust counterpart identification and high quality spectroscopy, make the NSS80 catalogue a very worthwhile sample with which to carry out obscured AGN analysis.

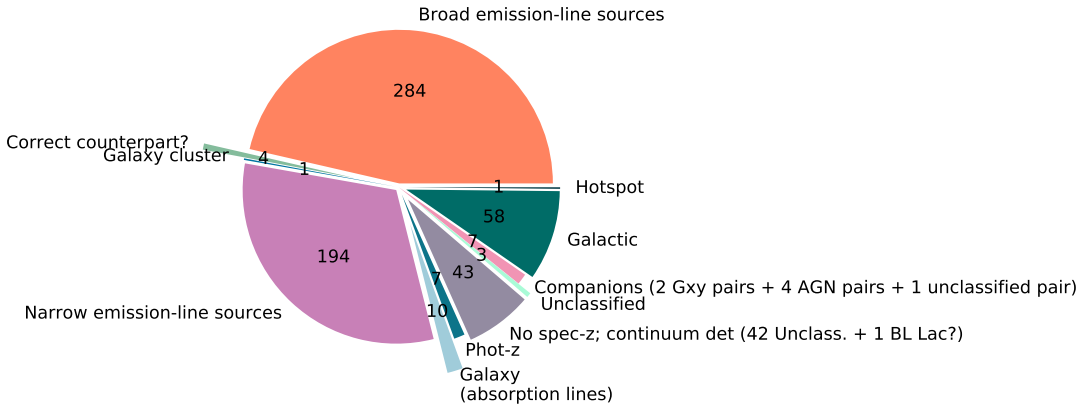

Figure 2.5 *Pie chart displaying the spectroscopic classifications of the sources in the NSS80 spectroscopic catalogue.*

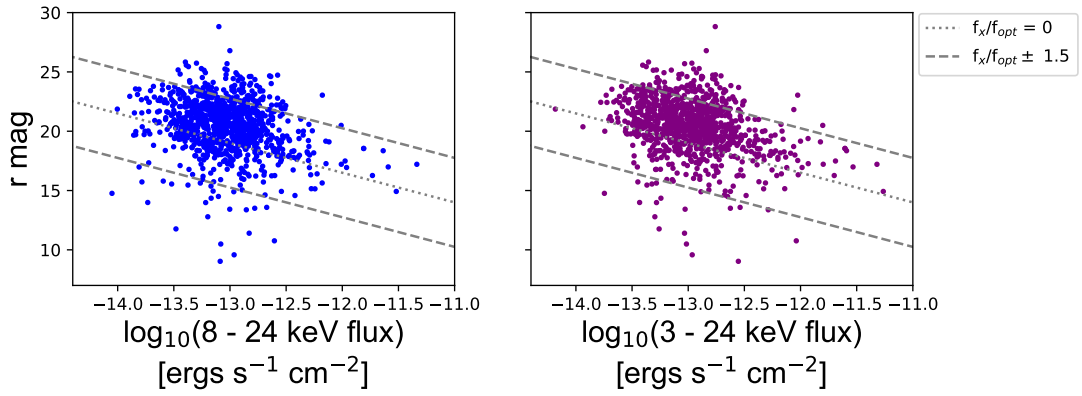

Figure 2.6 *Plot of r mag vs Hard band (left, log_{10} 8-24 keV) and Full band (right, log_{10} 3-24 keV) flux for the full NSS80 catalogue. Dotted and dashed grey lines are lines of constant f_X/f_{opt}, and illustrate the region within which most AGN are expected to lie: -1.5 $\leq f_X/f_{opt} \leq$ 1.5. These lines are defined for soft X-rays (0.5 - 2 keV), hence the offset from my harder X-ray sample.*

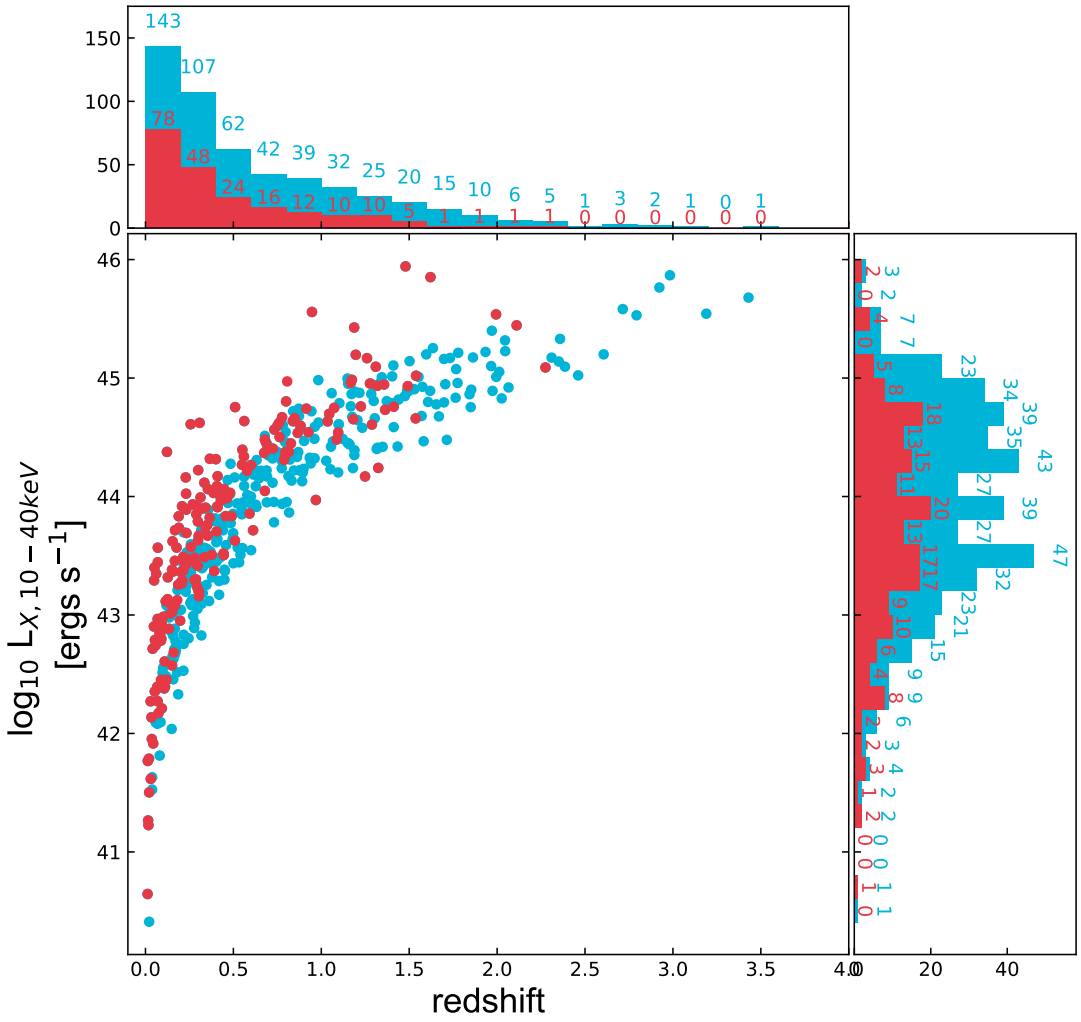

Figure 2.7 *Plot of luminosity vs redshift for the NSS22 source catalogue, with histograms for each axis, for hard (red, 8 - 24 keV) and full (blue, 3 - 24 keV) band detections. Luminosities are calculated from the observed frame NuSTAR fluxes, assuming an effective photon index of* $\Gamma = 1.8$.

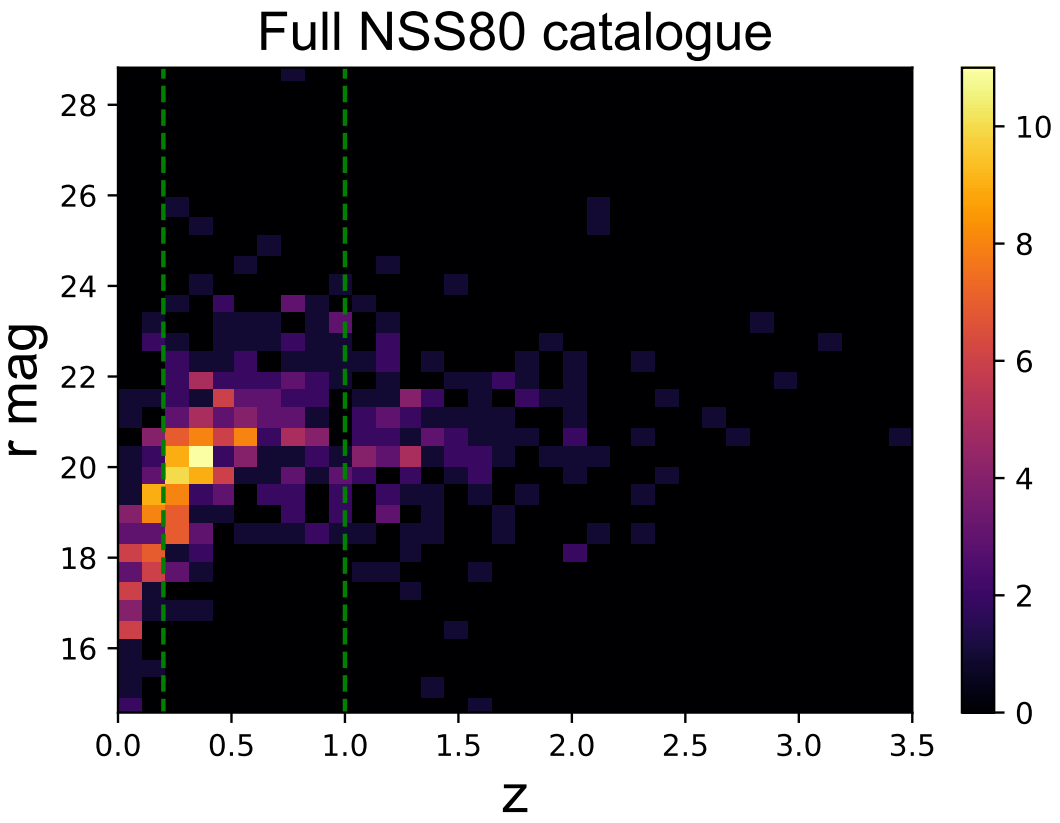

Figure 2.8 *2D histogram of r magnitude vs z for the full NSS80 catalogue. The dotted green lines illustrate the redshift range we are using for the XLF analysis described in Chapter 2.*

2.4 Constructing a sample

The NSS80 sources are a selection of astronomical objects which exist across a broad range of redshifts and have a range of optical properties. However, in order to constrain an accurate measurement of obscured versus unobscured black hole growth at z $\sim$ 0.2 - 1, it is essential to apply various cuts to this sample of sources. These cuts serve to create a highly spectroscopically complete sub-sample of sources with well matched multi-wavelength data, that we reasonably believe to be BL/NL AGN lying within our desired redshift range. Furthermore, after applying these cuts, we can correct for the incompleteness of the sample. It is only after carefully considering our sample and its corresponding completeness, that we can move on to generating new measurements of the obscured and unobscured AGN XLFs.

The first cut I applied to the full NSS80 sample was a Galactic plane cut. I essentially cut all sources that had a Galactic latitude between -10 and 10 degrees; i.e. cutting out the Galactic plane. This cut decreases the likelihood of Galactic sources such as stars or X-ray binaries contaminating our sample of extra-galactic AGN. Figure 2.9 illustrates the NSS80 sources which lie both within and outside of the Galactic plane. We can see from Figure 2.9 that making this cut within 10 degrees of the Galactic plane removes a large fraction of the unclassified sources, as well as most of the sources with Galactic origins. The next cut we initially applied was to remove all sources that lie at declinations of -30 degrees or less. This was so that we could focus on the NSS80 sources that lie in the northern hemisphere, as this is where the PanSTARRS coverage is greatest. When used in combination, these two cuts based on positional co-ordinates isolate the particular areas of sky within which we have a high probability of isolating the NSS80 sources which are extra-galactic AGN with good multi-wavelength coverage. However, upon checking the multi-wavelength coverage of our final sample before and after applying a -30 degrees declination cut, the multi-wavelength completeness of the resultant sample was not hugely affected. As such, we decided to not apply this cut so as to keep our sample size as large as possible (whilst maintaining high multi-wavelength completeness). After applying just the Galactic plane cut to the full NSS80 catalogue, we are left with 981 out of the original 1274 NSS80 sources: 303 of which are detected in the hard band and 825 of which are detected in the full band. These 981 sources make up our X-ray selected sample.

After making the positional cuts, the next cut we applied compares the

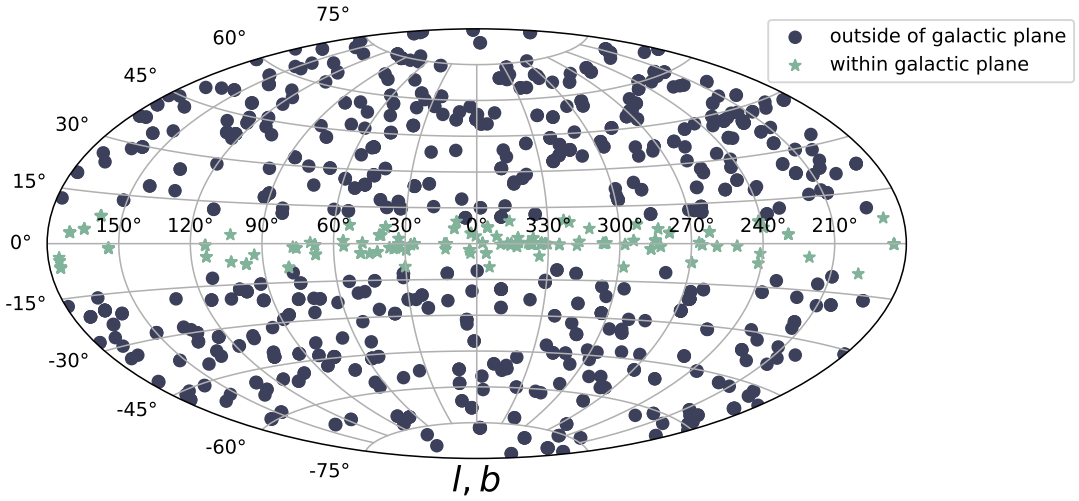

Figure 2.9 *Figure showing aitoff projection of NSS80 sources within 10 degrees (pale green stars) and outside of (dark blue circles) the Galactic plane, in Galactic co-ordinates. Sources within 10 degrees of the Galactic plane are removed from our sample.*

NWAY identified AGN (optical/IR NSS80 counterparts) to the spectroscopically confirmed AGN (NSS80 spectroscopic targets); this cut (listed as a flag in the NSS80 catalogue) identifies and excludes sources where there is a lower reliability of the spectroscopic target matching the adopted optical counterpart (or vice versa). The NSS80 team created this flag having assessed the reliability of each source's spectroscopic target. This involved, for each source, taking the positional offset between the best CatWISE20 counterpart and the followed up spectroscopic target. Any optical counterpart-spectroscopic target pair with $<$ 5" positional offset is deemed as being a 'reliable match' (5" broadly corresponds to the positional uncertainty of CatWISE20). To corroborate these matches and to investigate those which had a positional offset $>$ 5 ", the NSS80 team visually inspected all NSS80 sources to determine the optical counterpart-spectroscopic target matching reliability. The product of their analysis is the aforementioned flag, which identifies NSS80 sources which have a low reliability of their spectroscopic targets and optical counterparts being true matches. We apply this cut to our sample to only include sources which have high matching probabilities, reducing false associations and therefore any bias in our analysis. After applying this cut, we are left with 547 sources: 201 of which are detected in the hard band and 501 of which are detected in the full band.

Following this, we applied an additional cut which which accounts for the reliability with which NWAY matches optical counterparts to NSS80 X-ray

sources. Despite the careful multi-wavelength cross-matching procedure outlined in Section 2.3, not all NWAY matches will be true counterpart associations. This can be due to many factors: the dependence of accurate cross-matching on soft X-ray counterpart information, the Galactic latitude of sources (source confusion increases at high density fields within the Galactic plane), the depth of the optical/IR photometric surveys (shallow imagining can miss counterparts that could have been picked up by deeper surveys). By applying a cut which selects only the high-reliability counterpart matches, we can ensure all the sources in our sample have a high probability of correct multi-wavelength counterpart information, establishing a cleaner and more reliable sample. To do this, we applied the same cut used in Klindt et al. (in prep.) to identify high probability CatWISE20 and PS1-DR2 matches, in order to minimise false associations (and therefore biased results); selecting sources which have NWAY thresholds of $\rho_{Cat,best} > 0.4$ and $\rho_{Cat,any} > 0.5$ or, in cases where these thresholds are not met, $\rho_{PS,best} > 0.4$ and $\rho_{PS,any} > 0.5$. After applying this cut, we are left with 515 out of the 981 NSS80 sources in our X-ray selected sample: 188 of which are detected in the hard band and 478 of which are detected in the full band.

The final cuts applied to our NSS80 sample are (i) selecting sources based on whether they are detected in certain X-ray energy bands, and (ii) applying a suitable r magnitude cut such that our final samples have high spectroscopic completeness, whilst spanning a decent range of optical magnitudes, so we can then correct for the fainter counterparts. Regarding (i): I analyse sources with hard band (8-24 keV) detections and full band (3 - 24 keV) detections in separate subsamples; as such, my final NSS80 XLF analysis will be to constrain independent BL (unobscured) and NL (obscured) XLFs for the hard band detected sample, and then also for the full band detected sample. Regarding (ii): Applying an r band magnitude cut to the 981 sources in our X-ray selected sample will inevitably make our sample incomplete, and reduce the sample size that we use for our XLF measurements. However, the advantages of such a cut are that we can a) correct for the introduced incompleteness and b) the remaining r-limited sub-sample will have a high spectroscopic completeness.

In order to decide what r magnitude limit to apply to our sample, we need to first consider how our our sample's spectroscopic completeness changes with differing r magnitude cuts. Applying a low r magnitude cut would mean we have a very spectroscopically complete sample (as brighter sources tend to be easier to find and classify), but it would mean we miss out on including the fainter population

in our analysis. Conversely, applying a high r magnitude cut would mean we have a larger and more diverse sample, but with a lower overall spectroscopic completeness. In order to find the balance between pushing our sample to contain sources with fainter magnitudes, whilst maintaining a highly spectroscopically complete sample, we must assess our sample's spectroscopic completeness as a function of r magnitude. We can then use this to inform our choice of r magnitude cut.

Figures 2.10 and 2.11 illustrate the cumulative spectroscopic completeness of our hard and full band NSS80 samples after every cut we apply, as a function of r magnitude limit. Figures 2.12 and 2.13 illustrate the binned spectroscopic completeness of our hard and full band NSS80 samples after every cut we apply, as a function of r magnitude limit. These figures clearly illustrate how, with every cut we make to the NSS80 sample, the overall cumulative spectroscopic completeness of the resultant sample increases. In some r magnitude bins, our spectroscopic completeness pushes as high as 100% after our cuts - this occurs at brighter magnitudes, where high quality spectroscopy (and hence classifications/redshifts) is more easily obtained. The two red dashed lines on these plots emphasize our sample's spectroscopic completeness after enforcing specific r magnitude limits. An r magnitude cut of 22 would leave our hard and full band samples with an overall spectroscopic completeness of 98.2% and 97.5% respectively, with 94.1% and 96.8% in the respective r mag = 22 bins. Pushing to an r magnitude cut of 24 would leave our hard and full band samples with an overall spectroscopic completeness of 97.0% and 96.7% respectively, with 50.0% and 71.4% in the respective r mag = 24 bins. Despite the slight decrease in overall spectroscopic completeness, by applying the higher r magnitude cut of 24 we can increase our sample size and include fainter sources, whilst still maintaining a highly spectroscopically complete sample. In doing so, we have applied the final cut to our NSS80 sample.

It is worth noting that we are not only reliant on PanSTARRS data - most of our counterparts are identified via WISE and that, where necessary, dedicated follow-up r-band photometry was carried out to identify the optical sources associated with the previously identified WISE counterparts to the NuSTAR sources. Hence, an R cut of 24 is appropriate due to these reasons. 2.14 illustrates the 1/rmag error as a function of r band magnitude for our sample.

Our resultant NSS80 samples are as follows. In the hard band we have 163 sources selected in the X-ray, and after enforcing an r-band cut: 85 BL, 1 BL?, 64 NL,

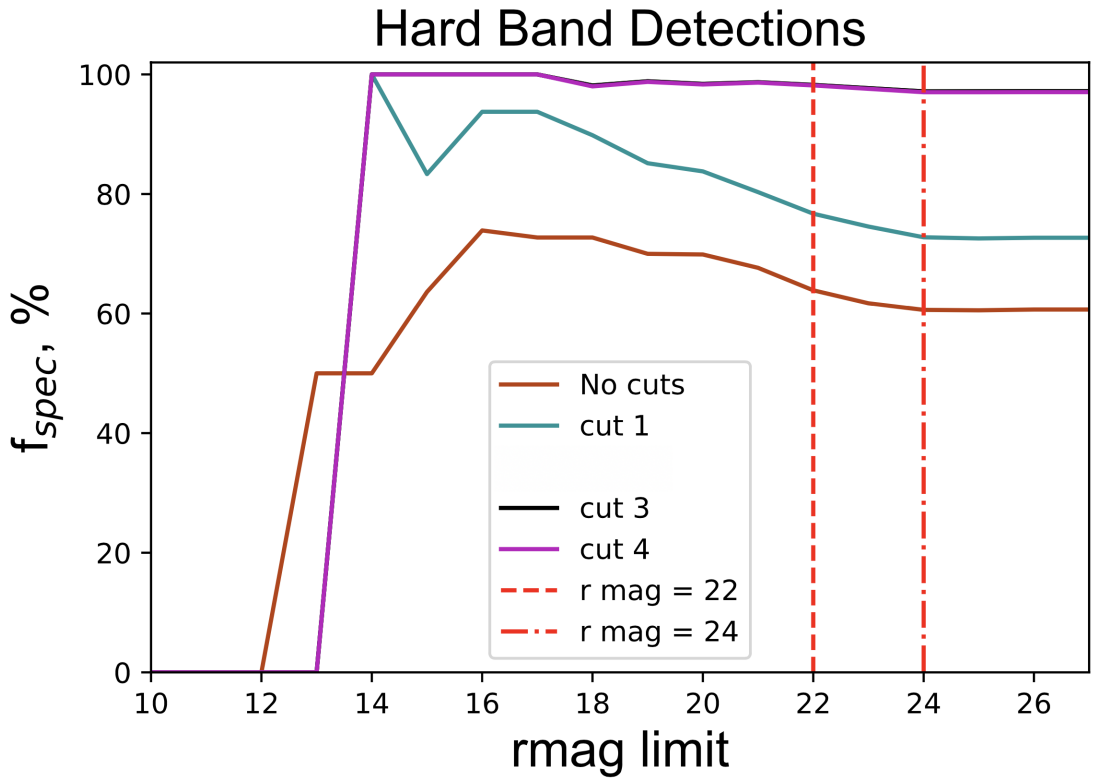

Figure 2.10 *The cumulative spectroscopic completeness of the hard band (8 - 24 keV) detected NSS80 sources, after specific cuts are made to the sample, as a function of r magnitude limit. The solid red line illustrates the full, uncut NSS80 sample; the solid teal line illustrates the sample remaining after applying our Galactic plane cut; the solid black line illustrates the sample remaining after excluding sources where there is a lower reliability of the spectroscopic targets matching the adopted optical counterpart; the solid purple line illustrates the sample remaining after excluding sources with lower X-ray to optical counterpart matching reliabilities (this is plot directly on top of the previous cut, as this step does not change the overall cumulative spectroscopic completeness of our remaining sample); the dashed red line indicates an r mag limit of 22; the dot-dashed red line indicates an r mag limit of 24.*

4 NL?, 5 unclassified, and 9 non-AGN contaminants (e.g. Galaxies/clusters). In the full band we have 412 sources selected after a r-band cut and using X-ray selection: 234 BL, 10 BL?, 125 NL, 10 NL?, 15 unclassified, and 32 non-AGN contaminants. Figures 2.15 to 2.18 show the X-ray luminosities, redshifts, classifications, optical magnitudes and X-ray/optical flux distributions of our resultant NSS80 sub-sample.

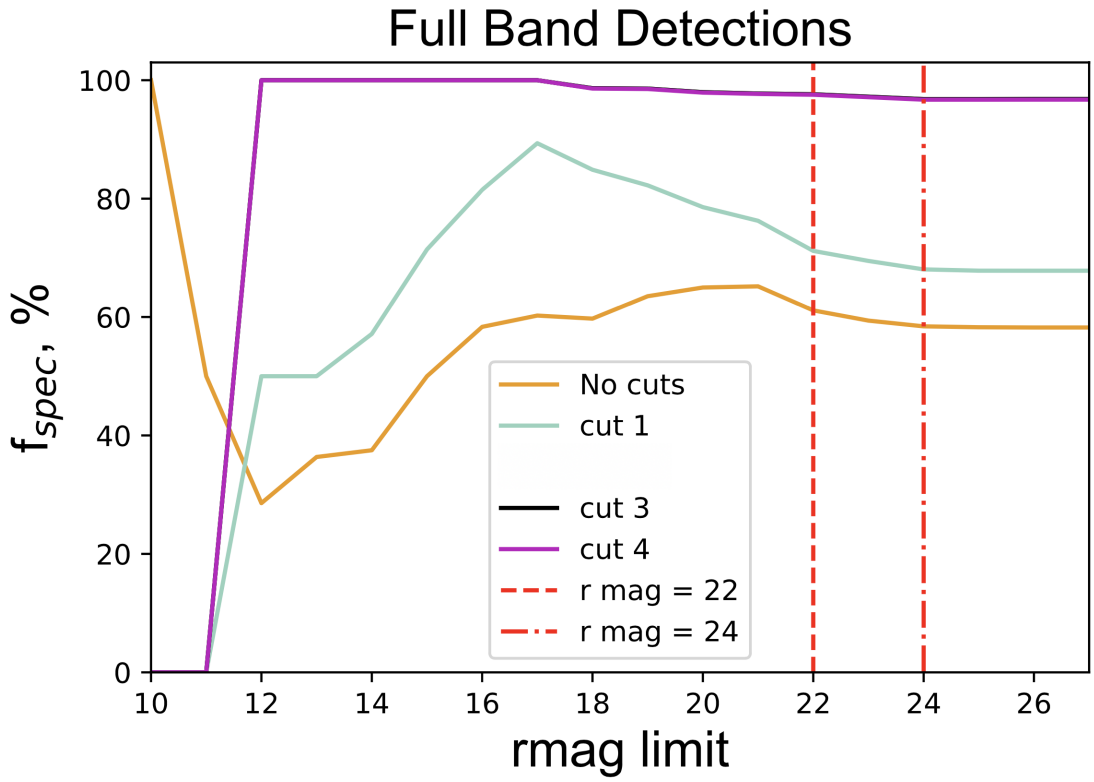

Figure 2.11 *The cumulative spectroscopic completeness of the full band (3 - 24 keV) detected NSS80 sources, after specific cuts are made to the sample, as a function of r magnitude limit. The solid yellow line illustrates the full, uncut NSS80 sample; the solid light blue line illustrates the sample remaining after applying our Galactic plane cut; the solid black line illustrates the sample remaining after excluding sources where there is a lower reliability of the spectroscopic targets matching the adopted optical counterpart; the solid purple line illustrates the sample remaining after excluding sources with lower X-ray to optical counterpart matching reliabilities (this is plot directly on top of the previous cut, as this step does not change the overall cumulative spectroscopic completeness of our remaining sample); the dashed red line indicates an r mag limit of 22; the dot-dashed red line indicates an r map limit of 24.*

2.5 Creating sensitivity curves

Having constructed our sample from the NSS80 catalogue, we moved on to creating sensitivity curves for this reduced sample. These are necessary for the X-ray Luminosity Function calculations described later in Section 2.6.

As previously mentioned in Section 2.4, the total survey area sensitive to a given flux (i.e. area curve, see the solid blue line in Figure 2.3) for the full NSS80 sample was determined by summing individual sensitivity curves for a given background and exposure map. These individual area curves were determined by calculating

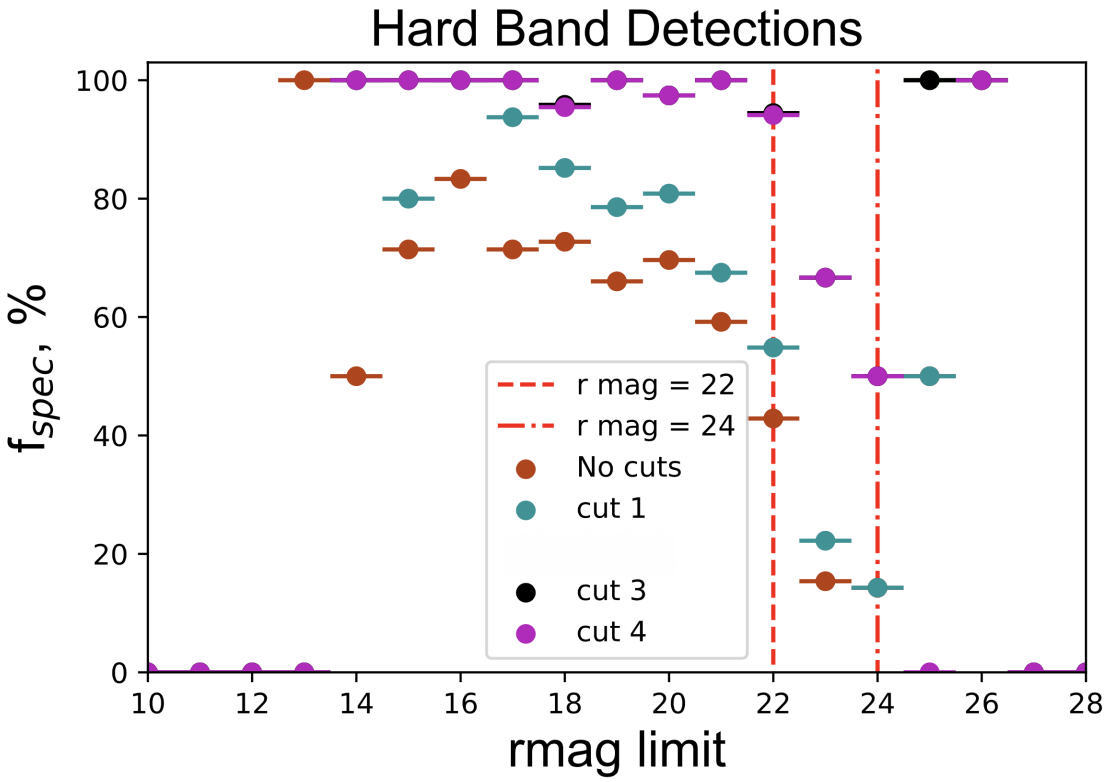

Figure 2.12 *The binned spectroscopic completeness of the hard (8 - 24 keV) band detected NSS80 sources, after specific cuts are made to the sample, as a function of r magnitude limit. The solid red points illustrate the full, uncut NSS80 sample; the solid teal points illustrate the sample remaining after applying our Galactic plane cut; the solid black points illustrate the sample remaining after excluding sources where there is a lower reliability of the spectroscopic targets matching the adopted optical counterpart; the solid purple points illustrate the sample remaining after excluding sources with lower X-ray to optical counterpart matching reliabilities (this is plot directly on top of the previous cut, as this step does not change the overall cumulative spectroscopic completeness of our remaining sample); the dashed red line indicates an r mag limit of 22; the dot-dashed red line indicates an r map limit of 24.*

the flux limit at the detection threshold for every point in the NuSTAR image. However, now we have cut down our sample from the full NSS80 catalogue, it is essential to constrain new hard band and full band area curves which are specific to our analysis.

To create these curves, we identified all the fields left after applying our Galactic plane cut to the full NSS80 catalogue; this accounts for all the NuSTAR fields which enter our sample. Using this list we identified the specific mosaics within which our sources were located, and extracted the pre-existing individual hard and full band sensitivity curves for those mosaics. The final hard and full band area curves for our sample was then determined by summing these individual

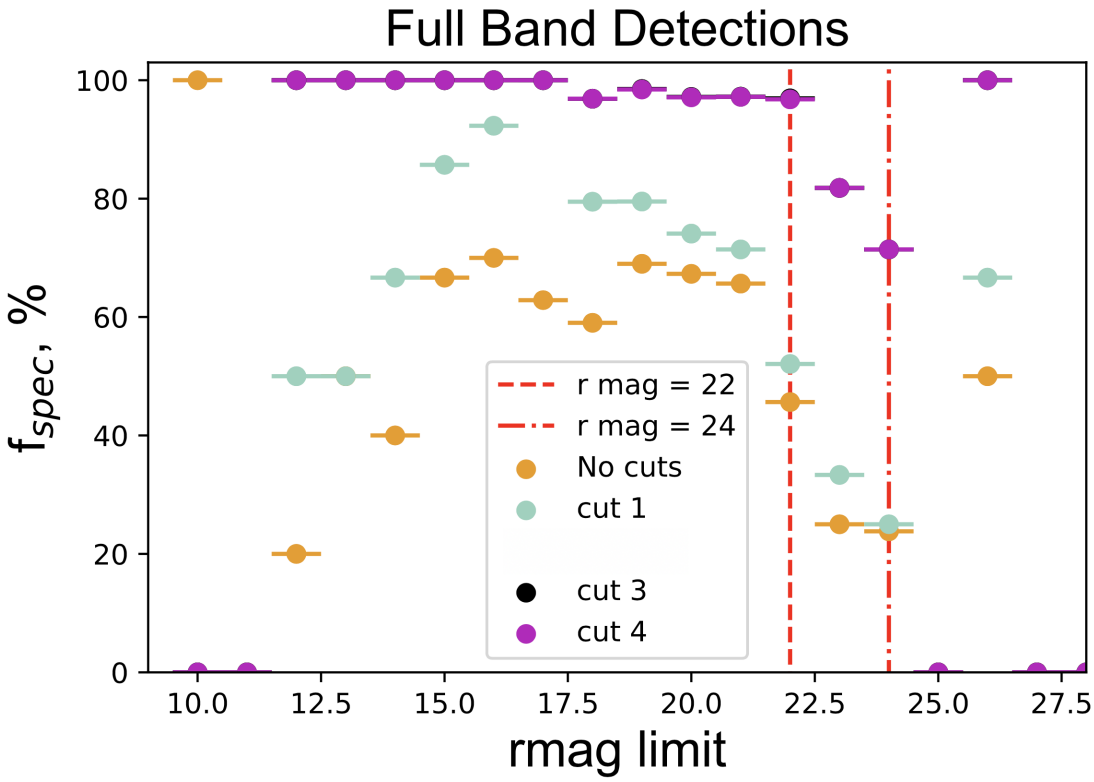

Figure 2.13 *The binned spectroscopic completeness of the full (3 - 24 keV) band detected NSS80 sources, after specific cuts are made to the sample, as a function of r magnitude limit. The solid yellow points illustrate the full, uncut NSS80 sample; the solid light blue points illustrates the sample remaining after applying our Galactic plane cut; the solid black points illustrate the sample remaining after excluding sources where there is a lower reliability of the spectroscopic targets matching the adopted optical counterpart; the solid purple points illustrate the sample remaining after excluding sources with lower X-ray to optical counterpart matching reliabilities (this is plot directly on top of the previous cut, as this step does not change the overall cumulative spectroscopic completeness of our remaining sample); the dashed red line indicates an r mag limit of 22; the dot-dashed red line indicates an r map limit of 24.*

sensitivity curves; Figure 2.19 and 2.20 illustrates the 3-24 keV and 8-24 keV area curves for our sample, and the previously mentioned Figure 2.3 shows these area curves in comparison to the full NSS80 catalogue, the area curve adopted in Aird et al. (2015), and the full NSS40 catalogue. Using our newly constrained area curves, we can move on to generating new calculations for the z $\sim$ 0.2 - 1.1 BL and NL AGN XLFs using our highly spectroscopically complete, robustly identified NSS80 sample.

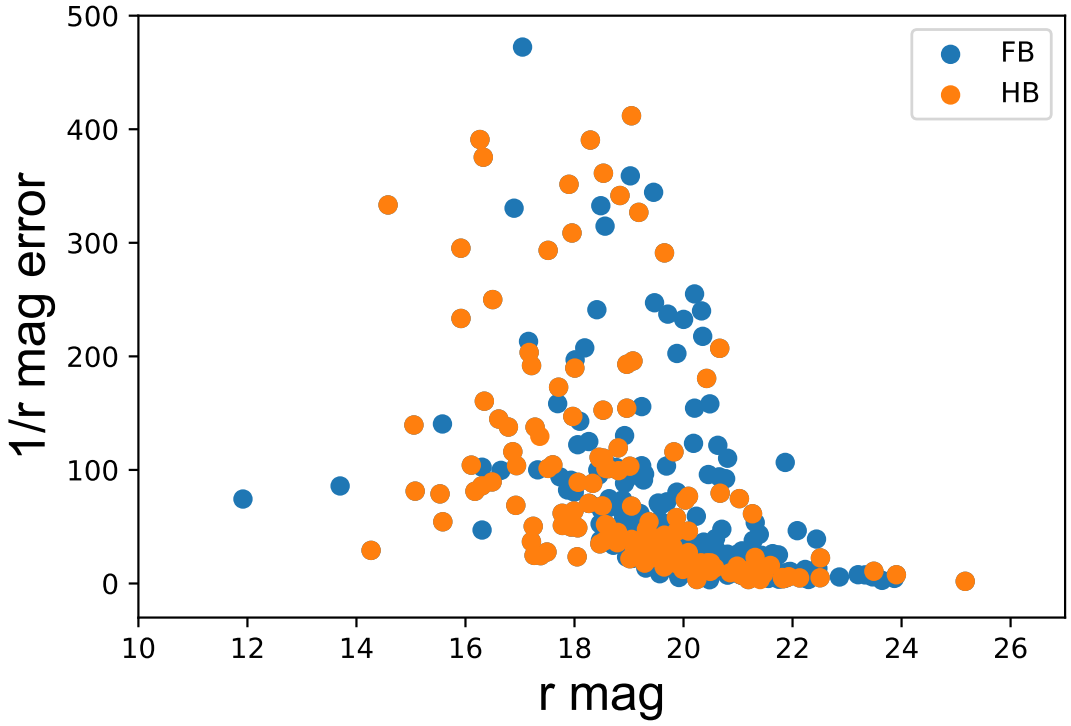

Figure 2.14 *Showing the 1/r mag error of our NSS80 subsample as a function of r band magnitude, for our hard (orange points) and full (blue points) band detections.*

2.6 Maximum Likelihood fitting approach to determine the completeness - corrected AGN XLF

In this work, I adopt a maximum likelihood fitting (referred to as MLF from this point onward) approach to generate my unbinned XLF determinations, as showcased in Aird et al. (2008) and Aird et al. (2010). Maximum likelihood fitting is a method of estimating the parameters of an assumed distribution, given some observed data. This is achieved by maximizing a likelihood function so that, under the assumed statistical model, the observed data is most probable. The point in the parameter space that maximizes the likelihood function is called the maximum likelihood estimate, and provides the best fitting parameters to describe the XLF, given the observed data.

In this work, I use such an approach to determine our luminosity functions, as using MLF allows me to fold in corrections for multiple selection effects present in

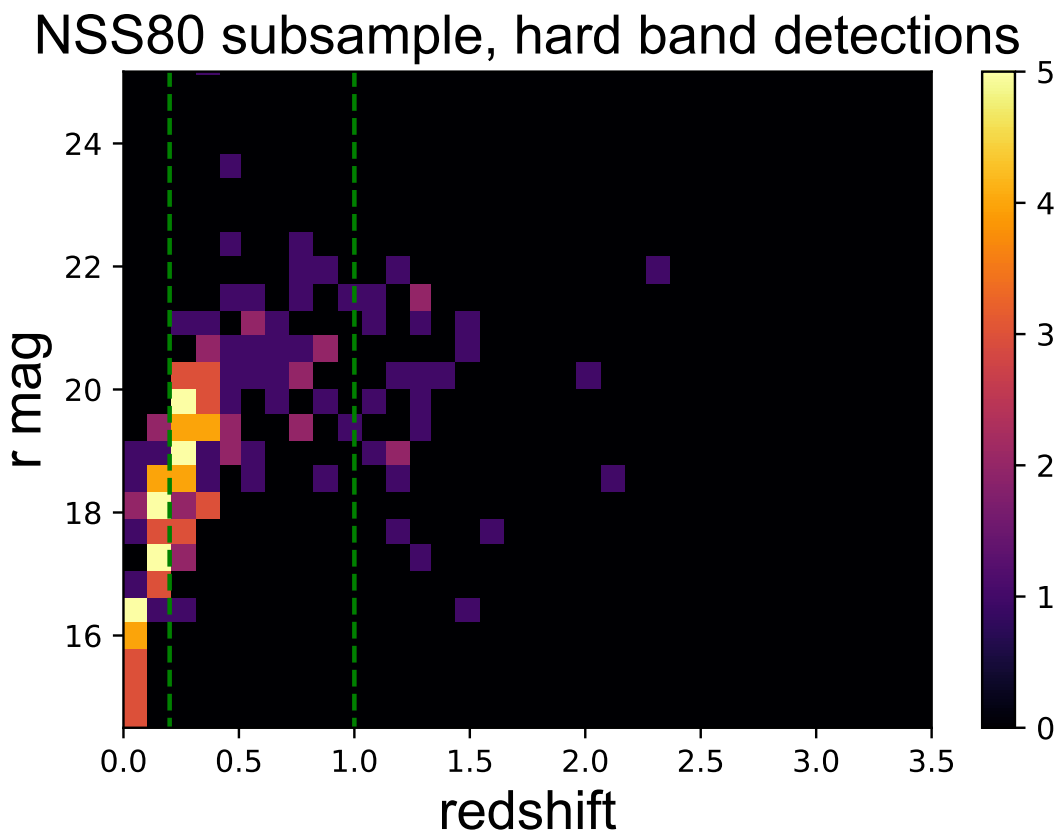

Figure 2.15 *r band magnitude vs. redshift for the hard (8-24 keV) band detections in our final NSS80 sub-sample.*

the overall sample and on a source-by-source basis. Furthermore, this approach does not involve any binning in luminosity, so all information from the individual sources is retained in my calculations.

In the following calculations, I assume independent XLFs for both the broad line and narrow line AGN populations. These XLFs are defined as $\phi_{BL}(L_x \mid \theta_{BL})$ for the broad line population and $\phi_{NL}(L_x \mid \theta_{NL})$ for the obscured population, where θ indicates a distinct set of model parameters for each population. These parameters characterise the double power law XLF space density that I am adopting:

$$\phi(L_X|\theta) = \frac{d\Phi(L_X)}{d\mathrm{logL}_X} = \frac{k_{norm}}{\left(\frac{L_X}{L_*}\right)^{\gamma_1} + \left(\frac{L_X}{L_*}\right)^{\gamma_2}} \tag{2.1}$$

where k_{norm} (normalisation factor), L_* (break luminosity), γ_1 and γ_2 (describing the slope of the XLF at low and high luminosities respectively) are the parameters

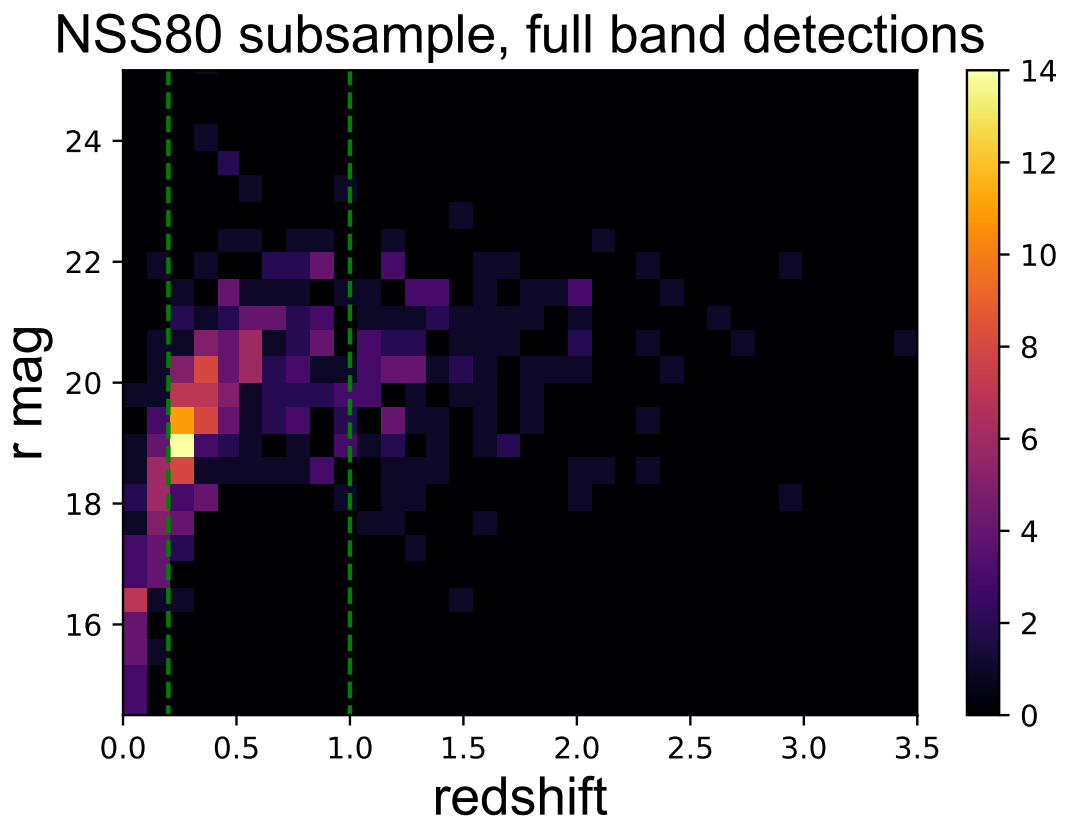

Figure 2.16 *r band magnitude vs. redshift for the full (3-24 keV) band detections in our final NSS80 sub-sample.*

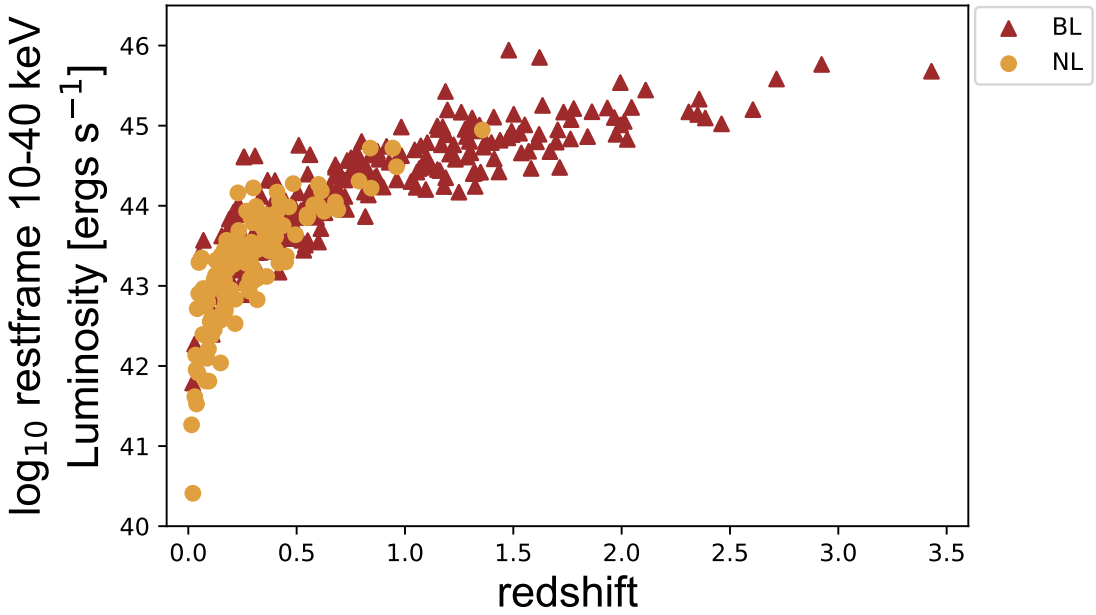

Figure 2.17 *10-40 keV X-ray luminosity vs. redshift for the full (3 - 24keV) band detections in our final NSS80 sub-sample. Yellow circles indicate NL sources and red triangles indicate BL sources.*

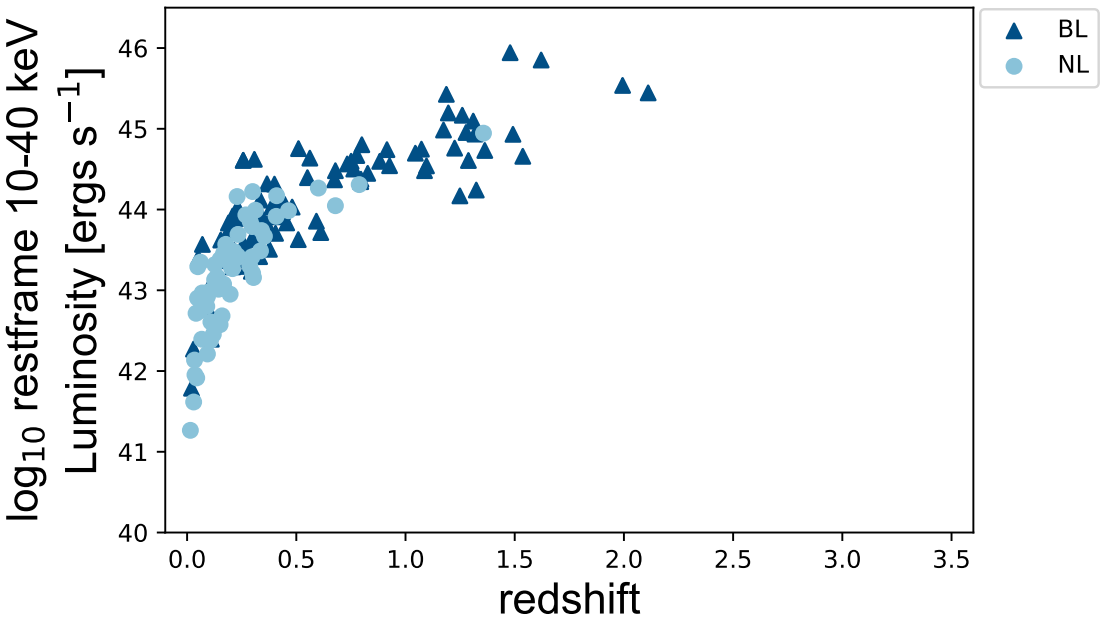

Figure 2.18 *10-40 keV X-ray luminosity vs. redshift for the hard (8-24 keV) band detections in our final NSS80 sub-sample. Light blue circles indicate NL sources and dark blue triangles indicate BL sources.*

(θ) to be evaluated to find those which satisfy the maximum likelihood estimate, and $\frac{d\Phi(L_X)}{dlogL_X}$ is the AGN number density per unit co-moving volume per logL_X as a function of X-ray luminosity L_X. This equation for the AGN XLF is represented by the term $\phi(L_X|\theta)$: the AGN XLF as a function of L_X, given parameters θ.

2.6.1 Defining the likelihood function

Having defined the functional form of the adopted model XLF, the next step is to define the likelihood of obtaining my real data di for a given set of model assumptions (θ):

$$\mathcal{L}(d_i \mid \theta) = e^{-\lambda(\theta)} \prod_{i=1}^{M} l_i(d_i \mid \theta) \tag{2.2}$$

where $\lambda(\theta)$ is the expected number of detected sources, θ are the parameters defining the assumed model XLF (see Equation 2.1) and $l_i(d_i \mid \theta)$ is the contribution of each individual source to the likelihood based on the data, d_i, for that source. This individual contribution to the likelihood of each of our

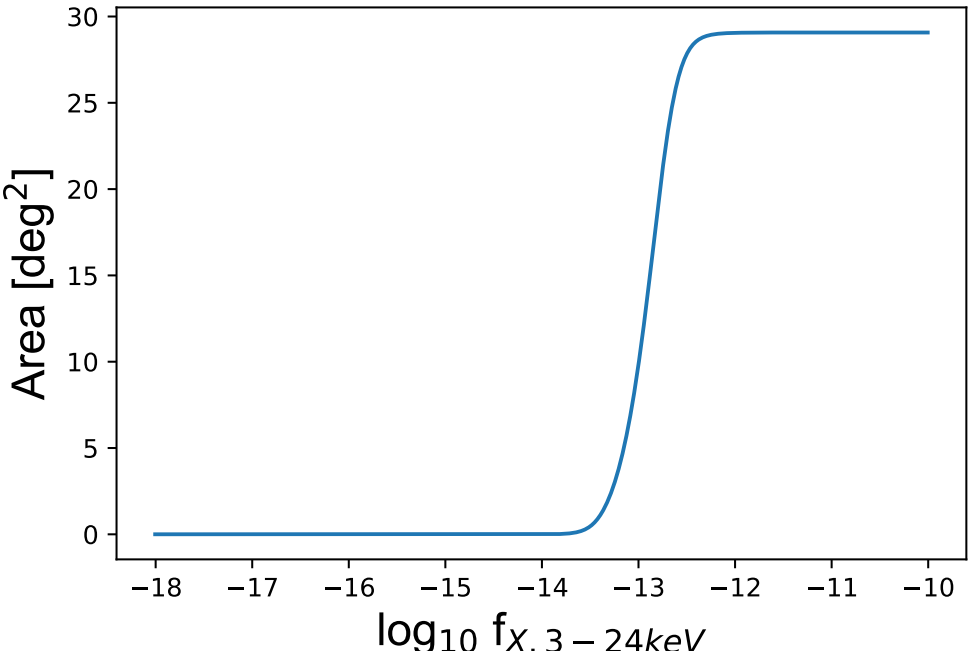

Figure 2.19 *3-24 keV curve for the NSS80 sub-sample we adopt in our XLF analysis.*

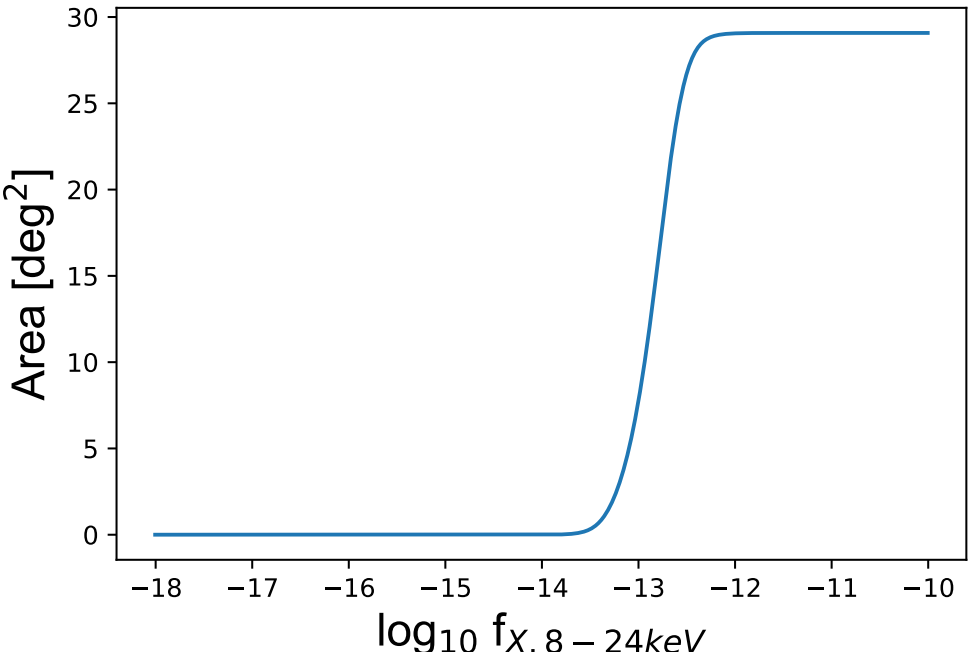

Figure 2.20 *8-24 keV area curve for the NSS80 sub-sample we adopt in our XLF analysis.*

sources is given by the XLF evaluated at $L_{X,i}$, z_i:

$$l_i(d_i \mid \theta) = \phi_i(L_{x,i}, z_i \mid \theta) \tag{2.3}$$

Next, I define the expected number of detectable sources, λ given a model of parameters θ. Over a given redshift range (between z_1 and z_2) and X-ray luminosity range (between $L_{X,1}$ and $L_{X,2}$), the expected number of detectable sources is given by integrating the XLF, folded through the relevant selection functions, over a defined L_X and z range. I start by defining it as such:

$$\lambda(\theta)_1 = \int_{L_{X,1}}^{L_{X,2}} d\log L_X \int_{z_1}^{z_2} dz \left[\phi(L_X, z \mid \theta)\ A(f_X(L_X, z))\ \frac{dV}{dz} \right] \tag{2.4}$$

where $\phi(L_x,\ z \mid \theta)$ is the assumed model luminosity function described by parameters θ; $A(f_x(L_X,\ z))$ is the term accounting for the area of survey sensitive to a source of a given X-ray flux, f_X, which depends on the intrinsic luminosity, L_X, and redshift, z (describing the probability of X-ray source detection, i.e. the survey area curves described in Figures 2.19 and 2.20); and $\frac{dV}{dz}$ is the differential co-moving volume per unit area as a function of z.

Whilst this expected number does account for the survey's X-ray sensitivity, it does not yet account for the incompleteness introduced by imposing our r magnitude ≤ 24 optical flux limit. In order to do this, we need to be able to characterise the distribution of our X-ray sources in terms of optical (r-band) magnitudes, and account for the number that are missed completely from our sample due to them being fainter than our imposed magnitude limit. Figure 2.21 illustrates this need to accurately characterise the underlying distribution of our sources in f_X-f_{opt} space; the green arrow above the enforced r-band magnitude limit highlights the graphical area in which sources will be 'missed', introducing incompleteness. Hence, if we can reasonably constrain the underlying source distribution across optical and X-ray fluxes, incompleteness introduced by such flux limits can be accounted for.

To constrain the optical/X-ray flux distribution of our sources, we calculate the ratio of their X-ray to optical flux emission; this is known in practise as $\log(f_X/f_{opt})$

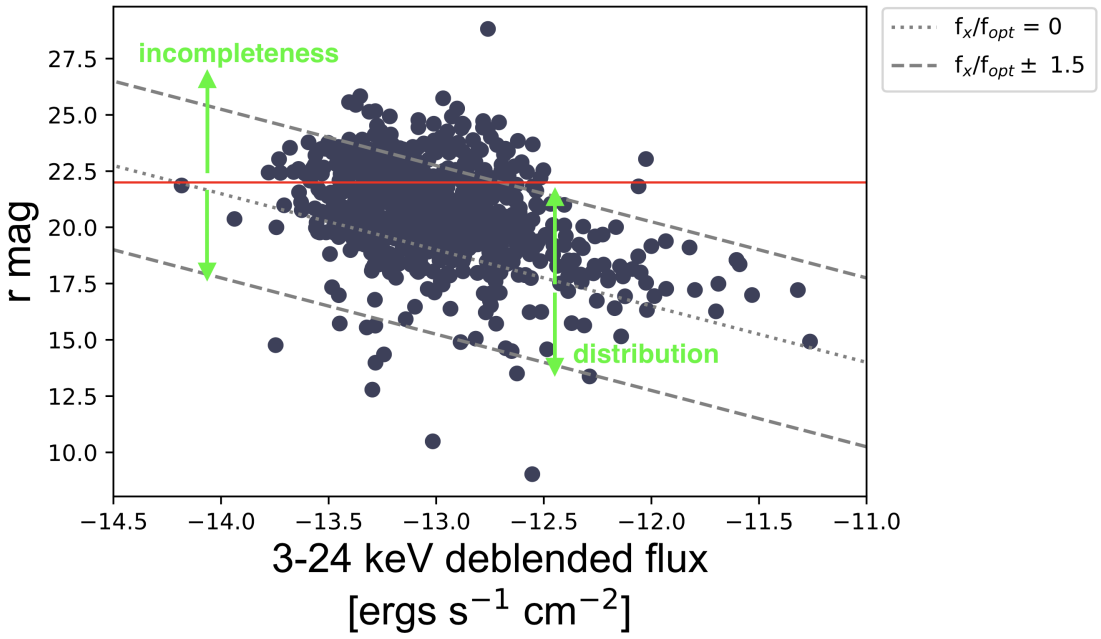

Figure 2.21 *Plot showing the optical and 3 - 24 keV X-ray flux distribution of all NSS80 X-ray sources, and how incompleteness can be introduced when magnitude limits are imposed. The dashed lines indicate lines of constant log(f_X/f_{opt}): dashed lines are log(f_X/f_{opt}) = ± 1.5, the solid line is log(f_X/f_{opt}) = 0. The solid red line indicates an imposed r band magnitude limit. The green arrows illustrate the distribution of sources in optical and X-ray flux space in general. As shown in this plot, any sources which fall above red r magnitude limit line will be missed from the the distribution and introduce incompleteness. Hence, it is important to model said distribution so we can correct for the incompleteness imposed by this r band limit.*

and is calculated using the following equation from Maccacaro et al. (1988):

$$\log\left(\frac{f_x}{f_{opt}}\right) = \log f_x + 5.4 + \frac{r}{2.5} \tag{2.5}$$

where f_X and r are the X-ray flux and r-band magnitude respectively. The bulk of the AGN population is generally seen to have log(f_x/f_{opt}) values ranging between -1.5 and 1.5. We calculated this log(f_X/f_{opt}) for our sample; Figures 2.22 and 2.23 illustrate the resultant optical-X-ray flux distribution for our hard band and full band sample. We fit a Gaussian approximation for both the full and hard band samples (*before* the r mag limit has been introduced), taking the sample mean, μ, and standard deviation, σ as provision for a reasonable description of these distributions; for our hard band sample: $\mu = 0.3475$ and $\sigma = 0.6228$, for our full band sample: $\mu = 0.2691$ and $\sigma = 0.5987$. These are illustrated by the solid blue lines in these Figures.

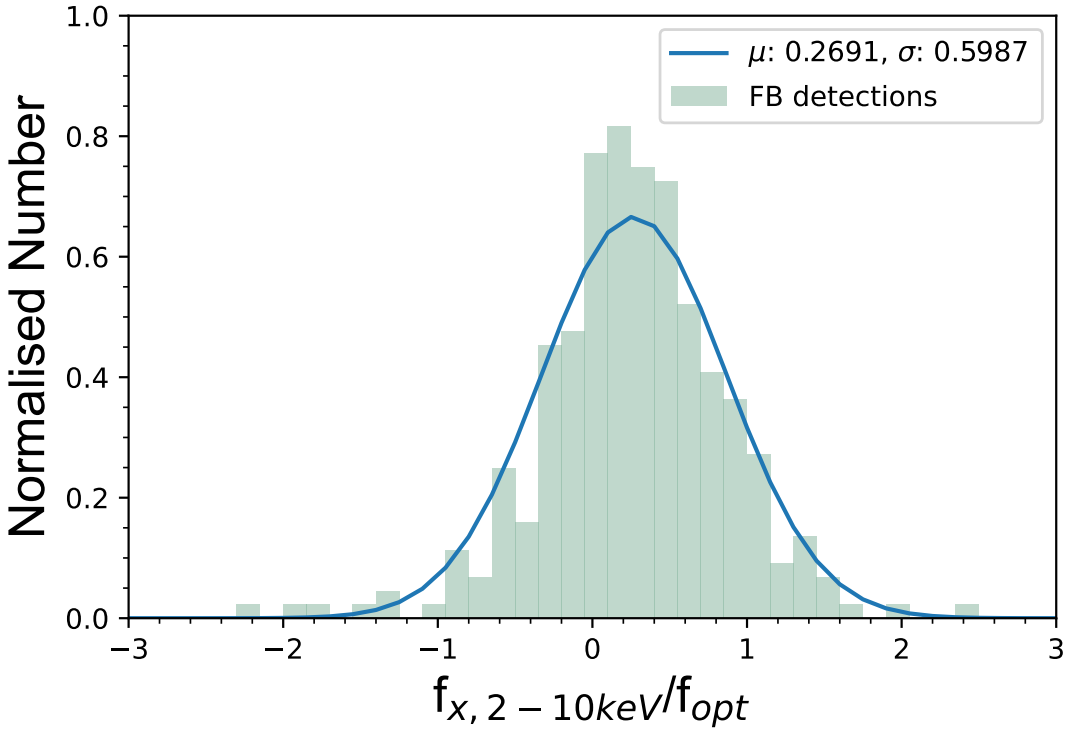

Figure 2.22 *log(f_x/f_{opt}) for our full band sample. The solid blue line illustrates the best fitting Gaussian for this distributions; this corresponds to a Gaussian with μ = 0.2691 and σ = 0.5987 for our full band sample.*

Having now constrained this $\log(f_X/f_{opt})$ for both our hard and full band samples, we can fold this Gaussian distribution describing the distribution of r-band magnitudes at given X-ray fluxes for our sample, into our expected number equation as a term we define as $\beta(r \mid \mu(f_X), \sigma)$. In doing so, we can effectively account for the distribution of the X-ray sources in terms of optical (r band) magnitudes and correct for the number that are missed completely, due to them being fainter than our imposed magnitude limits (i.e. correcting for incompleteness introduced when approaching optical flux limits). Folding this term into our expected number Equation 2.4, we get:

$$\lambda(\theta)_2 = \int_{L_{X,1}}^{L_{X,2}} d\log L_X \int_{z_1}^{z_2} dz \int_{r_1}^{r_2} dr \left[\phi(L_X, z \mid \theta)\ A(f_X(L_X, z))\ \beta(r \mid \mu(f_X), \sigma)\ \frac{dV}{dz} \right] \tag{2.6}$$

The final corrective term we then apply is a multiplicative factor which corrects for the spectroscopic incompleteness of our r mag < 24 sample, as constrained in Figures 2.10 to 2.13. With all of these corrections accounted for, this is the final

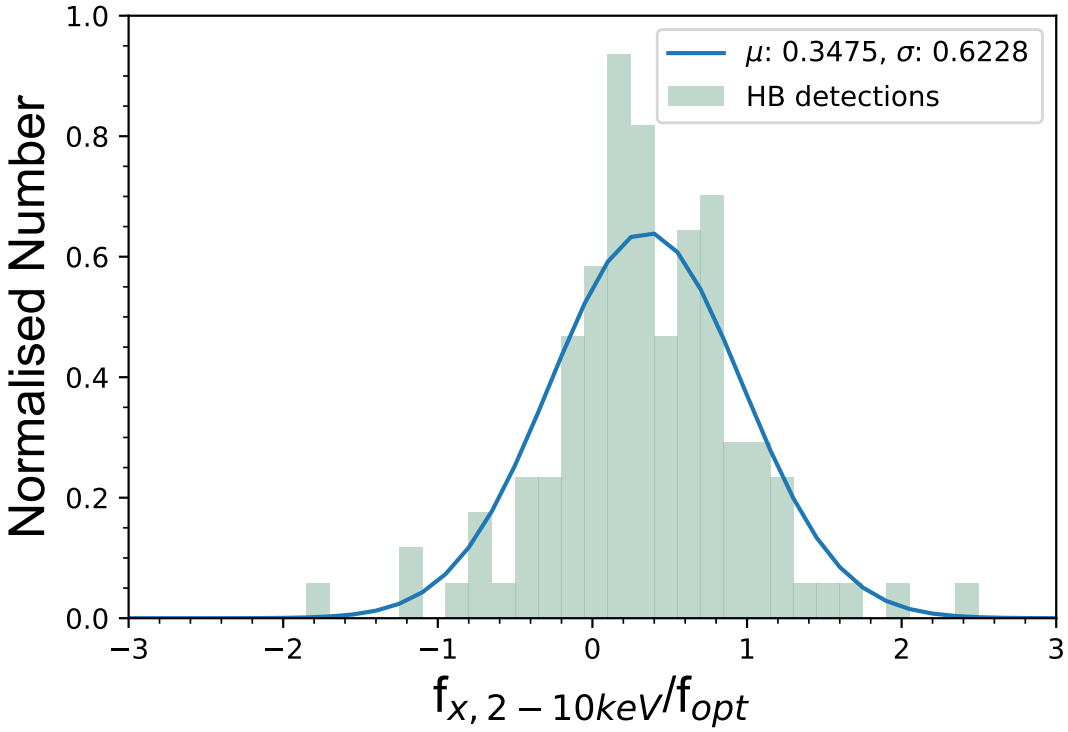

Figure 2.23 *$log(f_x/f_{opt})$ for our hard band sample. The solid blue line illustrates the best fitting Gaussian for this distributions; this corresponds to a Gaussian with μ = 0.3475 and σ = 0.6228 for our hard band sample.*

form of the expected number equation we adopt. Folding this into the original Likelihood function gives us the function which is maximised in an iterative way to find the model parameters which correspond to an XLF that best fits my observed data and accounts for completeness corrections every step of the way.

2.6.2 Maximum Likelihood Fitting (MLF)

In order to perform an initial test of the likelihood function and its behaviour over the XLF parameter space, I created a basic python function which manually performs the MLF process for my BL and NL AGN sources. Given a defined range over which each model parameter is to be maximised, this algorithm sweeps through each parameter in turn (whilst holding the other parameters constant) and calculates the likelihood for every interval in the given range. For example, the model XLF I adopted is defined by 4 parameters: L_*, γ_1, γ_2 and k. The function then takes the first parameter (lets say L_*) and varies it over the full range defined by the user whilst keeping the other parameters constant, and calculates the likelihood function for each variation of L_*. It then selects the

value of L_* that corresponds to the maximum of all the likelihoods, keeps that constant, and then moves on to the next parameter that needs to be evaluated (and so on and so forth). Once done, it returns the 4 best fitting parameters (the combination of which maximises the likelihood function) for my data. This algorithm essentially carries out an initial grid based search to test the likelihood function and its behaviour over the XLF parameter space, and provides us with sensible starting parameters and bounds with which to carry out a more comprehensive Maximum Likelihood Estimation (MLE) process.

Once I had verified that our likelihood function produced sensible values and that it was possible to identify a clear maximum, I moved on from this initial grid based search to using scipy's optimize.minimize function; this has the capability to iteratively minimise functions of multiple variables, within user-defined bounds. It is worth noting that this optimize module *minimises* functions, where we aim to *maximise* the likelihood. Maximising the likelihood is actually equivalent to *minimising the negative log likelihood*; as such, the optimize.minimize function enables us to find the maximum likelihood by computationally minimising the negative log likelihood.

The bounds I chose to use for each parameter in this minimisation process are as follows: $43 \leq \log_{10} L_* \leq 45$, $-10 \leq \log_{10}\gamma_1 \leq 1.5$, $1.5 \leq \log_{10}\gamma_2 \leq 4$, $1e\text{-}10 \leq \log_{10} k \leq 1e\text{-}2$. These bounds were chosen (informed by the aforementioned initial grid search) to encapsulate the typical parameter values found in the literature, whilst also allowing for full exploration of the parameter space. The bounds for γ_1 and γ_2 were also specifically chosen so that they didn't overlap; this was done upon realising, when doing some initial runs with overlapping γ_1 and γ_2 bounds, the minimisation function often got stuck on trying to find solutions that continually flipped between the slopes of γ_1 and γ_2 values - this behaviour is expected given the degeneracy of these two parameters.

Whilst using optimize.minimize alone would already result in valid XLF evaluations, we can take it one step further and use Markov Chain Monte Carlo (MCMC) methods. MCMC allows us to generate random draws from the likelihood to calculate posterior distributions and provide parameter uncertainties; by using MCMC methods we can sample across the full range of out model parameters' distributions and find an estimate of the posterior probability function (the distribution of parameters that is consistent with my dataset) for others. I can also extract parameter uncertainties based on the 16th and 84th percentiles of the samples in the marginalized distributions. We used the MCMC python package

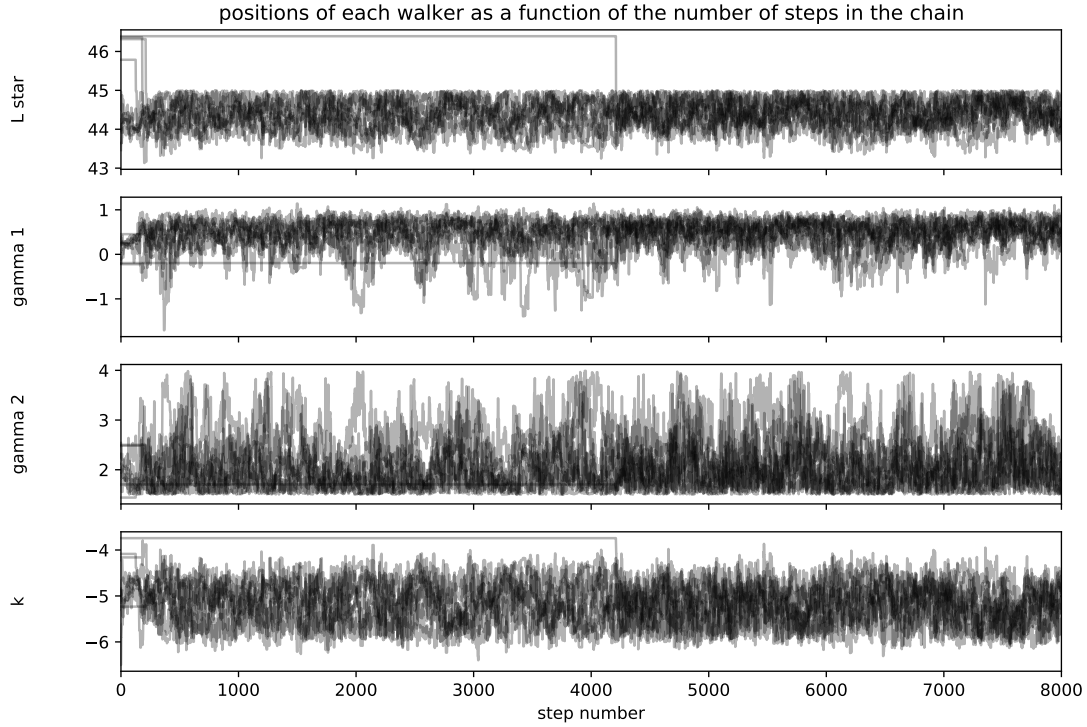

Figure 2.24 *Successive emcee samples spanning the full parameter space for each of my parameters (k, γ_1, γ_2 and L_*, across thousands of iterative steps in each chain.*

'emcee' (Foreman-Mackey et al., 2013) to carry out this analysis. The stopping criterion 'emcee' follows is outlined in Goodman & Weare (2010) and uses the integrated autocorrelation time to quantify the effects of sampling error on results and ensure that all chains are sufficiently converged; the autocorrelation time directly quantifies the Monte Carlo error (thus, sampler efficiency) on integrals computed using MCMC results and, as such, is the logical parameter to use when evaluating the robustness of such an MCMC analysis. (Foreman-Mackey et al., 2013; Goodman & Weare, 2010). I ran 50 - 100 chains for each parameter and, as recommended in the emcee documentation (Foreman-Mackey et al., 2013), assume chains are sufficiently converged when the integrated autocorrelation time is $\geq$ the number of samples/50. I also discard the first 25% of the steps for each parameter as a burn-in phase. Figure 2.24 shows how the MCMC process I adopt samples the full parameter space across thousands of iterations.

I ran this process for both my full and hard BL and NL AGN samples, producing the one and two dimensional projections of the posterior probability distributions of my parameters. As such, I constrained the 'best fitting' XLF parameters and parameter uncertainties for my BL and NL samples. Figure 2.25 demonstrates

all of the covariances between parameters for our hard band NL AGN sample. This figure show the marginalized distribution for each parameter independently in the histograms along the diagonal and then the marginalized two dimensional distributions in the other panels. The final results from this MLF analysis are presented and discussed in Section 2.7.

2.7 XLF measurements

Figures 2.26 and 2.27 below show my final, independent BL and NL 2 - 10 keV AGN XLFs for our hard and full band samples at z $\sim$ 0.2-1.1, with 1σ equivalent upper and lower confidence intervals based on MCMC maximum likelihood fitting. The sum of the BL and NL AGN XLF is overplotted as the total AGN XLF, in dashed black. All AGN XLFs generated in my thesis are converted to restframe 2 - 10 keV X-ray fluxes assuming a photon index of $\Gamma = 1.7$, to compare my results to the literature. The parameter uncertainties used on the plots are taken on the 16^{th} and 84^{th} percentiles of the samples in the marginalized distributions. Binned N_{obs}/N_{mod} XLF estimates are over-plotted in green, and comparison AGN XLFs from Aird et al. (2015) are over-plotted in red and yellow.

Looking at our full band AGN XLFs (Figure 2.26, we can see that there is not shown to be much difference between the NL and the BL populations. There appear to marginally be more NL AGN at lower luminosities, with this then decreasing (comparatively to the BL AGN number density) at intermediate luminosities, and then increasing once more at higher luminosities. Our results predict lower AGN number densities overall (mainly concentrated at lower luminosities) than the Aird (2015) XLFs, in both redshift ranges.

Looking at our hard band AGN XLFs (Figure 2.27, we can see a very similar picture: our results do not show much difference between the NL and the BL populations. Once again, there appear to marginally be more NL AGN at lower luminosities, with this then decreasing (comparatively to the BL AGN number density) at intermediate luminosities, and then increasing once more at higher luminosities. Our hard band XLFs are more in line with the z = 0.1 - 0.5 Aird et al. (2015) XLF than our full band XLFs, but still predict lower AGN number densities overall (mainly concentrated at lower luminosities) than the z = 0.5 - 1 Aird et al. (2015) XLF. For comparison, the median redshift of our sample is z $\sim$ 0.5.

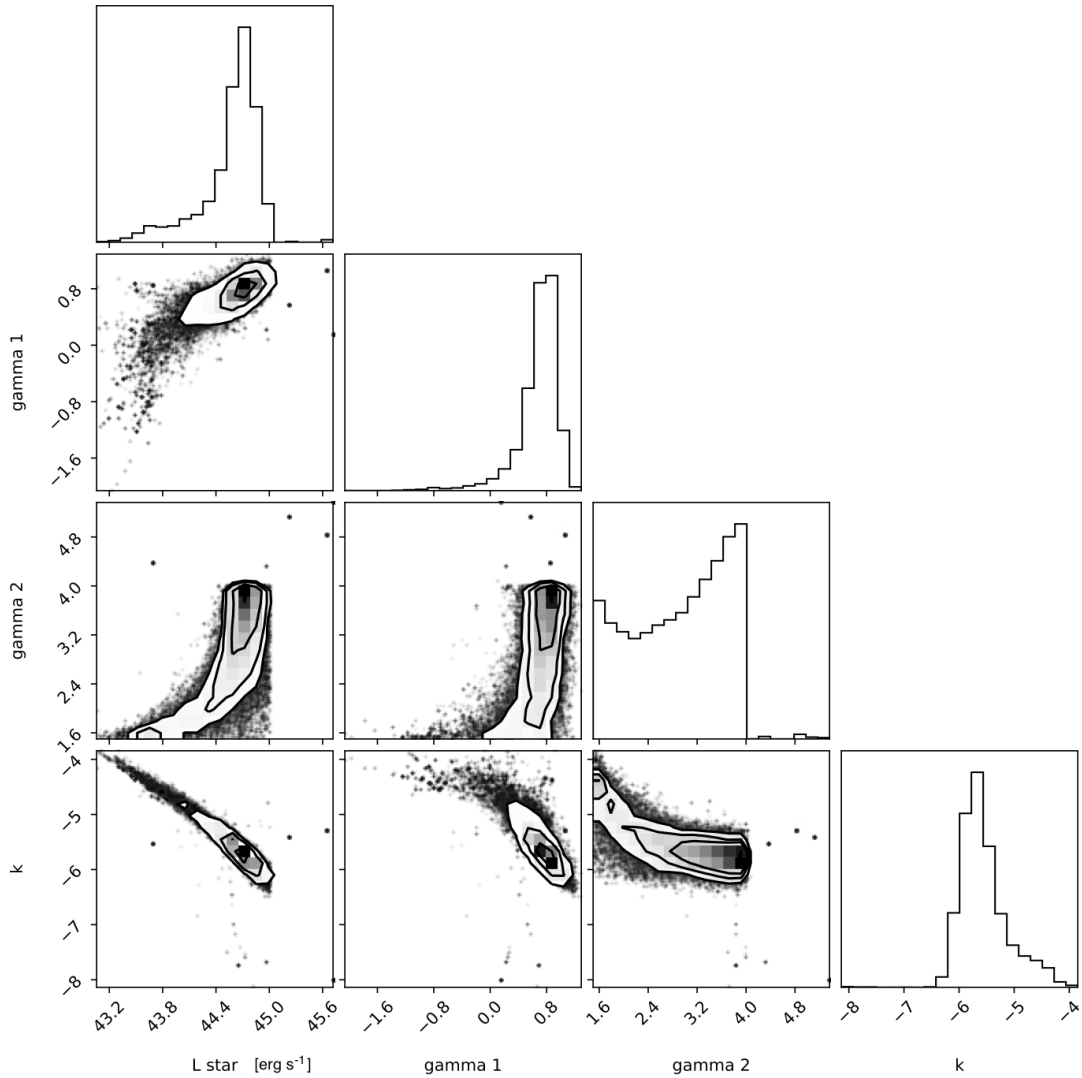

Figure 2.25 *Corner plot demonstrating the covariance between parameters for the NLAGN MCMC MLE process. Histograms along the diagonal show the marginalised distribution for each parameter independently, and the other panels show the marginalised 2-D distributions between parameters.*

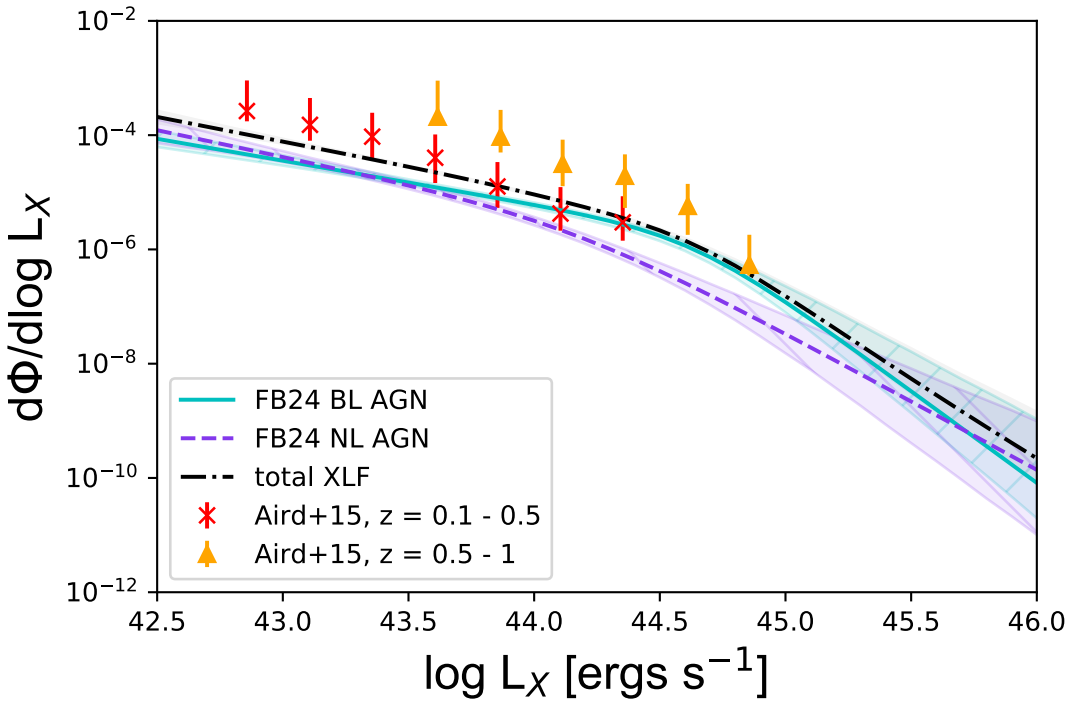

Figure 2.26 *BL (cyan solid), NL (purple dashed) and Total (black dot dashed) AGN 2 - 10 keV XLFs for our full band, r < 24 sample. Binned 2 - 10 keV XLF estimates for two comparative z ranges taken from Aird et al. (2015) are indicated by red exes and yellow triangles.*

2.7.1 The obscured AGN fraction at z ∼ 0.2 - 1.1

Another interesting way to look at these results is by looking at the fraction of obscured AGN as a function of luminosity; Figures 2.28 and 2.29 illustrate the obscured AGN fraction as a function of luminosity calculated in this work (blue points) in comparison to Hasinger (2008) (red crosses), a previous study which looked at Type 2 AGN fractions as a function of redshift and luminosity.

Our analysis finds that, in both bands, the fraction of obscured AGN at z ∼ 0.2 - 1.1 does not greatly depend on luminosity, within the errors of this study. This seems to refute the obscured fraction result of Hasinger (2008) (overplotted in red), which shows a comparatively steep and obvious dependence of obcured AGN fraction with luminosity. However, Hasinger (2008) uses a 2 - 10 keV selected sample (i.e. softer energies than my 3 - 24 and 8 - 24 keV samples), as well as using a 'mixed' obscuration classification approach (as opposed to my use of solely optical spectroscopic classifications). Furthermore, and perhaps most vitally, Hasinger (2008) combines data over a very wide z - range which, in combination with the aforementioned differences, could be introducing an

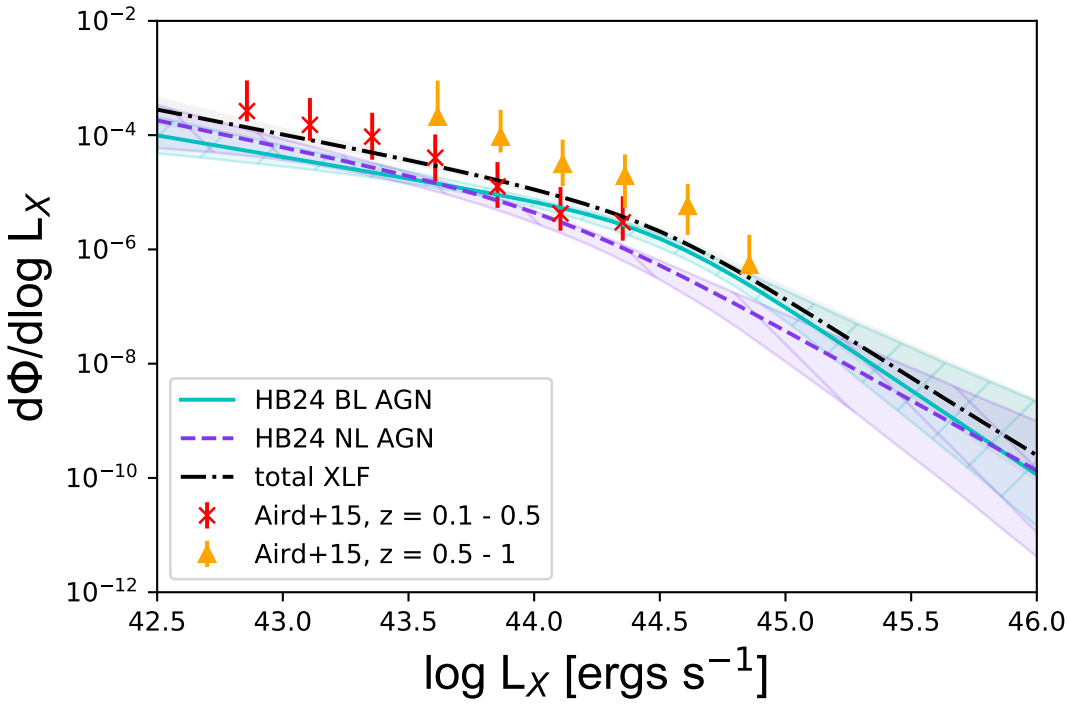

Figure 2.27 *BL (cyan solid), NL (purple dashed) and Total (black dot dashed) 2 - 10 keV AGN XLFs for our hard band, $r < 24$ sample. Binned 2 - 10 keV XLF estimates for two comparative z ranges taken from Aird et al. (2015) are indicated by red exes and yellow triangles.*

apparent L_X dependence.

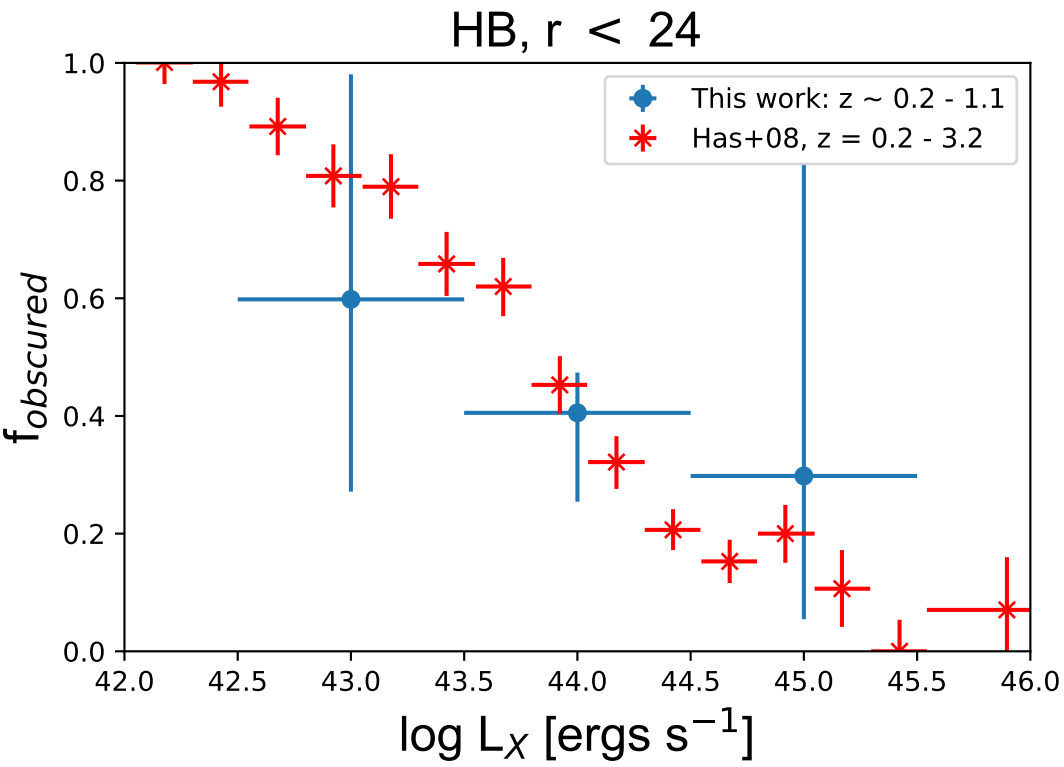

Figure 2.28 *Obscured AGN fraction as a function of luminosity for our HB, r < 24 sample (blue points) and that described in Hasinger (2008) (red crosses).*

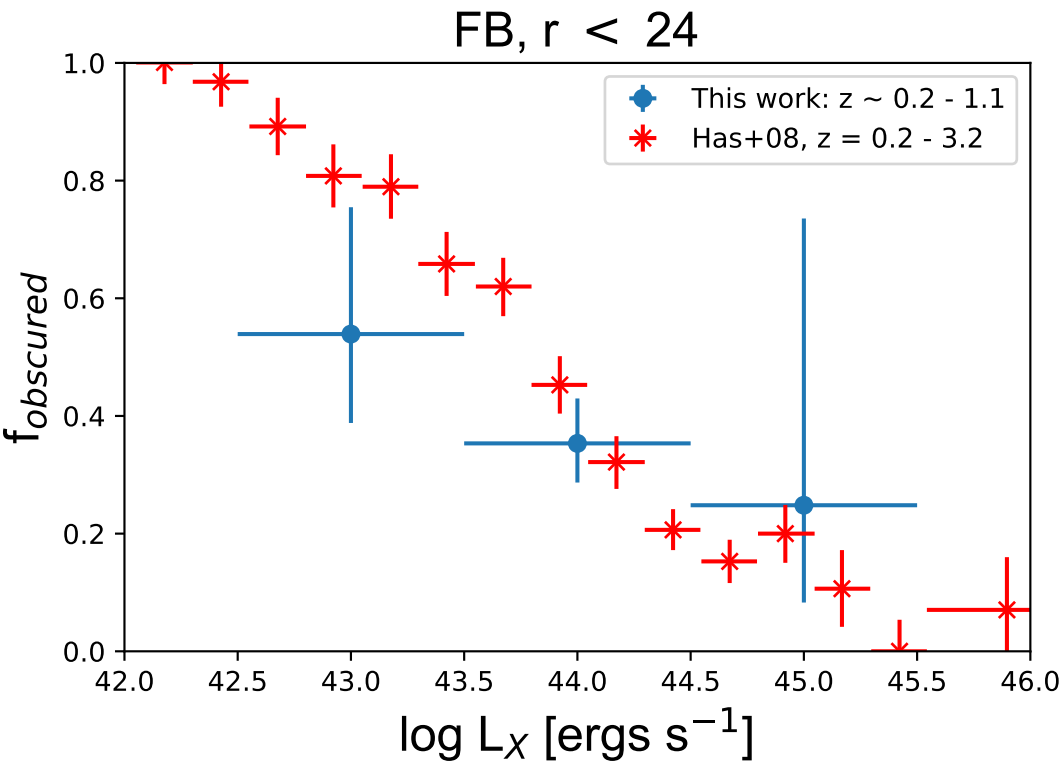

Figure 2.29 *Obscured AGN fraction as a function of luminosity for our FB, r < 24 sample (blue points) and that described in Hasinger (2008) (red crosses).*

Chapter 3

Classifications, z - estimations and completeness corrections for an X-ray and optically selected z $\sim$ 2 - 4 AGN sample.

In chapter 2 I presented new measurements of the obscured and unobscured AGN XLF at low redshifts. Now, in the next two chapters, I present a complementary higher-z analysis to determine the XLF for spectroscopically classified BL and NL AGN.

In this chapter I describe our X-ray selected AGN sample, which has been optically pre-selected to be at z $\sim$ 2 - 4, the optical spectroscopy of said sample, and a careful determination of the completeness of each stage of the selection process. Using this combination of X-ray and optical/UV data, I subsequently derive the obscured and unobscured z $\sim$ 2 - 4 AGN XLFs in chapter 4. The X-ray and optical pre-selection processes used to identify a suitable AGN sample are detailed in sections 3.1 and 3.2. The process of obtaining our optical/UV spectroscopy is detailed in section 3.3. The development and implementation of my automated spectral fitting and source classification code (used to fit, classify and estimate source redshifts) is detailed in section 3.4. The development of my spectral simulations, which are used to determine completeness corrections for the optical/UV spectroscopy, is detailed in section 3.5. The classifications, redshifts and completeness corrections derived in this chapter, as a direct result

of my spectral fitting and spectral simulations, will be folded into my AGN XLF measurements in chapter 4 to correct for completeness at every stage.

3.1 X-ray Pre-selection

When identifying both obscured and unobscured (absorbed and unabsorbed) AGN across a range of redshifts and luminosities, X-ray selection provides an effective alternative to the use of data from optical QSO surveys, which will mostly select unobscured and bright AGN (as discussed in chapter 1). The sources presented in this work were X-ray selected using data from *Chandra* and *XMM-Newton*, across the Canada France Hawaii Telescope Legacy Survey (CFHTLS) fields D1, D3 (otherwise known as the AEGIS field) and D4. D1 has $\sim$ 10–25 ks depth *XMM* coverage over $\sim$0.8 deg^2 of the CFHTLS field, AEGIS has 800ks depth *Chandra* coverage over 0.25 deg^2, and D4 has $\sim$ 260 ks deep *XMM* coverage from a central pointing with additional flanking pointings of depth $\sim$ 30ks that provide an overall coverage of $\sim$ 0.8 deg^2. For D3 (AEGIS) we use the source catalogue from Nandra et al. (2015) , selecting all sources detected in the full (0.5–7 keV), soft (0.5–2 keV) or hard (2–7 keV) energy bands. In D1 we select sources based on the X-ray point source catalogues from Pierre et al. (2007). In D4, we relied on the *XMM-Newton* pipeline processing to provide point source catalogues in each of the individual pointings, which we merge to produce a master source list. In D1 and D4 (our wider, shallower fields) we limit our sample to soft band selected sources with $f_{0.5-2\mathrm{keV}} > 5 \times 10^{-15}$ erg s^{-1} cm^{-2}, which reduces our number of targets and ensures these fields are restricted to high-luminosity AGN with $L_{\mathrm{X}} \gtrsim 10^{44}$ erg s^{-1} i.e. probe the bright-end of the XLF, above L_*.

To carry out measurements of the XLF requires a careful quantification of the X-ray completeness of the *Chandra* and *XMM-Newton* imaging, which is provided by the area curve (the area of sky covered as a function of the X-ray flux limit). Area curves in the AEGIS field for both the soft and hard observed energy bands are provided by Nandra et al. (2015) and calculated using the methodology described in Georgakakis et al. (2008). .

In the case of the D1 and D4 fields, a more involved approach was necessary in order to produce area curves which accurately describe our sensitivity/coverage across these two fields. First, we looked at the XMM-LSS pointings which lie in

the D1 field; Figure 3.1 shows the layout of these XMM pointings covering the D1 field and Figure 3.2 is a table indicating IDs, co-ordinates and exposure times for each pointing. As can be seen in Figure 3.2, only 11/12 of these XMM-LSS pointings have usable data. Each individual pointing nicely covers an area within radius $\sim$ 13 arcmin, with differing exposures as indicated. Thus, we need to determine the detection probability in each pointing, given its exposure time, out to a 13 arcmin radius.

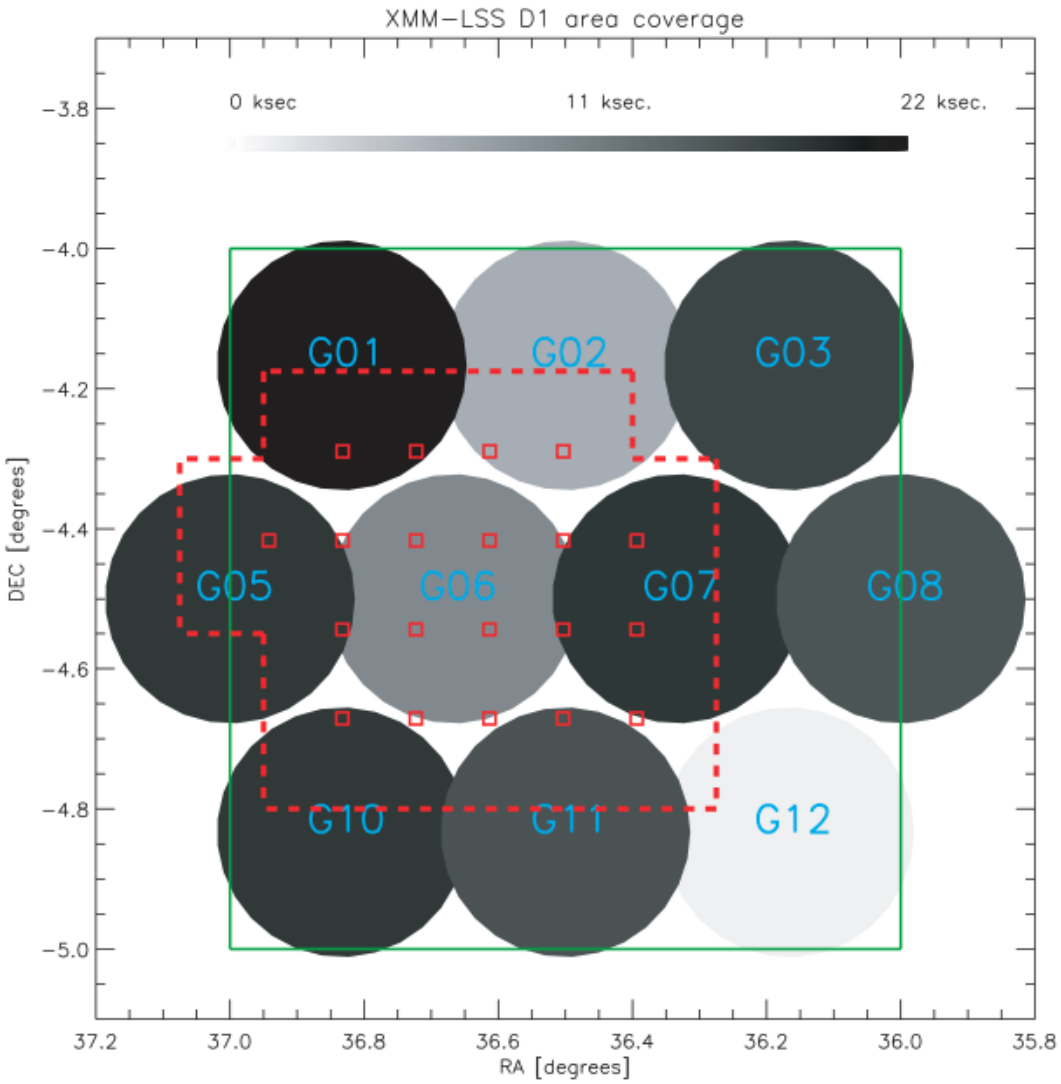

Figure 3.1 *Taken from Pierre et al. (2007); the XMM pointing mosaic over the D1 area (green square). The radius of the displayed pointings is 11'. The grey-scale indicates effective mean exposure time per detector, after removal of high background periods.*

Further data on these 11 XMM-LSS pointings were then extracted from Elyiv et al. (2012); this paper presents source detection probabilities for the 11/12

Internal ID	*XMM* ID	RA (J2000)	Dec. (J2000)	MOS1, MOS2, pn exposure times (ks)
G01	0112680201	02:27:20.0	−04:10:00.0	24.6, 25.3, 21.4
G02	0112680201	02:26:00.0	−04:10:00.0	10.1, 9.7, 6.7
G03	0112680301	02:24:40.0	−04:10:00.0	21.8, 21.7, 17.3
G05	0112680401	02:28:00.0	−04:30:00.0	23.5, 23.9, 12.5
G06	0112681301	02:26:40.0	−04:30:00.0	16.4, 16.6, 10.5
G07	0112681001	02:25:20.0	−04:30:00.0	22.5, 25.1, 18.6
G08	0112680501	02:24:00.0	−04:30:00.0	21.2, 21.3, 15.9
G10	0109520201	02:27:20.0	−04:50:00.0	24.7, 24.6, 18.5
G11	0109520301	02:26:00.0	−04:50:00.0	21.7, 21.8, 16.1
G12	0109520401	02:24:40.0	−04:50:00.0	Not usable because of very high flare rate

Figure 3.2 *Taken from Pierre et al. (2007); properties of individual XMM pointings within the D1 field. Quoted exposures are effective exposures computed after filtering high background periods.*

XMM-LSS pointings in the D1 field, as a function of various XMM pointing characteristics (exposure time, radius). This data is shown in Figures 3.3 and 3.4. In brief, the analysis in Elyiv et al. (2012) involves performing extensive simulations of each XMM pointing in the XMM-LSS field (within which the D1 field lies) and applying a source detection algorithm to these simulations. This is the same source detection algorithm used in the catalogue from which our sources were chosen for the D1 field (Pierre et al., 2007), thus matching our source catalogue in that field. The efficiency of said algorithm as a function of pointing depth and radius is then returned as a detection probability, as shown in Figures 3.3 and 3.4. Figure 3.3 shows, from simulations, the detection probability within a 5 arcmin radius for different exposure times and Figure 3.4 shows the detection probability within annuli spanning different radii for a single (10 ks) exposure time.

Having extracted these detection probabilities, the next step was to use them as the basis for our area curve calculations for the specific XMM-LSS pointings that overlap with our D1 optical imaging. We achieved this by translating the Elyiv et al. (2012) detection probabilities within an annulus, to the overall probability within a radius of the aim-point. Our process thus involves: i) Calculating the overall probability within a 13 arcmin circular area for a 10 ks exposure time, which can be described via the following equation:

$$\mathrm{p_{det}}(< \mathrm{R}) = \frac{\left(\sum_{\mathrm{i}}(\pi \mathrm{r}^2_{\mathrm{outer,\ i}} - \pi \mathrm{r}^2_{\mathrm{inner,\ i}})\right) \mathrm{p_{det,\ i}}}{\pi \mathrm{R}^2} \tag{3.1}$$

where $\mathrm{p_{det}}(< \mathrm{R})$ is the total probability of detection within some radius R, $\mathrm{r_{outer,\ i}}$ and $\mathrm{r_{inner,\ i}}$ are the outer and inner radii of some annulus i, $p_{det,i}$ is the probability

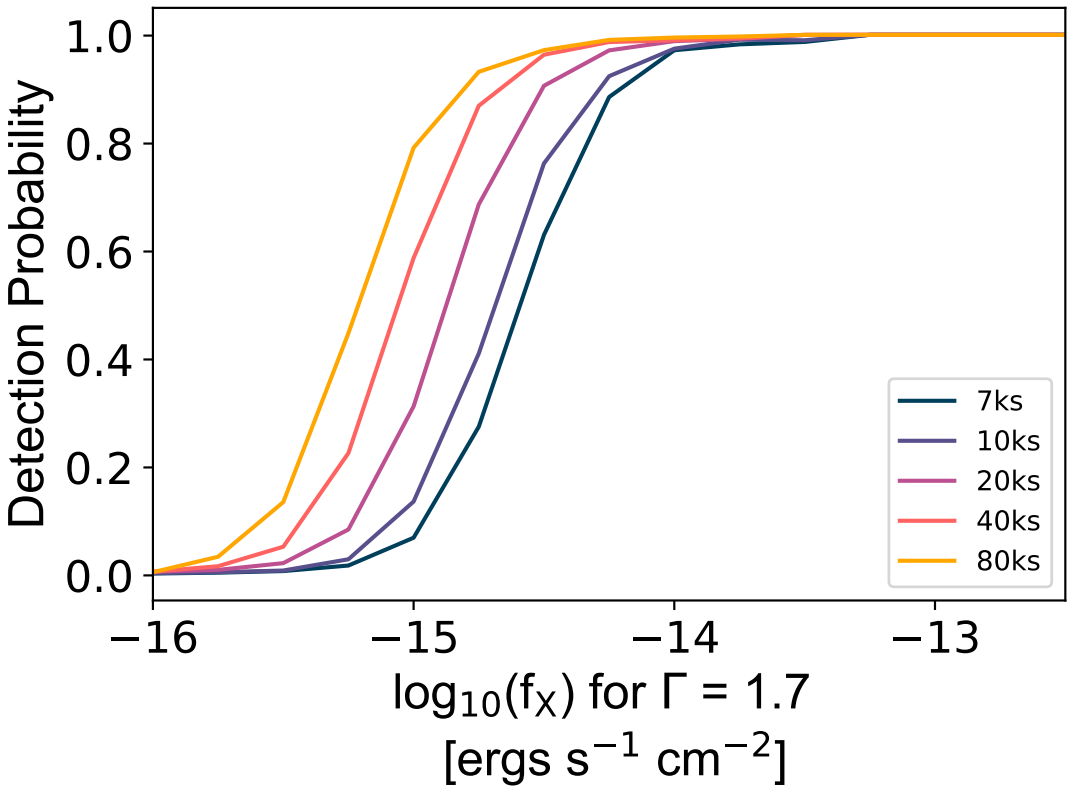

Figure 3.3 *Adapted from Elyiv et al. (2012); the efficiency of their source detection algorithm in the innermost annulus (0–5') of their simulated soft band XMM observations for differing exposure times, as a function of X-ray flux*

of detection with annulus i, and $\sum_i$ is summing over each annulus i, for $r_{outer,\ i} <$ R.

First, we used the information in Figure 3.4 to create 10ks detection probability functions for circular pointings of varying radii (as seen in Figure 3.5), corresponding to the 10ks detection probability functions for differing off-axis distances (i.e. for varying annuli) shown in Figure 3.4. We then isolated the specific XMM-LSS pointings that were within the D1 area, extracting their exposure times (see figure 3.2 for the properties of the individual XMM-LSS pointings which lie in the D1 field) and calculated the total area of the D1 field they covered: 0.89 deg^2. Next, using the exposure times for each XMM-LSS D1 pointing, we interpolated between the detection probability functions in Figure 3.3 and determined the detection probability as a function of flux for each specific XMM-LSS D1 pointing exposure time, as seen in Figure 3.6. Similarly to those in Figure 3.3, the detection probabilities for each XMM-LSS D1 pointing exposure time in Figure 3.6 are for pointings of 5' radii. However, each XMM-LSS pointing is 13'. To account for this, we calculated the scaling factors between the newly determined

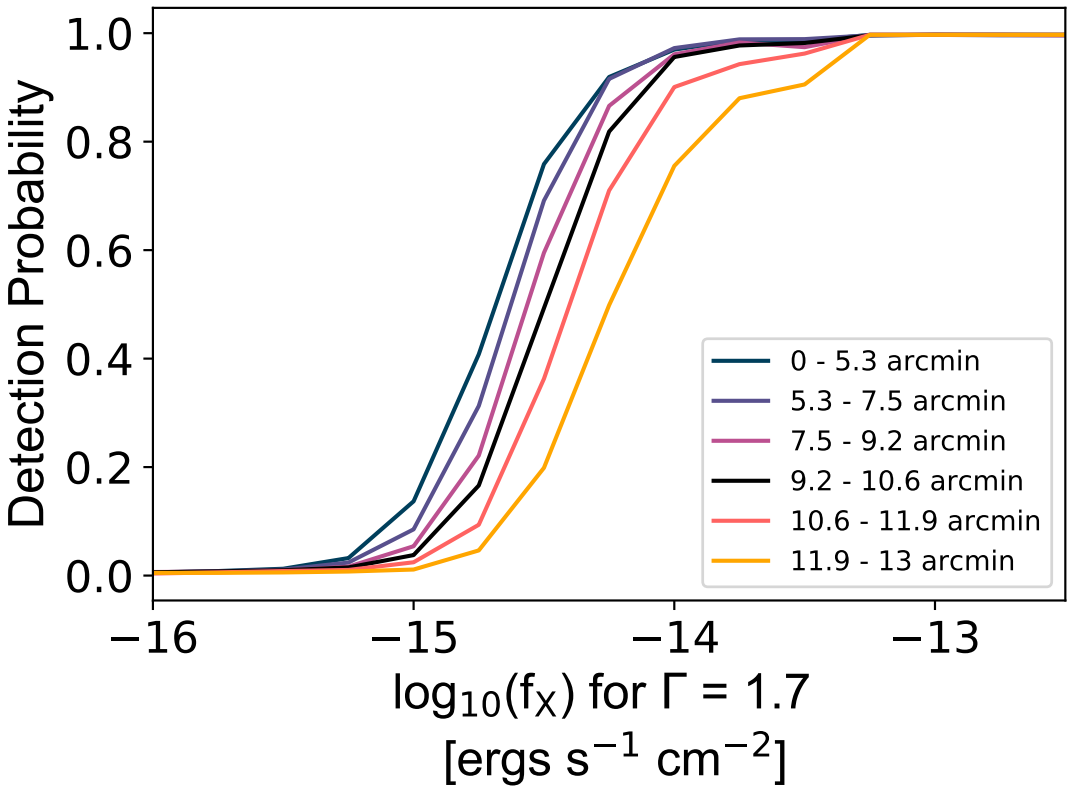

Figure 3.4 *Adapted from Elyiv et al. (2012), plot showing the efficiency of their source detection algorithm for differing off-axis distances of their simulated soft band XMM observations, as a function of flux . Exposure time is held at 10 ks in all cases.*

5' XMM-LSS D1 detection probabilities, and the 10ks 5' radius pointing detection probability in Figure 3.5; by applying these pointing-specific scaling factors to the 13' radius detection probability function in Figure 3.5, we effectively determined the detection probability for each individual 13' radius XMM-LSS pointing in the D1 field (i.e. area curves for each pointing as a function of exposure time and pointing radius), as shown in Figure 3.7. The final D1 area curve is then a combination of these pointings' area curves. In comparison to the D1 area curve seen in Pierre et al. (2007) our updated D1 area curve shows a slightly decreased sensitivity in the flux range 10^{-15} – 10^{-14} erg s^{-1} cm^{-2}. We employ the same method to calculate the D4 area curve. The final individual and combined X-Ray area curves for the D1, D4 and AEGIS fields can be seen in Figure 3.8, with the solid vertical red line indicating the soft band flux cut of $f_{0.5-2\mathrm{keV}} > 5 \times 10^{-15}$ erg s^{-1} cm^{-2}.

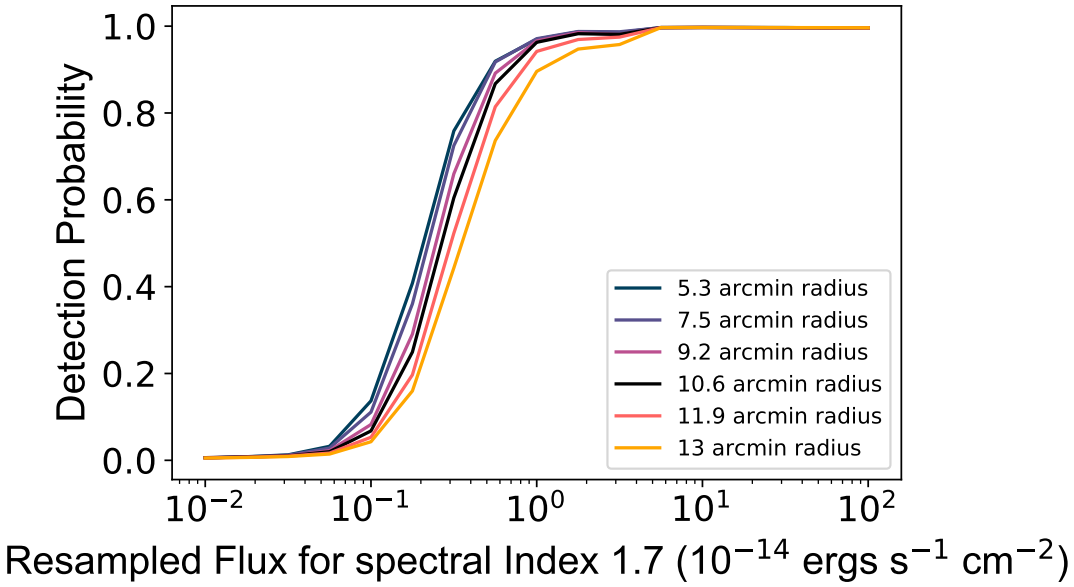

Figure 3.5 *The efficiency of the Elyiv et al. (2012) source detection algorithm for differing pointing radii of their simulated soft band XMM observations, as a function of flux. Exposure time is held at 10 ks in all cases.*

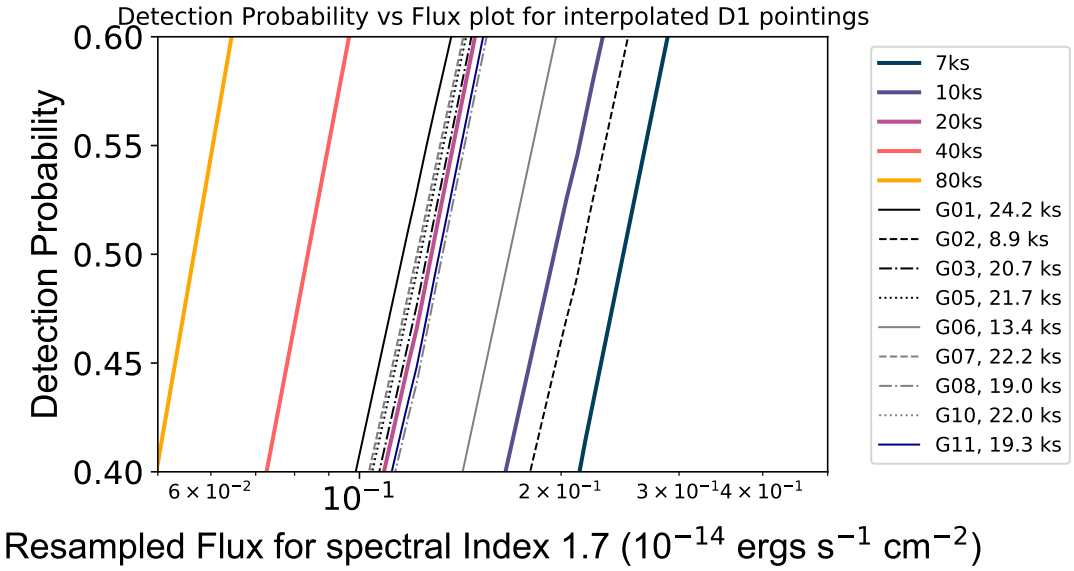

Figure 3.6 *A zoomed in section of the source detection probability functions for varying exposure times; The thicker, solid coloured lines are same functions shown in Figure 3.3 (depicting the exposure times defined by Elyiv et al. (2012)) and the broken lines are the detection probability functions for each XMM-LSS D1 pointing. Similarly to in Figure 3.3, these detection probabilities are for pointings of 5' radii.*

3.2 Optical Pre-selection

Spectroscopy of distant (high z) X-Ray selected AGN is especially challenging due to their optical faintness. Thus, spectroscopic follow-up of all X-ray sources

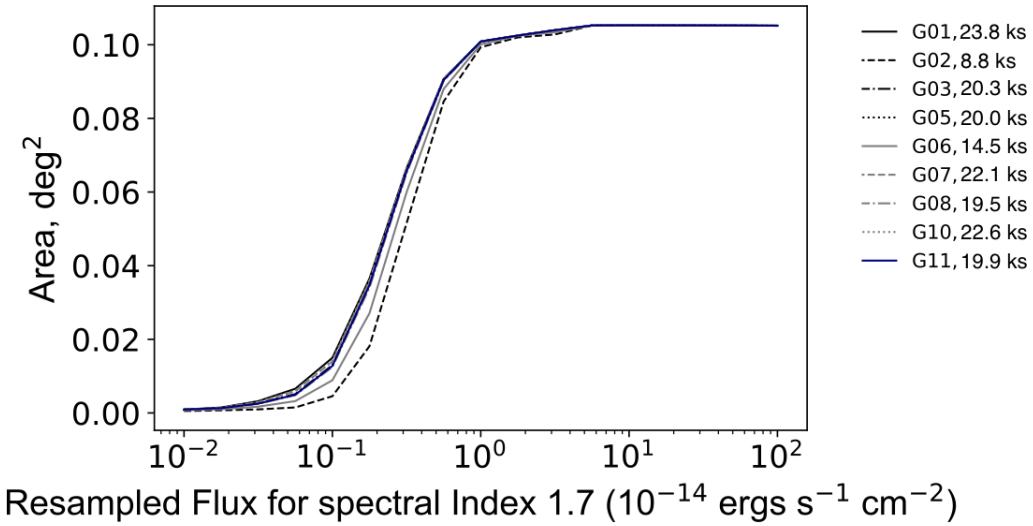

Figure 3.7 *The individual XMM-LSS D1 pointing area curves; these sum to create the total area curve for the D1 field.*

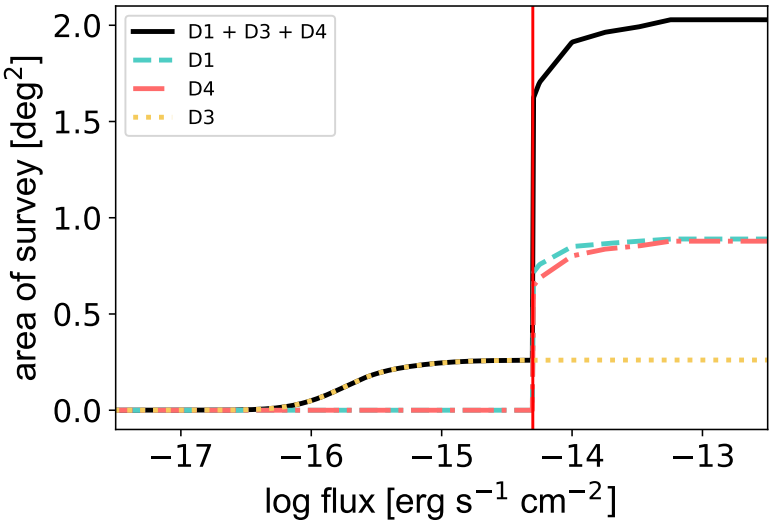

Figure 3.8 *X-ray area coverage as a function of flux for each of our fields and our overall survey. The green dashed lines show the area curve for the ∼10–25ks* XMM-Newton *coverage of the CFHTLS D1 field in the soft (0.5–2 keV) energy band. The red dot-dashed shows the soft band area curve for the ∼ 260ks and ∼ 30ks* XMM-Newton *coverage (for central and flanking pointings respectively) across CFHTLS D4. The dotted yellow shows the soft band area curve for the ∼800ks* Chandra *coverage across CFHTLS D3 field (AEGIS). The black line shows the total area curve for our overall survey, covering a total of ∼2 deg^2. The solid red vertical line shows the soft band flux cut of $f_{0.5-2\mathrm{keV}} > 5 \times 10^{-15}$ erg s^{-1} cm^{-2} applied to our source list in the D1 and D4 fields.*

in major survey fields is still not 100%. Furthermore, most samples used in spectroscopy are defined to a particular optical flux limit. This, coupled with the fact that the X-ray/optical flux fraction varies across all known AGN, means

that there is potential for X-ray sources which lie below the optical flux limit to be missed; this presents another form of incompleteness. When these problems are considered, along with the infrequency of high z objects appearing in most surveys, it becomes clear that accurate pre-selection of sources is necessary to obtain the desired sample of X-ray selected AGN at z $\sim$ 2 - 4.

To address this issue, we have used an optical pre-selection approach to identify candidate z $\sim$ 2 - 4 AGN for spectroscopic follow-up. We colour-selected AGN over an area of ~ 2 deg^2, from the Canada France Hawaii Telescope Legacy Survey (CFHTLS) fields with existing deep u* g' r' i' optical data.

Selecting for Lyman-Break Galaxies results in a selection of AGN at redshifts of ~ 3[1]. As this falls within our desired z range, we implemented LBG selection through using the following optical selection criteria, outlined in Steidel et al. (2003):

$$\begin{aligned} g' - r' &\leq 1.2 \\ u^* - g' &\geq g' - g' + 1.0 \\ 19.0 \leq g' &\leq 25.5 \end{aligned} \tag{1}$$

We used an additional selection criteria (BX) to select for marginally lower redshifts ~ 2, taken from Adelberger et al. (2004). This selects for objects that fall slightly below the colour space cordoned by the Lyman-break criteria:

$$\begin{aligned} g' - r' &\geq -0.2 \\ u^* - g' &\geq g' - r' + 0.2 \\ g' - r' &\leq 0.2(u^* - g') + 0.4 \\ u^* - g' &< g' - r' + 1.0 \end{aligned} \tag{2}$$

Using both of these criteria allows us to effectively isolate objects which are bright in the g band and faint in the u band, adequately selecting sources within the z

[1]Lyman-Break Galaxies, LBGs: radiation emitted from these galaxies at energies higher than those equivalent to 912 Å, the Lyman limit, is completely absorbed by molecular hydrogen between the source and observer. Emission is only seen in LBG spectra at wavelengths higher than the Lyman limit in the rest frame, forming a distinguishable 'break'.

$\sim$ 2 - 4 range.

Despite being versatile, these criteria alone would introduce incompleteness into our optical selection technique, resulting in some z $\sim$ 2 - 4 candidates being missed. We account for this by using optical selection functions. These optical selection functions have been derived to characterise the incompleteness in this colour selection process; by incorporating them into our XLF calculations, we can effectively correct for the incompleteness that our colour selection criteria introduce. The optical selection functions used are those outlined in Aird et al. 2010 . In brief, model u* g' r' colour distributions were generated from model template 'classification' spectra. These model colour distributions were used to generate fake sources and add them to the optical images. The same source detection and photometry procedures used on their real data were then used on these fake sources, in an attempt to recover them. The selection functions are then defined by the probability of recovering a source through this colour selection, with a given R magnitude as a function of redshift . Figure 3.10, showing the optical selection functions used (taken from Aird et al. 2010) illustrates how effective this optical selection technique is in picking out targets which lie in the desired redshift range $2 < z < 4$.

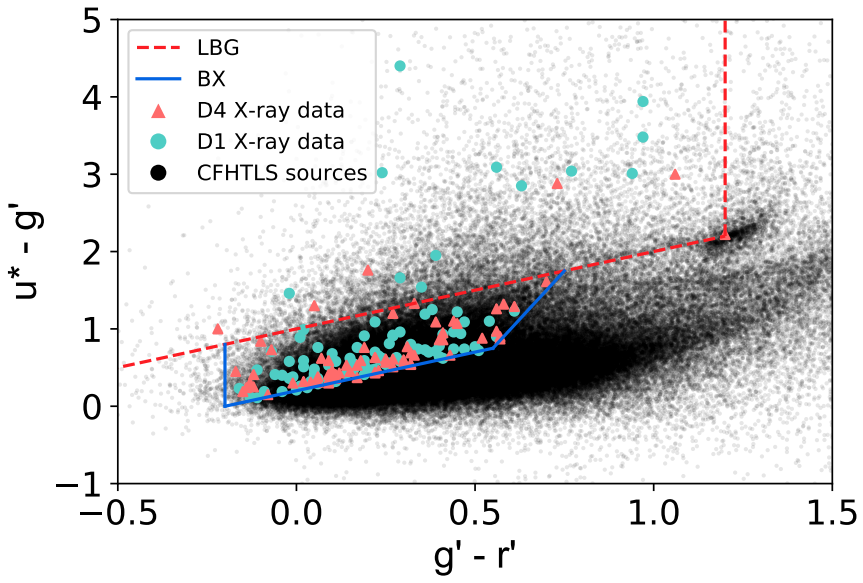

Figure 3.9 *u* g' r' colour - colour space graph with optical selection criteria overlaid, showing the candidates these particular u* g' r' criteria selected. Optical sources without matching X-ray data are marked as black points, and our sources with matching X-ray data are marked by blue circles and pink triangles for candidates in the D1 and D4 fields respectively. Dashed red lines outline the area enclosed by the LBG selection criteria (equation* **??***) and solid blue lines outline the area enclosed by the BX selection criteria (equation* **??***).*

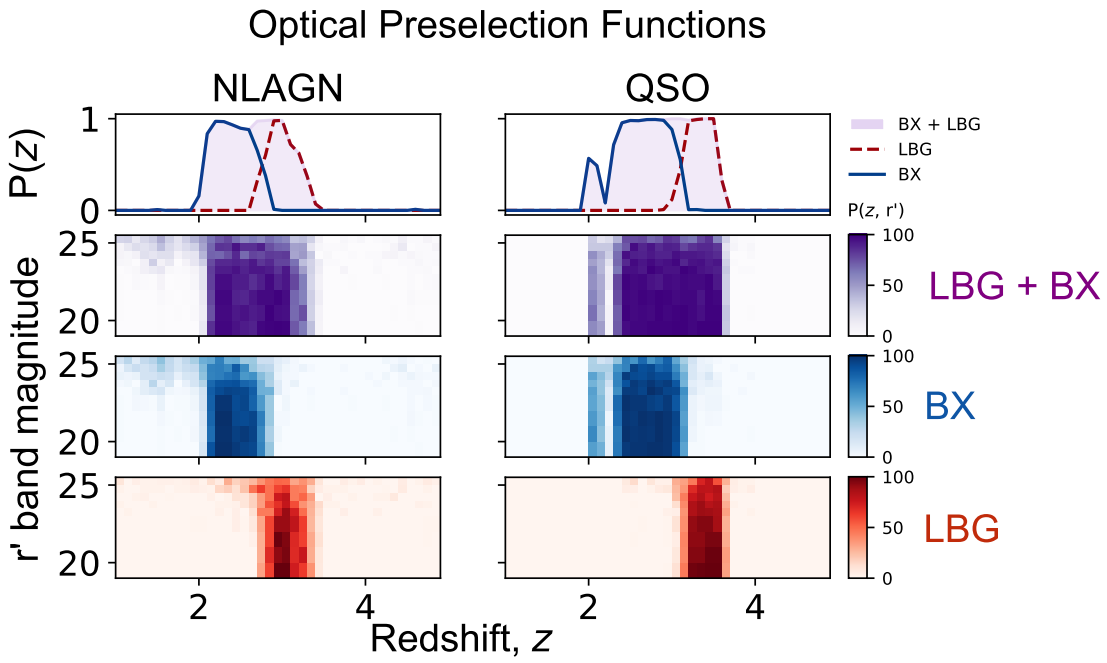

Figure 3.10 *Optical selection functions for the colour selection criteria for the CFHTLS-D3 optical data and u* g' r' i' filters, as described in Aird et al. 2010 .*

This method of colour selection is not the only way to isolate candidates at certain redshifts. As mentioned in Chapter 1, another route would be to try and get spectroscopic redshifts for all potential candidates (i.e. all X-ray sources) and filter through them afterwards. However, this would be an incredibly inefficient use of time and resources. If photometric redshifts can be obtained, a sophisticated approach would be to take into account the probability distribution function of these redshifts and use their uncertainties in the analysis, such as in Fotopoulou et al. (2016). That being said, when high-quality photometric redshift data is unavailable, using well-defined selection criteria and carefully modelling their completeness (as is done in this work) provides an effective and efficient alternative.

Using this combination of X-ray and optical pre-selection techniques, we identified 223 X-Ray selected AGN z $\sim$ 2 - 4 candidates in the CFHTLS D1, D3 (AEGIS) and D4 fields.

3.3 Keck Spectroscopy

Follow-up spectroscopy for these selected AGN targets was carried out with the Keck Observatory. The combination of Keck's large 10 m mirror, taking long exposure times, and use of Keck's Low Resolution Imaging Spectrometer (LRIS) allows us to obtain spectra for some of the faintest objects in the universe. Such a system is necessary when probing very distant and faint objects, such as our X-Ray selected AGN at high redshifts.

To make for a more efficient use of observing time, certain targets were given 'priority' status. Slitmasks were designed that typically included 1-3 of the primary targets (and a number of filler targets), with an assortment of single slits used to observe a few of the brightest sources. Observations were carried out on the following dates for the D1, D3 and D4 runs: 18/10/09, 19/10/09, 11/04/10, 12/04/10, 11/10/10, 12/10/10, 07/03/11, 09/03/11, 27/03/12, 15/05/12.

The spectra obtained are generally of very high quality, with obvious broad and narrow emission lines. However, some are significantly more noisy; this could be due to poor observing conditions or the fact that the source is incredibly faint/obscured, so much so that the noise is substantially greater than the signal. Figure 3.11 shows the range of spectra varying in both quality and classification obtained from our priority targets across the D1, AEGIS (D3) and D4 fields. Panels (a), (b) and (e) show good examples of broad line AGN, Lyman Break Galaxies (LBG) and narrow line AGN spectra respectively, with little noise and obvious emission lines. However, panel (e) particularly stands out as it looks like a low-redshift contaminant - a source that lies outside of our desired redshift range - due to the [OIII] doublet seen at ~ 5450 $\mathring{A}$, Hβ emission line at ~ 5300 $\mathring{A}$, and [OII] emission line at ~ 4100 $\mathring{A}$. These emission lines would place this AGN at z ~ 0.1. Panels (c) and (d) are seen to be noisier than the others shown, and with less prominent emission lines. The signal in (c) is particularly hard to extract emission lines from. It is this wide range in data quality, and the large potential for spurious candidates in our selected sample, that highlights the need for a more certain method of source and redshift classification than by-sight alone.

3.4 Automated Spectral Fitting and Source Classification

In order to accurately characterise the growth at $2 < z < 4$ of both the obscured (non broad line: NBL) and unobscured (broad line: BL) AGN populations in our sample individually, to ultimately obtain X-ray luminosity functions at $z \sim 2$ - 4 for BL and NBL AGN, an accurate redshift estimate and spectral classification must first be obtained for all priority spectra. We did this through creating a python code which compares the source spectra to AGN spectral templates redshifted at intervals across the desired range, carries out a reduced χ^2 test, and ultimately determines the best-fitting template and returns the estimated best-fitting redshift.

We adopted a range of spectral templates; Figures 3.12 and 3.13 illustrate those which we adopted. Three broad line QSO templates were taken from the selection used by SDSS for spectral cross-correlation; one is a standard QSO (henceforth known as the 'QSO' template) the other two contain broad absorption lines (henceforth known as 'QSOBAL1' and 'QSOBAL2')[2]. We adopted another broad line template from Bolton et al 2012 ; the first Eigenspectrum from their BOSS redshift and classification template basis set for Quasars. The final two templates we adopted were non-broad line templates taken from Hainline et al. (2011) and Shapley et al. (2003), both of which are created from BX/LBG selected sources at $z \sim 2$ - 3. The Hainline template is a stack of objects identified as AGN based on the presence of narrow emission lines, and the Shapley template is a stacked spectrum for large samples of galaxies, and as such shows strong narrow Lyman-α emission but generally lacks other emission lines in the relevant rest-frame UV wavelengths. This template allows for the presence of sources in our X-ray selected sample that lack signatures of AGN at optical (rest-frame UV) wavelengths. See Figures 3.12 and 3.13 for all templates used in this thesis.

As previously mentioned, the spectral fitting and source classification code we developed takes each spectral template and redshifts them at intervals defined by the user, through any desired redshift range. It then re-samples the templates to the wavelength grid of the data. Next, it carries out a reduced χ^2 test between every redshifted template and Keck spectrum. The χ^2 equation is shown by

[2]http://classic.sdss.org/dr5/algorithms/spectemplates/

Equation 3 below:

$$\chi^2 = \sum_i \left[\frac{N y_{\mathrm{mod,i}} - y_{\mathrm{dat,i}}}{\sigma_i} \right]^2 \tag{3}$$

where N is a normalisation factor, $y_{\mathrm{mod,i}}$ are the corresponding flux values for the redshifted template (model), $y_{\mathrm{dat,i}}$ are the flux values for our data and σ are the error values for our data. The normalisation factor is determined for each model-data pair by differentiating Equation 3 and solving for N. The solution is shown by Equation 4 below:

$$N = \sum_i \frac{2 y_{\mathrm{mod,i}} y_{\mathrm{dat,i}}}{{\sigma_i}^2} \div \sum_i \frac{2 y_{\mathrm{mod,i}}^2}{{\sigma_i}^2} \tag{4}$$

with all terms representing those as defined by Equation 3. Once calculated, the normalisation factor N is fed back into Equation 3, and a reduced χ^2 test is carried out:

$$\chi_\nu^2 = \frac{\chi^2}{\nu} \tag{5}$$

where χ^2 is calculated from Equation 3 and ν is the degrees of freedom for that particular Keck-redshifted template pair. What we then end up with is, for each model-data pair, a grid of $\tilde{\chi^2}$ as a function of redshift. Taking the lowest reduced χ^2 gives the nominal 'best-fitting' redshift and template (i.e. classification).

However, we discovered this alone was not enough to return an accurate redshift and source classification for all Keck data. Some of our Keck spectra are dominated by noise; as the code would attempt to fit these noise-dominated spectra anyway, it would return a 'best' reduced χ^2 even if it finds no unique solution. This could leave us with incorrect or 'false' classifications.

In other cases, where the Keck spectrum being fit is mainly comprised of a noisy continuum but still has a subtle emission line, taking the lowest reduced χ^2 is less reliable. In these cases, because the emission line is much more subtle than the noisy continuum, fitting one very flat spectral template to the noisy continuum can often return a slightly lower reduced χ^2 than another, less flat, template which picks up on the subtle emission line (this is hereafter referred to

as 'continuum fitting'). Once more, this would return less accurate z estimates and classifications.

We have developed a systematic method to identify cases where the spectral fitting results are failing to provide an accurate redshift and classification.

In the cases where the lowest reduced χ^2 for a template-Keck spectrum pair one where the corresponding normalisation factor (N, see Equation 4) is very small or negative, the fitting code deems that particular result unreliable. This can happen when the Keck spectrum being fit is solely comprised of noisy continuum and so, when the template normalisation factor N is being calculated in the reduced χ^2 evaluation, the 'best fitting' template is one with such a small normalisation that it is practically a flat line (or an inverted template, in those cases where $N < 0$). Taking this into account, we modified the fitting code to flag any template-Keck spectrum result where $N \leq 1 \times 10^{-6}$ as having an 'unreliable N'. If this happens with every template for a particular Keck spectrum, that source is classified as 'unclassifiable'; none of the spectral templates are able to reasonably fit these spectra without being flattened or inverted, so the spectrum is deemed unclassifiable.

To identify classifications that were the result of 'continuum fitting', we examined the reduced χ^2 vs z graphs for each Keck spectrum. In these cases, the reduced χ^2 vs z graph for those templates show a characteristic upwards "scoop" after their supposed global minima (see markers (a) - (d) in the top right panel of Figure 3.15), as opposed to the steeper, more obvious parabola observed around minima that are typical of accurately fitted emission lines (see marker (e) in the bottom right panel of Figure 3.15). This "scoop" tends to start when an ill-fitting template is probing the lower redshift space when doing reduced χ^2 calculations for a particular, noisy continuum-dominated Keck spectrum. At lower redshifts, our templates lack many discerning spectral emission lines between 3500 - 5500 $\mathring{A}$ (the usable wavelength range of our Keck spectroscopy). As such, in these cases and at these redshifts, a template with few features is being fit to a noisy Keck spectrum with few features - this can produce a relatively low reduced χ^2 (a supposed 'good-fit') (marker (a) in the top right panel of Figure 3.15). However, as the code then starts to increase the redshift of the template, more of the template's emission lines start to shift into the 3500 - 5500 $\mathring{A}$ range; with increasing redshift, the template increasingly becomes a worse and worse fit for this Keck spectrum with a noisy continuum and thus, the reduced χ^2 starts to slowly increase and "scoop" up (marker (b) in the top right panel of Figure 3.15).

It then reaches a peak at about z $\sim$ 3.7, after which it decreases again (markers (c) - (d) in the top right panel of Figure 3.15); at z $\sim$ 3.7, Lyman-α (the first and strongest emission line present in our template spectra) shifts out of the 3500 - 5500 $\mathring{A}$ range, and so the code is once more fitting a featureless template continuum to a noisy Keck spectrum continuum, resulting in the 'goodness' of the fit slowly increasing (corresponding to a decreasing reduced χ^2). The global minimum of the reduced χ^2 in these cases of 'continuum fitting' tends to correspond to the initial low reduced χ^2 at lower redshifts, when the a featureless template continuum is being fit to a noisy Keck spectrum continuum (marker (a) in the top right panel of Figure 3.15) - this means that such 'continuum fitting' can lead to false redshift estimates and classifications (as it can result in classifications which are based on no emission lines/spectral features). What's more, in some cases the global minimum produced by one template's 'continuum fit' can be slightly lower than the the more reliable global minima of other templates that have managed to fit actual emission lines - if our code simply selected the lowest reduced χ^2 in this case, the resulting classification and z estimate given to that source would be unreliable. It is therefore important that our code is able to discern such situations; we modified the code accordingly so that it can identify the scenarios where 'continuum fitting' was occurring and discount them as valid classifications (i.e. 'unclassifiable').

The last addition we made to our spectral classification code, was one that flagged all of our Keck sources with final redshift estimates of $\gtrapprox$ 3.7 as 'unreliable'. We did this because, as previously mentioned, Lyman-α is not seen within the wavelength range of our Keck spectroscopic data at these redshifts, so any classifications resulting from these fits will not be based on fitting emission lines.

In short, our code now effectively returns accurate z values for each of our source spectra, and classifies them as either BL/NBL (unobscured/obscured) or 'unclassifiable'.

The following spectral fitting results are seen illustrated in Figures 3.14 and 3.15:

1) Top left panel of Figure 3.14 - A BL AGN fit, with a spectroscopic z estimate of 2.54 and BL AGN classification. The Keck AGN spectrum is in black, the best fitting spectral template is overlaid in blue, and the noise spectrum is in red. You can see from the corresponding reduced χ^2 graph (top right panel of Figure 3.14) that the BLAGN Bolton template was chosen due to the evident minimum seen in the reduced χ^2 graph for that template.

2) Middle left panel of Figure 3.14 - A NL AGN fit, with a spectroscopic z estimate of 2.85 and NL AGN classification. The Keck AGN spectrum is in black, the best fitting spectral template is overlaid in blue, and the noise spectrum is in red. You can see from the corresponding reduced χ^2 graph (middle right panel of Figure 3.14) that the Shapley template was chosen due to the evident minimum seen in the reduced χ^2 graph for that template; although, it is worth commenting on the fact that a fit to a single emission line is less reliable than if multiple emission lines were present in the Keck spectrum.

3) Bottom left panel of Figure 3.14 - An AGN given the 'unclassifiable' classification and consequently no spectroscopic z estimate. Upon inspection, you can clearly see the source spectrum is just noise, and the resultant 'best fit' template overlaid in blue is not a meaningful fit.

4) Top panels of Figure 3.15 - A clear example of a case of 'continuum fitting', resulting in an 'unclassifiable' spectroscopic AGN classification. Markers (a) - (d) on the right-hand reduced χ^2 plot show the characteristic 'scoop' of the 'continuum-fitting' case.

5) Middle panels of Figure 3.15 - A clear example of a case where the normalisation N of the 'best-fitting' template is $\leq 10^{-6}$, resulting in an 'unclassifiable' spectroscopic AGN classification. Looking at the 'best-fit' template spectrum (solid blue line on the left panel), you can see how it is in fact an inverted template spectrum that is the 'best fit'.

6) Lower panels of Figure 3.15 - A clear example of a case where spectral emission lines are being distinctly picked up on and fit by the code. Marker (e) shows the distinct minimum which are characteristic of emission lines being fit, in contrast to the gentle 'scoop' feature seen in the reduced χ^2 plots which result from 'continuum-fitting' (see markers (a) - (d) in the top right subplot in Figure 3.15).

It is worth noting that some of the χ^2 values in the aforementioned figures span a surprisingly large range on the y-axis. This is due to the fact that our Keck spectra are very noisy outside of the 2000 - 6000 $\mathring{A}$ wavelength range. When a very noisy section of a spectrum is then compared to an already ill fitting template at an incorrect redshift, the resultant χ^2 is understandably very high. We therefore see high χ^2 values for ill-fitting template-spectrum pairs at incorrect redshifts and a largely dynamic y-axis in the aforementioned χ^2 vs. z Figures.

Figure 3.16 shows the $\log_{10}L_X$ - redshift distribution of our sources, with

histograms showing the projection of each axis. Dark green and yellow markers depict BL and NL AGN classifications respectively; solid triangles, crosses and pentagons illustrate sources which lie in the D4, D1 and D3 (aegis) fields respectively. Figure 3.17 shows the X-ray - optical flux distribution of our sources, with histograms showing the projections of each axis. Colours and markers correspond to those in Figure 3.16, with the addition of the red histograms and markers - these illustrate our 'unclassifiable' sources. The dashed red line at an r magnitude of 25.5 represents the optical magnitude cut applied to our sample, and grey points indicate a simulated sample of sources which we go on to create in order to constrain the completeness of our spectroscopic classification process; this is described later in detail in Section 3.5. Our sample consists of 47 BL AGN classifications, 45 NL AGN classifications, 2 'unreliable' BL AGN classifications, 3 'unreliable' NL AGN classifications and 8 'unclassifiable' sources.

3.5 Creating Spectral Simulations

With only ∼110 spectra, one could argue that developing a specific spectral fitting routine is not strictly necessary - classifications and redshifts could be derived by eyeballing each spectrum in turn. However, a significant advantage of creating and using such an automated routine is not only that it will automatically fit and perform z-estimations and AGN source classifications, but we will consequently be able to characterise the incompleteness and uncertainty of the spectroscopic classification and redshift estimation step of this analysis by running said fitting code on thousands of simulated spectra. Due to the obvious faintness and obscuration of our sources, our spectroscopic classification and redshift estimations won't be perfect; brighter, less noisy, less obscured spectra will likely be fit more accurately and the fainter, noisier, obscured AGN spectra may be harder to reliably identify a classification and redshift for. In running our spectral fitting code on a multitude of simulations, spanning broad magnitude and redshift ranges, with signal and noise properties representative of our real Keck spectra, we can accurately quantify my spectral classification success rate as a function of redshift, optical magnitude and source classification and constrain how well it works on our low luminosity, higher-z, obscured AGN sample. We can then apply these corrections in our subsequent XLF calculations, to generate as robust a result as possible.

In order to generate thousands of simulated AGN spectra which span the full

magnitude and signal properties of our real Keck sources, we need to start by creating basis AGN spectra, which we can then modify accordingly to reflect our real data. The AGN spectral templates used to fit our Keck data (see Figures 3.12 and 3.13) provide us with effective 'starting points' for our spectral simulations.

In order for our spectral simulations to span a full range of continuum emission strengths and wavelengths, we need a measure of the continuum component for each of our AGN spectral templates. We carried out by-eye continuum fits for each of the spectral templates to be used in the simulation process. Each spectral template continuum was fit with a simple power-law model:

$$F_\lambda = F_0(\lambda/\lambda_0)^{\alpha_\lambda} \tag{3.2}$$

where F_0 is the rest-frame flux normalisation at λ_0. This was done so that we could essentially create continuum-divided versions of each of these templates (obtaining the ratio of emission lines to continuum emission), so that we could then multiply the continuum component of these templates back in at varying strengths for a range of simulations. This essentially allows us to vary the continuum strength and emission line strengths across our spectral simulations, adding breadth and diversity to the simulated sample. Fitting the continuum emission for each template also allowed us to extrapolate said templates to bluer wavelengths, if they didn't already extend down to $750\mathring{A}$ - a necessary measure when hoping to produce spectral simulations at redshifts $\sim$ 2 - 4 that have visible emission (continuum or elemental line) at optical/UV wavelengths. Figure 3.18 below shows the aforementioned process for one of our template spectra (QSOBAL1).

In order to ensure that our spectral simulations were varied enough to reliably test to bounds of our spectral fitting code, we use varying combinations of pairs of our continuum-divided, extrapolated spectral templates as a basis for each individual simulation. Our simulation routine generates base AGN spectra from randomised ratios of pairs of the template spectra in Figures 3.12 and 3.13. So, for example, it may choose to generate a base spectrum using 43% of the signal from the SDSS QSO template spectrum and 57% from the SDSS QSOBAL1 template spectrum. This generates a mixed, extrapolated, continuum-divided basis spectrum. Figure 3.19 showsan example of how this process is done. In yellow we have a z = 2.55 simulated optical/UV AGN spectrum created using 100% of the SDSS QSO template and 0% of the SDSS QSOBAL1 template, with a continuum slope of 2.

This simulated spectrum also incorporates the signal/noise properties of one of our real Keck spectra, shown in dark blue.

Once a mixed, extrapolated, continuum-divided basis spectrum has been created, a continuum component of choice (following the form of Equation 3.2) is then folded back into this basis spectrum. Our simulation code randomly generates these continuum components over a wide range of α_ν values, drawn from a Gaussian distribution with μ = -0.5 and σ = 2.

Following this, our simulation code then redshifts the templates over any desired range of z values (for this work we created simulations ranging from z: 1.5 - 4.5) and accounts for intergalactic absorption using the Madau (1995) intergalactic attenuation model (see Figure 3.20 for graphical representation of the Madau (1995) intergalactic attenuation model at varying redshifts).

Once the redshifted, intergalactic absorption-applied basis spectra have been generated, the next step of the simulation process involves the re-scaling of these basis spectra and addition of noise, using the signal and noise properties of our real, individual Keck AGN spectra. In order to re-scale our redshifted simulated basis spectra to have the same signal properties as our observed Keck spectra, the expected flux, $\langle f_\nu \rangle_{Keck}$ for each Keck source was calculated using the following equation:

$$\langle f_\nu \rangle = \frac{\sum f_{\lambda,i} T_{p,i} \Delta\lambda_i}{\sum T_{p,i} \frac{c}{\lambda_i^2} \Delta\lambda_i} \tag{3.3}$$

where, for individual wavelength-flux spectral measurements i in any spectrum, f_λ is the flux density, T_p is the CFHTLS r-band filter transmission curve, c is the speed of light in a vacuum (2.998×10^8 ms^{-1}), λ is the wavelength, and $\Delta\lambda$ is the separation between wavelength data points. $\langle f_\nu \rangle_{template}$ can then be calculated for each basis template spectrum using the same equation (Equation 3.3), and the scaling factor A can be derived using:

$$\langle f_\nu \rangle_{Keck} = \langle f_\nu \rangle_{template} \times A \tag{3.4}$$

By deriving individual scaling factors between every single basis template spectrum and our Keck sources, we can fold these into our basis template spectra and effectively ensure that every simulation we produce is based on the signal

and magnitude properties of our real Keck data.

The noise properties of individual Keck spectra are similarly used to add realistic noise to our simulated spectra. The noise data for each Keck source is in the form of an array, sampled to the wavelength grid of said Keck source, with values denoting the standard deviation, σ, of a Gaussian distribution (with mean μ = the observed signal) which represents the statistical noise for the source at every measured wavelength value. It is therefore straightforward to use the noise properties of our Keck data in our simulations; noise is drawn from these distributions and added to our simulations as such. Figure 3.21 below broadly illustrates how our simulations are created, along with an example of one of our final simulated AGN spectra.

Thus, thousands of BL and NL AGN optical/UV spectroscopic simulations are produced over a range of redshifts, r magnitudes, continuum and emission line strengths, from varying combinations of templates, with realistic signal and noise properties representative of our sample. All of these steps go on to ensure that the vast number of spectral simulations we have produced vary across a wide range of r magnitudes, redshifts and continuum/emission line strengths, with signal/noise properties that reliably reflect our real AGN spectra. This provides a diverse and thorough testing space over which my spectral fitting code can be characterised.

We then run our spectral fitting and source classification code on these simulated spectra and obtain source classifications and redshift estimates for them. Using these results, we can go on to quantify our "spectroscopic success rate" as a function of z, r' mag, and classification.

3.6 Spectral Simulations: Results

Running our spectral fitting code on our diverse spectral simulations allowed us to constrain the probabilities of our code successfully (or unsuccessfully) classifying sources as functions of r band magnitude and redshift, for varying input and output classifications. Figures 3.22-3.43 illustrate these diverse results; I discuss each figure below, in turn.

Figures 3.22 and 3.23 are the inverse of each other. They show that my automated spectroscopic classification process has a high success rate when classifying brighter BL AGN across the entire redshift range, with near 100% success from z

$\sim$ 1.5 - 3.5 for spectra of 20 - 22 r magnitude. This only drops to between 60 - 80% as we extend to fainter sources, with an understandable decrease in success rate at the faintest r mags and the highest redshifts. The high failure rate at z $>$ 3.5 is due to the shifting of the Lyα emission line, out of the wavelength range visible in our optical-UV spectra. This results in a higher possibility of zero-emission lines being present in our spectra at these redshifts, making it difficult for my fitting procedure to accurately classify AGN at these redshifts.

Figures 3.24 to 3.27 show the probability of my automated spectroscopic classification routine, mis-classifying a true BL AGN as a (i) NL AGN, (ii) noise/unclassifiable, (iii) Unreliable BLAGN and (iv) Unreliable NLAGN. These results show that BLAGN are, uncerstandably, more often mis-classified at the faintest r magnitudes, and the highest redshifts.

Figures 3.28 and 3.29 are the inverse of each other, and are the NL equivalents of Figures 3.22 and 3.23. They similarly show that my automated spectroscopic classification process has a high success rate when classifying brighter BL AGN across the redshift range probed. However, these results highlight the increased difficulty in accurately classifying NL AGN over BL AGN - weaker narrow emission lines are much harder to identify in the optical-UV spectra of faint and high-z AGN, as signal-noise ratios are often poor enough for a narrow emission line to be hidden under noise. Once more, the high failure rate at z $>$ 3.5 is due to the shifting of the Lyα emission line, out of the wavelength range visible in our optical-UV spectra. This results in a higher possibility of zero-emission lines being present in our spectra at these redshifts, making it difficult for my fitting procedure to accurately classify AGN at these redshifts.

Figures 3.30 to 3.33 are the NL equivalents of Figures 3.24 to 3.27. They show the probability of my automated spectroscopic classification routine, mis-classifying a true NL AGN as a (i) BL AGN, (ii) noise/unclassifiable, (iii) Unreliable BLAGN and (iv) Unreliable NLAGN. These results show that NLAGN are similarly often mis-classified at the faintest r magnitudes and the highest redshifts.

Figures 3.34 and 3.35 are the inverse of each other, and show the probability that a source which has been given a BL AGN classification, is *truly* a BL AGN. These figures show that the success rate between z $\sim$ 2.5 - 3.5 is at least 70%, with this rate dipping only slightly outside this range.

Figures 3.36 to 3.38 show the probability that a source which has been given a (i) Noise/unclassifiable, (ii) Unreliable BL AGN or (iii) Unreliable NL AGN

classification, is *actually* a BL AGN. They show that these mis-classifications happen more often at higher redshifts and fainter r magnitudes.

Figures 3.39 and 3.40 are the inverse of each other, and are the NL equivalents of Figures 3.34 and 3.35. They show the probability that a source which has been given a NL AGN classification, is *truly* a NL AGN. Once again, these figures show a high success rate across redshift, which understandably decreases at the faintest r magnitudes.

Figures 3.41 to 3.43 show the probability that a source which has been given a (i) Noise/unclassifiable, (ii) Unreliable BL AGN or (iii) Unreliable NL AGN classification, is *actually* a NL AGN. Once more, they show that these mis-classifications happen more often at higher redshifts and fainter r magnitudes.

It is interesting to note that there are cases where our spectral fitting code fails to classify very bright simulated AGN spectra (i.e. r mag < 20). An example of one such case is illustrated by Figure 3.44. As seen in this figure, the combination of parameters and ratios used has resulted in a simulated spectrum which exhibits large absorption features, in addition to a simulated Lyman break (bottom left panel, yellow solid line). These more complex features are harder for our code to accurately fit and hence, sometimes result in a miss/failed classification. This figure aptly highlights why our code sometimes fails to classify bright, simulated sources - the intricate features found in some of these bright simulated spectra are difficult to fit and cause failures.

These results effectively characterise the success rate of our fitting code for varying AGN populations/classifications, as a function of r band magnitude and redshift, and are to be folded into the XLF calculations in Chapter 4 as completeness corrections.

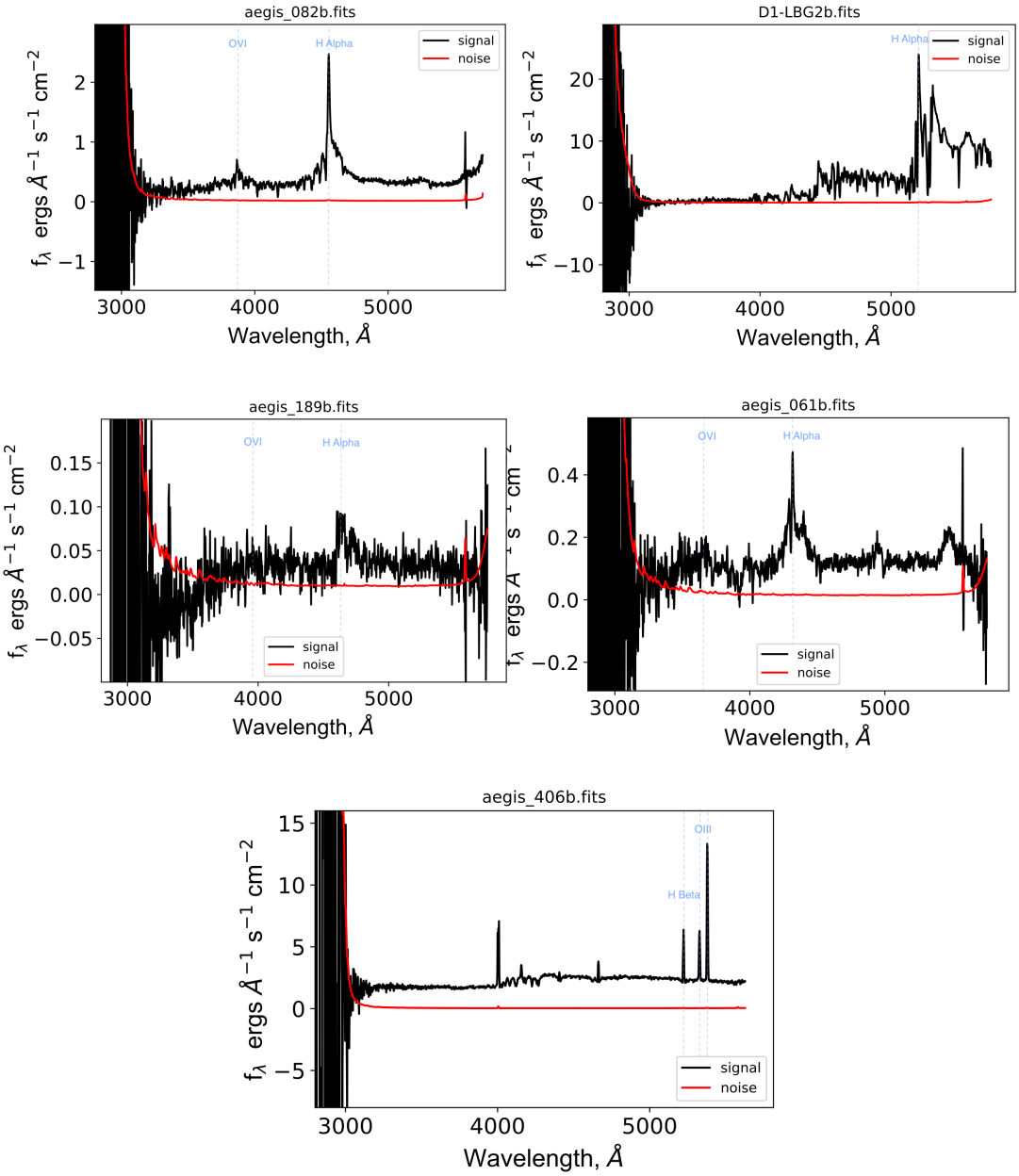

Figure 3.11 *A few examples of the spectroscopy obtained from Keck. In each of the above plots the black curve illustrates the signal, the red curve illustrates the noise, and prominent emission lines are marked in light blue.* ***(a)*** *Spectrum obtained from priority source located in the AEGIS (D3) field. Prominent and broad lines can be seen.* ***(b)*** *Spectrum obtained from priority source located in the D1 field. This spectrum resembles that which is typical of a LBG, with the Lyman break seen to occur at* $\approx$ *4250 Å. The Ly* α *emission line itself can be seen at* $\sim$ *5200 Å, corresponding to a rough redshift estimate of* $\sim$ *3.3.* ***(c)*** *Spectrum obtained from priority source located in the AEGIS (D3) field. This spectrum is decidedly more noisy than others in the sample. However, the 'lump' seen at* $\approx$ *4600 Å shows potential for an emission line.* ***(d)*** *Spectrum obtained from priority source located in the AEGIS (D3) field. Despite being slightly more noisy, this spectrum still resembles that of a typical broad line AGN, with obvious Ly*α *and CIV emission lines. However, spectra of these qualities (both c and d) and with this uncertainty in line identification (a) highlight the need for an automated, less-biased classification process.* ***(e)*** *Spectrum obtained from priority source located in the D3 field. Prominent emission lines can be seen. This particular spectrum resembles that of a lower-redshift contaminant in our sample.*

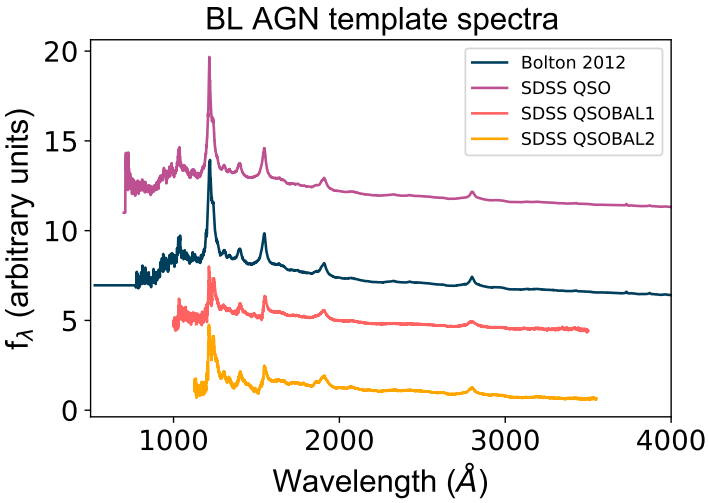

Figure 3.12 *The selection of BL AGN templates used for our spectral fitting process. Templates shown in purple, orange and yellow are all BL templates taken from the SDSS spectral cross-correlation template collection*[3]*. The template shown in dark blue is the first eigenspectrum of the BOSS redshift and classification template basis set for quasars, taken from Bolton et al. (2012)*

.

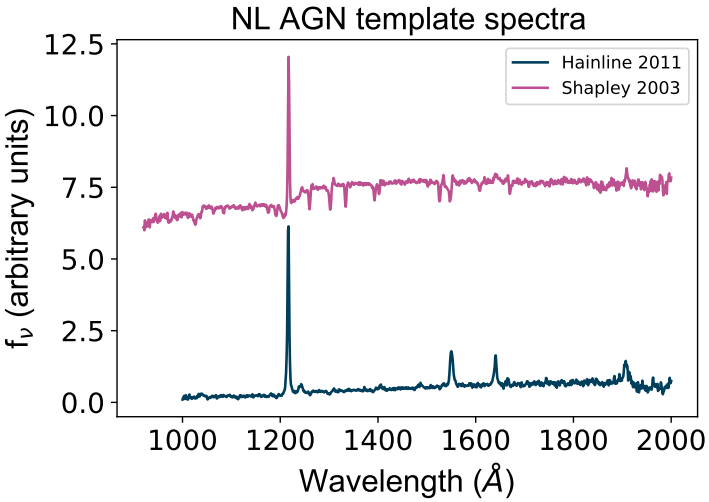

Figure 3.13 *The two NL AGN templates used for our spectral fitting process. The top (purple) and bottom (dark blue) template spectra are taken from Hainline et al. (2011) and* **?** *respectively. The Hainline template is a rest-frame UV composite spectrum for UV selected AGN at $z \sim 2$ - 3, created by the stacking of objects identified as AGN based on the presence of narrow emission lines. The Shapley template is a stacked composite rest-frame UV spectrum for large samples of galaxies at $z \sim 2$ - 3, and as such shows strong narrow Lyman-α emission but generally lacks other emission lines in the relevant rest-frame UV wavelengths. This template allows for the presence of sources in our X-ray selected sample that lack signatures of AGN at optical (rest-frame UV) wavelengths.*

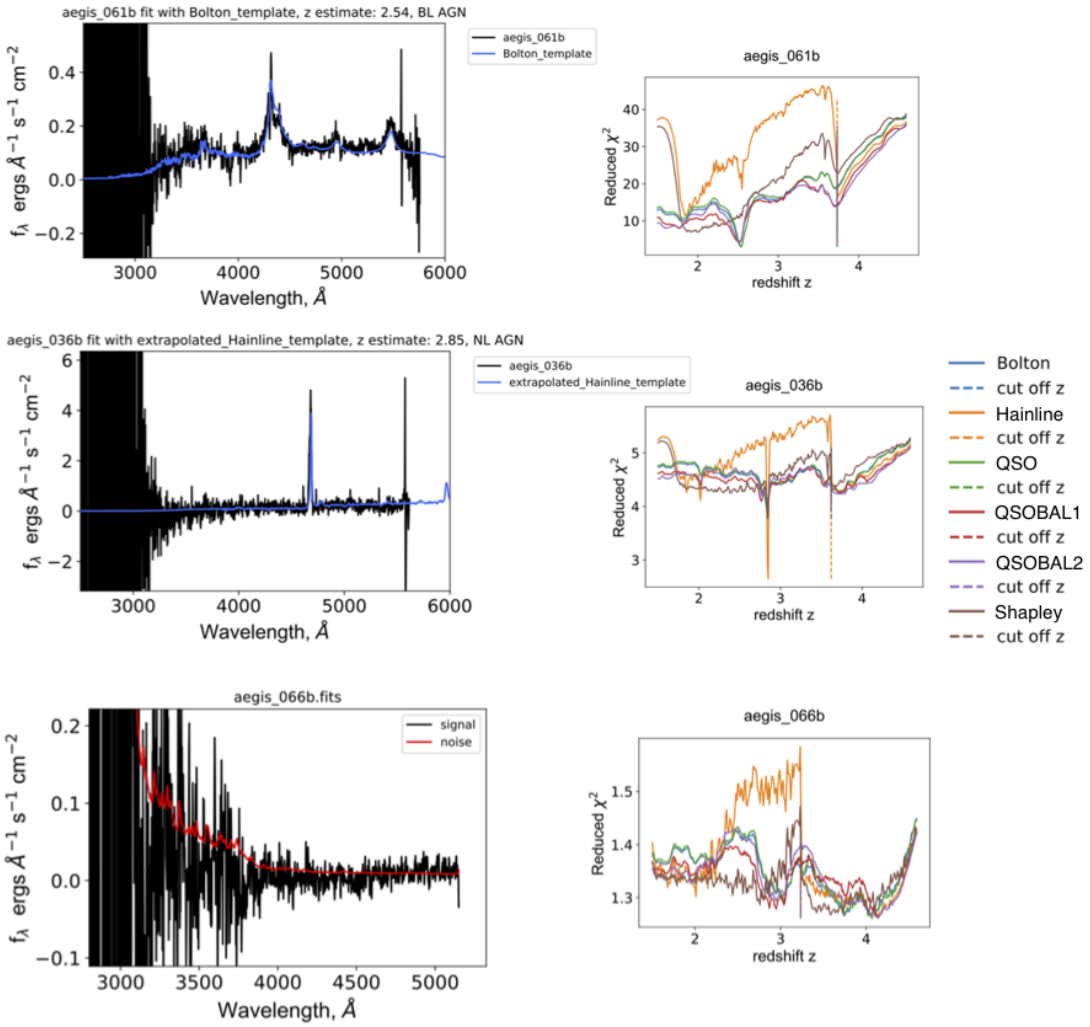

Figure 3.14 ***Right column:*** *Examples of Keck spectra and their resultant spectroscopic fits, classifications and z estimates. The Keck spectrum, corresponding noise spectrum, and the best-fitting template are shown by white, blue and red solid lines respectively. The top, middle and bottom panels are examples of sources which have been respectively classified as 'BL AGN', 'NL AGN' and 'unclassifiable'.*

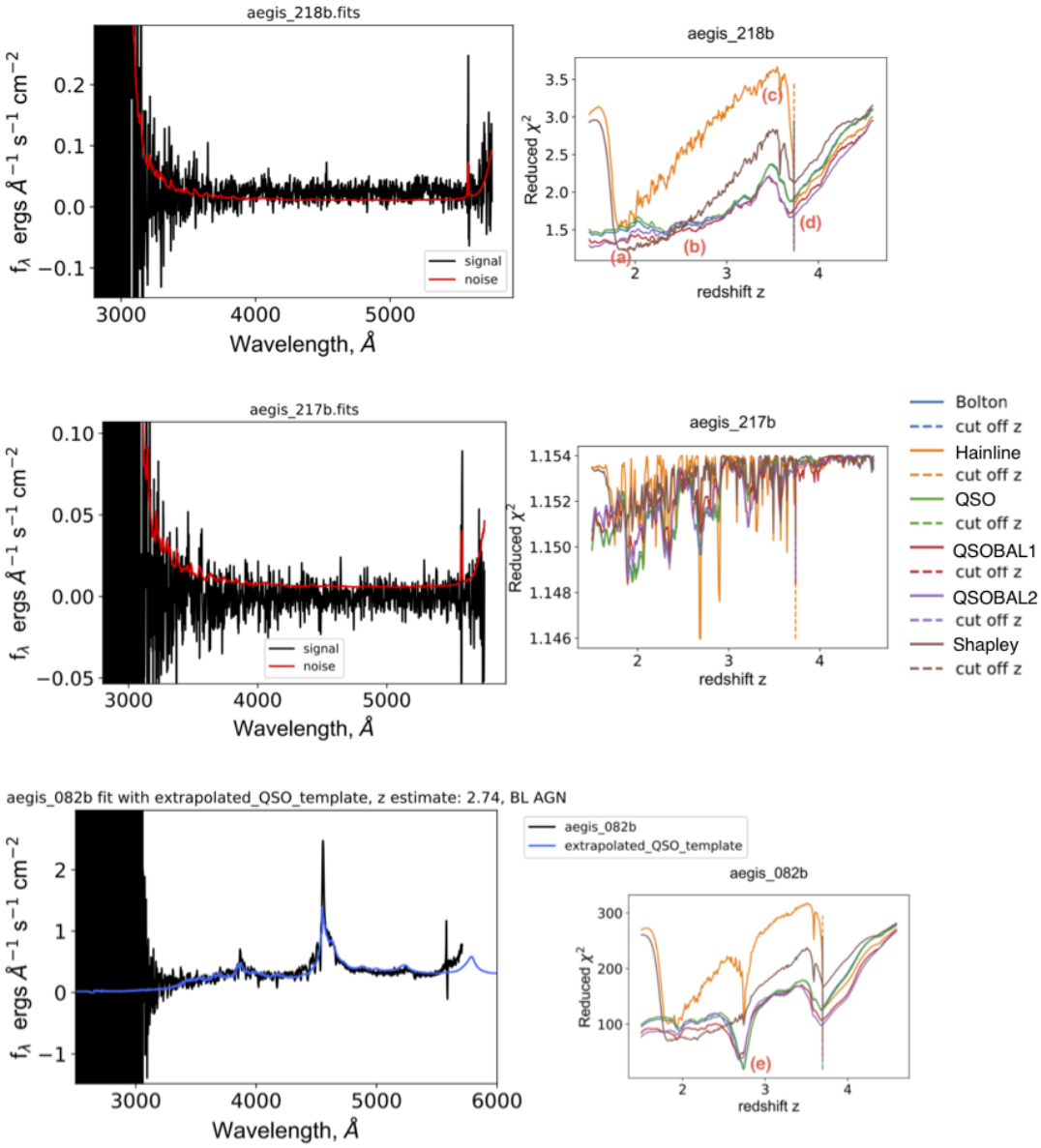

Figure 3.15 ***Right column:*** *Examples of Keck spectra and their resultant spectroscopic fits, classifications and z estimates. The Keck spectrum, corresponding noise spectrum, and the best-fitting template are shown by black, blue and red solid lines respectively. The top panel shows an example of 'continuum fitting', the middle panel shows an example where the normalisation N of the best-fitting spectrum is* $\leq 10^{-6}$*, and the bottom panel shows an example where the code has clearly fit the distinct spectral emission lines.* ***Left column:*** *Corresponding reduced* χ^2 *vs. z plots for the sources in the right column. Markers (a) - (d) in the top right panel show the characteristic 'scoop' shape of the 'continuum fitting' scenario, and marker (e) in the bottom right panel shows the distinct parabolas which are seen when the code is fitting to emission lines.*

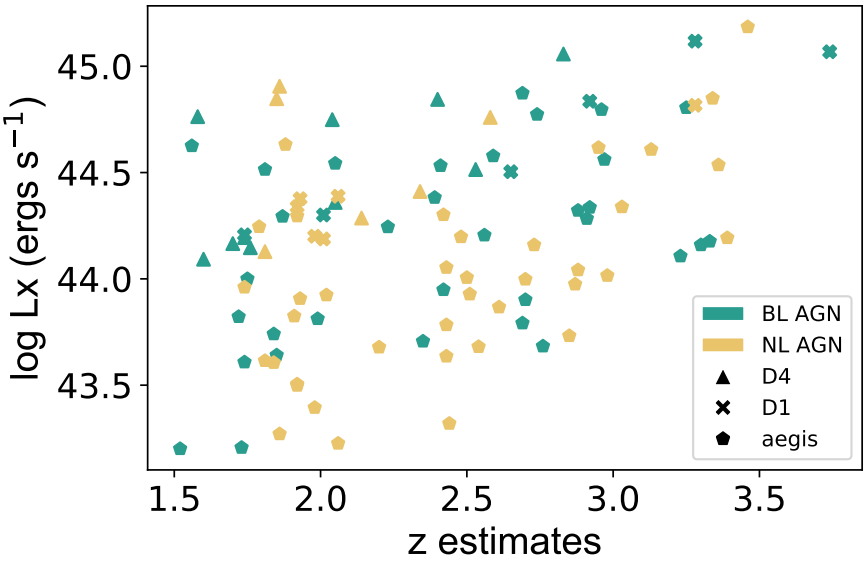

Figure 3.16 *The $log_{10}L_X$ - redshift distribution of our optically and X-ray preselected, spectroscopically classified $z \sim 2$ - 4 AGN sources, with histograms showing the projection of both axis. Dark green and yellow markers depict BL and NL AGN classifications respectively; solid triangles, crosses and pentagons illustrate sources which lie in the D4, D1 and D3 (aegis) fields respectively.*

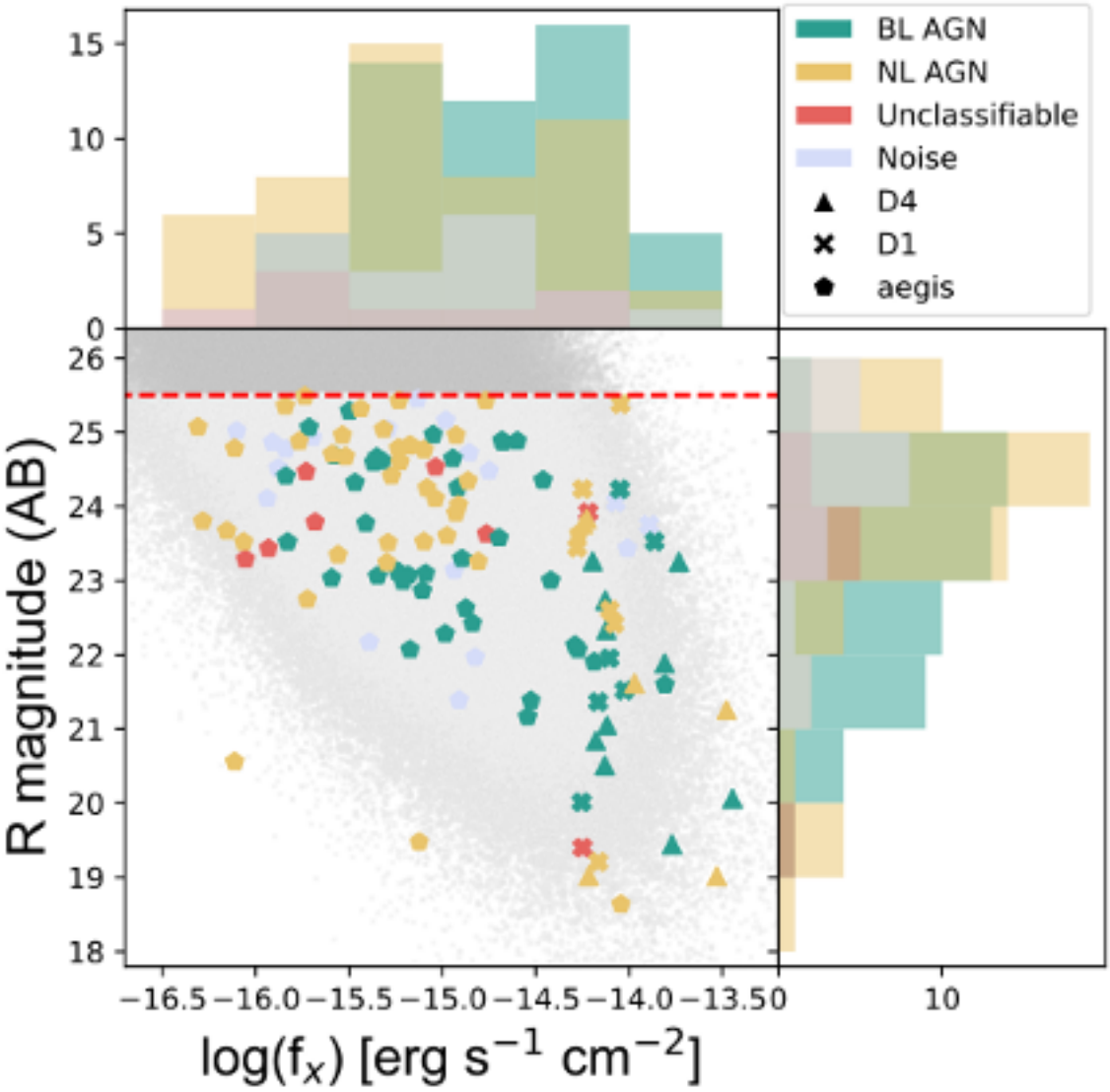

Figure 3.17 *The X-ray - optical flux distribution of our sources, with histograms showing the projections of each axis. Colours and markers correspond to those in Figure 3.16, with the addition of the red histograms and markers - these illustrate our 'unclassifiable' sources. The dashed red line at an r magnitude of 25.5 represents the optical magnitude cut applied to our sample, and grey points indicate a simulated sample of sources which we go on to create in order to constrain the completeness of our spectroscopic classification process; this is described in detail in Section 3.5.*

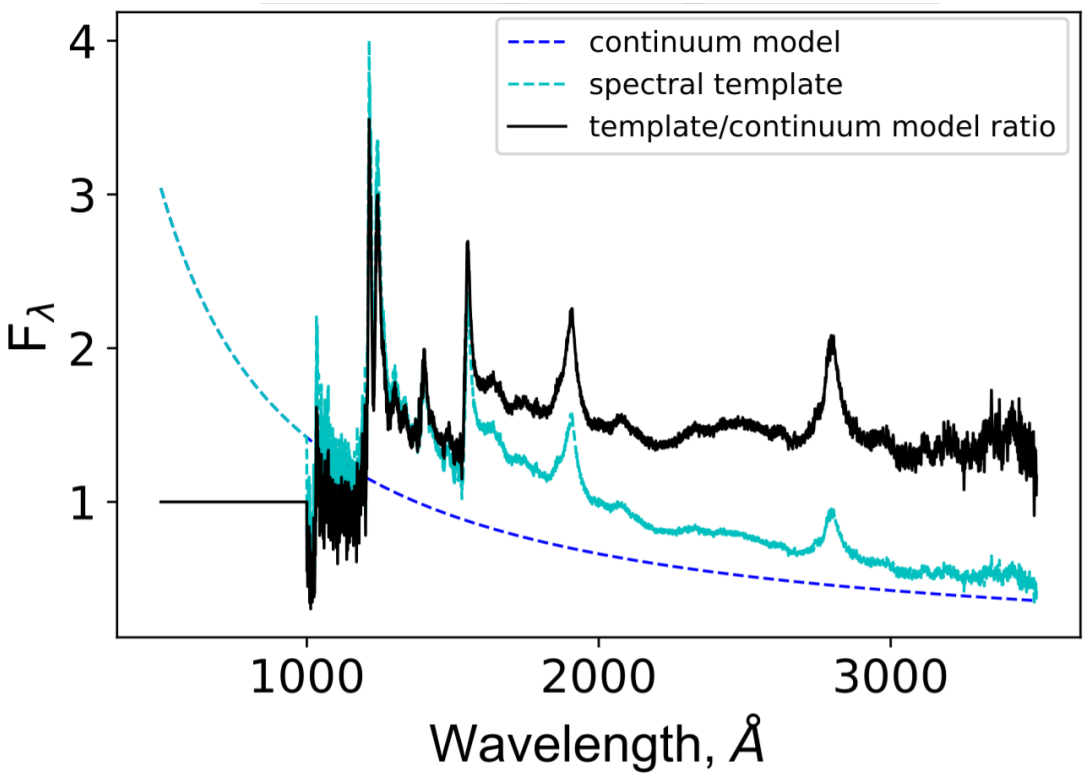

Figure 3.18 *Showing how the continuum is modelled and removed for our template spectra. The dashed light blue line shows the AGN spectral template in question (in this case it is one of our QSO templates that show broad absorption lines, QSOBAL1); the dashed dark blue line shows the continuum model used for this spectral template; the solid black line shows the resulting spectrum when the spectral template is divided by its fitted continuum and extrapolated to lower wavelengths, giving us a 'continuum divided' template (showing the emission-line:continuum ratio) used as a basis for our simulated spectra.*

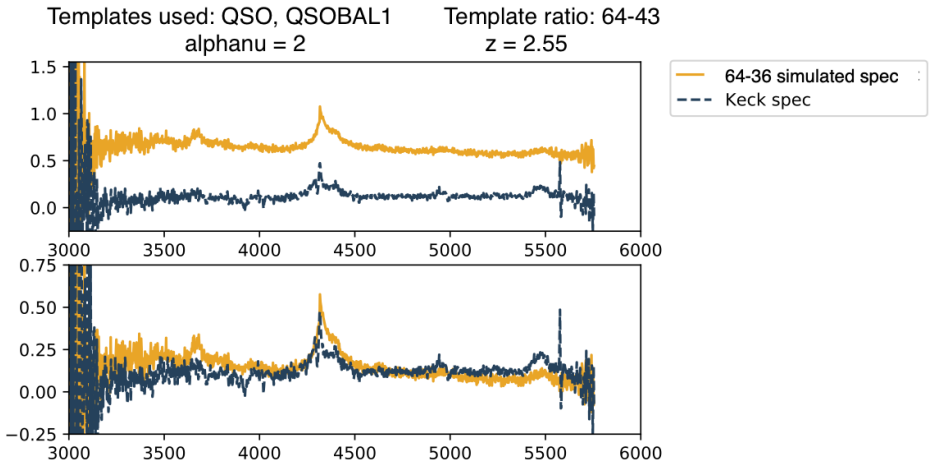

Figure 3.19 *An example of how combining varying ratios of spectral templates and a different continuum slope (α_ν) produces simulated optical/UV AGN spectra across a range of redshifts, continuum strengths and emission line ratios. The yellow solid line is one of these simulated spectra, and is compared to the dark blue dashed line showing our real Keck spectral data. In the top panel, a scale factor is added to ease comparison between both spectra. The bottom panel shows the real scale.*

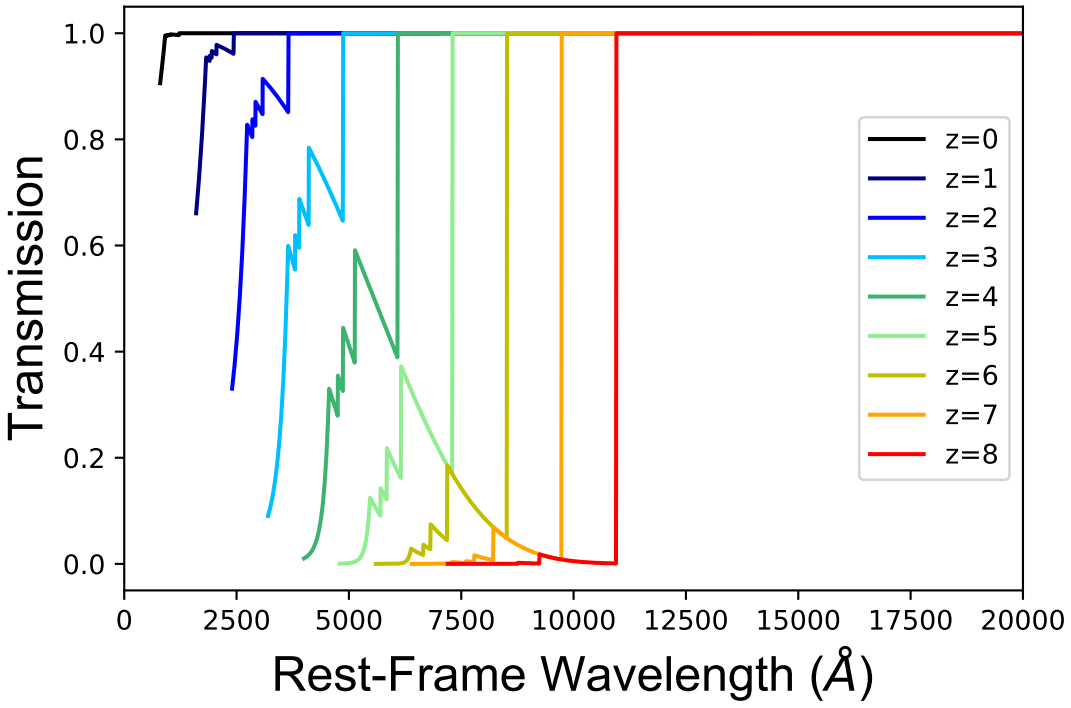

Figure 3.20 *Transmission of the universe as a function of rest-frame wavelength, for varying redshifts, taken/adapted from Madau (1995)*

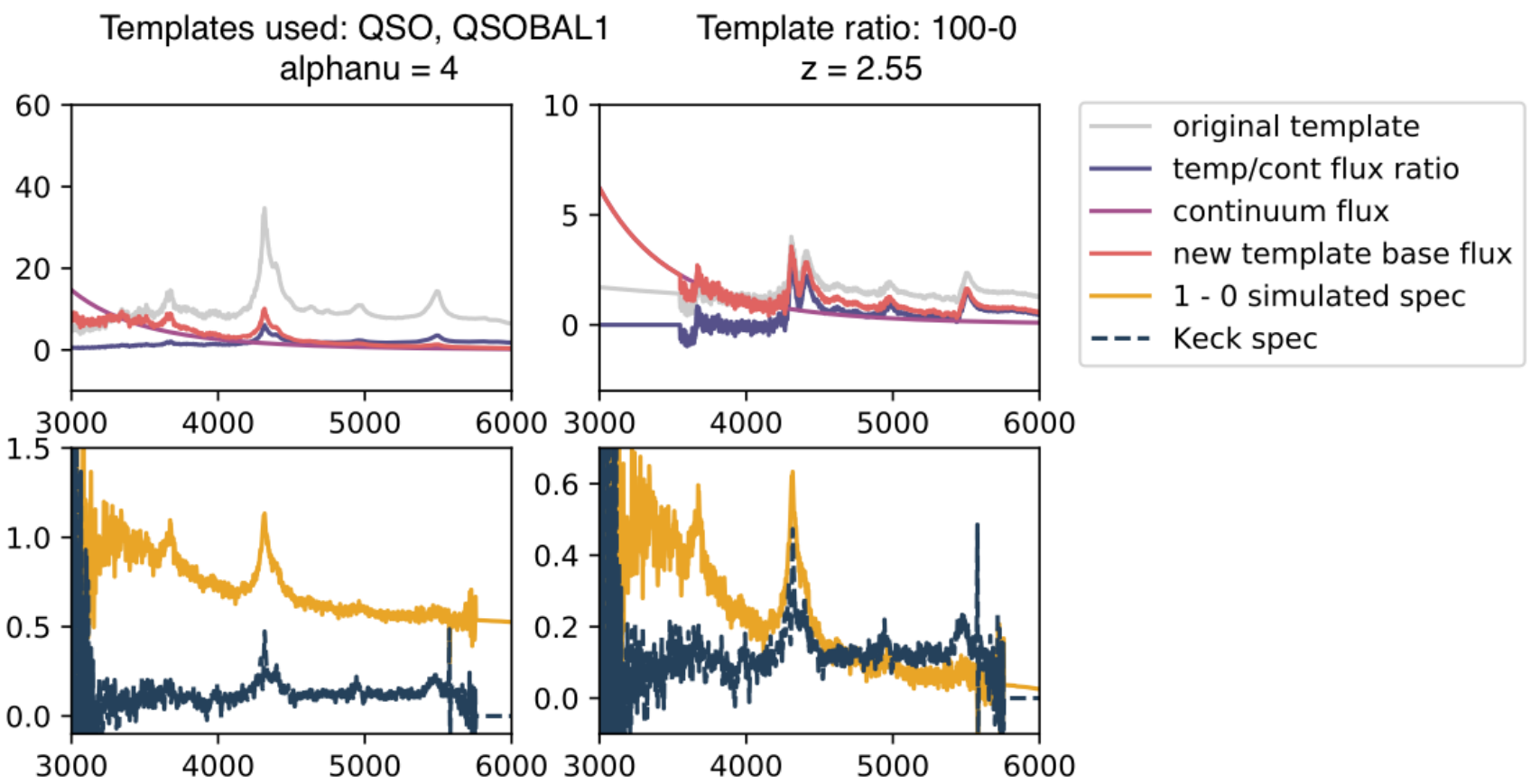

Figure 3.21 *Illustrating the basic spectral simulation generation process.* ***Top two panels:*** *show the spectral templates used, in combination, to produce an example simulated spectrum - in this case, the simulation is generated using a 1:0 ratio of the SDSS QSO template (top left) and the SDSS QSOBAL1 template (top right). In reality, a simulated spectrum is never generated from a 1:0 ratio of template spectra - this ratio is simply used in this figure for illustrative purposes. The original template spectra are shown by the solid grey lines and the continuum fits are shown by the solid light purple lines. The continuum-divided templates are shown by the dark purple lines and then the final basis spectral templates with new continuum components folded in (corresponding to the power law denoted in eq.3.4, with $\alpha_\nu = 4$ in this example) are shown by the solid pink lines.* ***Bottom two panels:*** *Showing the resultant simulated spectrum (yellow solid line) in comparison to the Keck spectrum (dark blue dashed line) from which this simulation's signal and noise properties were taken. The bottom left panel shows the simulated spectrum arbitrarily scaled up in the y-axis so that individual features can be seen clearly from the Keck spectrum, and the bottom right panel shows the same two spectra but plotted on top of each other, for comparison.*

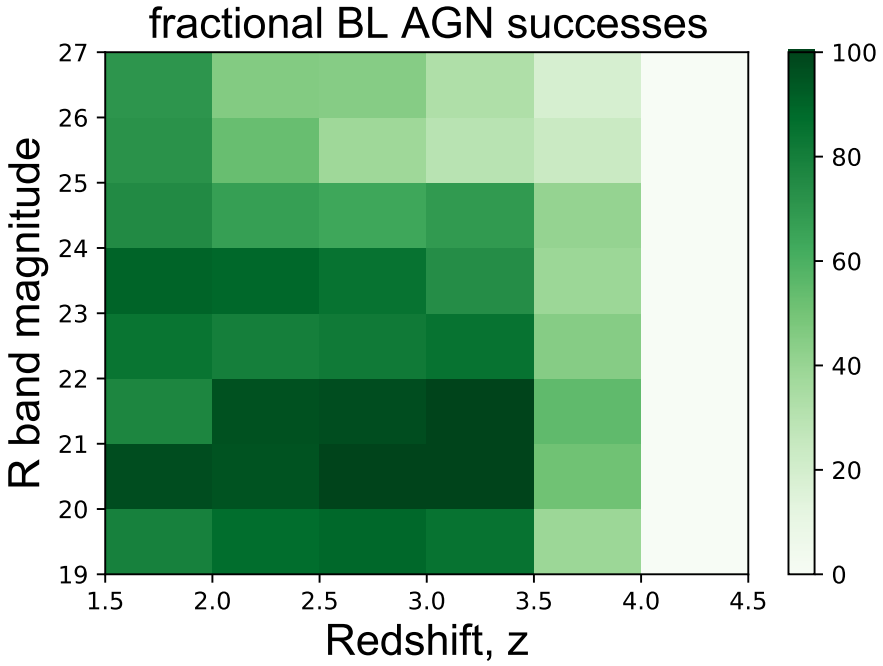

Figure 3.22 *Showing the probability of the spectral fitting code successfully classifying an input BL AGN spectrum as a BL AGN, as a function of output redshift and R band magnitude. Probabilities are shown as a percentage.*

Figure 3.23 *Showing the probability of the spectral fitting code failing to classify an input BL AGN spectrum as a BL AGN, as a function of output redshift and R band magnitude. Probabilities are shown as a percentage.*

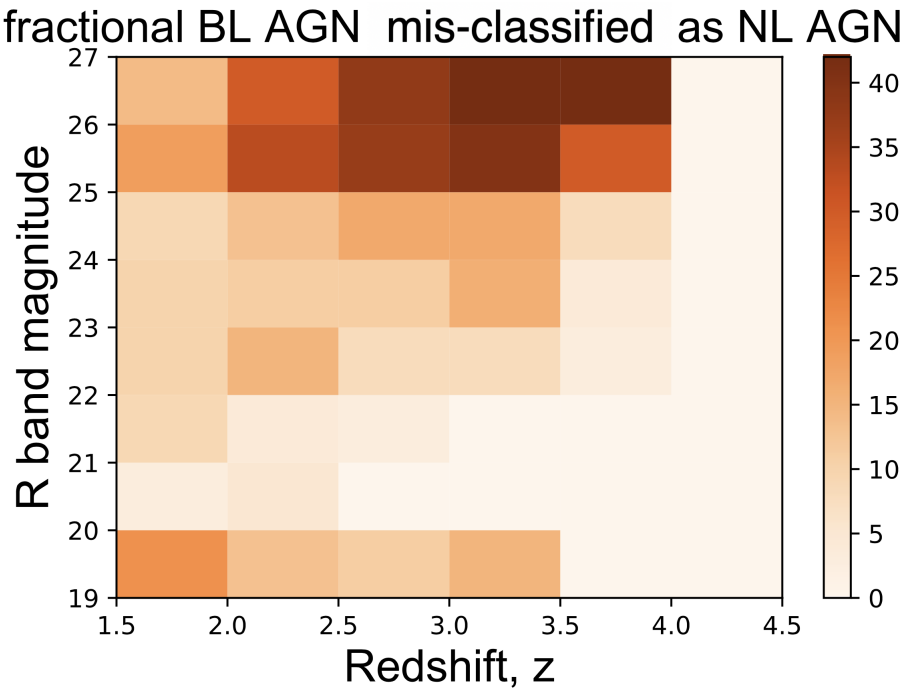

Figure 3.24 *Showing the probability of the spectral fitting code mis-classifying an input BL AGN spectrum as a NL AGN, as a function of output redshift and R band magnitude. Probabilities are shown as a percentage.*

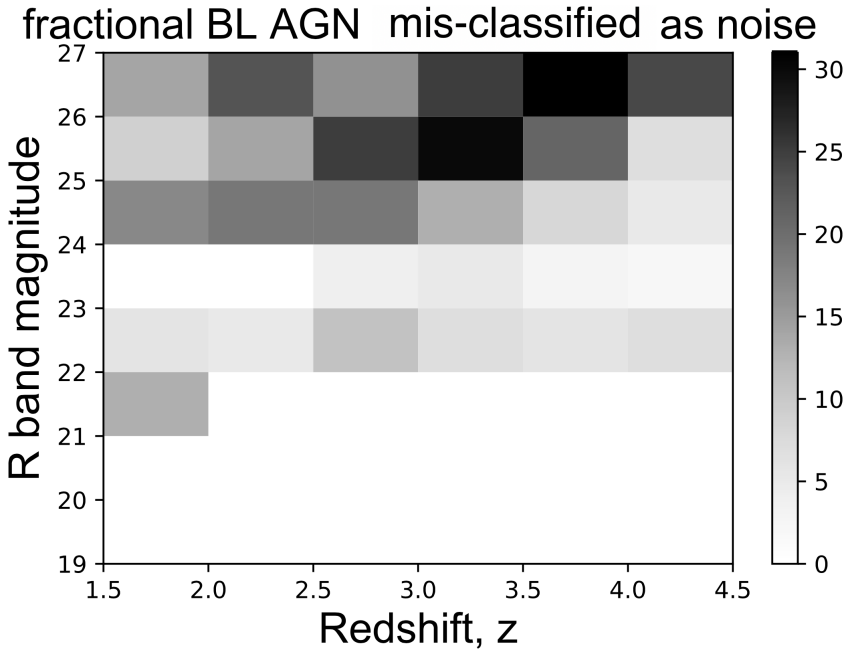

Figure 3.25 *Showing the probability of the spectral fitting code mis-classifying an input BL AGN spectrum as a noise, as a function of output redshift and R band magnitude. Probabilities are shown as a percentage.*

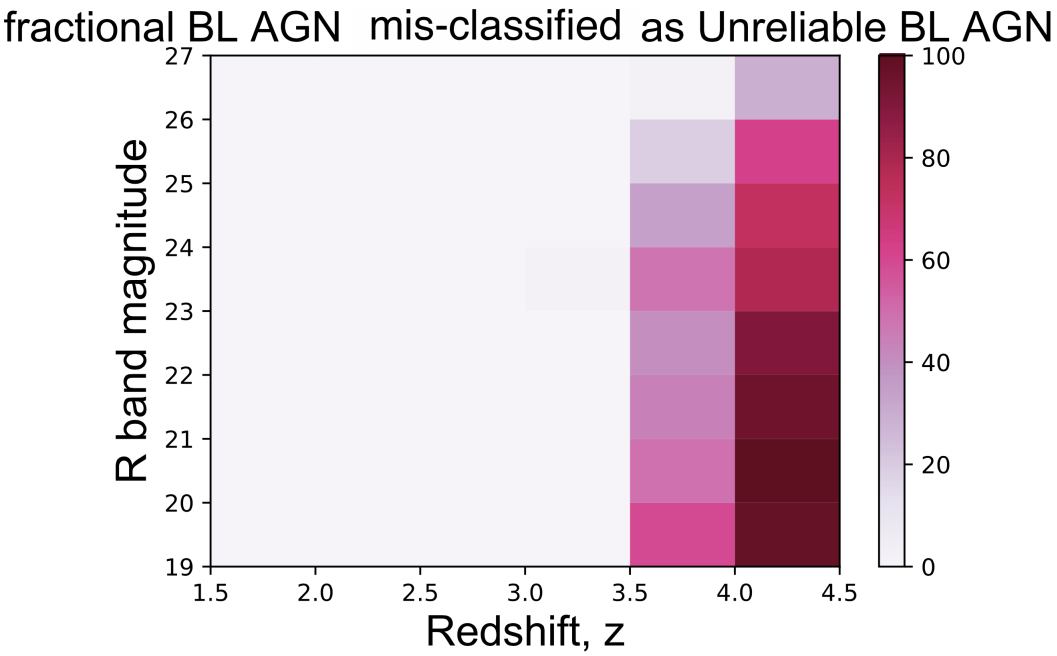

Figure 3.26 *Showing the probability of the spectral fitting code mis-classifying an input BL AGN spectrum as an Unreliable BL AGN, as a function of output redshift and R band magnitude. Probabilities are shown as a percentage.*

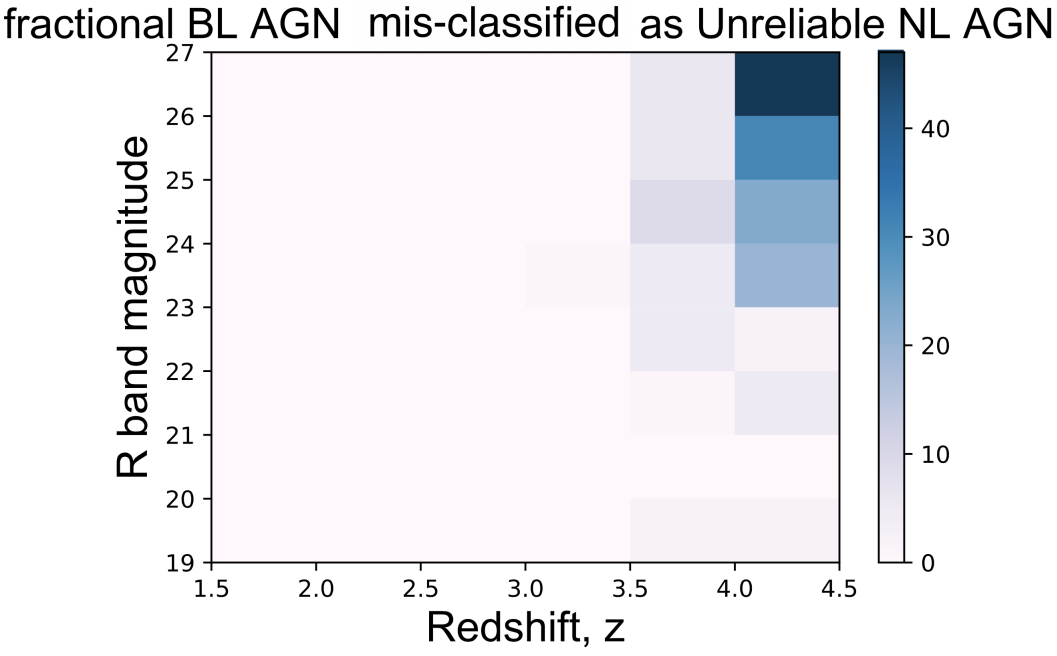

Figure 3.27 *Showing the probability of the spectral fitting code mis-classifying an input BL AGN spectrum as an Unreliable NL AGN, as a function of output redshift and R band magnitude. Probabilities are shown as a percentage.*

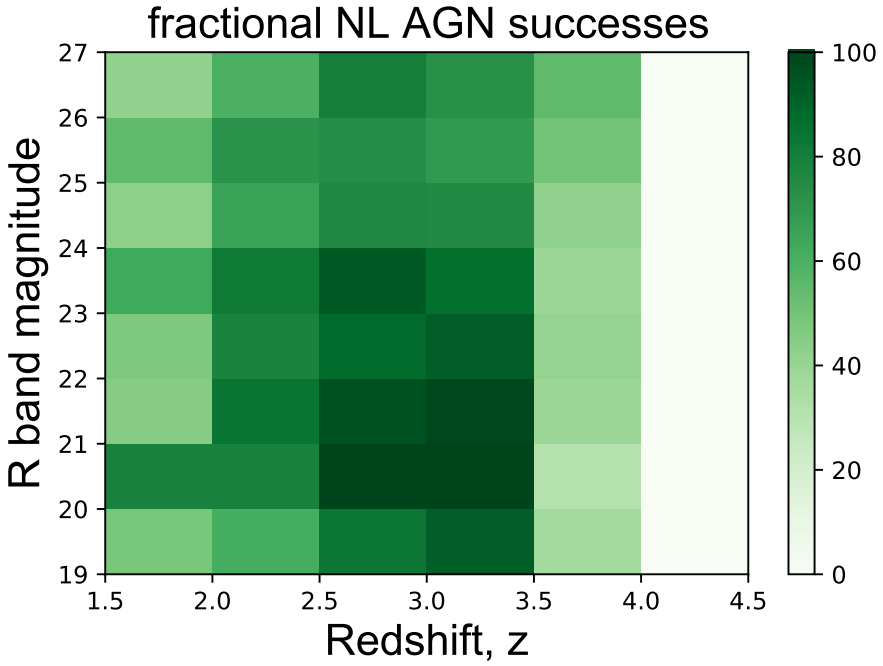

Figure 3.28 *Showing the probability of the spectral fitting code successfully classifying an input NL AGN spectrum as a NL AGN, as a function of output redshift and R band magnitude. Probabilities are shown as a percentage.*

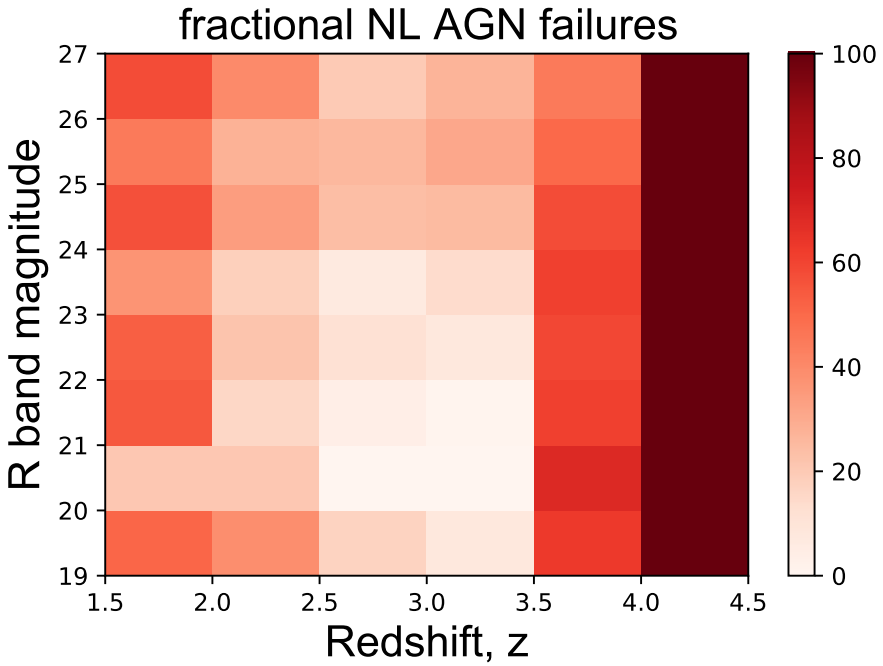

Figure 3.29 *Showing the probability of the spectral fitting code failing to classify an input NL AGN spectrum as a NL AGN, as a function of output redshift and R band magnitude. Probabilities are shown as a percentage.*

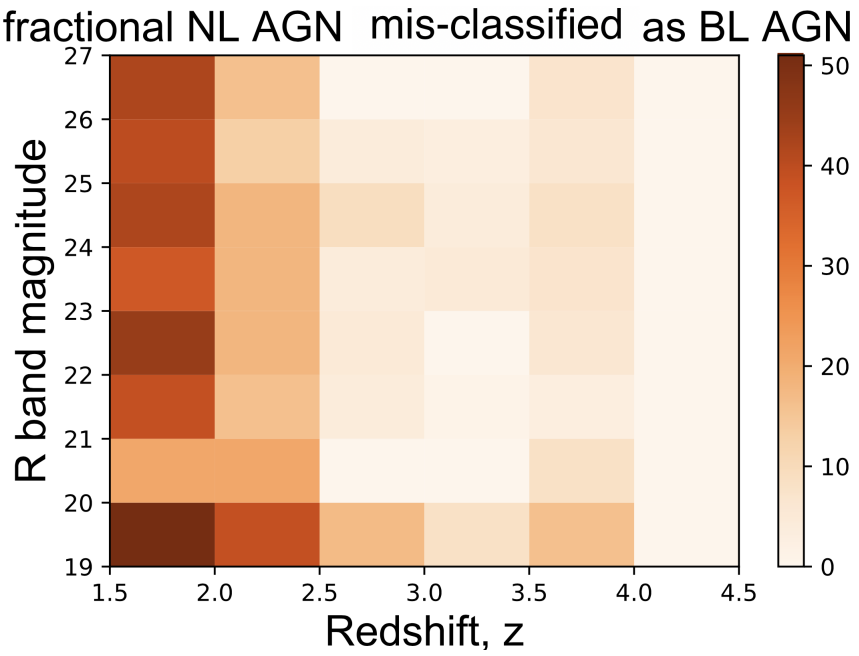

Figure 3.30 *Showing the probability of the spectral fitting code mis-classifying an input NL AGN spectrum as a BL AGN, as a function of output redshift and R band magnitude. Probabilities are shown as a percentage.*

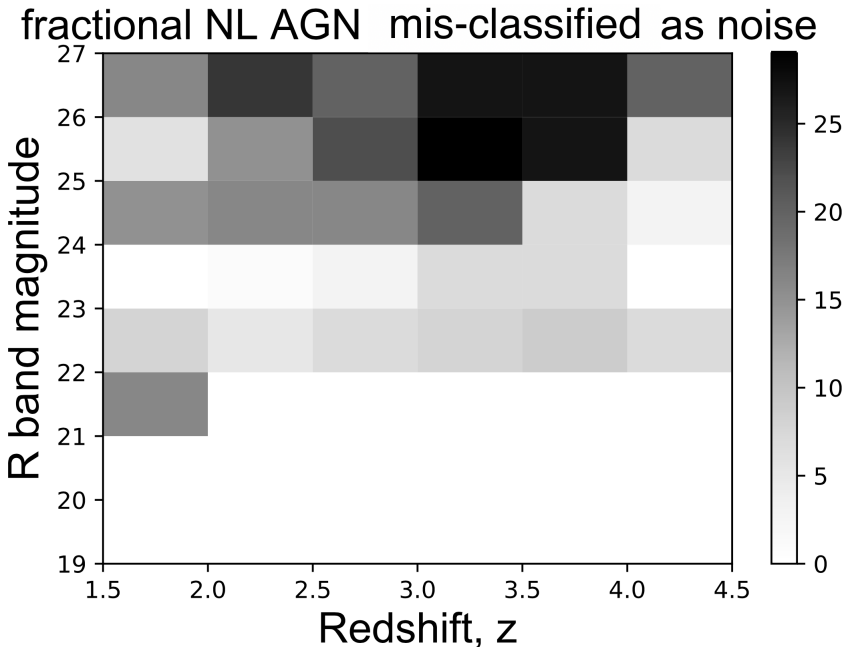

Figure 3.31 *Showing the probability of the spectral fitting code mis-classifying an input NL AGN spectrum as a noise, as a function of output redshift and R band magnitude. Probabilities are shown as a percentage.*

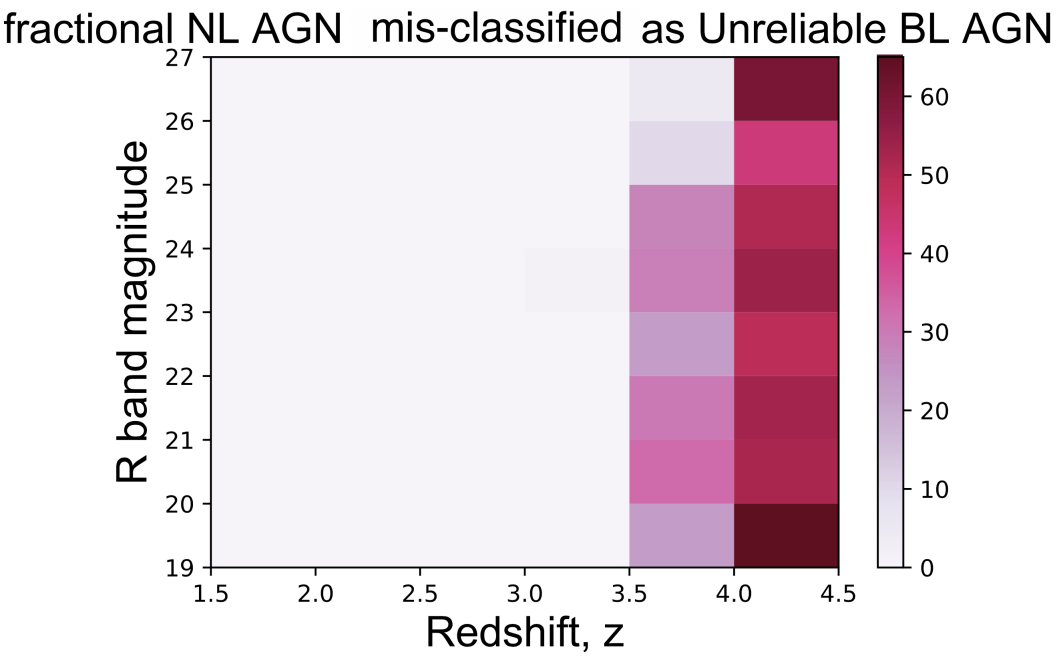

Figure 3.32 *Showing the probability of the spectral fitting code mis-classifying an input NL AGN spectrum as an Unreliable BL AGN, as a function of output redshift and R band magnitude. Probabilities are shown as a percentage.*

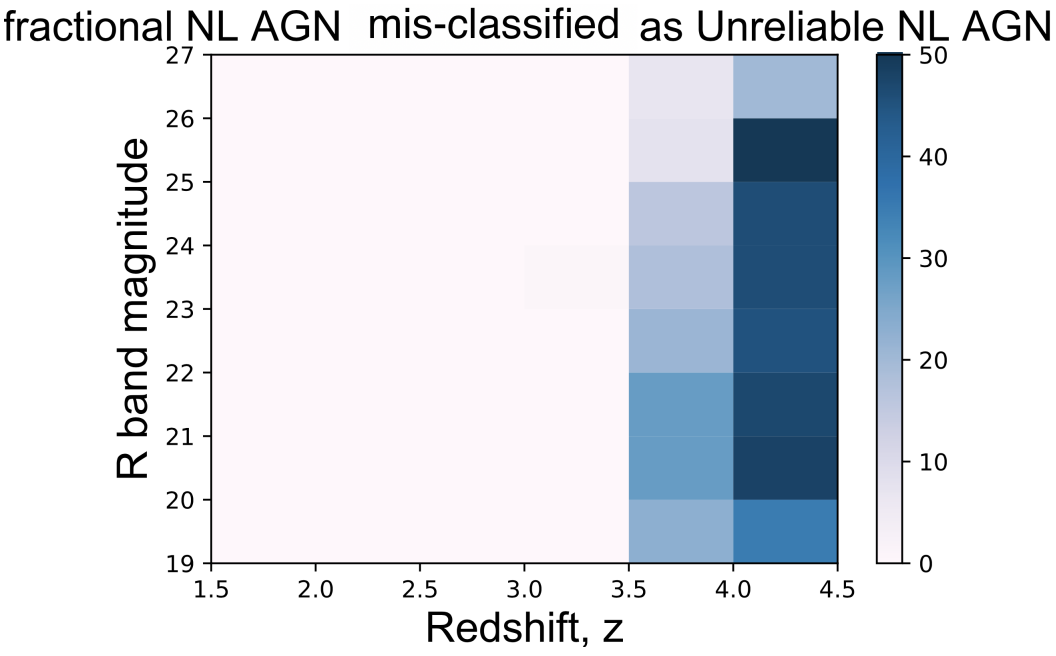

Figure 3.33 *Showing the probability of the spectral fitting code mis-classifying an input NL AGN spectrum as an Unreliable NL AGN, as a function of output redshift and R band magnitude. Probabilities are shown as a percentage.*

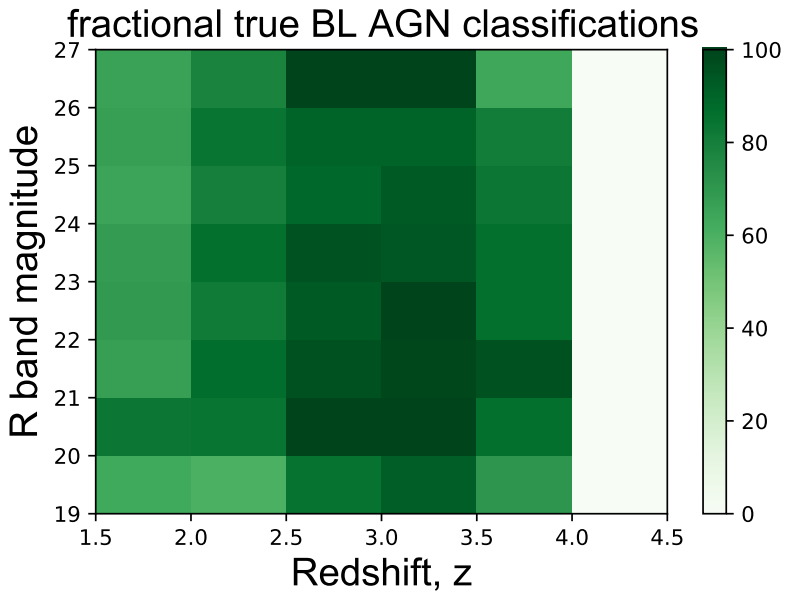

Figure 3.34 *Showing the probability that an output 'BL AGN' classification returned by the spectral fitting code was truly an input BL AGN, as a function of output redshift and R band magnitude. Probabilities are shown as a percentage.*

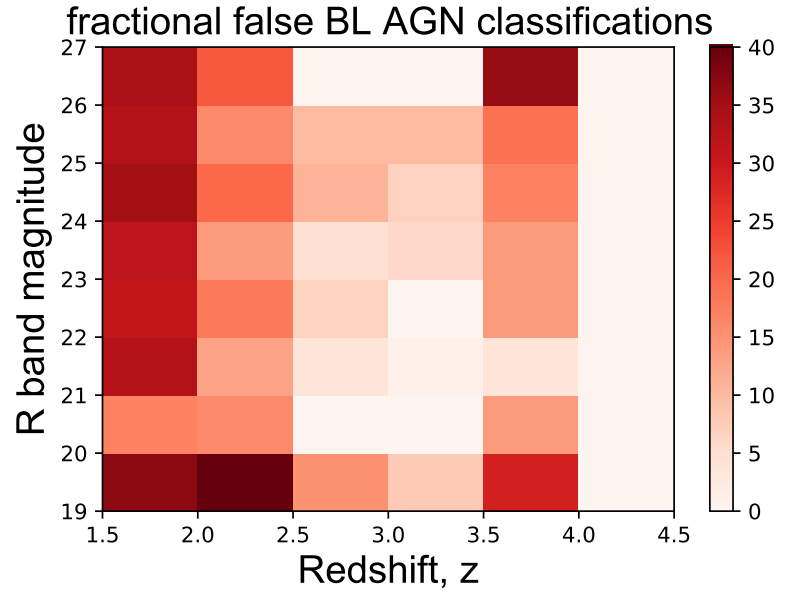

Figure 3.35 *Showing the probability that an output 'BL AGN' classification returned by the spectral fitting code was not truly an input BL AGN, as a function of output redshift and R band magnitude. Probabilities are shown as a percentage.*

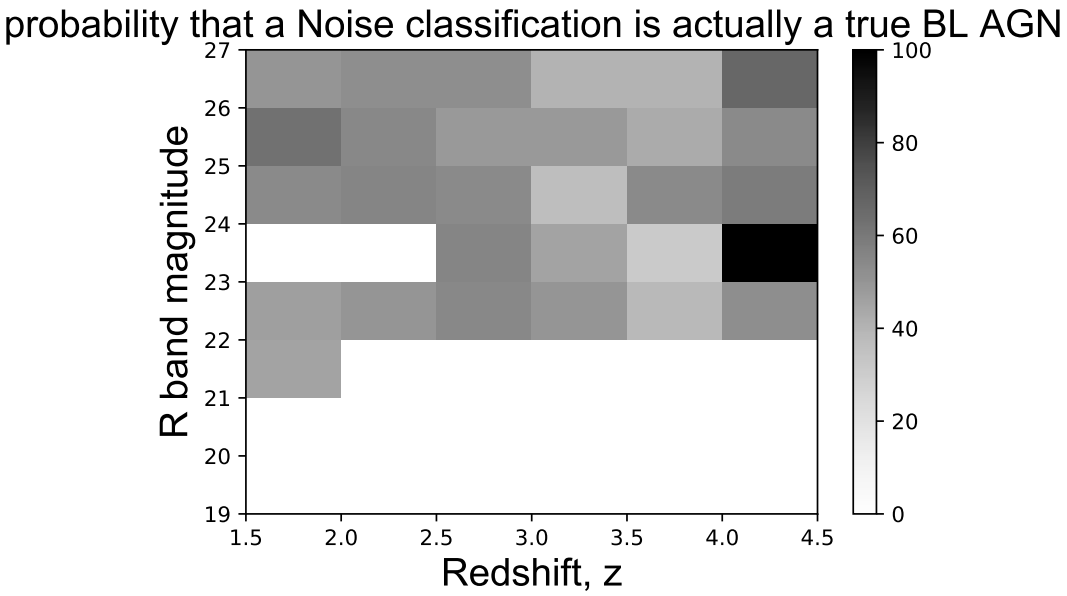

Figure 3.36 *Showing the probability that an output 'Noise' classification returned by the spectral fitting code was truly an input BL AGN, as a function of output redshift and R band magnitude. Probabilities are shown as a percentage.*

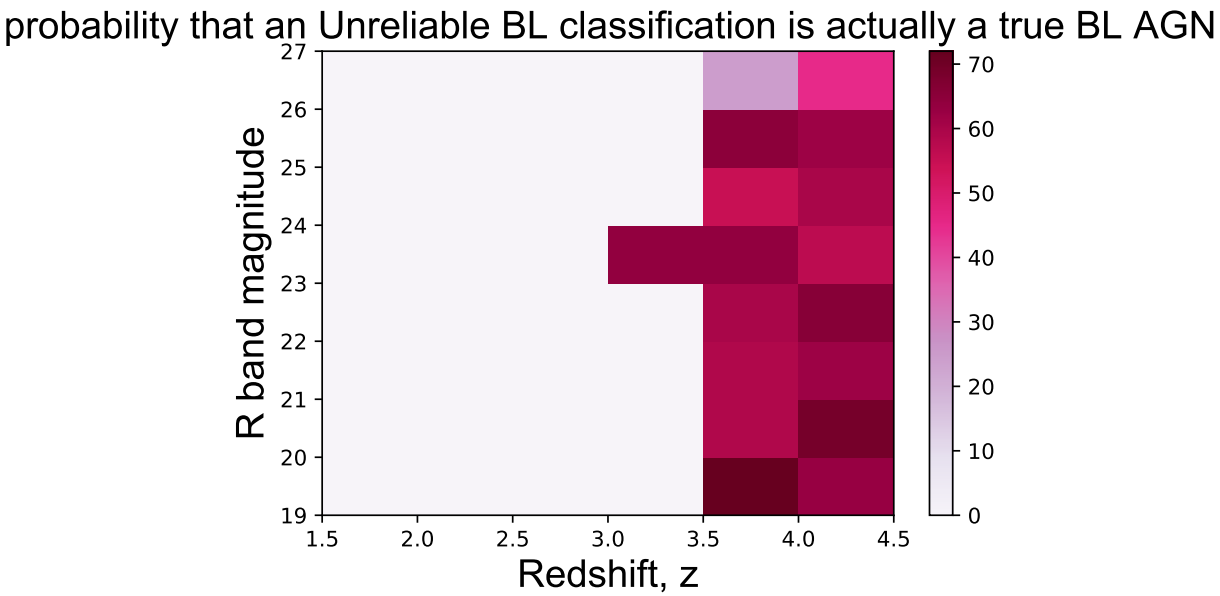

Figure 3.37 *Showing the probability that an output 'Unreliable BL AGN' classification returned by the spectral fitting code was truly an input BL AGN, as a function of output redshift and R band magnitude. Probabilities are shown as a percentage.*

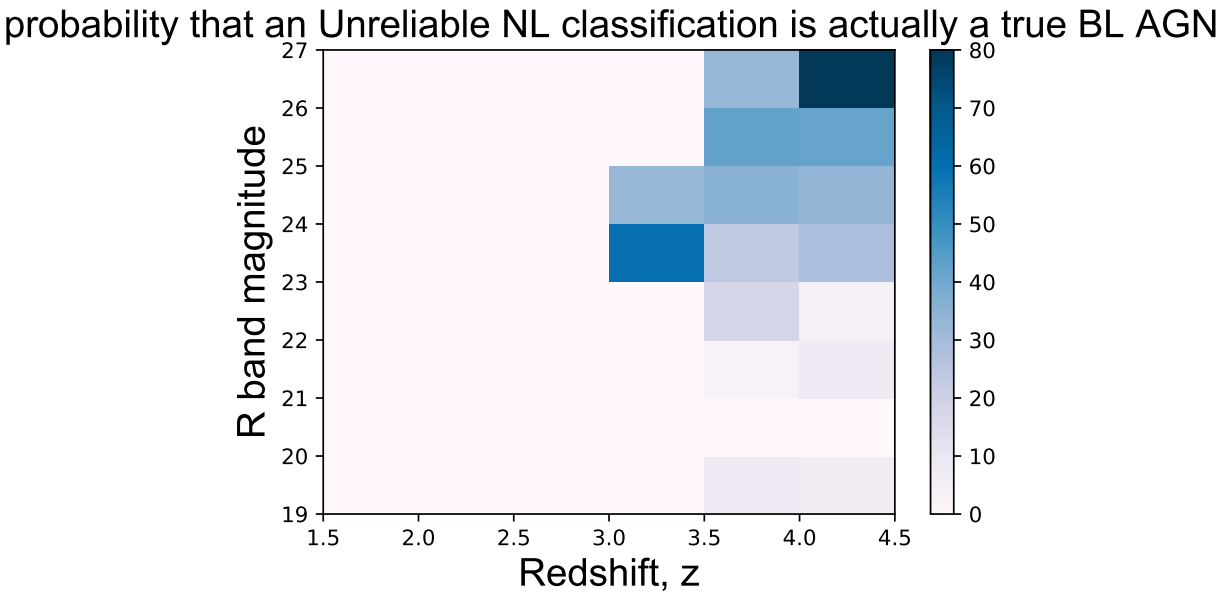

Figure 3.38 *Showing the probability that an output 'Unreliable NL AGN' classification returned by the spectral fitting code was truly an input BL AGN, as a function of output redshift and R band magnitude. Probabilities are shown as a percentage.*

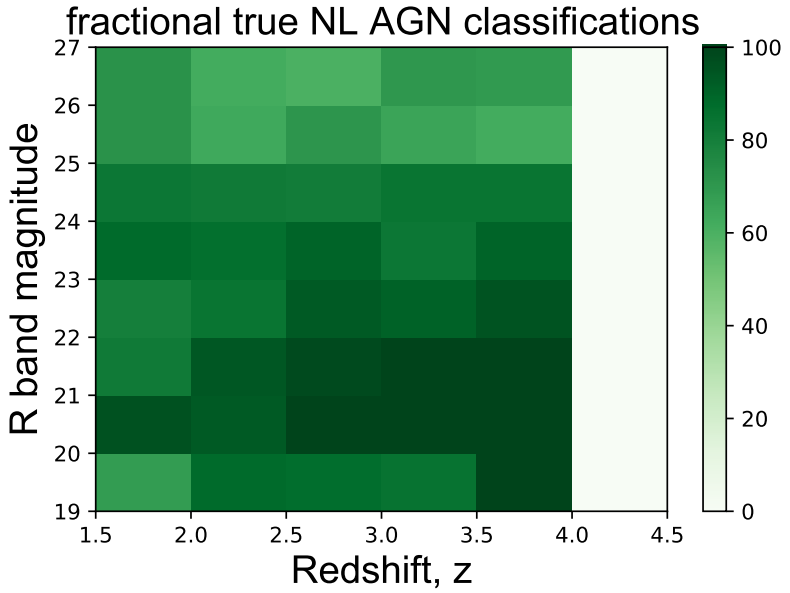

Figure 3.39 *Showing the probability that an output 'NL AGN' classification returned by the spectral fitting code was truly an input NL AGN, as a function of output redshift and R band magnitude. Probabilities are shown as a percentage.*

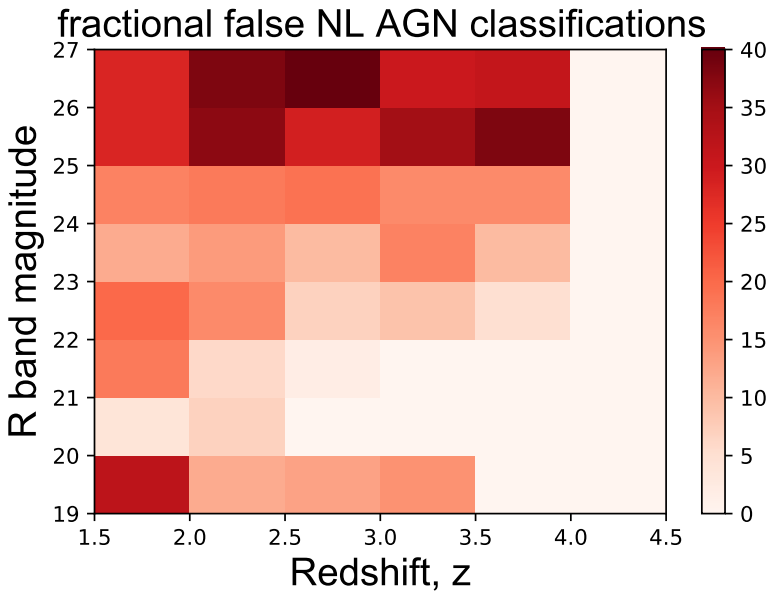

Figure 3.40 *Showing the probability that an output 'NL AGN' classification returned by the spectral fitting code was not truly an input NL AGN, as a function of output redshift and R band magnitude. Probabilities are shown as a percentage.*

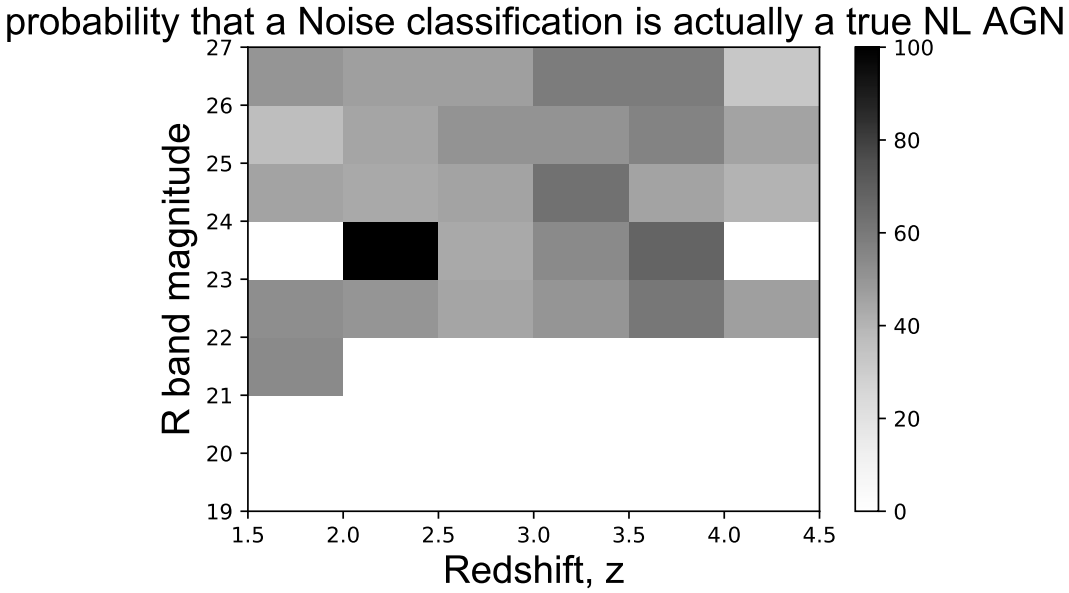

Figure 3.41 *Showing the probability that an output 'Noise' classification returned by the spectral fitting code was truly an input NL AGN, as a function of output redshift and R band magnitude. Probabilities are shown as a percentage.*

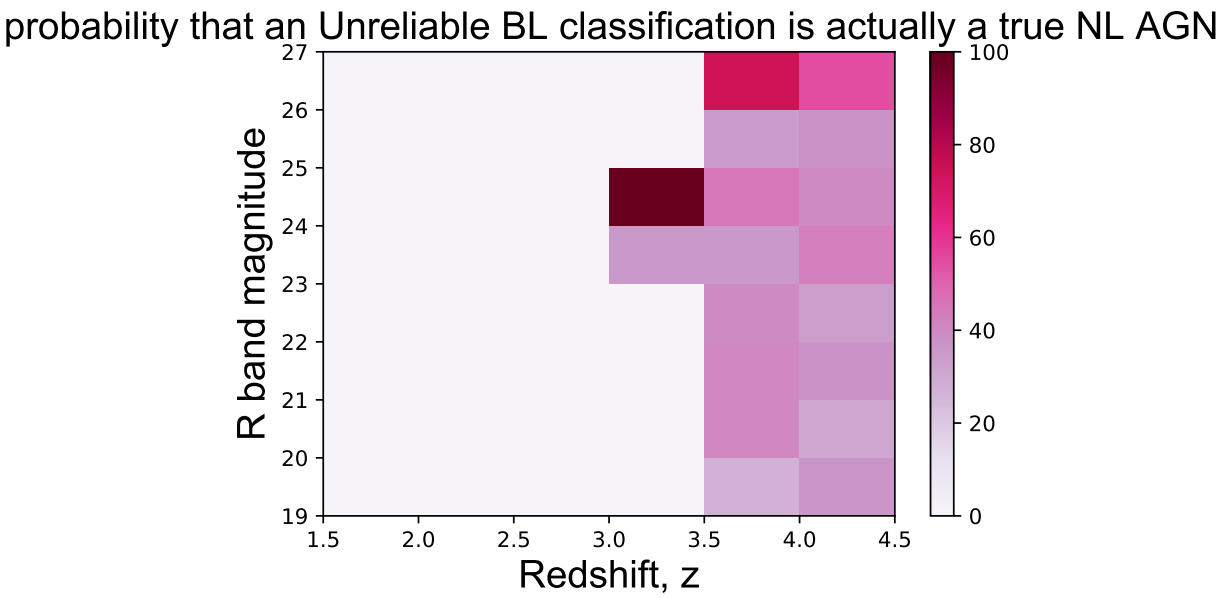

Figure 3.42 *Showing the probability that an output 'Unreliable BL AGN' classification returned by the spectral fitting code was truly an input NL AGN, as a function of output redshift and R band magnitude. Probabilities are shown as a percentage.*

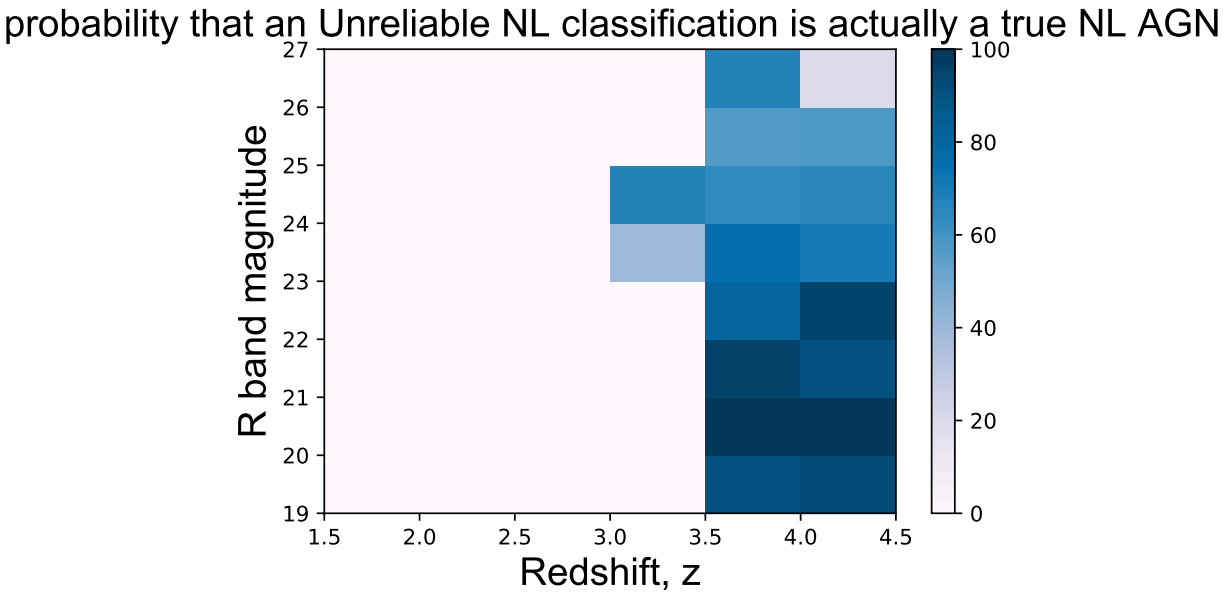

Figure 3.43 *Showing the probability that an output 'Unreliable NL AGN' classification returned by the spectral fitting code was truly an input NL AGN, as a function of output redshift and R band magnitude. Probabilities are shown as a percentage.*

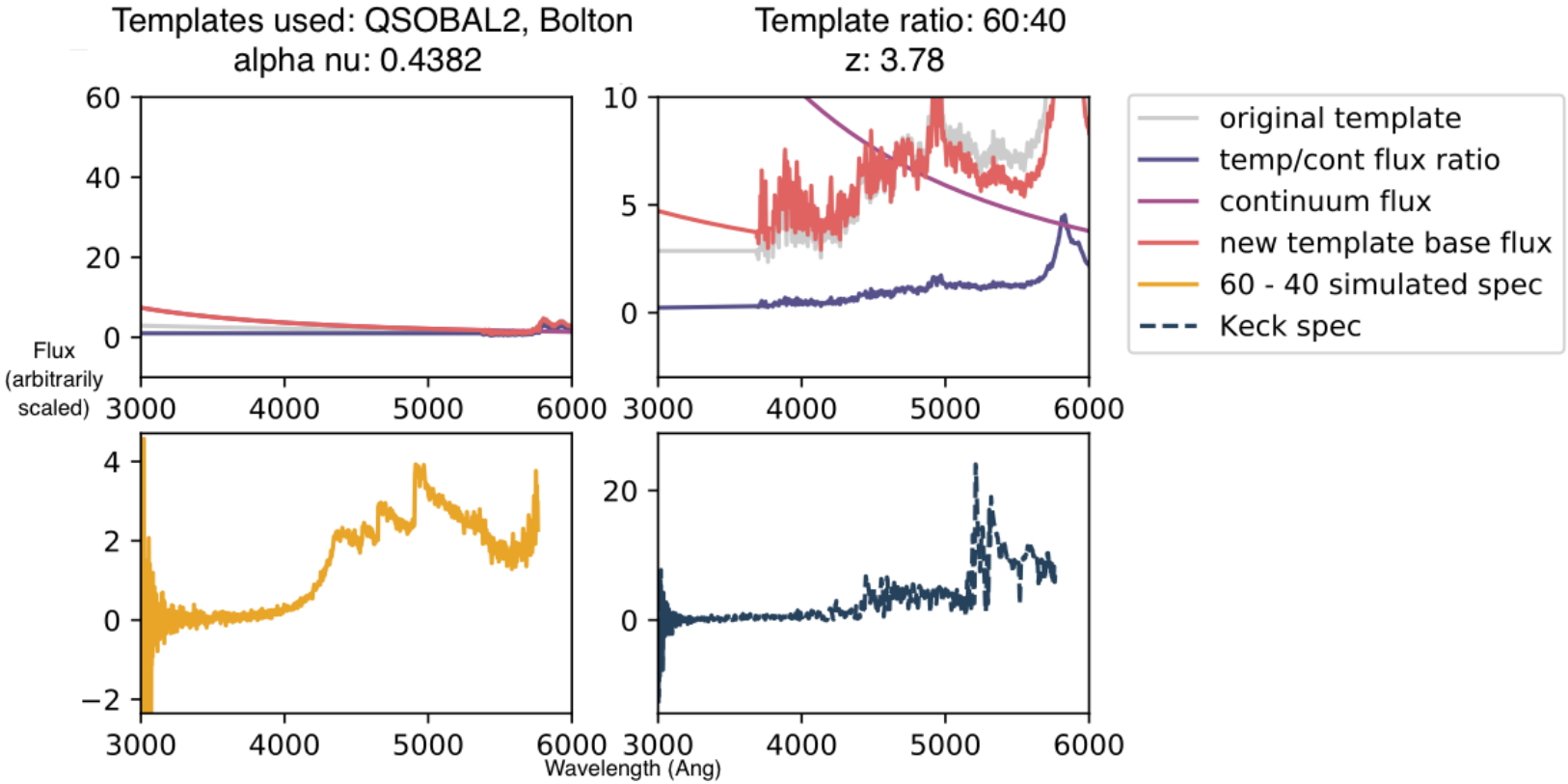

Figure 3.44 *Illustrating the spectral simulation generation process for a failed classification case.* ***Top two panels:*** *show the spectral templates used, in combination, to produce an example simulated spectrum - in this case, the simulation is generated using a 6:4 ratio of the SDSS QSOBAL2 (top left) and the Bolton template (top right). The original template spectra are shown by the solid grey lines and the continuum fits are shown by the solid light purple lines. The continuum-divided templates are shown by the dark purple lines and then the final basis spectral templates with new continuum components folded in (corresponding to the power law denoted in eq.3.4, with* $\alpha_\nu = 0.4382$ *in this example) are shown by the solid pink lines.* ***Bottom two panels:*** *Showing the resultant simulated spectrum (yellow solid line) in comparison to the Keck spectrum (dark blue dashed line) from which this simulation's signal and noise properties were taken.*

Chapter 4

Measuring the AGN X-ray Luminosity Function at high redshifts ($z \sim 2 - 4$)

We have now obtained spectroscopic classifications and redshift estimates for all our sources, and have constrained robust completeness corrections for our X-ray selection, colour pre-selection and spectroscopic classification processes. In the following chapter, I use the results of said analysis to perform new measurements of the AGN XLF at z $\sim$ 2 - 4 for: the independent BL (unobscured) AGN population, the independent NL (obscured) AGN population and the total AGN population. I then go on to calculate a new and robust measurement of the obscured AGN fraction at z $\sim$ 2 - 4, as a function of X-ray Luminosity.

The AGN population is one that has been seen to steadily evolve with luminosity and redshift; as discussed in Chapter 2, this means a more involved parameterisation is required to fit and model AGN XLFs over broad redshift ranges, if we want to gain accurate measurements of SMBH accretion and growth at varying cosmic times.

More is currently known about the very luminous AGN population, with a definite increase in number density seen at higher redshift (see Chapter 1). Unfortunately, the population of *lower* luminosity AGN tends to be overlooked due to a combination of reasons. We need the deepest X-ray data in order to investigate these lower luminosity objects at such high redshifts, leaving us with

a small number of samples from an even smaller number of appropriate fields. Furthermore, identification through spectroscopy can be difficult for these faint X-ray sources, as they also tend to have incredibly faint optical counterparts. This lead to fields becoming spectroscopically incomplete, which rendered established high z XLFs dangerously biased. Instances where this incompleteness has been corrected for has ultimately led to a higher AGN space density found at these higher redshifts (K.Nandra & Steidel, 2005), proving that including the faint AGN population is crucial in these studies.

In this chapter I use the faint, X-ray and colour selected z $\sim$ 2 - 4 AGN sample outlined in Chapter 3, along with the results of the careful modelling of incompleteness detailed in said chapter, to constrain the independent BL and NL AGN XLF's at z $\sim$ 2 - 4.

4.1 Maximum Likelihood fitting approach to determine the completeness-corrected AGN XLF

In this work, I again adopt a maximum likelihood fitting (referred to as MLF from this point onward) approach to generate my unbinned XLF determinations.

In this work, I build on the method presented in Chapter 2 and Aird et al. (2015) and use such an approach to determine our luminosity functions, as using MLF allows me to fold in corrections for multiple selection effects on a source-by-source basis. Furthermore, this approach does not involve any binning in Luminosity, so all information from the individual sources is retained in my calculations.

In the following calculations, I assume independent XLFs for both the broad line and narrow line AGN populations. These XLFs are defined as $\phi_{BL}(L_x \mid \theta_{BL})$ for the broad line population and $\phi_{NL}(L_x \mid \theta_{NL})$ for the obscured population, where θ indicates a distinct set of model parameters for each population. These parameters characterise the double power law XLF space density that I am adopting:

$$\phi(L_X|\theta) = \frac{d\Phi(L_X)}{d\mathrm{logL}_X} = \frac{k_{norm}}{\left(\frac{L_X}{L_*}\right)^{\gamma_1} + \left(\frac{L_X}{L_*}\right)^{\gamma_2}} \tag{4.1}$$

where k_{norm} (normalisation factor), L_* (break luminosity), γ_1 and γ_2 (describing the slope of the XLF at low and high luminosities respectively) are the parameters to be evaluated to find those which satisfy the maximum likelihood estimate.

4.1.1 Defining the likelihood function

Having defined the functional form of the adopted model XLF, the next step is to define the likelihood of obtaining my real data di for a given set of model assumptions (θ):

$$\mathcal{L}(d_i \mid \theta) = e^{-\lambda(\theta)} \prod_{i=1}^{M} l_i(d_i \mid \theta) \tag{4.2}$$

where $\lambda(\theta)$ is the expected number of detected sources, θ are the parameters defining the assumed model XLF (see eq. 4.1) and $l_i(d_i \mid \theta)$ is the contribution of each individual source to the likelihood based on the data, d_i, for that source. This individual contribution to the likelihood of each of our sources can be defined as follows:

$$l_i(d_i \mid \theta) = \phi_i(L_{x,i}, z_i \mid \theta) \tag{4.3}$$

Next, I define the expected number of detectable sources, λ given a model of parameters θ. Over a given redshift range (between z_1 and z_2) and X-ray luminosity range (between $L_{X,1}$ and $L_{X,2}$), the expected number of detectable sources is given by:

$$\lambda(\theta)_1 = \int_{L_{X,1}}^{L_{X,2}} d\mathrm{logL}_X \int_{z_1}^{z_2} dz \left[\phi(L_X, z \mid \theta)\ A(f_X(L_X, z))\ \frac{dV}{dz} \right] \tag{4.4}$$

where $\phi(L_x,\ z \mid \theta)$ is the assumed model luminosity function described by parameters θ; $A(f_x(L_X,\ z))$ is the term accounting for the area of survey sensitive to a source of a given X-ray flux, f_X, which depends on the intrinsic luminosity, L_X, and redshift, z (describing the probability of X-ray source detection, i.e. the survey area curves described in Chapter 3, Figure 3.8); and $\frac{dV}{dz}$ is the differential co-moving volume per unit area as a function of z.

Whilst this expected number does account for the X-ray selection process, it does not yet account for any of the other selection or classification processes I applied to my sample.

To correct for the optical pre-selection criteria applied to my sample, I define the following term:

$$g(\mathrm{L}_X, z, r \mid \mu, \sigma) = p_{sel}(z \mid r)\beta(r \mid \mu(f_X), \sigma) \tag{4.5}$$

which is, in itself, a combination of two terms. This first, $\mathrm{p}_{sel}(\mathrm{z} \mid \mathrm{r})$, is the optical pre-selection correction term at a given r band magnitude, and represents the optical pre-selection functions described in Chapter 3 Figure 3.10. These optical pre-selection functions (and hence this term) account for the reduction in source numbers seen as a result of using optical colour selection criteria, as a function of redshift and r band magnitude.

The second term, $\beta(\mathrm{r})$, is a Gaussian distribution describing the distribution of r-band magnitudes at a given X-ray flux, f_X; effectively accounting for the observed f_X-f_{opt} (X-ray and optical flux) distribution of AGN seen in our sample. This correction allows us to accurately account for incompleteness introduced when approaching optical flux limits. Figure 4.1 below illustrates the need to accurately characterise the underlying distribution of AGN in f_X-f_{opt} space; the green arrow above the enforced r-band magnitude limit highlights the graphical area in which sources will be 'missed', introducing incompleteness. Hence, if we can reasonably constrain the underlying source distribution across optical and X-ray fluxes, incompleteness brought in due to such flux limits can be accounted for.

For a particular source, the term which describes the ratio of its X-ray to Optical flux emission (known in practise as $\log(\mathrm{f}_x/\mathrm{f}_{opt})$ is calculated using the following equation:

$$\log\left(\frac{f_x}{f_{opt}}\right) = \log f_x + 5.4 + \frac{r}{2.5} \tag{4.6}$$

where f_X and r are the X-ray flux and r-band magnitude respectively. The bulk of the AGN population is generally seen to have $\log(\mathrm{f}_x/\mathrm{f}_{opt})$ values ranging between -1.5 and 1.5. To account for this distribution, I adopt the Gaussian $\log(\mathrm{f}_x/\mathrm{f}_{opt})$

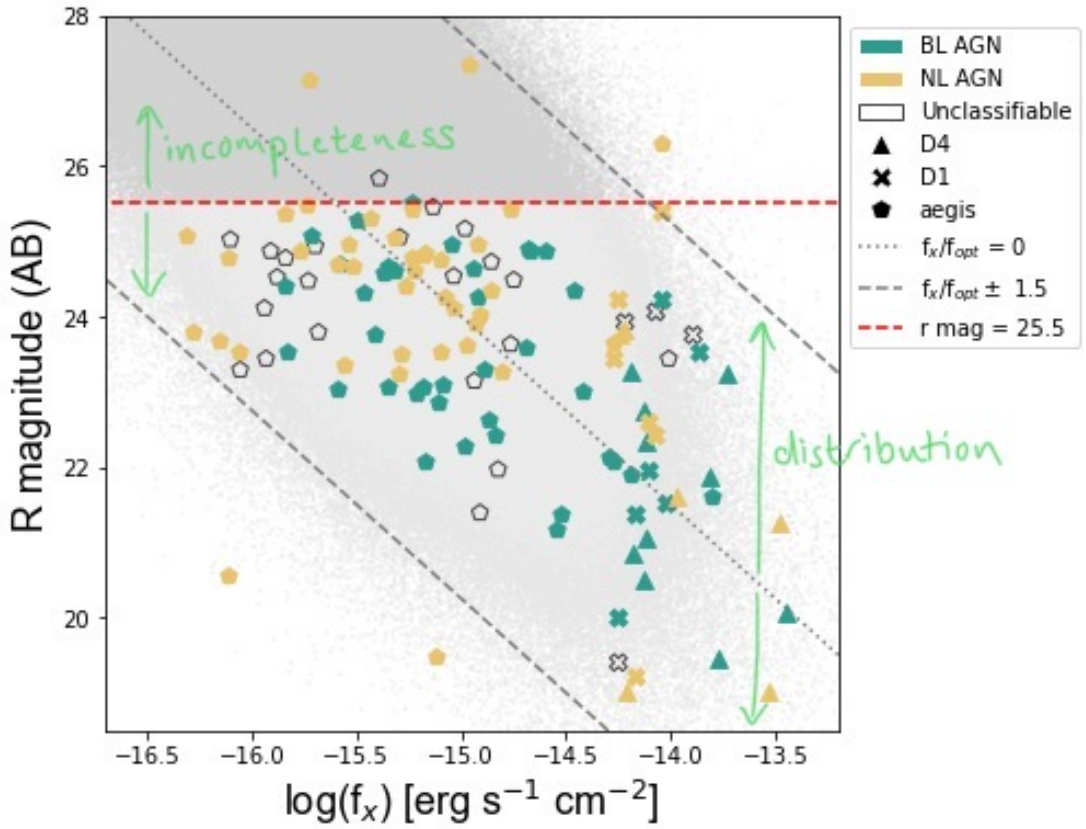

Figure 4.1 *Showing the log(f_x) - r magnitude distribution of our real and simulated data. The grey dotted line is where log(f_x/f_{opt}) = 0, and the grey dashed lines are where log(f_x/f_{opt}) ± 1.5 (AGN are generally seen to have log(f_x/f_{opt}) values within ± 1.5). The red dashed line shows our r magnitude = 25.5 cut. The $\beta(r)$ term in Equation 4.5 accounts for the distribution of sources in this 2D space; it encapsulates the fact that we expect to see more sources within log(f_x/f_{opt}) ± 1.5, and accounts for the incompleteness that flux limits introduce to our sample. Green and yellow markers indicate sources classed as BL and NL AGN respectively, and empty markers indicate 'unclassifiable' sources. Triangles, crosses and pentagons indicate sources from the D4, D1 and aegis (D3) fields respectively.*

distribution derived by Aird et al. (2010), with mean $\mu = 0.24 \pm 0.01$ and standard deviation $\sigma = 0.68 \pm 0.01$. This is shown the solid green line in Figure 4.2 below. Aird et al. (2010) describe how they derive a Gaussian by calculating the observed $\log(f_x/f_{opt})$ distribution of their similarly X-ray and colour selected sample, and then correct for biases due to incomplete optical identifications, as well as the distribution of X-ray fluxes. In their work, these corrections result in a slight shift towards higher $\log(f_x/f_{opt})$ ratios.

To check that using Aird et al. (2010)'s $\log(f_x/f_{opt})$ Gaussian in my analysis is reasonable, I calculated the $\log(f_x/f_{opt})$ ratios for my real sample of AGN, as well as for the sample of simulated AGN sources I created (mentioned in Chapter 3), before and after an r-band magnitude cut of 25.5. By comparing the $\log(f_x/f_{opt})$ distribution of my real and simulated data to the Aird et al. (2010) Gaussian, I can deduce whether the distributions are analogous enough to justify using the Aird et al. (2010) Gaussian.

Figure 4.2 below shows the resultant Aird et al. (2010) $\log(f_x/f_{opt})$ Gaussian as a solid green line, in comparison to histograms showing the distributions of the Aird et al. (2010) data (green dots), my real data (orange), and my simulated data (grey hashed).

The first thing of note when looking at Figure 4.2, is how well the $\log(f_x/f_{opt})$ of our simulated data (yellow histogram) emulates that of of our real data (grey, single hashed histogram). This confirms that our simulations do indeed provide a sizeable and favourable sample of sources, which are truly representative of our real dataset. In addition to this, a shift (similar to that seen in Aird et al. (2010)) towards higher $\log(f_x/f_{opt})$ ratios is seen when comparing the corrected Aird et al. (2010) Gaussian to our real data. This further justifies our use of Aird et al. (2010)'s Gaussian in our analysis.

It is combination of the Aird et al. (2010) $\log(f_x/f_{opt})$ Gaussian distribution term ($\beta(r|\ \mu(f_X), \sigma(f_X))$, a function of f_X and r-band magnitude for a given μ and σ), with the optical pre-selection function term ($p_{sel}(z \mid r)$, a function of z and r-band magnitude) is what produces the resulting $g(L_x, z, r \mid \mu, \sigma)$ correction term shown in Equation 4.5.

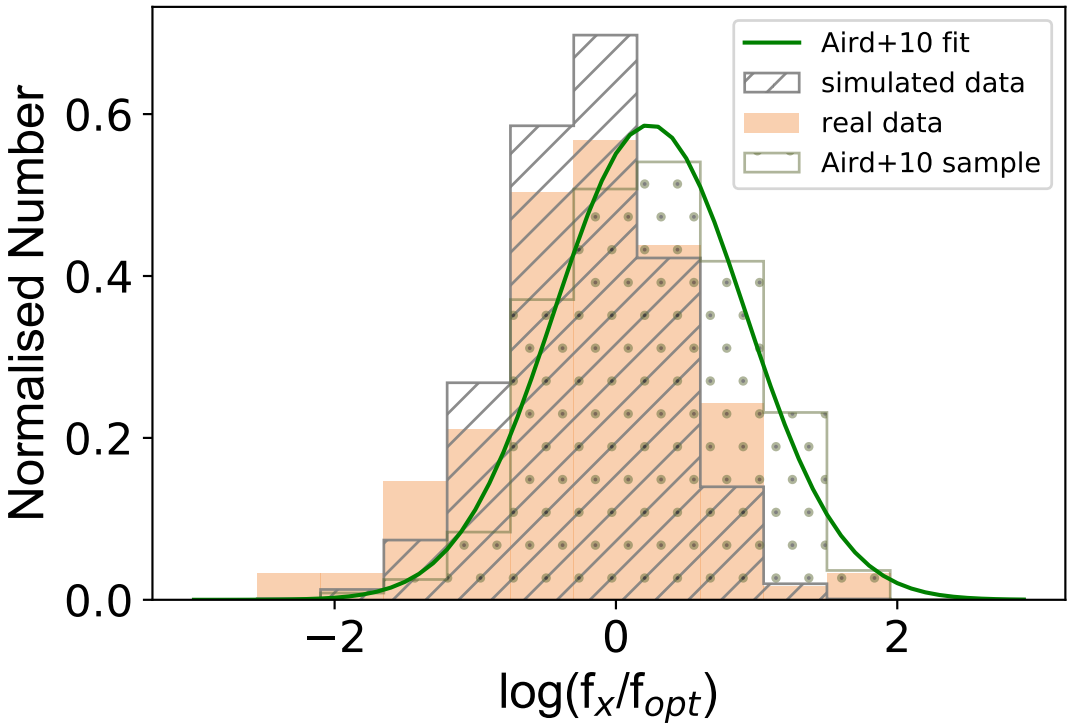

Figure 4.2 *Showing the $\log(f_x/f_{opt})$ distribution of our real data (orange histogram); our simulated data (single hashed, grey histogram histogram); the Aird et al. (2010). data set (green dotted histogram); and theAird et al. (2010), corrected for biases due to incomplete optical identifications and the distribution of X-ray fluxes (solid green line).*

Folding Equation 4.5 into Equation 4.4, we get:

$$\lambda(\theta)_2 = \int_{L_{X,1}}^{L_{X,2}} d\mathrm{logL}_X \int_{z_1}^{z_2} dz \int_{r_1}^{r_2} dr \left[\phi(L_X, z \mid \theta)\ A(f_X(L_X, z))\ \mathrm{g(L}_X, \mathrm{z, r} \mid \mu, \sigma)\ \frac{dV}{dz} \right] \tag{4.7}$$

The final correction that needs to be added to the expected number equation is the spectral classification correction term. This term is taken from the results of my spectral simulations in Chapter 3, the BL or NL (depending on which population I am determining the XLF for, unobscured or obscured) AGN classification success rate of my spectral classification code as a function of redshift and r-band magnitude. This term is essentially the probability of my code correctly classifying a NL or BL AGN, as a function of redshift and r-band magnitude, and is defined as p_{class}(z, r). The final expected number equation, with all correction terms folded in, is subsequently defined as:

$$\lambda(\theta)_3 = \int_{L_{X,1}}^{L_{X,2}} d\mathrm{logL}_x \int_{z_1}^{z_2} dz \int_{r_1}^{r_2} dr \left[\phi(L_X, z \mid \theta)\ A(f_X(L_X, z))\ \mathrm{g(L}_X, \mathrm{z, r} \mid \mu, \sigma)\ p_{class}(z, r) \frac{dV}{dz} \right] \tag{4.8}$$

As each correction term is folded into the expected number equation, it reduces the expected detectable number of BL and NL AGN for a fixed set of XLF parameters, until we reach expected numbers that are more similar to the number of BL and NL AGN we find in our X-ray and colour pre-selected, spectroscopically classified, observed sample; this decrease of expected number corresponds to an increased estimate of the true underlying XLF, in the maximum likelihood fitting process. Figure 4.6 in Section 4.3 illustrates how folding in each of the correction terms decreases the expected delectable number of AGN for set of fixed XLF parameters. The impact of different completeness corrections on the XLFs I derive is discussed in more depth later in Section 4.3.

Having successfully determined the final form of our 'expected number', I fold in one last term to the likelihood equation 4.3; this term defines, given the output classification of that particular source, the probability of the source itself *truly* being that source classification: a weighting term, $\mathrm{w}_i = \mathrm{p}_i(\mathrm{class}_{i,true} \mid \mathrm{class}_{i,returned})$. These probabilities are, once again, taken from the

results of my simulations in Chapter 3 . Being able to constrain the probability of every single one of my sources actually being a NL or BL AGN, no matter what classification my spectral fitting code has given them, as a function of z, r mag and L_x means I can reliably use every source I have in my XLF calculations if I weight them by the $p_i(\text{class}_{i,true} \mid \text{class}_{i,returned})$ term. Folding this term into eq. 4.3, we get:

$$l_i(d_i \mid \theta) = \phi_i(L_{x,i}, z_i \mid \theta)^{w_i} \tag{4.9}$$

which, when substituted into eq. 4.2, gives the final form of the likelihood function:

$$\mathcal{L}(d_i \mid \theta) = e^{-\lambda(\theta)} \prod_{i=1}^{M} \phi_i(L_{x,i}, z_i \mid \theta)^{w_i} \tag{4.10}$$

It is this function which is maximised in an iterative way to find the model parameters which correspond to an XLF that best fits my observed data and accounts for completeness corrections every step of the way. The β(r) term in Equation 4.5, along with the A($f_x(L_X, z)$) term in Equation 4.4, were the only completeness corrections I applied to my XLF analysis in Chapter 2. I have now described how I take even more corrections into account in this higher-z analysis, in an effort to robustly constrain new higher-z XLFs.

4.1.2 Maximum Likelihood Fitting (MLF)

I once again perform an initial test of the likelihood function and its behaviour over the XLF parameter space using the same python function I created and utilised in Chapter 2. In this way, I carry out an initial grid based search to test the likelihood function and its behaviour over the XLF parameter space, and provides us with sensible starting parameters and bounds with which to carry out a more comprehensive MLE process.

Once I had verified that our likelihood function produced sensible values and that it was possible to identify a clear maximum, I again moved on from this initial grid based search to using scipy's optimize.minimize function; this has the capability to

iteratively minimise functions of multiple variables, within user-defined bounds. It is again worth noting that this optimize module *minimises* functions, where we aim to *maximise* the likelihood. Maximising the likelihood is actually equivalent to *minimising the negative log likelihood*; as such, the optimize.minimize function enables us to find the maximum likelihood by computationally minimising the negative log likelihood.

The bounds I chose to use for each parameter in this minimisation process are as follows: $43 \leq \log_{10} L_* \leq 45$, $-10 \leq \log_{10}\gamma_1 \leq 1.5$, $1.5 \leq \log_{10}\gamma_2 \leq 4$, $1e^{-10} \leq \log_k \leq 1e^{-2}$. These bounds were chosen (informed by the aforementioned initial grid search) to encapsulate the typical parameter values found in the literature, whilst also allowing for full exploration of the parameter space. The bounds for γ_1 and γ_2 were also specifically chosen so that they didn't overlap; this was done upon realising, when doing some initial runs with overlapping γ_1 and γ_2 bounds, the minimisation function often got stuck on trying to find solutions that continually flipped between the slopes of γ_1 and γ_2 values.

I again move on to use Markov Chain Monte Carlo (MCMC) methods (specifically 'emcee' once more, Foreman-Mackey et al. (2013), so as to achieve even more thorough XLF estimates, in addition to evaluating their uncertainties. As detailed in Chapter 2 of this thesis, the stopping criterion 'emcee' follows is outlined in Goodman & Weare (2010) and uses the integrated autocorrelation time to quantify the effects of sampling error on results and ensure that all chains are sufficiently converged; the autocorrelation time directly quantifies the Monte Carlo error (thus, sampler efficiency) on integrals computed using MCMC results and, as such, is the logical parameter to use when evaluating the robustness of such an MCMC analysis. (Foreman-Mackey et al., 2013; Goodman & Weare, 2010). I ran 50 - 100 chains for each parameter and, as recommended in the emcee documentation (Foreman-Mackey et al., 2013), assume chains are sufficiently converged when the integrated autocorrelation time is $\geq$ the number of samples/50. I also discard the first 25% of the steps for each parameter as a burn-in phase.

Figure 4.3 shows how the MCMC process I adopt samples the full parameter space across 8000 iterations.

I ran this process for both my BL and NL AGN samples, producing the one and two dimensional projections of the posterior probability distributions of my parameters. As such, I constrained the 'best fitting' XLF parameters and

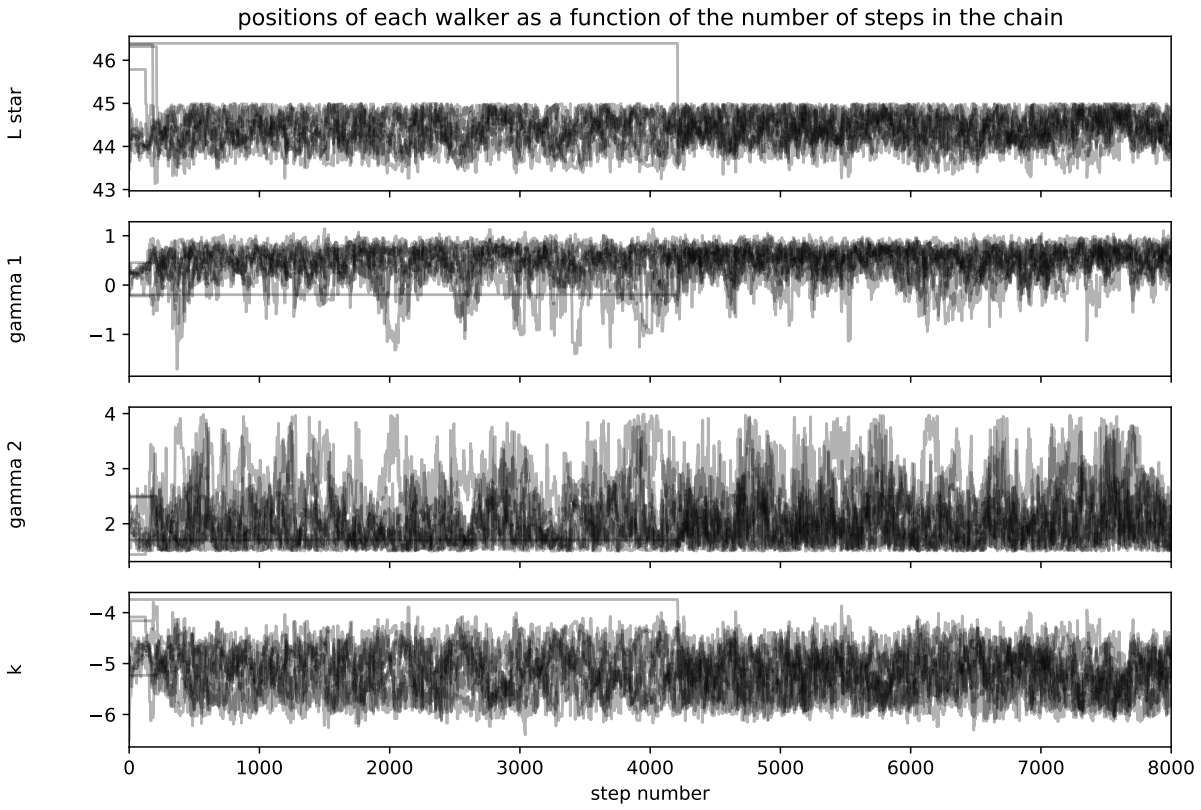

Figure 4.3 *Showing how emcee samples the full parameter for each of my parameters (k, γ_1, γ_2 and L_*, space across thousands of iterative steps in each chain.*

parameter uncertainties for my BL and NL samples. Figures 4.4 (for my BL sample) and 4.5 (for my NL sample) demonstrate all of the covariances between parameters, showing the marginalized distribution for each parameter independently in the histograms along the diagonal and then the marginalized two dimensional distributions in the other panels. The final results from this MLF analysis are presented and discussed in Section 4.3.

4.2 Binned XLF Estimates

Binned XLF estimates give us a direct comparison between data at a given X-ray luminosity and our overall model fit. Binned luminosity functions have previously been calculated using a number of approaches. The $1/\mathrm{V}_{max}$ Schmidt (1968) has proven to be a relatively general and popular technique. This method calculates the cosmological volume over which a given source could be detected given the flux limits of the survey. Unfortunately, this approach is not without its pitfalls and using it as a binned XLF estimator can cause significant biases: it performs poorly when dealing with a small sample size, it can introduce errors when a bin approaches or spans the pre-defined survey flux limits (Page & Carrera, 2000). This is a big problem in particular for X-ray surveys with complicated flux limits. In addition, this method assumes a constant luminosity function within both L_X and z bins.

Taking all of this into account, I use the improved $\mathrm{N}_{obs}/\mathrm{N}_{model}$ estimator (see eq. below) to calculate my binned XLFs. This estimator was derived by Miyaji et al. (2000) and is free from most of the unavoidable biases present in the $1/\mathrm{V}_{max}$ method. The binned $\mathrm{N}_{obs}/\mathrm{N}_{model}$ method involves assuming some best fit model function which describes the behaviour of the XLF within a given redshift range; in this work, I use the XLF results I derived in my Sections 4.1.1 to 4.1.2 as the best fit model. The estimated value for the XLF in a given bin is then:

$$\frac{\mathrm{d}\Phi}{\mathrm{dLog}L_x}(L_{x,i}, z_i) = \frac{d\Phi^{mdl}}{\mathrm{dLog}L_x}(L_{x,i}, z_i) \cdot \frac{N_i^{obs}}{N_i^{mdl}} \tag{4.11}$$

where $\mathrm{L}_{x,i}$ and z_i are the luminosity and redshift in bin i, $\frac{d\Phi^{mdl}}{\mathrm{dLog}L_x}(L_{x,i}, z_i)$ is the best fit analytical expression evaluated at this point, N_i^{obs} is the actual number of AGN observed in bin i, and N_i^{mdl} is the number of AGN predicted by the best fit model in bin i. We calculate N_i^{mdl} using the expected number equation 4.7, but

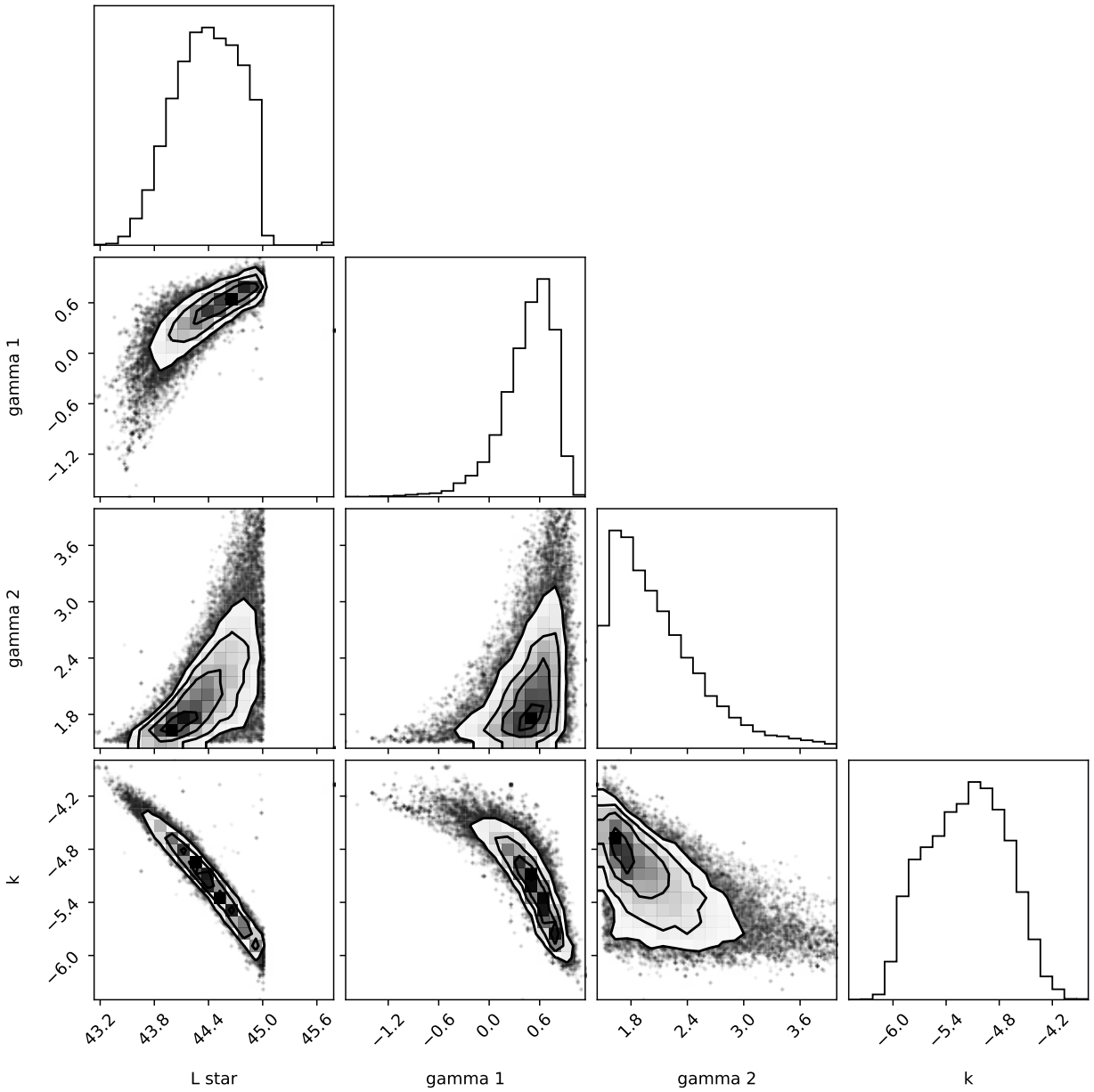

Figure 4.4 *Corner plot demonstrating the covariance between parameters for the BLAGN MCMC MLF process; histograms along the diagonal show the marginalised distribution for each parameter independently, and the other panels show the marginalised 2-D distributions between parameters.*

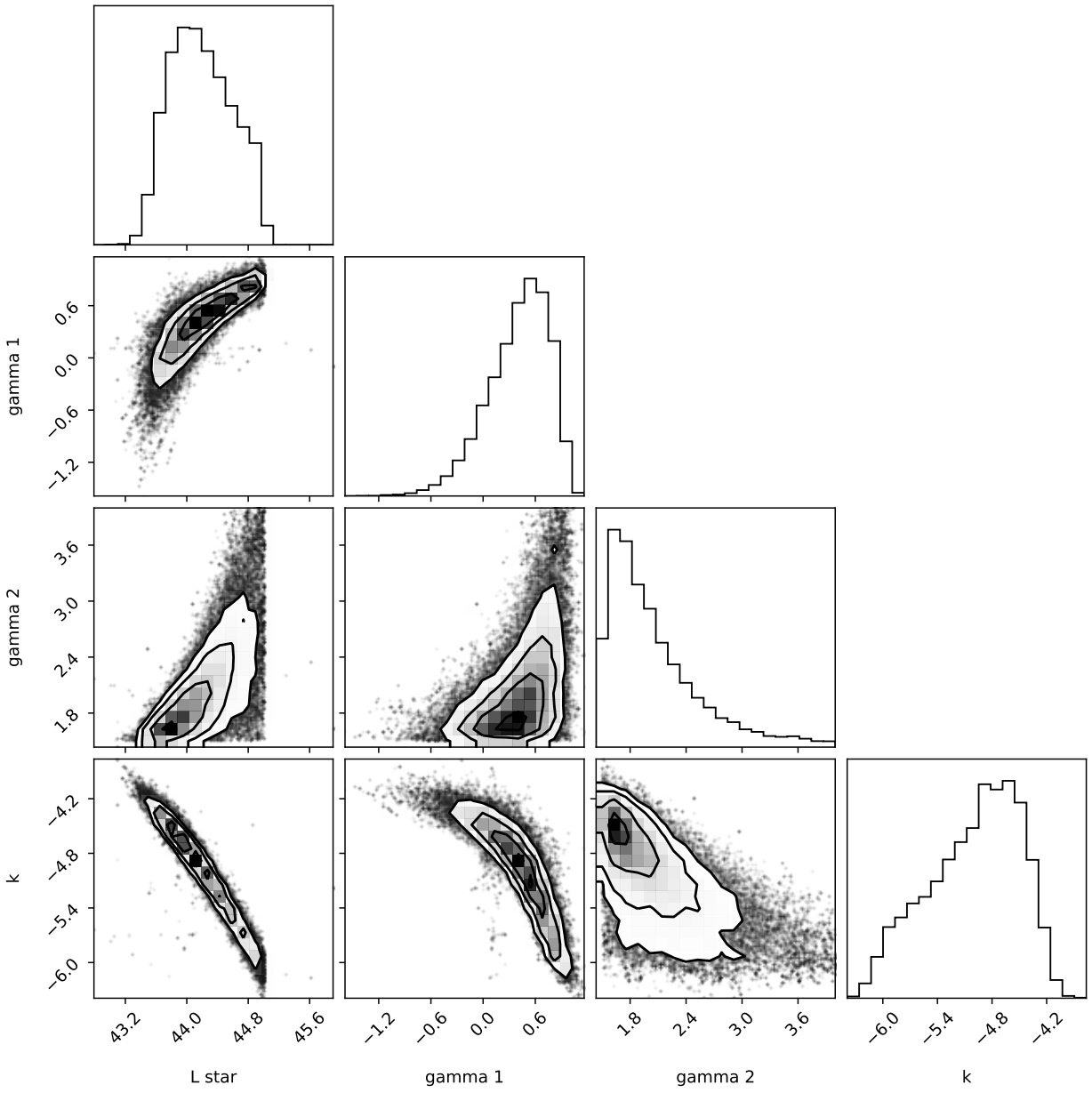

Figure 4.5 *Corner plot demonstrating the covariance between parameters for the NLAGN MCMC MLE process; histograms along the diagonal show the marginalised distribution for each parameter independently, and the other panels show the marginalised 2-D distributions between parameters.*

over limited ranges (i.e. luminosity bins). This means that our values for N_i^{mdl} incorporate all the completeness effects accounted for in Sections 4.1.1 and 4.1.2.

Following Miyaji et al. (2000), I calculate the uncertainty in the N^{obs}/N^{mdl} estimator by using the Poissonian errors given in Gehrels (1986):

$$\begin{aligned} \lambda_u &\approx n + S\sqrt{n + \frac{3}{4}} + \frac{S^2 + 3}{4} \\ \lambda_l &\approx n - S\sqrt{n - \frac{1}{4}} + \frac{S^2 - 1}{4} \end{aligned} \tag{4.12}$$

where, for confidence level of 84% (i.e. 1σ- equivalent uncertainties), the value of S is 1.000 and n is the number of data points in the bin being evaluated. The 1 sigma upper and lower uncertainties are therefore given by:

$$\begin{aligned} \sigma_u &= \frac{\lambda_u - N^{obs}}{N^{mdl}}\ \Phi^{mdl} \\ \sigma_l &= \frac{N^{obs-}\lambda_l}{N^{mdl}}\ \Phi^{mdl} \end{aligned} \tag{4.13}$$

4.3 XLF measurements and exploring the impact of different completeness corrections

The XLFs constrained in this chapter have involved the careful consideration of various correction terms, unlike many other XLF evaluations. It is therefore important to consider how each correction term affects the resulting XLFs. This section explores the impact of the completeness corrections I have folded into my calculations of the XLF, before discussing the final BL and NL z $\sim$ 2 - 4 XLFs.

As discussed in Section 4.1.1, I fold multiple correction terms into expected number term (equation 4.8), as well as a weighting term, to the likelihood equation (equation 4.10). As each correction term is folded into the expected number equation, it reduces the expected detectable number of BL and NL AGN for a fixed set of XLF parameters, until we reach expected numbers that are more similar to the number of BL and NL AGN we find in our X-ray and colour

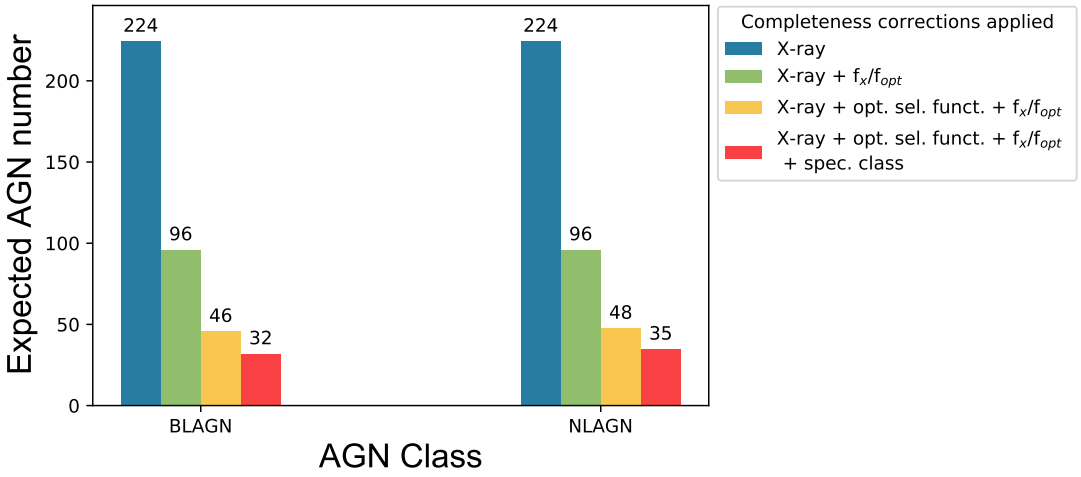

Figure 4.6 *Showing how the expected detectable number of BL and NL AGN decreases as each correction term is folded in. The colour scheme is as follows: blue is when purely X-ray selection is accounted for; green is when X-ray section and the f_x/f_{opt} distribution is accounted for; yellow is when X-ray selection, optical pre-selection and the f_x/f_{opt} distribution is accounted for; red is when X-ray selection, optical pre-selection, the f_x/f_{opt} distribution and spectroscopic classification corrections are accounted for. The expected detectable number is clearly seen to decrease as each correction term is folded in, until it reaches values similar to that observed in my real sample: 32 expected BL AGN and 35 expected NL AGN, totalling 67 X-ray selected, optically pre-selected AGN successfully spectroscopically classified as either BL or NL (unobscured or obscured).*

pre-selected, spectroscopically classified, observed sample. Figure 4.6 illustrates how folding in each of the correction terms decreases the expected detectable number of AGN for set of fixed XLF parameters. The colour scheme is as such: blue bars illustrate the expected number when X-ray selection completeness corrections are applied (i.e. Eq.4.4); green bars illustrate the expected number when X-ray selection and f_X/f_{opt} completeness corrections are applied; yellow bars illustrate the expected number when X-ray selection, f_X/f_{opt} and optical selection completeness corrections are applied (i.e. Eq. 4.7); red bars illustrate the expected number when X-ray selection, f_X/f_{opt}, optical selection and spectroscopic classification completeness corrections are applied (i.e. Eq. 4.8).

As each completeness correction is folded into the expected number calculation, the number of AGN one expects to detect decreases for both the BL and NL (unobscured and obscured) populations, with the final expected numbers being more representative of the actual number of BL and NL AGN we see in our real data. This decrease of expected number corresponds to an increased estimate of the true underlying XLF, in the maximum likelihood fitting process; figures

4.7 and 4.8 show how the BL and NL XLFs evolve as each correction term is folded into the expected number calculation. These XLFs were derived only using scipy's optimize.minimize module (not using emcee) as they aim to represent how, as each completeness correction is accounted for, it broadly changes the overall shape of the XLF. The colour scheme is analogous to that in Figure 4.6 and shows how each correction term affects the MLF XLF derived; as each correction term is folded in, the AGN number density increases in varying ways due to each of the terms accounting for the varying selection and classification effects. For both the BL and NL AGN populations, when only X-ray selection completeness corrections are applied (the blue line), the faint end slope of the XLF is quite shallow (i.e. less faint AGN sources are accounted for). Then, when the f_X/f_{opt} completeness correction is folded in (the green line), we see an increase in the number density of AGN at fainter luminosities as the underlying f_X/f_{opt} distribution is accounted for, correcting for the sources missed when approaching optical flux limits. This effect is seen again when the optical selection completeness corrections are folded in (the yellow line), as sources that were missed out when applying the optical selection criteria are effectively accounted for. Finally, the AGN number density increases once more as the spectroscopic classification completeness corrections are folded in (the red line), accounting for incorrect source classifications and the classification weightings. It is interesting to note that, when this final correction is applied, the BL AGN number density at brighter luminosities actually decreases, with more of an increase seen in the NL AGN XLFs. This is due to the spectroscopic classification correction accounting for the miss-classification of bright AGN as BL. It is worth noting that these example XLFs were created to show the importance and effectiveness of applying correction terms when carrying out MLF AGN XLF analysis. As previously stated, these XLFs were derived from example parameters, only using scipy's optimize.minimize module (not using emcee) as their aim is to represent how, as each completeness correction is accounted for, it broadly changes the overall shape of the XLF. The parameters of these XLFs are not our final results and, as such, may appear to disagree with those in the literature. However, the intention of the figure still rings true: applying careful corrections to this process changes the shape of the XLF and, not applying adequate corrections may lead to inaccurate estimates of the various XLF parameters (k, γ_1, γ_2 and L_*).

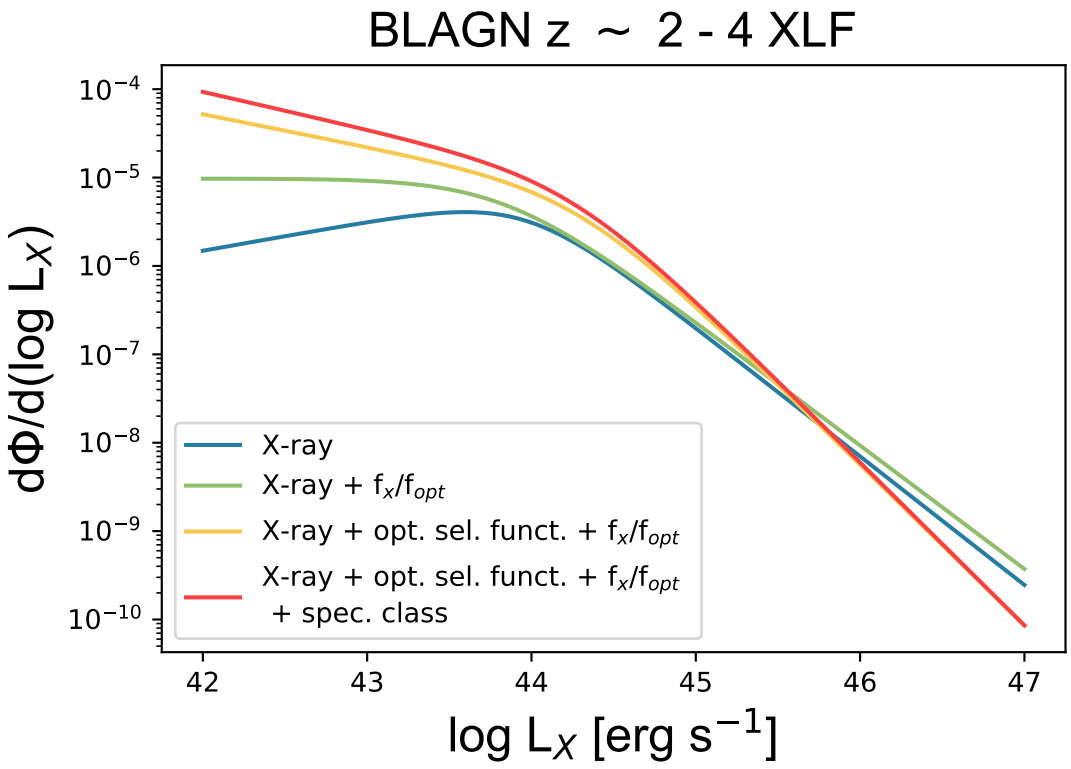

Figure 4.7 *Showing how the $z \sim$ 2- 4 BLAGN XLF changes as differing correction terms are folded in. The colour scheme is analogous to that in Fig. 4.6, and is as follows: blue is when purely X-ray selection is accounted for; green is when X-ray section and the f_x/f_{opt} distribution is accounted for; yellow is when X-ray selection, optical pre-selection and the f_x/f_{opt} distribution is accounted for; red is when X-ray selection, optical pre-selection, the f_x/f_{opt} distribution and spectroscopic classification corrections are accounted for. As each correction is folded in, the over AGN number density increases as more sources are accounted for, especially those which lie at fainter luminosities.*

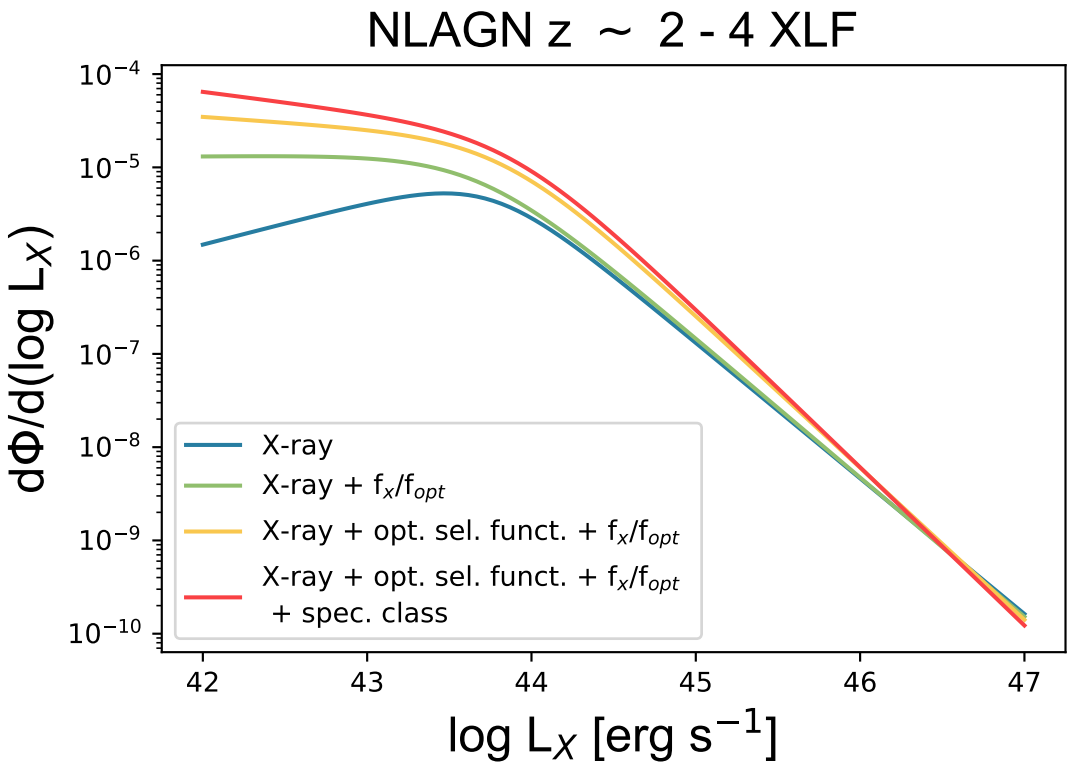

Figure 4.8 *Showing how the $z \sim$ 2- 4 NLAGN XLF changes as differing correction terms are folded in. The colour scheme is analogous to that in Fig. 4.6, and is as follows: blue is when purely X-ray selection is accounted for; green is when X-ray section and the f_x/f_{opt} distribution is accounted for; yellow is when X-ray selection, optical pre-selection and the f_x/f_{opt} distribution is accounted for; red is when X-ray selection, optical pre-selection, the f_x/f_{opt} distribution and spectroscopic classification corrections are accounted for. As each correction is folded in, the over AGN number density increases as more sources are accounted for, especially those which lie at fainter luminosities.*

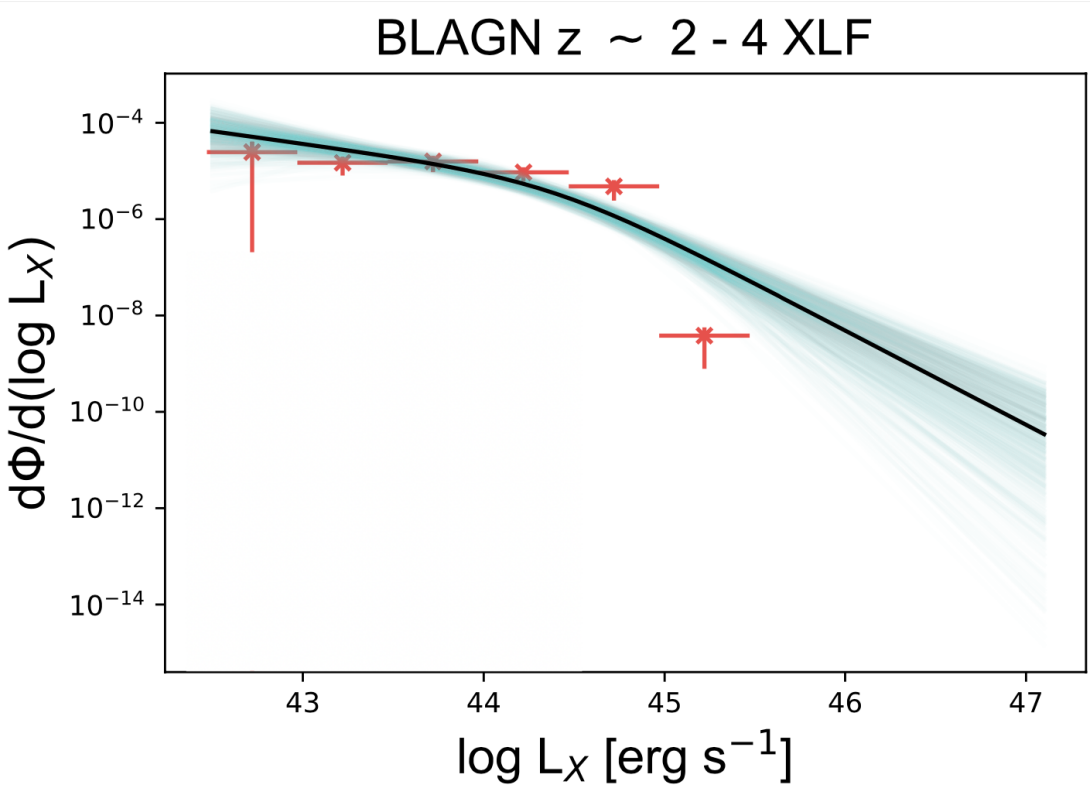

Figure 4.9 *$z \sim 2$ - 4 BL (unobscured) AGN XLF derived using MCMC and MLE. The blue region around and the parameter uncertainties quoted are taken the 16^{th} and 84^{th} percentiles of the samples in the marginalized distributions. Binned N_{obs}/N_{mod} XLF estimates are over-plotted in red.*

4.4 Final XLF measurements

Figures 4.9 and 4.10 below show my final, independent BL and NL AGN XLFs at z $\sim$ 2-4, with 1σ equivalent upper and lower confidence intervals based on MCMC maximum likelihood fitting. The parameter uncertainties quoted on the plots are taken on the 16^{th} and 84^{th} percentiles of the samples in the marginalized distributions. Binned N_{obs}/N_{mod} XLF estimates are over-plotted in red.

Figure 4.11 is a comparison plot my z $\sim$ 2 - 4 BL (blue solid line) and NL (purple dashed line) AGN XLFs. Upon comparing the two, there is not a significant difference between the NL (obscured) and BL (unobscured) AGN XLFs at z $\sim$ 2 - 4 within the 1σ uncertainties, suggesting that the number densities of obscured and unobscured AGN at z$\sim$2 - 4 are relatively equal. The sum of the BL and NL AGN XLF is overplotted as the total AGN XLF (grey dot-dashed line).

An interesting comparison to make is to newer X-ray data, namely the eROSITA

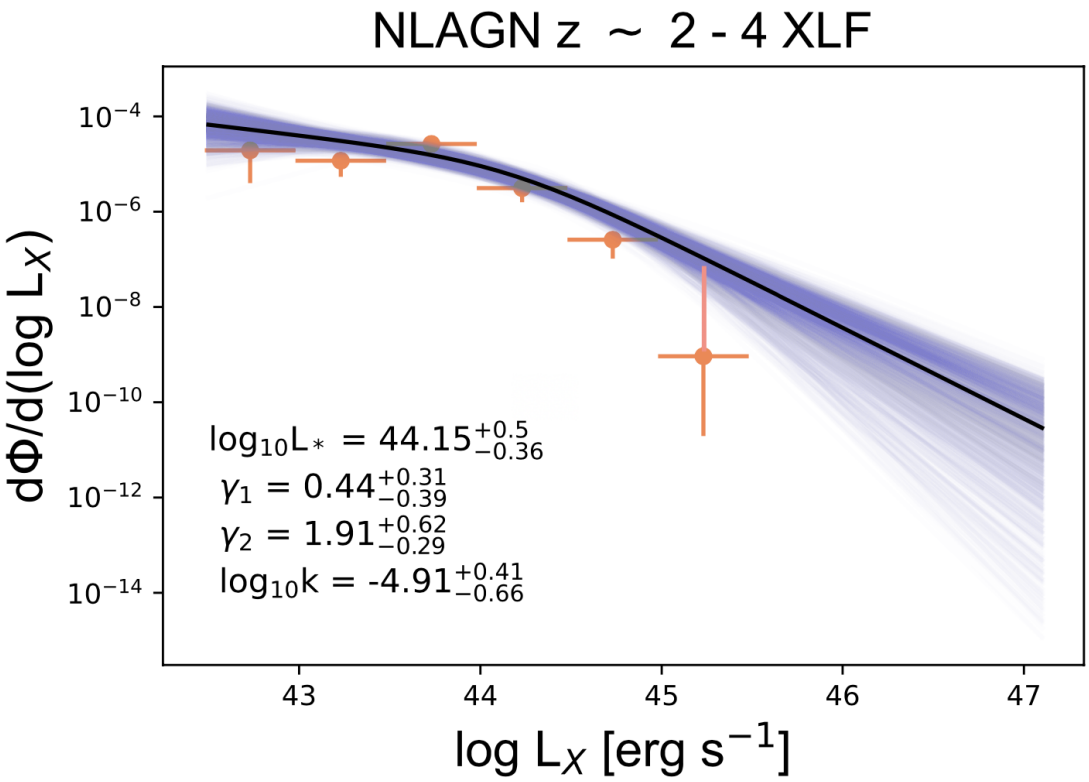

Figure 4.10 *$z \sim 2$ - 4 NL (unobscured) AGN XLF derived using MCMC and MLE. The purple region around and the parameter uncertainties quoted are taken the 16^{th} and 84^{th} percentiles of the samples in the marginalized distributions. Binned N_{obs}/N_{mod} XLF estimates are over-plotted in red.*

Final Equatorial Depth Survey, eFEDS. My work predicts $\approx 2 \times 10^3$ AGN, given my survey area of ≈ 2 deg^2 down to a flux limit of $\approx 10^{-14}$ erg s^{-1} cm^{-2}. eFEDS covers 140 deg^2 down to a similar flux limit; given this, my work predicts 1.4×10^5 AGN in eFEDS. This is an order of magnitude larger than the 22 079 AGN published in the the eFEDS AGN catalog (Liu, Teng et al., 2022). However, this is an overestimate as my survey area discussed in Chapter 3 is sensitive to fluxes $\approx$ half an order of magnitude fainter than eFEDS. In addition, the whole point of my work highlights how applying careful completeness corrections to observational data can greatly change or, in this case, increase the number of sources predicted at various redshifts and luminosities; hence, an overestimate.

Figures 4.12 and 4.13 compare the XLFs derived in this work, with those previously constrained at z$\sim$2 - 4.

Figure 4.12 compares the BL AGN XLF calculated in this work (in black) with the z = 3 (in pink) and z = 2 (in blue) unabsorbed AGN XLFs from Aird et al. (2015). Our work predicts higher BL AGN number densities at fainter and brighter luminosities than both the z = 3 and z = 2 Aird et al. (2015) XLFs. However, the z = 3 Aird et al. (2015) XLF predicts a higher L_*, which corresponds to a higher BL AGN number density in the mid-luminosity range.

Figure 4.13 compares the NL AGN XLF calculated in this work (in black) with the z = 3 (in pink) and z = 2 (in blue) absorbed AGN XLFs from Aird et al. (2015). Our work predicts higher NL AGN number densities which lie between those showcased by the z = 3 and z = 2 Aird et al. (2015) XLFs at fainter luminosities. Our work also predicts a lower L_* than both the z = 3 and z = 2 Aird et al. (2015), corresponding to a lower NL AGN number density in the mid-luminosity range.

4.5 The obscured AGN fraction at z $\sim$ 2 - 4

Another interesting way to look at these results is by looking at the fraction of obscured AGN as a function of luminosity; Figure 4.14 illustrates the obscured AGN fraction as a function of luminosity calculated in this work (black points) in comparison to previous estimates: Treister & Urry (2006)'s z $\sim$ 2.5-3 obscured fraction (not calculated as a function of luminosity) in orange; Buchner et al. (2015)'s z = 2.7-4 obscured fraction in blue; Aird et al. (2015)'s z = 3 obscured

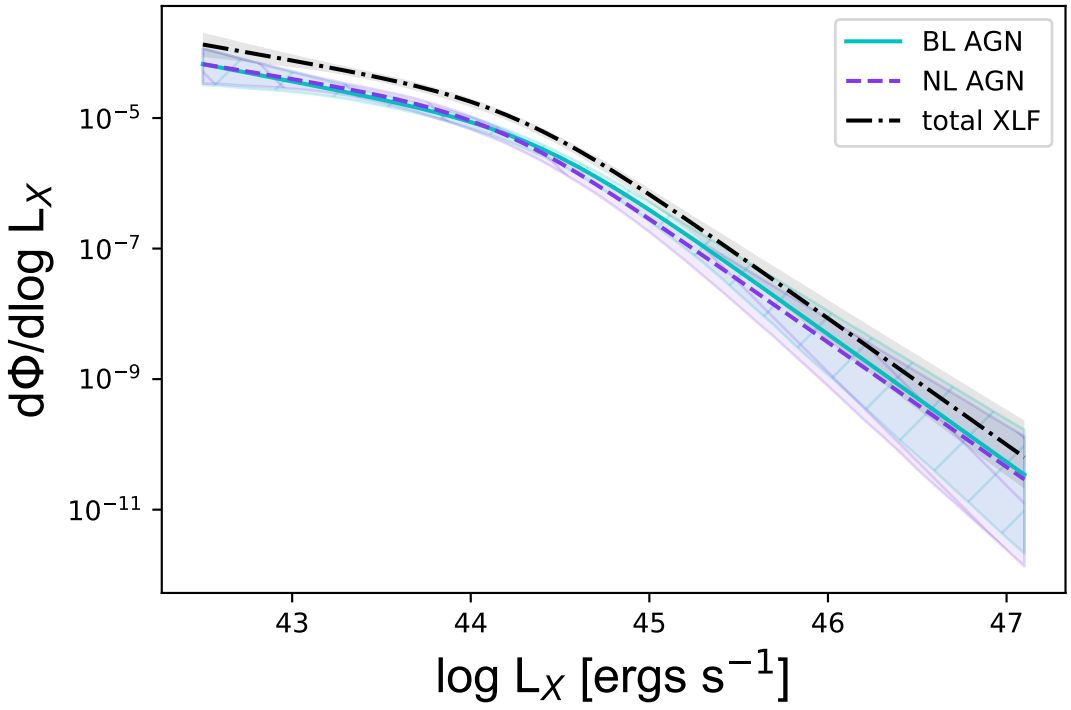

Figure 4.11 *Comparing the BL (light blue), NL (purple) and total (black) AGN XLFs at z∼ 2 - 4, derived in this work.*

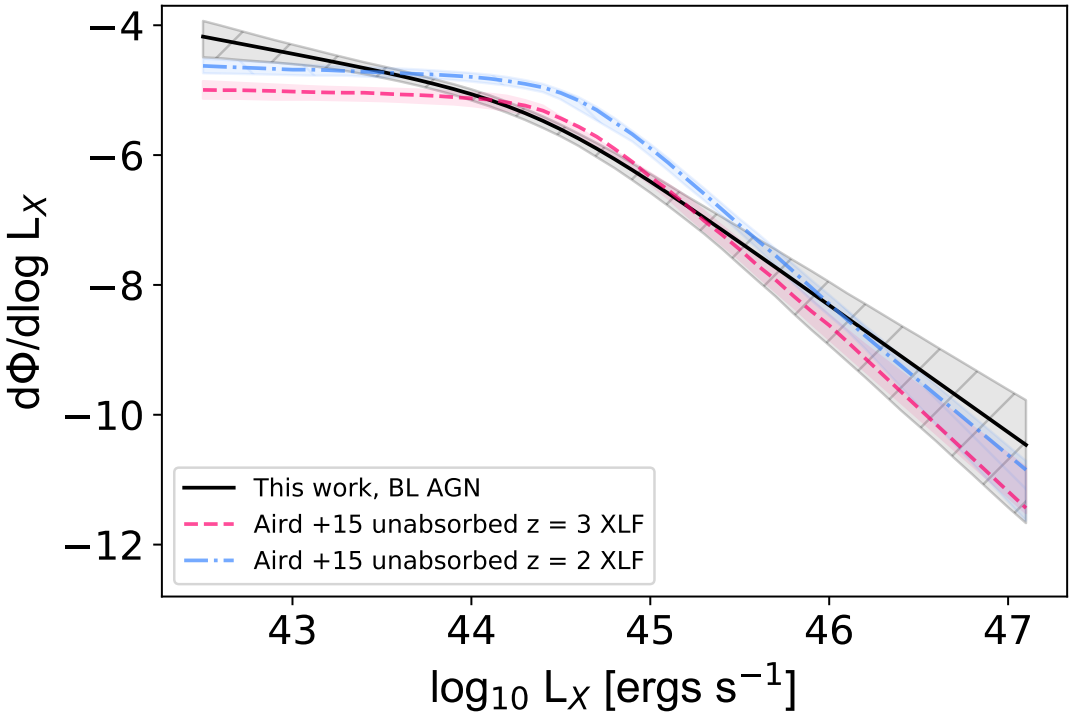

Figure 4.12 *Comparing the BL (unobscured) AGN XLF as derived in this work, with previous unobscured AGN XLF calculations at z∼ 2 - 4. Black illustrates our BL AGN XLF, blue illustrates Aird et al. (2015)'s unabsorbed z = 2 AGN XLF, and pink illustrates Aird et al. (2015)'s unabsorbed z = 3 AGN XLF.*

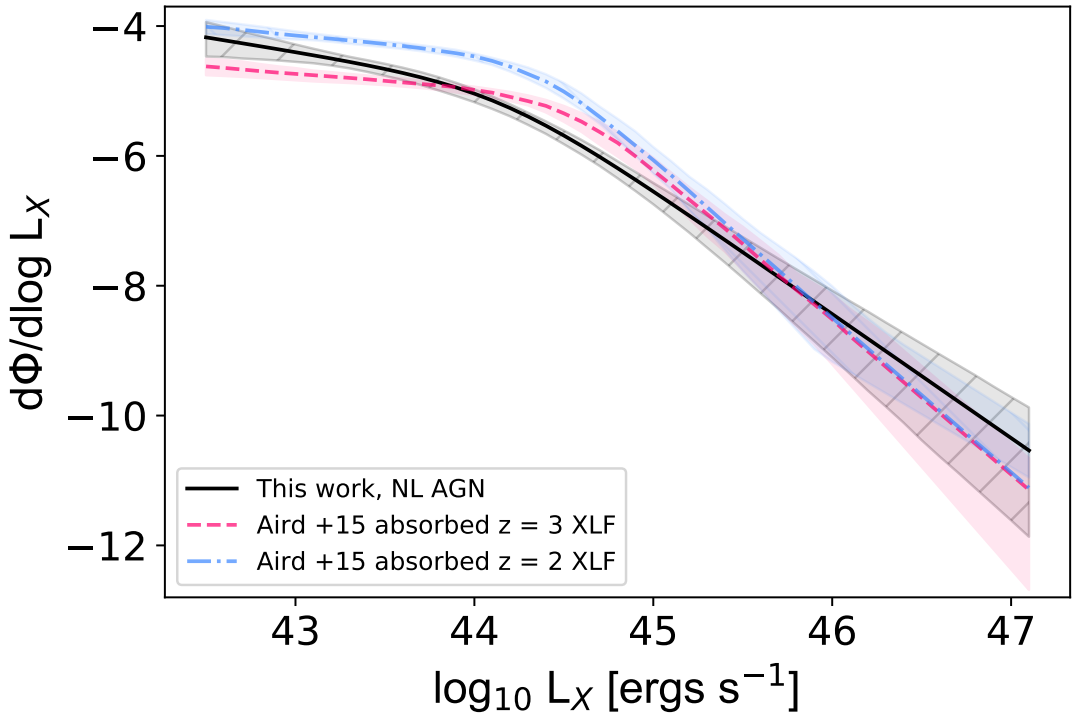

Figure 4.13 *Comparing the NL (obscured) AGN XLF as derived in this work, with previous obscured AGN XLF calculations at z∼ 2 - 4. Black illustrates our NL AGN XLF, blue illustrates Aird et al. (2015)'s absorbed z = 2 AGN XLF, and pink illustrates Aird et al. (2015)'s absorbed z = 3 AGN XLF.*

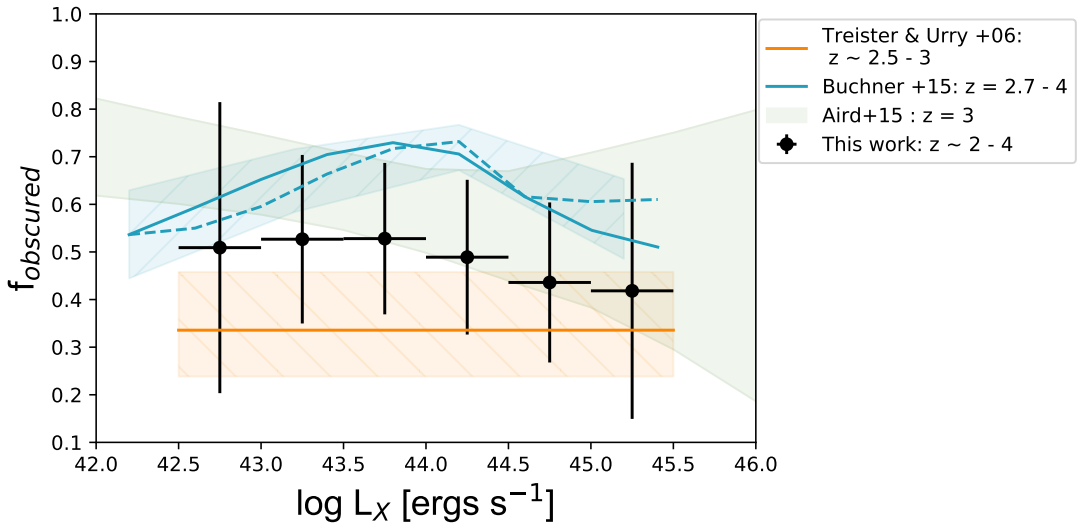

Figure 4.14 *Obscured AGN fraction as a function of X-ray Luminosity, at z ∼ 2 - 4. The colour scheme is as follows: black points illustrate the f_{obs} z ∼ 2 - 4 derived in this work; the orange region illustrates Treister & Urry (2006)'s z ∼ 2.5 - 3 f_{obs}; the blue region illustrates Buchner et al. (2015)'s z = 2.7 - 4 f_{obs}; the green region illustrates Aird et al. (2015)'s z = 3 f_{obs}.*

fraction in light green. Our analysis finds the AGN obscuration fraction to not be dependant on AGN X-ray Luminosity at z ∼ 2 - 4, within the uncertainties of this study. Whilst our obscured fraction seems to overlap with that of Treister & Urry (2006), we predict a higher overall obscured AGN fraction than that calculated in their analysis. Our obscured AGN fraction is not seen to evolve with luminosity in the way that Buchner et al. (2015)'s does; their analysis shows a peak of obscured AGN fraction at $\log_{10} L_X \sim 44.25$, whereas we predict lower obscured AGN fractions at those X-ray luminosities. It is worth noting that the definitions and selection of obscuration is not consistent across the works compared in this figure, due to the variety in obscured AGN selection techniques and definitions in the field (as previously discussed in Chapter 1 of this thesis). The implications of this are discussed in Chapter 5 of this work.

Chapter 5

Discussion

5.1 Introduction

In Chapter 2 of this thesis, I have generated new measurements of the obscured, unobscured and combined AGN XLFs at z $\sim$ 0.2 - 1.1. From these new XLF measurements, I determined new estimates of the obscured AGN fraction at z $\sim$ 0.2 - 1.1. In Chapters 3 and 4 I extended this to higher redshifts, z $\sim$ 2 - 4, whilst carrying out a more in-depth analysis and accounting for multiple completeness corrections. I constrain new and robust obscured and unobscured z $\sim$ 2 - 4 AGN XLF measurements. From these, I determined new estimates of the obscured AGN fraction at z $\sim$ 2 - 4.

In this chapter, I link my work from chapters 2 - 4. First, I discuss the evolution of obscured and unobscured AGN XLFs across cosmic time, as indicated by my combined AGN XLF results from previous chapters (Section 5.2). I then move on to discussing what my results infer about the evolution of the obscured AGN fraction with cosmic time (Section 5.3). Following this, I discuss the obscured and unobscured AGN space density evolution inferred from my work in chapters 2 - 4 (Section 5.4). Throughout, I discuss what all of these constraints tell us about obscured and unobscured AGN evolution across cosmic time.

5.2 The evolution of the obscured and unobscured AGN XLF from low to high redshifts

The work in this thesis provides new calculations and measurements of the AGN XLF across cosmic time and, as such, gives updated insight into the evolution of the obscured and unobscured AGN XLF with redshift. First, I summarise the results from Chapters 2 - 4 of this thesis, before introducing independent measurements of the XLFs of Type I and Type II AGN at z $\sim$ 0.01 - 0.3 to examine the evolution across a broad swathe of cosmic time.

In Chapter 2 I constrain new measurements for the obscured, unobscured and total z $\sim$ 0.2 - 1.1 AGN XLF using data from the new NSS80 survey catalogue. For both the FB and HB samples, these results show a higher NL AGN number density than BL AGN number density at faint luminosities ($42.5 < \log_{10}L_X < 43.5$), indicating more obscured than unobscured AGN at these faint L_X. At $\log_{10}L_X > 43.5$ we then see a higher BL AGN number density, meaning a higher incidence of unobscured AGN at these brighter luminosities. The break luminosity of the BL AGN XLF at this redshift is also higher than that of the NL AGN XLF. This indicates a higher incidence of unobscured AGN than obscured AGN at these redshifts. The faint end slope of our NL AGN XLFs are steeper than those of our BL AGN XLFs, which would indicate a steeper decline in obscured AGN number density (compared to unobscured) from faint to intermediate luminosities. In contrast to this, the BL AGN XLF at these redshifts has a steeper bright end slope, indicating a steeper decline in unobscured AGN number density from intermediate to bright luminosities.

I constrained obscured and unobscured AGN XLFs for both HB (8 - 24 keV) and FB (3 - 24 keV) X-ray samples. Our results have shown that, if we adopt the FB sample, our results have better constrained errors due to the increased number of sources. However, using the HB sample increases the overall obscured AGN XLF (selecting in higher energies allows us to pick up on the more obscured sources and estimate luminosities more accurately). This highlights the importance of being aware of how selecting in certain energy bands can influence results and conclusions drawn from selective population analysis.

I compare my z $\sim$ 0.2 - 1.1 XLF results in Chapter 2 to work previously done by Aird et al. (2015). In their paper, they present the first direct measurements of the rest-frame 10-40 keV AGN XLF based on a sample of 94 sources at $0.1 < z < 3$,

selected at 8-24 keV (comparable to our hard band sample) energies from sources in the NuSTAR extragalactic survey program. Their measurements show strong evolution of the AGN population, consistent with previous, lower-energy studies of the AGN XLF. However, they also show that adopting different models of the intrinsic distribution of absorption (these models are used to correct for selection biases) give significantly different predictions for the total number of sources in their sample. This difference leads to minor, but systematic differences in their binned estimates of the AGN XLF. I compare my obscured, unobscured and total $z \sim 0.2$ - 1 AGN XLF to the Aird et al. (2015) binned total AGN XLFs at $z \sim 0.1$ - 0.5 and $z \sim 0.5$ - 1. The comparison of my results to Aird et al. (2015) is shown in Figure 2.29, with the binned Aird et al. (2015) $z \sim 0.5$ - 1 total AGN XLF being significantly higher than my constrained total AGN XLF, particularly at fainter luminosities. This changes at brighter luminosities, where we start to see more of an agreement between the studies. Furthermore, Aird et al. (2015)'s $z \sim 0.1$ - 0.5 AGN XLF is wholly in agreement with my results within the errors of both studies. This could be due to the fact that my sample is dominated by AGN in the $z \sim 0.1$ - 0.5 range.

In Chapters 3 and 4 of this work I move onto a higher redshift AGN XLF analysis: constraining the $z \sim 2$ - 4 obscured and unobscured AGN XLFs, and folding in more careful completeness corrections. My results show the obscured and unobscured AGN XLFs to be very similar at these redshifts, from low to high luminosities, even within the errors (see Figure 4.11). This infers that we find there to be roughly equal numbers of obscured vs unobscured AGN at these redshifts, contrary to prior findings.

I then compare these results with previous estimates of the $z \sim 2$ - 4 AGN XLF (namely Aird et al. (2015)). In their paper, they present measurements of the evolution of the XLFs of unabsorbed and absorbed AGNs out to $z \sim 5$. They construct samples at 2 - 7 keV and 0.5 - 2 keV X-ray energies from a compilation of Chandra surveys supplemented by wide-area surveys from ASCA and ROSAT. They independently consider their hard and soft X-ray samples and, with and without accounting for the effect of absorption on the X-ray emission, find the best parametrization for the AGN XLF.

Figures 4.12 and 4.13 compare the unabsorbed and absorbed $z = 2$ and $z = 3$ AGN XLFs from Aird et al. (2015) with my $z \sim 2$ - 4 optically classified unobscured and obscured AGN XLFs. Looking at Figure 4.12, my BL AGN XLF has a steeper faint end slope than both Aird et al. (2015) unabsorbed AGN XLFs, predicting

higher unobscured AGN number densities than both Aird et al. (2015) unabsorbed XLFs at $42.5 < \log_{10}L_X < 43.5$. It then dips below Aird et al. (2015)'s unabsorbed XLFs at $43.5 < \log_{10}L_X < 45$, with a lower break luminosity predicting fewer AGN at intermediate X-ray luminosities at these redshifts. The bright end slope of our unobscured AGN XLF is then flatter than Aird et al. (2015)'s, indicating higher unobscured AGN number densities at $45.5 < \log_{10}L_X < 47$. The slightly different distribution of BLAGN across X-ray luminosities shown by my work could be explained by the additional completeness corrections I have folded into my XLF calculations, and the fact that we use an optical tracer of obscuration.

Moving onto Figure 4.13, my NL(obscured) z $\sim$ 2 - 4 AGN XLF predicts similar AGN number densities to Aird et al. (2015)'s at $42.5 < \log_{10}L_X < 43.5$, but the lower break luminosity and steeper faint end slope leads to lower obscured AGN number densities at $43.5 < \log_{10}L_X < 44.5$. The flatter bright end slope of my unobscured XLF predicts similar (if not slightly higher towards brighter L_X) AGN number densities to both the z = 2 and z = 3 Aird et al. (2015) absorbed AGN XLF.

To provide a local redshift comparison, I refer to a recent study undertaken by Ananna et al. (2022). In their paper, they determine the low-redshift X-ray luminosity function for both unobscured (Type 1) and obscured (Type 2) active galactic nuclei (AGNs), using the BAT AGN Spectroscopic Survey (BASS) data release 2. Their results included parameters for the z $\sim$ 0.01 - 0.3 AGN XLFs of both Type 1 and Type 2 AGNs, based on optical spectroscopic classifications. This adds a nice comparison to my obscured/unobscured AGN XLFs and, when used in combination with my results, presents an updated view of obscured and AGN XLF across redshift. Figures 5.1 and 5.2) show this AGN XLF evolution from z $\sim$ 0.01 - 4, combining Ananna+22 Type I/II AGN XLFs with my obscured/unobscured AGN XLFs.

Looking at fig 5.1: the left panel displays Ananna et al. (2022)'s z $\sim$ 0.01 - 0.2 AGN XLF results (solid dark blue line for total AGN XLF, dashed light blue line for NL/Type II AGN, dot-dashed medium blue line for BL/Type I AGN), the middle panel shows the z $\sim$ 0.2 - 1.1 AGN XLFs I constrain in Chapter 2 of this Thesis . The right panel shows the z $\sim$ 2 - 4 AGN XLFs I constrain in Chapter 3 and 4 of this Thesis . The Ananna et al. (2022) XLFs (5.1 left panel) show higher NL AGN number densities at low L_X ($42.5 < \log_{10} L_X < 43.25$) and higher BL number densities at high L_X ($43.5 < \log_{10} L_X < 45$), with the total AGN XLF peaking at 10^{-4} at the lowest X-ray luminosities. My z $\sim$ 0.2 -

1.1 results from chapter 2 (5.1 middle panel) show a similar pattern; higher NL AGN number densities at low L_X ($42.5 < \log_{10} L_X < 43.5$) and higher BL AGN number densities at high L_X ($43.7 < \log_{10} L_X < 45.5$), albeit with slightly higher total AGN number densities at the lowest X-ray luminosities. However, my z $\sim$ 2 - 4 results from chapter 4 (5.1 right panel) shows almost equal BL and NL AGN number densities across L_X, with a lower total AGN number densities at the lowest X-ray luminosities.

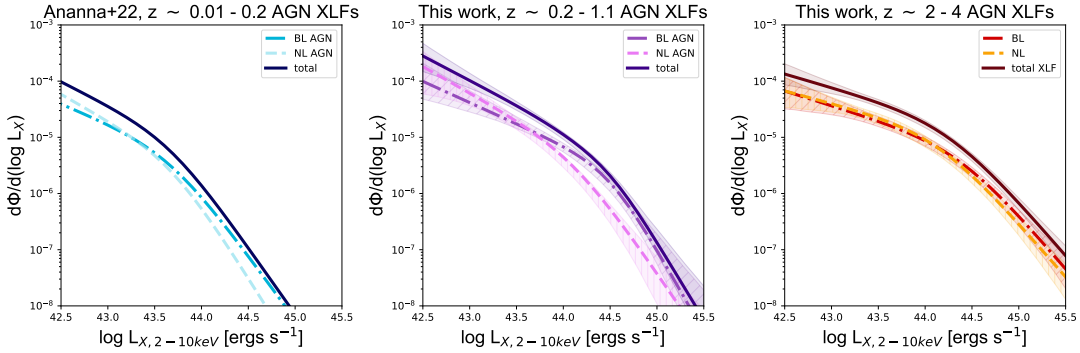

Figure 5.1 *BL, NL and total AGN XLFs for different redshift ranges.* ***Left****: Ananna et al. (2022)'s z $\sim$ 0.01 - 0.2 AGN XLF results; solid dark blue line for total AGN XLF, dashed light blue line for NL/Type II AGN, dot-dashed medium blue line for BL/Type I AGN.* ***Middle****: z $\sim$ 0.2 - 1.1 AGN XLFs constrained in Chapter 2 of this Thesis; solid dark purple line with dark purple un-hashed shaded error region for total AGN XLF, dashed pink line with pink vertically-hashed shaded error region for NL/obscured AGN, dot-dashed medium purple line with medium purple diagonally-hashed shaded error region for BL/unobscured AGN.* ***Right****: z $\sim$ 2 - 4 AGN XLFs constrained in Chapters 3 and 4 of this Thesis; solid dark brown line with brown un-hashed shaded error region for total AGN XLF, dashed yellow line with yellow vertically-hashed shaded error region for NL/obscured AGN, dot-dashed red line with red diagonally-hashed shaded error region for BL/unobscured AGN.*

Figure 5.2 illustrates another way to easily see features of the total, BL and NL AGN XLF's evolution with redshift, using a combination of Ananna et al. (2022)'s local AGN XLF work with the work done in this Thesis. The left panel shows the evolution of the total AGN XLF with redshift. The middle panel shows the evolution of the BL (unobscured) AGN XLF with redshift. The right panel shows the evolution of the NL (obscured) AGN XLF with redshift

Looking at fig 5.2: the left panel (total AGN XLF evolution across redshift) shows positive luminosity evolution (i.e. XLF moving to the right of the plot with redshift) and very slight positive density evolution (i.e. XLF moving up with redshift) from low redshift to intermediate redshifts. This indicates higher number densities of AGN at intermediate redshifts, in comparison to the number

densities we see in our local universe. We then see a flattening of the faint end of the XLF at higher redshifts, as well as the higher-z XLF having a higher break luminosity. This evolution bring the total AGN number density down at faint L_X, whilst showing an increase in the AGN number density at intermediate to brighter L_X.

The middle panel of figure 5.2: illustrates the BL (unobscured) AGN XLF evolution across redshift. Once more, we see a positive luminosity evolution (i.e. BL AGN XLF shifts to the right) and a positive density evolution (BL AGN XLF shifts up slightly) from low to intermediate redshift, in addition to the flattening of the faint end and bright end slopes at high redshifts. This evolution similarly predicts an increase in unobscured AGN number density across all luminosities from low to intermediate redshifts, with brighter and more luminous unobscured AGN seen at intermediate redshifts than at lower redshifts. Then, as we go back in time to the higher redshifts, we see a decrease in the number densities of fainter unobscured AGN, and an increase in the number densities of intermediate to bright unobscured AGN.

Finally, looking at the right panel of figure 5.2 (the NL, obscured AGN XLF evolution across redshift) we see the same type of evolution with redshift: the XLF moves to the right with increasing redshift and the faint end slope flattens from intermediate to higher z. Our intermediate redshift obscured AGN XLF does also predict higher AGN number densities at the faintest luminosities than our unobscured AGN XLF does. This indicates a larger number density of obscured to unobscured AGN even at intermediate redshifts.

Another interesting feature to note from Figure 5.2 is the difference in the rate of evolution of the BL and NL AGN XLF. Looking at the middle and right panels of figure 5.2, the rate of evolution of the BL AGN XLF from low to intermediate redshift seems to be much greater than that of the NL AGN XLF. This seems to indicate a more rapid decrease in the incidence unobscured SMBHs from $z \sim 0.2$ - 1.1 to $z \sim 0.01$ - 0.3. This rate then seems to be much slower when going from $z \sim 2$ - 4 to $z \sim 0.2$ - 1.1. The evolution of the incidence of NL AGN with redshift seems to be more constant/even.

Comparing the above evolution with previous works, Figure 5.3 shows the evolution of the total, unabsorbed and absorbed AGN XLFs with redshift, constrained in Aird et al. (2015). They see a similar luminosity evolution (the XLF moves out to the right as you go back in time) and also a similar flattening

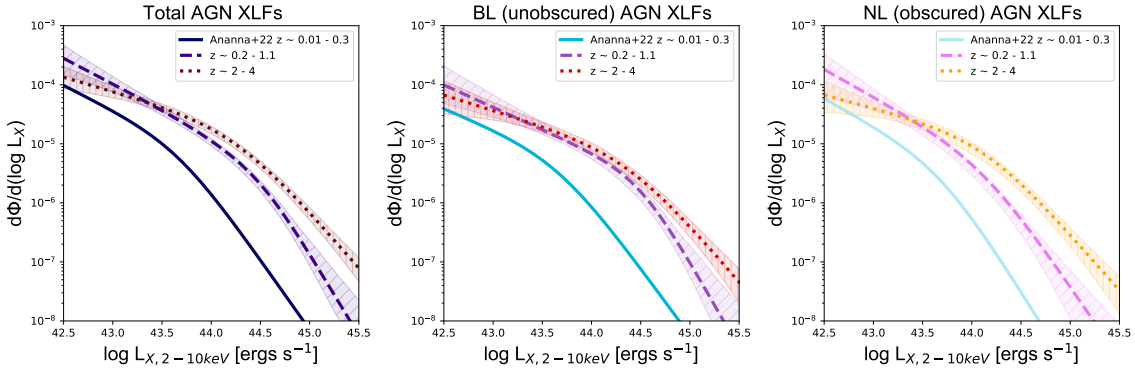

Figure 5.2 *The evolution of the total, BL and NL AGN XLF with redshift.* ***Left panel:*** *the evolution of the total AGN XLF with redshift - the solid dark blue line indicates the Ananna+22 z ∼ 0.01 - 0.3 XLF, the dashed dark purple line with diagonally hashed purple shaded regions shows my z ∼ 0.2 - 1.1 XLF, and the dotted dark brown line with vertically hashed brown shaded region shows my z ∼ 2 - 4 XLF.* ***Middle panel:*** *the evolution of the BL (unobscured) AGN XLF with redshift: the solid blue line indicates the Ananna et al. (2022) z ∼ 0.01 - 0.3 XLF, the dashed purple line with diagonally hashed purple shaded regions shows my z ∼ 0.2 - 1.1 XLF, and the dotted red line with vertically hashed red shaded region shows my z ∼ 2 - 4 XLF.* ***Right panel:*** *the evolution of the NL (obscured) AGN XLF with redshift: the solid light blue line indicates the Ananna et al. (2022) z ∼ 0.01 - 0.3 XLF, the dashed pink line with diagonally hashed pink shaded regions shows my z ∼ 0.2 - 1.1 XLF, and the dotted yellow line with vertically hashed yellow shaded region shows my z ∼ 2 - 4 XLF.*

of the faint end of the XLF as you approach the z ∼ 3 times. However, there are some important differences to note when comparing figures 5.2 and 5.3. Aird et al. (2015)'s intermediate redshift XLFs in figure 5.3 have much flatter faint end slopes than the ones constrained in this work in figure 5.2. This indicates that we predict a higher incidence of less luminous AGN across the board that previous studies (regardless of obscuration). This effect could be due to the careful completeness corrections carried out in this work, coupled our selection criteria and the capabilities of using Keck's LRIS to target the faintest sources. This highlights the importance of conducting such studies and accounting for these faint and hard to see sources - in doing so, we have find more of them than previous studies. This effect is also seen in the higher redshift XLF comparison between our work in figure 5.2 and Aird et al. (2015)'s work in figure 5.3 - we see a higher incidence of both obscured and unobscured AGN at these faintest luminosities.

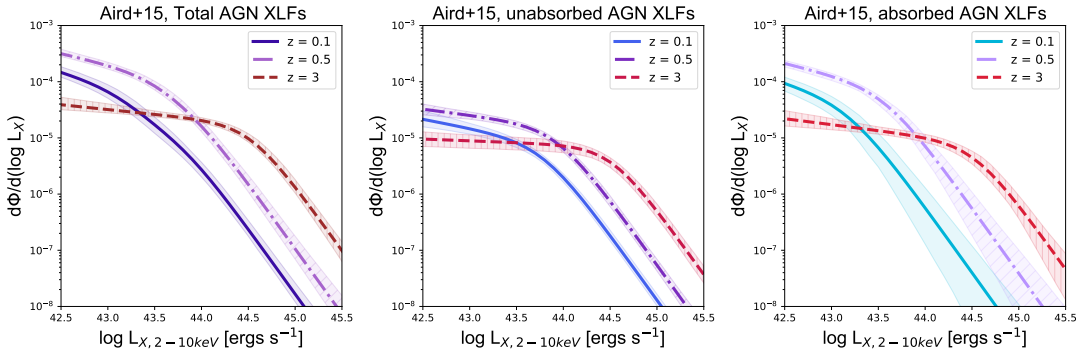

Figure 5.3 *The evolution of the total, BL and NL AGN XLF with redshift, as presented in Aird et al. (2015).* ***Left panel:*** *the evolution of the total AGN XLF with redshift - the solid dark blue line and blue shaded uncertainty regions indicates the z = 0.1 XLF, the dot-dashed purple line with diagonally hashed purple shaded regions shows the z = 0.5 XLF, and the dashed dark red line with vertically hashed brown shaded region shows the z = 3 XLF.* ***Middle panel:*** *the evolution of the unabsorbed AGN XLF with redshift: the solid blue line and blue shaded uncertainty regions indicates the z = 0.1 XLF, the dot-dashed purple line with diagonally hashed purple shaded regions shows the z = 0.5 XLF, and the dashed red line with vertically hashed red shaded region shows the z = 3 XLF.* ***Right panel:*** *the evolution of the absorbed AGN XLF with redshift: the solid light blue line with shaded blue uncertainty regions indicates the z = 0.1 XLF, the dot-dashed purple line with diagonally hashed purple shaded regions shows the z = 0.5 XLF, and the dotted red line with vertically hashed red shaded region shows the z = 3 XLF.*

5.3 The evolution of the obscured AGN fraction from low to high redshifts

In Chapter 2 of this work, I use the z $\sim$ 0.2 - 1.1 obscured and unobscured AGN XLFs I constrained, to calculate the obscured AGN fraction at z $\sim$ 0.2 - 1.1 as a function of X-ray luminosity. I do this for both my FB and HB detected sample and find no significant dependence of obscured fraction with X-ray luminosity for both samples, within the uncertainties of this work. I find using the hard band detected sample gives us higher obscured fractions than the full band sample. This is once more due to the fact that higher energy X-rays are better at penetrating all but the most obscuring material surrounding, allowing us to find more of these obscured sources. This again highlights the benefits of selecting in the harder X-ray bands when undertaking obscured AGN analysis. The full band sample results gives us better error constraints, due to its larger sample size. I compare my obscured fractions to those constrained in Hasinger (2008). In their study they similarly grouped AGN into obscured and

unobscured based on their optical spectroscopic classification. However, when this wasn't possible, they used hardness ratios to define obscuration based on source X-ray absorption properties. They used their sample of $\sim$ 1400 X-ray selected AGN across z $\sim$ 0.2 - 3.2 to generated obscured AGN fractions as a function of luminosity. The comparison of my results in chapter 2 and the Hasinger (2008) results is shown in Figure 2.30. Hasinger (2008) shows a very strong luminosity dependence in the obscured AGN fraction, whereas my results show less of this relationship. This difference, along with the tighter constrained error bars of Hasinger (2008), are the result of a variety of reasons. Firstly, their sample size is more than 5 times larger than ours, which would naturally result in smaller errors. Additionally, they include sources across a much larger redshift range than we do. This also increases their sample size (decreasing error bars), along with introducing potential redshift effects. Furthermore, they don't solely base AGN obscuration on their optical spectroscopic source properties - they fill in their gaps of missing redshift and emission line data by using photometric redshifts and X-ray absorption properties of their sample. Our work solely uses optical spectroscopy to define and classify AGN obscuration and, in doing so, hones in on AGN obscuration of the broad emission line region, giving us a truer representation of this specific and localised form of obscuration.

In Chapter 3 and 4 of this work, I use the z $\sim$ 2 - 4 obscured and unobscured AGN XLFs I constrained, to calculate the obscured AGN fraction at z $\sim$ 2 - 4 as a function of X-ray luminosity. We find, even more so, that the obscured AGN fraction at these redshifts does not change with X-ray luminosity, within the errors of this study; we find it to stay roughly constant around an obscured fraction of 0.5. This indicates almost equal numbers of obscured vs unobscured AGN at these cosmic times, across X-ray luminosities. I compare these results to Aird et al. (2015); Buchner et al. (2015); Treister & Urry (2006) at comparable redshift ranges (see figure 4.14), finding them to be broadly consistent with previous works whilst still predicting less of an obscured fraction evolution with X-ray luminosity.

To provide a more local universe comparison and to showcase my new constraints of the AGN obscured fraction across a broad range of cosmic time, I combine my results once more with the obscured fractions derived in Ananna et al. (2022). Using their Type I and Type II AGN XLF parameters, I calculate an obscured AGN fraction as a function of X-ray luminosity for z $\sim$ 0.01 - 0.3. I then combine this result with the work done in this Thesis to illustrate obscured AGN evolution across cosmic time, shown in Figures 5.4 and 5.5.

Looking at Figure 5.4: the left panel shows Ananna et al. (2022)'s obscured AGN fraction at z $\sim$ 0.01 - 0.3 in solid blue, my hard band sample's z $\sim$ 0.2 - 1.1 obscured AGN fraction in purple, and my z $\sim$ 2 - 4 obscured AGN fraction in red, with binned estimates on the left and shaded contours displayed on the right. Figure 5.5 shows the same thing, but using larger bins. From these plots we can clearly see the evolution of obscured AGN fraction across cosmic time. In the local universe: there appears to be higher fractions of obscured AGN at faint luminosities ($f_{frac} \sim 0.6$ at $\log_{10}L_X \sim 42.5$) and lower fractions of brighter obscured AGN ($f_{frac} \sim 0.2$ at $\log_{10}L_X \sim 45.5$), with a strong dependence of obscured fraction with X-ray luminosity. Then, moving out to intermediate redshifts, we predict a slight increase of the average obscured fraction at the lowest luminosities ($f_{frac} \sim 0.7$ at $\log_{10}L_X \sim 42.5$) and at higher luminosities ($f_{frac} \sim 0.5$ at $\log_{10}L_X \sim 45.5$), with only a hint (if that) of a dependence of obscured fraction on X-ray luminosity. Finally, moving out to higher redshifts, we see a decrease in the average obscured fraction at the lowest X-ray luminosities ($f_{frac} \sim 0.6$ at $\log_{10}L_X \sim 42.5$), with no real evolution with luminosity shown even for the brightest sources ($f_{frac} \sim 0.5$ at $\log_{10}L_X \sim 45.5$).

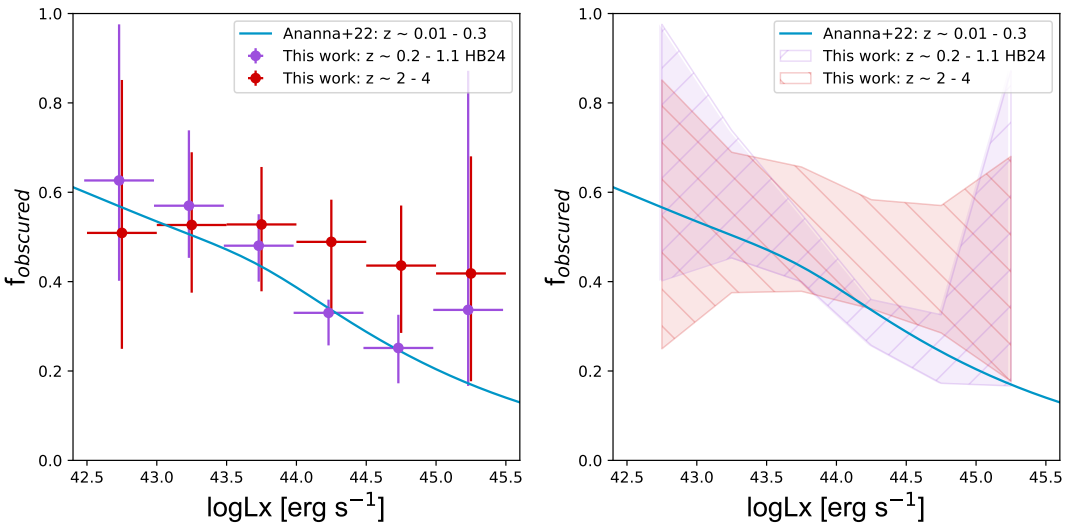

Figure 5.4 *Obscured AGN fraction as a function of log_{10} 2 - 10 keV X-ray luminosity, in bins of size 0.5 dex.* ***Left:*** *z $\sim$ 0.01 - 0.3 obscured AGN fraction calculated from Ananna et al. (2022) AGN XLF parameters in solid blue, binned z $\sim$ 0.2 - 1.1 obscured AGN fraction estimates calculated in this work in purple, z $\sim$ 2 - 4 obscured AGN fraction estimaes calculated in this work in red.* ***Right:*** *same as left, but with transparents shaded error regions. Purple right-leaning diagonal hashes represent our z $\sim$ 0.2 - 1.1 obscured AGN fractions and red left-leaning diagonal hashes represent our z $\sim$ 2 - 4 obscured AGN fractions.*

To compare the evolution of the obscured AGN fraction across cosmic time shown

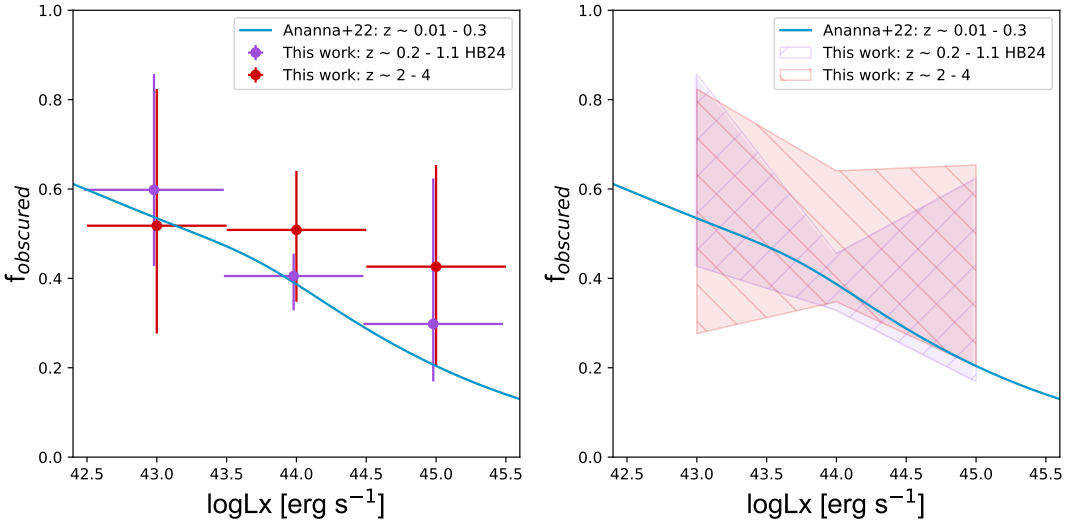

Figure 5.5 *Obscured AGN fraction as a function of log_{10} 2 - 10 keV X-ray luminosity, in bins of size 1 dex. Colour scheme is consistent with that shown in figure 5.4.*

by the results of my work, I look at the results from Aird et al. (2015). Figure 5.6, we can see that Aird et al. (2015)'s absorbed fraction as a function of luminosity for different z shows quite a strong redshift evolution and luminosity dependence. However, given that these are based on the X-ray obscuration of their sources as opposed to optical spectroscopic classifications, it is interesting that we still see similar trends: less of an obvious luminosity dependence as redshift increases, and a flattening and lowering of the overall obscured fraction around cosmic noon. This hints at a welcome concurrence between X-ray and optical obscured fraction evolution.

Signs of a luminosity dependence would be in favour of the receding torus model (as described in Chapter 1). Whilst our results do not show such an extreme a dependence as Aird et al. (2015)'s, following the evolution of the equal obscured AGN fractions at cosmic noon which we have constrained, down to the more luminosity dependent obscured fraction at lower redshifts, could highlight interesting AGN evolutionary phases. At higher redshifts (z $\sim$ 2 - 4) our results show more equal numbers of obscured and unobscured AGN. Then, as we travel through cosmic time to more intermediate redshifts (z $\sim$ 0.2 - 1.1), we see that the more luminous AGN have already started to recede their obscuring tori over cosmic time. Whilst the obscured fractions we measure for this z show only a mild luminosity dependence at best, it could be that the rate at which the brightest AGN start to recede their tori is less extreme than previously thought,

and happens over a longer period of time across all luminosities (which would result in less of an obvious steep luminosity dependence, as seen in our work). This then takes us to the more local universe, where we have more of an obvious dependence on luminosity as the brighter AGN have successfully receeded their tori.

At intermediate z: the same luminosity dependence is in favour of a receding torus model. Obscured fraction only shows a marginal increase (if any) from low to intermediate redshifts. At higher z: only a mild luminosity dependence (if any) is seen - within the errors, there is no significant luminosity dependence shown. Obscured AGN fractions at higher z are a bit higher than at lower z, which points to there being more obscured agn at earlier cosmic times. The mild lum dependence could favour a receding torus model, but we see no obvious dependence within the errors of this study. This points to there being more equal numbers of obscured vs unobscured AGN at these earlier cosmic times, than other works have suggested. Maybe at the receding torus model kicks in at lower redshifts - at these higher redshifts there could be equal numbers of obscured and unobscured agn, and then the more luminous AGN start to 'recede the torus' over cosmic time, hence us seeing more of a luminosity dependence at lower and more local redshifts.

5.4 The evolution of the obscured and unobscured AGN space density from low to high redshifts

My results provide new constraints on the space density of obscured and unobscured AGN as a function of redshift, providing new insights into these phases of BH growth over cosmic time, which I explore here. To derive an AGN space density from my XLF results, I integrate each XLF over luminosity. I do this for my obscured, unobscured and total AGN XLFs at z $\sim$ 0.2 - 1.1 and z $\sim$ 2 - 4, along with the Ananna et al. (2022) XLFs at z $\sim$ 0.01 - 0.3. I present results for luminosity bins of various sizes, as can be seen in Figures 5.7 and 5.8.

Taking a look at figure 5.7, the left panel shows measurements of BL (black points), NL (blue points) and total (red points) AGN space densities I've calculated from the low-z Ananna+22, my intermediate-z and my high-z XLFs. These space densities have been integrated over the full L_X range (42.5

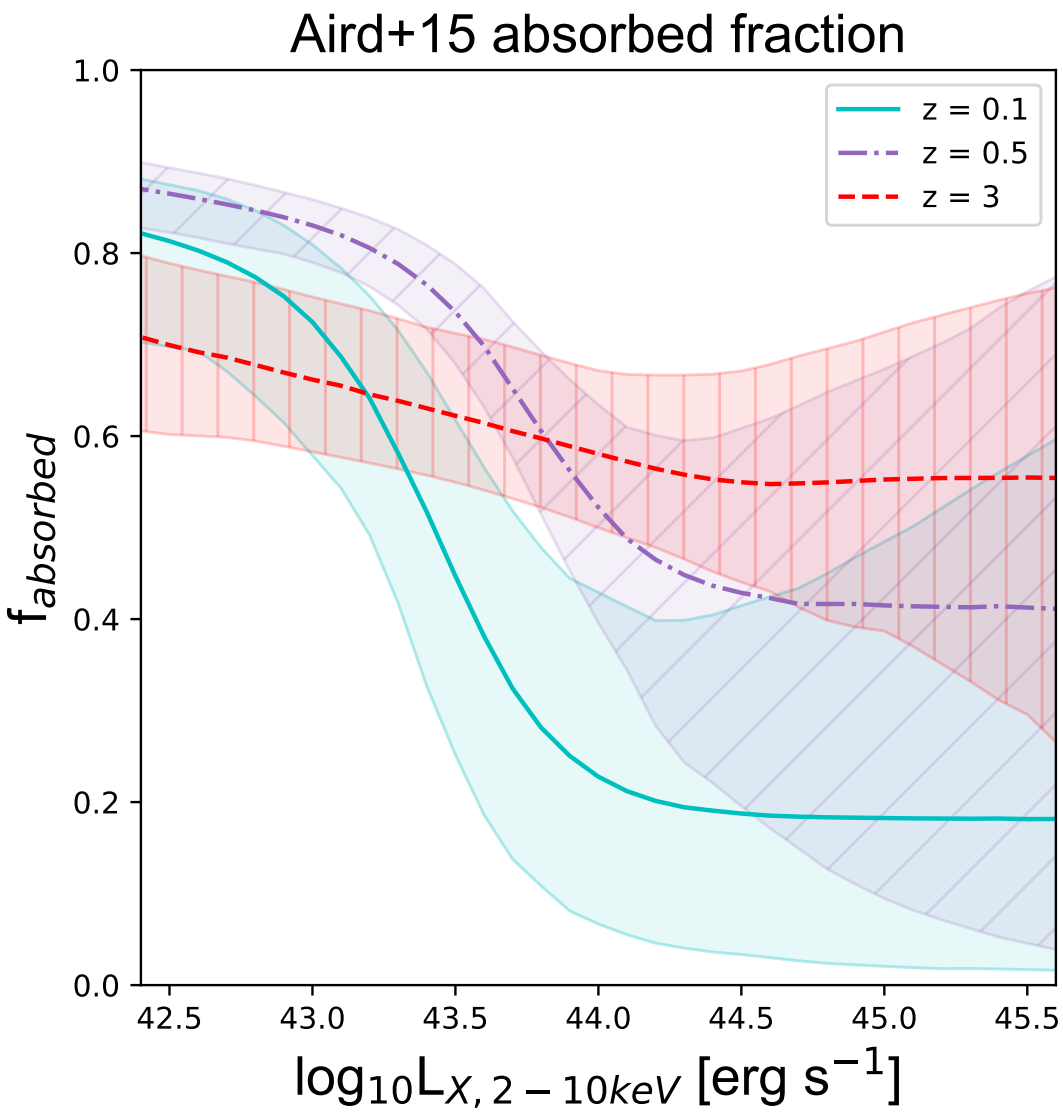

Figure 5.6 *Aird et al. (2015)'s fraction of absorbed AGNs, defined as the fraction of AGN with 22 < log_{NH} < 24 relative to all Compton-thin, log_{NH} < 24, AGNs based on their best-fitting model of the XLF evaluated at a number of redshifts. Taken and adapted from Aird et al. (2015).*

< $\log_{10}$ L_X < 45.5). The right panel shows the same points, but with total/absorbed/unabsorbed AGN space densities taken from Aird et al. (2015), overplotted in lines as a comparison. Looking at my results, we see an increase in BL, NL and total AGN space density from low to intermediate redshifts. We then see a decrease across the board from intermediate to high redshifts. Looking at the BL and NL populations individually, the decrease in AGN space density from intermediate to high redshifts is less acute for the BL AGN population than it is for the NL AGN population; we see a higher incidence of obscured AGN than unobscured AGN across all z, but the space density of the unobscured AGN

population does not seem to evolve as rapidly as the obscured AGN population as we go back in cosmic time from intermediate to high z.

Comparing my optical-classification based results to Aird et al. (2015)'s X-ray absorption based results (i.e. the right panel of figure 5.7), we see that they are mostly in agreement, within the errors of this study. Ananna et al. (2022)'s optical-classification based results predict lower AGN space densities across the board at low-z. My intermediate-z NL (obscured) and combined AGN space densities are in agreement with Aird et al. (2015)'s absorbed and total AGN space density. However, my BL AGN results predict higher space densities for the unobscured AGN population, than Aird et al. (2015)'s unabsorbed space density measurements - in this thesis, my constrained z $\sim$ 2 - 4 BL AGN XLF predicts higher unobscured AGN space densities at intermediate and high-z. This indicates that there could be more unobscured AGN than previously predicted at these redshifts. Another important result of mine to note is that, at all redshifts, at least 50% of the AGN space density is due to obscured AGN. In fact, at low and intermediate redshifts, obscured AGN make up the greater proportion, and the relative space densities become more comparable towards the higher redshifts. This same trend is apparent in the Aird+15 measurements, and corresponds to my results highlighting evolution of the AGN XLF; higher obscured AGN numbers indicated at low and intermediate redshifts, with the obscured and unobscured AGN XLFs becoming more comparable at higher redshifts.

A way to further examine these AGN space density results is to split them up into differing X-ray luminosity bins. Figure 5.8 shows exactly that; the AGN space densities for my work combined with Ananna et al. (2022) (solid points), and compared to Aird et al. (2015)'s (lines) in various X-ray luminosity bins.

First, we consider the left panel of figure 5.8: the unobscured/unabsorbed AGN space density measurements. Looking at the lowest luminosity bin ($42.5 < \log_{10} L_X < 43.5$, in black), the lowest-z space density measurement I calculated from Ananna et al. (2022)'s results is in agreement with that of Aird et al. (2015). However, my intermediate and high-z space density predictions are higher than Aird et al. (2015)'s; my intermediate-z results are higher by a factor of $\sim 2\sigma$, and my high-z results are significantly higher than Aird et al. (2015)'s ($> 3\sigma$). This indicates that we predict higher numbers of very faint, unobscured AGN at intermediate to high redshifts than previous studies, with a much shallower drop-off seen at higher redshift than seen in Aird et al. (2015). This in itself is a significant result, and is due to i) my use of optical classifications and ii) the very

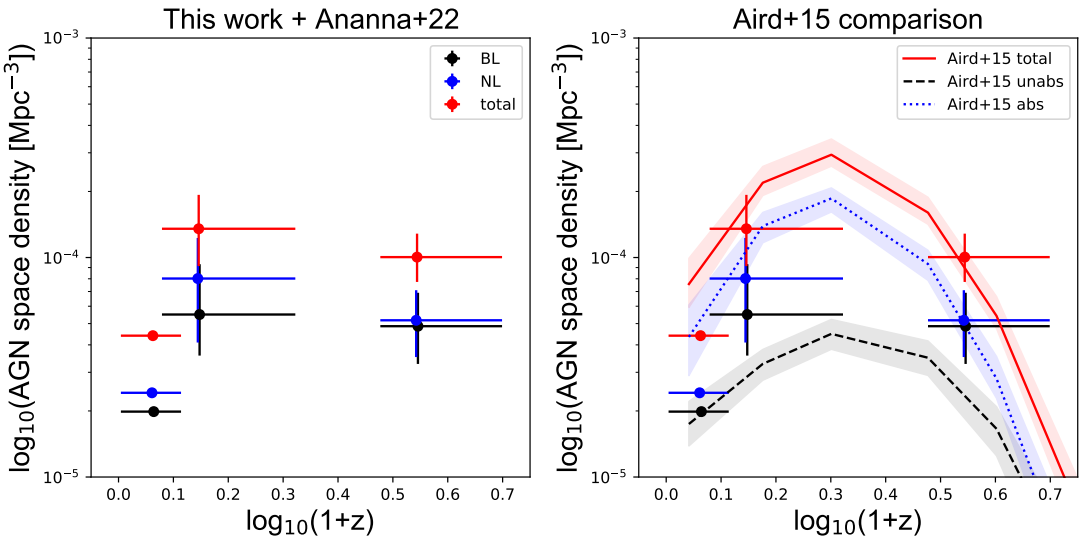

Figure 5.7 *AGN space densities as a function of redshift, across the full 42 < log_{10} L_X < 45.5 range.* ***Left:*** *BL (black points), NL (blue points) and total (red points) AGN space densities as a function of redshift, calculated from Ananna et al. (2022), and the work done in this thesis.* ***Right:*** *The same as in the left, but with Aird et al. (2015)'s total (red solid line), X-ray unabsorbed (black dashed line) and X-ray absorbed (blue dotted line) overplotted for comparison.*

careful corrections for incompleteness I have accounted for at low fluxes and low optical magnitudes.

Looking at the intermediate luminosity bin (L_X 43.5 - 44.5, in blue), the space density result calculated from Ananna et al. (2022) predicts lower numbers at the lowest redshifts than Aird et al. (2015)'s does, with my intermediate to high-z measurements being in agreement with Aird et al. (2015)'s results. Finally, taking the highest luminosity bin (L_X 44.5 - 45.5, in red), my points are once more in agreement (within the errors) with Aird et al. (2015)'s.

Moving onto the middle panel of figure 5.8: the obscured/absorbed AGN space density measurements. Looking at all luminosity bins, my results are in agreement with Aird et al. (2015)'s (within the errors of this study) - an abundance of the faintest, obscured AGN across z, with a slight decrease in space density as we get to higher-z; intermediate-luminosity obscured AGN are the next most prevalent, with space density seemingly increasing with z; and the brightest obscured AGN are the least common, with space density once more increasing with z. Finally, looking at the right panel of figure 5.8: the total AGN space density measurements. Once more, when examining the results across all luminosity bins, our results are in agreement with Aird et al. (2015)'s (within the errors of

this study).

Considering all panels of figure 5.8 as a whole, there are some key observations we can make. The first is that low-luminosity AGN make up the majority of all AGN, across all z. This bolsters our need to be able to push analyses down to the faintest fluxes and luminosities, so we can better study this bulk of the AGN population. Another interesting observation is that we can clearly see evidence of 'downsizing' seen in terms of AGN luminosity - we see that the space densities of the more luminous AGN (both obscured and unobscured) peak at higher-z than the lower luminosity AGN. It is also worth noting that, even though the relative rates of evolution of the space densities are different (i.e. obscured and unobscured populations evolve differently), both of these populations show a downsizing pattern. This differential evolution of faint and bright AGN with z, implies that, at low-z, AGN activity is dominated by either low-mass SMBHs growing rapidly or high-mass SMBHs accreting at low rates. It has been proposed by Hopkins et al. (2005) and Babić et al. (2007) that the faint end of the XLF is comprised of high mass BHs accreting quiescently, whilst the bright end is due to SMBHs accreting near their Eddington limit.

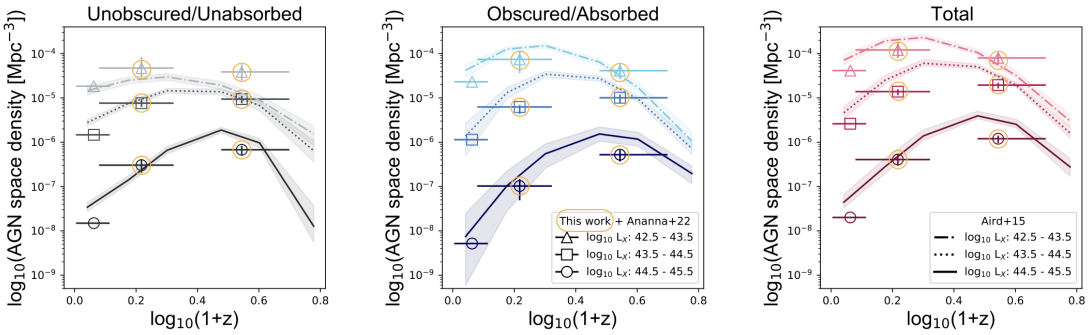

Figure 5.8 *Unobscured/unabsorbed (left panel) obscured/absorbed (middle panel) and total (right panel) AGN space density as a function of redshift, in different luminosity bins. Space density measurements calculated in this work, combined with the lowest-z measurements from Ananna et al. (2022), are illustrated by unfilled markers, with the data points produced from my work doubly outlined with a solid yellow circle in each panel.. Space density measurements from Aird et al. (2015) are in lines.*

Chapter 6

Conclusions and future work

6.1 Conclusions

In this thesis, I have explored the growth and evolution of obscured and unobscured AGN across cosmic time by producing new and robust calculations of the AGN XLFs, AGN obscured fractions and AGN space densities, for both the obscured and unobscured AGN populations.

In my lower redshift analysis, I have used X-ray and optical spectroscopic data from the 80 month NuSTAR Serendipitous Survey catalogue to define a sample of $\sim$ 100 AGN at z $\sim$ 0.2 – 1 detected in the hard (8 - 24 keV) band and $\sim$ 200 AGN at z $\sim$ 0.2 – 1 detected in the full (3 - 24 keV) band. I have accounted for incompleteness in the X-ray survey and counterpart identification, and used a Maximum Likelihood Fitting method to generate new measurements of the obscured, unobscured and combined AGN XLFs at z $\sim$ 0.2 – 1. With these results, I have determined new measurements of the fraction of obscured SMBH growth as a function of X-ray Luminosity, at these lower redshifts, as well as new measurements of the obscured and unobscured AGN space density across X-ray luminosities, as a function of redshift.

In my higher redshift analysis, I have used an optical colour pre-selection approach coupled with X-ray imaging from Chandra and XMM-Newton to identify a sample of $\sim$137 X-ray selected AGN at z $\sim$ 2 - 4, with optical spectroscopic follow-up data obtained using Keck LRIS. I have developed an automated spectral fitting code which classifies AGN sources as either broad-line or non-broad-line and

accurately identifies their redshift. I have quantified incompleteness in this optical follow-up and classification process by applying this fitting algorithm to large samples of self-generated simulated data. I have applied this completeness correction, along with careful corrections accounting for the X-ray selection and optical pre-selection process, to my results to recover XLFs and obtain robust measurements of the distribution of obscured and unobscured black hole growth in the early universe. I have used my new XLF results to constrain measurements of the obscured SMBH fraction as a function of X-ray luminosity at these higher redshifts, in addition to new measurements of the z $\sim$ 2 - 4 obscured/unobscured AGN space density across X-ray luminosity and redshift.

Throughout this work I have developed original techniques and methods to facilitate such thorough analyses. These include:

- Creating an AGN UV/optical spectral fitting, classification and redshift estimation tool suitable for characterising faint X-ray selected high-z AGN, including obscured populations.

- Creating a spectral simulation tool which generates simulated AGN UV/optical spectra over a broad range of magnitudes, redshifts, wavelengths and continuum strengths/emission line properties. This simulation tool has subsequently allowed me to characterise the completeness of the aforementioned AGN spectral classification/z-estimation procedure, enabling the inclusion of vital corrections to each stage of my statistical analysis.

- Determining new and robust calculations of the independent X-ray luminosity functions for obscured and unobscured AGN across a range of redshifts; involving the development of statistical techniques to correct for incompleteness, including X-ray selection, optical pre-selection, spectroscopic classification and redshift corrections.

The combination of my low-redshift and high-redshift analysis has allowed me to paint an updated picture of how the obscured and unobscured AGN populations evolve as a function of X-ray luminosity, over a broad cosmic time period, in an effort to determine the role obscuration may play in AGN growth and evolution. My primary results are as follows:

X-ray Luminosity function evolution: My new XLF measurements show an independent evolution for the obscured and unobscured AGN populations across $0.2 < z < 4$. Starting in the local universe and going out to intermediate

redshifts, we see an increase in the number density of both the obscured and unobscured AGN populations, with the majority being the faint and obscured AGN population. However, the rate of change of this increase is larger for the mid-luminosity unobscured AGN population than it is for their obscured counterparts. Then, as we push out to z $\sim$ 4, the faint end slopes of both the obscured and unobscured AGN XLFs flatten, and the obscured and unobscured AGN XLFs become comparable at these higher redshifts, contrary to other studies. This suggests the following evolution: the number densities of obscured and unobscured AGN as a function of luminosity are comparable at cosmic noon (i.e. there are relatively equal numbers of obscured and unobscured AGN at z $\sim$ 2 - 4). Then, as we approach intermediate redshifts (z $\sim$ 0.2 - 1.1), we see an increase in the number density of the faintest obscured and unobscured AGN, and a decrease in the number density of the intermediate-luminosity and the brightest AGN. The greater increase is seen the number density of the faintest, obscured AGN and the greater decrease is seen in the number density of the brighter, obscured AGN between these cosmic times. Then, as we go from intermediate to local redshifts, the number density of obscured and unobscured AGN decreases across all luminosities. Another key result of mine is an increase in the incidence of faint AGN at high redshift when compared to previous work. This is due to the addition of careful completeness corrections incorporated throughout my XLF calculations, which highlights the importance of such methods.

Obscured AGN fraction evolution: My obscured AGN fraction results show a mild evolution of the obscured AGN fraction as a function of X-ray luminosity, with redshift. At z $\sim$ 0.2 - 1.1 my results show a minor evolution of obscured AGN fraction with X-ray luminosity, with $f_{obs} \sim 0.7$ at $\log_{10}L_X \sim 42.5$ decreasing to $f_{obs} \sim 0.5$ at $\log_{10}L_X \sim 45.5$. At z $\sim$ 2 - 4 my results show no substantial dependence of obscured AGN fraction on X-ray luminosity, with $f_{obs} \sim 0.6$ at $\log_{10}L_X \sim 42.5$ decreasing to $f_{obs} \sim 0.5$ at $\log_{10}L_X \sim 45.5$. Combining these with the low redshift obscured AGN fractions I have derived from Ananna+22, my results paint a mildly evolving picture of obscured AGN fraction as a function of X-ray luminosity, across cosmic time; we see a strong luminosity dependence in the obscured AGN fraction in the local universe, with high fractions of the faint AGN population being obscured and low fractions of the bright AGN population being obscured. Then, as we push out to intermediate redshifts, we see the same luminosity dependence but to a lesser extent, with a slight increase in the fraction of faint, obscured AGN. Finally, when we arrive at cosmic noon, we see equal fractions of obscured and unobscured AGN across X-ray luminosity.

The lack of luminosity dependence on the obscured AGN fraction seen at higher redshifts could go to support the unified AGN model. However, the evolution we see as we move to lower redshifts supports a receding torus model; the decrease in obscured fraction with increasing X-ray luminosity could be due to the more powerful AGN sublimating their obscuring material and becoming less obscured over cosmic time. When considered together, these results point to the possibility of there being evolutionary phases of AGN obscuration over time.

AGN Space Density evolution: My AGN space density results show an increase of obscured and unobscured AGN space density from low to intermediate z, and then a decrease from intermediate to high z - this aligns with previous studies. When examined in different luminosity bins, my results show clear evidence of downsizing seen in terms of AGN luminosity - in both the obscured and unobscured populations, the space densities of more luminous AGN peak at higher-z than their lower luminosity counterparts.

When combined, these results support the possibility of there being evolutionary phases of AGN obscuration over time, with independent evolution of the obscured and unobscured population with luminosity and redshift in varying degrees. As mentioned in this Thesis' introduction, these phases could be instigated by violent AGN feedback events or mergers. However, we cannot rule out potential viewing angle effects and biases - the change in obscured fraction over cosmic time my results illustrate, do not completely rule out the Unified Model of AGN. Futhermore, being able to accurately characterise the nature and geometry of the obscuring material surrounding AGN, through theoretical models, would elevate studies like the ones displayed in this Thesis. As such, in order to further probe these objects and unite theory with observations, continued thorough multiwavelength study of obscured AGN across cosmic time is necessary to comprehend the nature and evolution of AGN obscuration.

6.2 Future work

As shown by my work in this thesis, the faint, obscured AGN population accounts for the majority of supermassive black hole growth in our universe. As they are more likely to reside in dust and gas-rich hosts, studying these obscured AGN can provide valuable insight AGN-galaxy co-evolution and the environments that are conducive to black hole growth.

However, there remain significant unresolved questions regarding the true extent of obscured SMBH growth and the role obscuration may play in AGN formation, growth and evolution. During my PhD, I have undertaken extensive and robust analyses of low-luminosity, obscured and unobscured AGN at high and low redshifts, including new calculations of independent X-ray luminosity functions for the obscured and unobscured AGN populations up to z $\sim$ 4. These statistical characterisations of said populations place vital constraints on the relative importance and contribution to black hole growth of these phases.

Multi wavelength surveys, combined with extensive spectroscopic follow-up, have enabled us to study AGN and host galaxy demographics up to high redshifts, illuminating how accretion and the associated physics at all scales evolves from late times to the earlier universe. However, due to difficulties in identifying AGN down to low-luminosities and accurately determining their obscuration properties, these surveys often favour bright and unobscured sources over their faint and dusty counterparts. Only through careful consideration and analysis of obscured and less luminous sources, combined with sophisticated corrections for survey limitations and incompleteness, can we obtain results that are representative of the true SMBH population and their host galaxies.

Throughout this thesis I have demonstrated the original techniques and methods I have developed to facilitate thorough analyses of faint and obscured sources. My results have provided an updated and accurate constraint on obscured and unobscured SMBH growth from late times to the early universe, shedding further light on the growth and evolution of AGN and their hosts, as well as the impact obscuration has and the role it plays therein.

Expanding on my former work, there are various potential avenues for future research into the links between quasar and host galaxy evolution:

Robustly quantifying AGN populations and their host galaxies with next generation spectroscopic surveys: Robust analyses of multi-wavelength quasar survey data is a vital component in survey creation and subsequent science. From generating and targeting source samples to determining spectroscopic success and distinct source properties, it is essential to be able to competently extract true quantifications of the underlying AGN and host galaxy populations from multi-wavelength observational data. Approaches such as the spectral fitting, simulation and completeness quantification procedures I have developed will allow us to effectively interpret and make robust scientific contributions when

working with spectroscopic data from current and upcoming facilities, such as VISTA 4MOST and VLT MOONS.

AGN obscuration, feedback and environment connection: It is well accepted that quasars can influence their hosts via powerful outflows, driving away dust, gas and other star forming ingredients. The measure of AGN obscuration can be an observational constraint of feedback models 'in action'; powerful AGN clean their sight-lines more rapidly than low luminosity AGN, therefore the fraction of obscured AGN can be viewed as a measure of the timescale over which the nuclear feedback is at work. If the fundamental correlation between the fraction of obscured AGN and luminosity is due to different timescales over which nuclear feedback is at work, this would have broad implications for our understanding of host galaxy evolutionary pathways. This renders obscuration an important factor to consider when studying AGN-galaxy co-evolution. Combining longer wavelength data from facilities such as ALMA and UV/optical AGN spectra will facilitate investigation into this; for example, using robust XLF calculations and evaluating the extent of obscured AGN growth relative to dust masses and galaxy star formation rates. By considering populations and quantifying properties such as these, will we be able to fully exploit multi-wavelength data to understand the potential connection between AGN obscuration, feedback and environment, which will in turn help unravel the connection between distant supermassive black holes and their host galaxies.

In all these avenues, if we want to accurately quantify obscured and unobscured AGN down to low luminosity with next-generation surveys, it will be essential to undertake careful and extensive completeness corrections. Such methods, like the techniques I have developed, will facilitate the robust and precise study of elusive populations and must be applied to future surveys if we want to wholly understand these astronomical objects and capitalise on the promising data future surveys will bring.

Chapter 7

Appendix

The following Appendix contains the optical/UV spectra of all Keck spectra used in Chapters 3 and 4 of this Thesis.

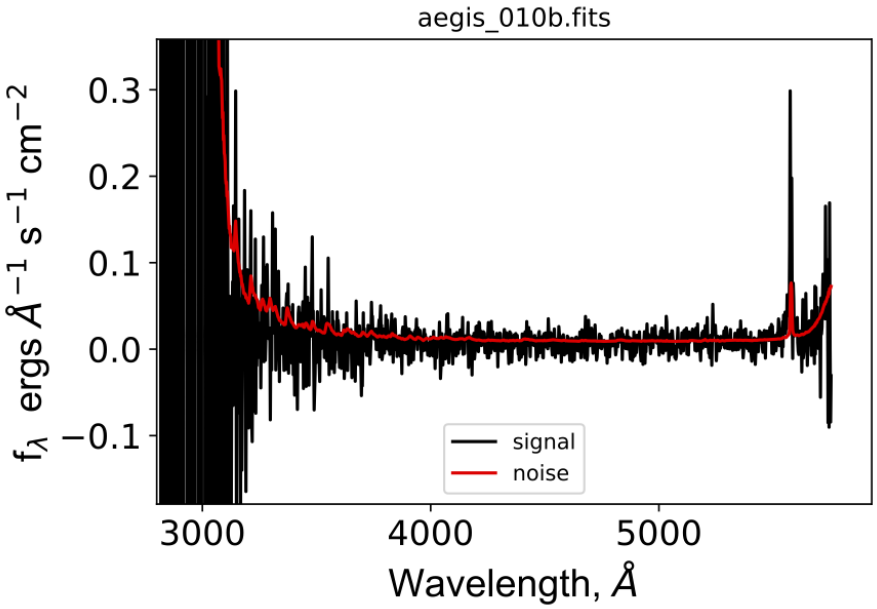

Figure 7.1 *Optical/UV spectrum of source 'aegis_010b', located in the CFHTLS D3 field.*

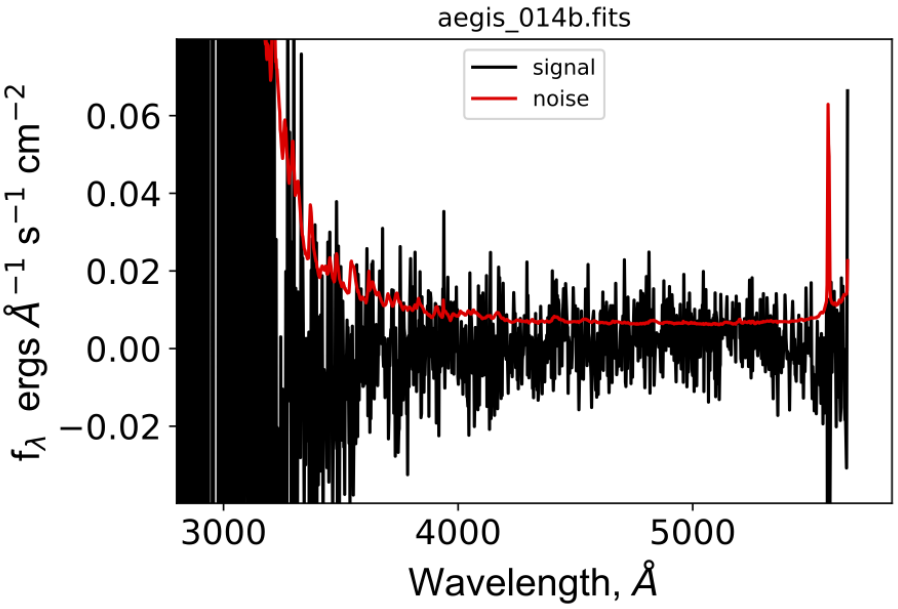

Figure 7.2 *Optical/UV spectrum of source 'aegis_014b', located in the CFHTLS D3 field.*

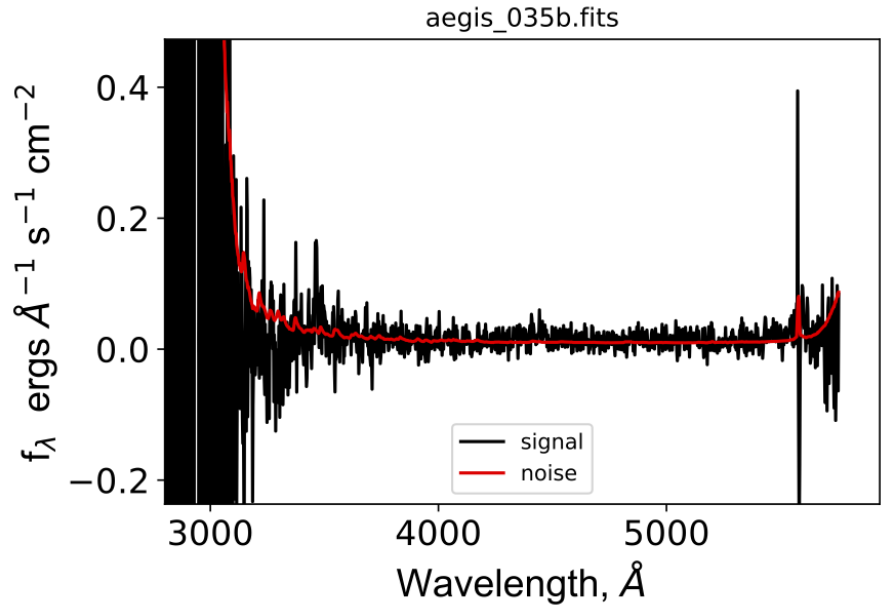

Figure 7.3 *Optical/UV spectrum of source 'aegis_035b', located in the CFHTLS D3 field.*

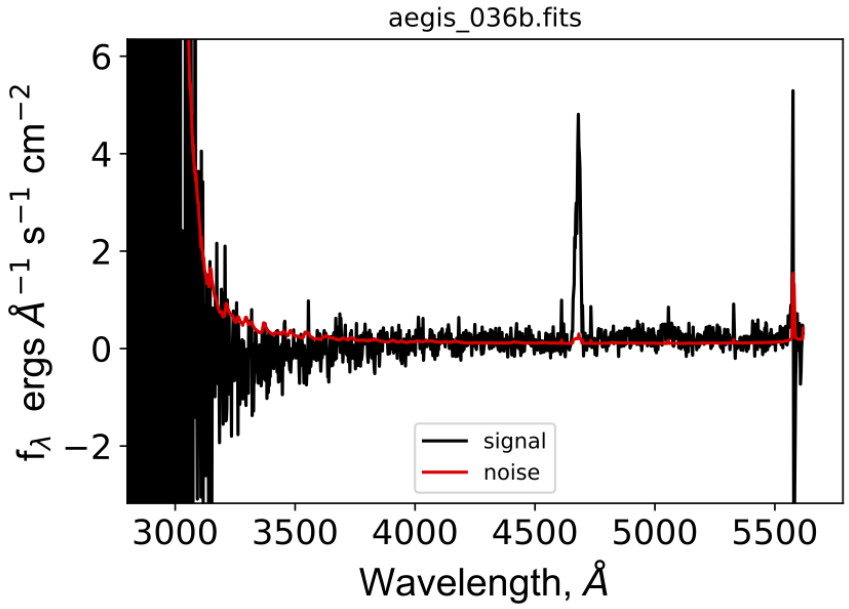

Figure 7.4 *Optical/UV spectrum of source 'aegis_036b', located in the CFHTLS D3 field.*

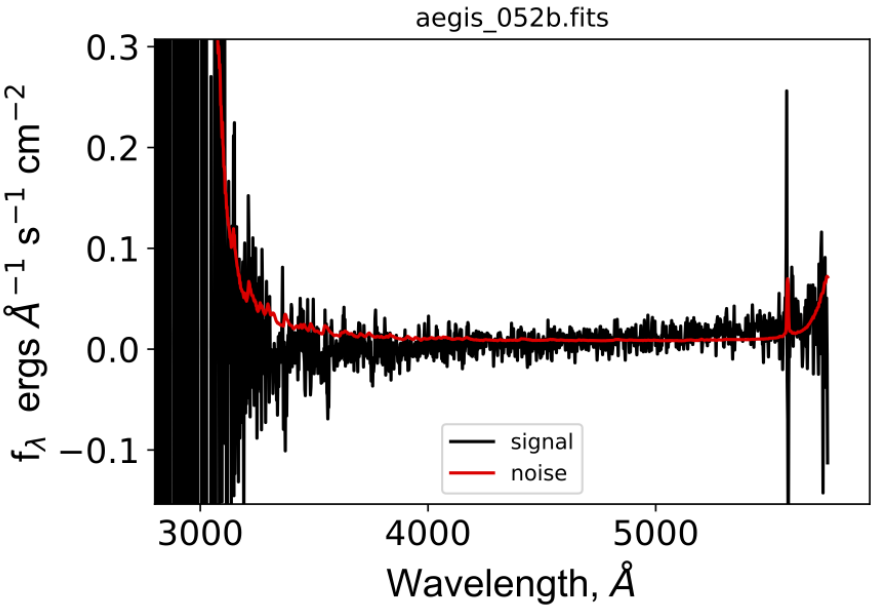

Figure 7.5 *Optical/UV spectrum of source 'aegis_052b', located in the CFHTLS D3 field.*

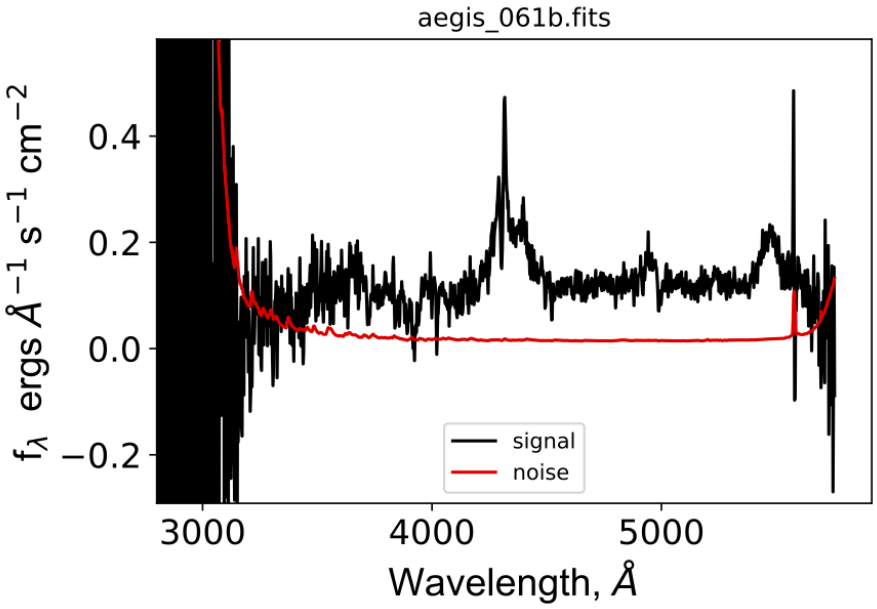

Figure 7.6 *Optical/UV spectrum of source 'aegis_061b', located in the CFHTLS D3 field.*

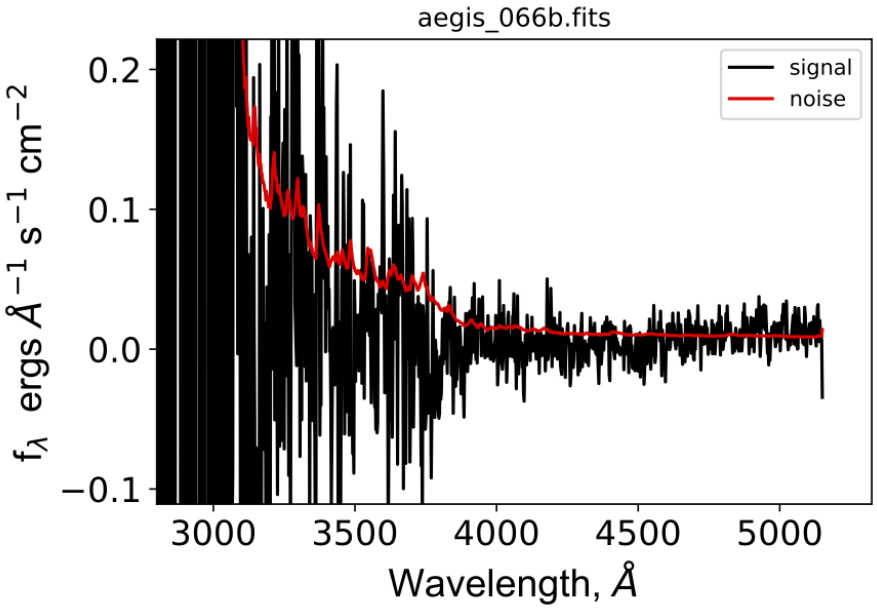

Figure 7.7 *Optical/UV spectrum of source 'aegis_066b', located in the CFHTLS D3 field.*

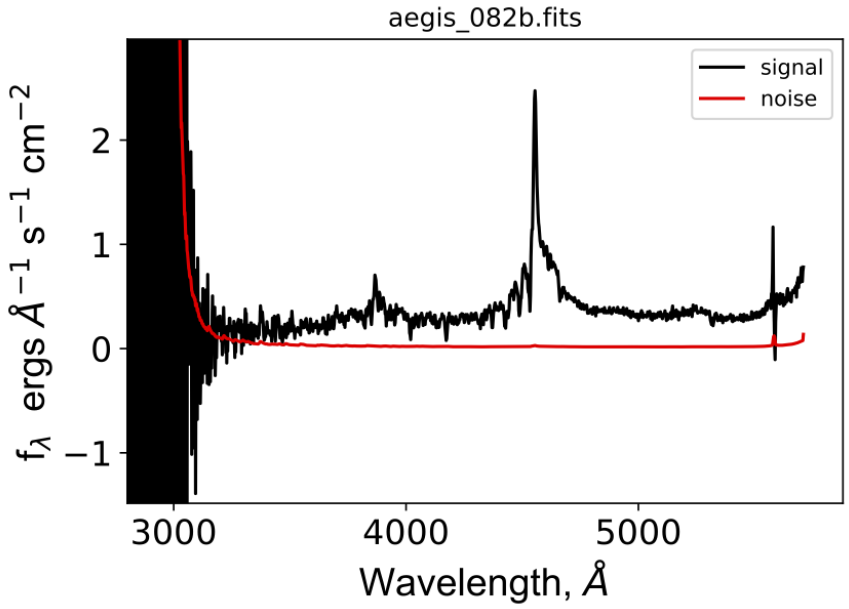

Figure 7.8 *Optical/UV spectrum of source 'aegis_082b', located in the CFHTLS D3 field.*

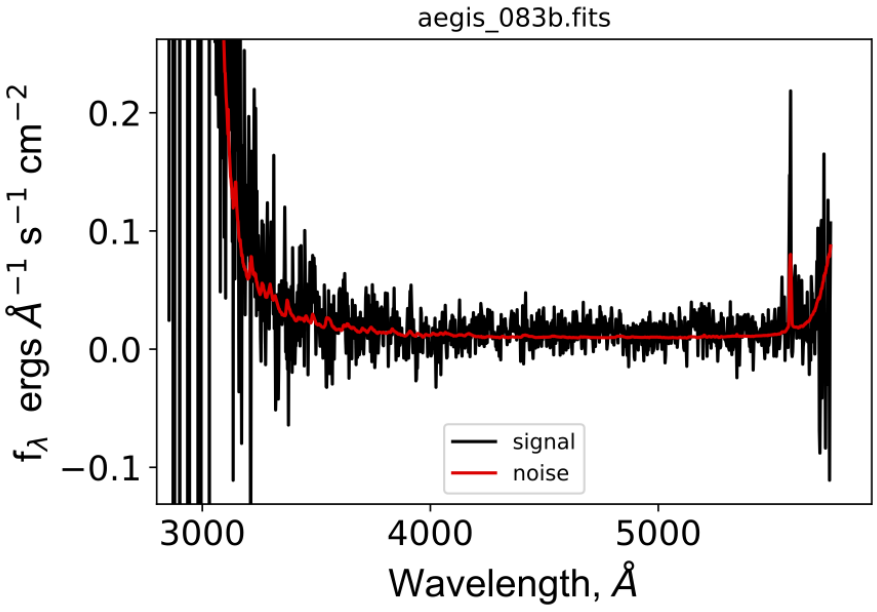

Figure 7.9 *Optical/UV spectrum of source 'aegis_083b', located in the CFHTLS D3 field.*

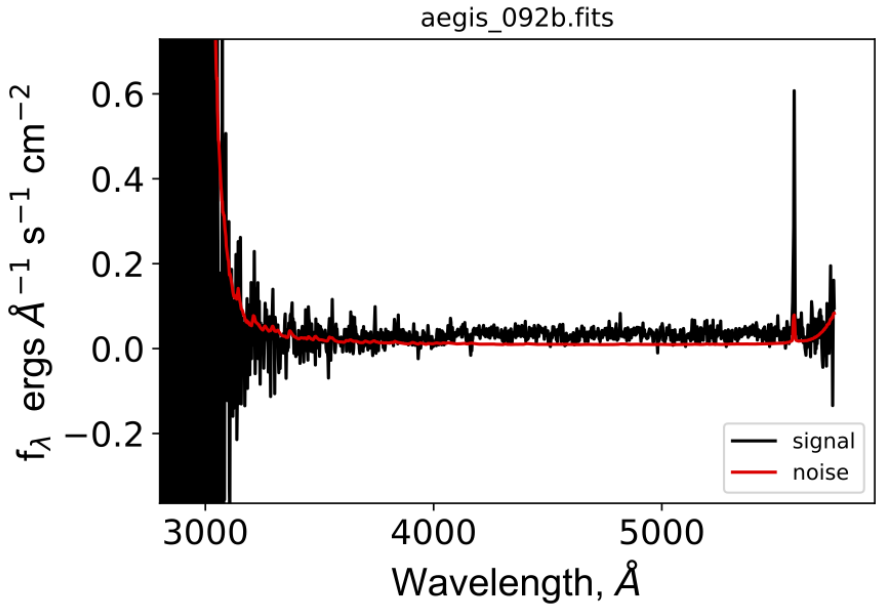

Figure 7.10 *Optical/UV spectrum of source 'aegis_092b', located in the CFHTLS D3 field.*

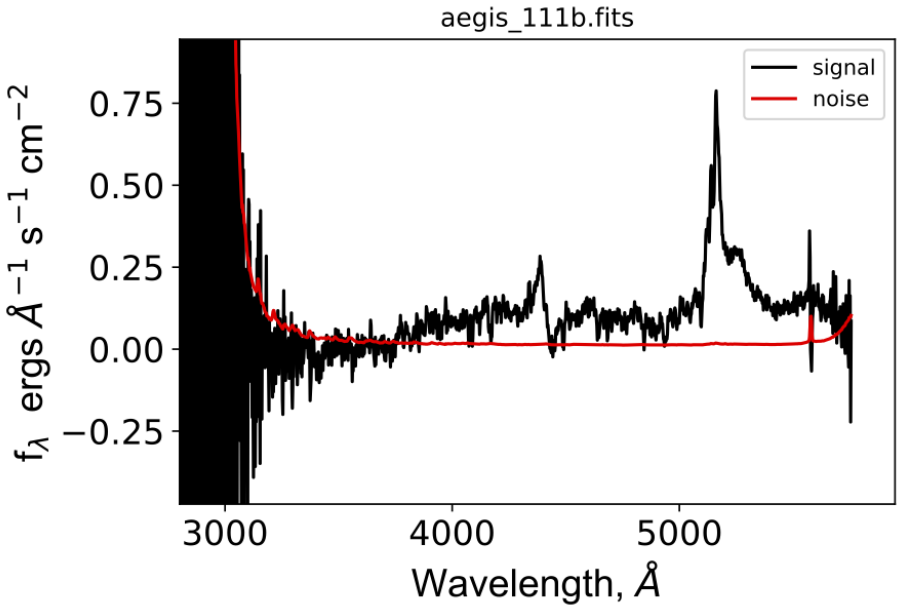

Figure 7.11 *Optical/UV spectrum of source 'aegis_111b', located in the CFHTLS D3 field.*

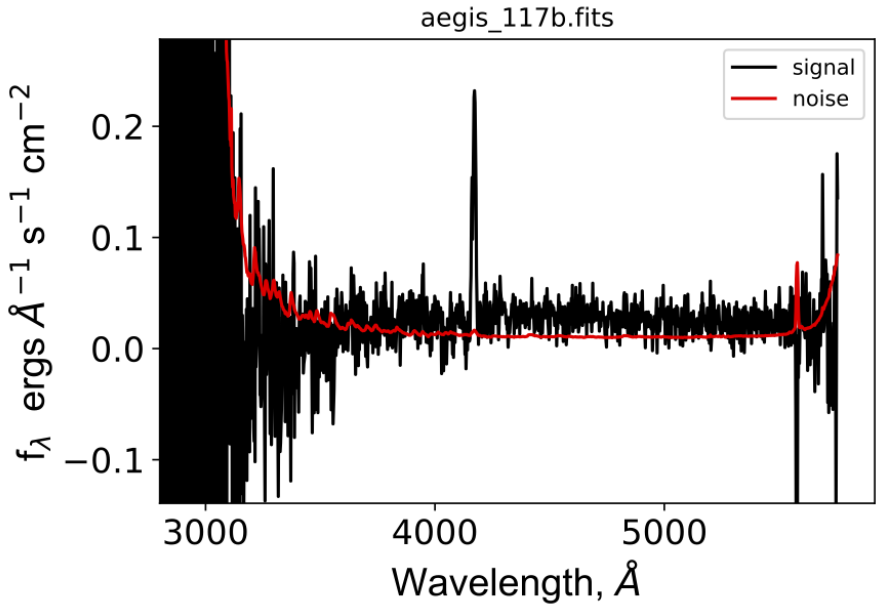

Figure 7.12 *Optical/UV spectrum of source 'aegis_117b', located in the CFHTLS D3 field.*

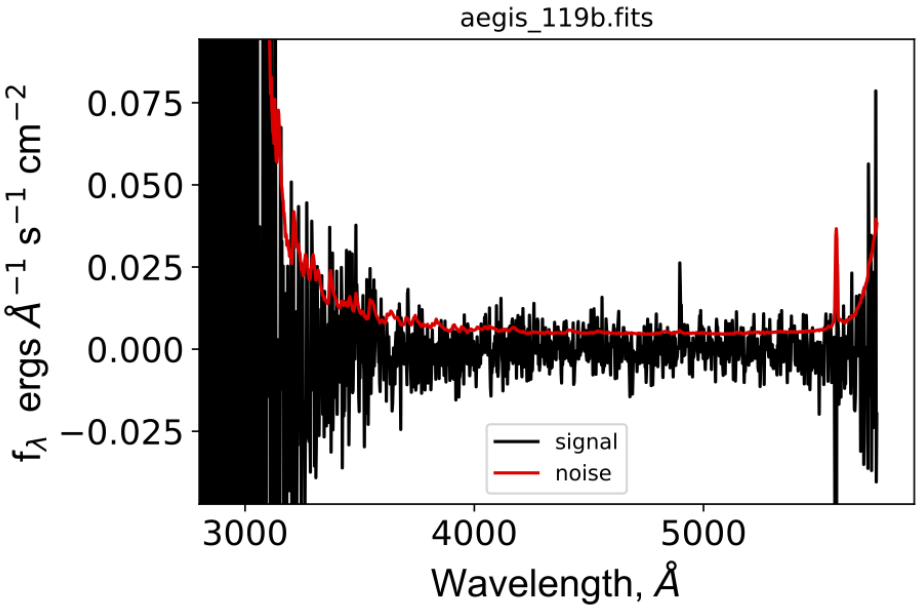

Figure 7.13 *Optical/UV spectrum of source 'aegis_119b', located in the CFHTLS D3 field.*

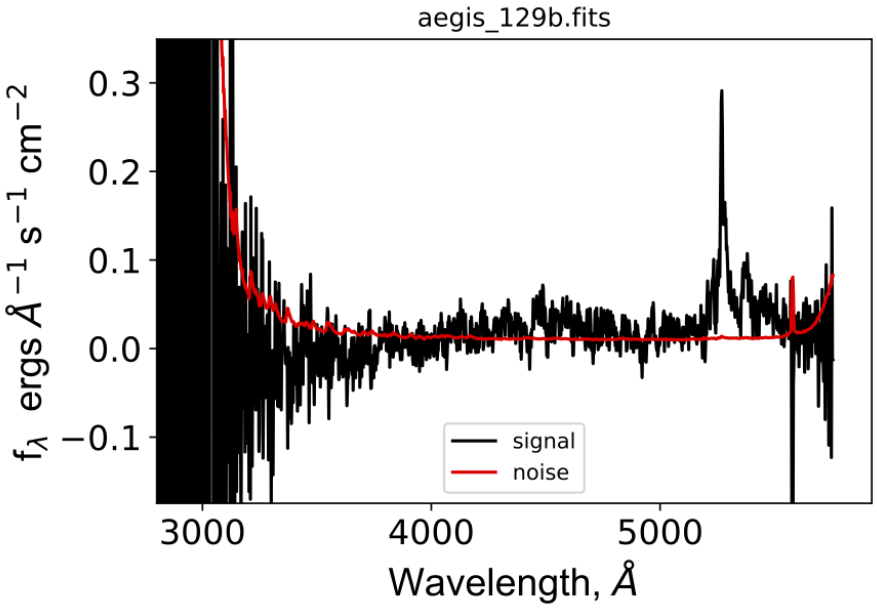

Figure 7.14 *Optical/UV spectrum of source 'aegis_129b', located in the CFHTLS D3 field.*

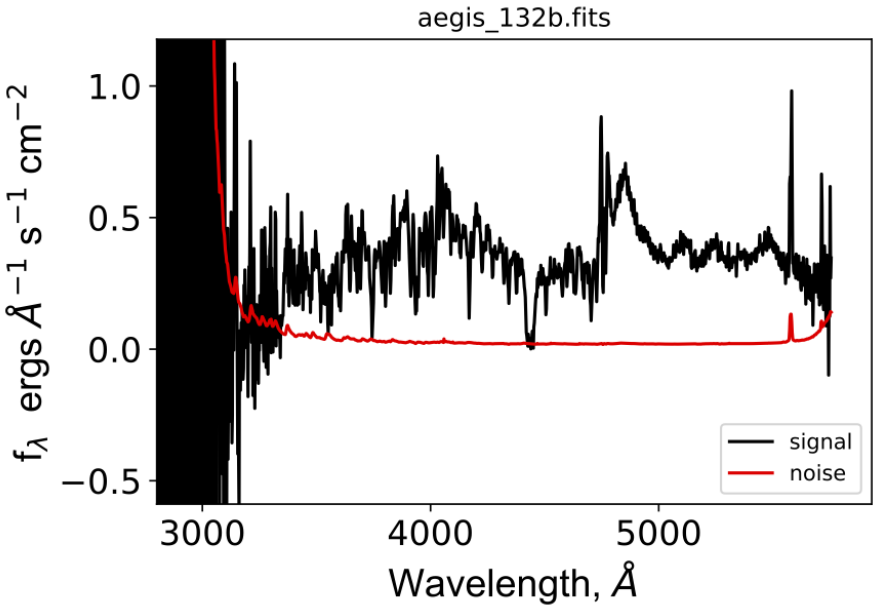

Figure 7.15 *Optical/UV spectrum of source 'aegis_132b', located in the CFHTLS D3 field.*

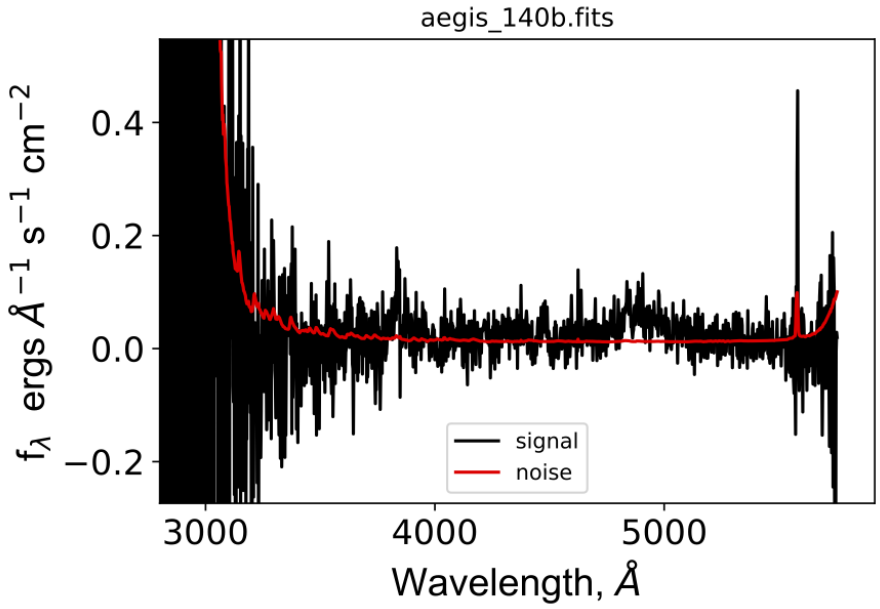

Figure 7.16 *Optical/UV spectrum of source 'aegis_140b', located in the CFHTLS D3 field.*

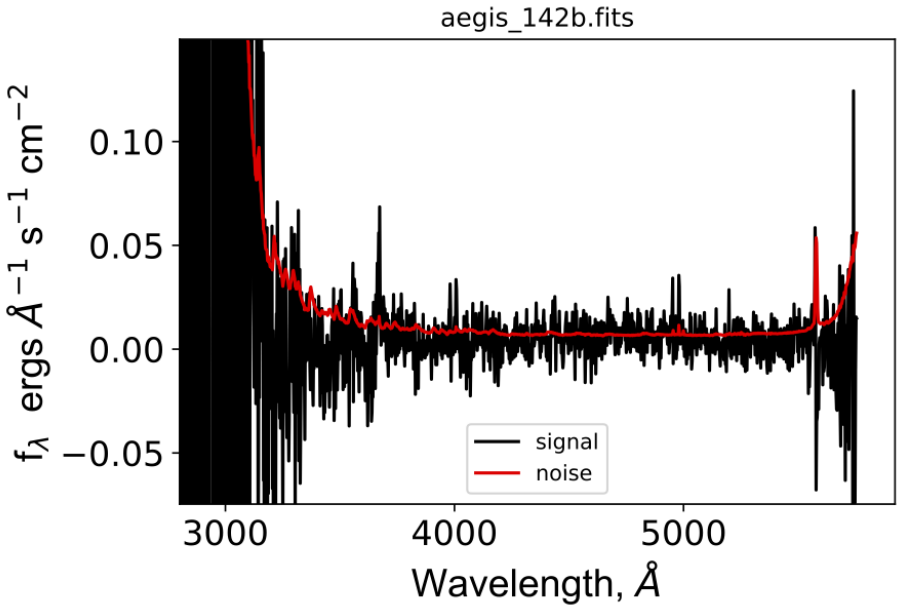

Figure 7.17 *Optical/UV spectrum of source 'aegis_142b', located in the CFHTLS D3 field.*

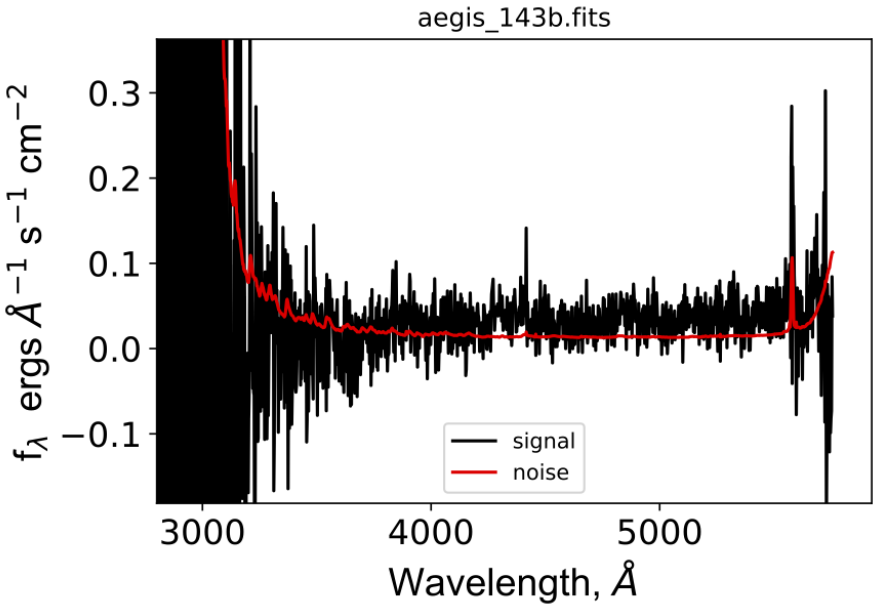

Figure 7.18 *Optical/UV spectrum of source 'aegis_143b', located in the CFHTLS D3 field.*

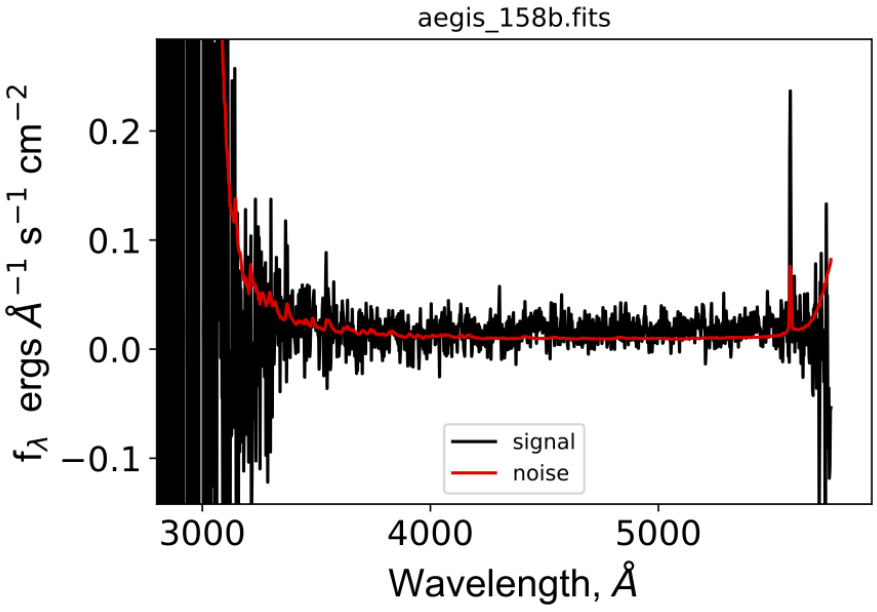

Figure 7.19 *Optical/UV spectrum of source 'aegis_158b', located in the CFHTLS D3 field.*

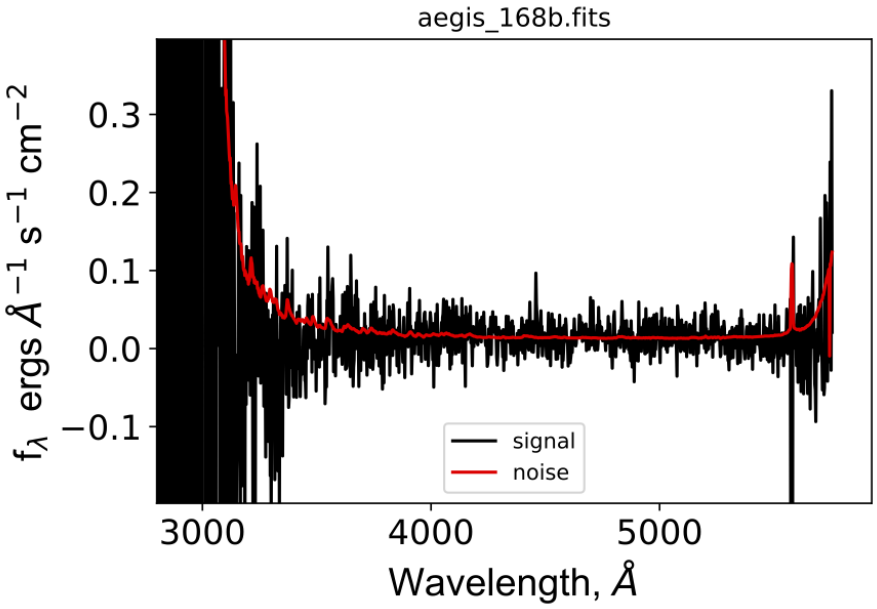

Figure 7.20 *Optical/UV spectrum of source 'aegis_168b', located in the CFHTLS D3 field.*

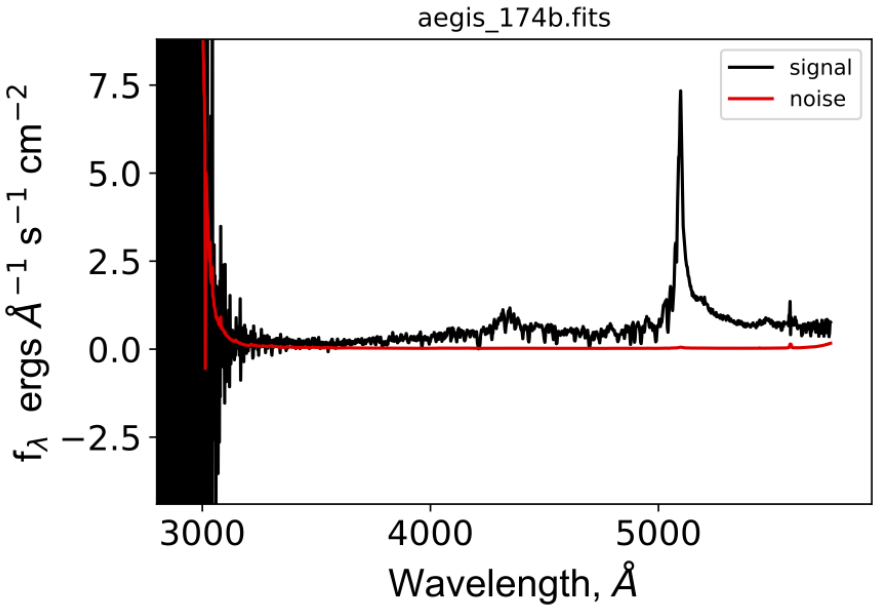

Figure 7.21 *Optical/UV spectrum of source 'aegis_174b', located in the CFHTLS D3 field.*

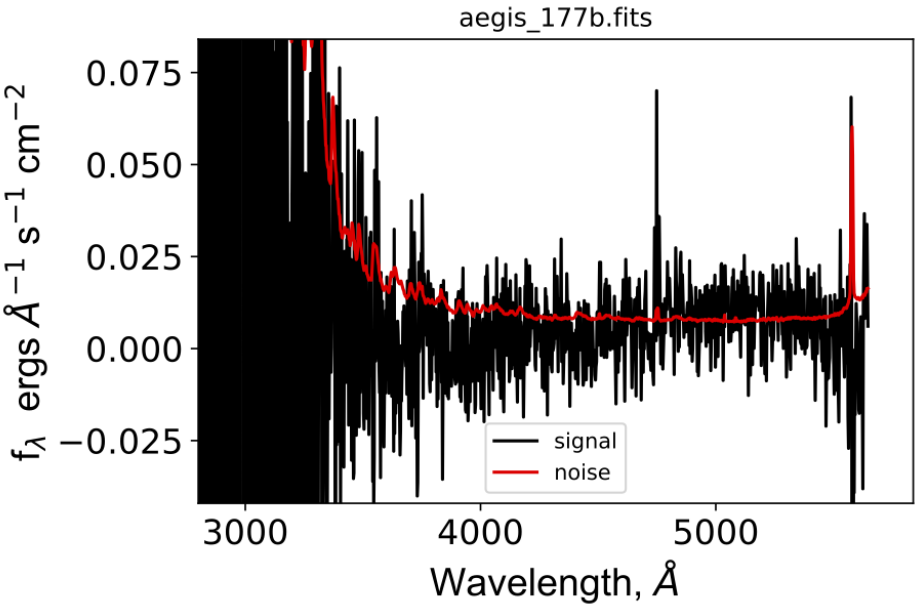

Figure 7.22 *Optical/UV spectrum of source 'aegis_177b', located in the CFHTLS D3 field.*

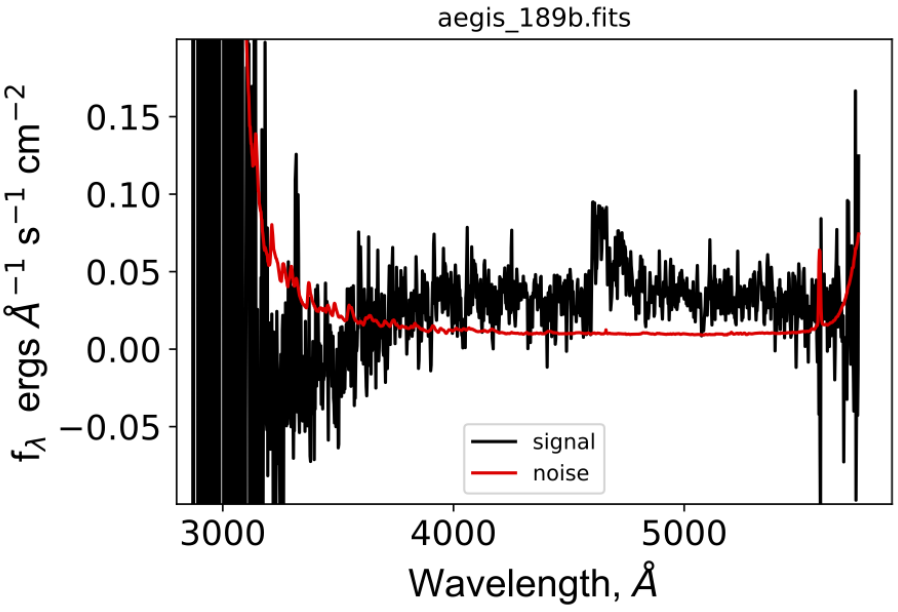

Figure 7.23 *Optical/UV spectrum of source 'aegis_189b', located in the CFHTLS D3 field.*

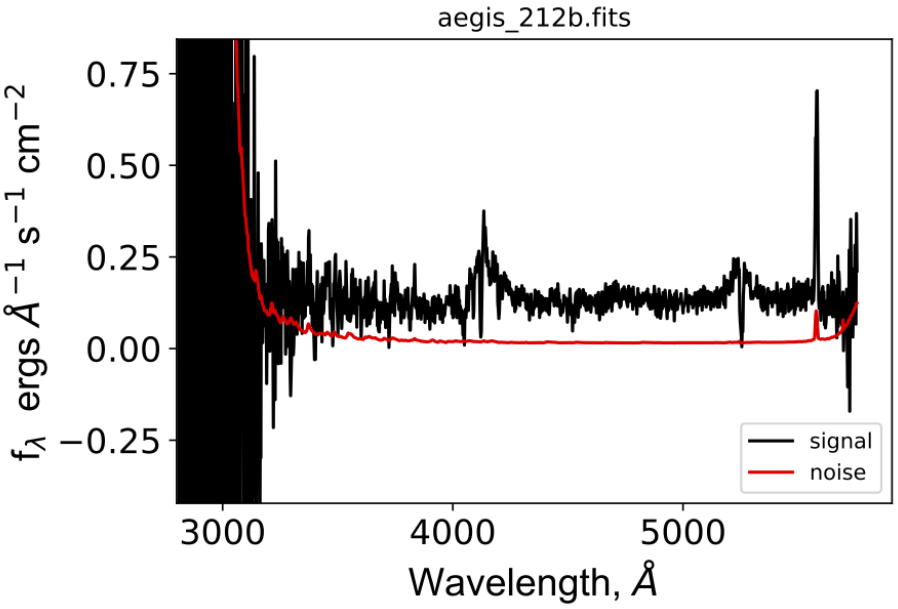

Figure 7.24 *Optical/UV spectrum of source 'aegis_212b', located in the CFHTLS D3 field.*

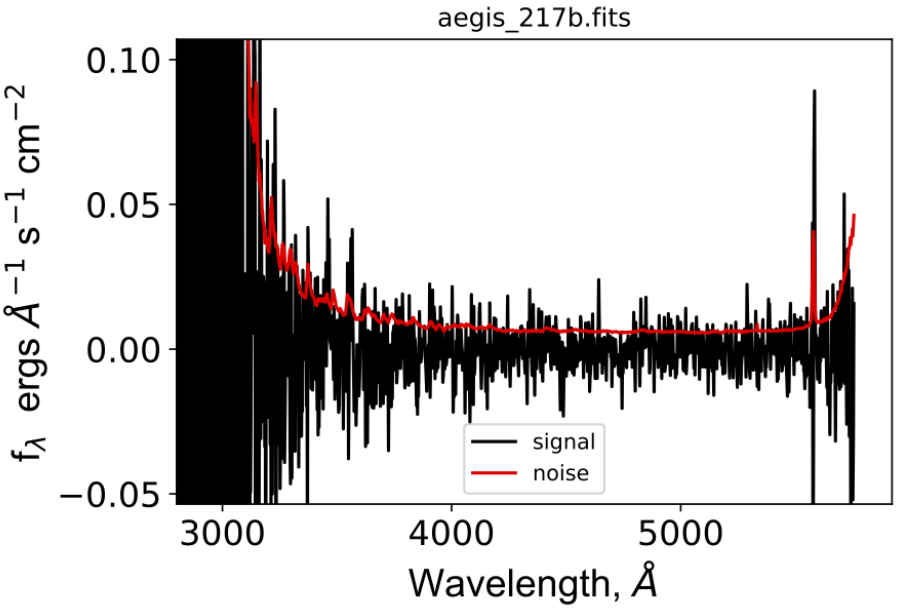

Figure 7.25 *Optical/UV spectrum of source 'aegis_217b', located in the CFHTLS D3 field.*

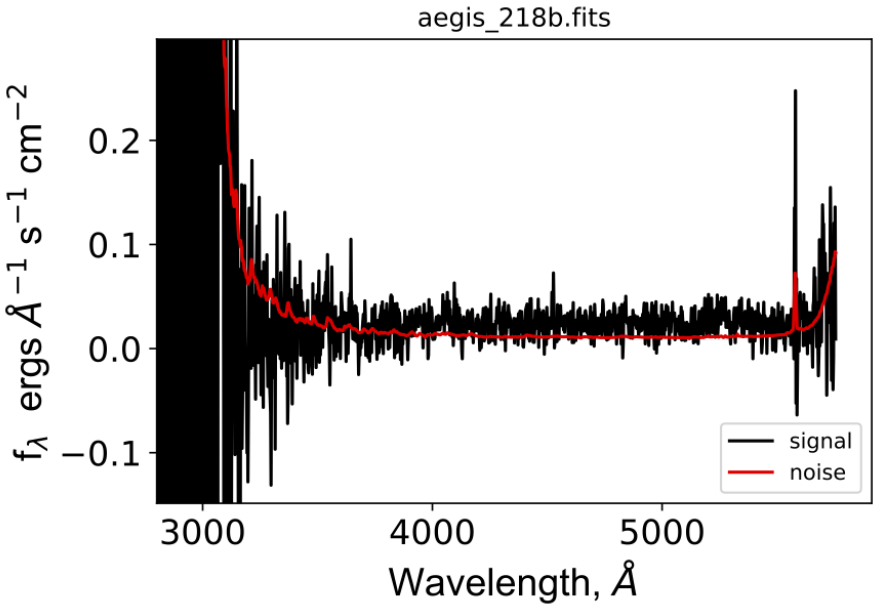

Figure 7.26 *Optical/UV spectrum of source 'aegis_218b', located in the CFHTLS D3 field.*

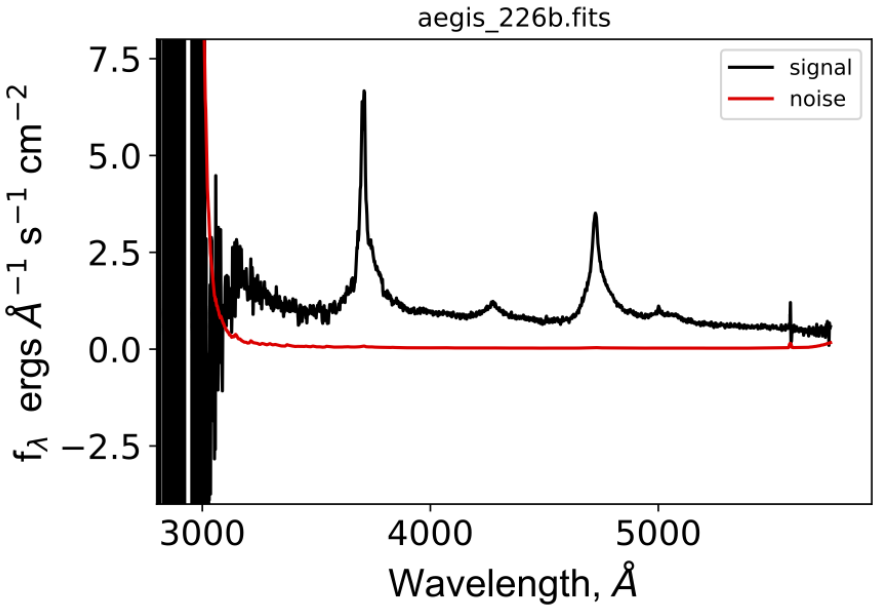

Figure 7.27 *Optical/UV spectrum of source 'aegis_226b', located in the CFHTLS D3 field.*

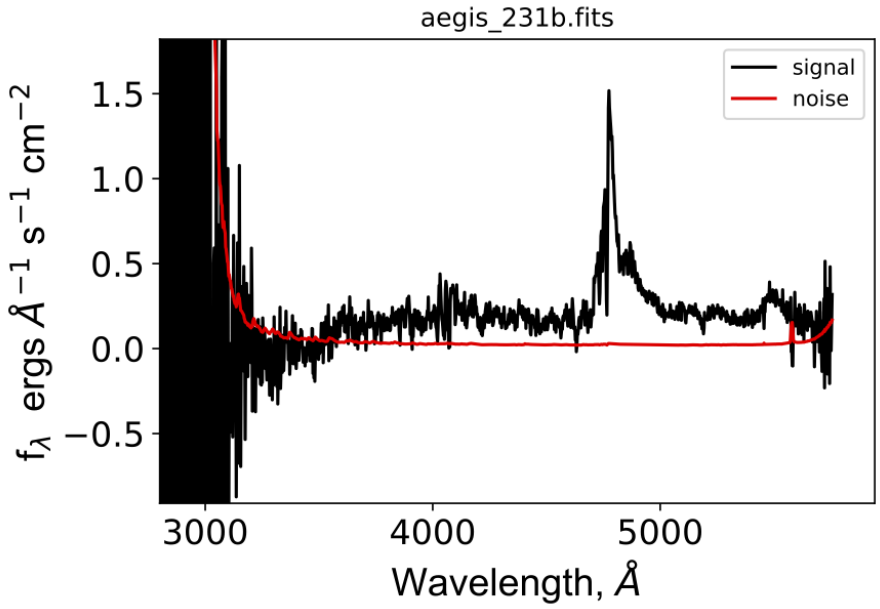

Figure 7.28 *Optical/UV spectrum of source 'aegis_231b', located in the CFHTLS D3 field.*

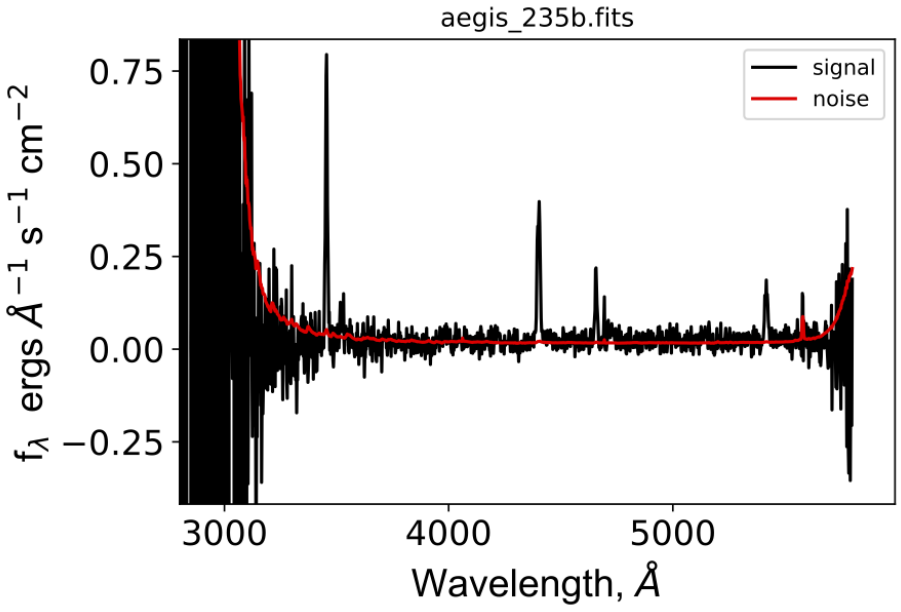

Figure 7.29 *Optical/UV spectrum of source 'aegis_235b', located in the CFHTLS D3 field.*

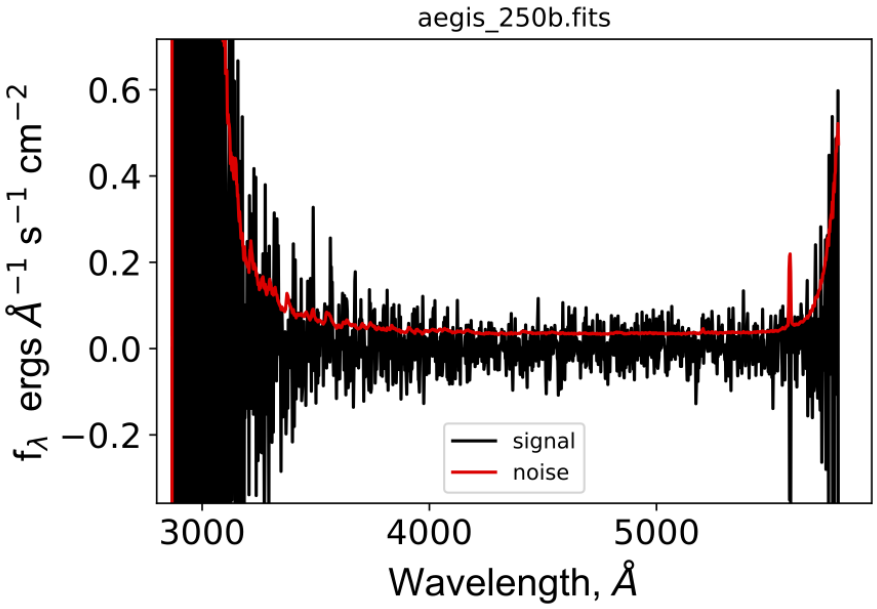

Figure 7.30 *Optical/UV spectrum of source 'aegis_250b', located in the CFHTLS D3 field.*

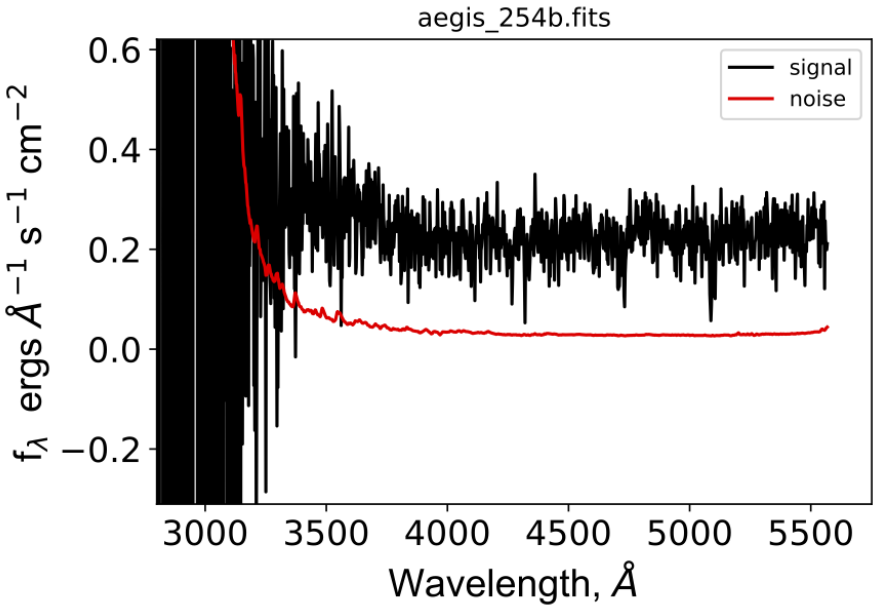

Figure 7.31 *Optical/UV spectrum of source 'aegis_254b', located in the CFHTLS D3 field.*

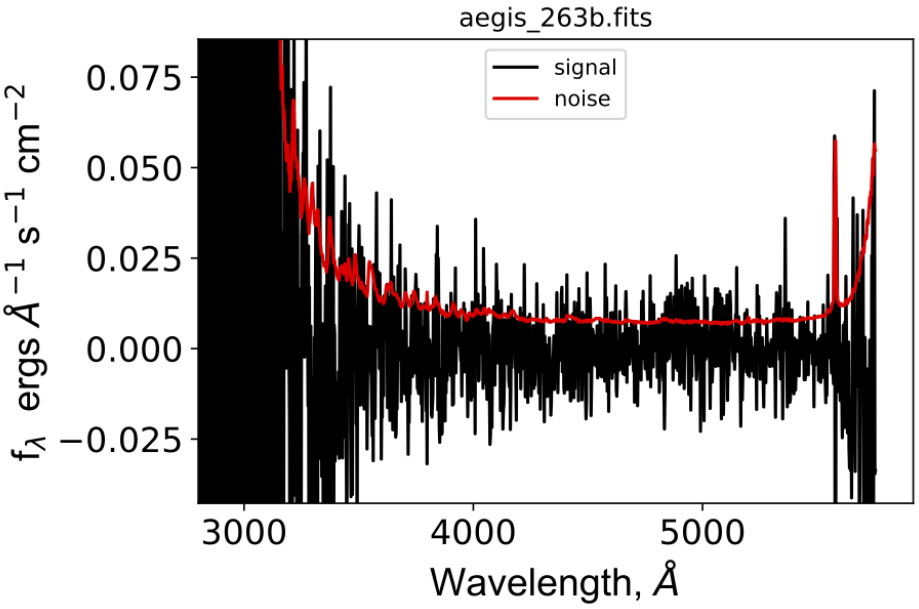

Figure 7.32 *Optical/UV spectrum of source 'aegis_263b', located in the CFHTLS D3 field.*

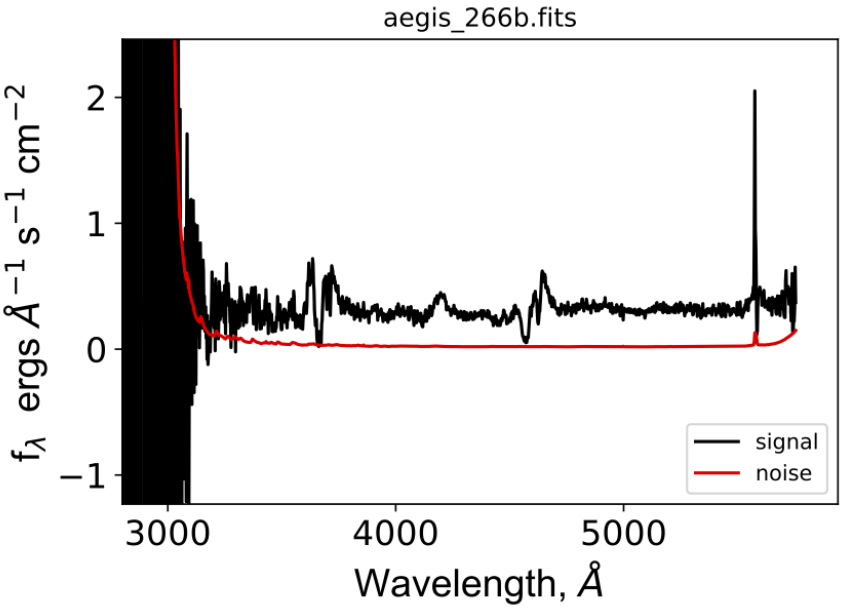

Figure 7.33 *Optical/UV spectrum of source 'aegis_266b', located in the CFHTLS D3 field.*

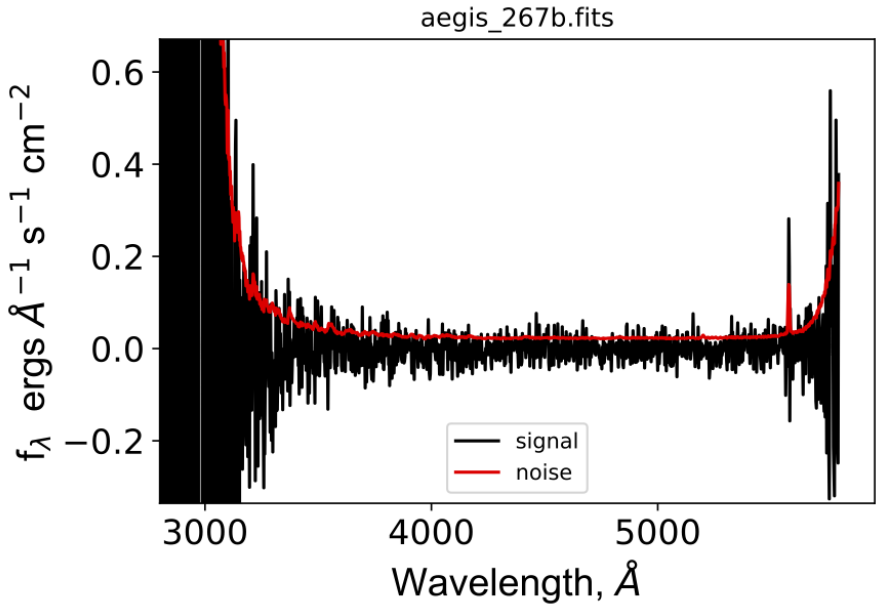

Figure 7.34 *Optical/UV spectrum of source 'aegis_267b', located in the CFHTLS D3 field.*

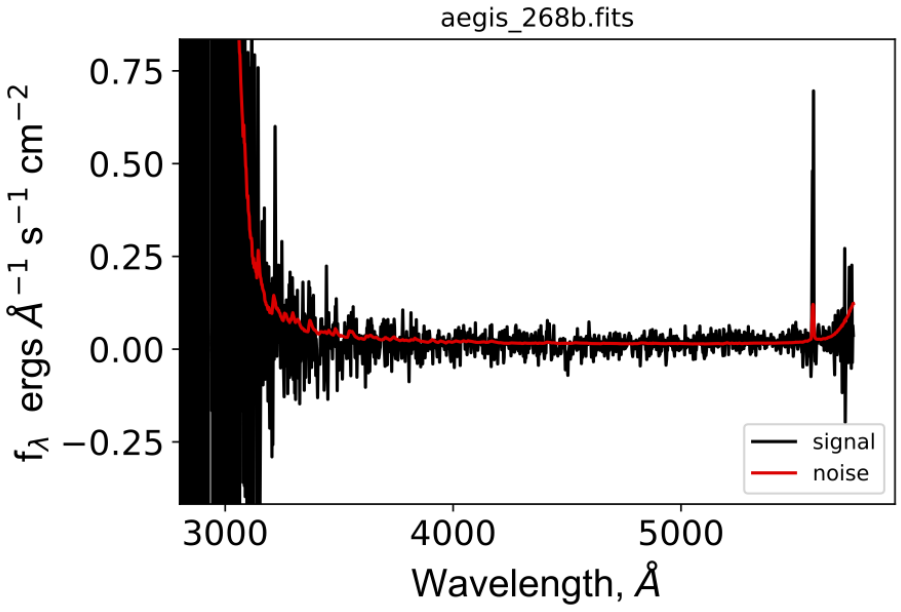

Figure 7.35 *Optical/UV spectrum of source 'aegis_268b', located in the CFHTLS D3 field.*

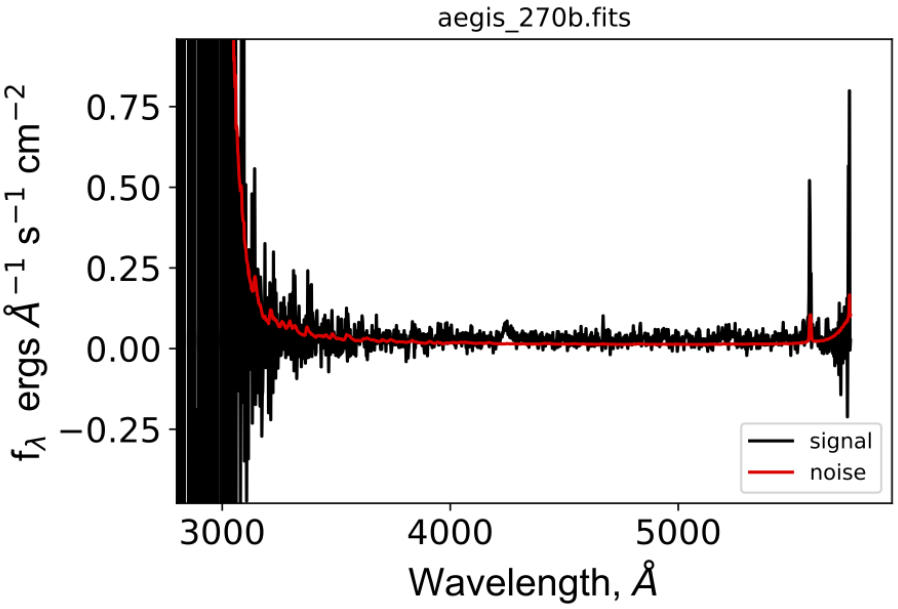

Figure 7.36 *Optical/UV spectrum of source 'aegis_270b', located in the CFHTLS D3 field.*

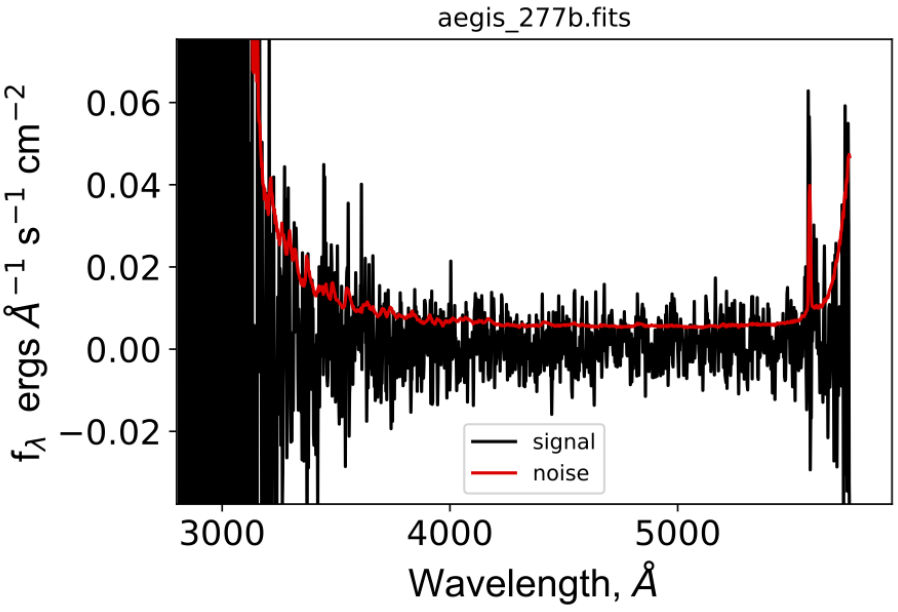

Figure 7.37 *Optical/UV spectrum of source 'aegis_277b', located in the CFHTLS D3 field.*

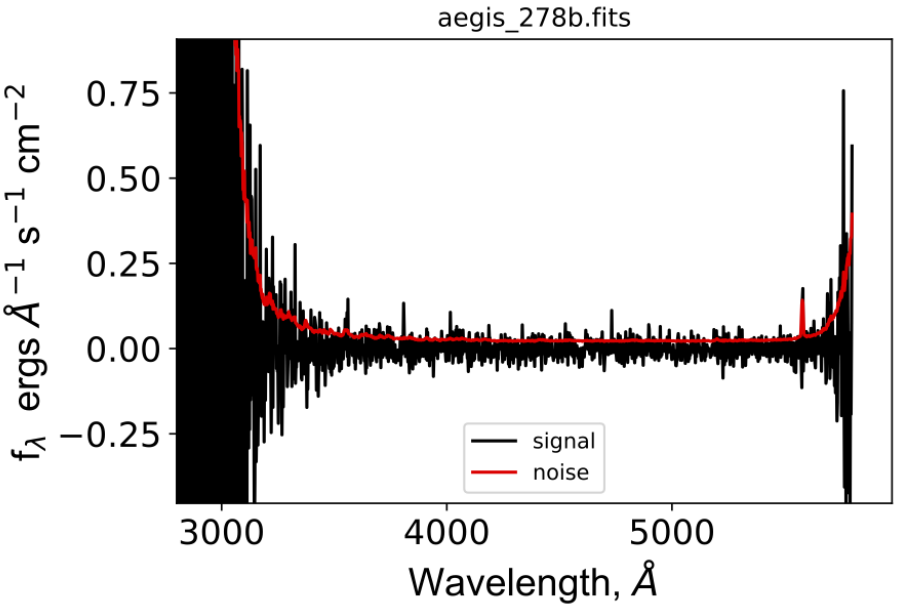

Figure 7.38 *Optical/UV spectrum of source 'aegis_278b', located in the CFHTLS D3 field.*

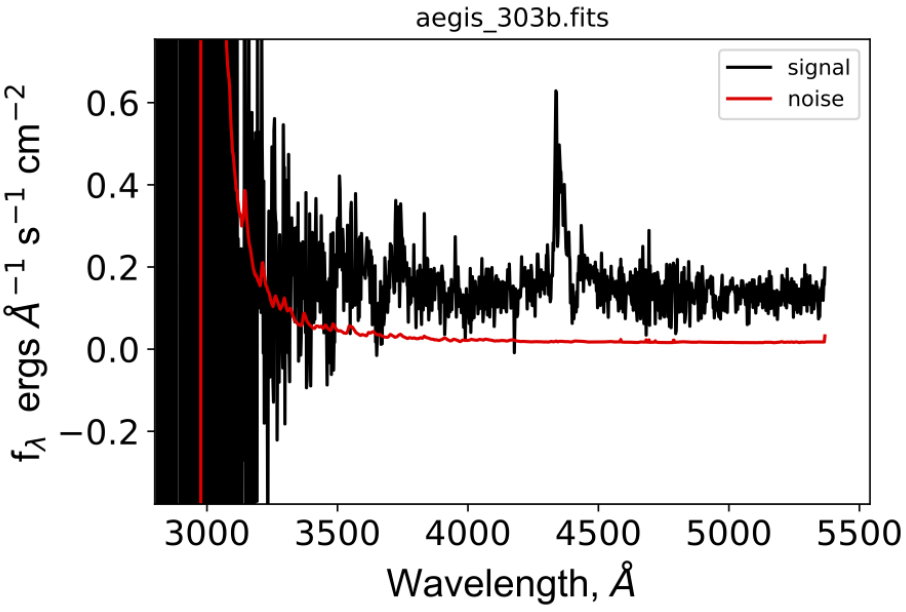

Figure 7.39 *Optical/UV spectrum of source 'aegis_303b', located in the CFHTLS D3 field.*

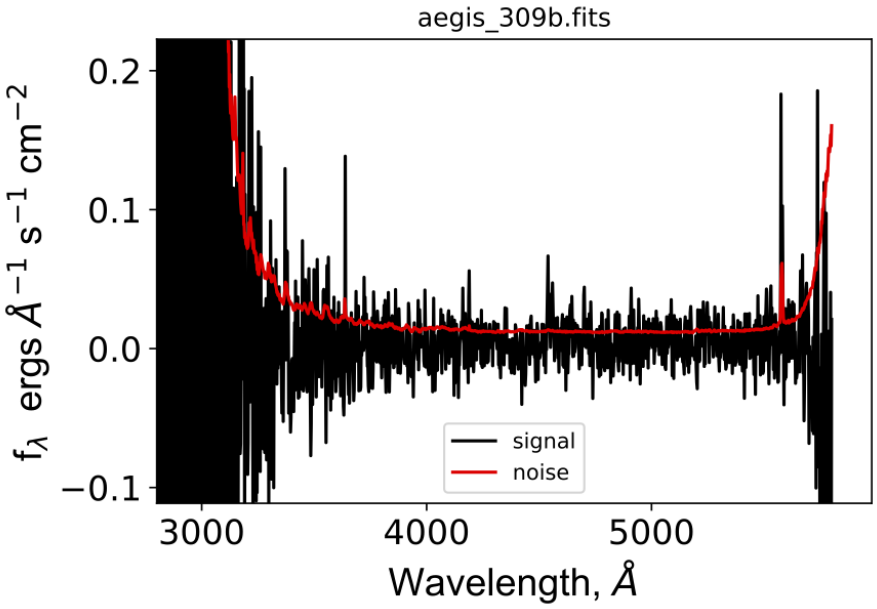

Figure 7.40 *Optical/UV spectrum of source 'aegis_309b', located in the CFHTLS D3 field.*

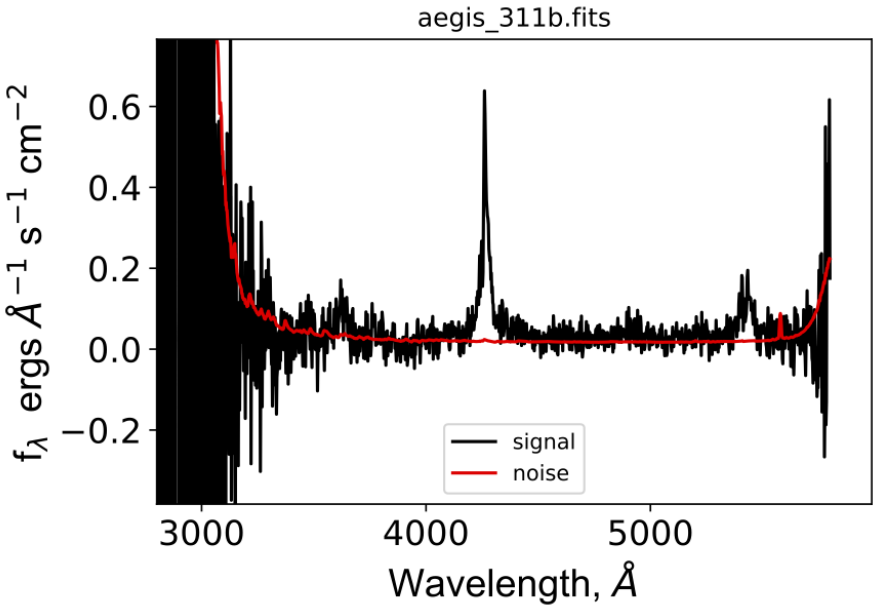

Figure 7.41 *Optical/UV spectrum of source 'aegis_311b', located in the CFHTLS D3 field.*

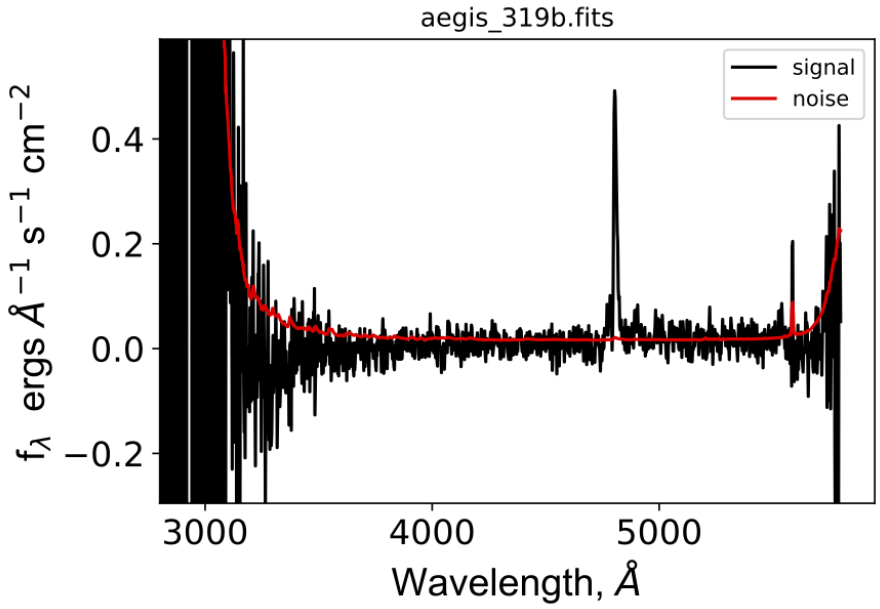

Figure 7.42 *Optical/UV spectrum of source 'aegis_319b', located in the CFHTLS D3 field.*

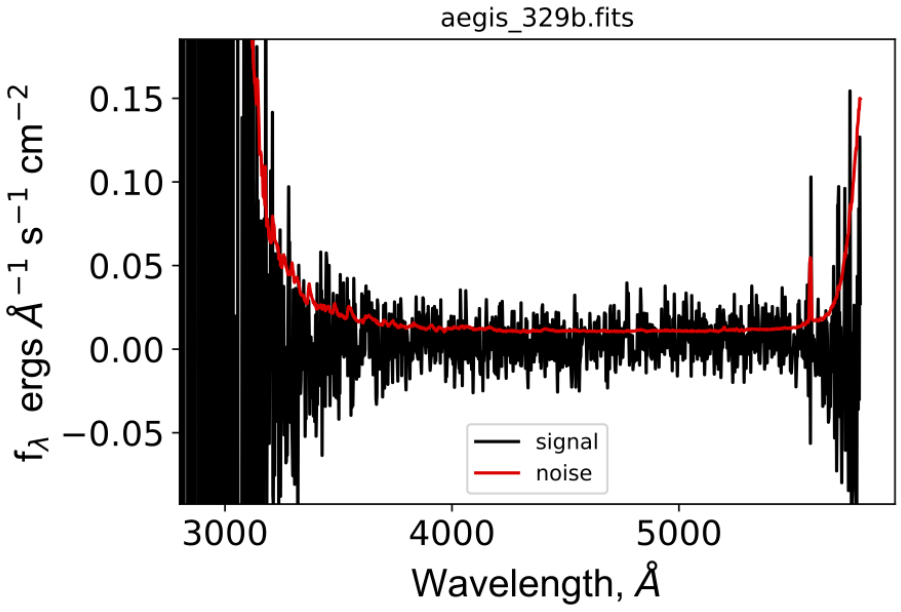

Figure 7.43 *Optical/UV spectrum of source 'aegis_329b', located in the CFHTLS D3 field.*

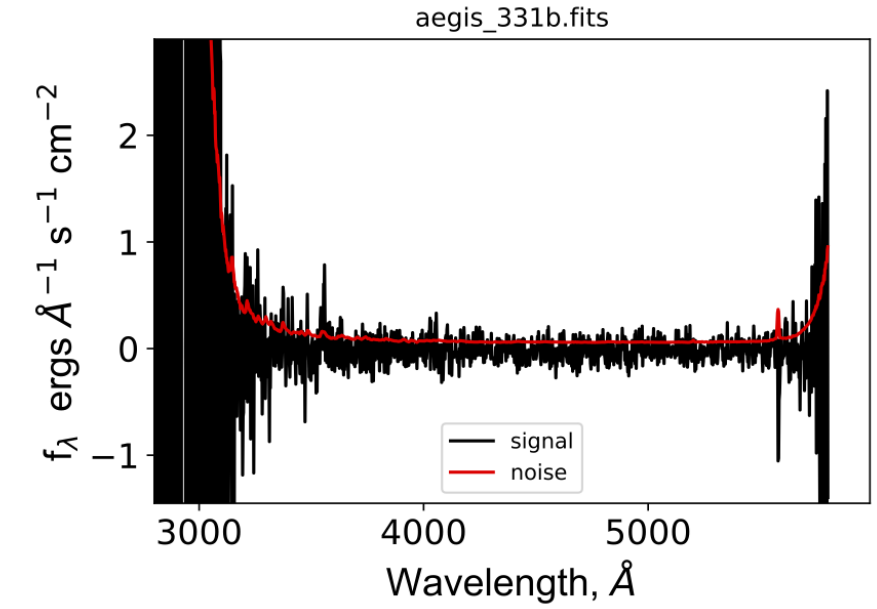

Figure 7.44 *Optical/UV spectrum of source 'aegis_331b', located in the CFHTLS D3 field.*

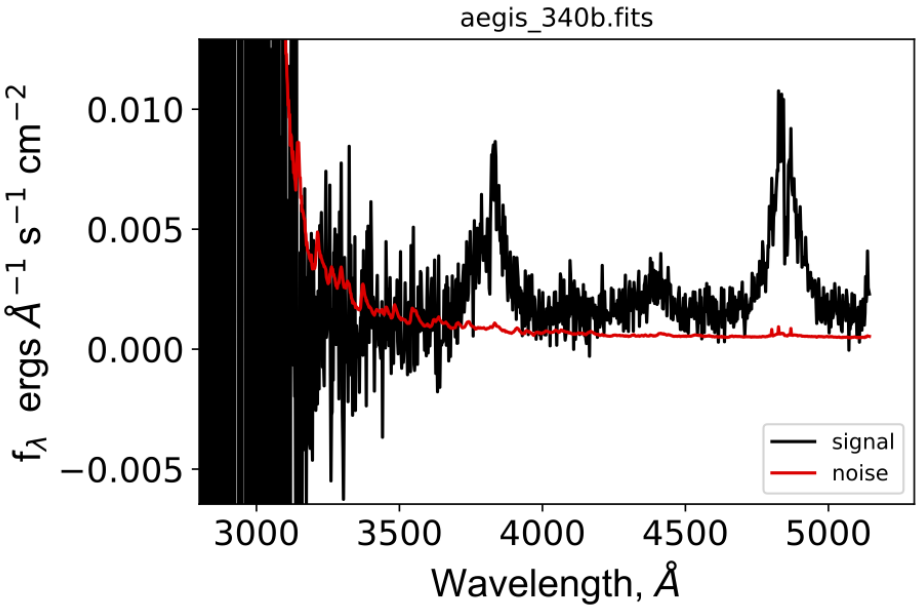

Figure 7.45 *Optical/UV spectrum of source 'aegis_340b', located in the CFHTLS D3 field.*

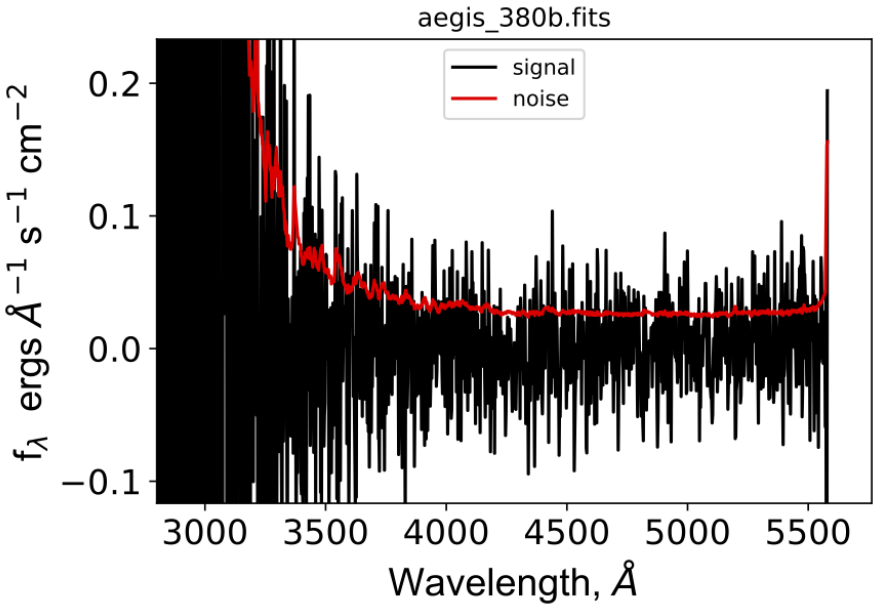

Figure 7.46 *Optical/UV spectrum of source 'aegis_380b', located in the CFHTLS D3 field.*

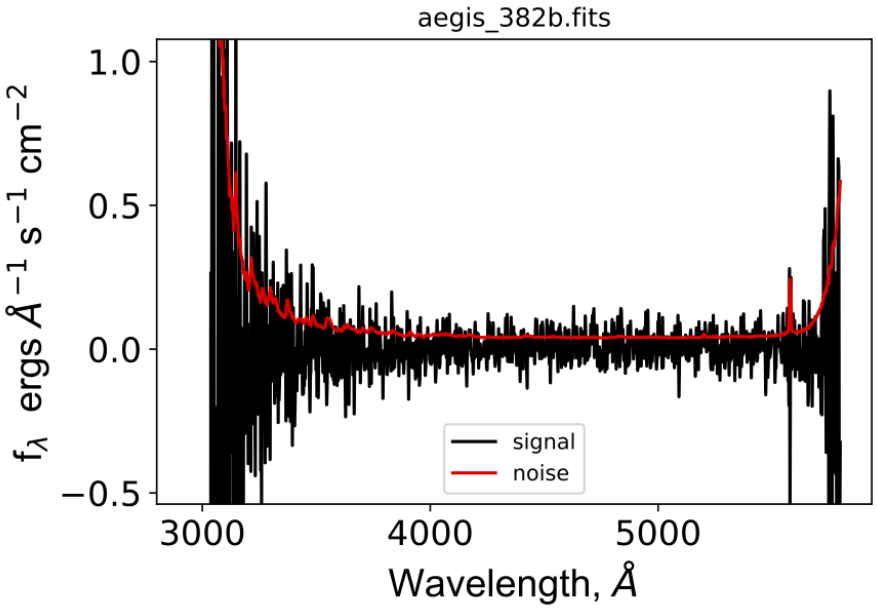

Figure 7.47 *Optical/UV spectrum of source 'aegis_382b', located in the CFHTLS D3 field.*

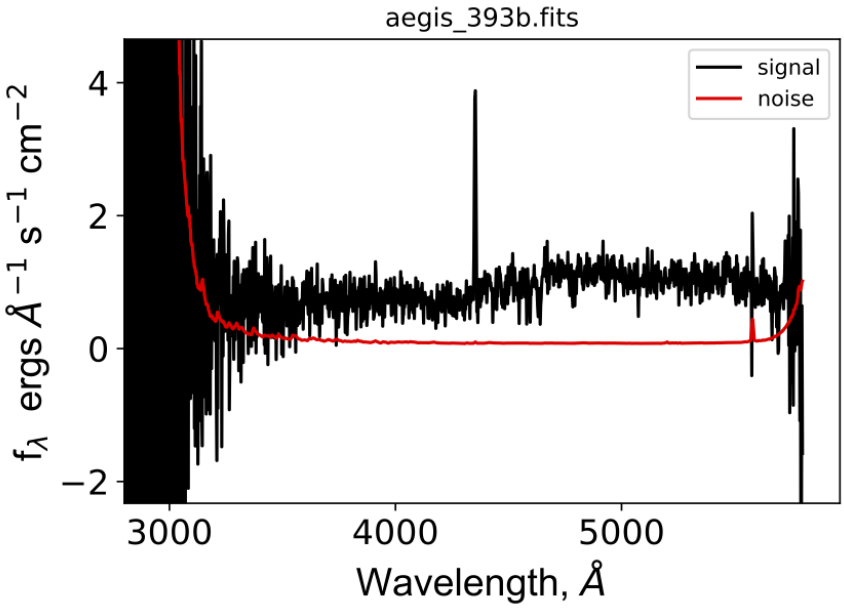

Figure 7.48 *Optical/UV spectrum of source 'aegis_393b', located in the CFHTLS D3 field.*

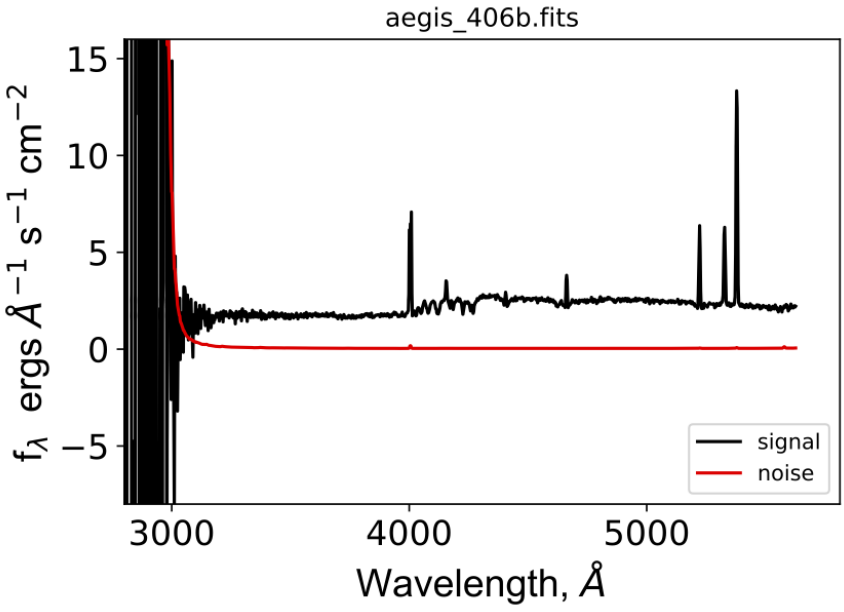

Figure 7.49 *Optical/UV spectrum of source 'aegis_406b', located in the CFHTLS D3 field.*

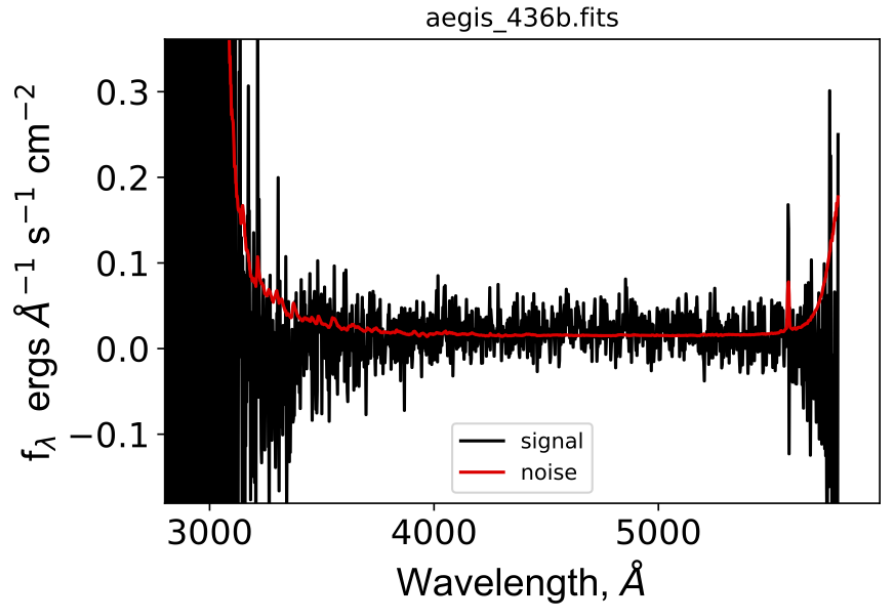

Figure 7.50 *Optical/UV spectrum of source 'aegis_436b', located in the CFHTLS D3 field.*

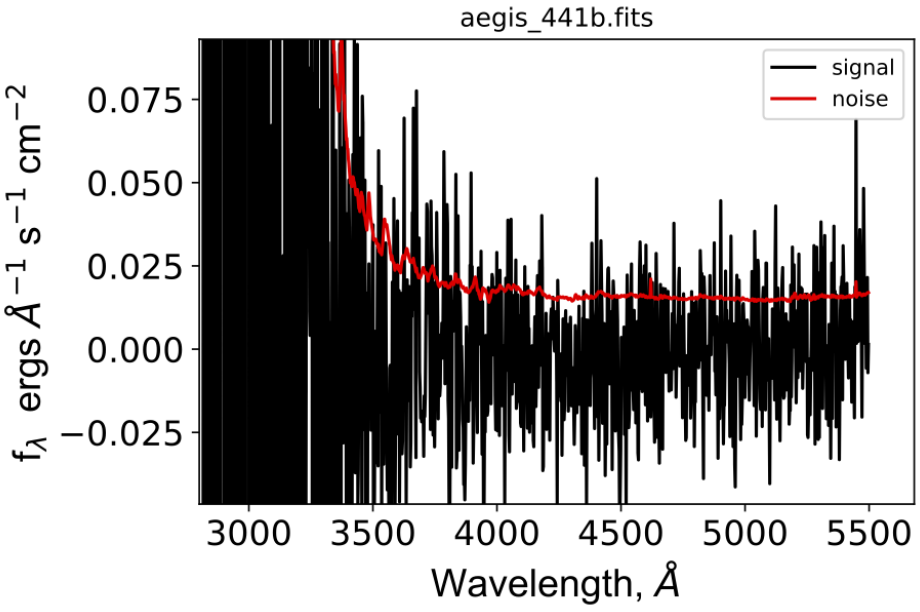

Figure 7.51 *Optical/UV spectrum of source 'aegis_441b', located in the CFHTLS D3 field.*

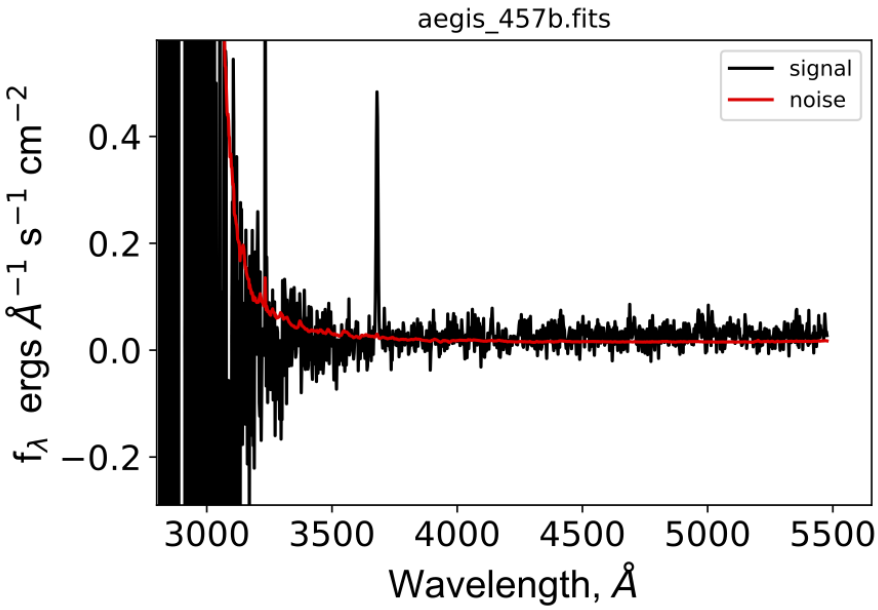

Figure 7.53 *Optical/UV spectrum of source 'aegis_457b', located in the CFHTLS D3 field.*

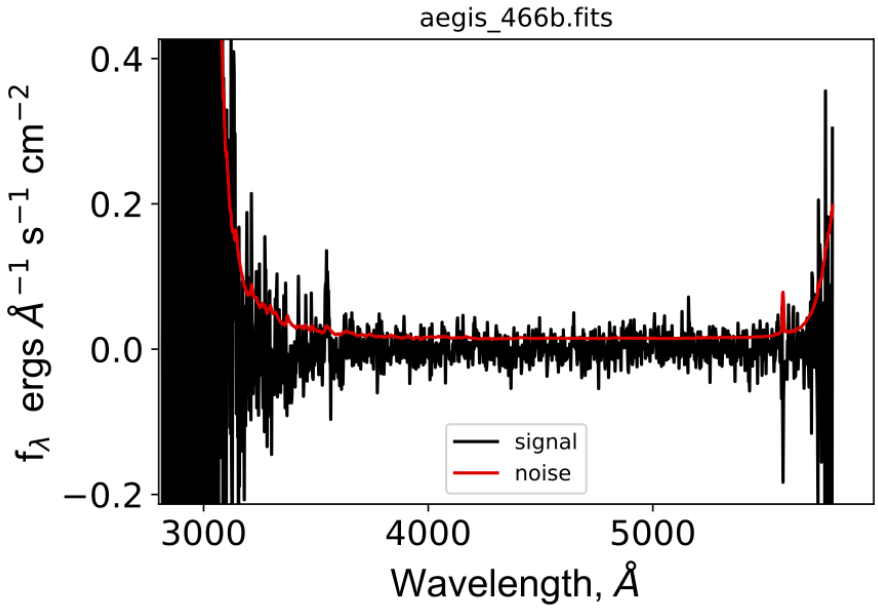

Figure 7.54 *Optical/UV spectrum of source 'aegis_466b', located in the CFHTLS D3 field.*

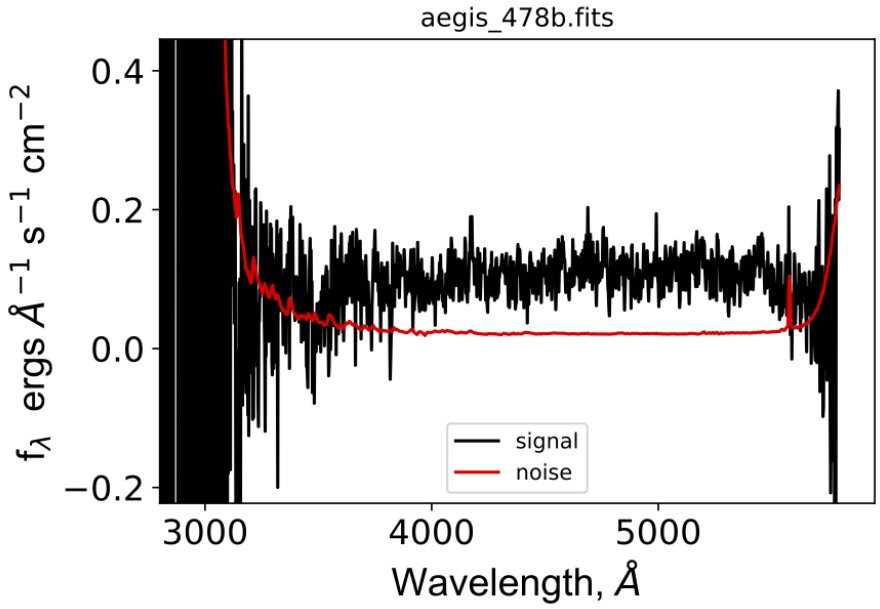

Figure 7.55 *Optical/UV spectrum of source 'aegis_478b', located in the CFHTLS D3 field.*

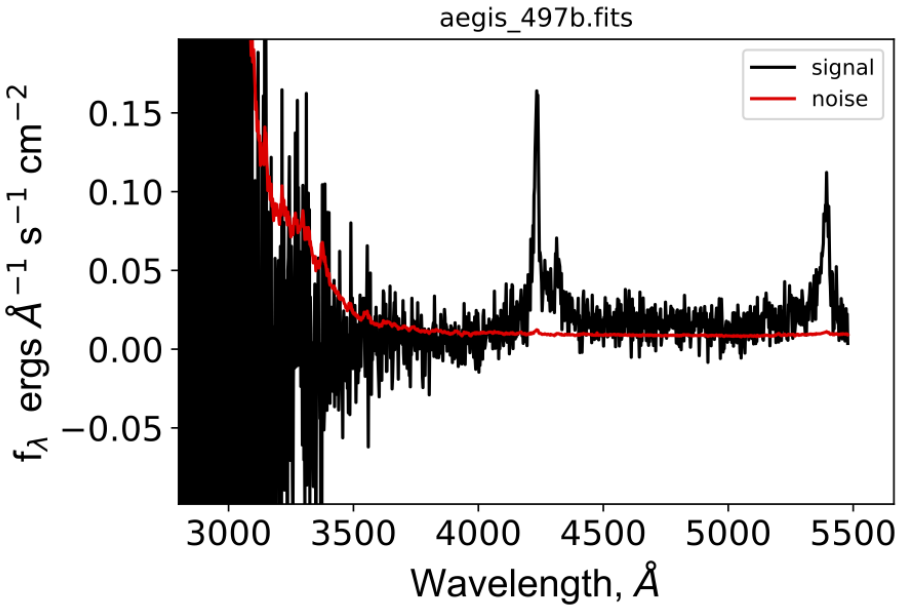

Figure 7.56 *Optical/UV spectrum of source 'aegis_497b', located in the CFHTLS D3 field.*

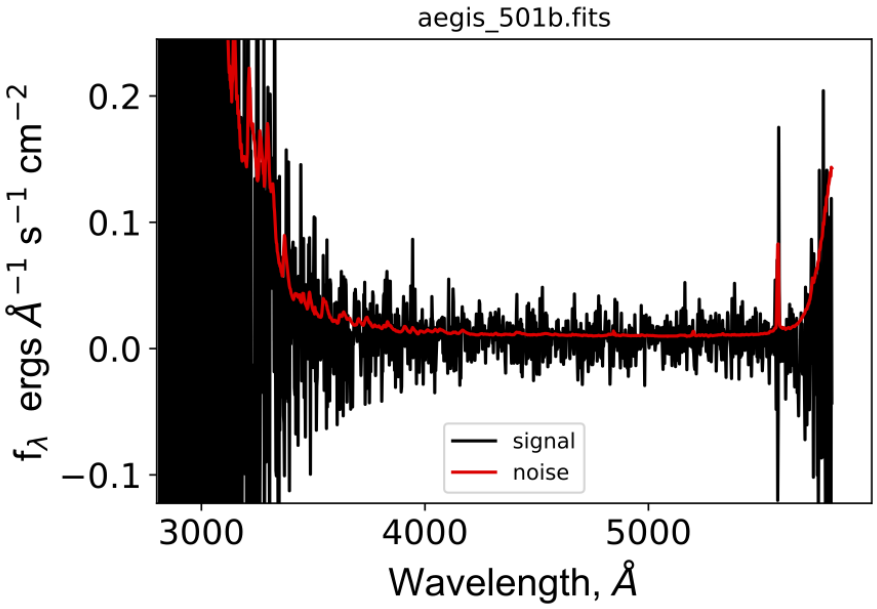

Figure 7.57 *Optical/UV spectrum of source 'aegis_501b', located in the CFHTLS D3 field.*

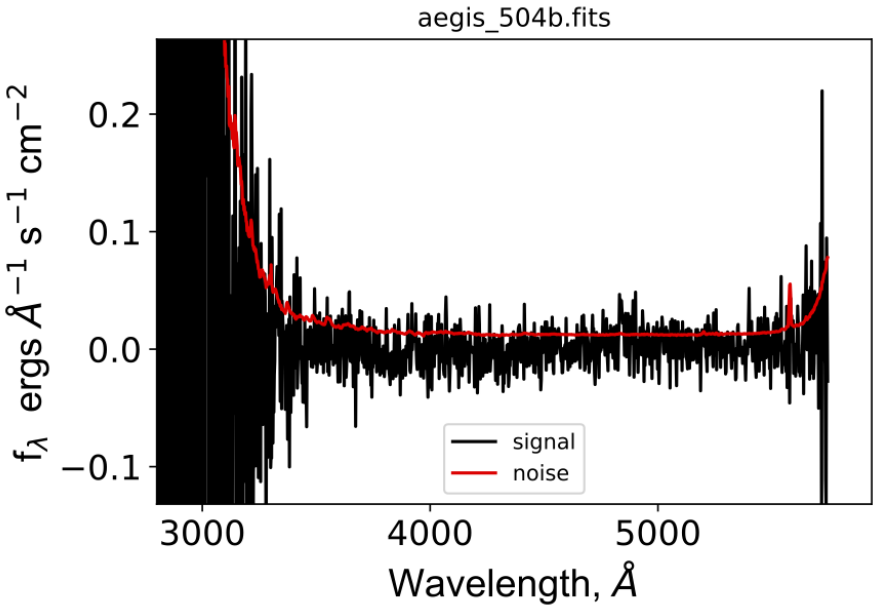

Figure 7.58 *Optical/UV spectrum of source 'aegis_504b', located in the CFHTLS D3 field.*

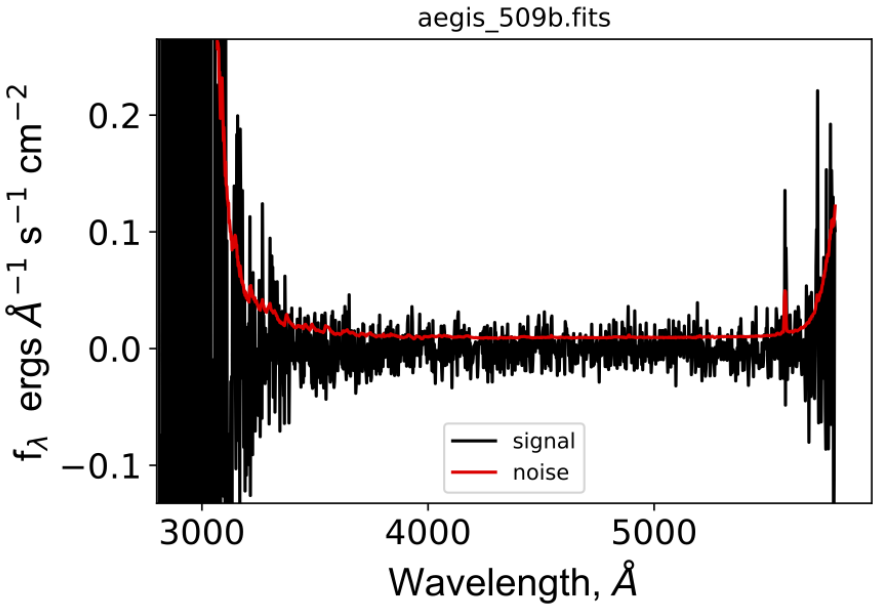

Figure 7.59 *Optical/UV spectrum of source 'aegis_509b', located in the CFHTLS D3 field.*

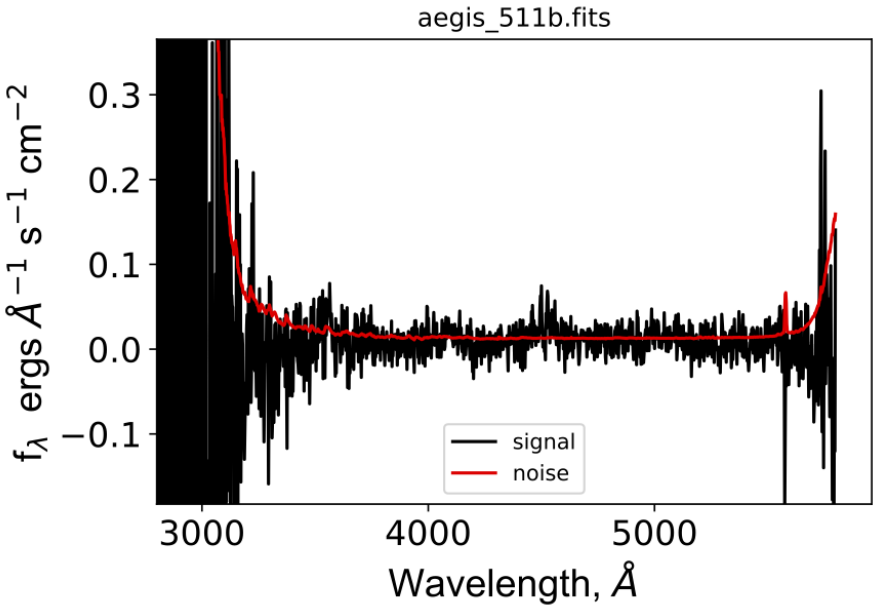

Figure 7.60 *Optical/UV spectrum of source 'aegis_511b', located in the CFHTLS D3 field.*

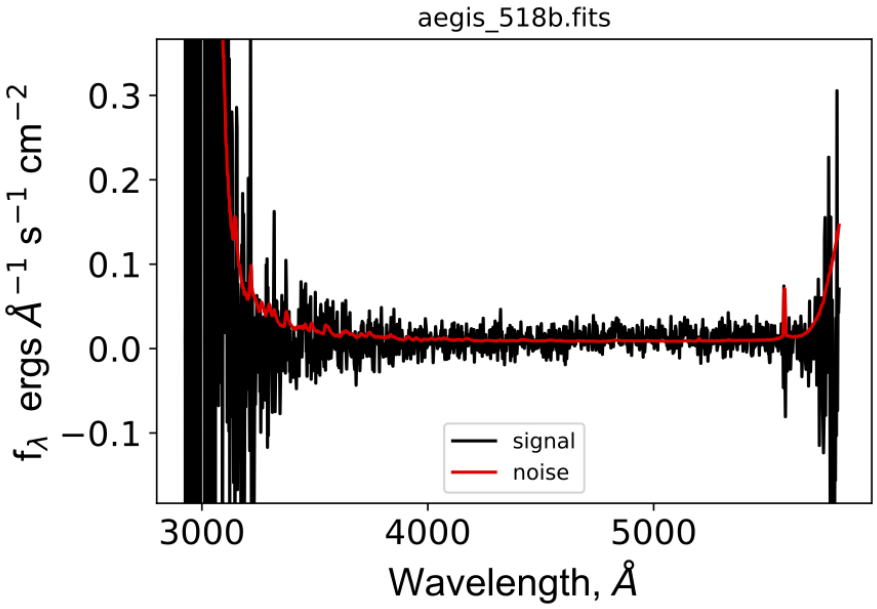

Figure 7.61 *Optical/UV spectrum of source 'aegis_518b', located in the CFHTLS D3 field.*

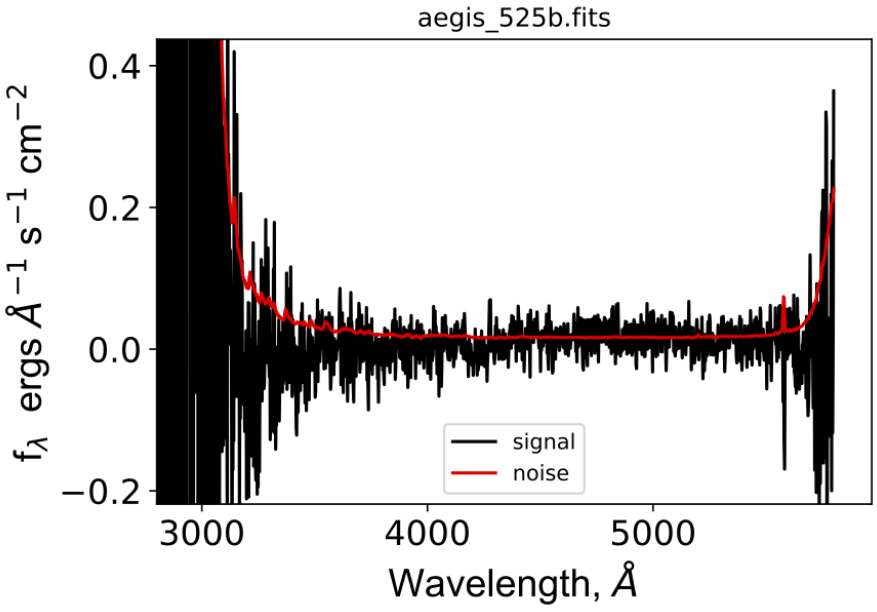

Figure 7.62 *Optical/UV spectrum of source 'aegis_525b', located in the CFHTLS D3 field.*

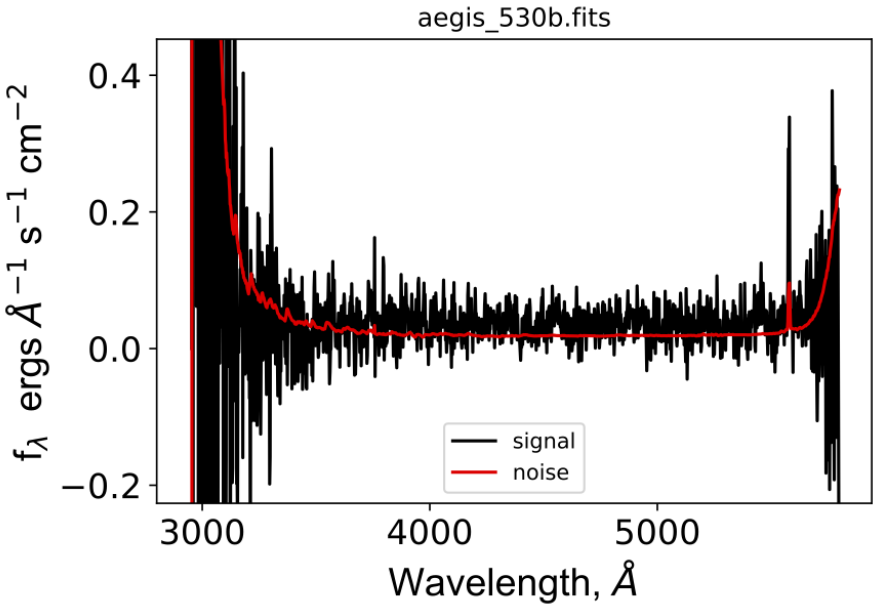

Figure 7.63 *Optical/UV spectrum of source 'aegis_530b', located in the CFHTLS D3 field.*

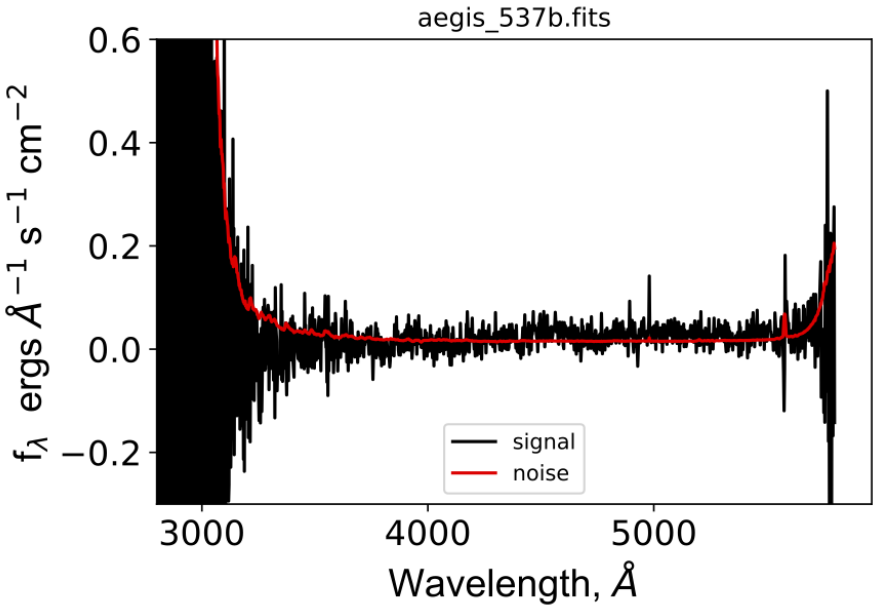

Figure 7.64 *Optical/UV spectrum of source 'aegis_537b', located in the CFHTLS D3 field.*

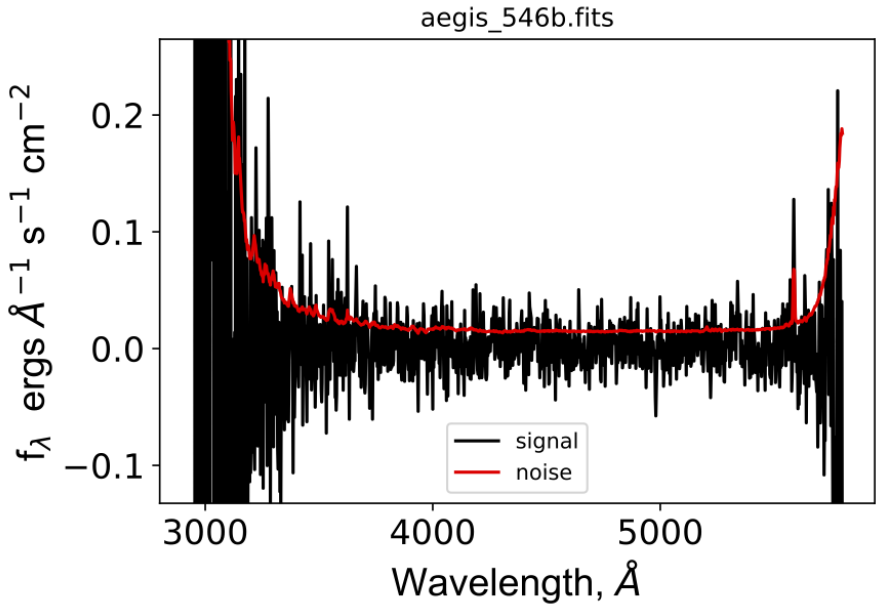

Figure 7.65 *Optical/UV spectrum of source 'aegis_546b', located in the CFHTLS D3 field.*

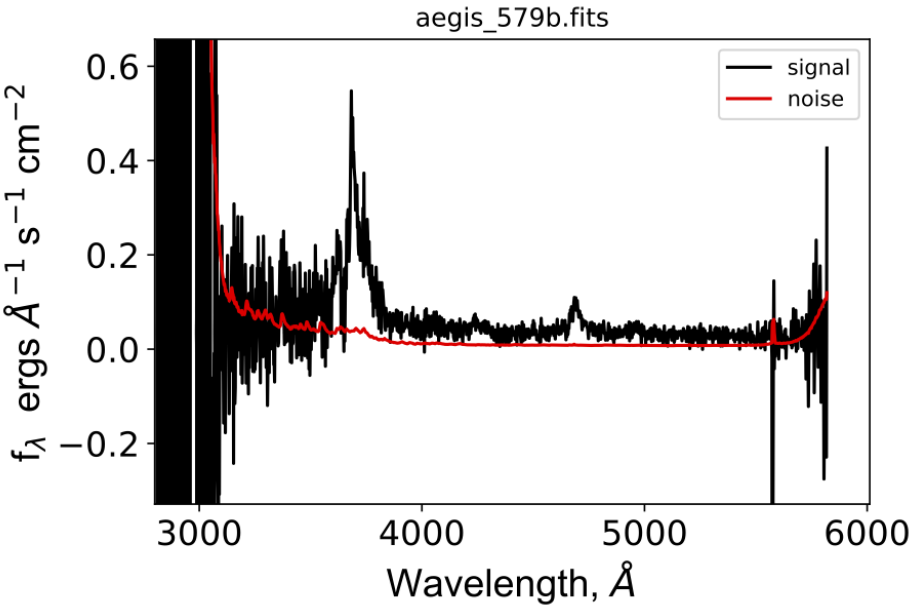

Figure 7.66 *Optical/UV spectrum of source 'aegis_579b', located in the CFHTLS D3 field.*

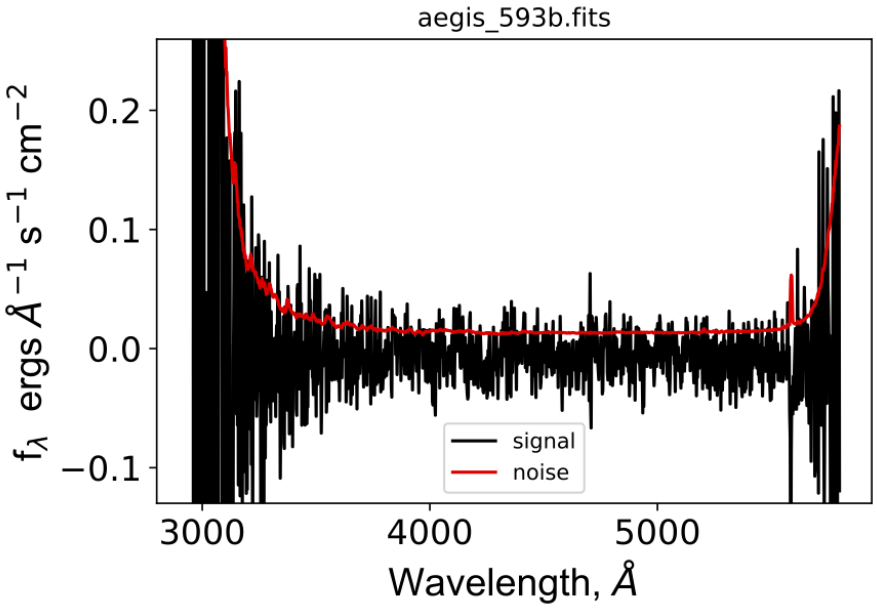

Figure 7.67 *Optical/UV spectrum of source 'aegis_593b', located in the CFHTLS D3 field.*

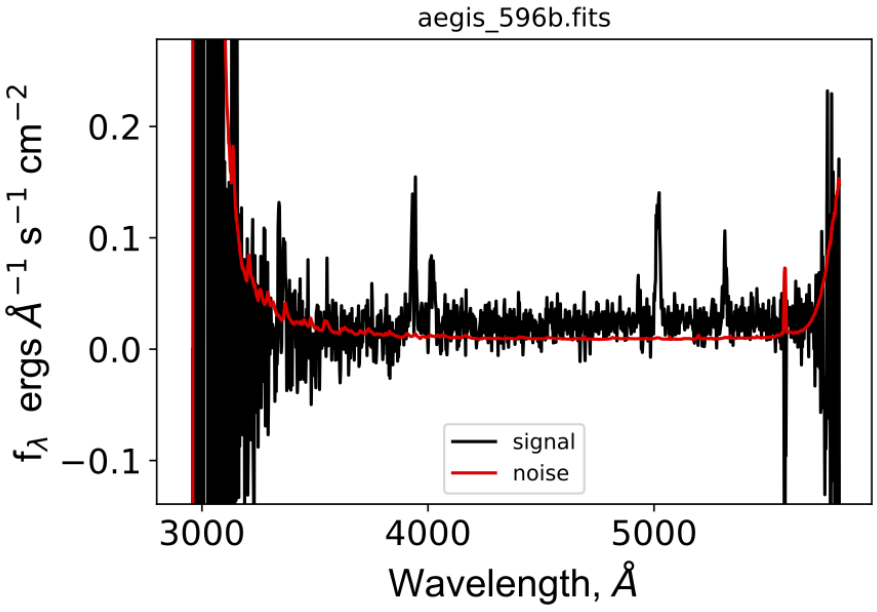

Figure 7.68 *Optical/UV spectrum of source 'aegis_596b', located in the CFHTLS D3 field.*

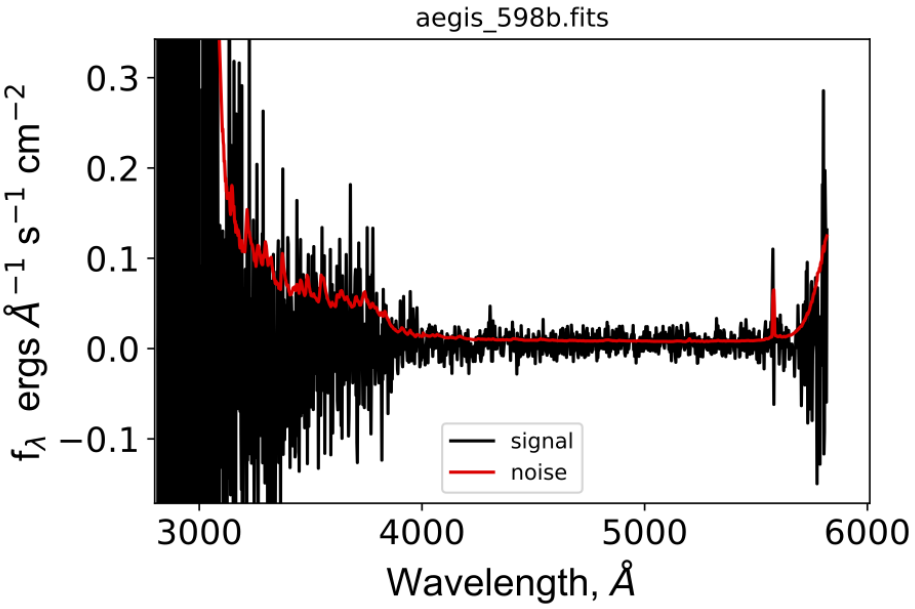

Figure 7.69 *Optical/UV spectrum of source 'aegis_598b', located in the CFHTLS D3 field.*

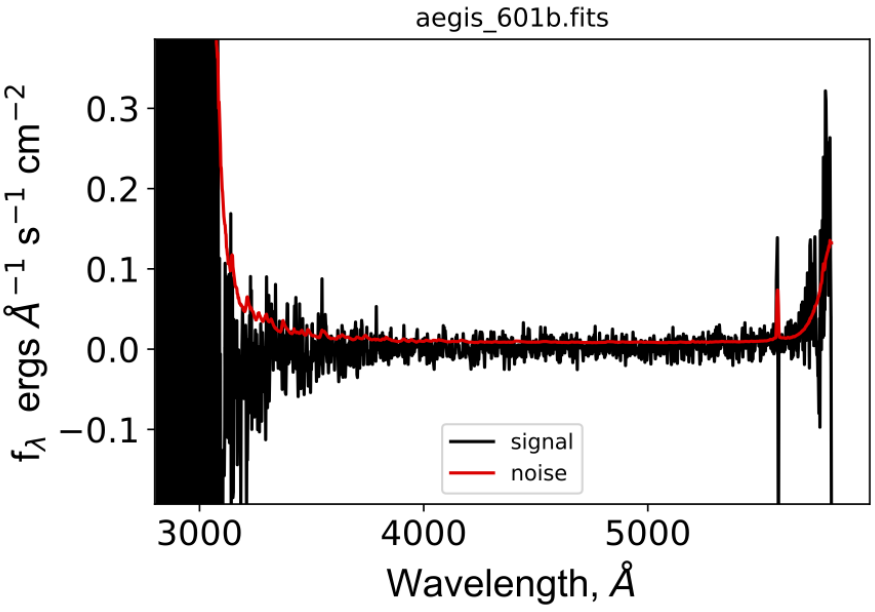

Figure 7.70 *Optical/UV spectrum of source 'aegis_601b', located in the CFHTLS D3 field.*

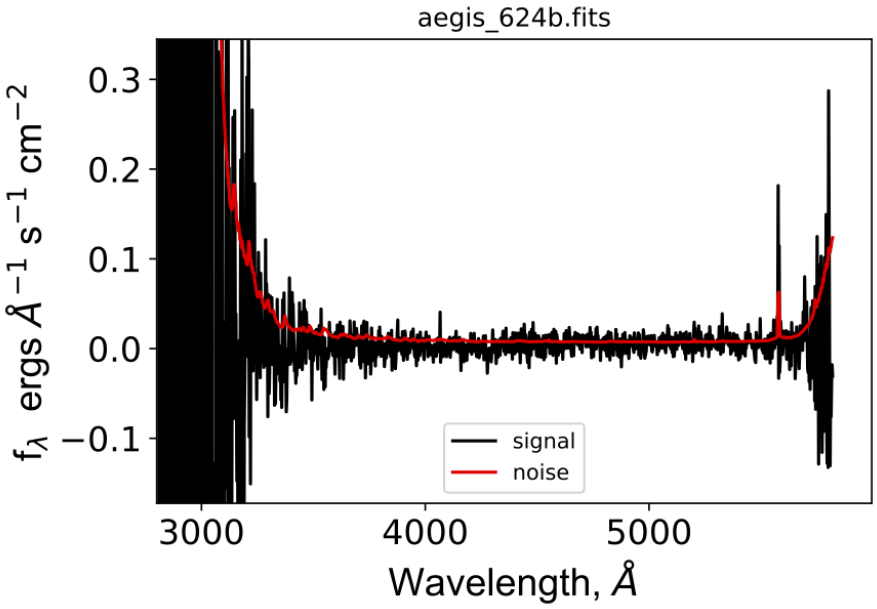

Figure 7.71 *Optical/UV spectrum of source 'aegis_624b', located in the CFHTLS D3 field.*

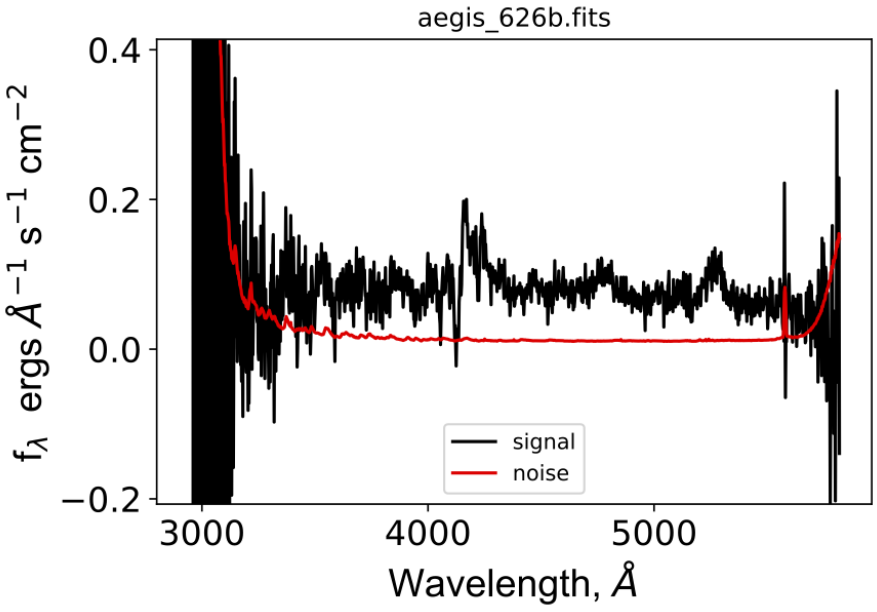

Figure 7.72 *Optical/UV spectrum of source 'aegis_626b', located in the CFHTLS D3 field.*

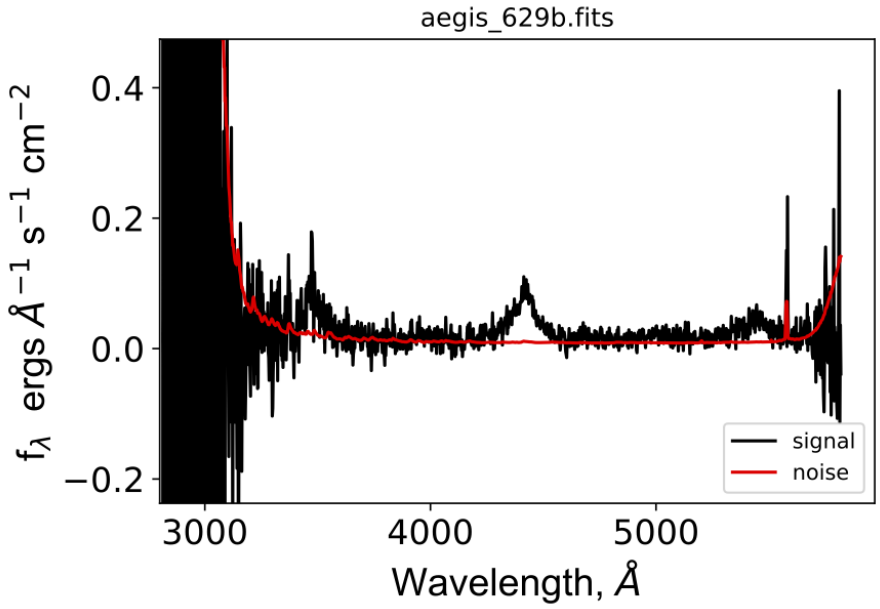

Figure 7.73 *Optical/UV spectrum of source 'aegis_629b', located in the CFHTLS D3 field.*

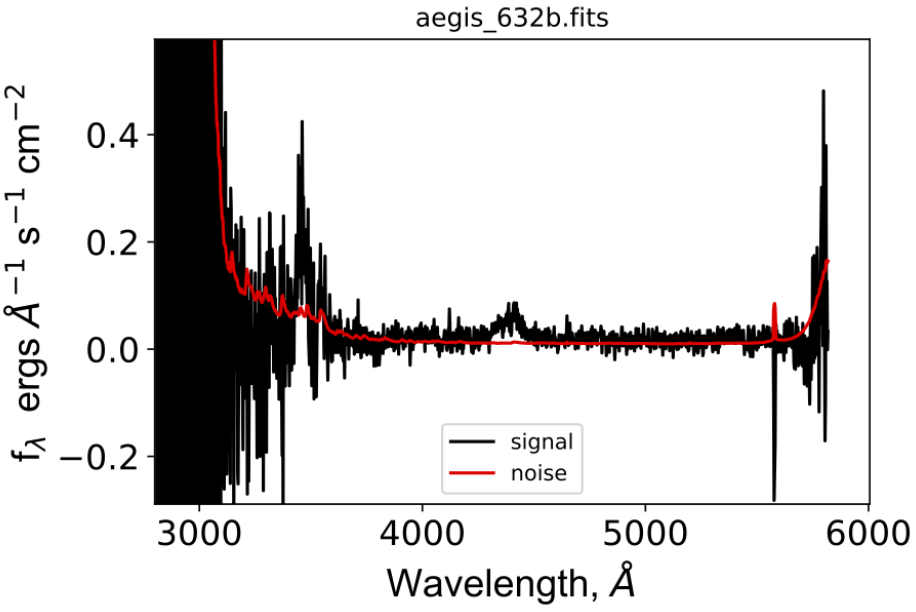

Figure 7.74 *Optical/UV spectrum of source 'aegis_632b', located in the CFHTLS D3 field.*

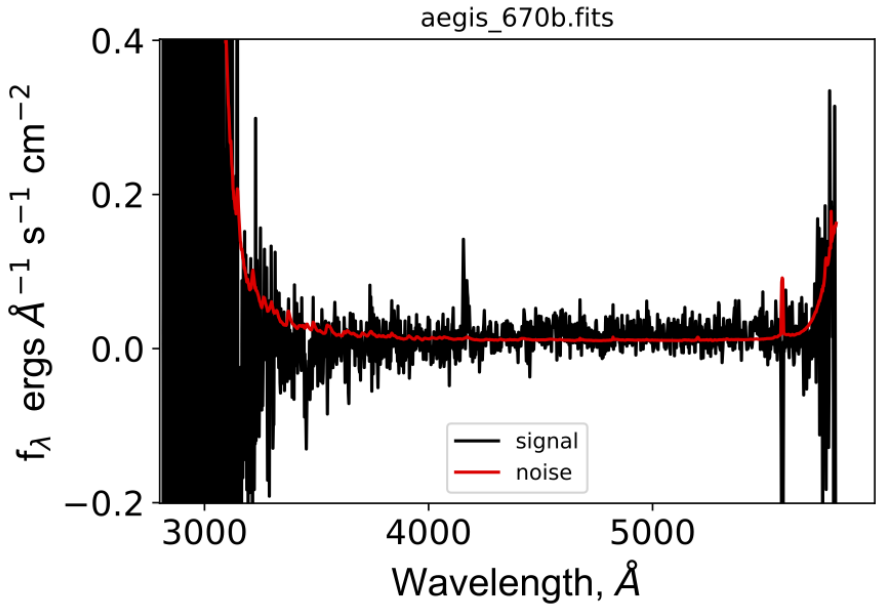

Figure 7.75 *Optical/UV spectrum of source 'aegis_670b', located in the CFHTLS D3 field.*

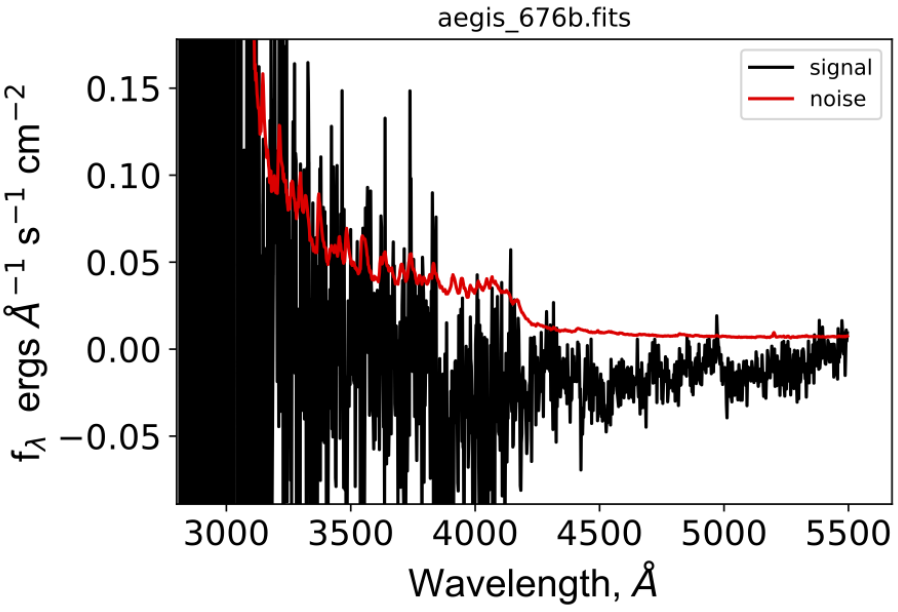

Figure 7.76 *Optical/UV spectrum of source 'aegis_676b', located in the CFHTLS D3 field.*

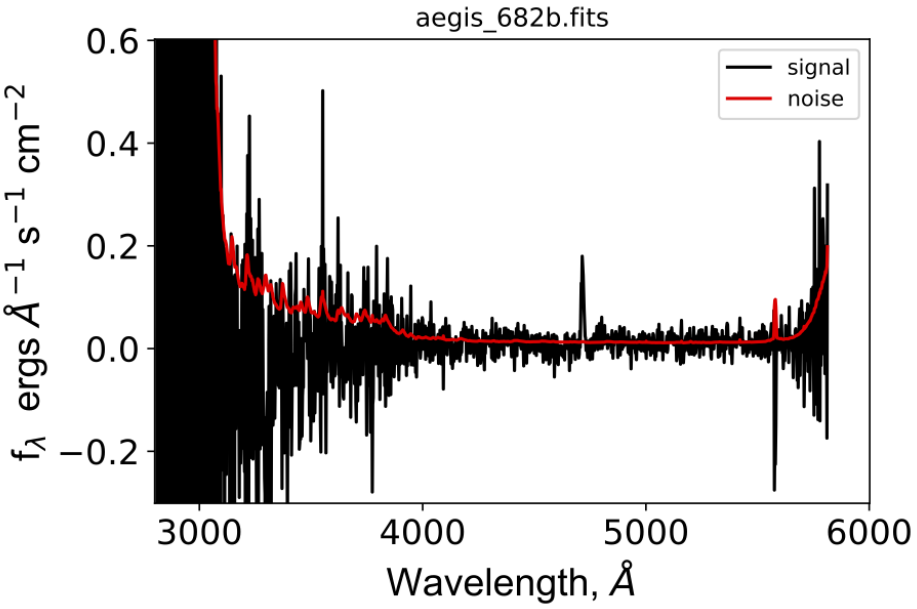

Figure 7.77 *Optical/UV spectrum of source 'aegis_682b', located in the CFHTLS D3 field.*

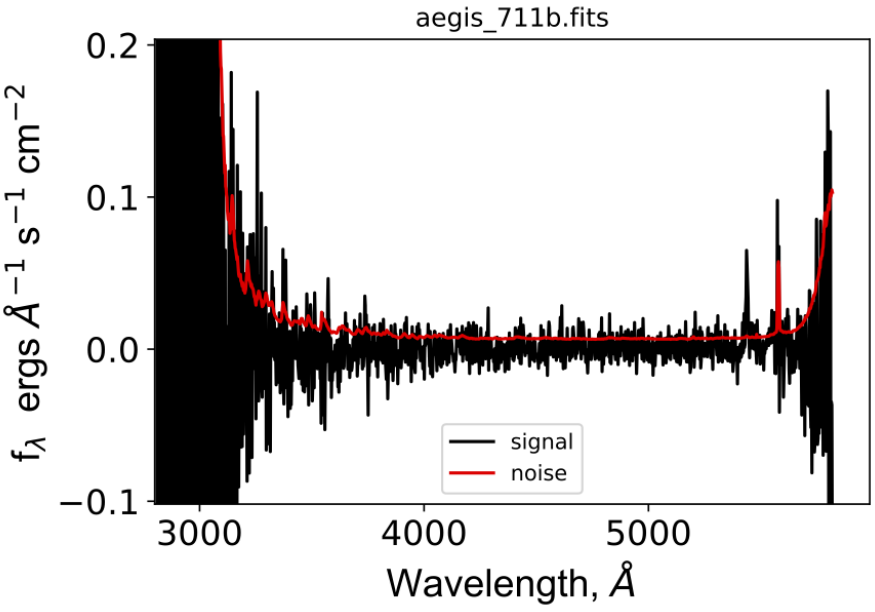

Figure 7.78 *Optical/UV spectrum of source 'aegis_711b', located in the CFHTLS D3 field.*

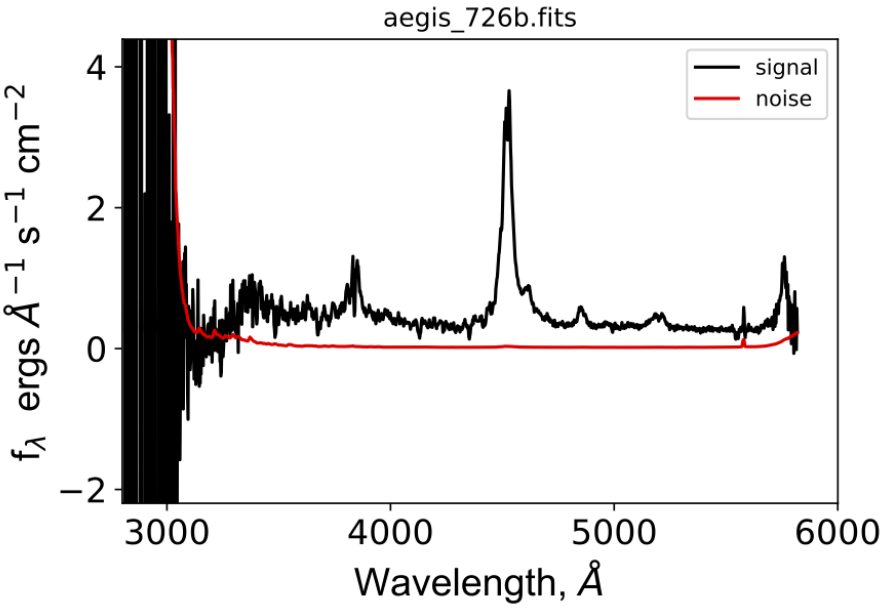

Figure 7.79 *Optical/UV spectrum of source 'aegis_726b', located in the CFHTLS D3 field.*

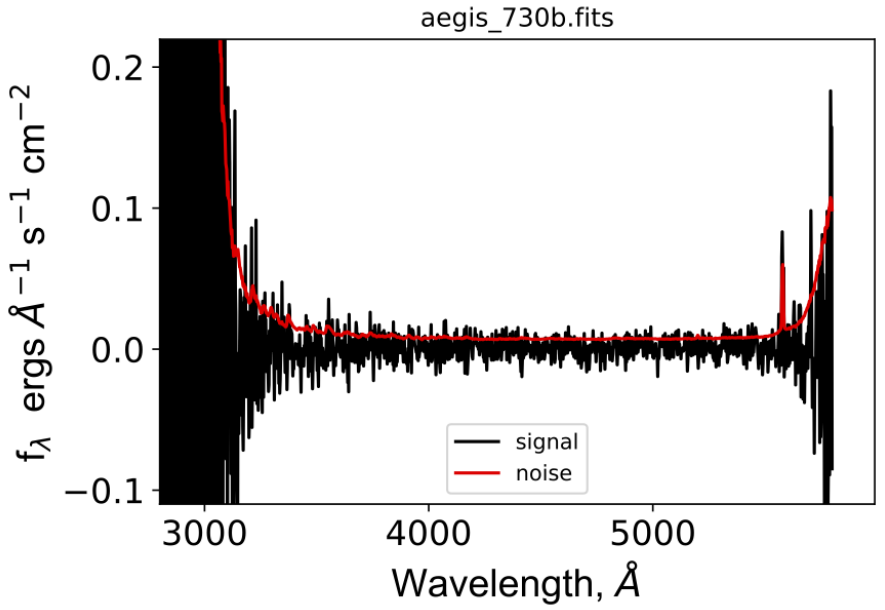

Figure 7.80 *Optical/UV spectrum of source 'aegis_730b', located in the CFHTLS D3 field.*

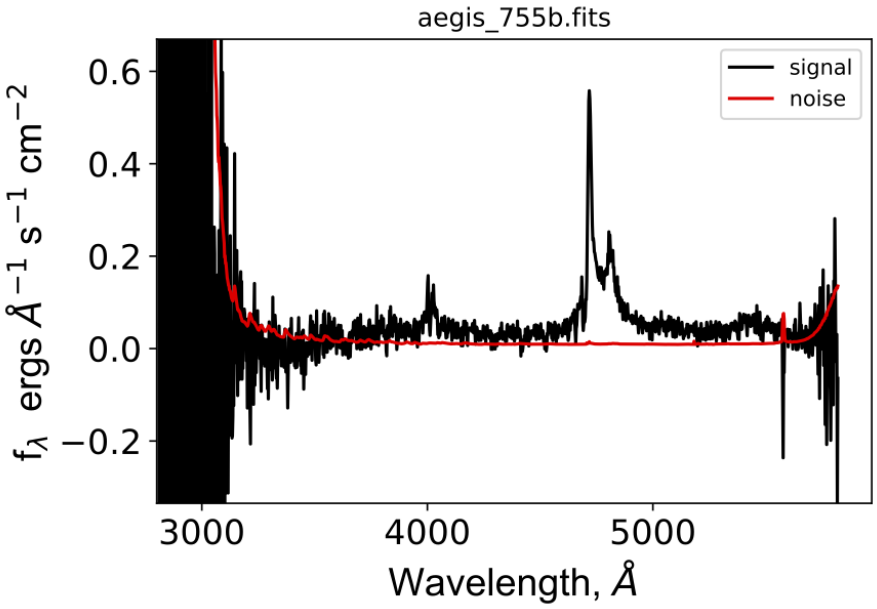

Figure 7.81 *Optical/UV spectrum of source 'aegis_755b', located in the CFHTLS D3 field.*

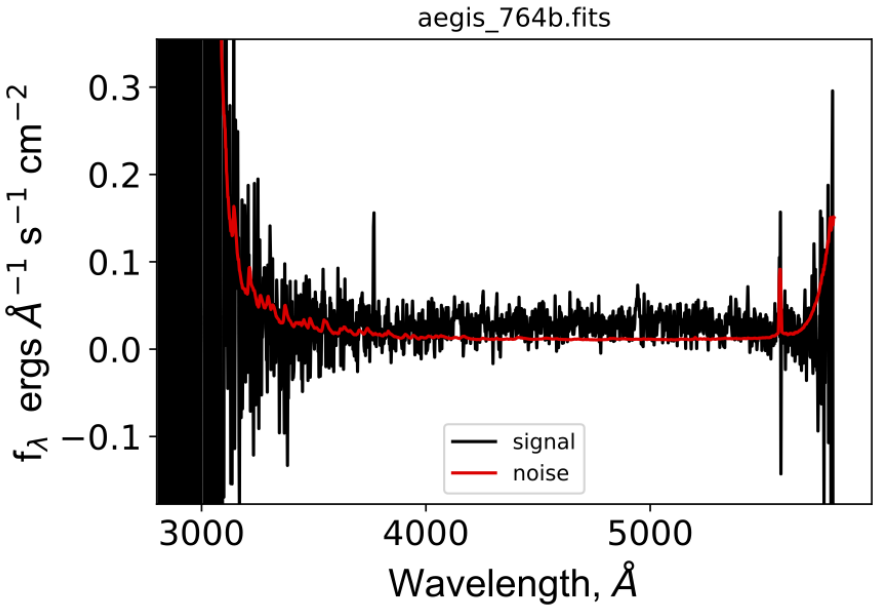

Figure 7.82 *Optical/UV spectrum of source 'aegis_764b', located in the CFHTLS D3 field.*

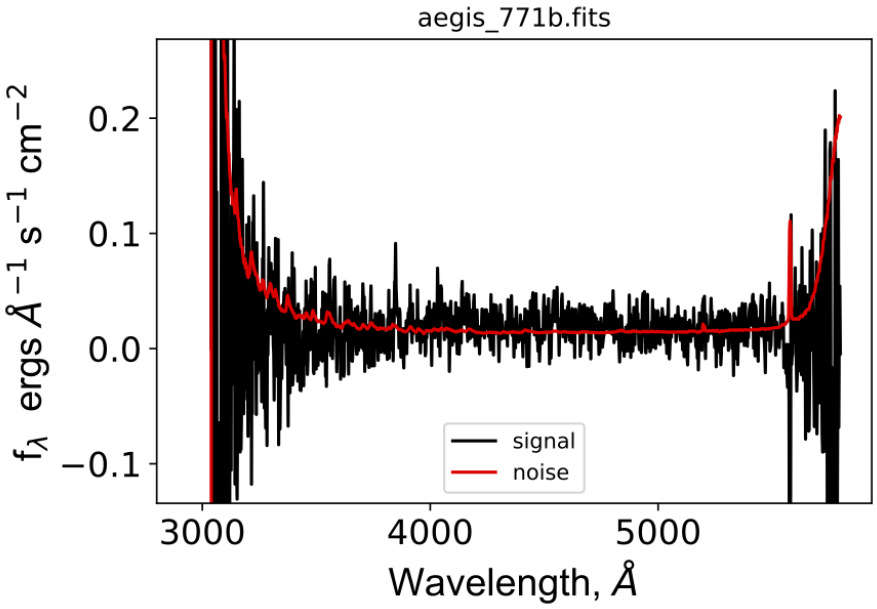

Figure 7.83 *Optical/UV spectrum of source 'aegis_771b', located in the CFHTLS D3 field.*

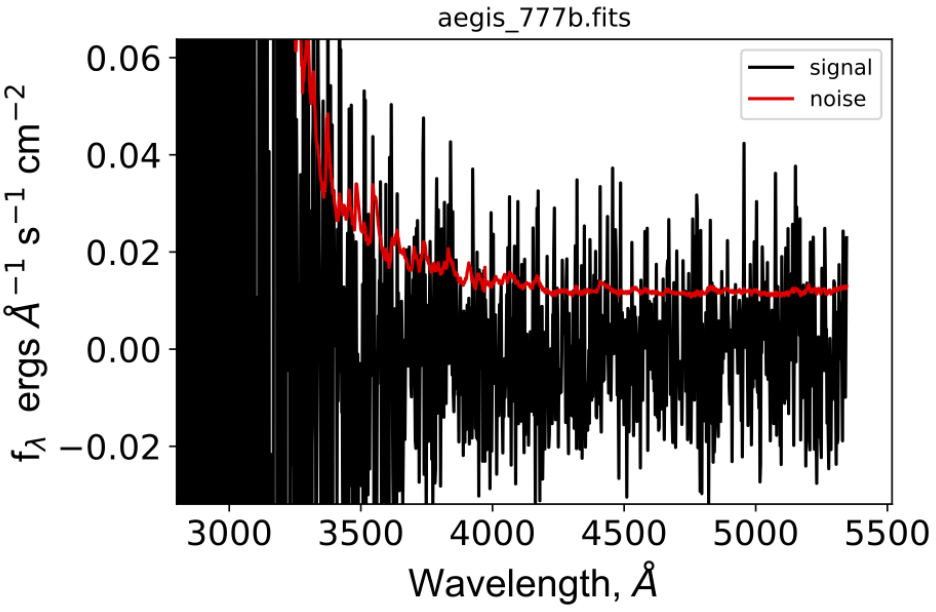

Figure 7.84 *Optical/UV spectrum of source 'aegis_777b', located in the CFHTLS D3 field.*

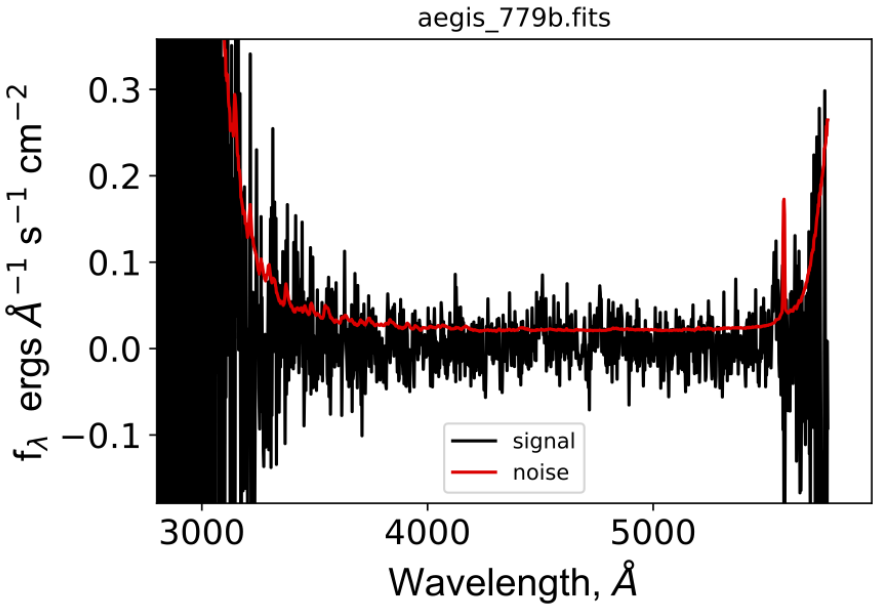

Figure 7.85 *Optical/UV spectrum of source 'aegis_779b', located in the CFHTLS D3 field.*

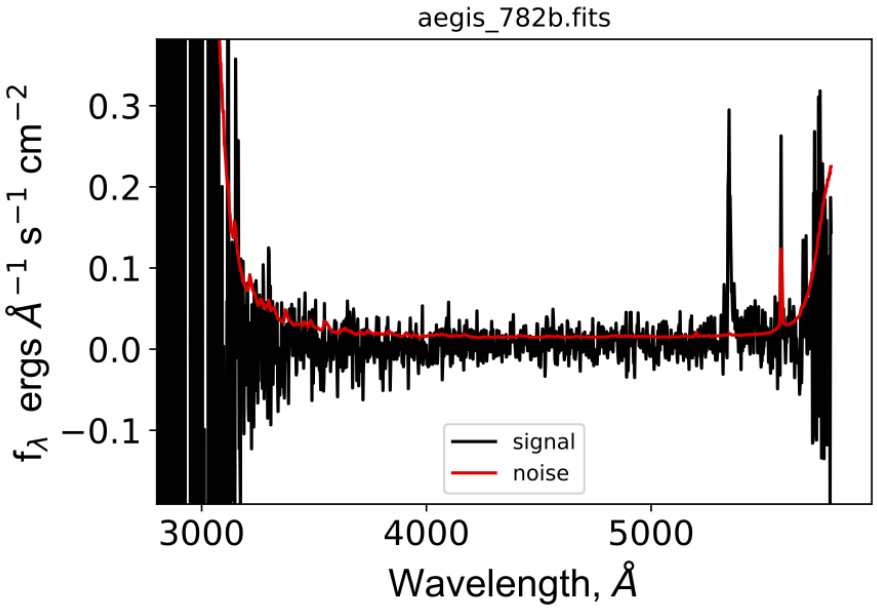

Figure 7.86 *Optical/UV spectrum of source 'aegis_782b', located in the CFHTLS D3 field.*

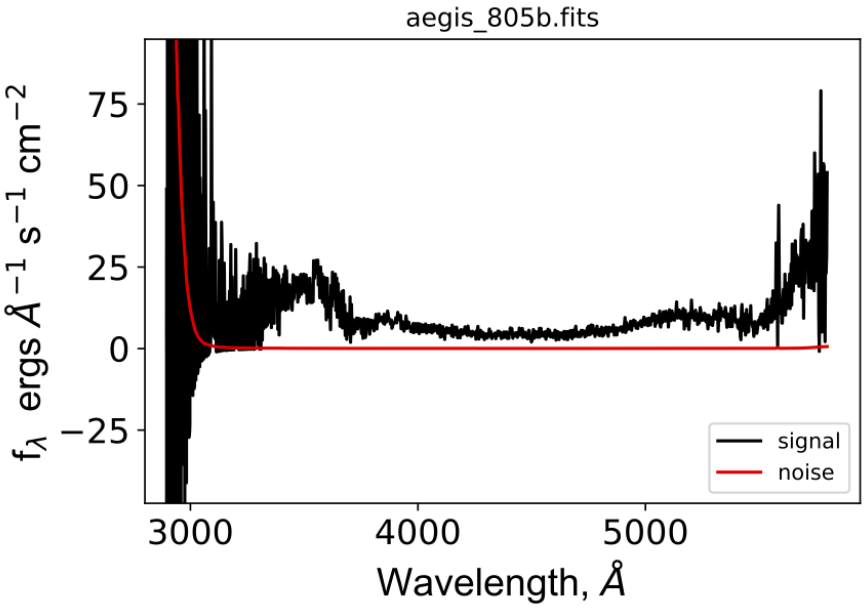

Figure 7.87 *Optical/UV spectrum of source 'aegis_805b', located in the CFHTLS D3 field.*

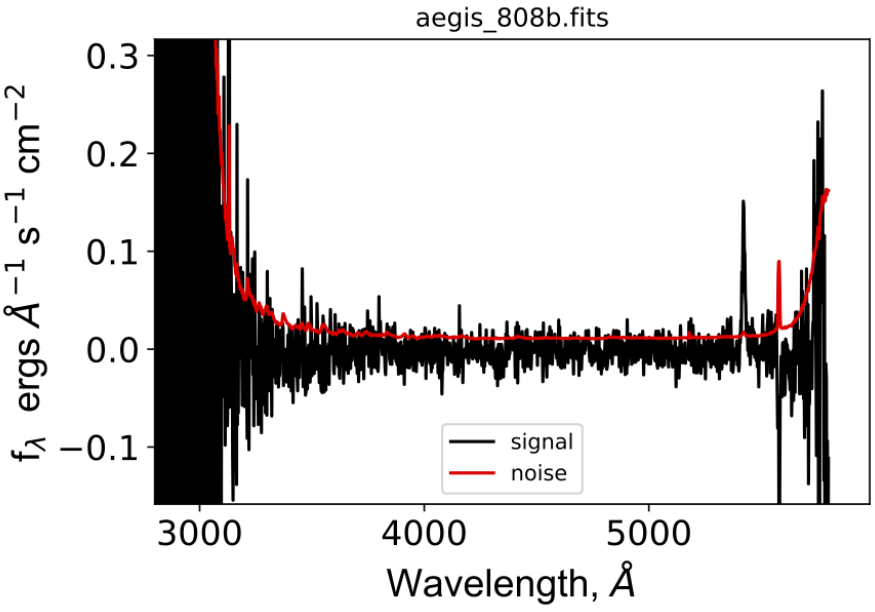

Figure 7.88 *Optical/UV spectrum of source 'aegis_808b', located in the CFHTLS D3 field.*

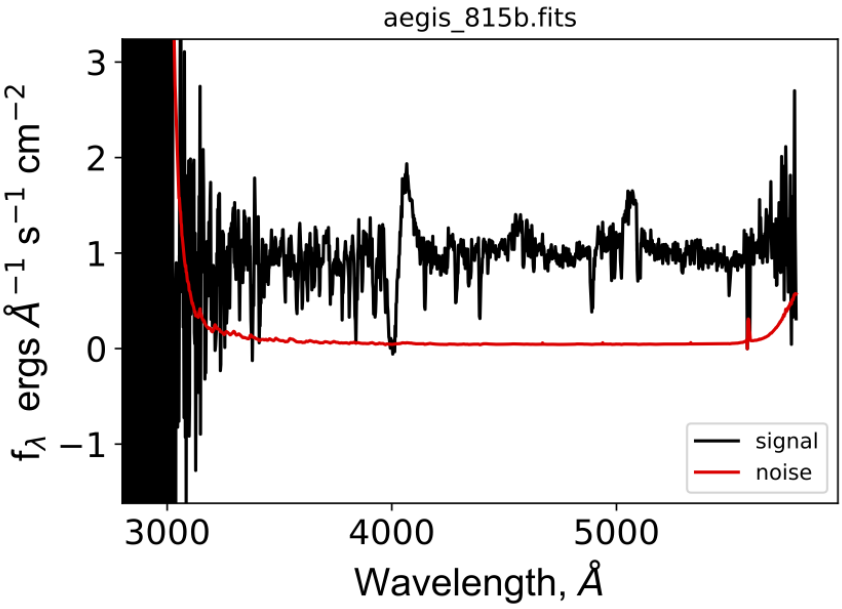

Figure 7.89 *Optical/UV spectrum of source 'aegis_815b', located in the CFHTLS D3 field.*

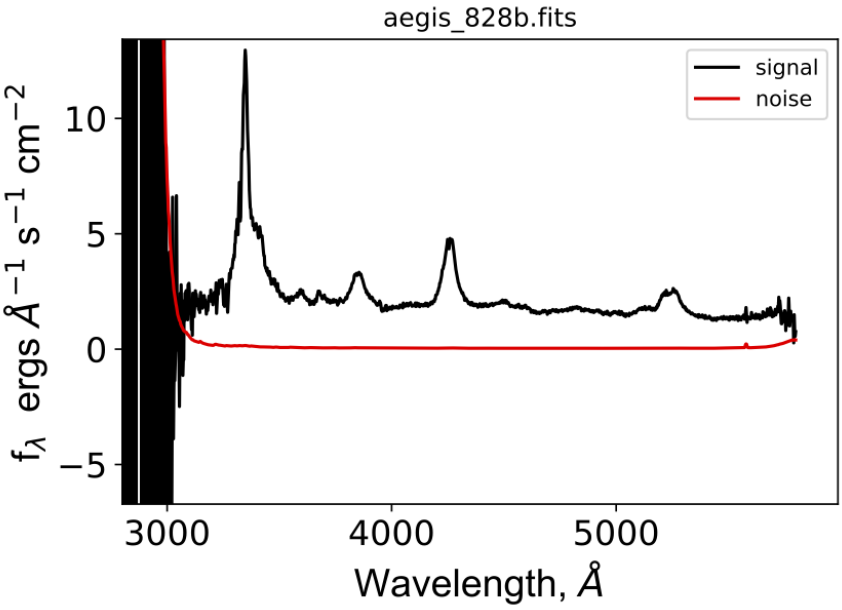

Figure 7.90 *Optical/UV spectrum of source 'aegis_828b', located in the CFHTLS D3 field.*

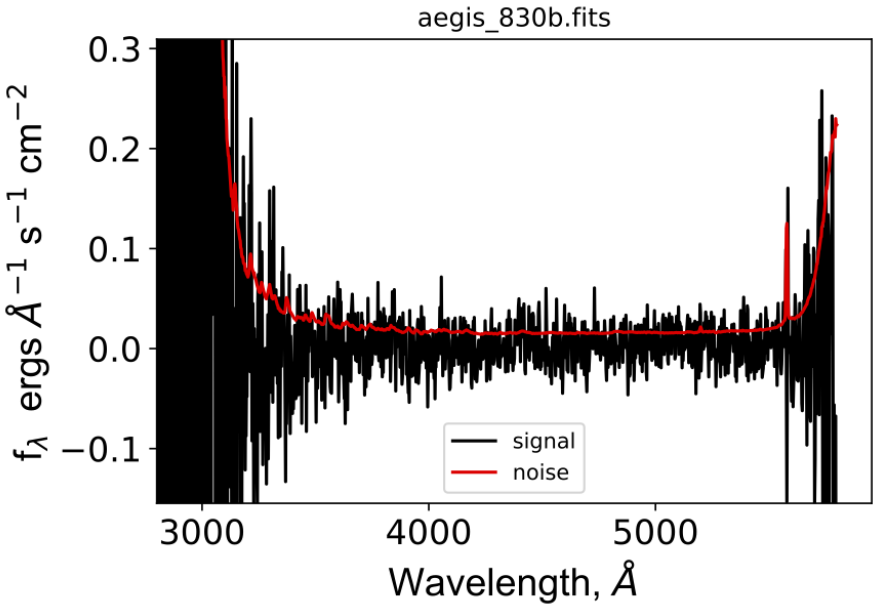

Figure 7.91 *Optical/UV spectrum of source 'aegis_830b', located in the CFHTLS D3 field.*

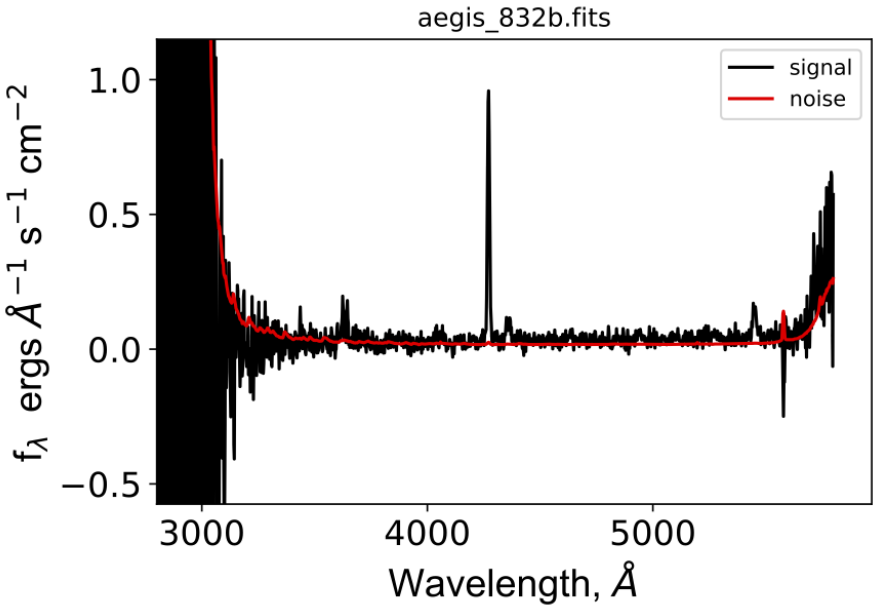

Figure 7.92 *Optical/UV spectrum of source 'aegis_832b', located in the CFHTLS D3 field.*

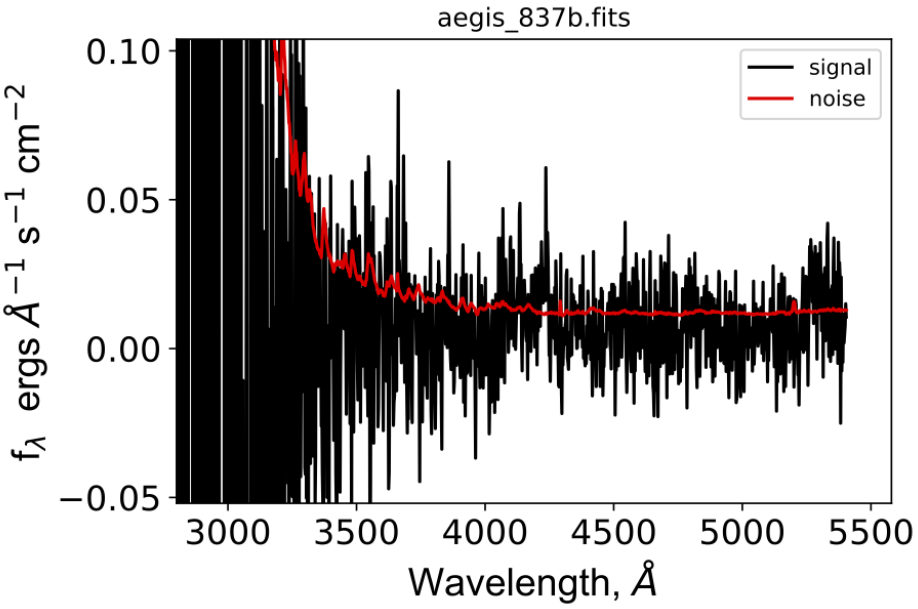

Figure 7.93 *Optical/UV spectrum of source 'aegis_837b', located in the CFHTLS D3 field.*

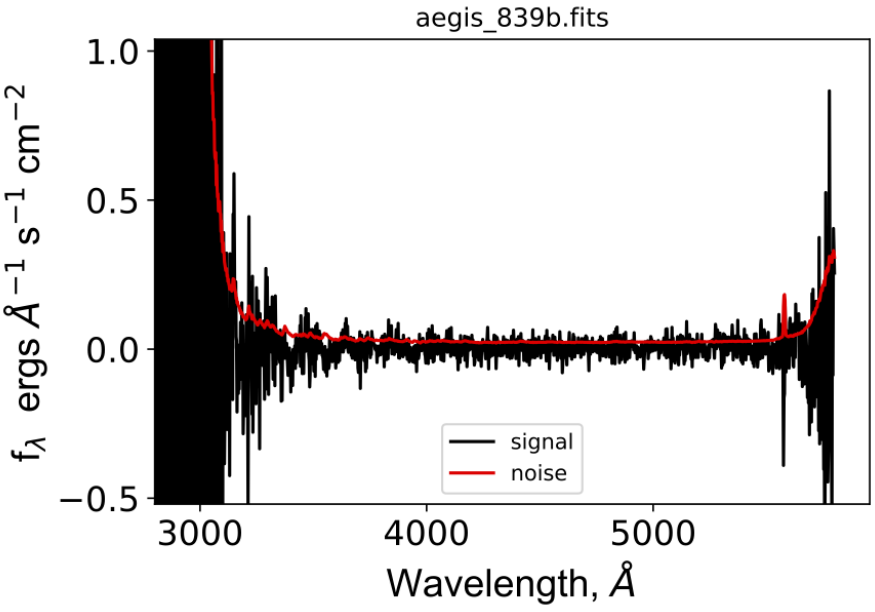

Figure 7.94 *Optical/UV spectrum of source 'aegis_839b', located in the CFHTLS D3 field.*

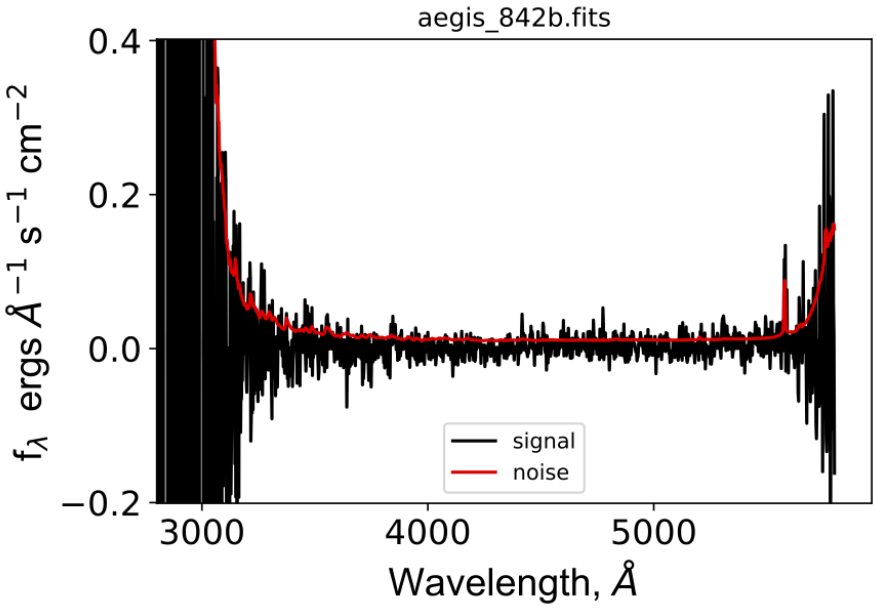

Figure 7.95 *Optical/UV spectrum of source 'aegis_842b', located in the CFHTLS D3 field.*

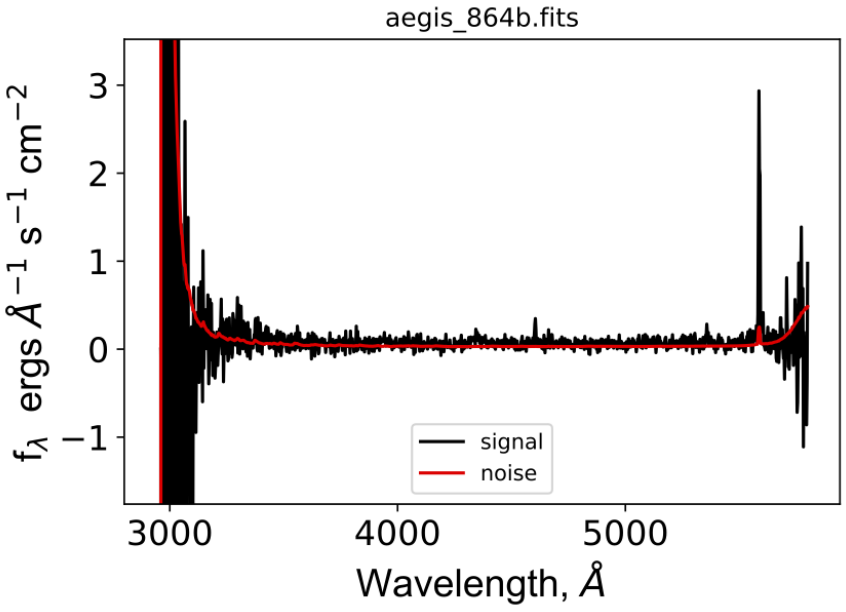

Figure 7.96 *Optical/UV spectrum of source 'aegis_864b', located in the CFHTLS D3 field.*

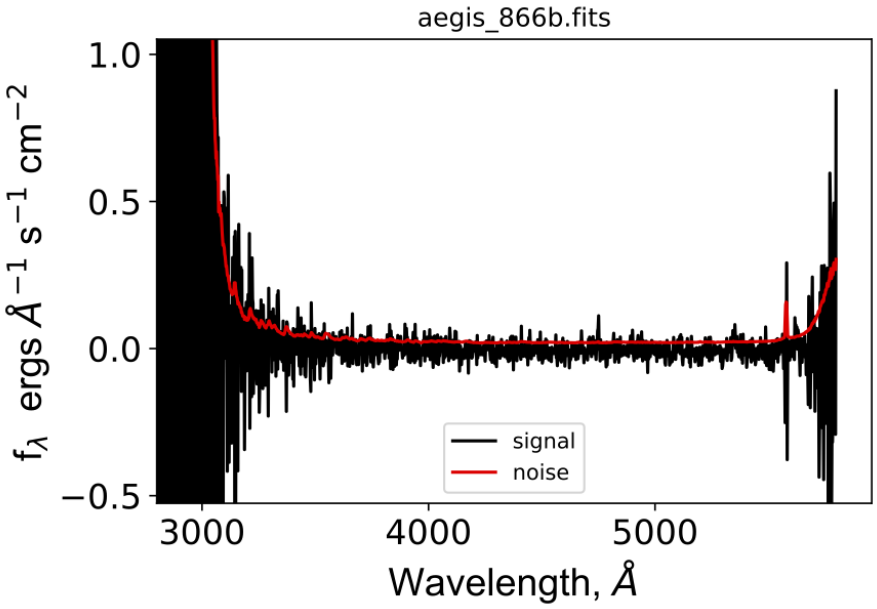

Figure 7.97 *Optical/UV spectrum of source 'aegis_866b', located in the CFHTLS D3 field.*

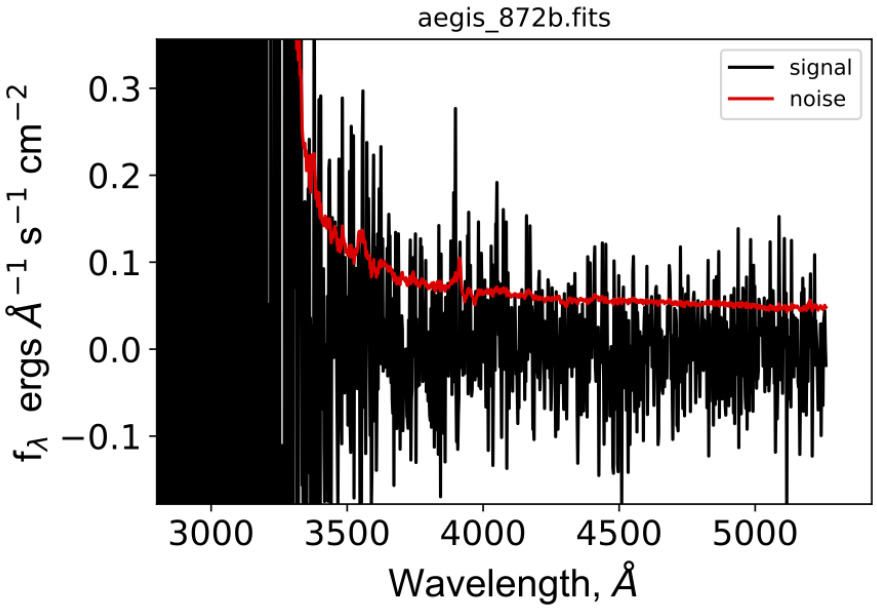

Figure 7.98 *Optical/UV spectrum of source 'aegis_872b', located in the CFHTLS D3 field.*

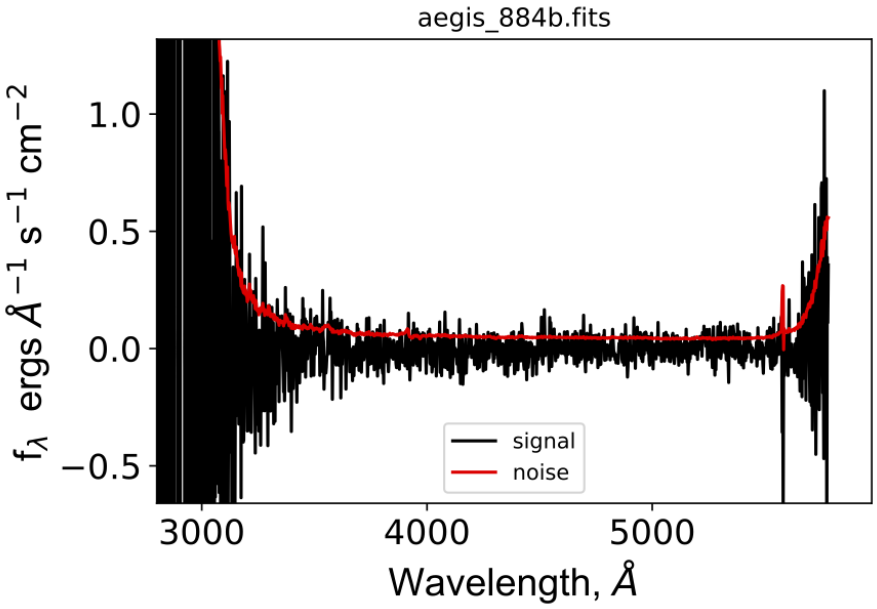

Figure 7.99 *Optical/UV spectrum of source 'aegis_884b', located in the CFHTLS D3 field.*

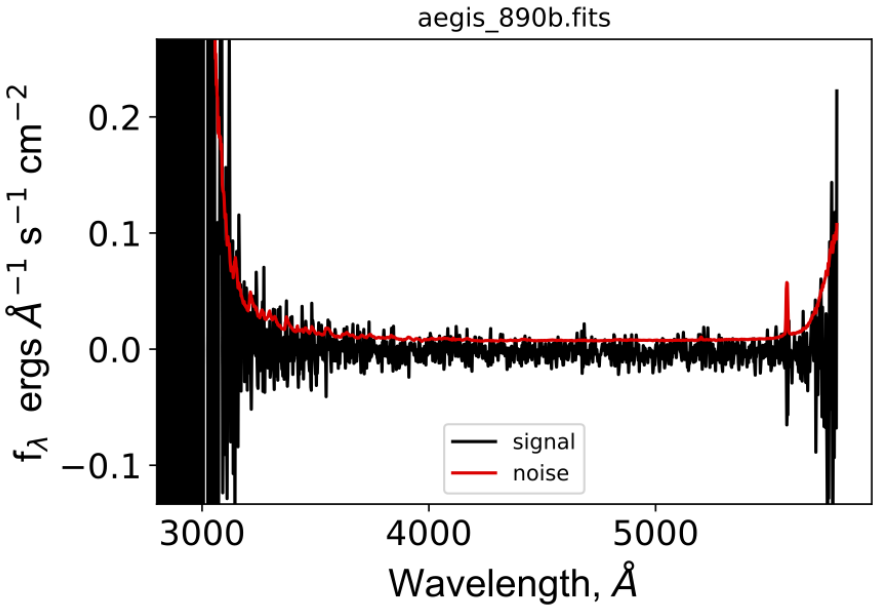

Figure 7.100 *Optical/UV spectrum of source 'aegis_890b', located in the CFHTLS D3 field.*

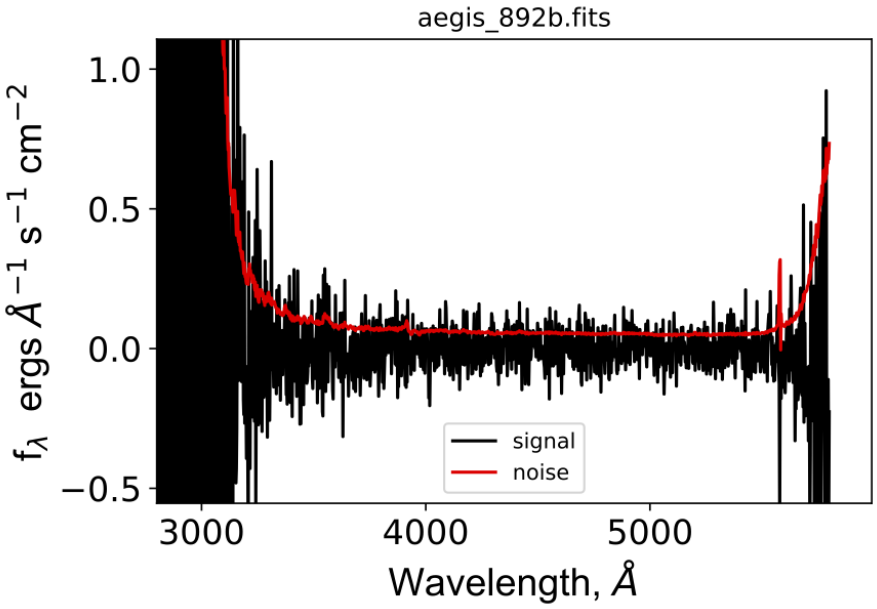

Figure 7.101 *Optical/UV spectrum of source 'aegis_892b', located in the CFHTLS D3 field.*

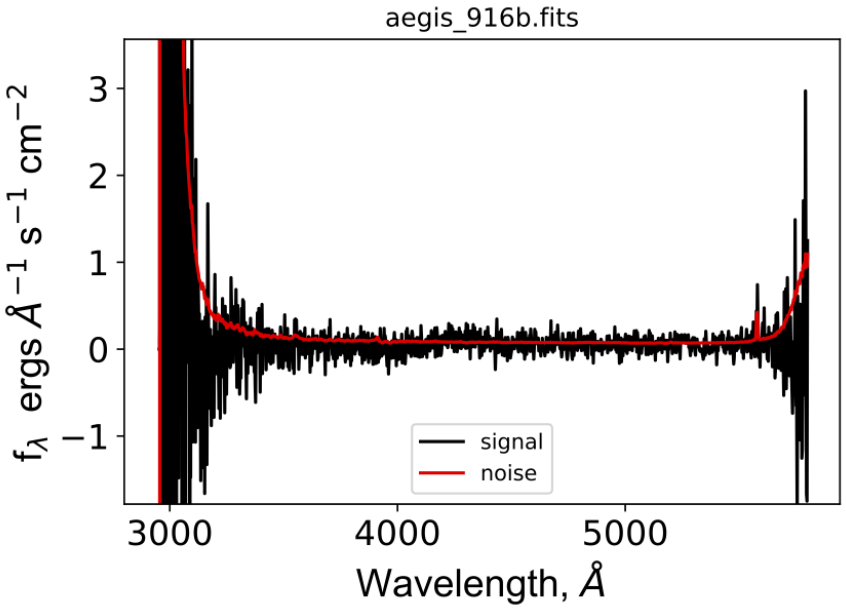

Figure 7.102 *Optical/UV spectrum of source 'aegis_916b', located in the CFHTLS D3 field.*

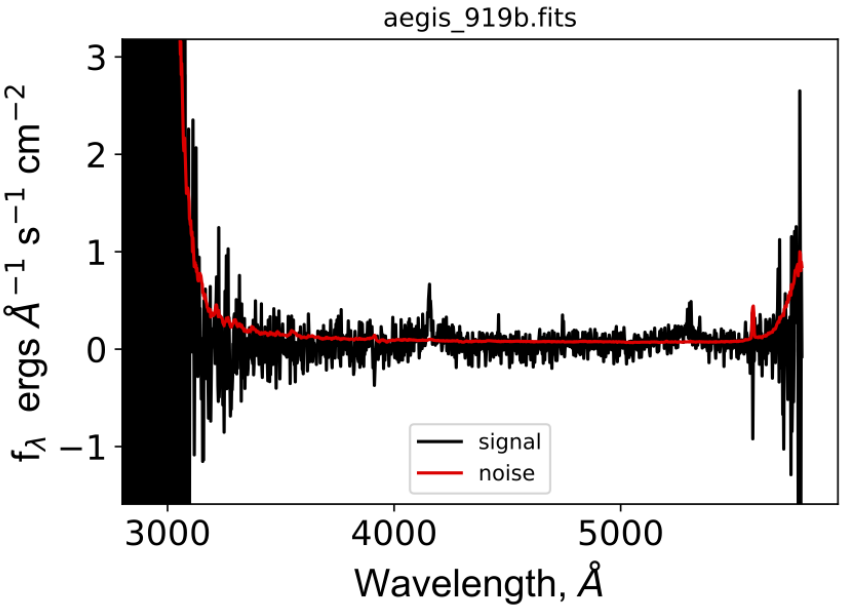

Figure 7.103 *Optical/UV spectrum of source 'aegis_919b', located in the CFHTLS D3 field.*

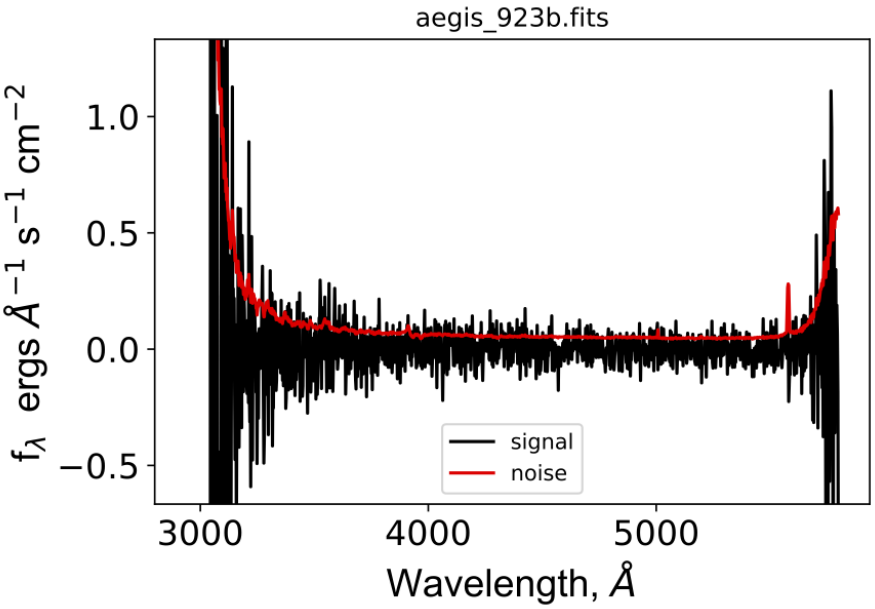

Figure 7.104 *Optical/UV spectrum of source 'aegis_923b', located in the CFHTLS D3 field.*

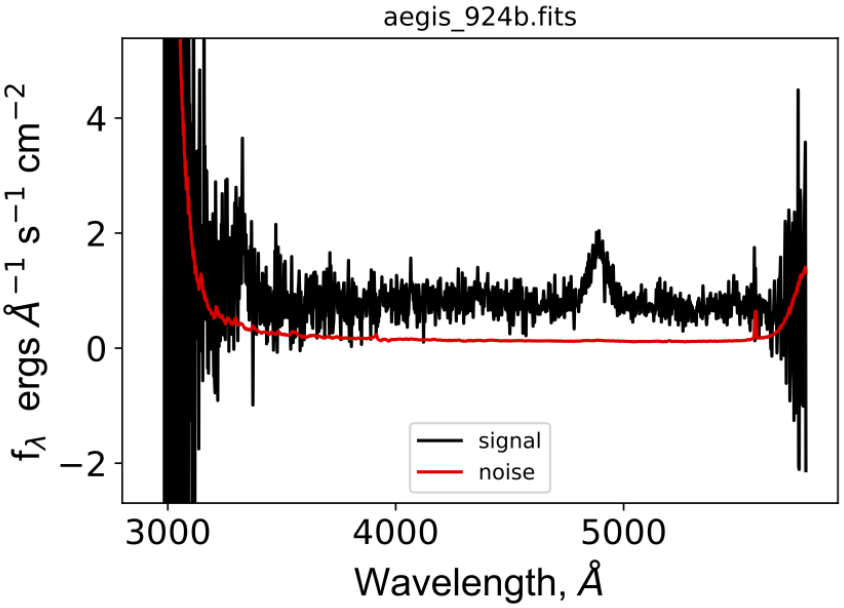

Figure 7.105 *Optical/UV spectrum of source 'aegis_924b', located in the CFHTLS D3 field.*

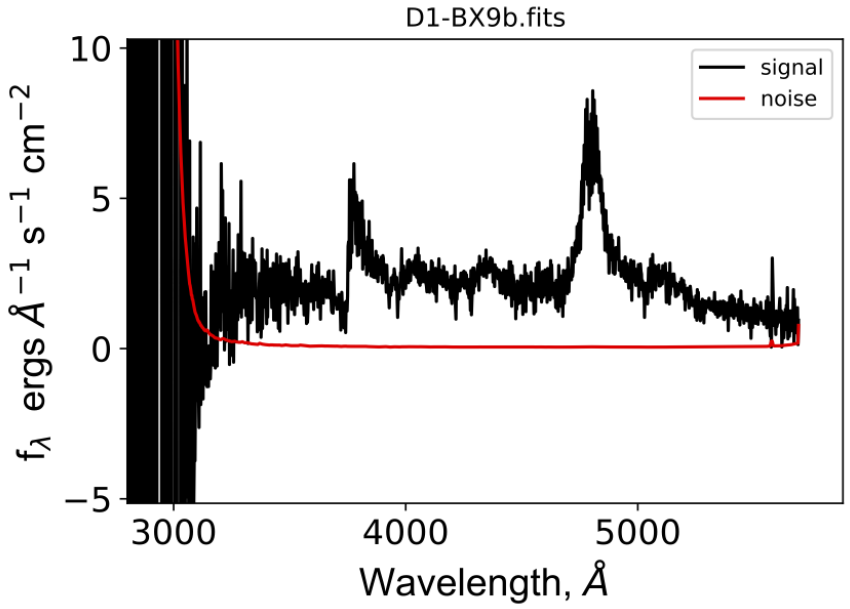

Figure 7.106 *Optical/UV spectrum of source 'D1-BX9b', located in the CFHTLS D1 field.*

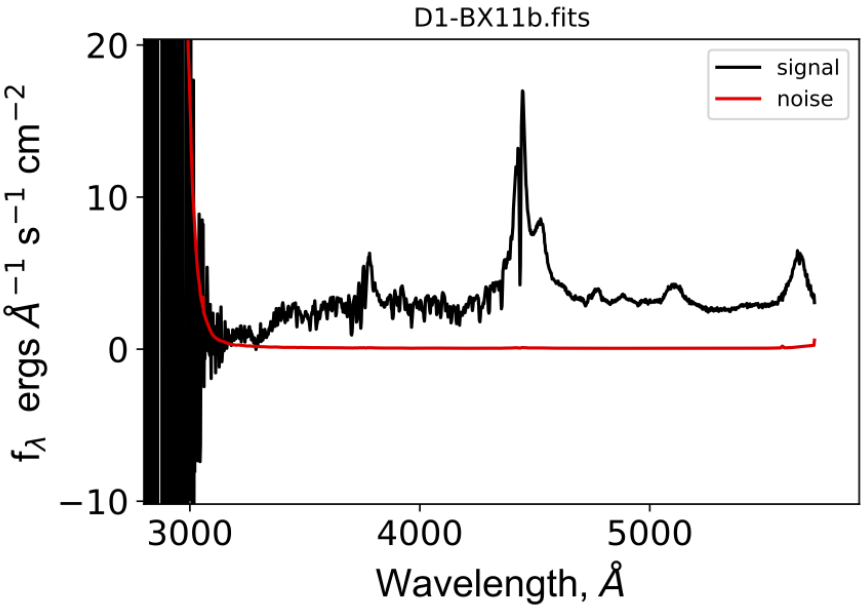

Figure 7.107 *Optical/UV spectrum of source 'D1-BX11b', located in the CFHTLS D1 field.*

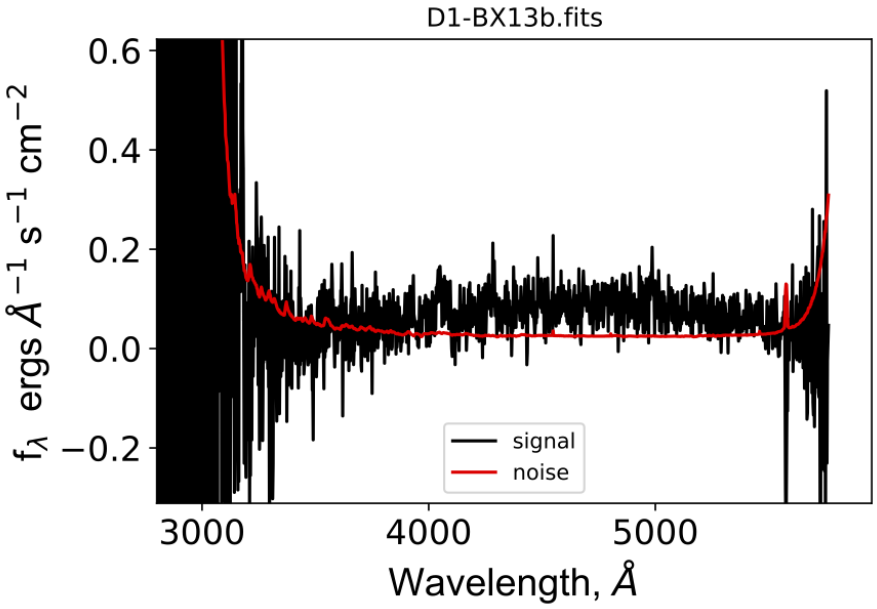

Figure 7.108 *Optical/UV spectrum of source 'D1-BX13b', located in the CFHTLS D1 field.*

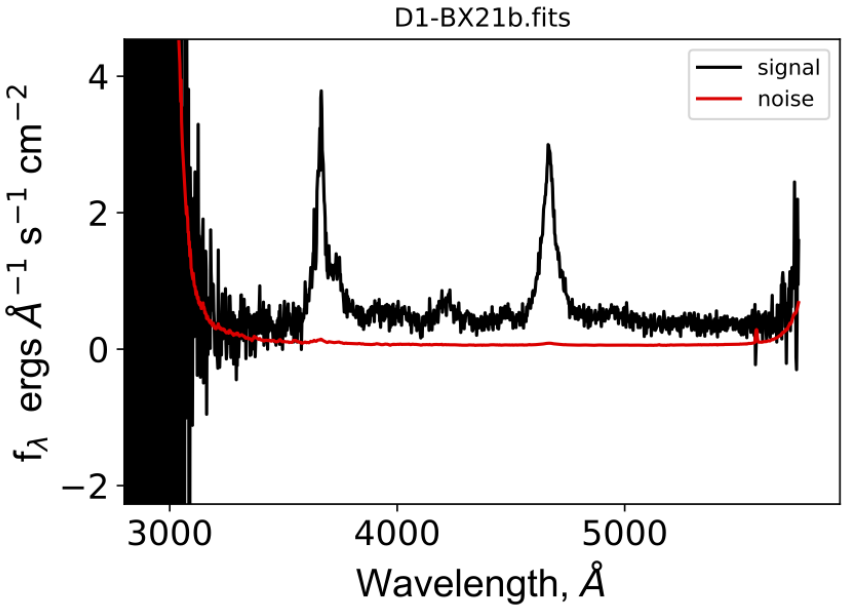

Figure 7.109 *Optical/UV spectrum of source 'D1-BX21b', located in the CFHTLS D1 field.*

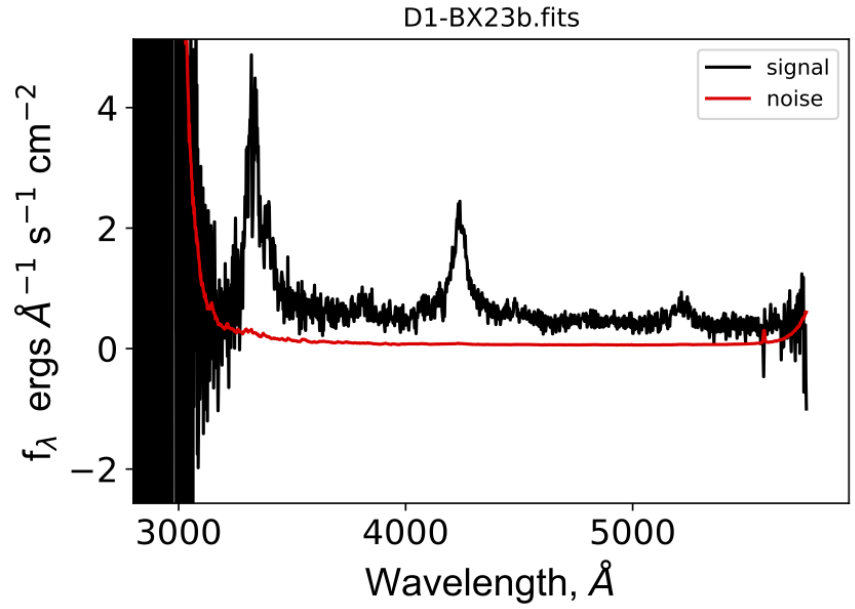

Figure 7.110 *Optical/UV spectrum of source 'D1-BX23b', located in the CFHTLS D1 field.*

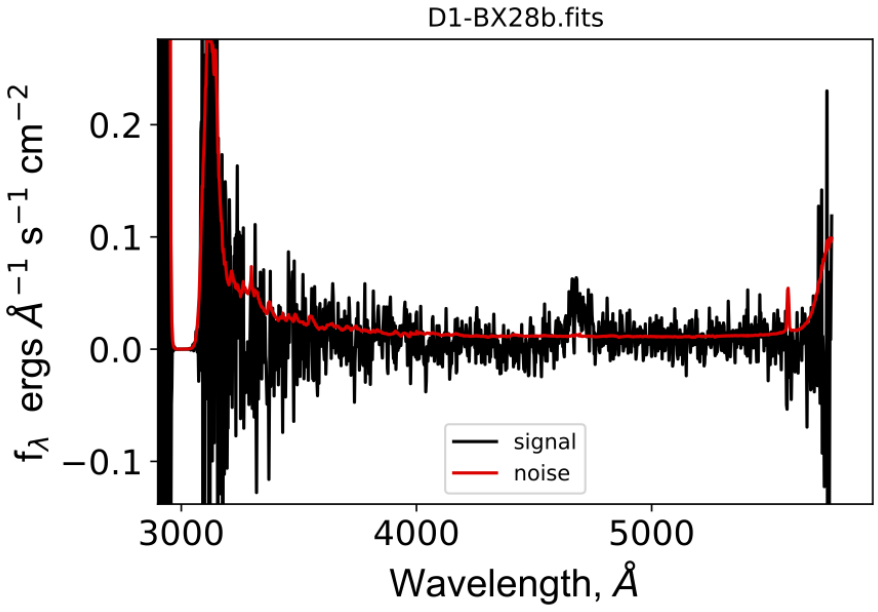

Figure 7.111 *Optical/UV spectrum of source 'D1-BX28b', located in the CFHTLS D1 field.*

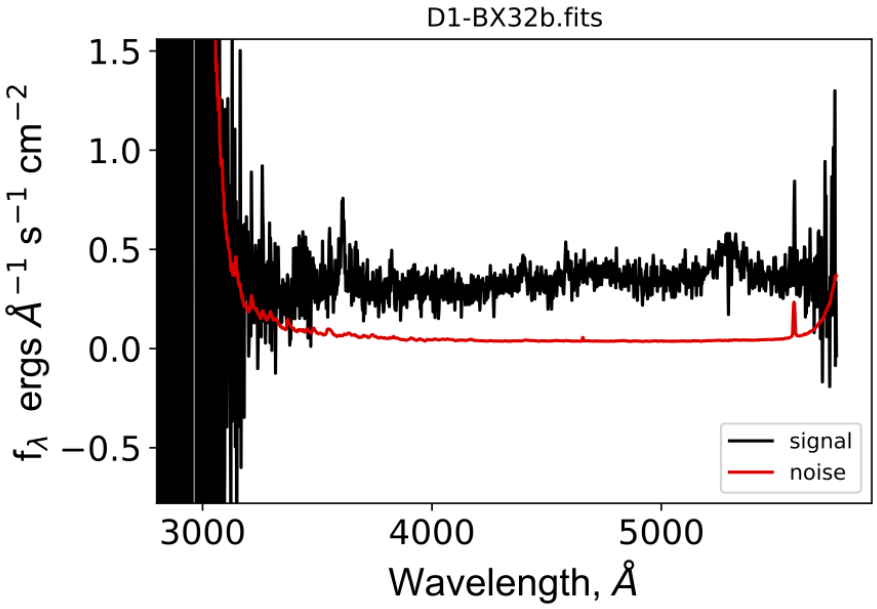

Figure 7.112 *Optical/UV spectrum of source 'D1-BX32b', located in the CFHTLS D1 field.*

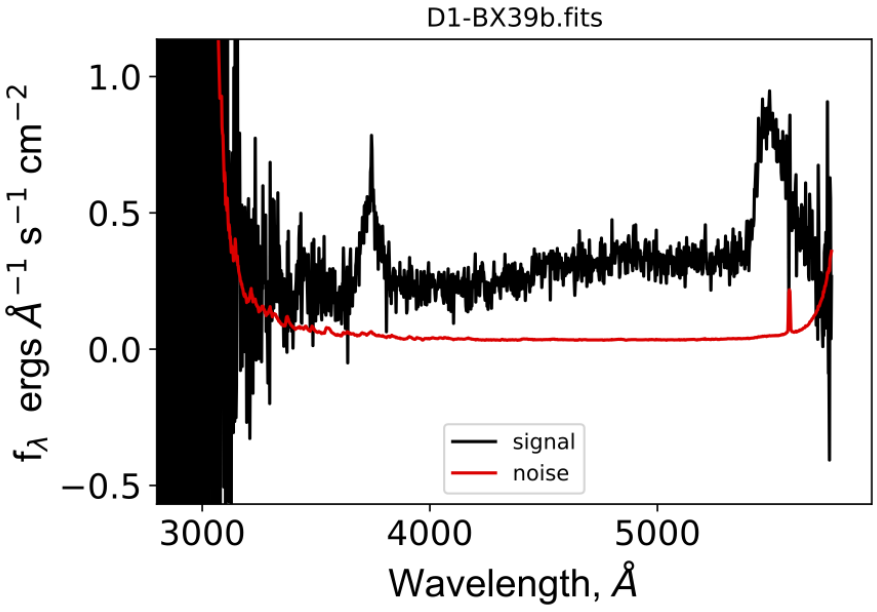

Figure 7.113 *Optical/UV spectrum of source 'D1-BX39b', located in the CFHTLS D1 field.*

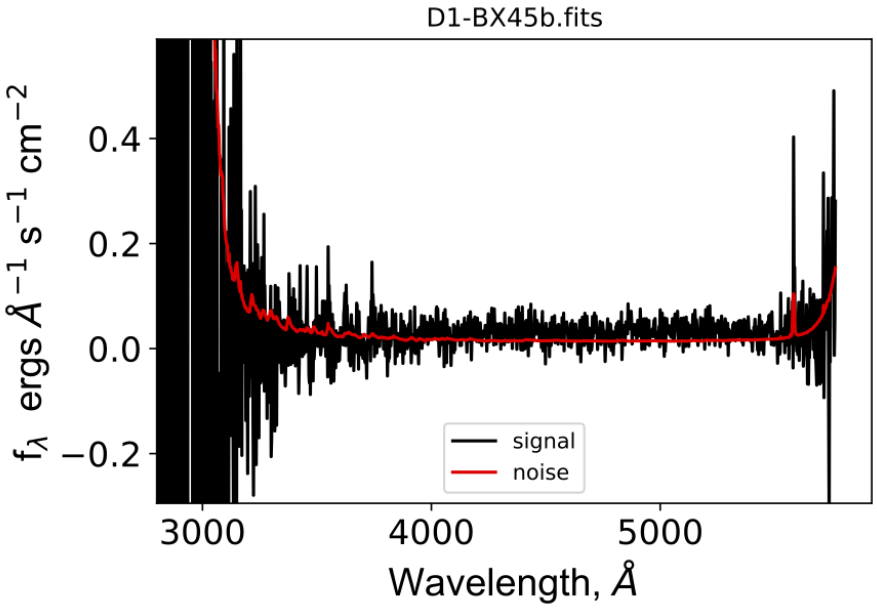

Figure 7.114 *Optical/UV spectrum of source 'D1-BX45b', located in the CFHTLS D1 field.*

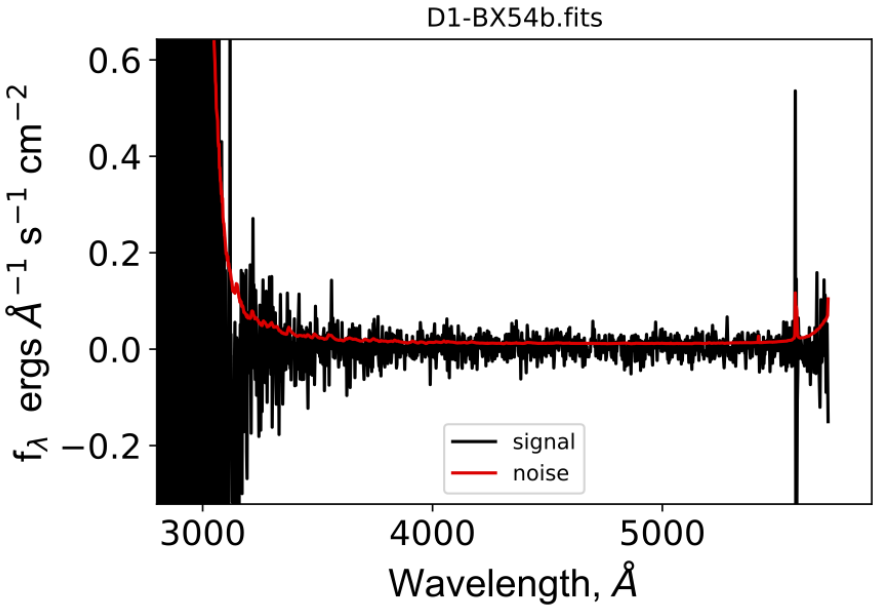

Figure 7.115 *Optical/UV spectrum of source 'D1-BX54b', located in the CFHTLS D1 field.*

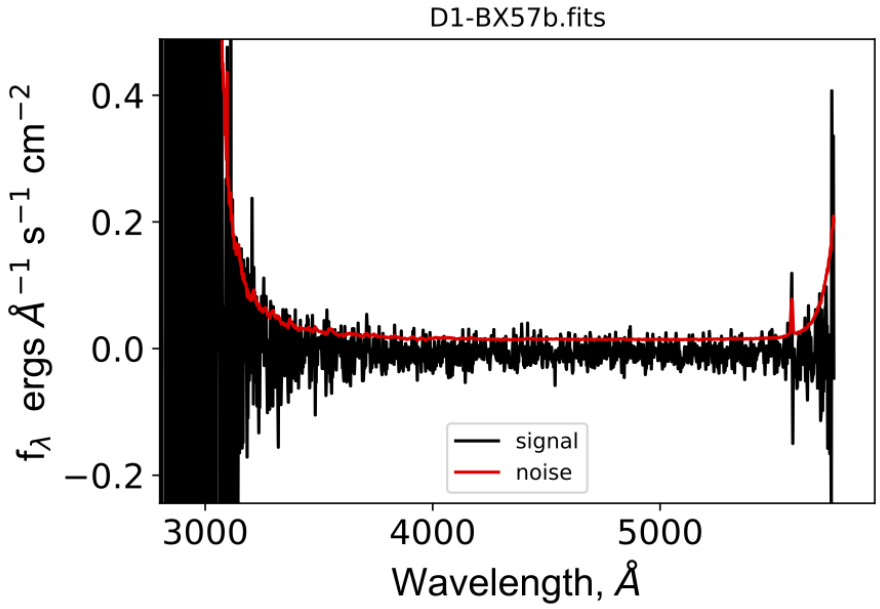

Figure 7.116 *Optical/UV spectrum of source 'D1-BX57b', located in the CFHTLS D1 field.*

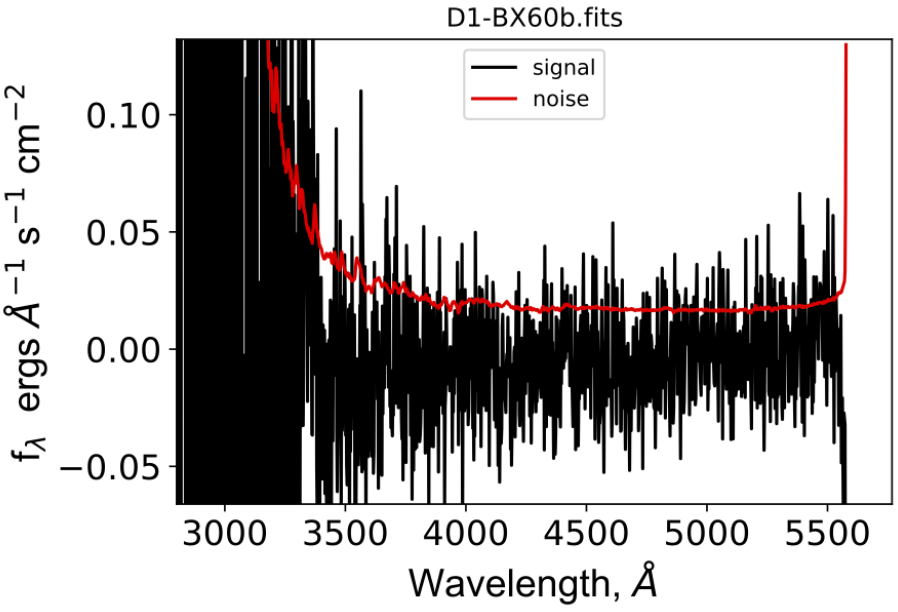

Figure 7.117 *Optical/UV spectrum of source 'D1-BX60b', located in the CFHTLS D1 field.*

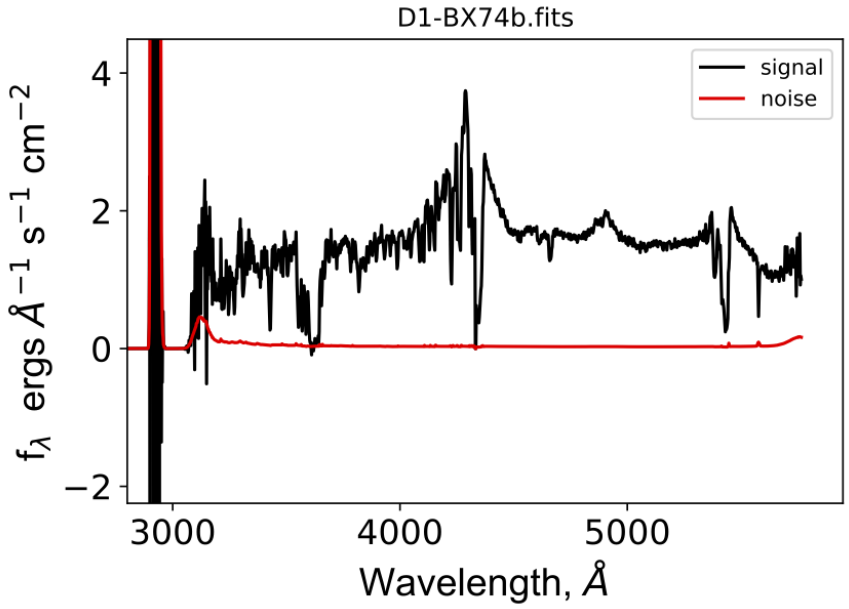

Figure 7.118 *Optical/UV spectrum of source 'D1-BX74b', located in the CFHTLS D1 field.*

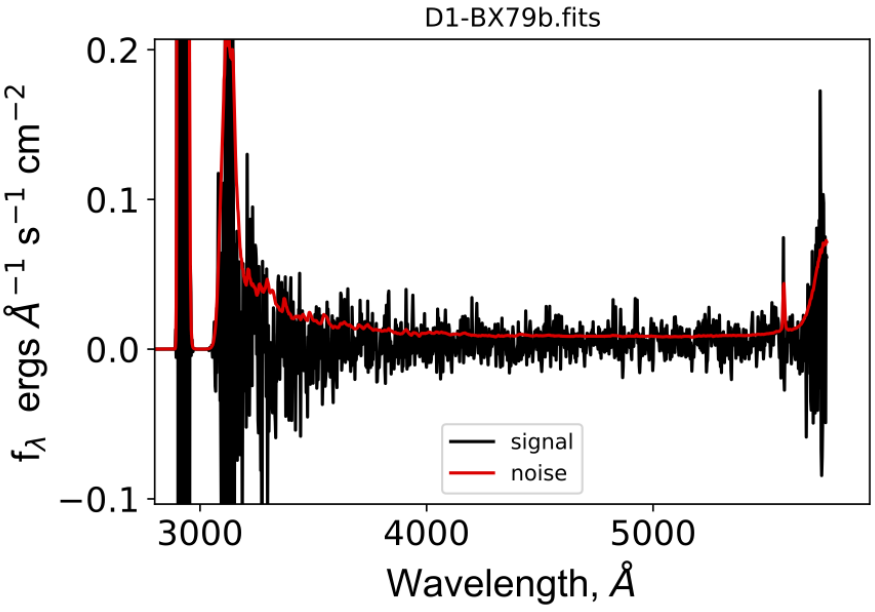

Figure 7.119 *Optical/UV spectrum of source 'D1-BX79b', located in the CFHTLS D1 field.*

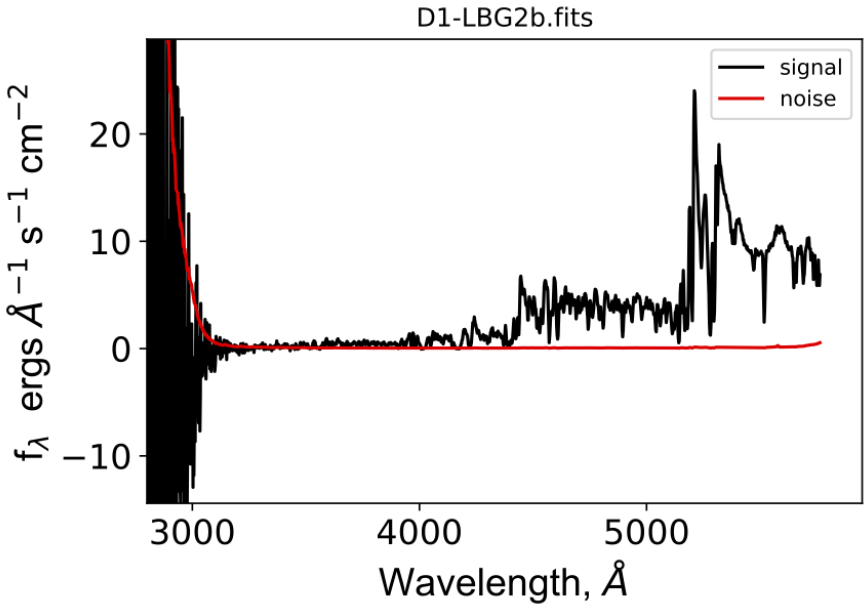

Figure 7.120 *Optical/UV spectrum of source 'D1-LBG2b', located in the CFHTLS D1 field.*

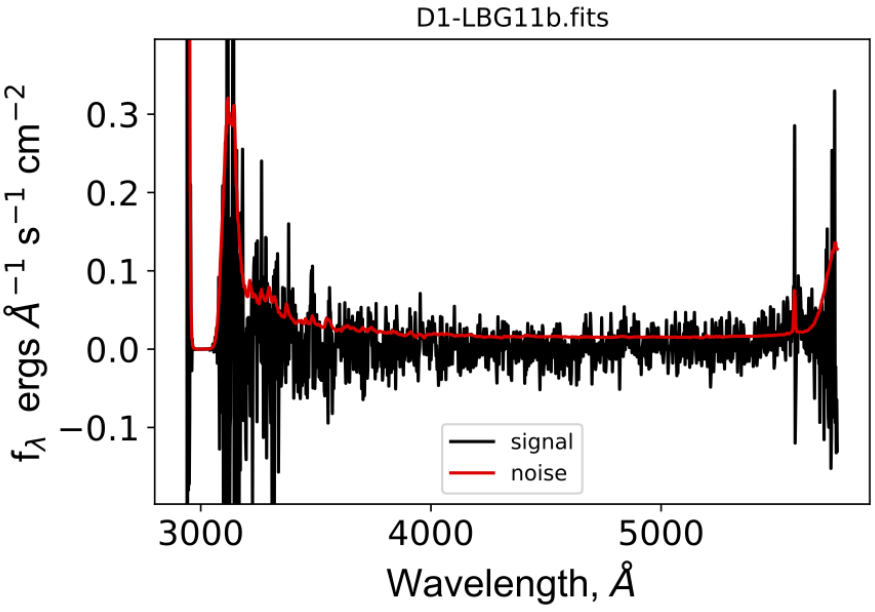

Figure 7.121 *Optical/UV spectrum of source 'D1-LBG11b', located in the CFHTLS D1 field.*

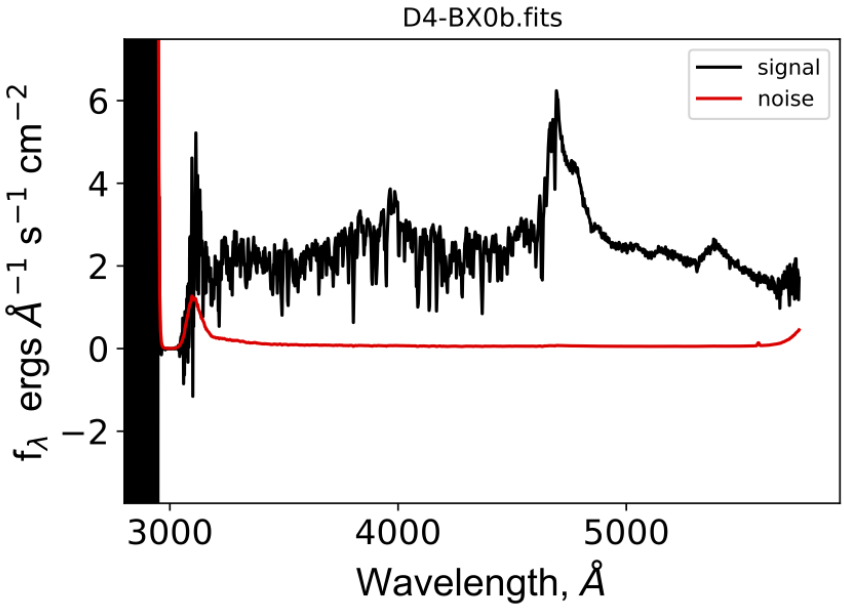

Figure 7.122 *Optical/UV spectrum of source 'D4-BX0b', located in the CFHTLS D4 field.*

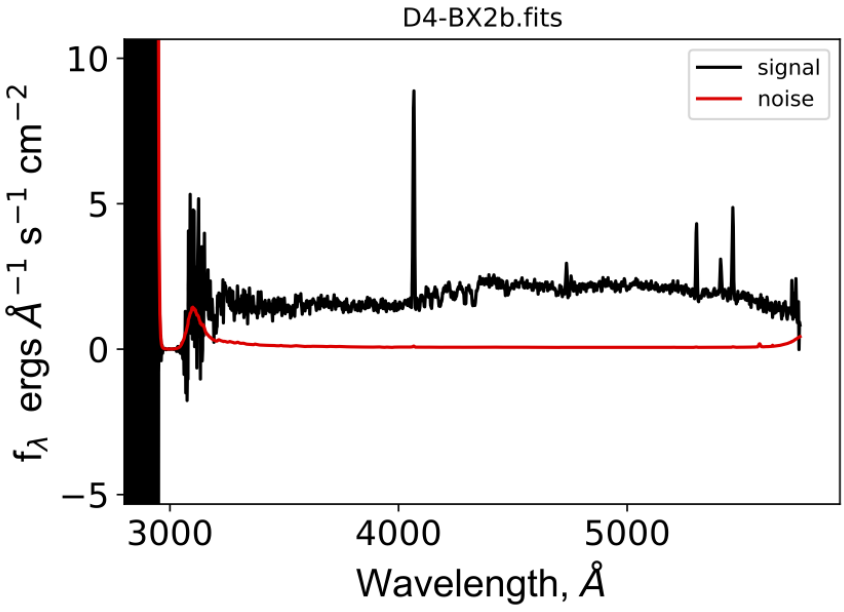

Figure 7.123 *Optical/UV spectrum of source 'D4-BX2b', located in the CFHTLS D4 field.*

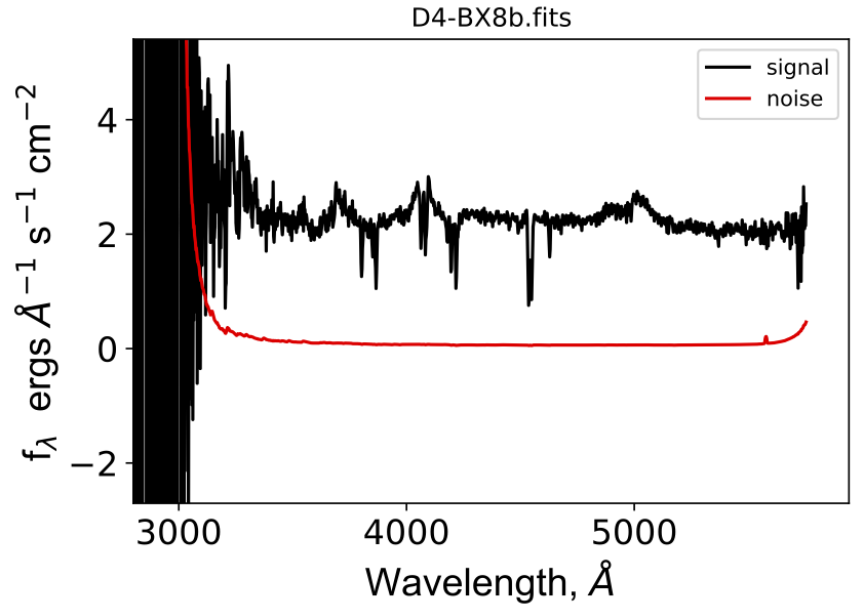

Figure 7.124 *Optical/UV spectrum of source 'D4-BX8b', located in the CFHTLS D4 field.*

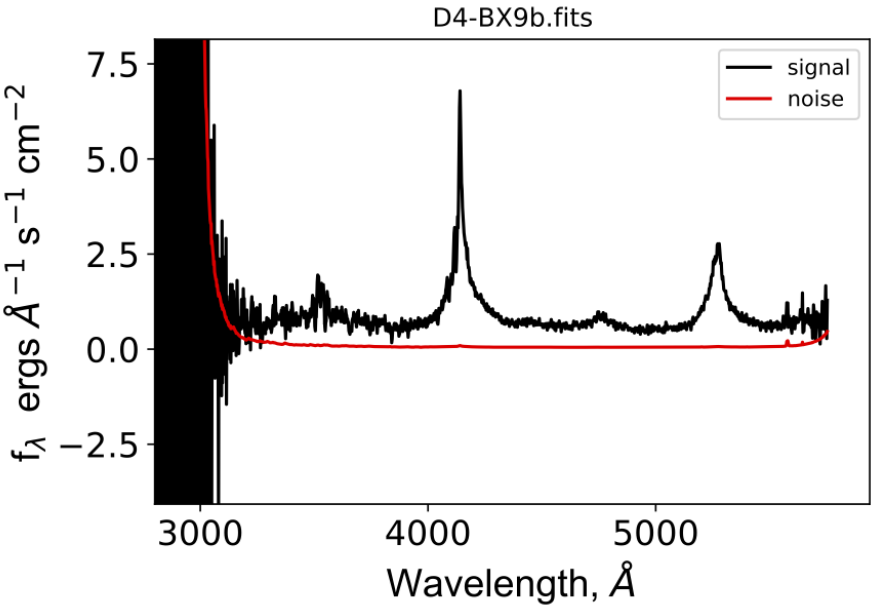

Figure 7.125 *Optical/UV spectrum of source 'D4-BX9b', located in the CFHTLS D4 field.*

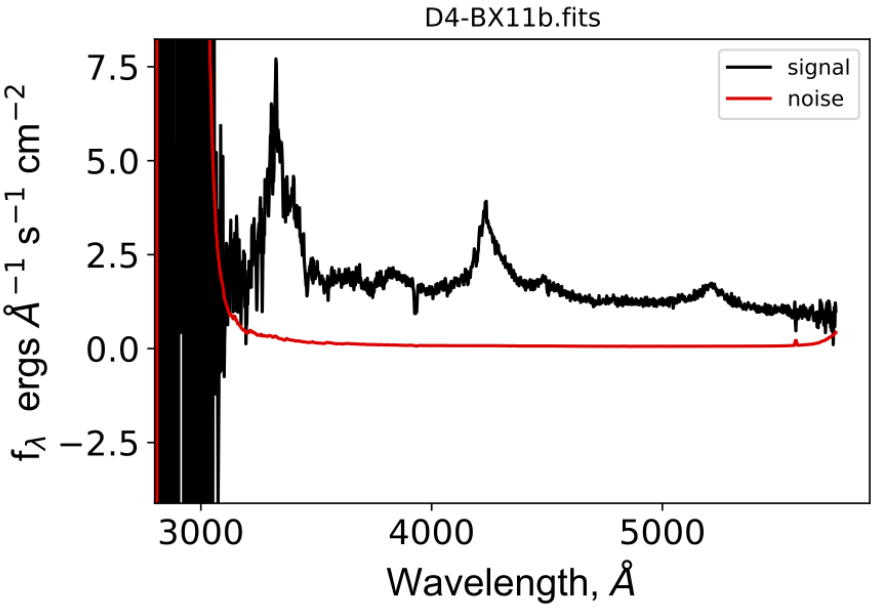

Figure 7.126 *Optical/UV spectrum of source 'D4-BX11b', located in the CFHTLS D4 field.*

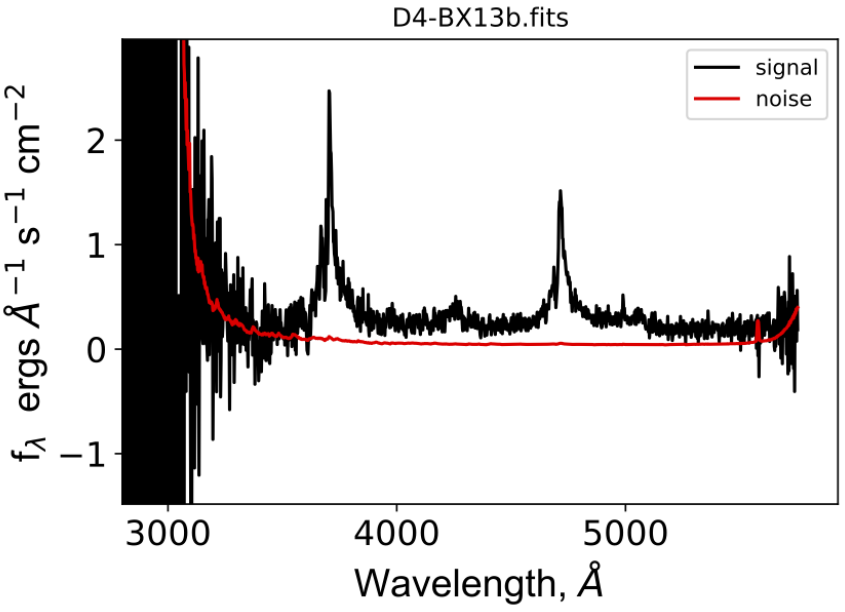

Figure 7.127 *Optical/UV spectrum of source 'D4-BX13b', located in the CFHTLS D4 field.*

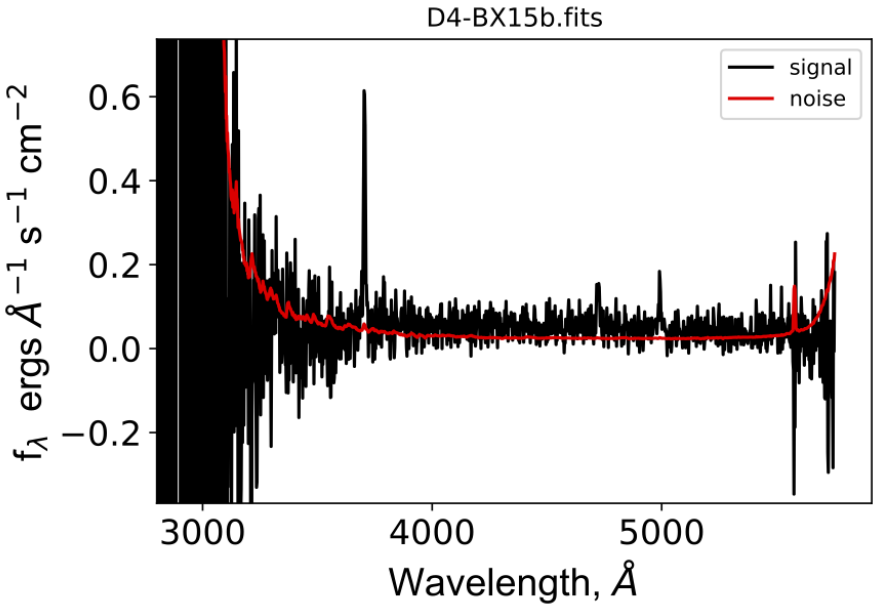

Figure 7.128 *Optical/UV spectrum of source 'D4-BX15b', located in the CFHTLS D4 field.*

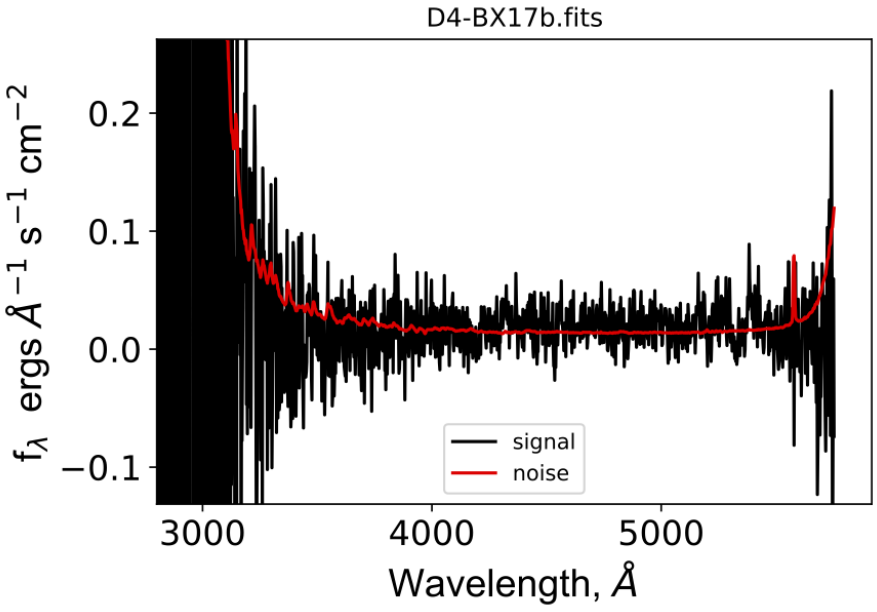

Figure 7.129 *Optical/UV spectrum of source 'D4-BX17b', located in the CFHTLS D4 field.*

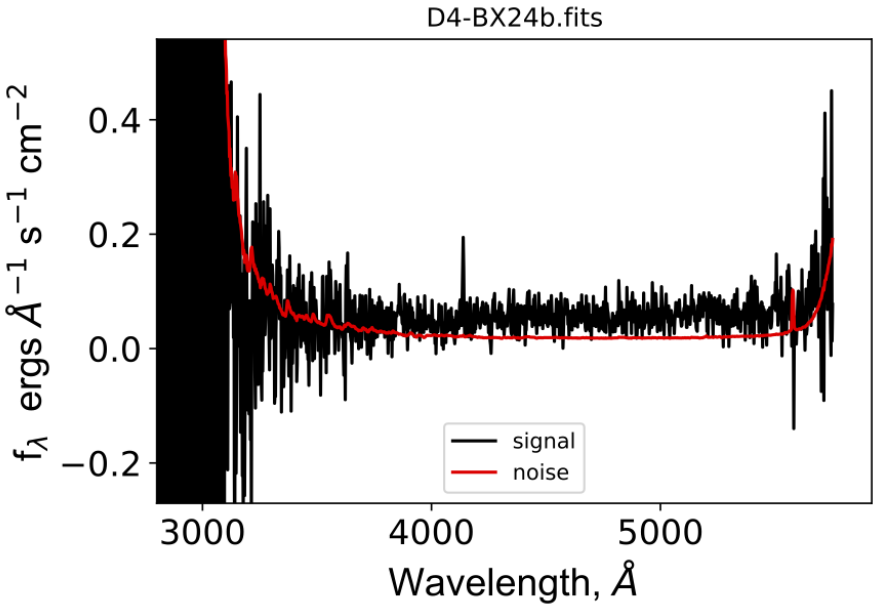

Figure 7.130 *Optical/UV spectrum of source 'D4-BX24b', located in the CFHTLS D4 field.*

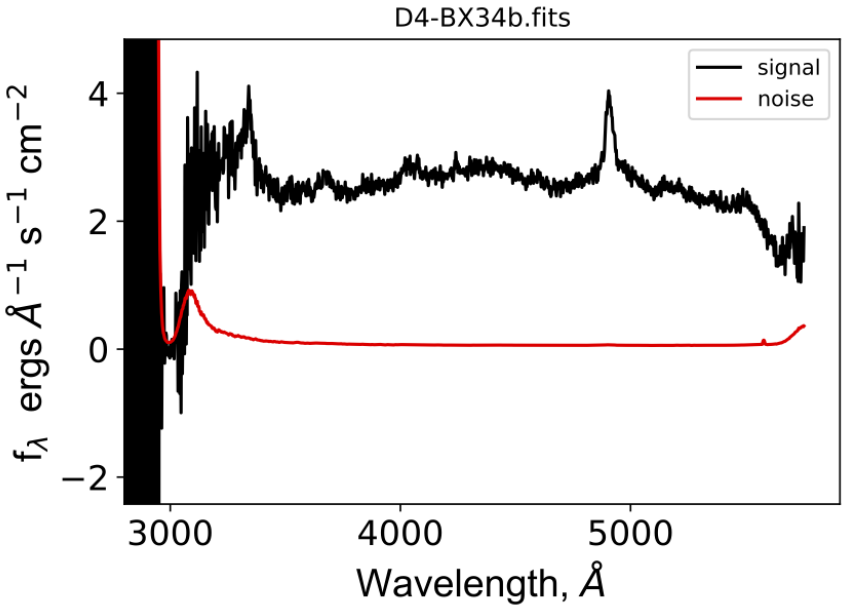

Figure 7.131 *Optical/UV spectrum of source 'D4-BX34b', located in the CFHTLS D4 field.*

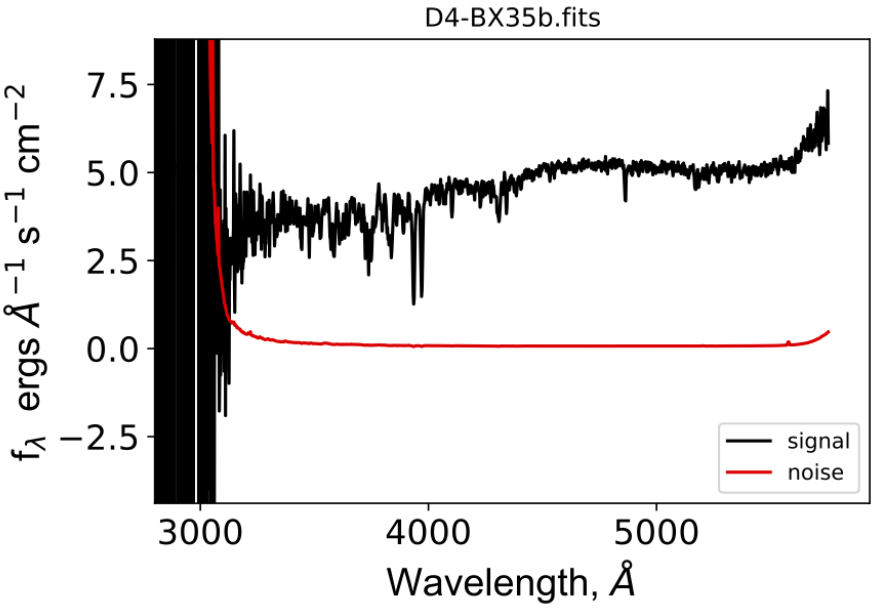

Figure 7.132 *Optical/UV spectrum of source 'D4-BX35b', located in the CFHTLS D4 field.*

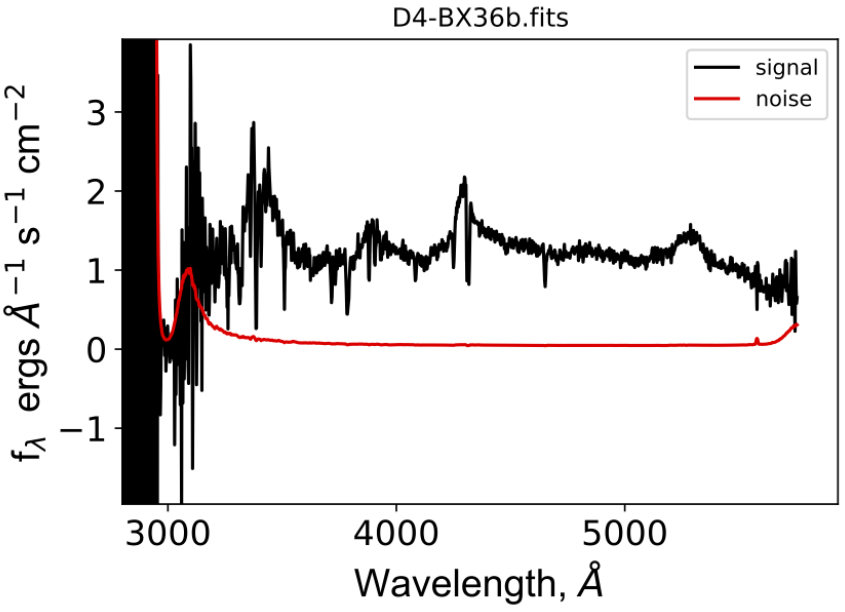

Figure 7.133 *Optical/UV spectrum of source 'D4-BX36b', located in the CFHTLS D4 field.*

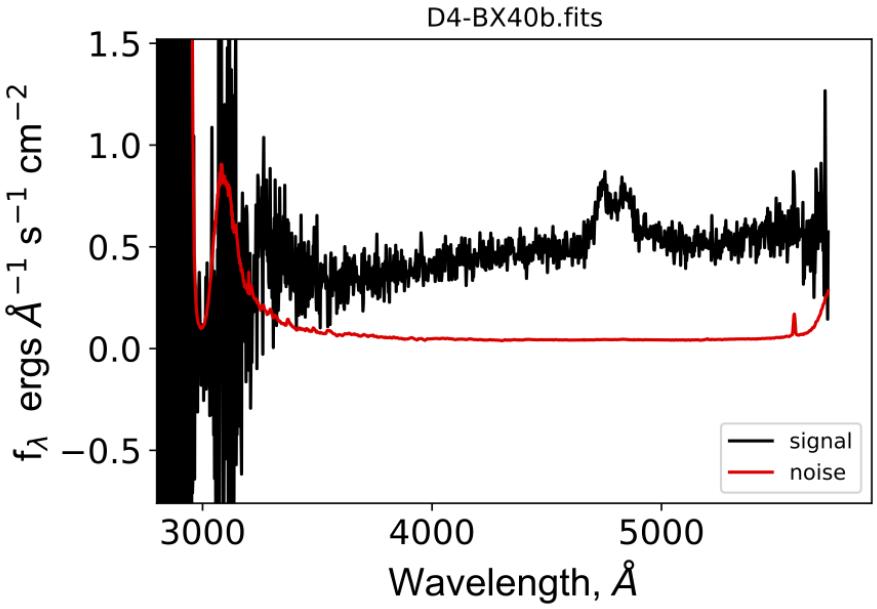

Figure 7.134 *Optical/UV spectrum of source 'D4-BX40b', located in the CFHTLS D4 field.*

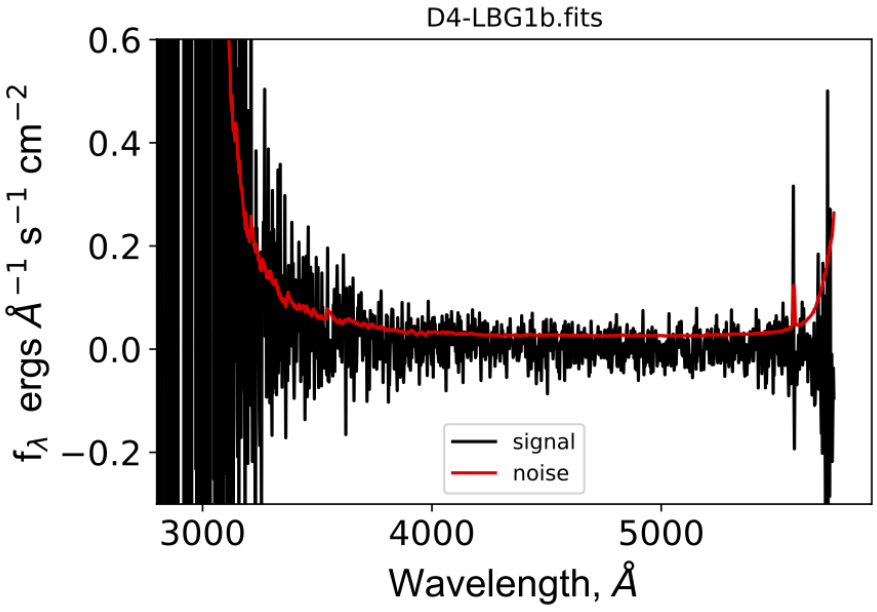

Figure 7.135 *Optical/UV spectrum of source 'D4-LBG1b', located in the CFHTLS D4 field.*

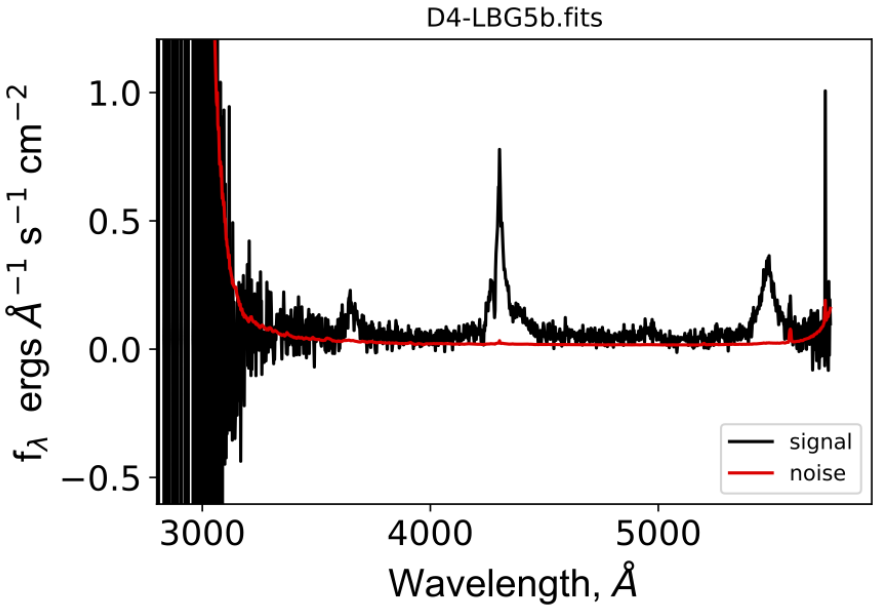

Figure 7.136 *Optical/UV spectrum of source 'D4-LBG5b', located in the CFHTLS D4 field.*

Bibliography

Abbott T., et al., 2016, Physical Review D, 94, 022001

Adelberger K. L., Steidel C. C., Shapley A. E., Hunt M. P., Erb D. K., Reddy N. A., Pettini M., 2004, ApJ, 607, 226

Aird J., 2015, The Astrophysical Journal, 815, p. 66

Aird J., Nandra K., Georgakakis A., Laird E. S., Steidel C. C., Sharon C., 2008, Monthly Notices of the RAS, 387, 883

Aird J., et al., 2010, Monthly Notices of the RAS, 401, 2531

Aird J., Coil A. L., Georgakakis A., Nandra K., Barro G., Pérez-González P. G., 2015, Monthly Notices of the RAS, 451, 1892

Alam S., et al., 2015, Astrophysical Journal, Supplement, 219, 12

Alexander D. M., et al., 2013, ApJ, 773, 125

Alig C., Schartmann M., Burkert A., Dolag K., 2013, ApJ, 771, 119

Ananna T. T., et al., 2022, Astrophysical Journal, Supplement, 261, 9

Antonucci R., 1993, Annual Review of Astron and Astrophysics, 31, 473

Balick B., Brown R. L., 1974, ApJ, 194, 265

Beckert T., Duschl W. J., 2004, AAp, 426, 445

Bertin E., Arnouts S., 1996, AAps, 117, 393

Blecha L., Snyder G. F., Satyapal S., Ellison S. L., 2018, Monthly Notices of the RAS, 478, 3056

Bolton A. S., et al., 2012, Astronomical Journal, 144, 144

Booth C. M., Schaye J., 2009, Monthly Notices of the RAS, 398, 53

Brandt W. N., Alexander D. M., 2015, The Astronomy and Astrophysics Review, 23

Brightman M., Ueda Y., 2012, Monthly Notices of the RAS, 423, 702

Buchner J., et al., 2015, ApJ, 802, 89

Buchner J., Salvato M., BudavaÌri T., Fotopoulou S., 2021, nway: Bayesian cross-matching of astronomical catalogs, Astrophysics Source Code Library, record ascl:2102.014 (ascl:2102.014)

Carilli C. L., Walter F., 2013, Annual Review of Astron and Astrophysics, 51, 105

Cisternas M., et al., 2011, ApJ, 726, 57

Civano F., et al., 2015, ApJ, 808, 185

Croton D. J., et al., 2006, Monthly Notices of the RAS, 365, 11

Di Matteo T., Springel V., Hernquist L., 2005, , 433, 604

Dovčiak M., Muleri F., Goosmann R. W., Karas V., Matt G., 2011, ApJ, 731, 75

Dubois Y., Pichon C., Devriendt J., Silk J., Haehnelt M., Kimm T., Slyz A., 2012, Monthly Notices of the Royal Astronomical Society, 428, 2885–2900

Dullemond C. P., van Bemmel I. M., 2005, AAp, 436, 47

Ebrero J., et al., 2009, AAp, 493, 55

Elitzur M., Ho L. C., 2009, ApJl, 701, L91

Elitzur M., Shlosman I., 2006, ApJl, 648, L101

Elyiv A., et al., 2012, AAp, 537, A131

Emmering R. T., Blandford R. D., Shlosman I., 1992, ApJ, 385, 460

Evans I., 2020, Chandra News, 28, 6

Evans P. A., et al., 2020, Astrophysical Journal, Supplement, 247, 54

Event Horizon Telescope Collaboration et al., 2022, ApJl, 930, L12

Fabian A. C., 2012, Annual Review of Astron and Astrophysics, 50, 455

Feltre A., Hatziminaoglou E., Fritz J., Franceschini A., 2012, Monthly Notices of the RAS, 426, 120

Ferrarese L., Merritt D., 2000, ApJl, 539, L9

Feruglio C., et al., 2015, AAp, 583, A99

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, , 125, 306

Fornasini F. M., et al., 2017, Astrophysical Journal, Supplement, 229, 33

Fotopoulou S., et al., 2016, AAp, 587, A142

Fritz J., Franceschini A., Hatziminaoglou E., 2006, Monthly Notices of the RAS, 366, 767

Gebhardt K., et al., 2000, ApJl, 539, L13

Gehrels N., 1986, ApJ, 303, 336

Georgakakis A., Nandra K., Laird E. S., Aird J., Trichas M., 2008, Monthly Notices of the RAS, 388, 1205

Gilli R., Comastri A., Hasinger G., 2007, AAp, 463, 79

Goodman J., Weare J., 2010, Communications in Applied Mathematics and Computational Science, 5, 65

Goulding A. D., Alexander D. M., 2009, Monthly Notices of the RAS, 398, 1165

Hainline K., Shapley A. E., Greene J. E., Steidel C. C., 2011, in American Astronomical Society Meeting Abstracts #217. p. 430.15

Häring N., Rix H.-W., 2004, ApJl, 604, L89

Harrison F. A., et al., 2013, ApJ, 770, 103

Harrison C. M., Alexander D. M., Mullaney J. R., Swinbank A. M., 2014, Monthly Notices of the RAS, 441, 3306

Harrison F. A., et al., 2016, ApJ, 831, 185

Hasinger G., 2008, AAp, 490, 905

Hickox R. C., Alexander D. M., 2018, Annual Review of Astron and Astrophysics, 56, 625

Hinton S., 2016, MARZ: Redshifting Program, Astrophysics Source Code Library, record ascl:1605.001 (ascl:1605.001)

Hong J., et al., 2016, ApJ, 825, 132

Hönig S. F., Beckert T., 2007, Monthly Notices of the RAS, 380, 1172

Hönig S. F., Kishimoto M., 2010, AAp, 523, A27

Hopkins P. F., Hernquist L., Cox T. J., Di Matteo T., Robertson B., Springel V., 2005, ApJ, 630, 716

Hopkins P. F., Hernquist L., Cox T. J., Kereš D., 2008, Astrophysical Journal, Supplement, 175, 356

Ishibashi W., Fabian A. C., 2016, Monthly Notices of the RAS, 463, 1291

K.Nandra E. L., Steidel C., 2005, Mon. Not. R. Astron. Soc, 360, L39

Khim H., Yi S. K., 2017, ApJ, 846, 155

King A., Pounds K., 2015, Annual Review of Astron and Astrophysics, 53, 115

Konigl A., Kartje J. F., 1994, ApJ, 434, 446

Krolik J. H., 2007, ApJ, 661, 52

La Franca F., et al., 2005, ApJ, 635, 864

Lagos C. D. P., Baugh C. M., Lacey C. G., Benson A. J., Kim H.-S., Power C., 2011, Monthly Notices of the RAS, 418, 1649

Lansbury G., et al., 2017a, in Ness J.-U., Migliari S., eds, The X-ray Universe 2017. p. 126

Lansbury G. B., et al., 2017b, ApJ, 836, 99

Lawrence A., 1991, Monthly Notices of the RAS, 252, 586

Lawrence A., Elvis M., 1982, ApJ, 256, 410

Lawrence A., Elvis M., 2010, ApJ, 714, 561

Liu, Teng et al., 2022, A&A, 661, A5

Lusso E., et al., 2013, ApJ, 777, 86

Lynden-Bell D., Rees M. J., 1971, Monthly Notices of the RAS, 152, 461

Maccacaro T., Gioia I. M., Wolter A., Zamorani G., Stocke J. T., 1988, ApJ, 326, 680

Madau P., 1995, ApJ, 441, 18

Magorrian J., et al., 1998, Astronomical Journal, 115, 2285

Malizia A., Bassani L., Bazzano A., Bird A. J., Masetti N., Panessa F., Stephen J. B., Ubertini P., 2012, Monthly Notices of the RAS, 426, 1750

Malkan M. A., Gorjian V., Tam R., 1998, Astrophysical Journal, Supplement, 117, 25

Marconi A., Hunt L. K., 2003, ApJl, 589, L21

Marocco F., et al., 2021, Astrophysical Journal, Supplement, 253, 8

Marsden C., Shankar F., Ginolfi M., Zubovas K., 2020, Frontiers in Physics, 8, 61

Masini A., et al., 2018, ApJ, 867, 162

Mateos S., et al., 2017, ApJl, 841, L18

Merloni A., et al., 2014, Monthly Notices of the RAS, 437, 3550

Miyaji T., Hasinger G., Schmidt M., 2000, AAp, 353, 25

Miyaji T., Hasinger G., Schmidt M., 2001, AAp, 369, 49

Miyaji T., et al., 2015, ApJ, 804, 104

Monet D. G., et al., 2003, Astronomical Journal, 125, 984

Mori K., et al., 2015, ApJ, 814, 94

Morton D. C., 1991, Astrophysical Journal, Supplement, 77, 119

Mullaney J. R., et al., 2012, Monthly Notices of the RAS, 419, 95

Mullaney J. R., et al., 2015, ApJ, 808, 184

Nandra K., et al., 2015, VizieR Online Data Catalog, p. J/ApJS/220/10

Nenkova M., Sirocky M. M., Ivezić Ž., Elitzur M., 2008, ApJ, 685, 147

Netzer H., 2015, Annual Review of Astron and Astrophysics, 53, 365

Nidever D. L., et al., 2021, Astronomical Journal, 161, 192

Padovani P., et al., 2017, AApr, 25, 2

Page M. J., Carrera F. J., 2000, Monthly Notices of the RAS, 311, 433

Peca A., et al., 2023, ApJ, 943, 162

Peterson B. M., 1997, An introduction to active galactic nuclei. Cambridge University Press, Cambridge

Pier E. A., Krolik J. H., 1992, ApJ, 401, 99

Pierre M., et al., 2007, Monthly Notices of the RAS, 382, 279

Prieto M. A., Mezcua M., Fernández-Ontiveros J. A., Schartmann M., 2014, Monthly Notices of the RAS, 442, 2145

Richardson M. L. A., Scannapieco E., Devriendt J., Slyz A., Thacker R. J., Dubois Y., Wurster J., Silk J., 2016, The Astrophysical Journal, 825, 83

Roseboom I. G., Lawrence A., Elvis M., Petty S., Shen Y., Hao H., 2013, Monthly Notices of the RAS, 429, 1494

Ross N. P., et al., 2013, ApJ, 773, 14

Sanders D. B., Soifer B. T., Elias J. H., Madore B. F., Matthews K., Neugebauer G., Scoville N. Z., 1988, ApJ, 325, 74

Schmidt M., 1968, ApJ, 151, 393

Schmidt M., 1983a, in Philip A. G. D., Upgren A. R., eds, IAU Colloq. 76: Nearby Stars and the Stellar Luminosity Function. p. 155

Schmidt M., 1983b, in Philip A. G. D., Upgren A. R., eds, IAU Colloq. 76: Nearby Stars and the Stellar Luminosity Function. p. 155

Schulze A., et al., 2015, Monthly Notices of the RAS, 447, 2085

Shapley A. E., Steidel C. C., Pettini M., Adelberger K. L., 2003, ApJ, 588, 65

Silk J., Mamon G. A., 2012, Research in Astronomy and Astrophysics, 12, 917

Silk J., Rees M. J., 1998, AAp, 331, L1

Simpson C., 2005, Monthly Notices of the RAS, 360, 565

Stalevski M., Fritz J., Baes M., Nakos T., Popović L. Č., 2012, Monthly Notices of the RAS, 420, 2756

Steidel C. C., Adelberger K. L., Shapley A. E., Pettini M., Dickinson M., Giavalisco M., 2003, ApJ, 592, 728

Traulsen I., et al., 2020, VizieR Online Data Catalog, p. IX/61

Treister E., Urry C. M., 2006, ApJl, 652, L79

Ueda Y., Akiyama M., Ohta K., Miyaji T., 2003, ApJ, 598, 886

Ueda Y., Akiyama M., Hasinger G., Miyaji T., Watson M. G., 2014, ApJ, 786, 104

Urry M., 2003, in Collin S., Combes F., Shlosman I., eds, Astronomical Society of the Pacific Conference Series Vol. 290, Active Galactic Nuclei: From Central Engine to Host Galaxy. p. 3 (arXiv:astro-ph/0301309), doi:10.48550/arXiv.astro-ph/0301309

Urry C. M., Padovani P., 1995, , 107, 803

Vietri G., et al., 2022, AAp, 659, A129

Vito F., et al., 2013, Monthly Notices of the RAS, 428, 354

Vollmer B., Beckert T., Davies R. I., 2008, AAp, 491, 441

Wada K., 2012, ApJ, 758, 66

Webb N. A., et al., 2020, AAp, 641, A136

Yencho B., Barger A. J., Trouille L., Winter L. M., 2009, ApJ, 698, 380

et al. A. B., 2007, Astronomy and Astrophysics, 474, 755