CONSTRAINING THE HIGH-REDSHIFT QUASAR LUMINOSITY FUNCTION

JASMINE LILIANE BARNARD FINER

SUPERVISOR: DR. DANIEL MORTLOCK

Submitted in partial fulfilment of the requirements for the degree of

MASTER OF PHILOSOPHY

Astrophysics Group
Department of Physics
Imperial College London

COPYRIGHT DECLARATION

DECLARATION OF ORIGINALITY

I declare that the work contained in this thesis is my own original work, and any sources used in the writing of this thesis are appropriately referenced.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. Daniel Mortlock, Imperial College London, for his continued support throughout my studies. I would also like to thank my internal examiner Prof. Steve Warren, Imperial College London, and my external examiner Dr. Elizabeth Stanway, University of Warwick for their efforts.

I extend my sincere gratitude to the members of the Imperial College Astrophysics Group, the academic staff and the administrative staff without whom I would not have been able to submit this thesis.

I would also like to thank the Science and Technology Facilities Council for providing me with the funding without which I could not have completed this research.

Finally, I would like to thank Thomas Kershaw for his endless encouragement and support.

ABSTRACT

Accretion onto supermassive black holes powers quasars so luminous that they can be observed at high redshifts. However, mass estimates of these quasars are in conflict with current theories of black hole growth during these early epochs. The root of this inconsistency lies within our theoretical understanding of black hole formation and evolution at high redshift, methods of estimating high redshift black hole masses, and the statistical approach to constraining the high redshift black hole population. This thesis focuses on the latter. We propose a correction to the completeness of the $z \sim 6$ Stripe 82 survey (Jiang et al., 2009) in the form of a shift in magnitude limit of $\Delta M_{1450} = -0.47$. We present an updated formalism of parameter inference for the quasar luminosity function, using a 6 parameter double power law model

$$\Phi(M_{1450}, z) = \frac{10^{k(z-z_{\rm ref})}\Phi^*}{10^{0.4(\alpha+1)(M_{1450}-M^*_{1450}-j(z-z_{\rm ref}))} + 10^{0.4(\beta+1)(M_{1450}-M^*_{1450}-j(z-z_{\rm ref}))}}, \tag{1}$$

and find maximum a posteriori estimates $\beta = -3.76,\ M^*_{1450} = -26.5,\ j = 0.0849,\ \Phi^* = 3.25 \times 10^{-9}\ \mathrm{Mpc}^{-3},\ \alpha = -1.73,\ k = -0.165$. In comparison to previous work on the redshift 6 quasar luminosity function (Willott et al., 2010a), our modelling prefers a steeper fall off at the QLF's bright end, a brighter break magnitude which evolves slowly with redshift and a more gradual decrease in number density with redshift.

CONTENTS

1 Introduction **9**

1.1 Characteristics of active galactic nuclei 9

1.2 Supermassive black holes at high redshift 11

2 The Quasar Luminosity Function **17**

2.1 Double power law QLF 18

2.2 The black hole mass function 20

3 Quasar Surveys **22**

3.1 The Sloan Digital Sky Survey 22

3.2 $z \sim 5$ quasar surveys 23

3.3 $z \sim 6$ quasar surveys 25

4 QLF parameter estimation methodology **29**

4.1 The QLF likelihood 29

4.2 Simulating quasar surveys 32

4.3 Maximum likelihood estimation 32

4.3.1 The ΔS Statistic 34

4.4 Bayesian Parameter Inference 35

4.4.1 Highest posterior density regions 36

4.4.2 Markov-Chain Monte Carlo 36

4.5 Grid method for low numbers of parameters 41

5 Parameter estimation of the QLF **43**

5.1 Results from maximum likelihood estimation 43

5.2 Stripe 82 Completeness 45

5.3 Treatment of overlapping surveys 46

5.4 Maximum likelihood estimates for corrected selection function 47

5.5 Prior Distribution on QLF parameters 48

5.6 MCMC analysis of QLF parameters 48

6 Discussion **58**

List of Figures

1.1 The spectrum of the first quasar, 3C 273. 10

1.2 The geometry of reverberation mapping. 14

2.1 The evolution of the comoving spatial density of quasars with $M_{1450} < -26.8$. 18

2.2 The Eddington ratio distribution at $z \sim 6$. 21

3.1 Spectra of 9 high redshift quasars. 23

3.2 The photometric bands of the Sloan Digital Sky Survey. 24

3.3 Selection functions of the $z \sim 5$ quasar surveys. 25

3.4 Selection functions of $z \sim 6$ quasar surveys. 28

4.1 Results from simulating high redshift quasar surveys. 33

4.2 Highest posterior density regions. 37

4.3 MCMC trace plots. 38

4.4 Log-likelihood of the MCMC chain. 39

4.5 Autocorrelation function calculated using the emcee algorithm. 40

4.6 The autocorrelation function of an MCMC chain. 41

5.1 Results of maximum likelihood analysis of the 3 parameter QLF model. 43

5.2 Results of maximum likelihood analysis of the 4 parameter QLF model. 44

5.3 The effective completeness of the deep Stripe 82 survey. 45

5.4 The corrected effective completeness of the deep Stripe 82 survey. 46

5.5 Positions of quasars discovered in the two Stripe 82 surveys. 47

5.6 Corrected selection functions after accounting for overlapping area and shift in magnitude limit of the deep Stripe 82 survey. 47

5.7 Marginalised posterior distributions of the 3 parameter QLF at $z \sim 6$. 49

5.8 Marginalised posterior distributions of the 3 parameter QLF at $z \sim 6$ parameterising the normalisation with $\Phi_{26} = \Phi(-26, 6)$ rather than $\Phi^* = \Phi(M^*_{1450}, 6)$. Contours are $(99.7\%, 95.4\%, 90\%, 68.3\%$ highest posterior density regions. Blue points are the maximum likelihood values of $M^*_{1450} = -25.13,\ \beta = -2.81,\ \Phi_{26} = 0.1980.1$ Gpc^{-3} with α fixed at $\alpha = -1.5$. 50

5.9 Results from MCMC analysis of 3 parameter QLF at $z \sim 6$. 51

5.10 Marginalised posterior distributions of the 4 parameter QLF at $z \sim 5$. 52

5.11 Results from MCMC analysis of 4 parameter QLF at $z \sim 5$. 53

5.12 Corrected selection functions and simulated quasar samples. 54

5.13 Corner plot showing results from MCMC analysis of double power law QLF with 6 free parameters, and simulated quasar samples. 55

5.14 Corner plot showing results from MCMC analysis of double power law QLF with 6 free parameters. 56

5.15 Plots of the QLF for different sets of parameters sampled from the posterior distribution. 57

LIST OF TABLES

3.1 $z \sim 6$ SDSS main sample. 26

3.2 Complete sample of $z \sim 6$ quasars discovered in the SDSS Stripe 82. 26

3.3 $z \sim 6$ quasars from the CHFQS Survey. 27

1 Introduction

This thesis is concerned with an outstanding question of quasar formation. Quasars are some of the most energetic objects in our universe. They are so luminous that they can be seen in the very early universe, where other astrophysical sources are typically too faint to be observed, leaving large uncertainty about astrophysical processes during this epoch. It is thought that quasars form by accretion onto supermassive black holes, as such, the luminosity of a quasar is closely related to the mass of the central black hole and observing quasar luminosities and emission lines allows us to place estimates on their respective black hole masses. At the same time, theoretical models of black hole evolution and quasar formation have been formulated and in general are unable to explain their respective black hole masses. This thesis investigates possible origins of this inconsistency.

1.1 Characteristics of active galactic nuclei

In 1963, Maarten Schmidt discovered an unresolved 13^{th} magnitude radio source 3C 273 (Schmidt, 1963). He found that the spectrum of the object (Fig. 1.1) was most easily explained by a high redshift, $z \sim 0.16$, meaning the radio source must have an extremely high luminosity. This new object was the first quasar. Within a few years quasars had been discovered up to $z \sim 2$ and were understood to be some of the most energetic and luminous objects in the universe. The mechanism leading to the formation of quasars was not understood for several years after their discovery when it was realised that quasars could be described by accretion of material onto supermassive black holes (e.g. Rees (1978, 1984)).

Quasars are believed to be subset of a class of objects called active galactic nuclei (AGNs). All AGNs exhibit certain characteristics and it is believed that differences between different subsets arise from viewing angle rather than physical processes (Antonucci, 1993). All AGNs exhibit continuum emission due to thermal emission from the accretion disk surrounding the central black hole. Quasars are the very brightest of these objects with the most massive central supermassive black holes. Assuming the geometrically thin, optically thick model of an accretion disk around a black hole of mass $M_{\rm BH}$, the energy released when a mass m falls from radius $r + \Delta r$ to r is

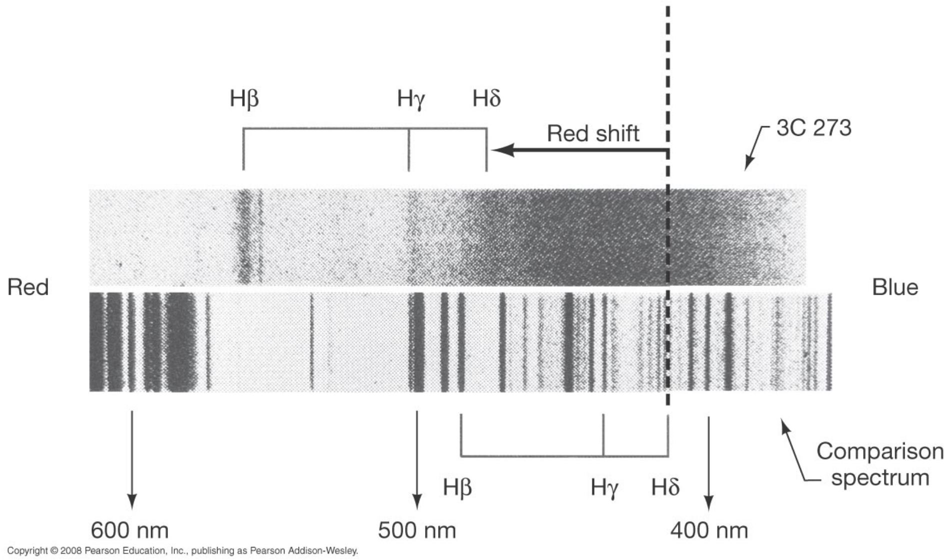


Figure 1.1: The spectrum of quasar 3C 273 discovered by Schmidt (1963). This figure is reproduced from Schmidt (1963)

$$\Delta E \approx \frac{GM_{\rm BH}m}{r}\frac{\Delta r}{r}.$$

From the virial theorem, half of this energy is converted to heat with corresponding lumionsity

$$\Delta L = \frac{GM_{\rm BH}\dot{m}}{2r^2}\Delta r, \tag{1.1}$$

where $\dot{m}$ denotes the accretion rate. If the disk is optically thick, we can assume black body emission and a differential ring between radius r and $r + \Delta r$ has a luminosity

$$\Delta L = 2 \times 2\pi\ r\ \Delta r\ \sigma_{\rm SB}T^4(r), \tag{1.2}$$

where $\sigma_{\rm SB} = 5.67 \times 10^{-8}$ W m^{-2} K^{-4} is the Stefan-Boltzmann constant and $T(r)$ is the radial temperature profile of the disk. The factor of 2 comes from the fact that the disc radiates from both sides.

Combining eq. (1.1) and eq. (1.2) we can obtain the radial temperature profile of the accretion disk,

$$T(r) = \left(\frac{3GM_{\rm BH}\dot{m}}{8\pi\sigma_{\rm SB}r^3}\right)^{\frac{1}{4}} = \left(\frac{3GM_{\rm BH}\dot{m}}{8\pi\sigma_{\rm SB}r_S^3}\right)^{\frac{1}{4}}\left(\frac{r}{r_S}\right)^{-\frac{3}{4}}, \tag{1.3}$$

where in the last step we have replaced r with the Schwarzchild radius of a black hole $r_s = 2GM_{\rm BH}/c^2 \approx 3 \times 10^3 M_{\rm BH}/M_\odot$ m, and the extra factor of 3 comes from considering dissipation by friction in the disk.

The thermal emission of a geometrically thin, optically thick accretion disk produces a very broad spectrum of continuum emission with large flux at infrared, optical and ultraviolet wavelengths, peaking in the UV. This continuum emission is variable on timescales, $t_{\rm var}$ as short as one week. This provides empirical evidence for size of the nuclear region of the quasar to be less than $r = c\, t_{\rm var} \simeq 1000$ AU, where the last equality holds for $t_{\rm var} = 1$ week. These sizes are much smaller than that of the host galaxy, and are extremely difficult to resolve, thus these objects are seen as point sources.

Another important characteristic of all AGNs (except for one subset, BL Lacs whose spectrum shows almost no emission lines) is the presence of strong emission lines, which are extremely broad in quasars and arise from regions close to the central engine, the broad line region (BLR). These lines typically have $\Delta\lambda/\lambda \gtrsim 0.03$ corresponding to $\Delta v \lesssim 10,000$ km s^{-1}. This value of Δv is inconsistent with thermal line broadening requiring temperatures $T \sim 10^{10}$K at which all atoms would be fully ionised and no emission lines would be produced. Instead, these lines are thought to be Doppler broadened as a result of rotational velocities from Keplerian motion,

$$v_{\rm rot} \sim \sqrt{\frac{GM_{\rm BH}}{r}} = \frac{c}{\sqrt{2}}\left(\frac{r}{r_S}\right)^{-\frac{1}{2}},$$

and hence must arise from regions close to the black hole where orbital speeds are the greatest. To obtain the required broadening $v_{\rm rot} = 10,000$ km s^{-1}, the radial distance must be $r \sim 500\, r_S$.

By examining ratios of allowed and forbidden transitions that occur in the BLR (described in detail in Peterson (1997), but beyond the scope of this thesis), it is possible to determine the particle density, $n_e \sim 1 \times 10^{10}$ cm^{-3}, of the BLR as well as the proportion of the volume that contains gas that produces emission lines, the filling factor, $\epsilon \lll 1$. The general picture of the BLR is then that it is a small region close to the central engine, made up of dense clouds of gas. These clouds absorb and re-emit continuum radiation from the source, and Doppler broadened emission lines are observed in the spectrum of the AGN.

Most AGNs (again with the exception of BL Lacs) also exhibit narrow emission lines with typical $\Delta v \sim 400$ km s^{-1}. This suggests that the narrow line region (NLR) is much larger than the BLR. Similar arguments lead to the belief that the NLR is also made up of clouds of gas that are less dense than those in the BLR , with typical values of $n_e \sim 10^4$ cm^{-3}.

1.2 SUPERMASSIVE BLACK HOLES AT HIGH REDSHIFT

Currently, over $200,000$ quasars have been catalogued, with most discovered in the Sloan Digital Sky Survey (York et al., 2000). Most of these quasars have been discovered with redshifts $z < 4$, however there are now ~ 200 quasars with $z \gtrsim 5$ including the most distant known quasar at $z \sim 7.1$ (Mortlock et al., 2011).

If mass estimates of the $z > 6$ supermassive black holes powering these objects (Willott et al., 2010; Wu et al., 2015) are correct, then they suggest that the black holes have been able to accrete up to $\sim 10^{10}\ M_{\odot}$ between hydrogen recombination at $z \sim 1100$ and $z \sim 7$, less than 10^9 years. This would require either a very large seed black hole of $M \gtrsim 10^5\ M_{\odot}$ (Begelman et al., 2006) or a very high accretion rate (Abramowicz, 2005).

BLACK HOLE SEEDS

Black hole seeds are currently thought to be the remnants of the first generation stars (Pop III stars), which formed out of the primordial hydrogen and helium gas and therefore have zero metallicity. A second possibility would be that black hole seeds form as a result of direct gravitational collapse of gas and dust in high density regions of the universe (Begelman et al., 2006) without ever having undergone stellar burning. The former of these is thought to produce seeds of $\lesssim 100\ M_{\odot}$ (Zhang et al., 2008). The latter model would provide a sufficiently massive seed ($M_{\rm BH} \sim 10^5\ M_{\odot}$), however these objects are predicted to be rarer and unlikely to make up the entire population of high redshift supermassive black holes.

THE EDDINGTON LIMIT

The Eddington limit defines a maximum luminosity for a spherically symmetric system of accretion of matter onto a central dense source. This limit is the result of radiation pressure due to photons released as potential energy of the in-falling material is converted to heat and light, which then streams outwards, away from the black hole. The radiation pressure for a spherically symmetric system is given by,

$$P_{\rm rad} = \frac{L}{c}\frac{1}{4\pi R^2},$$

where L is the luminosity of the source. This pressure exerts a force $F_{\rm rad} = P_{\rm rad}\kappa m$, where κ is the opacity of the surrounding medium and m is the mass of in-falling material. If we make the assumption that the accreted material is mostly ionised hydrogen, i.e. protons coupled to electrons, then the opacity is provided by Thomson scattering, with cross section σ_T, by electrons. The force then has the largest effect on protons, with mass $m_{\rm p} \gg m_{\rm e}$. The opacity is then given by $\kappa = \sigma_T / m_{\rm p}$. The Eddington limit becomes important when this force is comparable to the gravitational force exerted on the in-falling material,

$$F_{\rm grav} = \frac{GMm}{R^2}.$$

The Eddington Luminosity is defined as the value of $L = L_{\rm edd}$ for which these forces are equal. When $L \geq L_{\rm Edd}$ the force due to radiation pressure from photons is equal to or larger than the gravitational force and matter is unable to reach the surface of the black hole. $L_{\rm edd}$ is given by

$$L_{\rm edd} = \frac{4\pi G M m_{\rm p} c}{\sigma_T}.$$

We can relate the mass accretion rate to the luminosity of the quasar via

$$L = \epsilon \dot{M}_{\rm acc} c^2 = \frac{\epsilon}{1-\epsilon} \dot{M}_{\rm BH} c^2,$$

where $\dot{M}_{\rm acc}$ is the rate at which matter is accreted onto the black hole and ϵ is the fraction of rest-mass energy that is radiated away. The fraction of mass that is therefore gained by the black hole is given by: $M_{\rm BH} = M_{\rm acc}/(1-\epsilon)$.

We can obtain an upper bound on the mass of the SMBH as a function of time powering a quasar by assuming Eddington limited accretion for the lifetime of the quasar. The mass of the black hole as a function of time is given by (Volonteri, 2010),

$$M_{\rm BH}(t) = M_{\rm BH}(t_0) \exp\left\{ \frac{1-\epsilon}{\epsilon} \frac{t-t_0}{T_{\rm Edd}} \right\},$$

where $T_{edd} \sim 0.45$ Gyr. For a radiative efficiency, $\epsilon = 0.1$ for a spherically symmetric non-rotating black hole, and an initial seed mass of $M_{\rm BH}(t_0) = 10^5\ M_\odot$ (Begelman et al., 2006), it will take ~ 500 Myr for the black hole to reach $10^9\ M_\odot$. We know of the existence of supermassive black holes at redshift 7, which corresponds to ~ 780 Myr after the Big Bang. As such, it would, theoretically, be possible for these objects to form in time via accretion if they accrete at the Eddington limit for their entire life and if a large enough seed can be produced sufficiently early (i.e. within the first 280 Myr of the Big Bang).

SUPER EDDINGTON ACCRETION

In principle, this is not a hard limit and theories of super Eddington accretion have been developed (Paczynski, 1998; Abramowicz, 2005) in which rather than a friction dominated thin accretion disk, advection dominated accretion flows (ADAFs) are present in which the radiation pressure from photons is suppressed and super Eddington accretion rates can be achieved, even when luminosities remain sub-Eddington.

MASS ESTIMATES OF SUPERMASSIVE BLACK HOLES

As discussed in section 1.1, the continuum emission of quasars is variable even on short timescales. Thus, it is possible to exploit the response of the BLR to this continuum variability to determine its size using reverberation mapping (Vestergaard & Peterson, 2006; Peterson, 2010). Using a simplistic model of the BLR as a shell of gas clouds surrounding the central engine, it becomes clear that a time lag, τ, will be measured between the response of clouds closer to the observer and those further away.

Consider a cloud a distance R away from the central black hole at position $(0,0)$, Fig. 1.2.

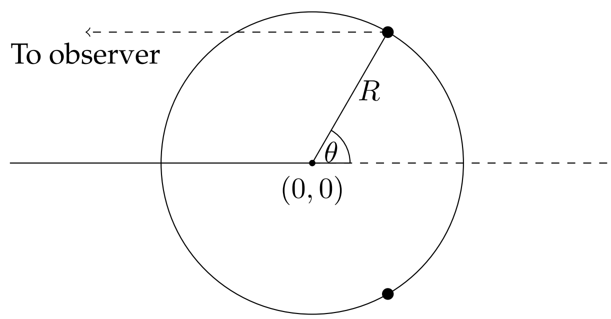


Figure 1.2: Graphic depicting a shell of gas clouds surrounding the central black hole, at position $(0,0)$ and the response of gas clouds at position (R,θ) to a continuum burst, (after Peterson (1997)).Sent

This cloud will reemit a continuum burst at a time lag, τ, corresponding to the path difference, δ between the light from the central engine and the emission lines from the BLR cloud, i.e.

$$\delta = R + R\cos\theta$$

The time lag will therefore be given by

$$\tau = \frac{R}{c}(1+\cos\theta).$$

By measuring these time delays it is possible to produce maps of the BLR and estimate its size, allowing us to make virial mass estimates of the central black hole, defined as

$$M_{\rm BH} = f\frac{R\Delta V^2}{G}, \tag{1.4}$$

where R is the size of the BLR, ΔV is the emission line width and f is a scale factor that depends on the geometry of the BLR and is of order unity (Vestergaard & Peterson, 2006).

Reverberation mapping has been used to estimate the masses of many black holes at low redshift

(Vestergaard & Peterson, 2006; Vestergaard & Osmer, 2009; Peterson, 2010), and various scaling relations have been determined allowing mass estimates to be made without having to use reverberation mapping, a very resource expensive process. These scaling relations rely on observation that there is a relation between the characteristic size, R of a region producing certain broad emission lines and the continuum luminosity that excites the gas in this region, for example, $R(\mathrm{H}\beta) \propto L_{1450}^{0.53}$ (Kaspi et al., 2005). This $R - L$ relation allows for the quasar luminosity to be used as a proxy for R in equation eq. (1.4). The line width of the MgII line can then be related to black hole mass, e.g., by the following empirical relation (Vestergaard & Osmer, 2009),

$$\log M_{\mathrm{BH}} = 6.68 + 2\log\frac{\mathrm{FWHM(MgII)}}{1000\ \mathrm{km\ s^{-1}}} + 0.5\log\frac{L_{3000}}{10^{37}\ W}.$$

This relation was used by Willott et al. (2010) to estimate black hole masses of $z \sim 6$ quasars. Note other empirical scaling relations exist for different broad lines and continuum luminosities (Vestergaard & Peterson, 2006). However, the assumption that these relations can be extrapolated to high redshift may not hold and it is a complicated process to treat the uncertainties that arise in all variables.

We now have an overview of the assumed physical mechanism driving observed quasar luminosities. As the observed luminosities lead to inconsistencies with theoretically predicted mass ranges, it is clear that somewhere an assumption has to be violated. These might be one or more of the following:

- BLACK HOLE SEEDS The mechanisms by which black hole seeds are produced are not well understood. If it is possible to produce very massive black hole seeds (i.e. by direct gravitational collapse) less time is needed for accretion.
- EVOLUTIONARY MODELS The development of accretion models that allow for radiatively inefficient accretion onto black holes or super-Eddington accretion would allow black holes to reach $10^9\ M_\odot$ in the time between the Big Bang and $z \sim 6$. These models would need to reproduce exactly the population of accreting black holes observed in the early universe.
- MASS ESTIMATES If scaling relations relating black hole masses to emission line equivalent widths are, in fact, overestimates of black hole mass, this could result in a partial resolution to inconsistencies between black hole masses and evolutionary models.

The physical processes and mechanisms outlined above determine the observed number density of quasars in the sky at a given redshift with a given magnitude. Therefore, any models determining the evolution and growth of supermassive black holes must in turn recreate this distribution, the quasar luminosity function (QLF). An accurate model for this distribution based on observations will allow us to pinpoint where our assumptions in the Black Hole evolutionary models have broken down. Previous efforts (detailed in section 4.3) to constrain the high redshift quasar luminosity function have used maximum likelihood estimates and have not treated surveys covering overlapping areas consistently and an improved statistical analysis will also aid in the understanding of this problem. As such, the remainder of this thesis is concerned with statistical analysis of the high redshift quasar population. In the following we introduce a self consistent approach to determining the distribution of quasars at high redshift, which must then be reproduced by any model of active supermassive black hole evolution at high redshift.

2 The Quasar Luminosity Function

The quasar luminosity function (QLF) describes the expected number, N, of quasars in a given comoving volume, V, with magnitudes in the range $[M_\lambda, M_\lambda + \Delta M_\lambda]$ and redshifts in the range $[z, z + \Delta z]$,

$$\mathrm{d}N = \Phi(M_\lambda, z)\mathrm{d}V\mathrm{d}M_\lambda,$$

where M_λ denotes the magnitude of a quasar at a given wavelength λ.

The QLF has been measured at a range of different redshifts so it is possible to see the evolution of the spatial density of quasars, $\rho(M_\lambda < M_{\mathrm{ref}}) = \int_{-\infty}^{M_{\mathrm{ref}}} \Phi(M_\lambda, z)\, \mathrm{d}M$ as a function of redshift. This number density is observed to increase with redshift up until $z \sim 2.5$ (Schmidt, 1968) after which a steep decrease with redshift is observed (Osmer, 1982; Warren et al., 1994; Schmidt et al., 1995b; Fan et al., 2001c; Richards et al., 2006a). Fig. 2.1 shows a plot of the cumulative spatial density $\rho(M_{1450} < -26.8)$ from Fan et al. (2001b) with the best fit values of Fan et al. (2001b); Boyle et al. (2000); Schmidt et al. (1995a) as shown in Fan et al. (2001c). It can be seen from this figure that the number density of bright quasars is low for $z \sim 0$ today and increases rapidly up to a peak at $z = 2.5$. For greater redshifts the number density of bright quasars decreases rapidly with redshift.

At high redshift, there is an upper limit on the mass of a supermassive black hole due to the limited time available for accretion. Observationally, this manifests itself as an upper limit on the absolute magnitude of a supermassive black hole. As a result we expect a break magnitude such that the number density of quasars with magnitudes brighter than this break decreases extremely rapidly. This break magnitude will move towards the faint end with increasing redshift as there is even less time for the black hole to accrete matter. Quasars with magnitudes fainter than this break have not had to experience such rapid accretion but will there will still be decrease in number density with magnitude due to the limited time available for accretion, but this rate will be slower.

Hopkins et al. (2007b) combines estimations of the QLF across different wavelength bands (X-ray, (Steffen et al., 2003; Ueda et al., 2003); radio, (Willott et al., 2001); optical, (Fan et al., 2001b; Croom et al., 2004; Willott et al., 2009)) to estimate the bolometric QLF and its evolution. The remainder of this report will be concerned with the QLF at 1450 Å measured by (Fan et al., 2001b,c,a; Jiang

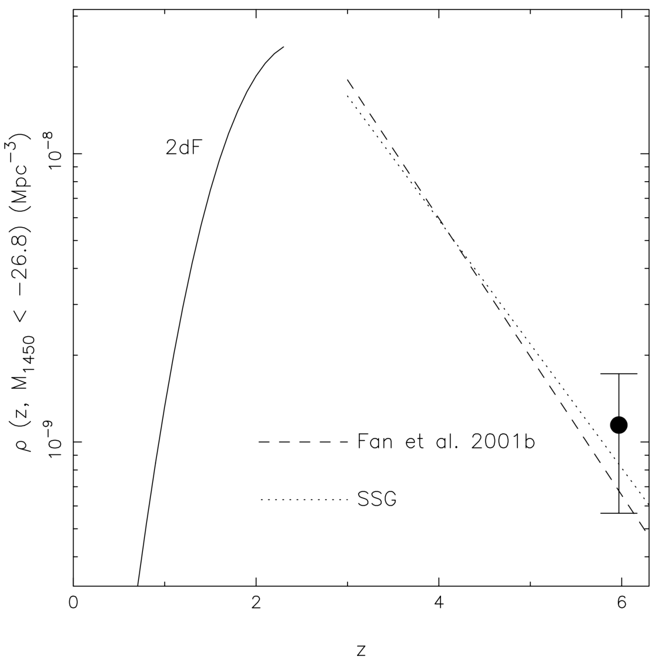


Figure 2.1: Evolution of comoving spatial density of quasars for $M_{1450} < -26.8$. Solid line is the spatial density measured by Boyle et al. (2000). Dashed line is that of Fan et al. (2001b). The dotted line is that of Schmidt et al. (1995b). This figure clearly shows the observed peak in the spatial density of quasars at $z \sim 2.5$. Taken from Fan et al. (2001b).

et al., 2008; Jiang et al., 2009; Willott et al., 2010a; Willott et al., 2010; McGreer et al., 2013). This wavelength is useful as there are no strong emission lines at $\lambda_{\rm rest} = 1450$ Å and photons with this wavelength are not absorbed by neutral hydrogen.

2.1 DOUBLE POWER LAW QLF

The double power law quasar luminosity function is motivated by observations of low-redshift quasars. It has been well established since it was introduced (Boyle et al., 1988, 2000; Pei, 1995) and is now used extensively to model the QLF (Fan et al., 2004; Hopkins et al., 2007b; Richards et al., 2006a).

The double power law luminosity function as used here takes the form

$$\Phi(M_{1450}, z) = \frac{\Phi^*(z)}{10^{0.4(\alpha(z)+1)(M_{1450}-M^*_{1450}(z))} + 10^{0.4(\beta(z)+1)(M_{1450}-M^*_{1450}(z))}}. \tag{2.1}$$

Where the power law is a function of redshift, z, and absolute magnitude at $\lambda = 1450$ Å. α and β parameterise the slope at faint and bright magnitudes respectively. The magnitude at which the turnover from α to β occurs is the break magnitude, denoted by M^*_{1450} and Φ^* parameterises the normalisation of the QLF and is the number density of quasars at the break magnitude. In principle all parameters of the QLF can evolve with redshift and additional parameters can then be included to characterise this evolution.

Two canonical evolutionary models are pure density evolution (PDE) and pure luminosity evolution (PLE) (Boyle et al., 2000; Croom et al., 2004; Ueda et al., 2003; Richards et al., 2005). These only consider evolution with redshift of the normalisation parameter, Φ^*, and the break magnitude, M^*_{1450}, respectively. It is also common (Hopkins et al., 2007b) to combine these models to obtain luminosity dependent density evolution (LDDE), as well as to add evolution for the faint end slope, α and bright end slope β (Hopkins et al., 2007b).

At high redshift, the number density of luminous quasars is limited by the lack of time available for accretion. Therefore, as redshift increases we would expect the number density to decrease. In a similar vein, we would also expect the break magnitude to shift towards the faint end of the QLF. As such, the parameterisation that I will be implementing for the $z \sim 6$ quasar luminosity is

$$\Phi^*(z) = 10^{k(z-z_{\text{ref}})}\Phi^* \tag{2.2}$$

$$M^*_{1450}(z) = M_{1450} + j(z - z_{\text{ref}}) \tag{2.3}$$

with j and k left as parameters to be fit to the data and $z_{\text{ref}} = 6$. The parameterisation of $\Phi^*(z)$ is adopted from Fan et al. (2001b). The parameterisation of M^*_{1450} is a linear expansion of $M^*_{1450}(z)$ to first order and is similar to that used in Croom et al. (2004); Richards et al. (2005) to model the low redshift QLF.

We chose not to include evolution of α and β with redshift, as from results at $z \sim 5$ and $z \sim 6$ M^*_{1450} and Φ^* have the clearest evolution. Both Willott et al. (2010a), using data at $z \sim 6$ and McGreer et al. (2013), using quasars at $z \sim 5$ fix $\alpha = -1.5$ and find that β is consistent within the range of their uncertainty. However this is not the case for M^*_{1450} or Φ^*, hence we chose to evolve these parameters as they are the ones for which there is evidence of strong evolution.

We therefore have up to 6 parameters, $\beta, \alpha, \Phi^*, k, M^*_{1450}, j,$ which we need to estimate, and a full QLF with the form

$$\Phi(M_{1450}, z) = \frac{10^{k(z-z_{\text{ref}})}\Phi^*}{10^{0.4(\alpha+1)(M_{1450}-M^*_{1450}-j(z-z_{\text{ref}}))} + 10^{0.4(\beta+1)(M_{1450}-M^*_{1450}-j(z-z_{\text{ref}}))}}. \tag{2.4}$$

2.2 THE BLACK HOLE MASS FUNCTION

The population of black holes in the universe is described by the black hole mass function (BHMF). This gives the number of black holes per comoving volume element with masses in the range $[M, M + \mathrm{d}M]$. The characteristic shape of the BHMF has a sharp cut off at high mass (Shankar et al., 2009; Vestergaard et al., 2008) due to AGN feed back and in the case of high redshift, the lack of time for black hole growth (Volonteri & Rees, 2005). (Willott et al., 2010) adopts a Schechter function (Schechter, 1976) form of the BHMF:

$$\Phi(M_{\mathrm{BH}}) = \Phi^{*}\left(\frac{M_{\mathrm{BH}}}{M^{*}_{\mathrm{BH}}}\right)^{\alpha} \exp\left(-\frac{M_{\mathrm{BH}}}{M^{*}_{\mathrm{BH}}}\right) \tag{2.5}$$

This has the same power law form as the QLF at the faint end, but a sharper cut off at the bright end as is theoretically favoured by e.g. Shankar et al. (2009); Willott et al. (2010).

THE EDDINGTON RATIO DISTRIBUTION

The Eddington ratio is defined as the ratio of a quasar's bolometric luminosity to the Eddington luminosity determined by the mass of the SMBH at the centre of the quasar.

$$\lambda = \frac{L_{\mathrm{bol}}}{L_{\mathrm{Edd}}} \tag{2.6}$$

By measuring this quantity for observed quasars it is possible to determine the distribution in λ. This distribution tells us how active the population of quasars we observe are and by measuring this over a range of redshifts can tell us how the behaviour of the population evolves. Willott et al. (2010) have made the first determination of the Eddington ratio distribution $P(\lambda)$ at $z = 6$. They found that the distribution can be approximated by a lognormal distribution with a peak at $\lambda = 0.6$ and a narrow dispersion of 0.3 dex (Fig. 2.2).

This observation implies that there are very few black holes at $z = 6$ with $\lambda < 0.1$, suggesting that there are extremely few inactive black holes at $z = 6$ and that most of them are accreting close to their Eddington rates. Large hydrodynamical simulations by Sijacki et al. (2009) have also reached similar conclusions.

The BHMF is directly related to the QLF. If one is only interested in the mass function of SMBH that power quasars, then it is possible to transform the BHMF to the QLF by convolution of the BHMF with the Eddington ratio distribution

$$\frac{\mathrm{d}n}{\mathrm{d}L_{\mathrm{bol}}} = \int\int \frac{\mathrm{d}n}{\mathrm{d}M}\, P(\lambda)\, \delta_D[M - M(\lambda, L_{\mathrm{bol}})]\, \mathrm{d}\lambda \mathrm{d}M, \tag{2.7}$$

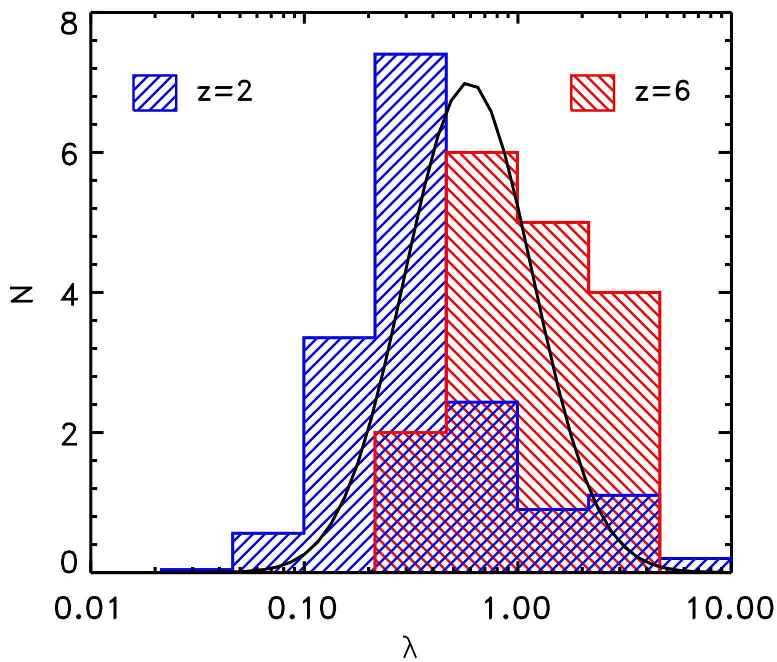


Figure 2.2: The Eddington ratio distribution. Red histogram shows the Eddington ratio for $z = 6$ while blue histogram shows that of $z = 2$. The intrinsic distribution (black curve) is a lognormal distribution with and intrinsic peak at $\lambda = 0.6$ and dispersion 0.3 dex. Figure taken from Willott et al. (2010)

where

$$M(\lambda, L_{\rm bol}) = \frac{L_{\rm bol}}{\beta\lambda},\ \beta = \frac{4\pi G m_p c}{\sigma_T}.$$

This relationship has the following implication. If we can provide a self consistent formalism with which to estimate the parameters of the QLF, then we can use mass estimates to understand how the population of active black holes behaves at high redshift. Any theoretical model or observation of black hole mass must reproduce the QLF, thus it will provide us a bench mark with which to compare future models to data.

3 | Quasar Surveys

Quasars emit light at all wavelengths. However, neutral hydrogen (HI), which is abundant in the high redshift intergalactic medium, absorbs light at rest frame wavelengths $\lambda < 1216$Å corresponding to the Lyman-α absorption, observed when an electron in a neutral hydrogen atom moves between orbitals $n = 1$ and $n = 2$. Due to such a high abundance of HI, absorption of continuum emission from quasars occurs at all redshifts along the line of sight to an observer. This means a significant fraction of light blue-wards of $\lambda_{\rm obs} = (1 + z)1216$Å is absorbed as can be clearly seen in Fig. 3.1.

Because of the existence of the Lyman-α absorption at high redshift, the flux of the quasar at $\lambda_{\rm obs} \leq (1 + z)1216$Å is greatly supressed. The quasar will therefore be observed in all bands except for those that lie blue-ward of the Lyman-α break. At $z \sim 6$ this break coincides with the top end of the i-band (centred on 7840Å Fukugita et al. (1996)) and quasars are therefore detected as i-band dropouts and this break moves into the z band. Fig. 3.2 shows the photometric bands implemented in the SDSS survey (section 3.1) demonstrating the width of the i-band, centred at 7840Å.

3.1 The Sloan Digital Sky Survey

The Sloan Digital Sky Survey (SDSS) (York et al., 2000) was a multi-filter imaging and spectroscopic survey. The survey used a dedicated 2.5 m optical telescope located at the Apache Point Observatory (APO) in New Mexico. The SDSS imaging camera surveyed an area of π sr above Galactic latitude $30°$ in five optical bands (u' g' r' i' z', Fukugita et al. (1996), Fig. 3.2). The imaging survey consisted of $\sim 6000\ \rm deg^2$ of the northern galactic cap, a region that lies above Galactic latitude $30°$ and centred on $\alpha = 12^{\rm h}20^{\rm m},\ \delta = +32°.5$. In addition, three stripes in the southern Galactic cap were observed, two north and south of the equator and one along the equator (York et al., 2000). The equatorial stripe was be imaged repeatedly to search for variable objects and to reach a magnitude limit 2 mag deeper than the northern imaging survey (York et al., 2000).

The SDSS began data collection in 2000 and since has obtained photometric data for $\sim 5 \times 10^8$ objects and spectra for 3×10^6 objects. Of those, around $300,000$ quasars have been discovered in the SDSS (Pâris et al., 2014) with 30 at $z \sim 6$ (Jiang et al., 2008; Jiang et al., 2009; Fan et al., 2006;

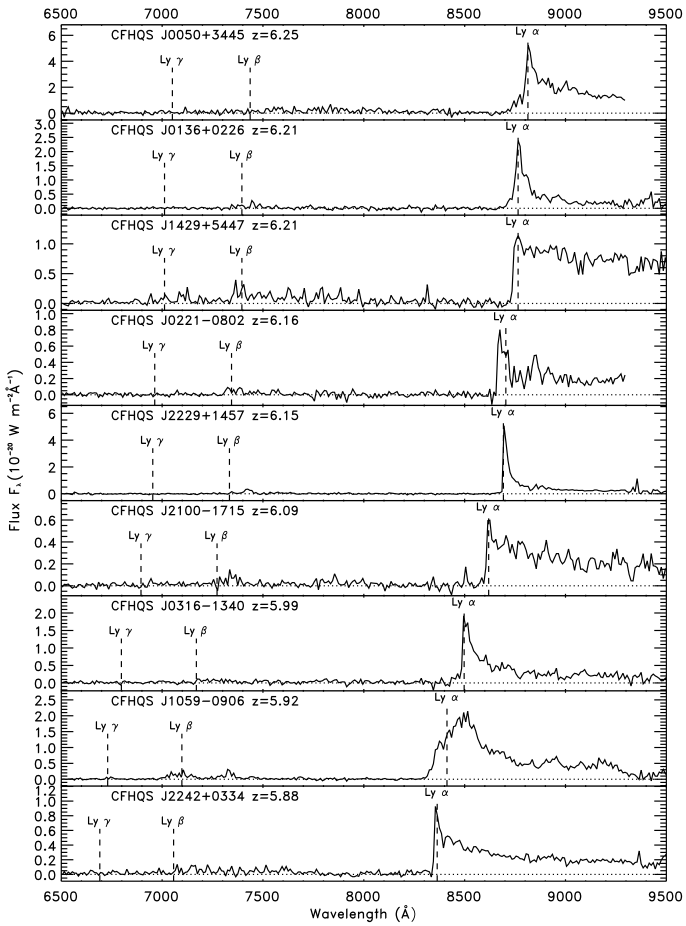


Figure 3.1: Spectra of 9 high redshift quasars found in the CFHQS survey (Willott et al., 2009). Wavelengths are λ_{obs}. Figure taken from Willott et al. (2010a).

Jiang et al., 2015). The SDSS pioneered the search for redshift 6 quasars and was followed by the Canada France High-z Quasar survey (CHFHQS, Willott et al. (2009)) and the UKIRT Infrared Deep Sky Survey (UKIDSS, Lawrence et al. (2007)).

3.2 $z \sim 5$ QUASAR SURVEYS

$z \sim 5$ SDSS MAIN SAMPLE

The bright end of the $z \sim 5$ luminosity function is best sampled by the seventh data release (DR7) of the SDSS main survey (DR7QSO) (Abazajian et al., 2009; McGreer et al., 2013). The selection algorithm (Richards et al., 2002) for this survey was not complete before the survey was already in progress (McGreer et al., 2013). Using the method prescribed in Richards et al. (2006b) DR7 quasars were identified from regions with uniform target selection. Colour cuts were also applied (McGreer et al., 2013) resulting in 146 quasars with $4.7 < z < 5.1$ in an effective area of $6222\ \text{deg}^2$. These

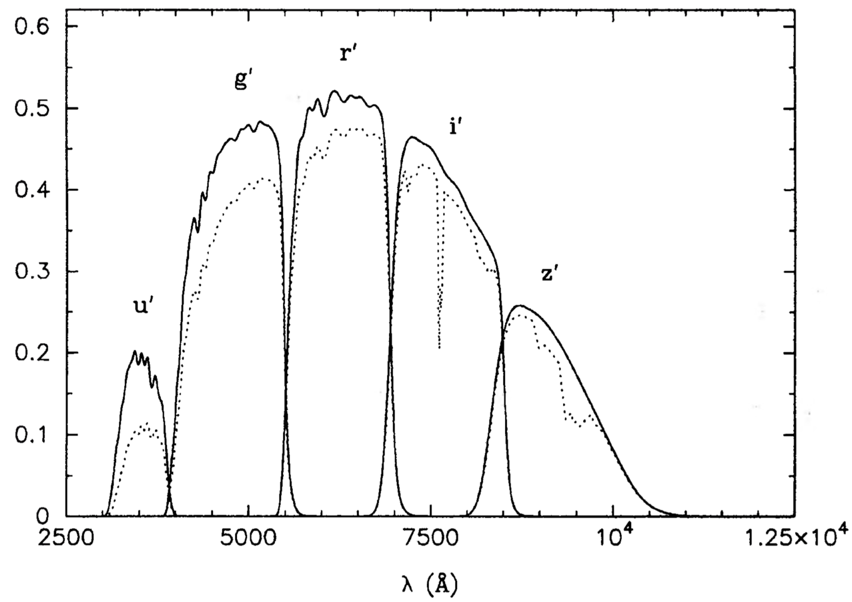


Figure 3.2: The 5 photometric bands implemented in the SDSS survey. Figure taken from Fukugita et al. (1996).

quasars are shown in the left hand panel of Fig. 3.3.

$z \sim 5$ Stripe 82 sample

11 new quasars were found in SDSS DR7 (Abazajian et al., 2009), which selects quasars with $4.7 < z < 5.1$ to a magnitude limit of $i \lesssim 20.2$. A total of 14 new quasars with $z > 4.6$ were found in BOSS DR9 (Ross et al., 2012). This survey was primarily designed to find quasars with $2.2 < z < 3.5$ (McGreer et al., 2013) however the survey covered a large section of Stripe 82 and the search was extended to find $z > 4.6$ quasars in Stripe 82. Four quasar candidates were observed on the Magellan Spectrograph, of which one was a star and one had previously been discovered in BOSS, leaving a total of 2 new high redshift quasars. The bulk of the spectroscopic observations were made at the MMT 6.5 m telescope. Over the course of several observing runs a total of 57 spectroscopically confirmed quasars were observed at MMT with $4.3 < z < 5.46$.

Of the 84 observed quasars, 73 are spectroscopically confirmed with 52 within the redshift range $4.7 < z < 5.1$. The final area of Stripe 82 used for quasar selection was $235\ \mathrm{deg}^2$ (McGreer et al., 2013).

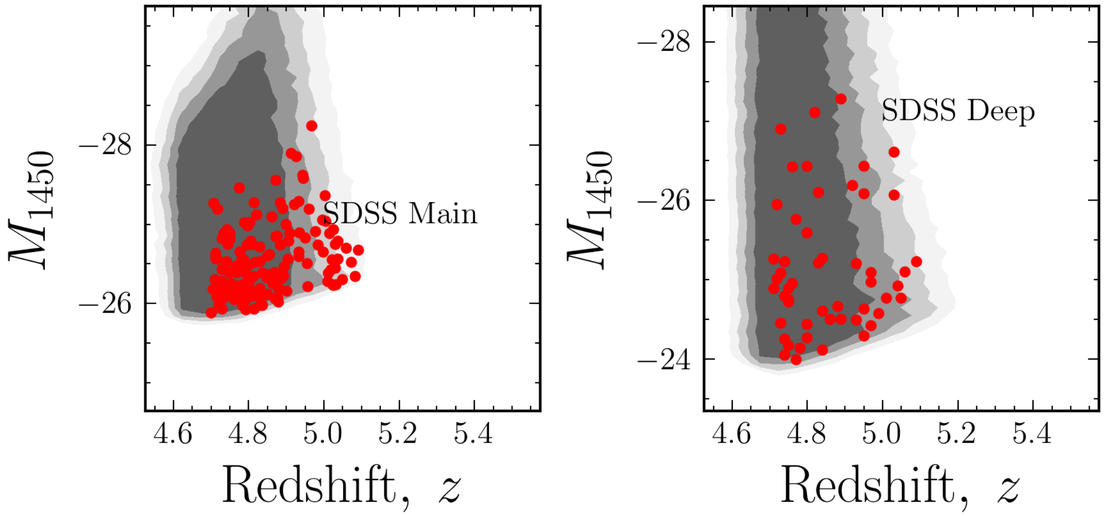


Figure 3.3: The selection functions and quasar samples from the SDSS main sample (left) and Stripe 82 redshift 5 quasar sample (right) (McGreer et al., 2013). Contours show $20\%,\ 40\%,\ 60\%$ and 80% completeness. Red points are the quasars selected from the surveys after a redshift cut of $4.7 < z < 5.1$ has been applied.

3.3 $z \sim 6$ QUASAR SURVEYS

SDSS MAIN SAMPLE

The most up to date description of the SDSS main high-z sample (Fan et al., 2001b,c,a, 2006) is given in Fan et al. (2006) and contains 14 quasars with $5.74 < z < 6.42$. The survey has a magnitude limit of $z' < 20.2$ and as a result the bright end of the quasar luminosity function is well sampled by the quasars in this survey. Quasars are selected from a effective area of ~ 6600 deg^2 and the selection function shown in the first panel of Fig. 3.4. The selection function for this survey was first described in Fan (2003), and adjusted to the cosmology used in this thesis by Willott et al. (2010a). Noticeably the quasars found in this survey are clustered close to the cut off in the selection function.

SDSS STRIPE 82

A small patch of the SDSS survey, Stripe 82, was imaged by Jiang et al. (2008) to search for fainter quasars.

6 new quasars were found in this new ~ 260 deg^2 area spanning $30° < \alpha < 310°$, $-1.5° < \delta < 1.5°$ with a magnitude limit of $20 < z' < 21$ (see panel 2 of Fig. 3.4), greater than that of the SDSS main sample due to repeated imaging of the same area of sky. This search was extended further in Jiang et al. (2009) on a slightly smaller patch ranging $60° < \alpha < 310°$ and an effective area of 195 deg^2

Quasar	Redshift, z	Magnitude, M^1_{1450}
$J1044-0125$	5.74	-27.55
$J0836+0054$	5.82	-27.93
$J1030+0524$	6.28	-27.24
$J1306+0356$	5.99	-27.20
$J1048+4637$	6.23	-27.60
$J1148+5251$	6.42	-27.87
$J1411+1217$	5.93	-26.80
$J1602+4228$	6.07	-26.87
$J1623+3112$	6.22	-26.72
$J0818+1722$	6.00	-27.42
$J0840+5624$	5.85	-26.69
$J0927+2001$	5.79	-26.83
$J1137+3549$	6.01	-27.13
$J1250+3130$	6.13	-27.16

Table 3.1: The complete sample of $z\sim 6$ quasars found in the SDSS survey (Fan et al., 2001b,c,a, 2006).
[1] All magnitudes are AB magnitudes (Oke & Gunn, 1983).

Quasar	Redshift, z	Magnitude, M_{1450}
$J0005-0006$	5.85	-25.87
$J0203+0012$	5.85	-25.77
$J0303-0019$	6.07	-25.48
$J0353+0104$	6.05	-26.54
$J2054+0005$	6.06	-26.16
$J2315-0023$	6.12	-25.43
$J2053+0047$	5.92	-25.52
$J2147+0107$	5.81	-25.05
$J2307+0031$	5.87	-24.98
$J2356+0023$	6.00	-24.97

Table 3.2: The complete sample of $z\sim 6$ quasars discovered in the SDSS Stripe 82 (Jiang et al., 2008; Jiang et al., 2009). The first 6 were found in the initial shallower survey (Jiang et al., 2008) while the remaining 4 were discovered in the deeper re-imaging. (Jiang et al., 2009).

resulting in the detection of 4 additional quasars with a magnitude limit of $21 < z' < 21.8$ (see panel 3 of Fig. 3.4).

CANADA-FRANCE HIGH-Z QUASAR SURVEY

Also included in the redshift 6 quasar sample are those found in the Canada France High-z Quasar Survey (Willott et al., 2007, 2009). The main sample consists of 12 quasars found in the Red-Sequence Cluster Survey 2 (RCS-2) (Yee et al., 2007) and 4 quasars found in the Canada-France Hawaii Telescope Legacy Survey (CFHTLS) Very Wide, the shallowest component of the CFHTLS.

Quasar	Redshift, z	M_{1450}
$\text{J}1509-1749$	6.12	-26.804
$\text{J}1641+3755$	6.05	-25.214
$\text{J}2329-0301$	6.42	-25.024
$\text{J}0033-0125$	6.13	-24.909
$\text{J}2318-0246$	6.05	-24.854
$\text{J}0055+0146$	5.98	-24.552
$\text{J}0102-0218$	5.95	-24.338
$\text{J}2329-0403$	5.90	-24.385
$\text{J}0050+3445$	6.25	-26.645
$\text{J}2229+1457$	6.15	-24.547
$\text{J}2100-1715$	6.09	-25.052
$\text{J}2242+0334$	5.88	-24.245
$\text{J}2315-0023$	6.12	-25.644
$\text{J}1059-0906$	5.92	-25.604
$\text{J}0136+0226$	6.21	-24.431
$\text{J}0316-1340$	5.99	-24.657
$\text{J}0216-0455$	6.01	-22.236

Table 3.3: Quasars found in the CFHQS $z \sim 6$ survey (Willott et al., 2009, 2010a). The first 16 were discovered in the main sample, the last was discovered in the deeper CFHQS survey. Note the Quasar J2315 quasar that was discovered both in the CFHQS survey and SDSS.

These surveys make up a total sky area of $494\ \text{deg}^2$. The deep sample of the CFHQS consists of 1 quasar found in the CFHTLS Deep survey/Subaru/XMM-Newton Deep Survey (SXDS). This forms a survey with a total area of $4.47\ \text{deg}^2$. The completeness tables of these surveys is shown in the last two panels of Fig. 3.4 and details of how these tables were obtained are described in Reylé et al. (2010).

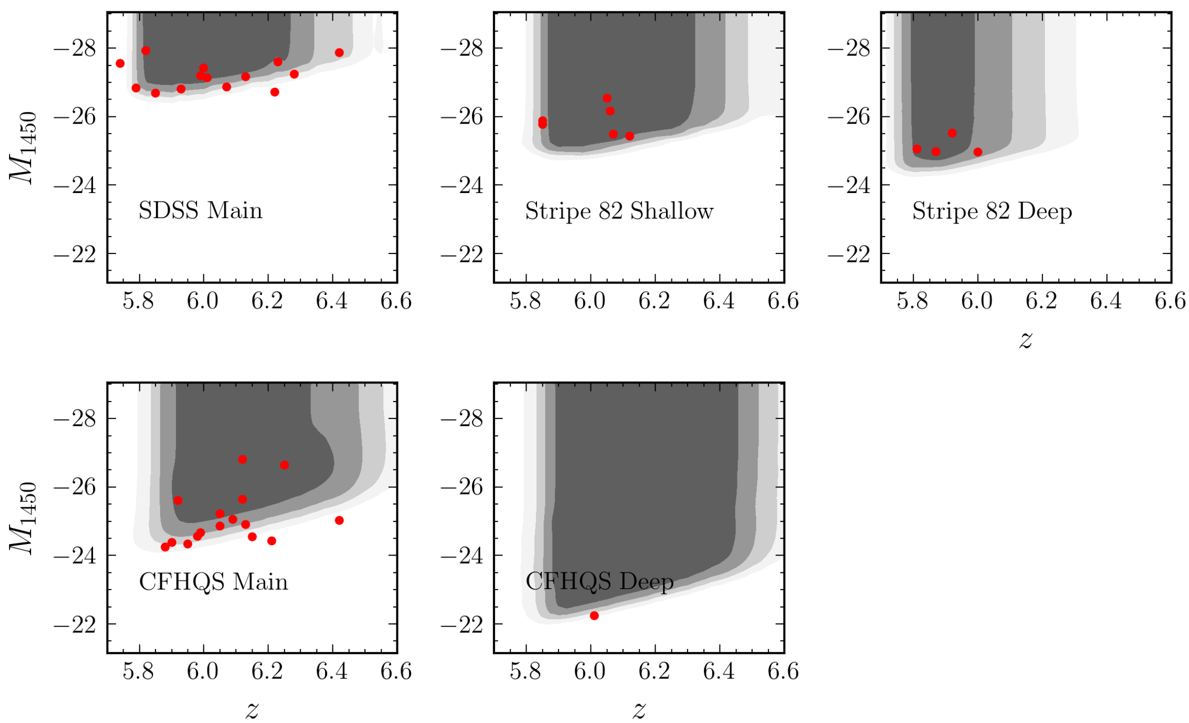


Figure 3.4: Selection functions and quasar samples of $z \sim 6$ surveys used by Willott et al. (2010a) to calculate the quasar luminosity function at $z \sim 6$. Contours are placed at $20\%,\ 40\%,\ 60\%$ and 80% completeness. Red points show the complete quasar samples from each survey.

4 QLF PARAMETER ESTIMATION METHODOLOGY

Parameter estimation is a statistical problem and as a result many statistical and algorithmic techniques must be used in order to treat the problem correctly. I start by constructing the likelihood of the quasar luminosity function, section 4.1, before moving on to discuss two methods of performing parameter inference, maximum likelihood estimation, with χ^2 confidence intervals to quantify uncertainties, section 4.3, and Bayesian parameter inference, section 4.4. I also outline a method for simulating quasar surveys for the purposes of comparing results from simulations to results with actual data, section 4.2. Finally, I describe the algorithms for producing results and caluclating the likelihood and posterior distributions for the quasar luminosity function. For the case of low dimensionality, $N_p \leq 3$, this can be done by evaluating the likelihood or posterior over a N_p dimensional array, section 4.5. However, for $N_p > 3$ we must utilize a Markov Chain Monte Carlo algorithm, section 4.4.2.

4.1 THE QLF LIKELIHOOD

Whether for use in a Bayesian or maximum likelihood approach to parameter inference, it is necessary to construct the likelihood of observing data given a model for the QLF. We want to construct the probability, $\Pr(\boldsymbol{d}|\boldsymbol{\theta}, M)$ of observing a set of N data points $\boldsymbol{d} = \{d_1, d_2, ..., d_N\}$ given a model, M with P parameters $\boldsymbol{\theta} = \{\theta_1, \theta_2, ..., \theta_P\}$. For the case of the QLF our data are the magnitudes and redshifts, $\{M_{1450,i}, z_i\}$ of each quasar and the total number N_q of quasars found in the survey, and the parameters are those of the QLF outlined in chapter 2 . This likelihood can be constructed via two different methods, outlined below.

The expected number of quasars to be detected with magnitudes and redshifts in the range $[M_{1450}, M_{1450} + \Delta M_{1450}],\ [z, z + \Delta z]$ in a survey is given by:

$$\begin{aligned} \mathrm{d}\bar{N} &= \frac{A}{4\pi}\,\Phi(M_{1450},z)\,p(M_{1450},z)\,\frac{\mathrm{d}V}{\mathrm{d}z}\mathrm{d}M_{1450}\,\mathrm{d}z && (4.1)\\ &= f(M_{1450},z)\,\mathrm{d}M_{1450}\,\mathrm{d}z, && (4.2)\end{aligned}$$

where $\Phi(M_{1450}, z)$ is defined in eq. (2.1), $p(M_{1450}, z)$ is the selection function of the specific survey and A is the area of the survey. The expected number $\bar{N}$ is then given by the integral of equation 4.1 over all M_{1450} and z. Equation 4.2 defines $f(M_{1450}, z)$.

Constructing the likelihood

Marshall et al. (1983) and Fan et al. (2001a) describe a method for calculating the likelihood of a complete sample of quasars using Poisson probabilities. The survey coverage is split into cells of size $\Delta M_{1450}\Delta z$. As long as the cells are infinitesimally small, only $n_c = \{0, 1\}$ quasars are detected in each cell. The likelihood for a given survey is a product over these probabilities for each cell. The rate in the Poisson probability is given by the expected number of quasars in the c^{th} cell, $\bar{N}_c$:

$$\bar{N}_c = f(M_{1450,c}, z_c)\,\Delta M_{1450}\,\Delta z \tag{4.3}$$

The likelihood, $\mathcal{L}$, of the full sample is then given by

$$\begin{aligned} \mathcal{L} &= \prod_{c=1}^{N_c} \frac{\bar{N}_c^{n_c}\exp -\bar{N}_c}{n_c!} && (4.4)\\ &= \prod_{c=1}^{N_c}\exp -\bar{N}_c \underbrace{\prod_{i=1}^{N_q}}_{\text{cells with } n_c=1} \bar{N}_i && (4.5)\\ &= \exp\underbrace{\sum_c -\bar{N}_c}_{=-\bar{N}}\prod_{i=1}^{N_q}\bar{N}_i. && (4.6)\end{aligned}$$

Combining 4.6 with 4.1 and 4.3, and noting that $\sum_c \bar{N}_c = \bar{N}$, we arrive at

$$\mathcal{L} = \prod_{i=1}^{N_q}\Phi(M_{1450,i}, z_i)\,p(M_{1450,i}, z_i)\,\frac{\mathrm{d}V}{\mathrm{d}z}\bigg|_{z=z_i}\Delta M_{1450,i}\,\Delta z_i \times \int\int \Phi(M_{1450}, z)\,p(M_{1450}, z)\,\frac{\mathrm{d}V}{\mathrm{d}z}\mathrm{d}M_{1450}\,\mathrm{d}z$$

The quantity of interest is $S = -2\ln\mathcal{L} + \text{const.}$ which is given by

$$S = -2\sum_{i=1}^{N_q} \ln\left[\Phi(M_{1450,i}, z_i)\, p(M_{1450,i}, z_i)\right] + 2\int\int \Phi(M_{1450}, z)\, p(M_{1450}, z)\, \frac{\mathrm{d}V}{\mathrm{d}z} \mathrm{d}M_{1450}\, \mathrm{d}z, \quad (4.7)$$

where the terms $\frac{\mathrm{d}V}{\mathrm{d}z}|_{z=z_i}\Delta M_{1450,i}\ \Delta z_i$ in the sum have been neglected as they only affect the normalisation of the likelihood, not the shape, so for the purposes of parameter estimation, they are not important. Note the distinction between S and $-2\ln\mathcal{L}$; the former neglects any constant terms with no parameter dependence, where as the latter must include all the constants.

This form of S matches that in Marshall et al. (1983) and is used by Fan et al. (2001a); Kelly et al. (2008); Willott et al. (2010a); McGreer et al. (2013)[1]

Alternatively, we can start from the likelihood $\mathcal{L}$ of a given survey for which N_q quasars are found:

$$\mathcal{L} = p(N_q, \{M_{1450,i}, z_i\}|M^*_{1450}, \Phi^*, \beta, \alpha, j, k),$$

where M_{1450}, z are the redshifts and magnitudes of each quasar in the survey. This likelihood can be split into the product of the Poisson probability for observing N_q quasars in a survey with the product over all quasars of $p(M_{1450}, z|\Phi(M_{1450}, z))$ to give

$$\mathcal{L} = \frac{\bar{N}^{N_q} \exp -\bar{N}}{N_q!} \times \prod_{i=1}^{N_q} p(M_{1450,i}, z_i|M^*_{1450}, \Phi^*, \beta, \alpha, j, k) \quad (4.8)$$

$$= \frac{\bar{N}^{N_q} \exp -\bar{N}}{N_q!} \times \prod_{i=1}^{N_q} \frac{f(M_{1450,i}, z_i)}{\int\int f(M_{1450}, z)\mathrm{d}M_{1450}\mathrm{d}z}, \quad (4.9)$$

where the integral extends over all M_{1450} and z.

Again, the quantity of interest is $S = -2\ln\mathcal{L} + \text{const.}$ Noting that the integral in the denominator of equation 4.9 is simply $\bar{N}$ we have

$$S = N_q \ln \bar{N} - \bar{N} - \ln N_q! + \sum_{i=1}^{N_q} \ln\left[f(M_{1450,z})\right] - N_q \ln \bar{N}.$$

Simplifying this equation and using equation 4.1 we are left with

[1]Although it is worth noting that in McGreer et al. (2013); Kelly et al. (2008) the term $p(M_{1450,i}, z_i)$ is missing from the logarithm. This term clearly affects the shape of the likelihood and is therefore important to include

$$S = -2\sum_{i=1}^{N_q} \ln\left[\Phi(M_{1450,i}, z_i)\, p(M_{1450,i}, z_i)\right] + 2\int\int \Phi(M_{1450}, z)\, p(M_{1450}, z)\, \frac{\mathrm{d}V}{\mathrm{d}z}\mathrm{d}M_{1450}\,\mathrm{d}z, \quad (4.10)$$

which is identical to 4.7 and again we have neglected terms for which there is no parametric dependence.

4.2 SIMULATING QUASAR SURVEYS

In order to simulate a quasar survey for the purposes of testing our analysis, it is necessary to sample from $f(M_{1450}, z)$. Plots of $f(M_{1450}, z)$ for the seven $z \sim 5$ & $z \sim 6$ surveys are shown as contours in Fig. 4.1. As this is not an analytic function and as the completeness is in the form of a 2-D array, we cannot use methods such as inverse transform sampling or rejection sampling to generate draws from this distribution.

To draw a sample, $\{M_{1450}, z\}$, from a distribution evaluated on an array, A, we perform the following steps.

1. Generate N_q, the number of samples to be drawn from the expected number of quasars in the survey: $N_q \sim \mathrm{Po}(N_{\mathrm{exp}})$
2. For $k = 0...N$
   - sample $Y_k \sim \mathcal{U}[0, \mathrm{sum}(\mathsf{A})]$
   - Perform a sorted search of the cumulative sum of A to find index $[i, j]$ where Y_k would be found in A
   - This index gives an interval in $[M_{1450,i}, M_{1450,i} + \Delta M_{1450}]$ and $[z_j, z_j + \Delta z]$.
   - Draw sample $M_{1450,k} \sim \mathcal{U}[M_{1450,i} \leq M_{1450,k}] \leq M_{1450,i} + \Delta M_{1450}$
   - Draw sample $z_k \sim \mathcal{U}[z_j \leq z_k \leq z_j + \Delta z]$

These simulations are shown as red points in Fig. 4.1. The real quasars are shown for reference as blue points, and contours show the distribution $f(M_{1450}, z)$.

4.3 MAXIMUM LIKELIHOOD ESTIMATION

Maximum likelihood estimation is a method of obtaining point estimates for the parameters of a model. As the likelihood is defined as the probability of observing a data set given the parameters

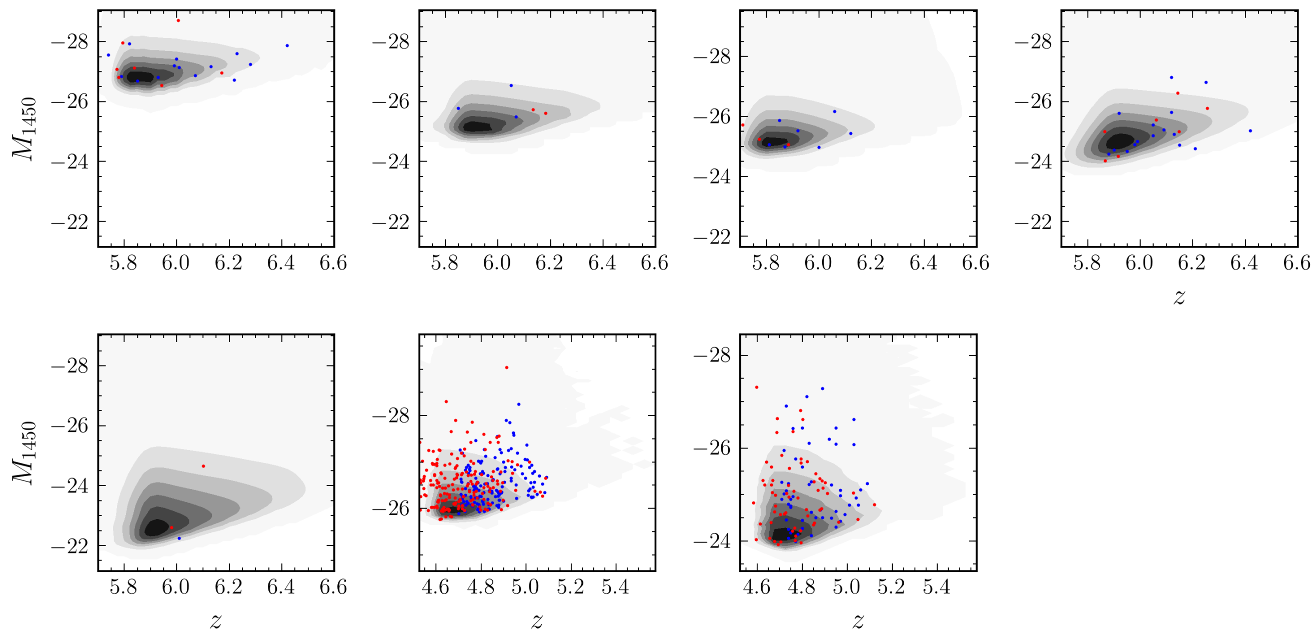


Figure 4.1: Contours in $f(M_{1450}, z)$ for each of the seven $z \sim 5$ and $z \sim 6$ quasar surveys. Red points are quasars sampled from this distribution using methodology prescribed above. Blue points are quasars found in the surveys. (Willott et al., 2010a; McGreer et al., 2013). The discrepancy between the real and simulated quasars in the last two panels is due to the fact that the redshift cut of $4.7 < z < 5.1$ applied to the data section 3.2 was not imposed in the simulation. Contours are evenly placed between the minimum and maximum of $f(M_{1450}, z)$

of the model, a maximum likelihood estimate intuitively involves finding the parameters of a model such that there is best agreement with the observed data.

However, simply performing a maximum likelihood estimate does not provide information about uncertainty within the calculation. In many cases it is possible to utilise χ^2 confidence regions in order to produce uncertainty estimates in the case where the data are gaussian distributed. If $X_i \sim N(\mu, \sigma^2)$ we have that

$$\mathcal{L} = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ \sum_{i=1}^{N} \left[-\frac{1}{2\sigma^2}(x_i - \mu)^2 \right] \right\}$$

For this special case it can be noted that the term in the exponent is related to the χ^2 distribution of the data,

$$\chi^2 = \sum_{i=1}^{N} \frac{(x_i - \mu)^2}{\sigma^2},$$

so we find that $\chi^2 = -2\ln\mathcal{L}$ which largely motivates the use of $-2\ln\mathcal{L}$ in the literature (Marshall et al., 1983; Fan et al., 2001b; Willott et al., 2010a; McGreer et al., 2013).

4.3.1 The ΔS Statistic

Lampton et al. (1976) provides a method for producing confidence intervals once the likelihood has been calculated, which I will summarise here. For a set of N data points whose deviations from the values expected by a given model are Gaussian, S will be distributed as χ^2 with N degrees of freedom. If we then minimise S with respect to p model parameters to obtain the $S_{\min}$ statistic, we can say that this is χ^2 distributed with $N - p$ degrees of freedom. So for the purposes of plotting confidence regions, it is sensible to talk about the statistic: $\Delta S = S - S_{\min}$ which will once again be χ^2 distributed with p degrees of freedom, i.e. $\Delta S \sim \chi^2_p$, so we can write that for any number, T,

$$\Pr(\Delta S > T) = \Pr(\chi^2_p > T)$$

We can then define contours S_L such that $S_L = S_{\min} + T$ where we have

$$\Pr(\Delta S > S_L - S_{\min}) = \Pr(\chi^2_p > S_L - S_{\min})$$

It can be noted that the left hand side is simply the probability, α, of a contour failing to enclose the true value, where α is defined via $\alpha = \Pr[\chi^2_p > \chi^2_p(\alpha)]$. This leaves the required contour for confidence $C = 1 - \alpha$ as,

$$S_L = S_{\min} + \chi^2_p(\alpha). \tag{4.11}$$

In many cases, however, one is only interested in a subset of the parameters used in the minimisation of S. Cases can include those where there are nuisance parameters which are not of interest in the analysis, or when for presentation purposes the confidence regions need to be in two dimensional space. In this case, one can project the initial p dimensional parameter space on to a subspace of q interesting parameters. The method of creating this subspace is as follows: at each point in the q dimensional subspace, minimise S w.r.t. the remaining $p - q$ parameters. Contours can then be placed again at positions given by eq. (4.11). Furthermore, numerical simulations (Avni, 1976) show that the level of confidence with which the p parameters are estimated can be accurately given by a χ^2_q function, reducing the size of the confidence regions to:

$$S_L = S_{\min} + \chi^2_q(\alpha). \tag{4.12}$$

This method of reduced numbers of parameters has been employed numerous times in evaluating confidence regions for the 4 parameters of the QLF (Willott et al., 2010a; McGreer et al., 2013). Fig. 5.1 shows confidence regions for the $z = 6$ QLF parameters from (Willott et al., 2010a). It does, however, rely on the assumption that input data are gaussian distributed random variables, which is not necessarily the case, and is therefore not a particularly robust method of producing reliable contours.

4.4 BAYESIAN PARAMETER INFERENCE

Bayes' theorem (Bayes & Price, 1763) states that the probability $P(A|B)$, is defined as

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)},$$

where $\mathrm{P}(A|B)$ is the probability of an event A occurring given the event B has already occurred.

A common application of Bayes' theorem is in data analysis, where some collected data is used to infer constraints on parameters of a given model. The quantity of interest is then the posterior, i.e. the probability of our model being correct, given that we've observed the data. For a set of N data points, $\boldsymbol{d} = \{d_1, d_2, ..., d_N\}$ that is to be fit by a model, M, with P parameters $\boldsymbol{\theta} = \{\theta_i = 1, \theta_2, ..., \theta_P\}$, Bayes' theorem can be written as

$$\Pr(\boldsymbol{\theta}|\boldsymbol{d}, M) = \frac{\Pr(\boldsymbol{\theta}|M)\Pr(\boldsymbol{d}|\boldsymbol{\theta}, M)}{\Pr(\boldsymbol{d}|M)}, \tag{4.13}$$

Once data has been included, $\Pr(\boldsymbol{\theta}|\boldsymbol{d}, M)$ is the posterior distribution in $\boldsymbol{\theta}$.

Using the law of total probability, $\Pr(B) = \sum_{i=1}^{N} \Pr(B|A_i)P(A_i)$ and generalising to the continuous case, we can write $\Pr(\boldsymbol{d}|M) = \int \Pr(\theta'|M)\Pr(\boldsymbol{d}|\theta', M)\mathrm{d}\theta'$ and so eq. (4.13) becomes

$$\Pr(\boldsymbol{\theta}|\boldsymbol{d}, M) = \frac{\Pr(\boldsymbol{\theta}|M)\Pr(\boldsymbol{d}|\boldsymbol{\theta}, M)}{\int \Pr(\theta'|M)\Pr(\boldsymbol{d}|\theta', M)\,\mathrm{d}\theta'}. \tag{4.14}$$

As introduced in 4.1, $\Pr(\boldsymbol{d}|\boldsymbol{\theta}, M)$ is the likelihood - the conditional probability of obtaining the data given the parameters of a model. $\Pr(\boldsymbol{\theta}|M)$ is the prior distribution on the model parameters and incorporates any prior knowledge. The normalised product of these two terms gives the posterior distribution $\Pr(\boldsymbol{\theta}|\boldsymbol{d}, M)$ on the possible parameter values.

In many cases, for reasons of displaying results on two-dimensional paper or if there are nuisance parameters in which we are not interested, it can be useful to reduce a multi-dimensional posterior distribution to just 1 or 2 dimensions. This can be done by marginalising (i.e. integrating) over the parameters we are not interested in. For example, suppose we want the posterior on simply θ_1 and θ_2 we can write

$$\Pr(\theta_1, \theta_2 | \boldsymbol{d}, M) = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} \Pr(\theta_1, \theta_2, \theta_3 ... \theta_P | \boldsymbol{d}, M) \mathrm{d}\theta_3 ... \mathrm{d}\theta_P$$

.

4.4.1 Highest posterior density regions

Even with 1 or 2 dimensional marginal posteriors, a result can still be difficult to interpret. The aim of any parameter estimation question is to place constraints on possible values of the parameters. Extracting a single value in parameter space that maximises the posterior distribution (the maximum a posteriori point estimate), gives the most concrete result to a question of parameter estimation, but in the process discards much of the information obtained by evaluating the posterior in the first place, and ignores uncertainty. Instead, much more information is obtained by defining a credible region, C, that encloses a chosen fraction, f, of the posterior distribution. So that

$$\int_C \Pr(\theta' | \boldsymbol{d}) \mathrm{d}\theta' = f. \tag{4.15}$$

It is important to note that C is not uniquely defined by eq. (4.15). Instead, we can impose the condition that all values within the credible region must have $P(\boldsymbol{\theta}|\boldsymbol{d}) \geq p$ such that f is redefined as

$$\int \Pr(\theta' | \boldsymbol{d}) \Theta(\Pr(\theta' | \boldsymbol{d}) - p) \mathrm{d}\theta' = f. \tag{4.16}$$

In general these regions can be defined to enclose any fraction, f, but common choices are the usual $1\sigma, 2\sigma, 3\sigma$ confidence intervals analogous to χ^2 confidence intervals, or alternatively $30\%,\ 60\%,\ 90\%$ credible regions.

4.4.2 Markov-Chain Monte Carlo

For a model with $N_P \gtrsim 4$ parameters, the grid method outlined in section 4.5 becomes computationally inefficient. Markov Chain Monte Carlo simulations draw samples θ from an approximate distribution to the actual posterior. The approximate distribution is then adjusted and a new draw is made. Each step of an MCMC chain is dependent only on the previous step[2] and so, in principle, the algorithm can uniformly and randomly sample the posterior.

In general the Markov chain simulation is initiated with a starting value θ^0. Subsequent draws θ^t are then sampled from the transition distribution $T_t(\theta^t|\theta^{t-1})$. This transition distribution is conditional only on the previous draw but must be chosen such that the Markov Chain converges

[2] This by definition forms the Markov Chain. A Markov chain is defined as a chain of values $\theta^0, \theta^1, \theta^2 ... \theta^N$ such that for any value of t, the distribution on θ^t is only conditional on the value of θ^{t-1}.

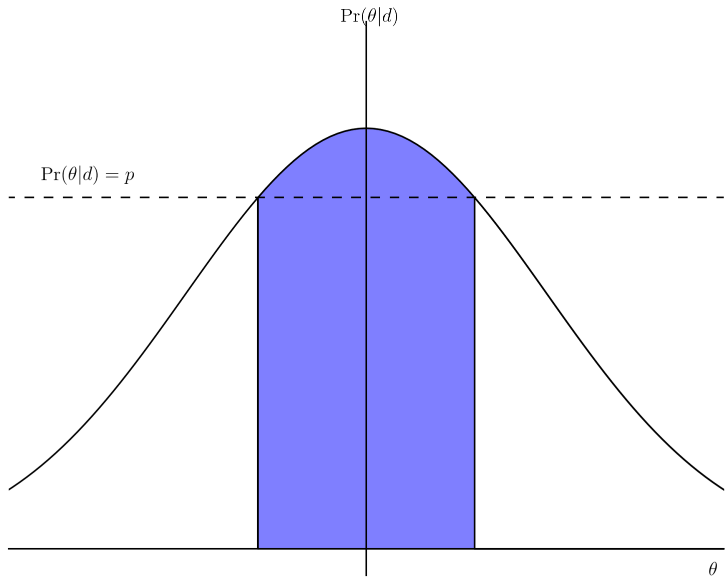


Figure 4.2: Figure illustrating a HPD region of a one dimensional distribution. Defining HPD regions via eq. (4.16) ensures contours of constant p are the smallest possible contours defined by eq. (4.15). In addition, values of θ on the boundary of C are all equally probable.

to the posterior distribution $p(\theta|d)$. As such, it is often useful to allow the transition distribution to be updated at each step of the chain, speeding up the rate of convergence.

Many algorithms implementing MCMC can be used. The most common are the Metropolis and Gibbs samplers. These algorithms differ most drastically in the choice of transition distribution and its dependence on step number. For reference, the algorithm for the Metropolis Hastings algorithm (Metropolis et al., 1953; Hastings, 1970) is presented here (Gelman et al., 2014).

1. Starting point θ^0 is drawn from the starting distribution $p_0(\theta)$ under the requirement that $p(\theta^0|d) > 0$.

2. For $t = 1, 2, \ldots$:

   i) Sample a proposal θ^* from a proposal distribution at time t, $J_t(\theta^*|\theta^{t-1})$.

   ii) Calculate the ratio of densities,

$$r = \frac{p(\theta^*|d)/J_t(\theta^*|\theta^{t-1})}{p(\theta^{t-1}|d)/J_t(\theta^{t-1}|\theta^*)}.$$

   iii) Set

$$\theta^t = \begin{cases} \theta^* & \text{with probability } \min(r,1) \\ \theta^{t-1} & \text{otherwise.} \end{cases}$$

The MCMC routine employed in the subsequent analysis is the emcee python package `emcee.py` written by Foreman-Mackey et al. (2013). This algorithm is a python implementation of the Affine Invariant MCMC Ensemble Sampler by Goodman & Weare (2010) and has been used extensively in cosmology and astrophysics . Trace plots of MCMC chains for parameters of the QLF are shown in Fig. 4.3

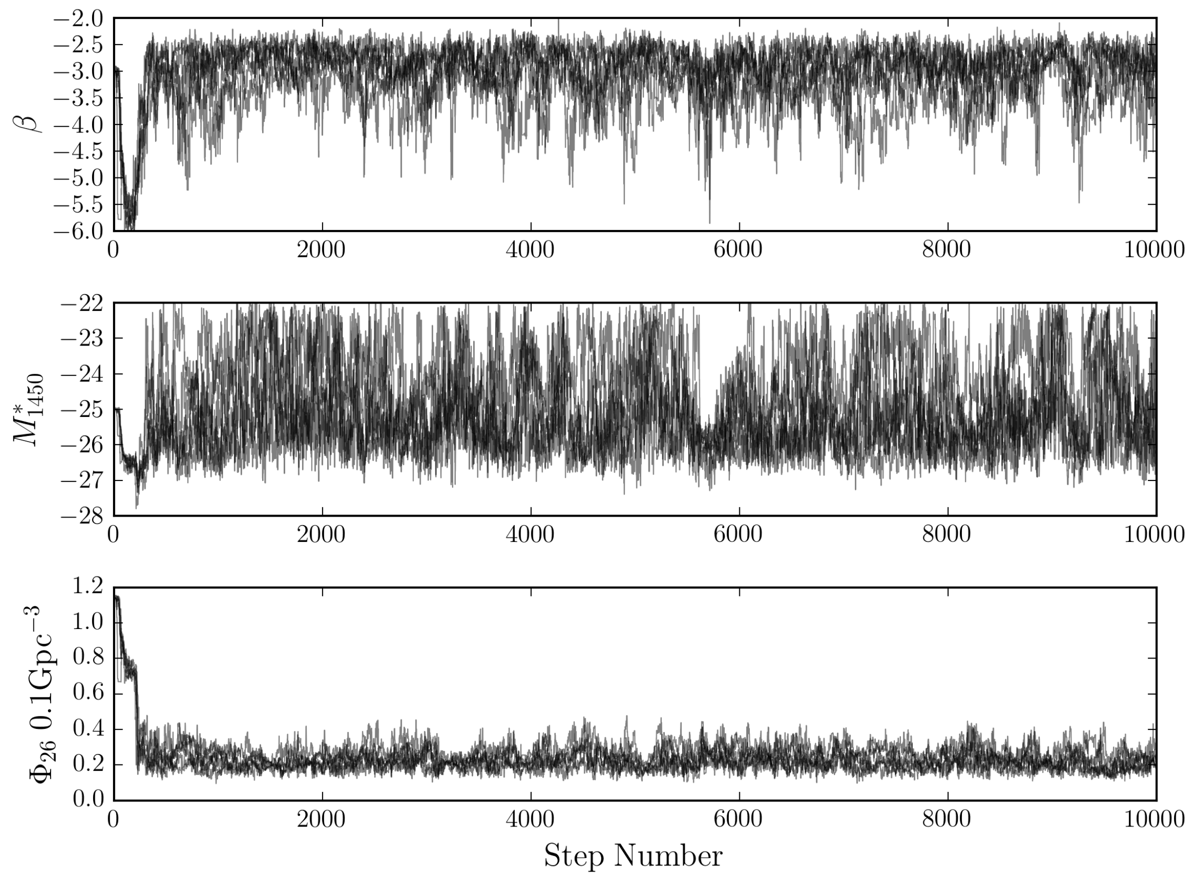


Figure 4.3: Trace plots of MCMC chain sampling from the posterior distribution of QLF parameters from the 3 parameter model as used by Willott et al. (2010a).

BURN-IN

During an MCMC sampling routine, the sampler can take time to find the peak of the posterior distribution. This is the burn-in period. As such, the first few samples are not useful in representing the posterior and can be discarded. In the following analysis initial draws are discarded until the first draw x_0 that satisfies $\mathcal{L}(x_0|\boldsymbol{\theta}) = 0.9 \times \max[\mathcal{L}(\boldsymbol{x}|\boldsymbol{\theta})]$ which ensures that the first retained sample is a plausible draw from the target distribution.

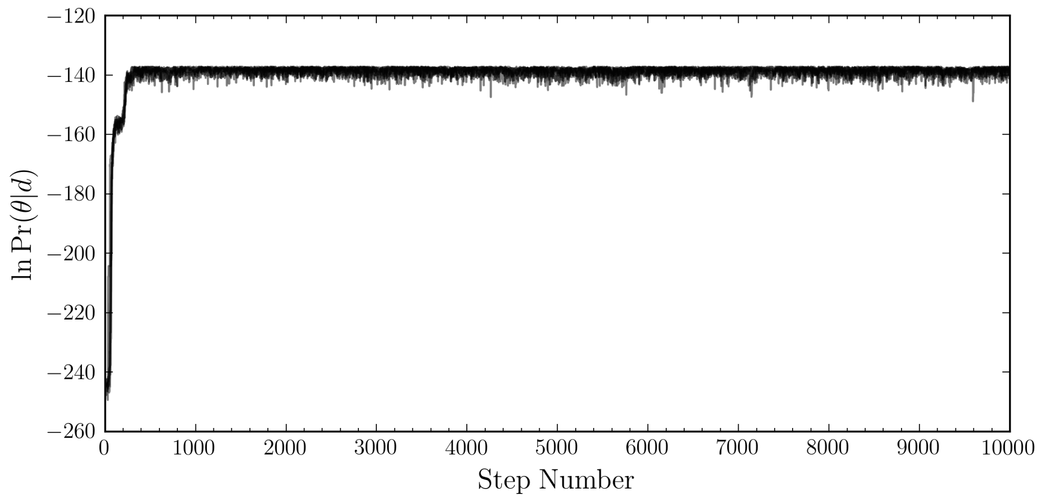


Figure 4.4: Plot of $\ln \mathcal{L}$ for each draw from the posterior distribution of the three parameter model of the QLF as used in Willott et al. (2010a)

AUTO-CORRELATION FUNCTION

An important check for behaviour of an MCMC chain is to look at the auto-correlation function. For every element in a chain, the autocorrelation function evaluates the correlation between the step and all the other points in the chain. It is therefore a function of time-lag, t, or distance in steps between two points in the chain, where

$$\rho_t = \frac{E[(\theta_i - \mu)(\theta_{i+t} - \mu)]}{E[(\theta_i - \mu)^2]E[(\theta_{i+t} - \mu)^2]},$$

where ρ_t is the autocorrelation function at lag t, θ_i are steps in the Markov chain with mean μ and E denotes the expected value operator.

The most computationally efficient way to evaluate the autocorrelation function for an MCMC chain is to use a Fast Fourier Transform (FFT) algorithm Foreman-Mackey et al. (2013). This however can underestimate the noise in ρ_t at large lags, due to the fact that FwaFTs use periodic boundary conditions. Once the FFT reaches the end of the chain, the algorithm simply loops round and evaluates the correlation with the first step again. This can be clearly seen by comparing figures Fig. 4.5 and Fig. 4.6. In the former there is little noise at large lag due to improper treatment of boundary conditions, whereas in the latter we see the level of noise increase at large time lag. In both the ACF at low lag is the same. This effect is not of great importance, however, because we are most interested in the ACF at small time lag, which gives an indication of how many points in the chain are correlated.

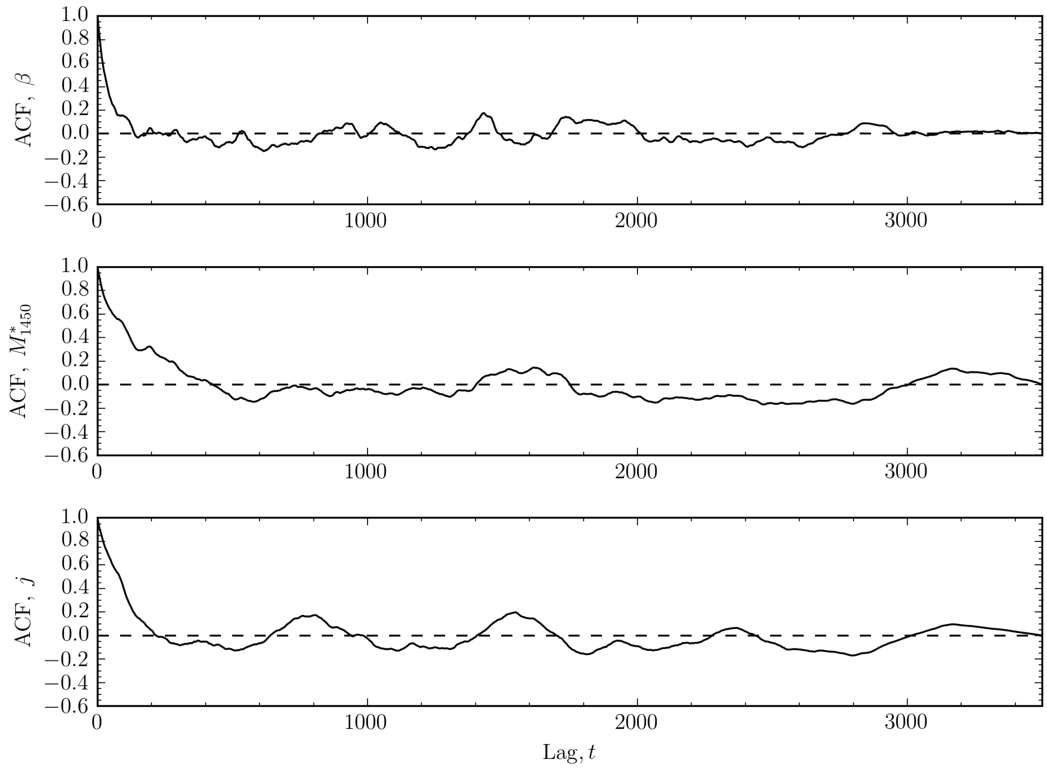


Figure 4.5: The autocorrelation function of MCMC chains for the 3 parameter QLF, constructed using the `emcee.py` algorithm utilising fast Fourier transforms. It is clear in these plots that there is a lack of noise at large lag, caused by improper treatment of boundary conditions.

EFFECTIVE SAMPLE SIZE

One can naively be led to believe that, when sampling a posterior with m chains of n samples, the total sample size is $N = n \times m$. However, correlations can arise between consecutive draws meaning that this number can be substantially lower than $n \times m$, and effectively one is only sampling $n_{\text{eff}} < N$ times. For example, for a MCMC run of 50 chains each sampling $10,000$ times (i.e. $N = 500,000$), if, on inspection of the ACF, draws are still correlated up until $n = 50$, then roughly only every 50^{th} draw is truly random. This means that each chain samples ~ 200 draws, resulting in an effective sample size of $n_{\text{eff}} \sim 10,000$.

STOPPING CRITERIA

There is no universally agreed stopping criteria for an MCMC sampling routine. As such, we chose to use 12 chains each sampling $10,000$ times from the posterior distribution. This was chosen because even with large correlations between samples, the effective sample size will still

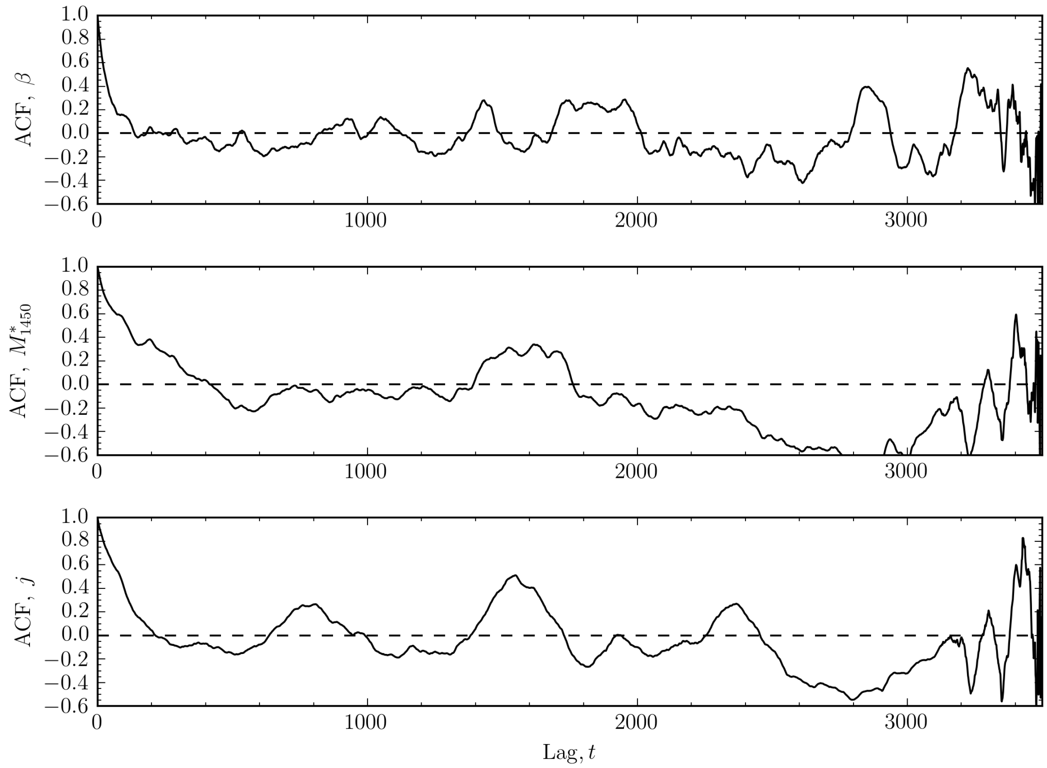


Figure 4.6: The autocorrelation function of MCMC chains for the 3 parameter QLF, constructed without periodic boundary conditions. It is clear in these plots that there is more noise at large lag than in Fig. 4.5.

be fairly large. Further we implement the Gelman-Rubin convergence diagnostic, R_c (Gelman & Rubin, 1992; Brooks & Gelman, 1998) to determine if our samples have converged to the true posterior distribution. This statistic compares the within-chain and between chain variances for each parameter of the model. A large value of R_c suggests large differences in these variances and indicates that the sampler has not converged on the true posterior distribution. We look for a value of $R_c < 1.2$ for all model parameters, as suggested by Brooks & Gelman (1998).

4.5 GRID METHOD FOR LOW NUMBERS OF PARAMETERS

If $N_p \leq 3$ the most computationally efficient way to estimate the maximum likelihood parameters is to evaluate an N_p dimensional array of $S = -2\log\mathcal{L}$ over a range of values of the parameters. This is particularly efficient for the case of the QLF as it is possible to factorise out the value of Φ^* such that S becomes

$$\underbrace{-2\sum_{i=1}^{N_q}\ln[\Phi_0(M_{1450,i},z_i)p(M_{1450,i},z_i)]} - 2N_q\ln[\Phi^*] + \Phi^*\times \underbrace{2\int\int \Phi_0(M_{1450},z)p(M_{1450},z)\frac{\mathrm{d}V}{\mathrm{d}z}\mathrm{d}M_{1450}\,\mathrm{d}z}\,.$$

Where $\Phi_0(M_{1450},z)$ is defined as $\Phi(M_{1450},z)/\Phi^*$. The two underlined terms need to be only evaluated for each of the N_p-1 parameters and then simply adjusted to calculate the likelihood for each value of Φ^* as well.

This means that S can be evaluated over N_p-1 dimensions and then linearly scaled to evaluate the likelihood at different values of Φ^*. To find the maximum likelihood parameters, we then find the coordinates of the minimum of the array of S.

To display confidence intervals, we can either use the ΔS prescription section 4.3.1 by minimising the array over the dimensions we wish to ignore. Alternatively, we can perform numerical integration over the same dimensions to display the marginal posterior distribution (section 4.4). In this case the limits of integration $\mathrm{p}_{min},\mathrm{p}_{max}$ provide a uniform prior distribution $\Pr(\theta)=\mathcal{U}[p_{\min},p_{\max}]$. We then place contours corresponding to highest posterior density regions (section 4.4.1).

5 | Parameter estimation of the QLF

5.1 Results from maximum likelihood estimation

Willott et al. (2010a) used quasars from the SDSS and CFHQS quasar surveys to estimate constraints on the $z \sim 6$ quasar luminosity function. The $1/V_a$ method prescribed in Avni & Bahcall (1980) was used to determine the shape of the QLF. Due to a lack of faint quasars in the sample, the faint end slope was fixed at $\alpha = -1.5$, a value quoted by Croom et al. (2004); Hunt et al. (2004); Richards et al. (2005); Croom et al. (2009) from lower redshift surveys where large numbers of faint quasars have been observed. Nevertheless it is worth noting that the faint end has been found to steepen towards high redshift (Glikman et al., 2010; Masters et al., 2012). Using a maximum likelihood fit the following parameters were obtained: $\Phi^* = 1.14 \times 10^{-8}\ \mathrm{Mpc}^{-3},\ M^*_{1450} = -25.13,\ \beta = -2.81$.

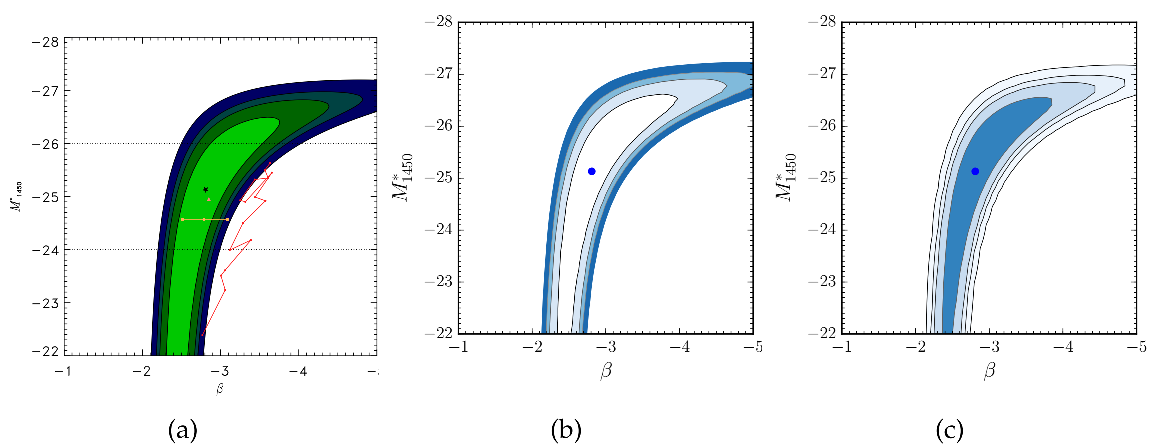
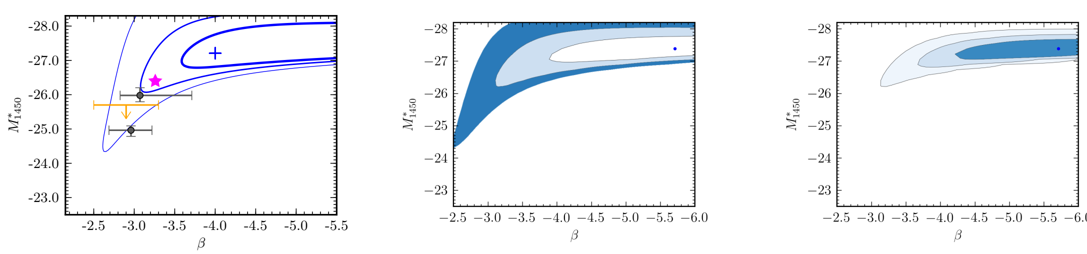


Figure 5.1: Confidence regions for the bright end slope, β and break magnitude M^*_{1450} contours show $(68.3\%, 90\%, 95.4\%, 99\%)$ confidence. In 5.1a and 5.1b, contours were obtained according to the prescription in Lampton et al. (1976). In 5.1c contours were obtained by marginalising over Φ^* and plotting highest posterior density regions corresponding to the same confidence. The star in 5.1a and blue dot in 5.1b and 5.1c corresponds to the maximum likelihood value of $M^*_{1450} = -25.13,\ \beta = -2.81$ as found by Willott et al. (2010a)

Using the method prescribed by Lampton et al. (1976) (see subsection 4.3.1) contour regions corresponding to a confidence of $(68.3\%, 90\%, 95.4\%, 99\%)$ were obtained Figure 5.1a. Also shown in Figure 5.1 is a recreation of the Willott results (Figure 5.1b) and the marginal posterior distribution in β and M^*_{1450} produced according to the description in section 4.5 (Figure 5.1c). Both methods of producing contour plots of credible regions give similar results, however marginalisation is a much more rigorous approach as we are not assuming Gaussian errors in our data.

McGreer et al. (2013) performed a similar analysis on the $z \sim 5$ data, with β fixed at $\beta = -4.0$, and obtained maximum likelihood values of $\log \Phi^*(z = 6) = -8.94,\ M^*_{1450} = -27.21,\ \beta = -4,\ \alpha = -2.03$ (Figure 5.2a). An analogous plot to Figure 5.2a was produced using the same analysis as Willott et al. (2010a) and is shown in Figure 5.2b below. There is a noticeable discrepancy between these two figures. As noted in Equation 4.1, the likelihood McGreer et al. (2013) use is slightly different from the one use here, as the term $p(M_{1450,i}, z_i)$ is missing from the logarithm in Equation 4.9. Fixing $\beta = -4.0$, we obtain maximum likelihood values of $M^*_{1450} = -27.01,\ \Phi^* = 0.1,\ \alpha = -1.79$. In Figure 5.2a all 4 model parameters are allowed to vary for calculation of the likelihood. However, the maximum likelihood values that McGreer et al. (2013) take correspond to setting $\beta = -4.0$. Therefore contours in Figure 5.2a are plotted with respect to the maximum likelihood at fixed β rather than the global maximum likelihood. In Figure 5.2b we have calculated confidence intervals with respect to the global maximum likelihood value of $M^*_{1450} = -27.39,\ \beta = -5.71,\ \Phi^*(z = 6) = 0.528 \times 10^{-8}\mathrm{Mpc}^{-3},\ \alpha = -1.94$. In both cases, the confidence intervals are $(68.3\%, 95.4\%, 99\%)$ Finally, we show the marginalised posterior distribution obtained using the method in section 4.5 (Figure 5.2c) and position contours at highest posterior density regions corresponding to the same confidence.



(a) (b) (c)

Figure 5.2: Confidence regions for the bright end slope, β and break magnitude M^*_{1450}. Contours show $(68.3\%, 95.4\%, 99\%)$ confidence intervals. In 5.2a and 5.2b, contours were obtained according to the prescription in Lampton et al. (1976). In 5.2c, contours were obtained by marginalising over Φ^* and plotting highest posterior density regions corresponding to the same confidence. In Figure 5.2a the blue cross corresponds to the McGreer et al. (2013) maximum likelihood estimates with β fixed at $\beta = -4$, while in Figure 5.2b and Figure 5.2c the blue points correspond to the maximum likelihood values calculated here.

5.2 STRIPE 82 COMPLETENESS

As noted in section 3.3 the two Stripe 82 surveys cover overlapping areas. The deeper of the two surveys a fraction of the area of the shallower with a fainter magnitude limit and discovered 4 new quasars. We can think of the 4 new quasars as being discovered in the area of sky where only the difference in completeness between the two stripe 82 surveys is high. Taking the difference between panels 2 & 3 of Fig. 3.4 we are left with the solid contours shown in the right hand panel of Fig. 5.3.

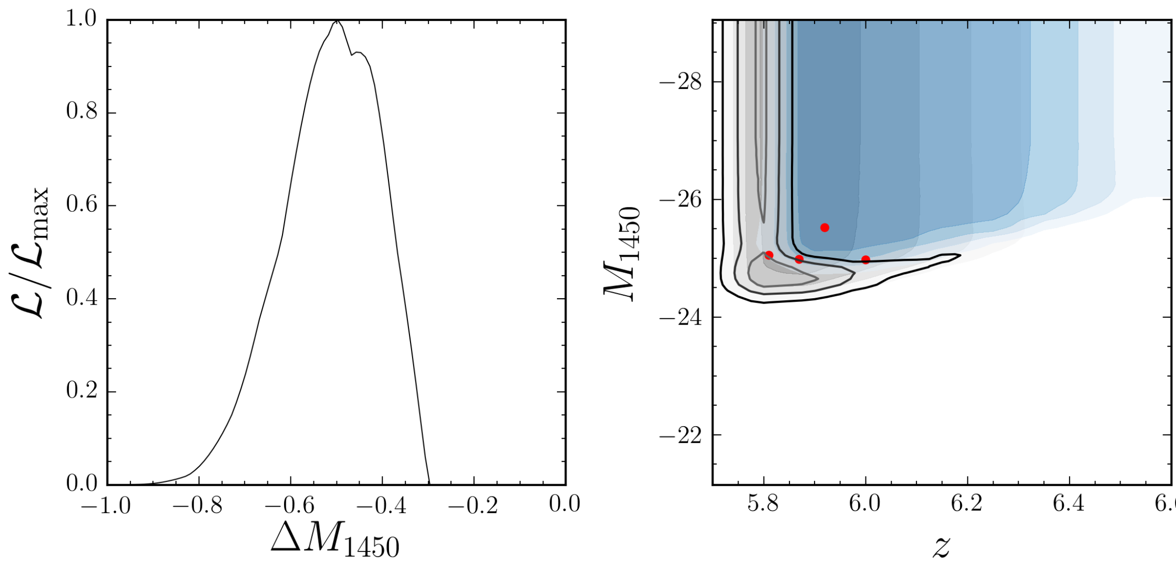


Figure 5.3: Right hand panel: Blue contours are the completeness of the inital Stripe 82 run (Jiang et al., 2008). Grey contours are those of the second survey (Jiang et al., 2009) and solid contours are those of the effective completeness, demonstrating that one quasar lies at a point with zero completeness. Left hand panel: The likelihood as a function of the parameter ΔM_{1450} which define the shift in magnitude limit of the survey. The maximum likelihood of $\Delta M_{1450} = -0.497$.

Upon inspection of this figure we can clearly notice a quasar in a position of zero completeness, i.e. with no probability of being found in this newly available region in M_{1450} and z space, suggesting a more conservative magnitude limit and that this quasar should have been found in the shallower survey. If we introduce a new parameter ΔM corresponding to a shift in magnitude limit of the Stripe 82 deep survey and perform a maximum likelihood analysis we find that the likelihood of the $\Delta M = 0.0$ (i.e. that the completeness is as quoted in Jiang et al. (2009)) is zero, see Fig. 5.3.

The maximum likelihood value for $\Delta M_{1450} = -0.497$. We then need to apply this new magnitude limit to the completeness table to and obtain the right hand panel of Fig. 5.4. Here we can clearly see that all quasars found in the survey are at points of non-zero completeness.

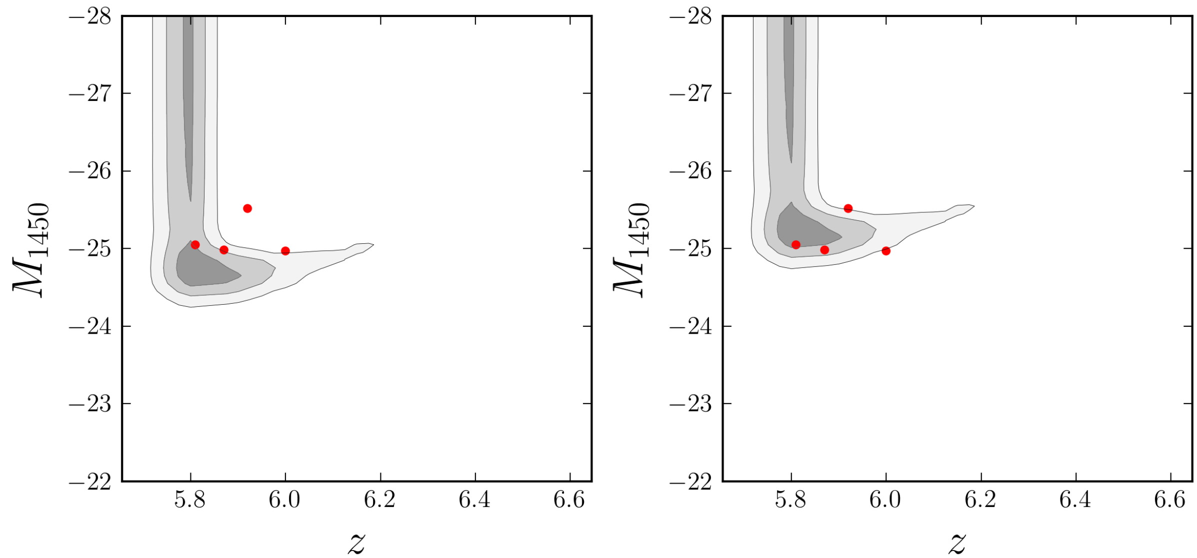


Figure 5.4: Figures showing the effective completeness of the second Stripe 82 survey (Jiang et al., 2009). It is clear from the left hand panel that one of the quasars here lies at a point of zero completeness. This suggests that the survey's magnitude limit was not as faint as originally thought. The right hand panel shows the new completeness after performing a maximum likelihood estimate of the value $\Delta M = -0.497$.

5.3 TREATMENT OF OVERLAPPING SURVEYS

When the Stripe 82 survey was conducted, an initial survey (Jiang et al. 2008) was carried out over $\sim 260\ \mathrm{deg}^2$ and 6 quasars were found. This survey covers a right ascension range of $30 < \alpha < 310$. Subsequently, a deeper survey (Jiang et al. 2009) surveyed $195\ \mathrm{deg}^2$ of the sky and found a subsequent 4 quasars at fainter magnitudes, with a right ascension range of $60° < \alpha < 310°$. The two surveys both span a declination range of $-1.5° < \delta < 1.5°$ (Fig. 5.5 summarises the coverage of the two surveys and the positions of the quasars that were discovered). The corresponding completeness of these two surveys are shown in panels 2 and 3 of Fig. 3.4.

The overlap between the two surveys (see Fig. 5.5) means that they cannot be treated independently. Instead one needs to treat the shallower of the two surveys as only surveying the range $4\mathrm{h} < \alpha < 2\mathrm{h}$ (the yellow shaded area in Fig. 5.5) where the surveys don't overlap. Three quasars will then be part of this sample with an effective area of $65\ \mathrm{deg}^2$ while the remaining 7 discovered in the range $30° < \alpha < 310°$ (blue shaded region in Fig. 5.5) will be part of the deep sample with an area $195\ \mathrm{deg}^2$.

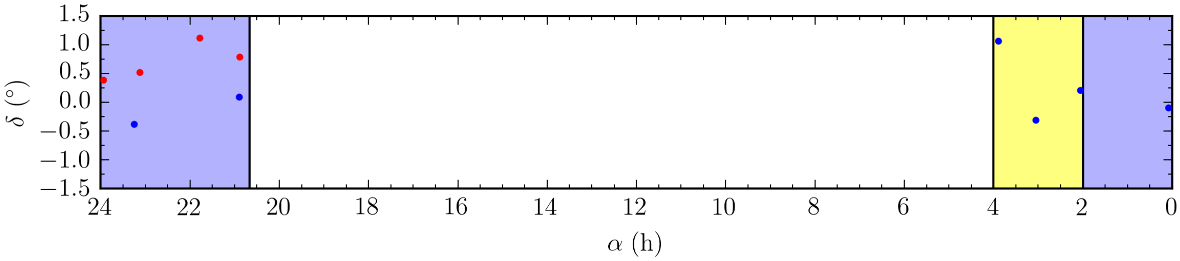


Figure 5.5: Positions of the quasars discovered in both Stripe 82 surveys. Blue points show those found in the shallow survey (Jiang et al., 2008) while red points show those found in the deep survey (Jiang et al., 2009). The blue shaded region is the area covered by both surveys, while the yellow shaded region is that was only covered by the shallow survey.

5.4 MAXIMUM LIKELIHOOD ESTIMATES FOR CORRECTED SELECTION FUNCTION

To properly treat these quasar surveys, we need to implement the maximum likelihood value for the shift in magnitude cut off for the Stripe 82 deep (Jiang et al., 2009) survey while also rearranging the quasars between the two (Jiang et al., 2008; Jiang et al., 2009) surveys according to section 5.3. We are then left with the following tables (Figure 5.6) of completeness with the quasars overlayed on top.

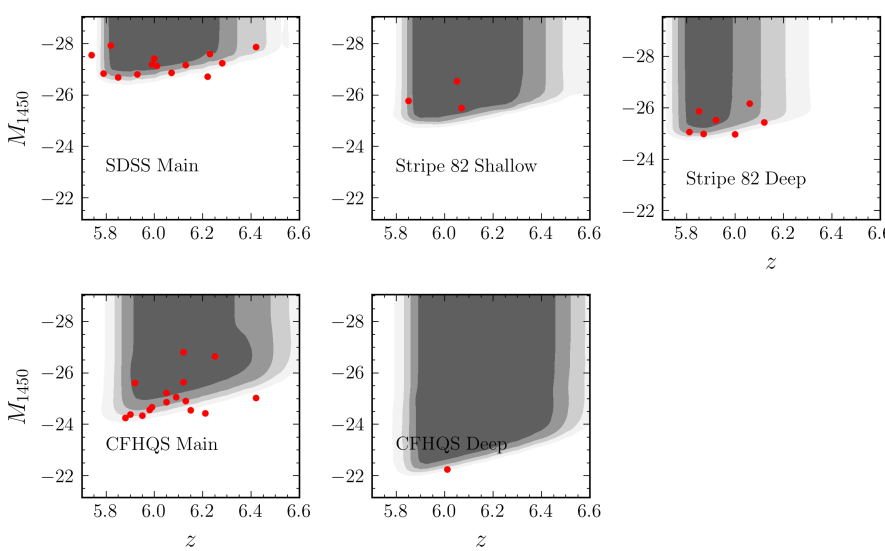


Figure 5.6: Completeness tables of the $z \sim 6$ quasar surveys correcting for those of the Stripe 82 surveys according to section 5.3 and section 5.2.

We re-estimate the maximum likelihood values for the three parameters and find the following, $\beta = -2.85,\ M^*_{1450} = -25.09,\ \Phi^* = 1.36 \times 10^{-8}\ \mathrm{Mpc}^{-3}$. Fig. 5.7 shows re-calculated confidence intervals in all parameters.

5.5 PRIOR DISTRIBUTION ON QLF PARAMETERS

On inspection of Figure 5.1 it becomes clear that there is a degeneracy between the break magnitude and bright end slope of the quasar luminosity function. This is a result of the small sample of faint quasars observed in the redshift 6 surveys. In principle they could all be fit by a double power law where the break magnitude was fainter than the faintest quasar and therefore in principle the break magnitude could have any value lower than ~ -27 as long as the bright end slope was shallow enough to pass through all the data points. With only the redshift 6 data, it is difficult to impose constraints that would break this degeneracy. However, there are a much larger number of quasars observed at $z \sim 5$ (McGreer et al. (2013), section 3.2), and including these in our sample will provide more data at faint magnitudes will allow us to place tighter constraints on the shape of the quasar luminosity function.

In order to account for the evolution of the QLF parameters, we need to introduce extra parameters into our analysis. Most importantly these include evolution of number density, parameterised by k, and evolution of the break magnitude, parameterised by j.

We can further constrain these parameters with the knowledge that there will be more quasars at $z = 5$ than $z = 6$ due to the extra time available for accretion to those at $z = 5$. Following with the same reasoning, we expect the break magnitude to move to the bright end as we increase the redshift. We therefore expect $k \leq 0$ and $j \geq 0$. We also impose uniform priors on all parameters, with the condition $\Phi^* \geq 0,\ M^*_{1450} < 0,\ \alpha < 0,\ \beta < \alpha$. With the redshift 5 data constraining the shape of the QLF and prior distributions on QLF parameters it is possible to implement an MCMC routine to sample from the posterior distribution on QLF parameters.

5.6 MCMC ANALYSIS OF QLF PARAMETERS

With prior distributions on QLF parameters and the inclusion of redshift 5 data, we are now in a position to use an MCMC routine to sample from the posterior distribution on 6 QLF parameters. We start by ensuring that the results detailed in section 4.3 are reproduced on the double power law model with 3 parameters using only the redshift 6 data. Fig. 5.7 shows the posterior distribution of QLF parameters obtained by evaluating the likelihood over a grid in β, M^*_{1450} and Φ^*.

Figure 5.7 shows that there are large degeneracies between all the parameters of the QLF with this parameterisation. As a result of these large degeneracies the MCMC routine is unable to sample correctly from the posterior distribution. If we chose to fix the number density at a given value of M_{1450}, rather than at M^*_{1450}, we can break the degeneracies between Φ^* and β and Φ^* and M^*_{1450}.

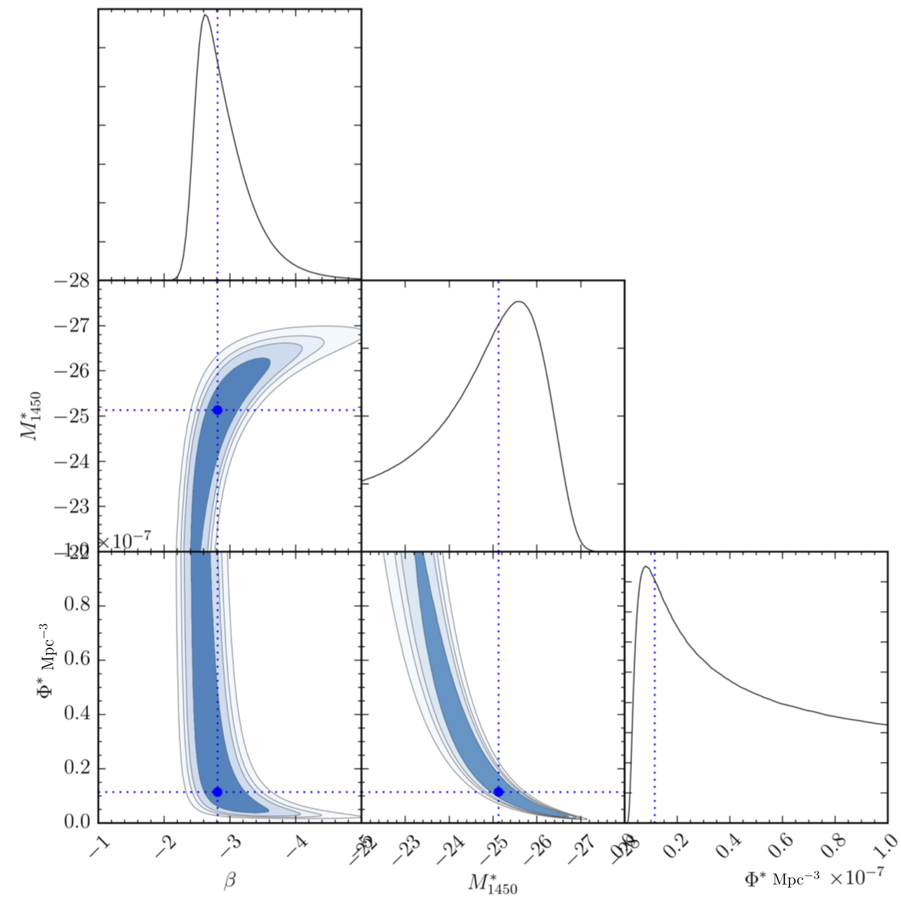


Figure 5.7: Marginalised posterior distributions of the 3 parameter QLF at $z \sim 6$ calculated using the corrected Stripe 82 completeness and quasar sample. Contours are highest posterior density regions and blue points show the maximum likelihood result calculated in section 5.4.

This can clearly be seen in Figure 5.8 where we see that Φ_{26} is now well constrained. The degeneracy between M^*_{1450} and β is still obvious, however this should be reduced by including the $z \sim 5$ data and reducing the number of parameters that are degenerate will allow the sampling routines to sample correctly from the posterior distribution. The conversion from Φ^* to Φ_{26} for a given set of model parameters ($\beta_0,\ M^*_{1450,0},\ \alpha_0$) is then,

$$\Phi_{26} = \Phi(z = 6, M_{1450} = -26) = \frac{\Phi^*}{10^{0.4(\alpha_0+1)(-26-M^*_{1450,0})} + 10^{0.4(\beta_0+1)(-26-M^*_{1450,0})}} \tag{5.1}$$

In both Figure 5.7 and Figure 5.8 the priors on Φ^* and Φ_{26} are uniform. This means that the adjustment from Φ^* to Φ_{26} is analogous to adding a non-uniform prior over Φ^* eliminating the possibilities of Φ^* become arbitrarily large as M^*_{1450} moves to fainter magnitudes.

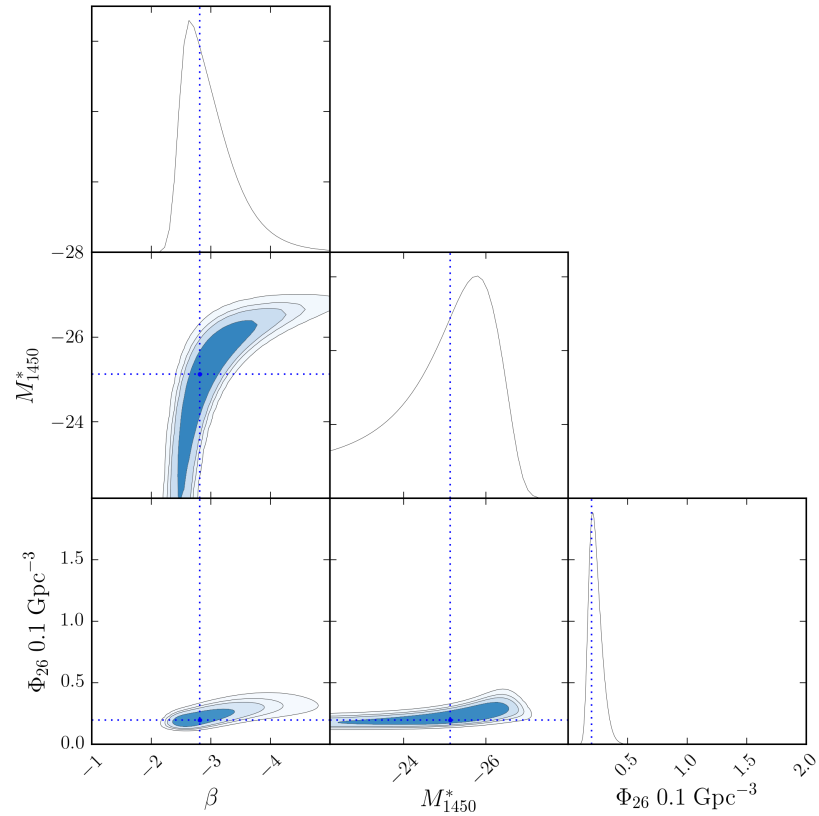


Figure 5.8: Marginalised posterior distributions of the 3 parameter QLF at $z \sim 6$ parameterising the normalisation with $\Phi_{26} = \Phi(-26, 6)$ rather than $\Phi^* = \Phi(M^*_{1450}, 6)$. Contours are $(99.7\%, 95.4\%, 90\%, 68.3\%$ highest posterior density regions. Blue points are the maximum likelihood values of $M^*_{1450} = -25.13,\ \beta = -2.81,\ \Phi_{26} = 0.1980.1\ \mathrm{Gpc}^{-3}$ with α fixed at $\alpha = -1.5$.

Having removed the degeneracies between M^*_{1450} and Φ^* and β and Φ^* we can now be confident that the posterior distribution can be sampled correctly by the MCMC routine. We run a sampler of 6 chains each taking $10,000$ samples from the posterior distribution of QLF parameters and obtain Figure 5.9.

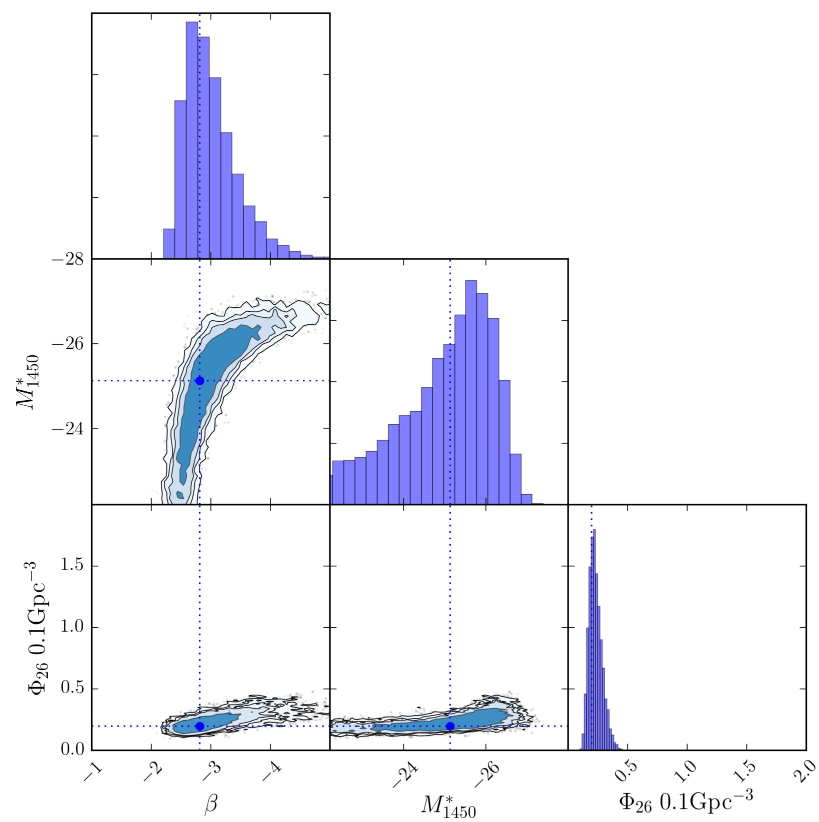


Figure 5.9: Results from MCMC analysis of 3 parameter QLF at $z \sim 6$. Contours are highest posterior density regions. Blue points show maximum likelihood result calculated in section 5.4.

We see that the MCMC sampler converges to the correct posterior distribution of QLF parameters. In particular, the degeneracy between M^*_{1450} and β is clear. This allows us to extend the analysis to higher dimensions where exactly calculating the posterior using the method prescribed in section 4.5 is computationally slow and inefficient.

For completeness, we repeat this analysis for the 4 parameter QLF model at $z \sim 5$. Figure 5.10 shows the posterior distribution of QLF parameters for the 4 parameter model adopted by McGreer et al. (2013) calculated using the grid method (section 4.5). The blue points are the maximum likelihood values of the parameters using the likelihood prescribed in section 4.1. Figure 5.11 shows results of MCMC analysis with 8 chains each sampling $10,000$ times from the 4 dimensional posterior distribution. Again, as in the 3 parameter case, the grid method and MCMC analysis produce consistent results.

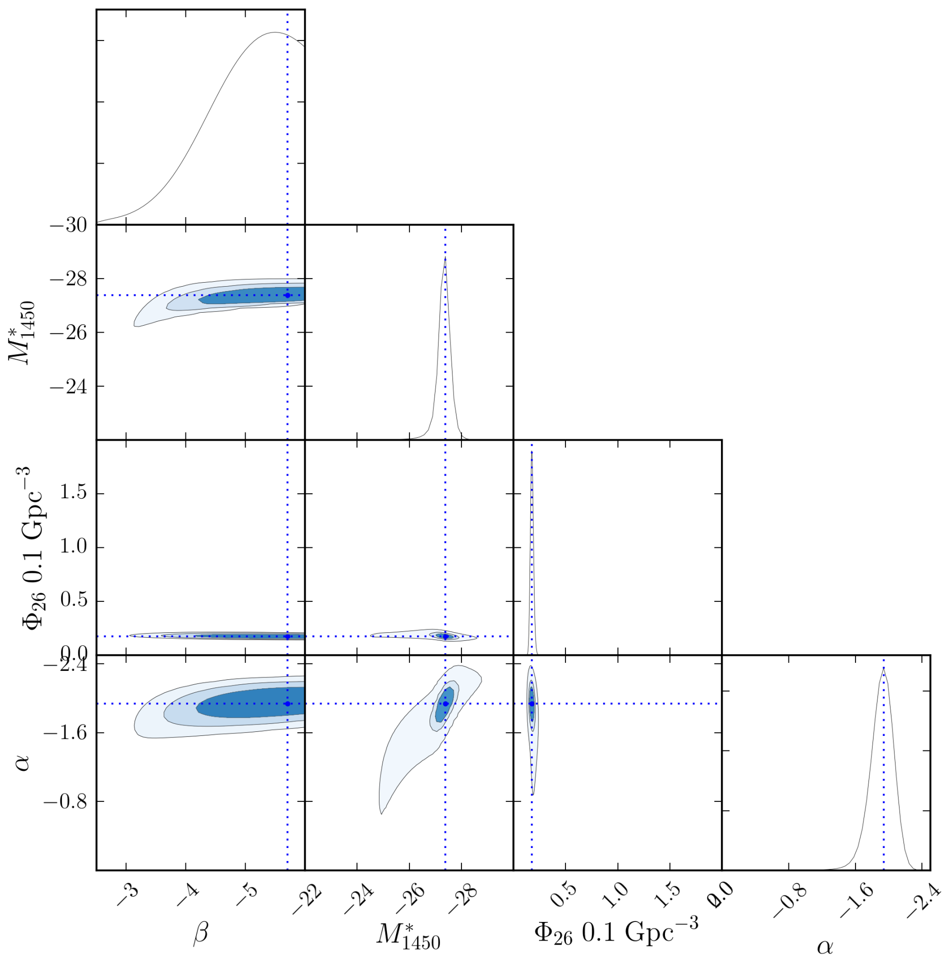


Figure 5.10: Marginalised posterior distributions of the 4 parameter QLF at $z \sim 5$. Contours are highest posterior density regions and blue points show the maximum likelihood result as calculated in McGreer et al. (2013).

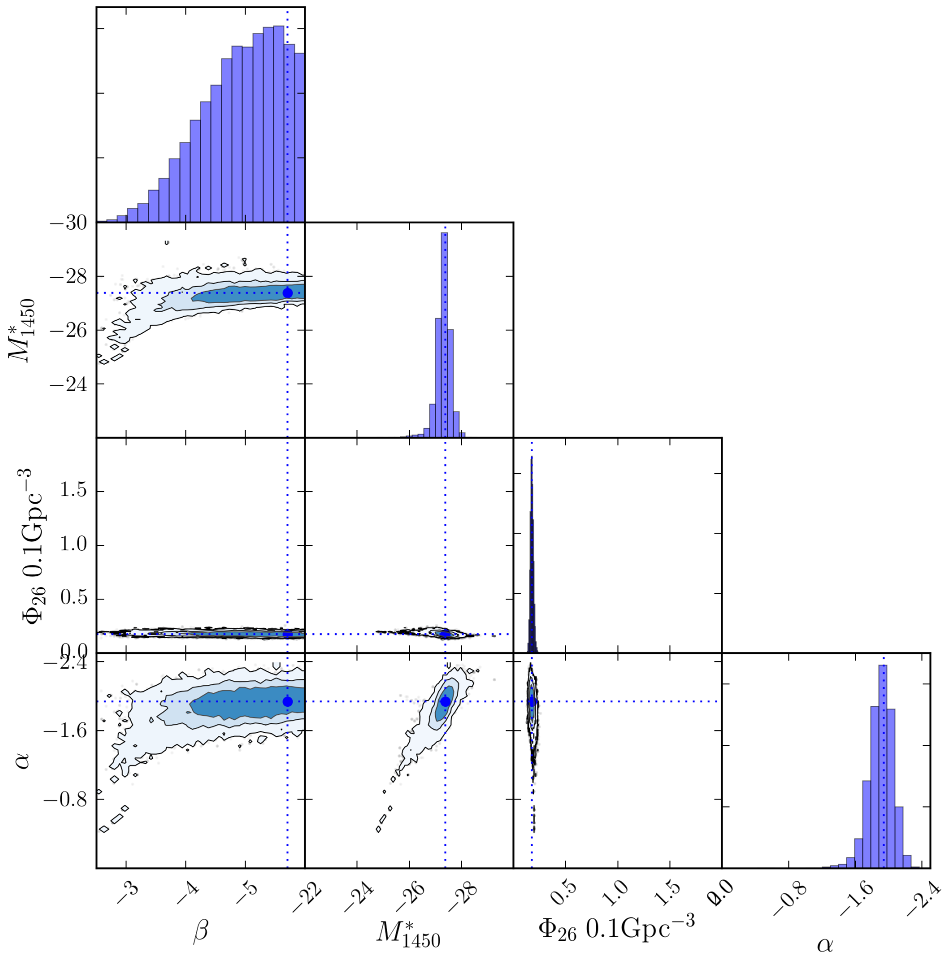


Figure 5.11: Results from MCMC analysis of 4 parameter QLF at $z \sim 5$. Contours are highest posterior density regions. Blue points show maximum likelihood result calculated in McGreer et al. (2013).

We then included both redshift 5 and 6 data in our analysis and allowed for 6 free parameters: $\beta, M^*_{1450}, \Phi^*, \alpha, k, j$. Initially, we ran our analysis on quasars sampled from $f(M_{1450}, z)$ with the following values for QLF parameters: $\alpha = -1.5, \beta = -2.81, M^*_{1450} = -25.13, \Phi^* = 11.4$ Gpc^{-3}, $k = -0.47, j = 0.0$. Quasars were simulated from the same completeness tables but the survey areas were increased by a factor of 10 to produce tighter constraints on the QLF parameters. The samples are shown in Fig. 5.12 and results from this analysis are shown in Fig. 5.13. From these results we see that the MAP estimates are consistent with the qlf model parameters from which the quasars are sampled.

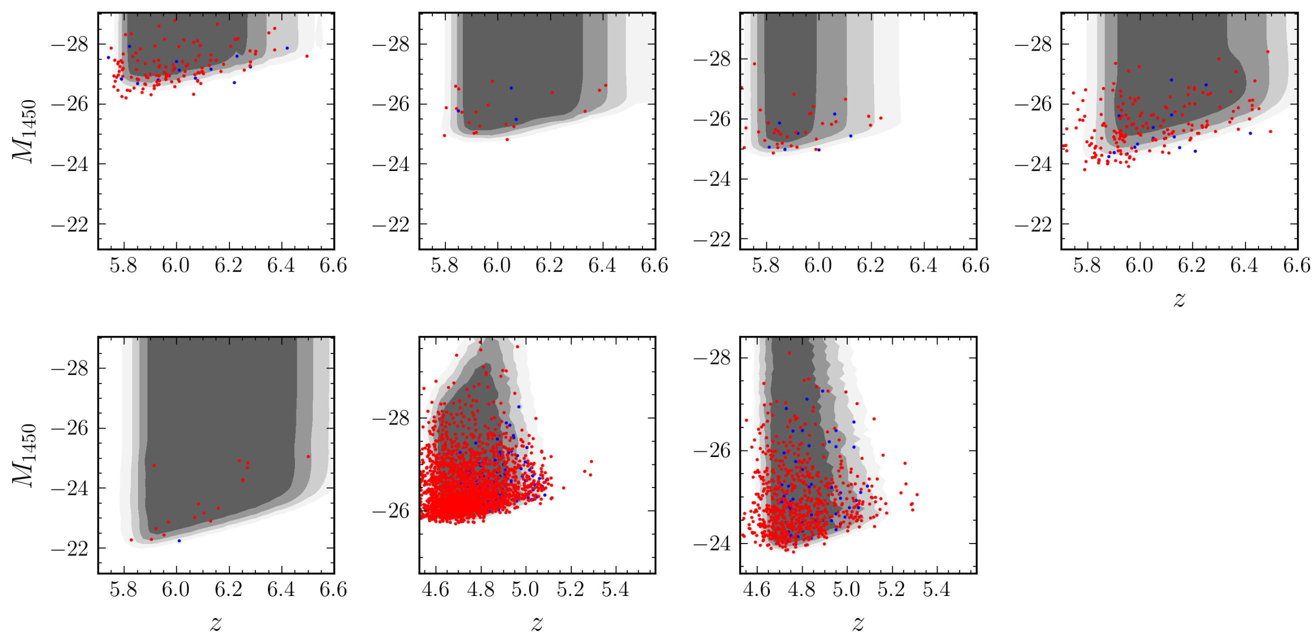


Figure 5.12: Figure showing contours of corrected completeness, as in Fig. 5.6 with simulated quasar surveys shown as red points. The true quasars are shown as blue points, corrected for overlapping survey areas.

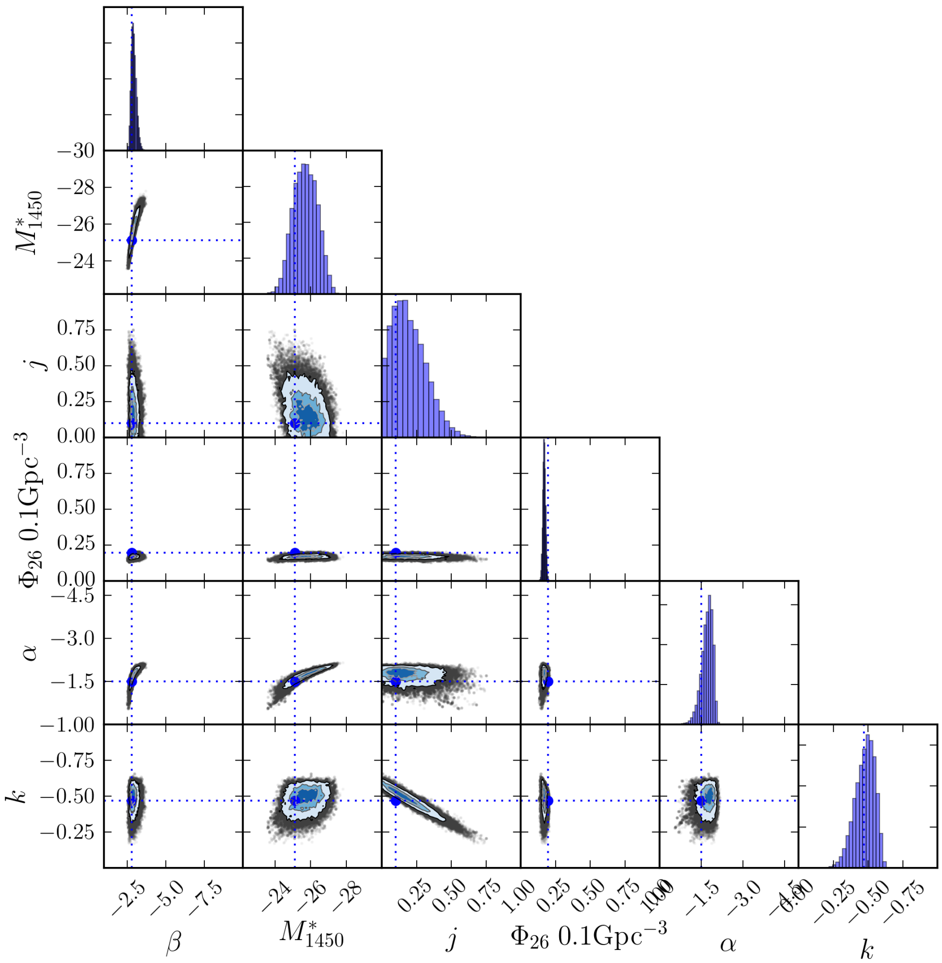


Figure 5.13: Corner plot showing results from MCMC analysis of double power law QLF with 6 free parameters, with quasars simulated from a double power law QLF with parameters ($\beta = -2.81,\ M^*_{1450} = -25.13,\ \Phi^* = 11.4\ \mathrm{Gpc}^{-3},\ \alpha = -1.5,\ j = 0.1,\ k = -0.47$).

Using the real data Fig. 5.14 is obtained. Maximum A Posteriori (MAP) estimates are $\beta = -5.45,\ M^*_{1450} = -27.36,\ j = 0.0198,\ \Phi^* = 3.0\ \mathrm{Gpc}^{-3},\ \alpha = -1.95,\ k = -0.25$. From this figure, it seems that bright break magnitudes with steep bright end slopes are favoured by the data. This could suggest the possibility of adopting a QLF of the form Schechter (1976),

$$\Phi(M_{1450}, z) = 10^{k(z-z_{\mathrm{ref}})}\Phi^*\ 10^{0.4(\alpha+1)(M_{1450}-M^*_{1450}-j(z-z_{\mathrm{ref}}))} \times \exp -10^{0.4(M_{1450}-M^*_{1450}-j(z-z_{\mathrm{ref}}))}], \tag{5.2}$$

which is characterised by an exponential fall off of the bright slope. Note that this is the same form as the BHMF as described in section 2.2.

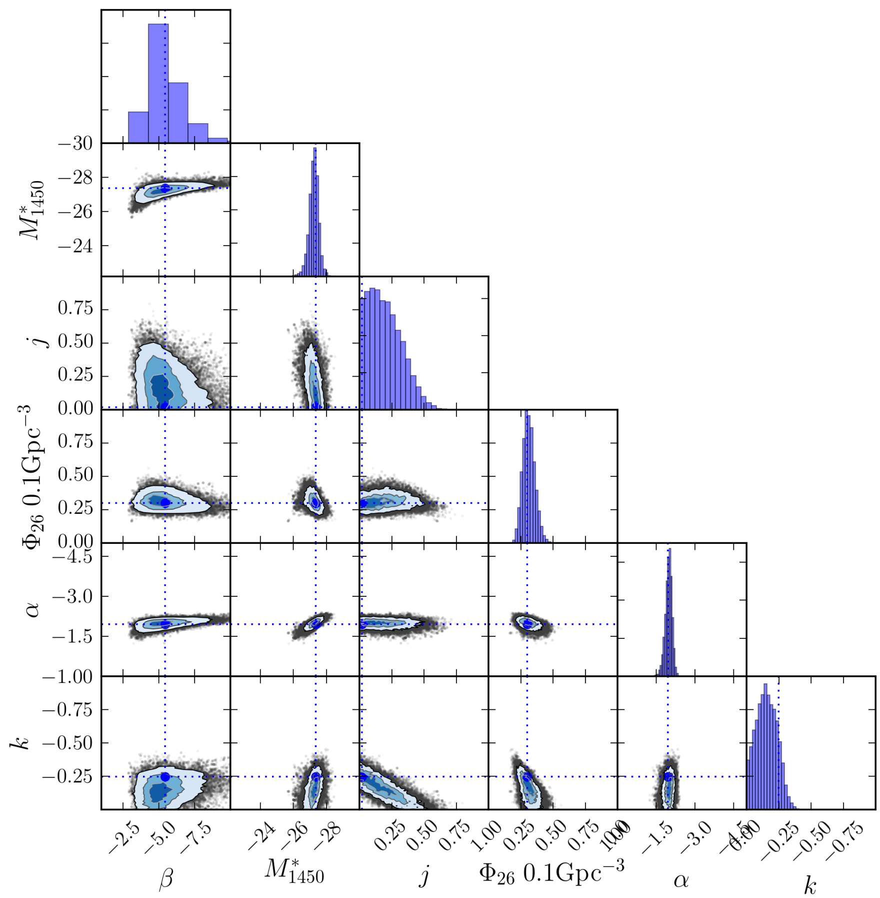


Figure 5.14: Corner plot showing results from MCMC analysis of double power law QLF with 6 free parameters. Maximum A Posteriori estimates are ($\beta = -5.45,\ M^*_{1450} = -27.36,\ j = 0.0198,\ \Phi^* = 3.0\ \mathrm{Gpc}^{-3},\ \alpha = -1.95,\ k = -0.25$), indicated in each panel. Note that the MAP estimates are global estimates and as a result will not always correspond to the peak of the marginalised posterior distributions.

To better interpret the results of our MCMC analysis, it is useful to visualise the possibilities for the luminosity function. Taking 50 draws from the posterior, we can calculate the luminosity function for different combinations of QLF parameters we have sampled, as shown in Fig. 5.15. In in the two panels, the line colours correspond to the same values of QLF parameters.

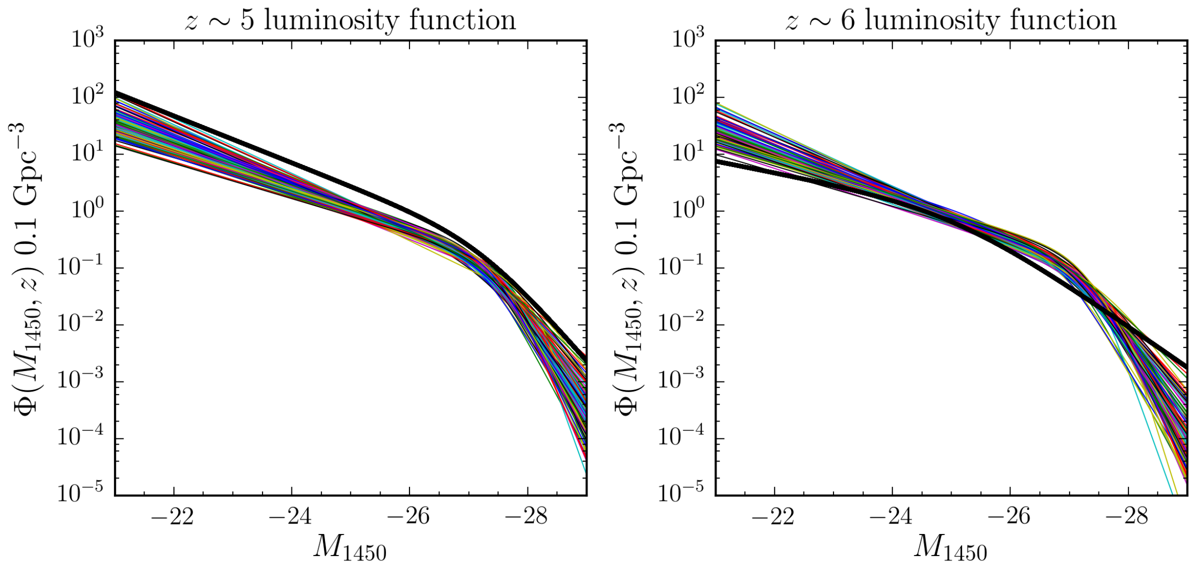


Figure 5.15: Figure showing plots of the QLF for different parameters sampled from the posterior in Fig. 5.14. Left hand panel shows the QLF at $z = 5$, the right hand panel shows the QLF at $z = 6$. Note that colours in each panel correspond to the same QLF parameters. The thick black lines correspond to the McGreer et al. (2013) best fit values and Willott et al. (2010a) best fit values for $z \sim 5$ and $z \sim 6$ respectively. The best fit QLF in the left hand panel is systematically above the samples, which is due to the fact that the likelihood used in McGreer et al. (2013) is slightly different to the one used in this thesis, as discussed in section 5.1

6 | Discussion

As we have seen, there is an outstanding issue with our understanding of quasar formation. Observations of quasars powered by black holes with masses up to $10^9 M_\odot$ at redshifts $\gtrsim 6$ leave unanswered the question of how black holes have evolved to reach such high masses within only 1 billion years of the Big Bang. This inconsistency could arise from a misunderstanding of the black hole seed formation, black hole growth or from uncertainties in the black hole mass estimates themselves. Observationally, these physical mechanisms all contribute to the observed QLF and its evolution with redshift. The aim of this thesis, therefore, was to present a self-consistent approach to obtaining posterior distributions on QLF parameters at $z \sim 6$. By improving the accuracy of statistical analysis we provide reliable constraints on the quasar luminosity function which evolutionary models and mass estimates must reproduce.

We first presented a method for treating the case of overlapping surveys and apply a correction to the completeness of the SDSS Stripe 82 shallow survey (Jiang et al., 2008). These adjustments have only a small effect on the maximum likelihood estimates of QLF parameters, in comparison to those found in Willott et al. (2010a). This is because approximately 4 out of 50 quasars in the sample were affected by the update. However, with larger high redshift quasar surveys in the future it will become all the more important to treat these problems carefully and correctly, using the correct formalism.

We then introduced the methodology (see section 4.4) for obtaining posterior distributions on the quasar luminosity function parameters by using MCMC. Due to the limited number of quasars at $z \sim 6$, we use quasars from $z \sim 5$ (McGreer et al., 2013) to impose what are effectively prior constraints on the $z \sim 6$ QLF parameters.

The following 6 parameter model of the QLF was used, described in detail in chapter 2,

$$\Phi(M_{1450}, z) = \frac{10^{k(z-z_{\rm ref})}\Phi_{26}}{10^{0.4(\alpha+1)(M_{1450}-M^*_{1450}-j(z-z_{\rm ref}))} + 10^{0.4(\beta+1)(M_{1450}-M^*_{1450}-j(z-z_{\rm ref}))}}, \tag{6.1}$$

where we have introduced a new parameterisation of the QLF normalisation, fixing it at a reference magnitude $M_{1450} = -26$ rather than at M^*_{1450}. This new parameterisation with uniform priors on Φ_{26} is equivalent to imposing non-uniform priors on the original normalisation parameter, Φ^*

and removes many of the degeneracies between the parameters of the model. This improves the performance of MCMC routines when sampling from the posterior distribution.

Maximum A-Posteriori point estimates for these parameters are ($\beta = -5.45,\ M^*_{1450} = -27.36,\ j = 0.0198,\ \Phi^* = 3\ \mathrm{Gpc}^{-3},\ \alpha = -1.95,\ k = -0.25$). As in Willott et al. (2010a), no uncertainties are quoted due to the existence of large degeneracies between parameters, however confidence regions can be seen in the corner plots in section 5.6. The bright end is noticeably steeper than quoted by Willott et al. (2010a) and this could lead to the possibility of using a Schechter (1976) QLF of the form

$$\Phi(M_{1450}, z) = 10^{k(z-z_{\mathrm{ref}})}\Phi^*\ 10^{0.4(\alpha+1)(M_{1450}-M^*_{1450}-j(z-z_{\mathrm{ref}}))} \times \exp -10^{0.4(M_{1450}-M^*_{1450}-j(z-z_{\mathrm{ref}}))}].$$

characterised by an exponential fall-off at the bright end.

The value of k is smaller than calculated previously by Fan et al. (2001c) suggesting the fall off in number density of quasars with redshift is slower than previously thought. The value of$j$ is small, as expected, and our results are consistent with a value $j = 0$. Further analysis could include model comparison to decide if the inclusion of this evolution is needed, and whether or not the data would be better fit by also including evolution in β and/or α.

With a self-consistent approach to obtaining posterior distributions on QLF parameters in place, two avenues of future work should be pursued to understand the formation of supermassive black holes at high redshift. Assuming current mass estimates of the black holes powering quasars in the early universe are accurate, it is possible to obtain the black hole mass function (BHMF) by de-convolving the QLF with the Eddington ratio, $\lambda = L/L_{\mathrm{Edd}}$ of high redshift black holes. The BHMF is defined analogously to the quasar luminosity function and L and M_{BH} are related via the Eddington ratio distribution, λ. Thus the BHFM and the QLF are related via a convolution with the distribution of Eddington ratios. Large hydrodynamical simulations of black hole growth have been performed (Di Matteo et al., 2012) and these can then be constrained by the necessity to recreate the BHMF at each given redshift.

It is also important to consider the possibility that the black hole mass estimates are inaccurate. If black hole masses are, in fact, less massive than currently thought, then current models of black hole growth may be able to explain their existence. As described in chapter 1, black hole masses at high redshift are estimated via scaling relations (Vestergaard & Peterson, 2006; Vestergaard & Osmer, 2009) observed at low-z. However, the assumption that these relations apply at high redshift may not be valid, and the complicated nature of the uncertainties means that classical regression is difficult. Implementation of a Bayesian hierarchical model would allow updated mass estimates and incorporate correct treatment of the large number of uncertainties arising from these measurements.

In conclusion, we have developed a rigorous method of statistical analysis with which QLF parameters can be estimated. We further discussed the appropriate adjustments that must be made when selecting quasars from overlapping surveys. The results of this work are intended to provide a reliable benchmark to which future evolutionary models should be compared, thereby aiding in the understanding of quasar formation.

Bibliography

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, ApJS, 182, 543

Abramowicz, M. A. 2005, in Growing Black Holes: Accretion in a Cosmological Context, ed. A. Merloni, S. Nayakshin, & R. A. Sunyaev, 257

Abramowicz, M. A., Czerny, B., Lasota, J. P., & Szuszkiewicz, E. 1988, ApJ, 332, 646

Alvarez, M. A., Wise, J., & Abel, T. 2009, ApJL, 701, L133

Antonucci, R. 1993, ARAA, 31, 473

Avni, Y. 1976, ApJ, 210, 642

Avni, Y., & Bahcall, J. N. 1980, ApJ, 235, 694

Bayes, M., & Price, M. 1763, Philosophical Transactions of the Royal Society of London Series I, 53, 370

Begelman, M. C., Volonteri, M., & Rees, M. J. 2006, MNRAS, 370, 289

Bondi, H., & Hoyle, F. 1944, MNRAS, 104, 273

Boyle, B. J., Shanks, T., Croom, S. M., et al. 2000, MNRAS, 317, 1014

Boyle, B. J., Shanks, T., & Peterson, B. A. 1988, MNRAS, 235, 935

Brooks, S. P., & Gelman, A. 1998, Journal of computational and graphical statistics, 7, 434

Burbidge, G. R., Burbidge, E. M., & Sandage, A. R. 1963, Reviews of Modern Physics, 35, 947

Croom, S. M., Smith, R. J., Boyle, B. J., et al. 2004, MNRAS, 349, 1397

Croom, S. M., Richards, G. T., Shanks, T., et al. 2009, MNRAS, 399, 1755

Di Matteo, T., Colberg, J., Springel, V., Hernquist, L., & Sijacki, D. 2008, ApJ, 676, 33

Di Matteo, T., Khandai, N., DeGraf, C., et al. 2012, ApJ, 745, L29

Fan, X. 2003, in Bulletin of the American Astronomical Society, Vol. 35, American Astronomical Society Meeting Abstracts, 1415

Fan, X., Narayanan, V. K., Lupton, R. H., et al. 2001a, AJ

Fan, X., Strauss, M. A., Richards, G. T., et al. 2001b, AJ

Fan, X., Strauss, M. A., Schneider, D. P., et al. 2001c, AJ

Fan, X., Hennawi, J. F., Richards, G. T., et al. 2004, AJ, 128, 515

Fan, X., Strauss, M. A., Richards, G. T., et al. 2006, AJ, 131, 1203

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

Fukugita, M., Ichikawa, T., Gunn, J. E., et al. 1996, AJ, 111, 1748

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. 2014, Bayesian Data Analysis

Gelman, A., & Rubin, D. B. 1992, Statistical science, 457

Glikman, E., Bogosavljević, M., Djorgovski, S. G., et al. 2010, ApJ, 710, 1498

Goodman, J., & Weare, J. 2010, CAMCOS, 5, 65

Hao, L., Strauss, M. A., Fan, X., et al. 2005, AJ, 129, 1795

Hastings, W. K. 1970, Biometrika, 57, 97

Hopkins, P. F., Lidz, A., Hernquist, L., et al. 2007a, ApJ, 662, 110

Hopkins, P. F., Richards, G. T., & Hernquist, L. 2007b, ApJ, 654, 731

Hunt, M. P., Steidel, C. C., Adelberger, K. L., & Shapley, A. E. 2004, ApJ, 605, 625

Ikeda, H., Nagao, T., Matsuoka, K., et al. 2011, ApJL, 728, L25

Jiang, L., McGreer, I. D., Fan, X., et al. 2015, AJ, 149, 188

Jiang, L., Fan, X., Annis, J., et al. 2008, AJ, 135, 1057

Jiang, L., Fan, X., Bian, F., et al. 2009, AJ, 138, 305

Kaspi, S., Maoz, D., Netzer, H., et al. 2005, ApJ, 629, 61

Kellermann, K. I. 2013, Bulletin of the Astronomical Society of India, 41, 1

Kelly, B. C., Fan, X., & Vestergaard, M. 2008, ApJ, 682, 874

Kelly, B. C., & Merloni, A. 2012, Advances in Astronomy, 2012, 970858

Kelly, B. C., & Shen, Y. 2013, ApJ, 764, 45

Kelly, B. C., Vestergaard, M., Fan, X., et al. 2010, ApJ, 719, 1315

Komatsu, E., Dunkley, J., Nolta, M. R., et al. 2009, ApJS, 180, 330

Lampton, M., Margon, B., & Bowyer, S. 1976, ApJ, 208, 177

Lawrence, A., Warren, S. J., Almaini, O., et al. 2007, MNRAS, 379, 1599

Marshall, H. L., Tananbaum, H., Avni, Y., & Zamorani, G. 1983, ApJ, 269, 35

Marshall, J. L., Burles, S., Thompson, I. B., et al. 2008, in SPIE, Vol. 7014, Ground-based and Airborne Instrumentation for Astronomy II, 701454

Masters, D., Capak, P., Salvato, M., et al. 2012, ApJ, 755, 169

McGreer, I. D., Jiang, L., Fan, X., et al. 2013, ApJ, 768, 105

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. 1953, Journal of Chemical Physics, 21, 1087

Mortlock, D. J., Warren, S. J., Venemans, B. P., et al. 2011, Nature, 474, 616

Netzer, H. 2015, ARAA, 53, 365

Oke, J. B., & Gunn, J. E. 1983, ApJ, 266, 713

Osmer, P. S. 1982, ApJ, 253, 28

Pacucci, F., Volonteri, M., & Ferrara, A. 2015, MNRAS, 452, 1922

Paczynski, B. 1998, AcA, 48, 667

Pâris, I., Petitjean, P., Aubourg, É., et al. 2014, A&A, 563, A54

Pei, Y. C. 1995, ApJ, 438, 623

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

Peterson, B. M. 2010, in Bulletin of the American Astronomical Society, Vol. 42, American Astronomical Society Meeting Abstracts #215, 581

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 2002, Numerical recipes in C++ : the art of scientific computing

Rees, M. J. 1978, The Observatory, 98, 210

—. 1984, ARA&A, 22, 471

Reylé, C., Delorme, P., Willott, C. J., et al. 2010, å, 522, A112

Richards, G. T., Fan, X., Newberg, H. J., et al. 2002, AJ, 123, 2945

Richards, G. T., Croom, S. M., Anderson, S. F., et al. 2005, MNRAS, 360, 839

Richards, G. T., Lacy, M., Storrie-Lombardi, L. J., et al. 2006a, ApJS, 166, 470

Richards, G. T., Strauss, M. A., Fan, X., et al. 2006b, AJ, 131, 2766

Ross, N. P., Myers, A. D., Sheldon, E. S., et al. 2012, ApJS, 199, 3

Schechter, P. 1976, ApJ, 203, 297

Schmidt, M. 1963, Nature, 197, 1040

—. 1968, ApJ, 151, 393

Schmidt, M., Schneider, D. P., & Gunn, J. E. 1995a, AJ, 110, 68

—. 1995b, AJ, 110, 68

Schneider, D. P., Schmidt, M., & Gunn, J. E. 1994, AJ, 107, 1245

—. 1999, AJ, 117, 40

Shankar, F., & Mathur, S. 2007, ApJ, 660, 1051

Shankar, F., Weinberg, D. H., & Miralda-Escudé, J. 2009, ApJ, 690, 20

—. 2013, MNRAS, 428, 421

Shen, Y., & Kelly, B. C. 2012, ApJ, 746, 169

Siana, B., Polletta, M. d. C., Smith, H. E., et al. 2008, ApJ, 675, 49

Sijacki, D., Springel, V., & Haehnelt, M. G. 2009, MNRAS, 400, 100

Steffen, A. T., Barger, A. J., Cowie, L. L., Mushotzky, R. F., & Yang, Y. 2003, ApJL, 596, L23

Tanaka, T., & Haiman, Z. 2009, ApJ, 696, 1798

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

Urry, C. M., & Padovani, P. 1995a, PASP, 107, 803

—. 1995b, PASP, 107, 803

Vestergaard, M., Fan, X., Tremonti, C. A., Osmer, P. S., & Richards, G. T. 2008, ApJL, 674, L1

Vestergaard, M., & Osmer, P. S. 2009, ApJ, 699, 800

Vestergaard, M., & Peterson, B. M. 2006, ApJ, 641, 689

Volonteri, M. 2010, AR&A, 18, 279

Volonteri, M., Habouzit, M., Pacucci, F., & Tremmel, M. 2015a, ArXiv e-prints, arXiv:1511.02588

Volonteri, M., Lodato, G., & Natarajan, P. 2008, MNRAS, 383, 1079

Volonteri, M., & Rees, M. J. 2005, ApJ, 633, 624

Volonteri, M., Silk, J., & Dubus, G. 2015b, ApJ, 804, 148

Warren, S. J., Hewett, P. C., & Osmer, P. S. 1994, ApJ, 421, 412

Willott, C. J., Rawlings, S., Blundell, K. M., Lacy, M., & Eales, S. A. 2001, MNRAS, 322, 536

Willott, C. J., Delorme, P., Omont, A., et al. 2007, AJ, 134, 2435

Willott, C. J., Delorme, P., Reylé, C., et al. 2009, AJ, 137, 3541

Willott, C. J., Albert, L., Arzoumanian, D., et al. 2010, AJ, 140, 546

Willott, C. J., Delorme, P., Reylé, C., et al. 2010a, AJ, 139, 906

Wu, X.-B., Wang, F., Fan, X., et al. 2015, Nature, 518, 512

Yee, H. K. C., Gladders, M. D., Gilbank, D. G., et al. 2007, in Astronomical Society of the Pacific Conference Series, Vol. 379, Cosmic Frontiers, ed. N. Metcalfe & T. Shanks, 103

York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, AJ, 120, 1579

Yu, Q., & Tremaine, S. 2002, MNRAS, 335, 965

Zhang, W., Woosley, S. E., & Heger, A. 2008, ApJ, 679, 639